ICIC Express Letters ICIC International ©2020 ISSN 1881-803X
Volume **14**, Number **10**, October **2020** pp. 1001–1008

ROAD DAMAGE DETECTION OVER ROAD SCANNER IMAGES USING DEEP CONVOLUTIONAL NEURAL NETWORK

Hwiyong Jo[1], Dongsoo Kim[2], Kyong-Won Pak[3] and Minsoo Kim[4,*]

[1]Graduate School of Management of Technology
Pukyong National University
365 Sinseon-ro, Nam-gu, Busan 48547, Korea
johwiyong@pukyong.ac.kr

[2]Department of Industrial and Information Systems Engineering
Soongsil University
369 Sangdo-ro, Dongjak-gu, Seoul 06978, Korea
dskim@ssu.ac.kr

[3]PAVETEC KOREA Co., Ltd.
66 Chungmin-ro, Songpa-gu, Seoul 05838, Korea
pkw4425@naver.com

[4]Division of Systems Management and Engineering
Pukyong National University
45 Yongso-ro, Nam-gu, Busan 48513, Korea
*Corresponding author: minsky@pknu.ac.kr

Received March 2020; accepted June 2020

Abstract. *This paper deals with the automatic detection and classification of road cracks. Determining the type and severity of a crack under the conventional PMS (Pavement Management System) is a challenging task in terms of process efficiency and consistency since it relies on human intervention of many trained workers. Many studies that reduce this human reliance have been conducted for the automation of crack detection, and thus improve the whole pavement management process significantly. Recently, neural networks have been actively tried for road damage detection and have achieved remarkable results. However, since many researches have been done on the basis of limited road images of a few countries or well-cleansed example data, more adaptive research is needed to obtain effective results for road conditions of individual countries. In this paper, 3 CNN-based models for detecting road cracks are tested using actual data from Korea. By doing multiple experiments for 6 types of cracks, several approaches to better detect cracks in Korean road conditions are examined. Based on the result of this study, many researches will be able to further enhance their studies while adapting their models to better fit to their respective domestic road conditions.*

Keywords: Road crack detection, Pavement distress detection, Image classification, Deep learning, Convolutional neural network

1. **Introduction.** The Pavement Management System (PMS) constantly monitors pavement distress and supports proper maintenance work depending on the type and severity of cracks. In metropolitan areas with heavy traffic and so many roads to manage, cracks occur very rapidly and thus significant resources are needed to smoothly detect road cracks and respond quickly. In most areas, however, crack detection is still done manually with lots of manpower. There are recent studies to obtain a 3D-scanned image of the road using 3D laser scanner as a way to clearly detect road cracks [1,2]. These latest approaches, however, have a practical problem that they require enormous investment to

DOI: 10.24507/icicel.14.10.1001

run the equipment in multiple vehicles. While conventional researches on road crack detection have been conducted through algorithmic analysis and structural interpretation of crack images [3-5], the focuses of those studies are now shifting to classification and object detection with neural networks that have shown significant advancement recently [6-8]. The introduction of Deep Neural Networks (DNNs) shows promising results for automatic crack detection, but is still in a lab environment or in a limited application phase with less crack types, which relies heavily on clean, limited sample input images that reflect road conditions of a few test countries. The main contribution of this paper lies in its extensive use of real image data obtained from the road scanner to distinguish 6 crack types under various test settings for field adaptation in Korea. In this study, 3 neural network models are reviewed and compared using road scanner images, and their accuracies are evaluated while changing the resolution of input images. Actual classification performance is also tested when multiple crack types coexist in the same input image. By referring this study, many researches will be able to further enhance their studies while adapting their models to better fit to their respective domestic road conditions.

This paper is organized as follows. Section 2 summarizes several related works on road crack detection. Section 3 presents general road crack management process, our test setups, and input data. Experiment results are given in Section 4 with brief explanations. Finally, in Section 5, our conclusion is stated along with some future research directions.

2. Related Works.

Researches to detect road damages or cracks can be roughly categorized into 2 approaches: an algorithmic approach and a DNN approach. The algorithmic approach basically focuses on the image processing and structural analysis, which includes parametric machine learning approaches too. The DNN approach has recently emerged in this field, which attempts to capture crack patterns by learning from the large amount of data. In this section, above two approaches are briefly described.

2.1. Algorithmic and analytic approaches.

This is a traditional approach that involves a variety of logical methods for detecting road cracks. In some studies, classification and segmentation were attempted using the threshold method for the distinction between crack pixels and non-crack background pixels [2,9]. Abdel-Qader et al. tried to classify cracks via edge detectors such as Canny detection and Sobel detection [10,11]. Tsai et al. studied a quantitative scoring method that can evaluate other crack segmentation methods objectively, and they applied their measure to six representative segmentation methods [5]. Subirats et al. conducted a study to print out the existence of crack in the form of a binary image using 2D CWT (Continuous Wavelet Transform) and wavelet coefficient maps [12]. Varadharajan et al. implemented an image-based detection method by applying SLIC (Simple Linear Iterative Clustering) superpixel algorithm in their classifier that can separate cracks from the backgrounds such as road and sky. In that study, they gained about the level of accuracy that was achieved with manual classification at that time [3]. Zalama et al. conducted a study to distinguish the longitudinal and transverse cracks using Gabor filter, but their approach showed some limitations that crosswalks or lanes were confused with cracks [4]. Nguyen et al. used CTA (Conditional Texture Anisotropy) and FFA (Free-Form Anisotropy) to enhance crack segmentation, and those two anisotropic measures showed better results than 2D CWT [13,14]. Hu and Zhao proposed a new LBP (Local Binary Pattern) algorithm for crack detection that was more robust to noises than the existing LBP algorithms [15]. Shi et al. presented the Crack-Forest method that used random structured forest approach. Their method performed better on several datasets with low-resolution images compared to other approaches. It also showed superior processing speed [16]. These traditional approaches, however, have shown some limitations that detection performance is sensitively affected by various kinds of noises, and as the complexity of algorithm increases, so do the implementation difficulty and computational workloads.

2.2. **DNN approaches.** Unlike the other analytic approaches, DNN approaches have the power to improve performances along with the growth of input data by constantly learning from the data. Among other techniques, CNN (Convolutional Neural Network) is one of the most widely used techniques in areas such as computer vision, speech recognition, and medical image processing that has solved many challenging problems. Unlike FCN (Fully Connected Network), CNN is a specialized network actively using the structure of local receptive fields that are suitable for extracting image features. Krizhevsky et al. applied deep CNN for ISBI (International Symposium on Biomedical Imaging) 2012 EM (Electron Microscopy) segmentation challenge, and they outperformed all other approaches in the competition, despite their network was not even being tailored to this particular segmentation task [17]. Gopalakrishnan et al. used transfer learning approach to detect pavement cracks by borrowing abilities from the other pre-trained networks [18]. While previous studies have used road images collected from special cameras, DNN can also be applied to various data acquired from DSLR camera, smartphone camera, and black-box installed in a vehicle since it shows good results even on low-resolution images [19,20]. Zhang et al. applied neural network for the first time to detecting cracks by building a very simple CNN architecture without using transfer learning, and showed better results than conventional SVM (Support Vector Machine) and boosting method [6]. After this, Pauly et al. improved the result of Zhang et al. by extending network architecture over the same dataset [8]. Zhang et al. even further improved previous result without using pooling layer by adding fully connected layer between convolution layers [21]. Bang et al. proposed an optimally configured encoder-decoder network for detecting cracks on a pixel-by-pixel basis in road images collected from black-box camera [7]. Carr et al. also obtained fairly good result from the CrackForest datasets by using RetinaNet which was based on FPN (Feature Pyramid Network) [22].

3. **Crack Management Process and Experiment Setups.** In this section, an overview of crack management process is given together with a description for experiment setups.

3.1. **Road crack management process.** The overall process for road crack management consists of four major stages, as shown in Figure 1, and is highly dependent on manual labor. At the data collection stage, vehicles equipped with road scanner periodically capture snapshots of the pavement while running the road, and send the snapshots to the central system. These transmitted data include original road image and location information. In the following crack detection stage, a number of workers manually mark the existence of cracks on the captured images, and each crack's bounding rectangle is also annotated on the image. The severity detection stage is also performed manually over the labeled cracks to determine their seriousness by calculating defected pixels' relative ratio within the bounding rectangle. Finally, in the repair decision stage, a repair plan that takes account of the availability of current resources is established for roads with

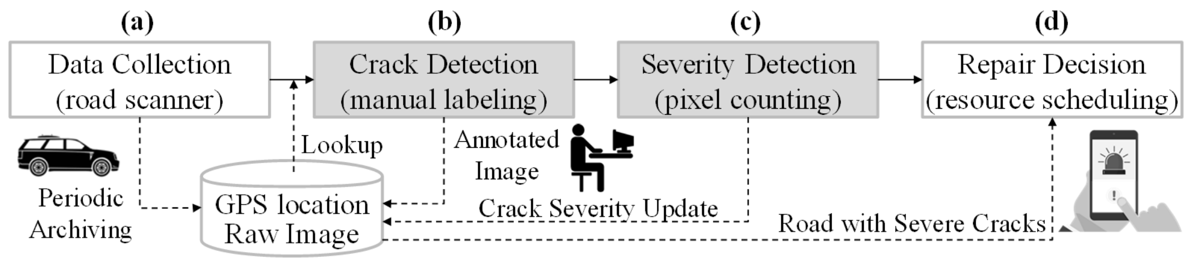

Figure 1. Overall road crack management process

severe damages. To facilitate effective handling of the entire crack management process, it is necessary to automate crack detection and severity detection phases, which take the longest time and incur the most processing cost. In this study, various experiments were conducted to change these labor intensive stages into automatic and integrated ones through neural network.

3.2. **Test model and experiment environment.** At first, authors have conducted a training based on Zhang et al.'s CNN model (5-CNN) that applied neural network to crack detection for the first time, while excluding dropout from the network [6]. After this, without any preprocessing to input data except z-normalization, authors have obtained better accuracies with our 7-CNN + BN model and well-known SqueezeNet model [24] than the accuracy of Zhang et al.'s original model for the same input data. Since the data used in our experiment is much more complex than the original dataset used in Zhang et al.'s work, authors have used deeper models than the original one. The models tested in this paper are summarized in Figure 2. You can see that BN (Batch Normalization) layer is added between convolution layer and activation layer in our CM (CNN Module). All experiments were done on the machine with Intel Xeon Silver 4116 CPU at 2.10 GHz running Ubuntu 16.04.5 (12 core/32GB RAM/GTX 1050Ti GPU) using Tensorflow and Keras high-API.

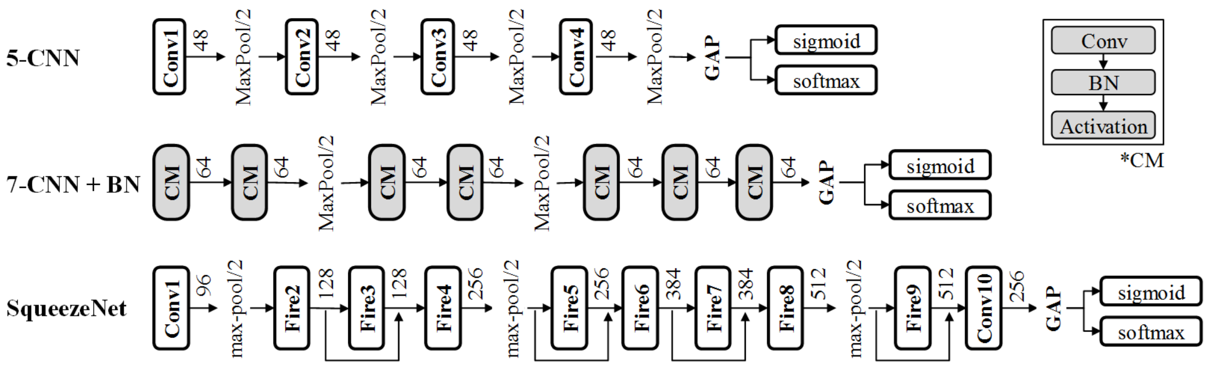

FIGURE 2. Tested model architecture

3.3. **Data and preprocessing.** The data used in the experiment are 640,000 road images acquired over 3 years (2015∼2017) from Korea. Each image has a resolution of 10,000 × 3,739, and has labels that indicate the type and severity of cracks. Labels are manually created for all 6 types of cracks which are AC (Alligator Crack), CN (Crack Construction Joint), TC (Transverse Crack), LC (Longitudinal Crack), PTCHG (patching) and POTHOLE. Input images for AC, CN, TC, and LC crack types are classified into 3 groups (HIGH, LOW, None), depending on the severity of crack type, respectively. For each of the above 4 types of cracks, 2,200 images are used for each severity group. That is, a total of 6,600 images are used for network training of AC, CN, TC, and LC crack types, respectively. Since the PTCHG and POTHOLE crack types neither have sufficient images nor distinguish crack severity, the inputs are classified into 2 groups (EXIST, None), and only 600 images are used for each group, which means that a total of 1,200 images are used for network training of PTCHG and POTHOLE crack types, respectively. White paint, commonly used for road signs, crosswalks, and lanes, is replaced with background pixels from the input image using thresholds. Log-scale conversion and pyramid down preprocessing are also applied to inputs before training. Since the resolution of original image is too high, authors have used low resolution images for experiments. There were concerns about distortion of crack shapes or deterioration of minor cracks while reducing the resolution. However, authors found a reasonable level of resolution that still kept the

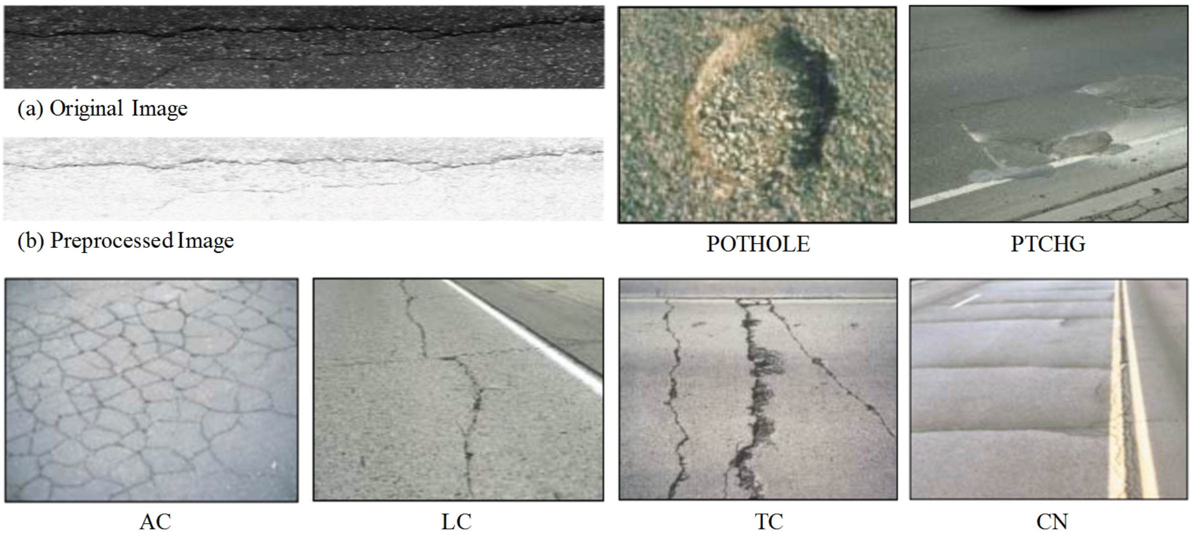

FIGURE 3. Sample input images and example crack types [23]

characteristics of major cracks from the input image without seriously affecting model accuracy at the experiment. In Figure 3, original input image and preprocessed image are given together with 6 sample images for each crack type.

4. **Experiment Result.** Our 7-CNN + BN network design has been largely inspired by the models of the previous two works [6,8]. Its network architecture is similar to the existing architecture, but the BN layer which shows good results in object detection field, is added between convolution layer and activation layer. Our 7-CNN + BN model is compared to two other well-known network models (i.e., 5-CNN and SqueezeNet) for 7 datasets. All the results are summarized in Table 1. For each dataset, best accuracies are highlighted with bold figures in the table. For 5-CNN model, all results are inferior to the other two models, except when using original dataset from [6]. This can be explained in terms of network depth and properties of the input image. Since the dataset in [6] is relatively simple and has less noise in the image than the other crack datasets, the performance of 5-CNN model that has less network depth than the other two models can be poor at noisy and complex input images. However, the 7-CNN + BN and SqueezeNet models that have enough network depth work very well on all crack datasets, with very similar performance results. Because the SqueezeNet has the least number of learning parameters and short execution time while showing the best result across all datasets, additional experiments were performed on SqueezeNet with 3 sets of input images with resolutions of 448 × 448, 224 × 224, and 250 × 100 to understand the effect of image resolution on accuracy. The final accuracies were 93.56%, 93.75% and 91.67% for each image resolution, and any significant accuracy degradation could not be found at this experiment. Based on this result, input images with 224 × 224 resolution that showed best accuracy were used throughout this research.

TABLE 1. Classification result from 3 test models over 7 crack datasets

Crack Dataset / Model	AC	CN	TC	LC	PTCHG	POTHOLE	Zhang et al. [6]
5-CNN	82.58%	71.29%	79.32%	84.02%	70.11%	75.00%	99.83%
7-CNN + BN	**94.22%**	79.10%	**93.41%**	**94.24%**	77.72%	**85.34%**	**99.90%**
SqueezeNet	93.75%	**80.68%**	89.58%	93.84%	**81.42%**	84.41%	99.85%

As shown in Table 1, it is relatively easy to detect a crack and its severity from individual crack dataset. The best accuracy for each crack type is always above 80% from the 7-CNN + BN and SqueezeNet model. When an experiment was conducted to determine the type of crack out of 6 crack types by mingling above crack dataset as an input, the resulting accuracy was just around 60%, which could not be used in real applications. This low accuracy of multiclass classification can be explained by intrinsic features of road damages that share similar patterns between different crack types, and thus bring about the confusion of network. The fact that an input can have multiple or redundant types of cracks inside an image also makes it hard for the network to clearly distinguish individual crack regions out of input image. It seems that a proper preprocessing or ad hoc network architecture is needed to achieve higher performance that detects multiple crack types from a single image.

In the whole training, activation layer primarily used ReLU activation function, and there was no significant change in the results while training with Leaky-ReLU. Learning rate was 0.0001 and the learning rate decay was set to 0.2. Adam optimizer which showed stable and good results was used in the experiment. In the SqueezeNet experiment, simply squeeze type with 0.5 SR (Squeeze Ratio) value was used. Since the outcome was to determine the class of input image as HIGH, LOW and None, one-hot encoding out of 3 digits was used as output format for AC, CN, TC, and LC crack datasets, and the categorical_crossentropy was used for loss function. For PTCHG and POTHOLE crack datasets, binary_crossentropy was used for loss function.

5. **Conclusion and Further Research.** Constant monitoring of road cracks and proper repair works are essential for traffic safety. However, many processes are still performed manually. Recently, researches have been actively performed to utilize CNN to automate pavement management process, especially for road crack detection. Since the results of many previous studies are based on limited datasets with well-cleansed sample images, researches using actual data, including various noises such as oil spot, white paint, and even wastes are critical for the actual automation of the process. In particular, different road conditions in individual countries can hugely affect the characteristics of input data, so experiments using real datasets suited to the actual working environment are more crucial. This study was conducted to obtain a more effective experimental environment and network configuration based on the real data using the well-known CNN models. Three models were tested to distinguish individual cracks from the input, and the best accuracy was roughly around 80∼90% for each crack type. Overall, our 7-CNN + BN model seems better than SqueezeNet model but its accuracy difference is somewhat minor. Under limited computing capabilities, SqueezeNet model that has more compact parameters is recommended. In case of SqueezeNet model, detection of multiple cracks in a single input image was also tested, but the accuracy remained at around 60%, so it is still not practical to automatically detect multiple cracks. The classification performance according to the resolution change of input image was also tested on SqueezeNet, and showed training time could be hugely reduced while retaining high level of accuracy. Finally, simple preprocessing methods such as thresholding, log-scale conversion, and pyramid down were very useful for improving the accuracy of crack detection. In the future, more advanced preprocessing is required to remove noise such as oil spots, manhole cover, and road lanes. To detect multiple cracks in a single image, object detection using a bounding box is required. Crack detection using low-resolution images of smartphone cameras needs to be studied to help field workers make on-site judgements.

Acknowledgment. This research was supported by the Technology Innovation Program (Graduate School of Management of Technology) funded by the Ministry of Trade, Industry and Energy (N0001613).

REFERENCES

[1] M. Gavilán, D. Balcones, O. Marcos, D. F. Llorca, M. A. Sotelo, I. Parra et al., Adaptive road crack detection system by pavement classification, *Sensors*, vol.11, no.10, pp.9628-9657, 2011.
[2] N. T. Sy, M. Avila, S. Begot and J. C. Bardet, Detection of defects in road surface by a vision system, *MELECON 2008 – The 14th IEEE Mediterranean Electrotechnical Conference*, pp.847-851, 2008.
[3] S. Varadharajan, S. Jose, K. Sharma, L. Wander and C. Mertz, Vision for road inspection, *IEEE Winter Conference on Applications of Computer Vision*, pp.115-122, 2014.
[4] E. Zalama, J. Gómez-García-Bermejo, R. Medina and J. Llamas, Road crack detection using visual features extracted by Gabor filters, *Computer-Aided Civil and Infrastructure Engineering*, vol.29, no.5, pp.342-358, 2014.
[5] Y. C. Tsai, V. Kaul and R. M. Mersereau, Critical assessment of pavement distress segmentation methods, *Journal of Transportation Engineering*, vol.136, no.1, pp.11-19, 2010.
[6] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, Road crack detection using deep convolutional neural network, *2016 IEEE International Conference on Image Processing (ICIP)*, pp.3708-3712, 2016.
[7] S. Bang, S. Park, H. Kim and H. Kim, Encoder-decoder network for pixel-level road crack detection in black-box images, *Computer-Aided Civil and Infrastructure Engineering*, 2019.
[8] L. Pauly, D. Hogg, R. Fuentes and H. Peel, Deeper networks for pavement crack detection, *Proceedings of the 34th ISARC*, pp.479-485, 2017.
[9] H. N. Koutsopoulos, I. E. Sanhouri and A. B. Downey, Analysis of segmentation algorithms for pavement distress images, *Journal of Transportation Engineering*, vol.119, no.6, pp.868-888, 1993.
[10] J. Canny, A computational approach to edge detection, *Readings in Computer Vision*, pp.184-203, 1987.
[11] I. Abdel-Qader, O. Abudayyeh and M. E. Kelly, Analysis of edge-detection techniques for crack identification in bridges, *Journal of Computing in Civil Engineering*, vol.17, no.4, pp.255-263, 2003.
[12] P. Subirats, J. Dumoulin, V. Legeay and D. Barba, Automation of pavement surface crack detection using the continuous wavelet transform, *2006 International Conference on Image Processing*, pp.3037-3040, 2006.
[13] T. S. Nguyen, S. Begot, F. Duculty and M. Avila, Free-form anisotropy: A new method for crack detection on pavement surface images, *2011 the 18th IEEE International Conference on Image Processing*, pp.1069-1072, 2011.
[14] T. S. Nguyen, M. Avila and S. Begot, Automatic detection and classification of defect on road pavement using anisotropy measure, *2009 the 17th European Signal Processing Conference*, pp.617-621, 2009.
[15] Y. Hu and C. X. Zhao, A novel LBP based methods for pavement crack detection, *Journal of Pattern Recognition Research*, vol.5, no.1, pp.140-147, 2010.
[16] Y. Shi, L. Cui, Z. Qi, F. Meng and Z. Chen, Automatic road crack detection using random structured forests, *IEEE Trans. Intelligent Transportation Systems*, vol.17, no.12, pp.3434-3445, 2016.
[17] A. Krizhevsky, I. Sutskever and G. E. Hinton, ImageNet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems*, pp.1097-1105, 2012.
[18] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary and A. Agrawal, Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection, *Construction and Building Materials*, vol.157, pp.322-330, 2017.
[19] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama and H. Omata, Road damage detection using deep neural networks with images captured through a smartphone, *arXiv preprint arXiv:1801.09454*, 2018.
[20] Y. J. Cha, W. Choi and O. Büyüköztürk, Deep learning-based crack damage detection using convolutional neural networks, *Computer-Aided Civil and Infrastructure Engineering*, vol.32, no.5, pp.361-378, 2017.
[21] A. Zhang, K. C. Wang, B. Li, E. Yang, X. Dai, Y. Peng et al., Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network, *Computer-Aided Civil and Infrastructure Engineering*, vol.32, no.10, pp.805-819, 2017.
[22] T. A. Carr, M. D. Jenkins, M. I. Iglesias, T. Buggy and G. Morison, Road crack detection using a single stage detector based deep neural network, *2018 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*, pp.1-5, 2018.
[23] J. S. Miller and W. Y. Bellinger, *Distress Identification Manual for the Long-Term Pavement Performance Program*, No. FHWA-HRT-13-092, Office of Infrastructure Research and Development, Federal Highway Administration, United States, 2017.

[24] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally and K. Keutzer, SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size, *arXiv preprint arXiv:1602.07360*, 2016.