Pavement Crack Mosaicking Based on Crack Detection Quality

Y. Yoon[a], S. Bang[a], F. Baek[b], and H. Kim[a]

[a]Department of Civil and Environmental Engineering, Yonsei University, Republic of Korea
[b]Industrial Technical Laboratory, Yonsei University, Republic of Korea
E-mail: yeosanyoon92@yonsei.ac.kr, bangdeok@yonsei.ac.kr, fbaek@yonsei.ac.kr, hyoungkwan@yonsei.ac.kr

Abstract –

A vehicle-mounted video camera, which is one of low-cost off-the-shelf devices, can be used economically for pavement crack monitoring. The pavement frames obtained by the video camera can be merged to form a mosaic image, from which road distress information can be extracted. However, quality of crack detection in the frames is different from one another. The different level of crack detection quality should be considered for accurate construction of crack mosaic. This paper proposes a new pavement crack mosaicking method based on quality of crack detection in each frame. A convolutional neural network is suggested as a way to evaluate the quality of crack detection in the video frames. The proposed method showed a promising mosaicking performance compared to other existing methods.

Keywords –

Convolutional Neural Network; Crack Detection Quality; Pavement Crack Mosaicking; Vehicle-mounted Camera

1 Introduction

Cracking is an important factor in evaluating pavement surface condition [1]. A crack can allow water to enter the pavement and cause potholes to develop. Because potholes can damage vehicles [2], it is crucial to perform pavement monitoring and timely maintenance [3]. Pioneering studies were conducted for automatic crack detection in 1990s [4,5,6]. Ever since, various inspection devices have been used for automatic pavement crack detection research, including a ground penetrating radar (GPR) [7], an unmanned aerial vehicle (UAV) [8], laser scanners [9], infrared spectrometers [10], and a red green blue depth (RGB-D) sensor [11]. Low-cost off-the-shelf devices have also been used for pavement crack detection research; examples include smartphones [12,13,14] and vehicle-mounted cameras [15,16,17].

There are several advantages to using a vehicle-mounted video camera, which is one of the low-cost off-the-shelf devices, for pavement monitoring. By using a device already installed in a general vehicle, it is not necessary to purchase an additional experimental device for data acquisition. Obtaining data with multiple vehicles can allow large areas to be observed in a short time. A pavement that is too long to be covered by a single image can be observed by making a mosaic from successive video frames [18,19]. However, quality of crack detection is different frame by frame. For generation of the mosaic from video data, the different level of crack detection quality should be considered.

In this paper, we propose a new pavement crack mosaicking method based on crack detection quality in images. A convolutional neural network plays a major role to evaluate the crack detection quality of each frame. The methodology and experiments are explained in Chapters 2 and 3, respectively, followed by conclusions in Chapter 4.

2 Proposed Methodology

Convolutional neural network (CNN), which has shown encouraging performance in recent years [20], was used for image data processing in this paper. An encoder-decoder network based on ResNet [21], which is the winner of ImageNet Large Scale Visual Recognition Challenge 2015 (ILSVRC2015), was used for the pixel-wise crack detection; the network was the preprocessing step for this study. An input image size for the crack detection network was 1920 x 1080. Input images were collected from a vehicle-mounted camera. The dataset consisted of 427 and 100 images for training and testing, respectively. Ground truth crack binary images were made by manual labelling. Pre-trained weights learned with ImageNet data were used as the initial weights of the crack detection network.

2.1 Crack Detection Quality Prediction

The core of the proposed method is a crack detection quality prediction network (CDQ-Net). The CDQ-Net calculates a crack detection quality (CDQ) of each frame. An f1-score calculated by the crack detection network is considered as the CDQ of a frame. The CDQ-Net consists of convolutional layers, a global average pooling layer, and fully connected layers. Raw input images (2560 x 1440) obtained from a vehicle-mounted camera included non-road area. By eliminating the non-road area, the amount of computation could be reduced. On the raw input images, only the road area were extracted to result in the image size of 1536 x 540. At the convolutional layers, the filter size and the stride were 7 x 7 and 4, respectively.

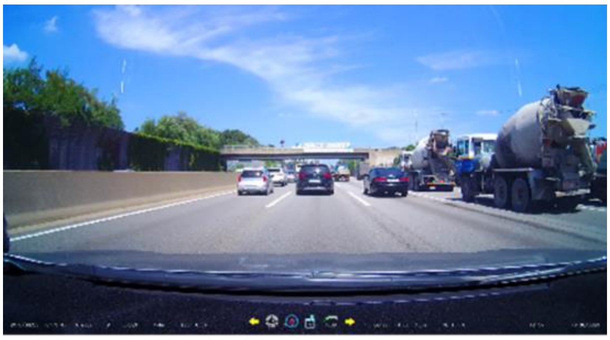
The motivation why the CDQ-Net is a shallow network and includes a big filter size and stride is as follows. Most of crack shapes are thin lines, which are expected to be more related to low-level features than high-level features. Low-level features are easy to be generated by a shallow network. The vehicle-mounted camera has a high resolution (2560 x 1440) compared with many applications with ImageNet (256 x 256). If a 3 x 3 filter is used, as in a typical CNN model, it may be difficult to capture a long line feature such as cracks. A relatively larger filter sizes and strides can be appropriate for the purpose of this study.

To avoid overfitting problems, batch normalization is adopted right after each convolution and before activation. The activation function is ReLU [22]. Dropout at 50% was applied to each fully connected layer. The number of channels in each convolutional layer is determined to be similar to ResNet [21]. The number of channels in the first convolutional layer is 64 and the number of channels in the second convolutional layer is 128. In the same way as above, the number of channels in a convolutional layer is twice the number of channels in the previous convolutional layer. Except for the last fully connected layer, each fully connected layer has the same number of neurons as the number of channels in the last convolutional layer. The number of neurons of the last fully connected layer is 1 because the output of the CDQ-Net should be a single scalar value.

2.2 Pavement Crack Mosaicking

Mosaicking pavement crack information from consecutive video frames was conducted in the following steps. A CDQ of each frame was calculated by the CDQ-Net. A crack probability, which is the output just before binarization, was obtained for each pixel of a frame through the crack detection network. An example of a crack probability image is shown in Figure 1. The crack probability image, which is an oblique-view image, was transformed into a top-view image. Consecutive crack probability images were then overlapped based on the geometric relationships, which is explained in the next paragraph. Crack probabilities of overlapped pixels corresponding to a specific pixel location were weighted averaged according to the CDQ of each frame. The weighted averaged crack probability is the crack probability of the mosaic. Next, a binary crack mosaic image was derived from the crack probability mosaic image based on a specific threshold value. The optimal threshold value was to maximize a mean of f1-scores for binary crack detection of multiple mosaic images. An example of a pavement crack mosaic is shown in Figure 2.

The geometric relationship between two consecutive frames was calculated in the following steps. Oblique-view frames obtained by a vehicle-mounted camera were first transformed into top-view frames. A top-view frame included lane markings on both sides. The center curve in the previous top-view frame was then estimated based on the lane markings. Geometric relationship candidates between the two consecutive frames were derived according to the position and tangential angle of the center curve. The optimal geometric relationship was finally selected when the mean squared intensity difference between the two consecutive frames was minimum.



(a) An image obtained by a vehicle-mounted camera

(b) A crack probability image derived from the upper image

Figure 1. An example of pixel-wise crack probability prediction

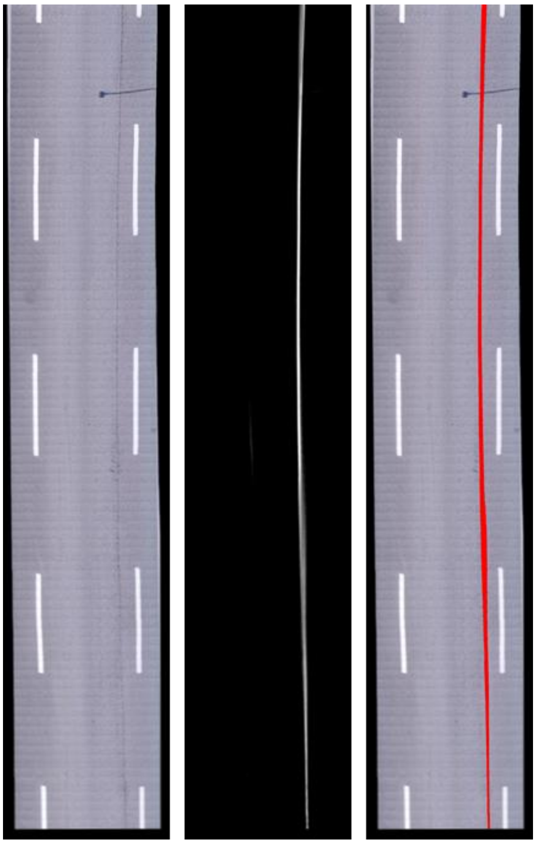


Figure 2. An example of pavement crack mosaicking (from left to right: pavement mosaic, weighted averaged crack probability mosaic, and binary crack mosaic)

3 Experiments and Results

The experimental environment was as follows. All of the image data were obtained by a dash-board camera Qrontech LK-919 QAD. The fps (frames per second) was 30 and the resolution was 2560 x 1440. Experiments were performed in a desktop with GeForce GTX 1080 Ti GPU and Intel Core i7-7700 CPU. The operating system of the desktop was Ubuntu 16.04.

3.1 Crack Detection Quality Prediction

The number of convolutional layers and the number of fully connected layers were changed to find the optimal CDQ-Net structure. The image dataset consisted of 270, 66, and 83 images for training, validation, and testing, respectively. Ground truth binary crack images were made by manual labelling. A CDQ of each frame was calculated through the crack detection network. Epochs were 500 for a majority of the CDQ-Nets, but larger epochs were used when the value of the loss function did not converge sufficiently in the validation process. Batch size was fixed to 4. The network with the lowest MAE (Mean Absolute Error) was chosen as the optimal CDQ-Net. The optimal CDQ-Net consisted of four convolutional layers and one fully connected layer. The experimental results are summarized in Table 1.

Table 1. MAE and epochs of various CDQ-Net structures: MAE (Epochs)

No. of convolutional layers \ No. of fully connected layers	1	2	3
2	8.61% (2000)	11.01% (1000)	9.99% (3000)
3	10.30% (500)	8.59% (1000)	10.98% (2000)
4	8.26% (500)	8.59% (500)	8.95% (1500)
5	10.11% (500)	10.01% (500)	10.72% (500)

The proposed CNN model was compared with a representative CNN model, the ResNet, in terms of its capability to predict the CDS. For the direct comparison, the experimental environment was set to the same conditions. Pre-trained weights learned with ImageNet data were used as the initial weights of ResNet50 [21]. In terms of MAE and duration for training, CDQ-Net's performance was superior to ResNet50. The experimental results are summarized in Table 2.

Table 2. MAE and duration on CDQ-Net and ResNet50

Performance	CDQ-Net	ResNet50
MAE (Epochs)	8.26% (500)	9.23% (1000)
Duration (Unit)	0.39 (hours)	9.77 (hours)

3.2 Pavement Crack Mosaicking

The proposed method for pavement crack mosaicking was tested using LOOCV (Leave-One-Out Cross-Validation) [23]. In the LOOCV, one mosaic image was selected for testing, while the others were selected for the determination of the optimal threshold value. The threshold value was selected to binarize each pixel of the test mosaic image into crack or non-crack category. Based on the ground truth images, the performance was measured by the f1-score. This process was repeated for every mosaic image of the dataset. In the end, the f1-scores were averaged to provide a scalar value of performance.

The performance of the proposed method was compared with two intuitively obvious methods: overwriting and accumulating. The overwriting method is to use the crack information of the new frame to its fullest extent, ignoring the crack information of the existing mosaic. The accumulating method is to conserve every crack information of the existing mosaic to result in the OR operation with the new information. The two methods were measured by the average f1-score.

The experimental dataset for pavement crack mosaicking consisted of five videos. Each video has 90 frames. The ground truth binary crack mosaic images were made by manual labelling. In terms of f1-score for crack detection, the performance of the proposed pavement crack mosaicking method was superior to the other mosaicking methods. The experimental results are summarized in Table 3. Examples of pavement crack mosaicking results are shown in Figure 3.

Table 3. Performance of the crack mosaicking methods

Mosaicking method	Precision	Recall	F1-score
CDQ	0.5175	0.6748	0.5838
Overwriting	0.8251	0.2634	0.3969
Accumulating	0.3427	0.8543	0.4870

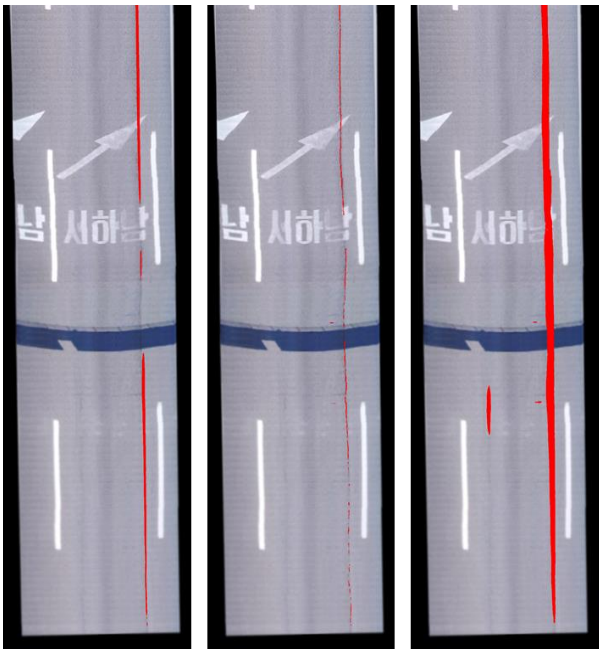


Figure 3. Examples of pavement crack mosaicking results (from left to right: CDQ based method, overwriting method, and accumulating method)

4 Conclusions

This paper presented the pavement crack mosaicking method based on crack detection quality. The proposed network, which is to predict a crack detection quality of each frame, performed better than ResNet50. The proposed pavement crack mosaicking method, which is based on the crack detection quality of each frame, showed better performance than the two intuitively obvious methods: overwriting and accumulating. The proposed method, when strengthened by a large number of crack image data, is expected to produce a significantly improved mosaic map of pavement cracks.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. 2018R1A6A1A08025348) and by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (No. 2018R1A2B2008600).

References

[1] Tong Z., Gao J., Han Z., and Wang Z. Recognition of asphalt pavement crack length using deep convolutional neural networks. *Road Materials and Pavement Design*, 19(6):1334–1349, 2018.

[2] Akagic A., Buza E., and Omanovic S. Pothole detection: An efficient vision based method using RGB color space image segmentation. In *40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 1104–1109, Opatija, Croatia, 2017.

[3] Walker D. Understanding asphalt pavement distresses – five distresses explained. On-line: http://asphaltmagazine.com/understanding-asphalt-pavement-distresses-five-distresses-explained, Accessed: 26/01/2019.

[4] Haas C., and Hendrickson C. Computer-based model of pavement surfaces. *Transportation Research Record*, 1260: 91–98, 1990.

[5] Haas C. Evolution of an automated crack sealer: a study in construction technology development. *Automation in construction*, 4(4): 293–305, 1996.

[6] Kim Y. S., Haas C. T., and Greer R. Path planning for machine vision assisted, teleoperated pavement crack sealer. *Journal of Transportation Engineering*, 124(2): 137–143, 1998.

[7] Tong Z., Gao J., and Zhang H. Recognition, location, measurement, and 3D reconstruction of concealed cracks using convolutional neural

networks. *Construction and Building Materials*, 146: 775–787, 2017.

[8] Zakeri H., Nejad F. M., and Fahimifar A. Rahbin: A quadcopter unmanned aerial vehicle based on a systematic image processing approach toward an automated asphalt pavement inspection. *Automation in Construction*, 72: 211–235, 2016.

[9] Guan H., Li J., Yu Y., Chapman M., and Wang C. Automated road information extraction from mobile laser scanning data. *IEEE Transactions on Intelligent Transportation Systems*, 16(1): 194–205, 2015.

[10] Miraliakbari A., Hahn M., and Maas H. G. Development of a multi-sensor system for road condition mapping. In *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(1): 265–272, Denver, United States, 2014.

[11] Jahanshahi M. R. Jazizadeh, F., Masri S. F., and Becerik-Gerber B. Unsupervised approach for autonomous pavement-defect detection and quantification using an inexpensive depth sensor. *Journal of Computing in Civil Engineering*, 27(6): 743–754, 2012.

[12] Mertz C., Varadharajan S., Jose S., Sharma K., Wander L., and Wang J. City-wide road distress monitoring with smartphones. *In Proceedings of 21st World Congress on Intelligent Transport Systems (ITSWC)*, pages 1–9, Michigan, United States, 2014.

[13] Zhang L., Yang F., Zhang Y. D., and Zhu Y. J. Road crack detection using deep convolutional neural network. In *Proceedings of International Conference on Image Processing (ICIP)*, pages 3708–3712, Arizona, United States, 2016.

[14] Bang S., Park S., Kim H., Yeo-san Y., and Kim, H. A deep residual network with transfer learning for pixel-level road crack detection. In *Proceedings of the 35th International Association for Automation and Robotics in Construction (ISARC)*, pages 753–756, Berlin, Germany, 2018.

[15] Radopoulou S. C., Brilakis I., Doycheva K., and Koch C. A framework for automated pavement condition monitoring. In *Proceedings of the 2016 Construction Research Congress (CRC)*, pages 770–779, San Juan, Puerto Rico, 2016.

[16] Yan W. Y., and Yuan X.-X. A low-cost video-based pavement distress screening system for low-volume roads. *Journal of Intelligent Transportation Systems*, 22(5): 376–389, 2018.

[17] Park S., Bang S., Kim H., and Kim H. Patch-based crack detection in black box road images using deep learning. In *Proceedings of the 35th International Association for Automation and Robotics in Construction (ISARC)*, pages 757–760, Berlin, Germany, 2018.

[18] Geiger, A. Monocular road mosaicing for urban environments. In *IEEE Intelligent Vehicles Symposium*, pages 140–145, Xi'an, China, 2009.

[19] Sikirić I., Brkić K., and Šegvić S. Recovering a comprehensive road appearance mosaic from video. In *33rd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 755–759, Opatija, Croatia, 2010.

[20] Gu J., Wang Z., Kuen J., Ma L., Shahroudy A., Shuai, B., Liu T., Wang X. Wang G. Cai J., and Chen T. Recent advances in convolutional neural networks. *Pattern Recognition*, 77: 354–377, 2018.

[21] He K., Zhang X., Ren S., and Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pages 770–778, Las Vegas, United States, 2016.

[22] Nair V., Hinton G. E. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814), Haifa, Israel, 2010.

[23] Stone M. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society*. Series B (Methodological), 36(2): 111–147, 1974.