UAV-Based Pavement Crack Detection Using Deep Convolutional Neural Networks

Ahmed Elamin , Ahmed El-Rabbany, Canada

Keywords: UAV, crack detection, deep convolutional network, classification, segmentation.

SUMMARY

Pavement maintenance and rehabilitation are vital to safe driving. Conventional manual crack detection is costly, time-consuming, and subjective. Recently, unmanned aerial systems (UASs) have shown great potential for pavement crack detection and monitoring. This study proposes a UAV-based pavement crack detection model based on a deep convolutional neural network (DCNN). This model combines the advantages of image classification and segmentation of VGG-19 and U-net DCNNs. First, the crack classification model VGG-19 is applied to classify the crack sample images into two classes, namely crack and non-crack. Then, the images classified as crack are fed into the segmentation network U-net to precisely segment the pavement cracks. For the purpose of this research, a DJI Matrice-600 UAV equipped with a Sony a7Rii camera was used for data collection. The classification model VGG-19 achieved 100% accuracy in classifying the images into crack and non-crack. The DCNN for segmentation used in this study is an improved version of the U-net, where the convolution blocks are replaced with residual blocks inspired by ResNet. To improve the segmentation results, the network was pretrained using the CRACK500 dataset due to the limited number of available crack samples on the collected data. Then to optimize the hyperparameters, the transfer learning technique was used to train the network on the crack samples. Following the optimization, the pavement crack segmentation accuracy was improved by about 10%. The proposed approach showed that it can provide accurate classification and segmentation for pavement cracks from UAV-based images, which can improve the detection and monitoring of pavement conditions.

UAV-Based Pavement Crack Detection Using Deep Convolutional Neural Networks

Ahmed Elamin , Ahmed El-Rabbany, Canada

1. INTRODUCTION

Pavement distress is one of the main factors that affect road safety. Up-to-date and accurate detection of pavement distresses is essential for pavement maintenance. Cracks are the initial index of various types of pavement damage. Pavement cracks affect the pavement appearance and driving comfort. Additionaly, it can escalate to cause pavement structural distress, which shortens the overall service life of the pavement (Zhang and Li, 2015, Dong et al., 2016). Therefore, early crack detection can decrease pavement maintainance costs and ensure safe driving conditions. Generally, crack detection and maintenance rely on manual detection, which is time-consuming, less accurate, and accompanies with associated risks (Cao, 2014). Recently, pavement crack image acquisition has taken advantage of more advanced techniques, such as unmanned aerial systems (UASs). UASs have been widely tested for pavement monitoring, in order to collect high-resolution images with low cost, flexibility, and without affecting normal traffic flow (Ersoz et al., 2017, Pan et al., 2018). For example, in (Ersoz et al., 2017), a crack identification method was proposed, which used a support vector machine (SVM) model from UAV-based RGB images. They classified the images as a crack or non-crack and reached 97% accuracy.

Several researchers have investigated crack extraction from images (Liu and Zhang, 2012, Oliveira and Correia, 2014, Yang et al., 2019). Traditional crack detection methods, such as the ones proposed in (Wang and Wu, 2014) and (Ren et al., 2015) were designed for a specific scenario. However, these models would not generalize to new datasets. In addition, these models suffer from a lack of automation. Recently, deep convolutional neural networks (DCNNs) have been widely applied in the field of crack classification and segmentation (Xu et al., 2008, Oliveira and Correia, 2014, Zou et al., 2018). Using deep learning techniques improved the accuracy and efficiency of crack detection (Silva and Lucena, 2018). DCNNs approaches for pavement crack detection can be divided into two main types, namely crack classification and crack segmentation. For crack classification, the output of the DCNN for each input image is a label. The input images might be classified as crack and non-crack or classified according to the types of pavement cracks, such as transversal, longitudinal, and alligator. For example, Zhang et al. (2016) proposed a road crack recognition technique based on deep learning. A deep CNN was trained using supervised learning to classify the images as crack and no-crack. Li et al. (2019) proposed a pavement crack classification method based on a DCNN using 5966 images taken at different angles and distances for training. They were able to detect cracks of various categories, including lateral crack, longitudinal crack, alligator crack, and pothole. Sun et al. (2020) proposed an approach to classify pavement expansion cracks with an improved Faster R-CNN. The aforementioned studies only classified pavement cracks, which identify their location in the images but cannot quantify crack characteristics, such as crack width.

On the other hand, for crack segmentation, a label is assigned to each pixel value within an image. Zhang et al. (2017) proposed CrackNet, which consisted of two fully connected layers, two convolution layers, and one output layer. However, their approach utilized a feature extractor as preprocessing step to generate feature maps and then feeded them as input to CrackNet, which is inefficient. Jenkins et al. (2018) proposed a DCNN for crack segmentation based on U-net, which consisted of two main parts: encoder and decoder (Ronneberger et al., 2015). The proposed network was tested using the crack forests detection dataset (CFD), which was created by (Shi et al., 2016) and is composed of 118 images. They achieved 92%, 82%, and 87% for precision, recall and F-measure, respectively. The fast pavement crack detection network (FPCNet) was proposed by Liu et al. (2019), which used encoder-decoder architecture networks for pixel-based crack segmentation. FPCNet was trained and tested using the CFD dataset and achieved an F-measure of 95%.

Most of the previous studies used datasets collected manually, such as the CRACK500, which was collected using a smartphone (Zhang et al., 2016). In addition, those studies addressed a single function only, which is classification or segmentation, and most of the classified cracks were roughly positioned using a bounding box. Consequently, their results cannot be used for evaluating road conditions. Given the aforementioned problems, this research proposes a deep convolutional neural network fusion model for pavement crack detection using UAV-based images. This model is capable of classifying the input images to crack and non-crack, and then the detected crack images can be segmented and used to calculate the geometric parameters of the crack. This paper is organized as follows. Section II, where the proposed methodology is introduced. The system used, the data acquisition and processing are explained in Section III. Section IV discusses the experimental results, and Section V concludes this work.

2. PROPOSED CRACK DETECTION METHOD

The images were first processed to create a georeferenced orthomosaic image of the area. To train the DCNN models, a total of 30 crack and 30 non-crack samples were created as subset images of the orthomosaic image. The images were classified to crack and non-crack and then the detected crack images were segmented. The DCNN used for crack classification is the VGG-19 model. This model was proposed by (Simonyan and Zisserman, 2014), which comprises 19 layers, 16 convolutional and 3 fully-connected layers with a total of 138 million parameters. The convolutional layers used a filter size of 3×3 and a stride and pad size of 1 pixel. The small kernel size decreases the number of parameters and enables them to cover the entire image. A 2×2 max pooling operation with a stride of 2 is considered. This model ranked first in positioning and second in classification at the 2014 ILSVRC (Ali et al., 2021). The DCNN used for crack segmentation is the improved version of the U-net proposed by (Lau et al., 2020). In this network, the convolution blocks were replaced with residual blocks inspired by ResNet (He et al., 2016). The overall methodology flowchart is shown in Figure 1.

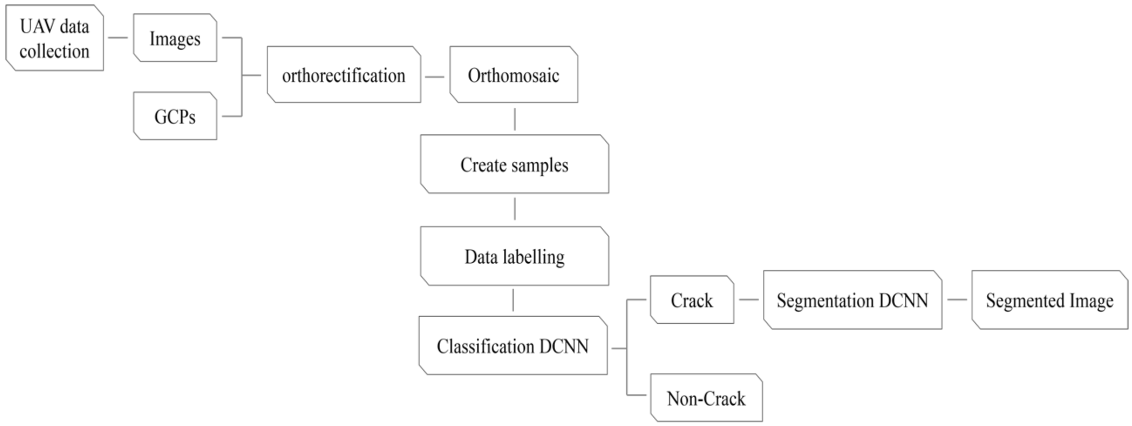

Figure 1. Flow chart of the proposed methodology.

3. DATA ACQUISITION AND PREPROCESSING

In this study, a UAV was used as the platform for imagery data acquisition. The UAV consists of a DJI Matrice 600 Pro carrying a Sony α7R II RGB camera. The Sony α7R II camera has a 42MP spatial resolution with a 7952×5304 image size. Figure 2 shows the top view of the flight trajectory over the area of interest. A total of 243 RGB images were captured. Twelve ground control points (GCPs) were used to georeference the collected images, as there was a problem the onboard GNSS system. The images were processed with the 12 GCPs using Pix4D mapper software to create a georeferenced orthomosaic image, as shown in Figure 3. To train the segmentation model, a total of 30 crack samples were labelled using the automated ground truth labeller application in the Matlab software

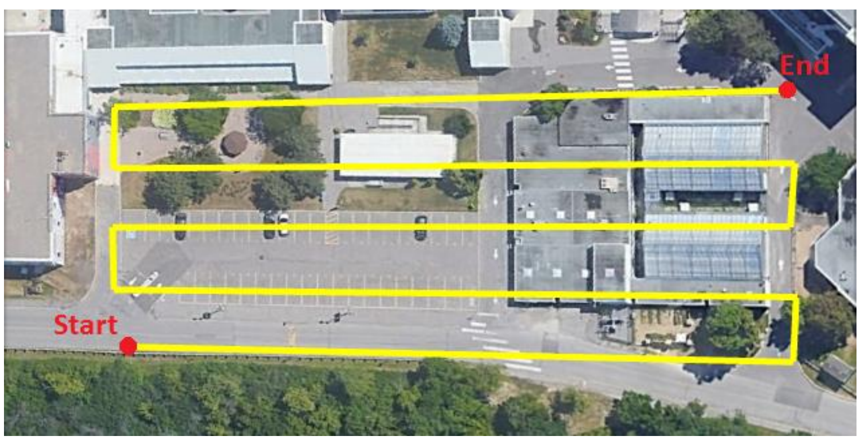

Figure 2. UAS flight lines over the area of interest.

Figure 3. Orthomosaic image of the area.

4. RESULTS AND DISCUSSION

In order to train and evaluate the VGG-19 network, the 60-image samples were randomly divided into 70% training, 15% validation, and 15% testing images. Figure 4 shows the loss function curves obtained by training network on the 60-image samples in the VGG-19 model. The orange line represents the training loss curve, while the blue line represents the validation loss curve. The overall accuracy reached 100%. Figure 5 shows the confusion matrix. The 9 images of the testing dataset were 4 non-crack and 5 crack images. The gradient-weighted class activation mapping (Grad-CAM) (Selvaraju et al., 2017) visualizations of the 4 non-crack images and the 5 crack images are shown in Figures 6 and 7, respectively. The Grad-CAM visualizations of the VGG-19 model predictions revealed that the model had learned to detect the shape of the crack correctly.

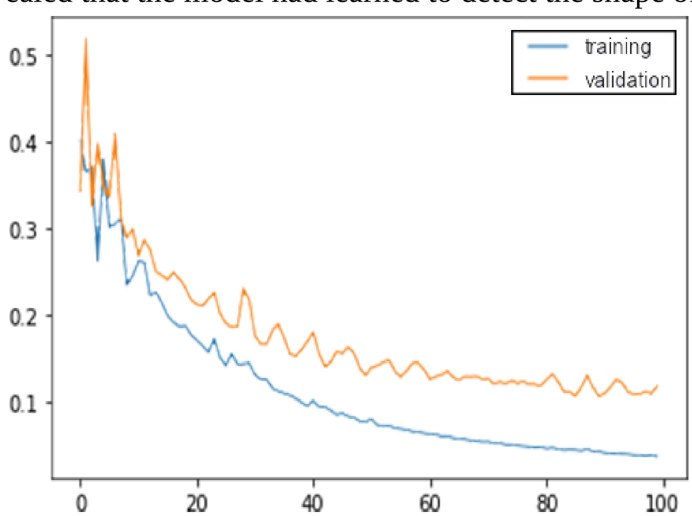

Figure 4. Training and validation loss curves for the classification network VGG-19.

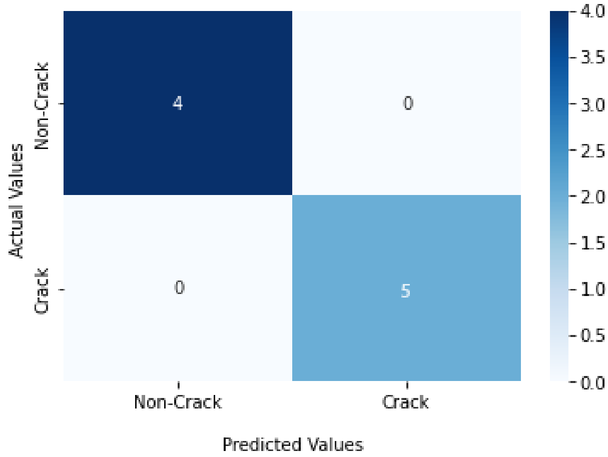

Figure 5. The confusion matrix for the testing dataset.

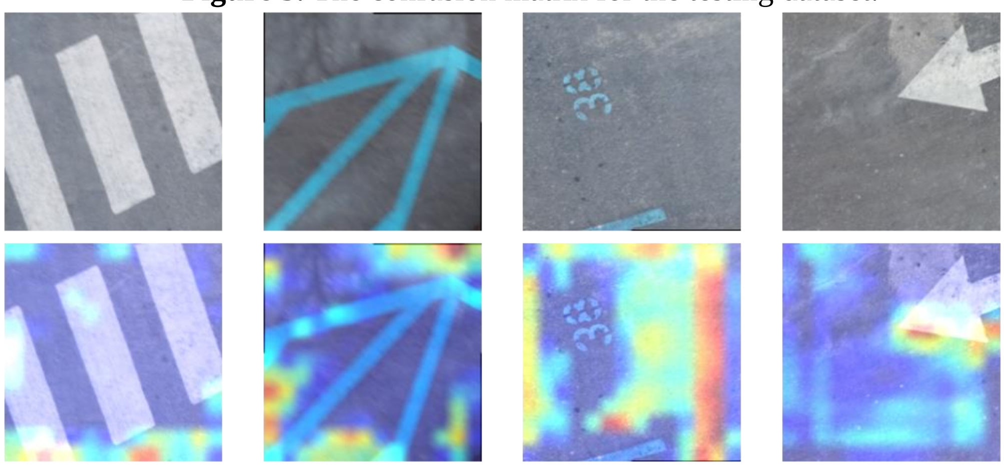

Figure 6. Grad-CAM visualizations of the non-crack images.

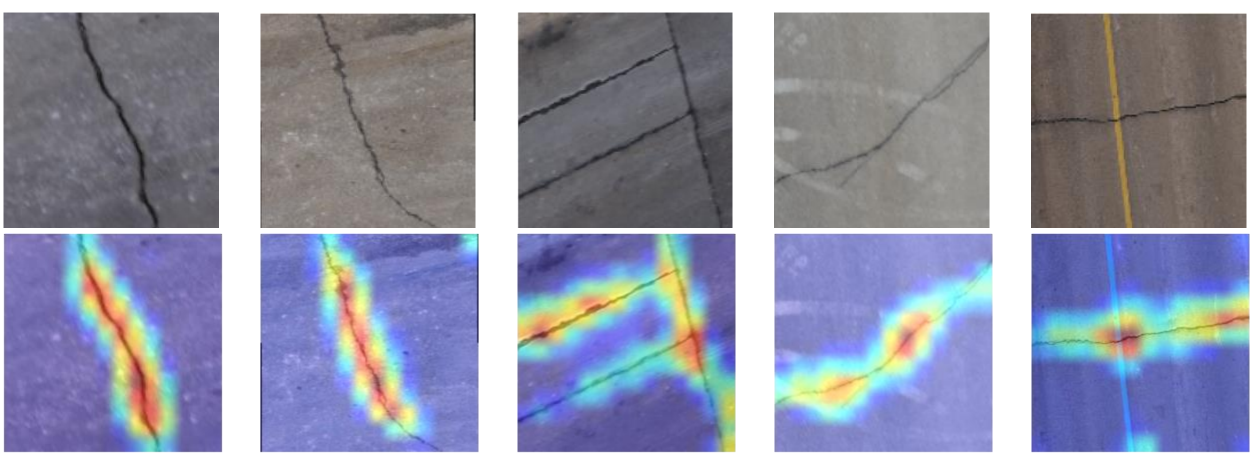

Figure 7. Grad-CAM visualizations of the crack images.

To carry out the segmentation task, the 30 crack image samples were divided into 20 training, 5 validation, and 5 testing images. Where the 5 testing images are the same testing images of the classification model. The network was initially trained and tested on the 30-image dataset. However, the obtained segmentation accuracy was relatively low due to the limited data size. Therefore, to optimize the network and to improve the segmentation results, the network was pretrained using the CRACK500 dataset. Figure 8 shows the training and validating loss curves and Table 1 shows the performance indicators obtained by training the network on the CRACK500 dataset. Figure 9 shows the prediction results of the testing dataset. Table 2 shows the performance indicators obtained after the optimization. Optimizing the segmentation network improved the pavement crack segmentation accuracy by about 10%. As shown in Table 2, the obtained precision, recall, and F-measure were 77.48%, 87.66%, and 82.26%, respectively.

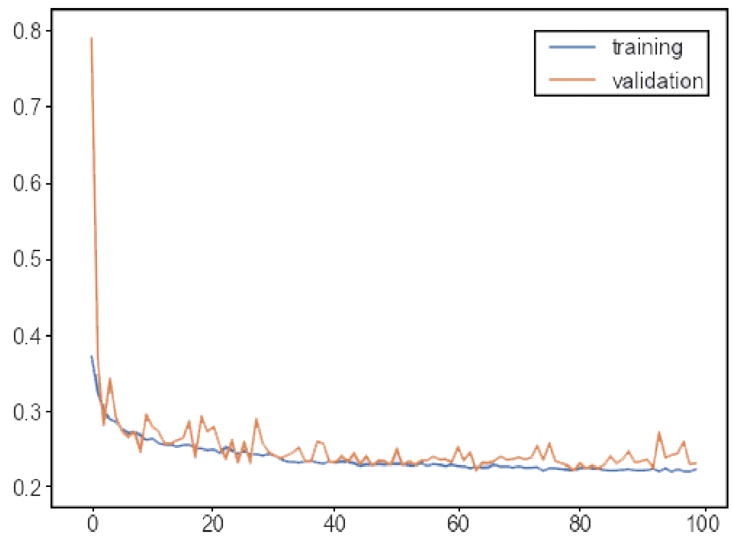

Figure 8. Training and validation loss curves for the pretrained network.

Table 1. The indicator parameters for the pretrained network on the Crack500 dataset.

Dataset	Recall(%)	Precision(%)	F-measure(%)
Crack500	77.54	81.71	79.57

Table 2. The indicator parameters for the segmentation results.

Combination	Recall(%)	Precision(%)	F-measure(%)
RGB	87.66	77.48	82.26

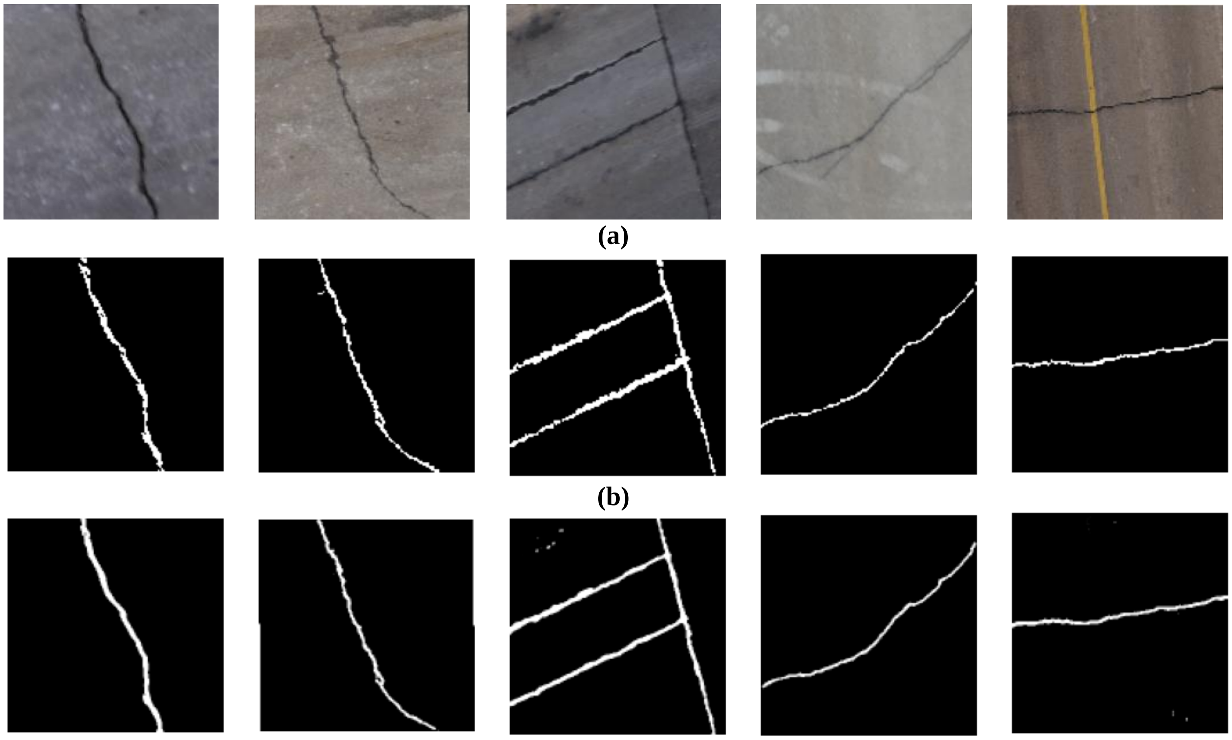

Figure 9. Prediction results of the segmentation model: (a) pavement crack images, (b) ground truths, (c) prediction

5. CONCLUSION

In this paper, a pavement crack detection model was proposed based on deep convolutional neural network models using UAV-based high spatial resolution camera images. The proposed model fuses crack classification and crack segmentation models, where the input images are first classified to crack and non-crack. Subsequently, the crack images are segmented by the segmentation model. To train the network, 60 crack and non-crack samples were extracted as a subset of the orthomosaic image. To train and evaluate the model, the samples were divided into 70% training, 15% validation, and 15% testing images. The VGG-19 network was used to classify the images into two classes, namely crack and non-crack. The classification accuracy reached 100%. An improved version of the U-net was used for crack segmentation. Initially, the accuracy was relatively low because of the limited data size. To improve the network, it was pretrained using the CRACK500 dataset, which improved the pavement cracks segmentation accuracy by about 10%. To assess the network, precision, recall, and F-measure were calculated in the test dataset. The precision, recall, and F-measure obtained are 77%, 87%, and 82%, respectively. The proposed model showed good performance for pavement crack detection, which can achieve accurate classification and segmentation.

6. ACKNOWLEDGMENTS

This research is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Ryerson University. The authors would like to thank Dr. Akram Afifi, a professor at the Faculty of Applied Sciences and Technology, Humber College, for providing the datasets used in this research.

7. REFERENCES

ALI, L., ALNAJJAR, F., JASSMI, H. A., GOCHOO, M., KHAN, W. & SERHANI, M. A. 2021. Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors,* 21**,** 1688.

CAO, J. 2014. Research on the Technology of Automatic Pavement Crack Detection Based on Digital Image Processing. Chang'an University, Xian, China.

DONG, Z., YE, S., GAO, Y., FANG, G., ZHANG, X., XUE, Z. & ZHANG, T. 2016. Rapid detection methods for asphalt pavement thicknesses and defects by a vehicle-mounted ground penetrating radar (GPR) system. *Sensors,* 16**,** 2067.

ERSOZ, A. B., PEKCAN, O. & TEKE, T. Crack identification for rigid pavements using unmanned aerial vehicles. IOP Conference Series: Materials Science and Engineering, 2017. IOP Publishing, 012101.

HE, K., ZHANG, X., REN, S. & SUN, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 770-778.

JENKINS, M. D., CARR, T. A., IGLESIAS, M. I., BUGGY, T. & MORISON, G. A deep convolutional neural network for semantic pixel-wise segmentation of road and pavement surface cracks. 2018 26th European Signal Processing Conference (EUSIPCO), 2018. IEEE, 2120-2124.

LAU, S. L., CHONG, E. K., YANG, X. & WANG, X. 2020. Automated pavement crack segmentation using u-net-based convolutional neural network. *IEEE Access,* 8**,** 114892-114899.

LI, J., ZHAO, X. & LI, H. Method for detecting road pavement damage based on deep learning. Health Monitoring of Structural and Biological Systems XIII, 2019. SPIE, 517-526.

LIU, W., HUANG, Y., LI, Y. & CHEN, Q. 2019. FPCNet: Fast pavement crack detection network based on encoder-decoder architecture. *arXiv preprint arXiv:1907.02248*.

LIU, Z. & ZHANG, Y. 2012. Detection technology of damaged road surface based on PDE and wavelet analysis. *Microcomputer and Application,* 31**,** 35-37.

OLIVEIRA, H. & CORREIA, P. L. CrackIT—An image processing toolbox for crack detection and characterization. 2014 IEEE international conference on image processing (ICIP), 2014. IEEE, 798-802.

PAN, Y., ZHANG, X., CERVONE, G. & YANG, L. 2018. Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11**,** 3701-3712.

REN, L., ZHIGANG, X. & ZHAO, X. 2015. Pavement crack connection algorithm based on Prim minimum spanning tree. *Computer Engineering,* 41**,** 31-36.

RONNEBERGER, O., FISCHER, P. & BROX, T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention, 2015. Springer, 234-241.

SELVARAJU, R. R., COGSWELL, M., DAS, A., VEDANTAM, R., PARIKH, D. & BATRA, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision, 2017. 618-626.

SHI, Y., CUI, L., QI, Z., MENG, F. & CHEN, Z. 2016. Automatic road crack detection using random structured forests. *IEEE Transactions on Intelligent Transportation Systems,* 17**,** 3434-3445.

SILVA, W. R. L. D. & LUCENA, D. S. D. 2018. Concrete cracks detection based on deep learning image classification. *Multidisciplinary digital publishing institute proceedings,* 2**,** 489.

SIMONYAN, K. & ZISSERMAN, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

SUN, Z., PEI, L., LI, W., XUELI, H. & YAO, C. 2020. Pavement encapsulation crack detection method based on improved Faster R-CNN. *Journal of South China University of Technology (Natural Science Edition),* 48**,** 84-93.

WANG, W. & WU, L.-C. 2014. Extraction of pavement cracks based on valley edge detection of fractional integral. *Huanan Ligong Daxue Xuebao/Journal of South China University of Technology (Natural Science),* 42**,** 117-122.

XU, G., MA, J., LIU, F. & NIU, X. Automatic recognition of pavement surface crack based on BP neural network. 2008 International Conference on Computer and Electrical Engineering, 2008. IEEE, 19-22.

YANG, F., ZHANG, L., YU, S., PROKHOROV, D., MEI, X. & LING, H. 2019. Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Transactions on Intelligent Transportation Systems,* 21**,** 1525-1535.

ZHANG, A., WANG, K. C., LI, B., YANG, E., DAI, X., PENG, Y., FEI, Y., LIU, Y., LI, J. Q. & CHEN, C. 2017. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. *Computer-Aided Civil and Infrastructure Engineering,* 32**,** 805-819.

ZHANG, D. & LI, Q. 2015. Overview of the development of highway pavement rapid detection technology. *Surveying and mapping geographic information,* 40**,** 1-8.

ZHANG, L., YANG, F., ZHANG, Y. D. & ZHU, Y. J. Road crack detection using deep convolutional neural network. 2016 IEEE international conference on image processing (ICIP), 2016. IEEE, 3708-3712.

ZOU, Q., ZHANG, Z., LI, Q., QI, X., WANG, Q. & WANG, S. 2018. Deepcrack: Learning hierarchical convolutional features for crack detection. *IEEE Transactions on Image Processing,* 28**,** 1498-1512.

BIOGRAPHICAL NOTES

Ahmed Elamin is a PhD candidate in Geomatics Engineering, the Department of Civil Engineering, Ryerson University, Canada. His Ph.D. research is on the development of UAV-based multi-sensor integration for precision mapping and object classification.

Ahmed El-Rabbany obtained his PhD degree in GPS Satellite Navigation from the Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada. At present, Dr. El-Rabbany is a full professor at Ryerson University, where he leads research projects in the areas of satellite navigation and multi-sensor integration.

CONTACTS

Ahmed Elamin
Ryerson University
350 Victoria Street, Toronto, Ontario, Canada M5B 2K3
e-mail: agelamin@ryerson.ca
Phone: (647) 673-0404
Web: http://www.civil.ryerson.ca

Professor Ahmed El-Rabbany
Ryerson University
350 Victoria Street, Toronto, Ontario, Canada M5B 2K3
e-mail: rabbany@ryerson.ca
Phone: (416) 979-5000 ext. 556472; fax: (416) 979-5122
Web: http://www.civil.ryerson.ca