Artificial Intelligence Research and Development
A. Cortés et al. (Eds.)

doi:10.3233/FAIA220342

A Curated Dataset for Crack Image Analysis: Experimental Verification and Future Perspectives

Ammar M. OKRAN [a,1], Mohamed ABDEL-NASSER [a,b], Hatem A. RASHWAN [a] and Domenec PUIG [a]

[a] *Department of Computer Engineering and Mathematics, Rovira i Virgili University, 43007 Tarragona, Spain*

[b] *Electrical Engineering Department, Aswan University, Egypt*

Abstract. Most crack image datasets are developed for crack segmentation or detection. They cannot be used to train a deep learning model to detect and segment cracks simultaneously. Most of existing datasets do not include a very accurate annotation. Besides, some crack images cannot be used to train deep learning models because of their inferior quality. In this paper, we propose a promising curated crack image dataset that allows the development of crack segmentation, detection, and classification on the same set of images simultaneously. There is no dataset for road crack that involves detection and segmentation tasks to the best of our knowledge. The current version of the curated database consists of 506 images derived from the RDD2020 dataset taken from multi-countries (Japan, Czech, and India). We use the curated dataset to build different deep learning-based crack detection and segmentation methods. Our experiments demonstrate that the proposed dataset yields promising results for crack detection and segmentation.

Keywords. Road Crack, Deep learning, Mask-RCNN, Object detection, Instance Segmentation

1. Introduction

Human visual inspection of cracks is a time-consuming operation that also comes at a higher expense. As a result, computer vision-based crack detection systems can provide reliable results at a reasonable cost. Crack detection and segmentation remain a difficult challenge because of the following factors: lighting variations, poor continuity, low contrast between cracks and backdrop (e.g., pavement), intensity inhomogeneity on crack regions, and similar crack shadows. Developing efficient deep learning-based crack segmentation and detection models requires an extensive crack image dataset, including images captured under the aforementioned conditions. Table 1 summarizes the widely used datasets for crack detection and segmentation in the last 5 years. As one can see, these datasets are developed for crack segmentation or detection. They cannot be used to train a deep learning model *to detect and segment cracks simultaneously*. Also, most of these

[1]Corresponding Author: E-mail: ammar.okran@urv.cat.

datasets do not include a very accurate annotation. Besides, some crack images cannot be used to train deep learning models because of their *very poor quality* (e.g., some images of RDD2020). There is *no dataset for road crack* that involves detection and segmentation tasks to the best of our knowledge.

In this paper, we propose a promising curated crack image dataset that allows the development of crack segmentation, detection, and classification on the same set of images simultaneously.

Table 1. Comparison between our curated crack image dataset and exiting ones.

Dataset	Year	Detection	Segmentation	Classification	Samples
GAPs v1 [1]	2017	×	×	✓	1969
RDD2018 [2]	2018	✓	×	✓	9053
GAPs v2 [3]	2019	✓	×	✓	2468
DeepCrack [4]	2019	✓	×	×	537
Angulo et al. [5]	2019	✓	×	✓	18034
RDD2020 [6]	2020	✓	×	✓	26620
Ours	2022	✓	✓	✓	506 (up-to-date)

2. Methodology

2.1. Data curation process

The curation process is shown in Figure 1. The curated dataset includes 506 images, carefully selected from the public dataset RDD2020 [6] which consists of 26,620 images of 3 countries (i.e., Japan, Czech, and India) taken by smartphone, with 850 road cracks and potholes. We selected images of acceptable quality and rejected the ones of bad quality. Also, when selecting the crack images, we tried keeping the same number of images per crack class (Longitudinal, Transverse, Alligator, and Pothole) to balance the dataset. An engineer has manually segmented each crack and drawn a bounding box around it. Quality control has been done on the generated annotations by AI engineers. Rejected annotations are repeated manually.

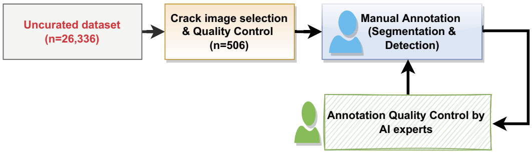

Figure 1. Data curation process.

2.2. Tested deep learning models for crack image analysis

In this study, we consider state-of-the-art deep learning models for crack segmentation and detection. In the task of *crack segmentation*, four segmentation models have been developed based on the Gated Skip Connections [7], High-Resolution Network (HR-Net), and Swin transformer [8], ConvNext [9], and the curated dataset. In turn, different *crack detection models* have been developed based on Mask R-CNN [10] and the curated dataset. Mask R-CNN [10] is a general framework for object instance segmentation that is simple to train, flexible, and adaptable. It recognizes objects in a photo while also creating a high-quality segmentation mask for each one.

3. EXPERIMENTS

Table 2 presents the results of different crack segmentation models trained and tested on the curated dataset. The developed crack segmentation models are HR-Net with fully connected layer (FC)—Fc+HRNet, UpperNet with Swin-Small as an encoder (UpperNet+Swin-Small), and UperNet with ConvNext-Base as an encoder (UperNet+ConvNext-Base), and Gated skip connections with ConvNext-small as an encoder (Gated skip connections+ConvNext-small). As one can see, the Gated skip connections crack segmentation model achieves the best segmentation results. It obtains an IoU of 48.8 and a Dice score of 61.98. The results of the UperNet+ConvNext-Base model are close to the Gated skip connections model. UpperNet+Swin-Small achieves the best precision.

Table 2. Performance of the segmentation models on the curated dataset.

Model	Acc	IoU	Acc	Fscore	Precision	Recall	Dice
Fc+HRNet	98.71	42.74	47.58	55.27	69.7	47.58	55.27
UpperNet+Swin-Small	**98.81**	43.94	48.27	56.58	**75.31**	48.27	56.58
UperNet+ConvNext-Base	98.71	48.48	55.1	61.84	72.61	55.1	61.84
Gated skip connections+ConvNext	98.8	**48.8**	**55.59**	**61.98**	72.08	**55.59**	**61.98**

Figure 2 shows quantitative results of the segmentation models, where each row shows a different crack—Longitudinal, Transverse, Alligator, and Pothole. As one can see, Fc+HRNet and UpperNet+Swin-Small under-segment the cracks. While both UpperNet+ConvNext and gated skip connections+ConvNext produce acceptable crack segmentation results in all shown examples. For crack detection, we develop different mod-

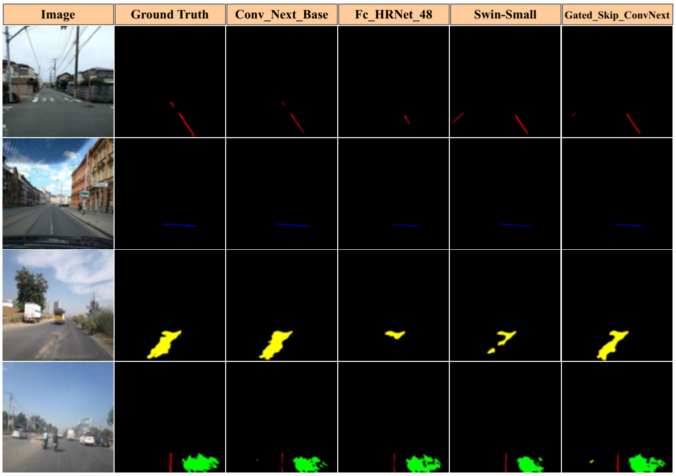

Figure 2. Segmentation results of different models. Each row presents a different crack—row1: Longitudinal Crack, row2: Transverse Crack, row3: Alligator Crack, and row4: Pothole.

els based on *Mask-RCNN*. Three models are shown here: standard *Mask-RCNN*, *Mask-RCNN_HRNet_32_1x*, and *Mask-RCNN_HRNet_40_1x*, where 32 and 40 stand for the width of the high-resolution convolution. It should be noted that the training of these models requires the bounding boxes of cracks and segmentation ground-truth masks, which are available in our curated dataset. Table 3 presents the results of the crack detection models trained on the curated dataset. Mask-RNN achieves a mean average precision (mAP) of 0.415 with the bounding box (i.e., *bbox_mAP*) and 0.392 with the segmentation branch (i.e, *segm_mAP*).

Table 3. Performance of the detection models on the curated dataset.

Model	bbox_mAP	segm_mAP
Mask-RCNN	0.415	0.392
Mask-RCNN_HRNet_32_1x	0.41	0.38
Mask-RCNN_HRNet_40_1x	0.402	0.394

4. Conclusion and Future work

This work introduced a promising curated crack image dataset that allows performing different crack image analyses simultaneously, like crack segmentation, detection, and classification, to be developed on the same collection of images. For crack segmentation and identification, we examined cutting-edge deep learning algorithms. Four segmentation models based on the Gated Skip Connections, High-Resolution Network, Swin transformer, ConvNext, and the curated dataset have been constructed for the task of crack segmentation. Future work will be focused on increasing the number of samples of the dataset and adding more crack classes.

References

[1] Markus Eisenbach, Ronny Stricker, Daniel Seichter, Karl Amende, Klaus Debes, Maximilian Sesselmann, Dirk Ebersbach, Ulrike Stoeckert, and Horst-Michael Gross. How to get pavement distress detection ready for deep learning? a systematic approach. In *2017 international joint conference on neural networks (IJCNN)*, pages 2039–2047. IEEE, 2017.

[2] Hiroya Maeda, Yoshihide Sekimoto, Toshikazu Seto, Takehiro Kashiyama, and Hiroshi Omata. Road damage detection and classification using deep neural networks with smartphone images. *Computer-Aided Civil and Infrastructure Engineering*, 33(12):1127–1141, 2018.

[3] Ronny Stricker, Markus Eisenbach, Maximilian Sesselmann, Klaus Debes, and Horst-Michael Gross. Improving visual road condition assessment by extensive experiments on the extended gaps dataset. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2019.

[4] Yahui Liu, Jian Yao, Xiaohu Lu, Renping Xie, and Li Li. Deepcrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing*, 338:139–153, 2019.

[5] Andres Angulo, Juan Antonio Vega-Fernández, Lina Maria Aguilar-Lobo, Shailendra Natraj, and Gilberto Ochoa-Ruiz. Road damage detection acquisition system based on deep neural networks for physical asset management. In *Mexican International Conference on Artificial Intelligence*, pages 3–14. Springer, 2019.

[6] Deeksha Arya, Hiroya Maeda, Sanjay Kumar Ghosh, Durga Toshniwal, Alexander Mraz, Takehiro Kashiyama, and Yoshihide Sekimoto. Deep learning-based road damage detection and classification for multiple countries. *Automation in Construction*, 132:103935, 2021.

[7] M Jabreel and M Abdel-Nasser. Promising crack segmentation method based on gated skip connection. *Electronics Letters*, 56(10):493–495, 2020.

[8] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10012–10022, 2021.

[9] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s. *arXiv preprint arXiv:2201.03545*, 2022.

[10] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.