Application of convolutional neural network for detecting concrete cracks

Tu T. Nguyen(1), Hiep H. Vu(2) and Kien T. Doan(3)

Abstract

Deep learning continues to growing in popularity and expanding for civil engineering applications thanks to easy access to massive sets of labeled data, increased computing power, and the availability of pre-trained models built by experts. In this paper, a Convolutional Neural Network (CNN) method is employed to classify the crack/non-crack aerial images captured on the surface of concrete structures. The CNN model was trained and validated using the available experimental data of 4000 previously published images. The trained CNN model was then tested with 330 unseen images. It was shown that the proposed CNN model can classify the crack/non-crack images with an accuracy level of 93%.

Key words: *Deep Learning, Convolutional Neural Network, Crack Detection, Concrete Crack*

(1) Dr., Faculty of Civil Engineering, Hanoi Architectural University,
Email: <tunt@hau.edu.vn>
(2) Assoc. Prof., Faculty of Civil Engineering, Hanoi Architectural University,
Email: <vuhoanghiep@hau.edu.vn >
(3) Dr., Faculty of Civil Engineering, Hanoi Architectural University,
Email: <kiendt@hau.edu.vn >

Date of receipt: 15/4/2022
Editing date: 6/5/2022
Post approval date: 5/9/2022

1. Introduction

Crack on concrete structures is a significant indication of possible reinforcement corrosion, spall development, or overload conditions. Thus, monitoring the cracks on the structure surface would provide important information to evaluate the safety level of the structure as well as to have an appropriate rehabilitation plan. Manual visual inspections using human labor are proven as an effective method to detect surface cracks in concrete, however, the method is time-consuming, labor-intensive, and sometimes exposes risks to the inspectors. With the development of aerial vehicle (AV) devices and machine learning-based techniques, more and more automated AV-based visual inspections are available at an affordable cost and high level of accuracy. The technique itself consists of two parts: (i) image data collection, and (ii) data processing.

The process of collecting aerial image data has been conducted by many investigators [1-6]. For example, Jong et al. [1] used unmanned aerial vehicles (UAV) to capture the images from the lower part of slab desks in bridges. Chen et al. [3] employed UAV to take aerial images of different types of typical ground targets namely buildings, roads, mountains, and riverways to study the aftermath of an earthquake strike. Li and Zhao [5] obtained the image dataset using a smartphone from the surface of a pylon and anchor room of a suspension bridge. In a recent study, Zhou and Song [6] utilized the high-resolution, vehicle - mounted to collect aerial images from the concrete roadways.

With regard to image data processing, various popular convolutional neural network systems such as VGG [7], GoogLeNet [8], and ResNet [9] have been proposed. In recent years, the applications of deep learning to address engineering issues have been widely used among researchers [10-15]. Related to the application of CNN, Zhang et al. [10] applied a deep learning technique for road crack detection using images captured from smartphones. Maeda et al. [11] developed a mobile phone application to detect road surface defects. The application of the CNN approach for defects detection was also found in a study by Tong et al. [12] with images from a ground-penetrating radar.

In this paper, the CNN-based model was developed to identify crack/non-crack images collected on the surface of a concrete structure. The CNN model was adapted from the pre-trained, open-sourced model developed by Google and distributed through TensorFlow. Available experimental data was collected from the concrete roadways with a vehicle-mounted laser imaging system. The CNN-based model was developed with the PYTHON environment.

2. Methodology

This section presents brief description of the data collection process, as well as the method to pre-process and generate aerial image data, were presented. In addition, a predictive model called CNN was employed to detect the surface concrete cracks using the datasets mentioned in the previous sections. The structure of the CNN model, the title of layers, their roles in the system, and some basic steps to train the model were also briefly discussed. Detailed information is presented in the subsequent sections.

2.1. Data set and data augmentation

Data used in this study were obtained from an available, published source [6]. Experimental data were collected from the concrete

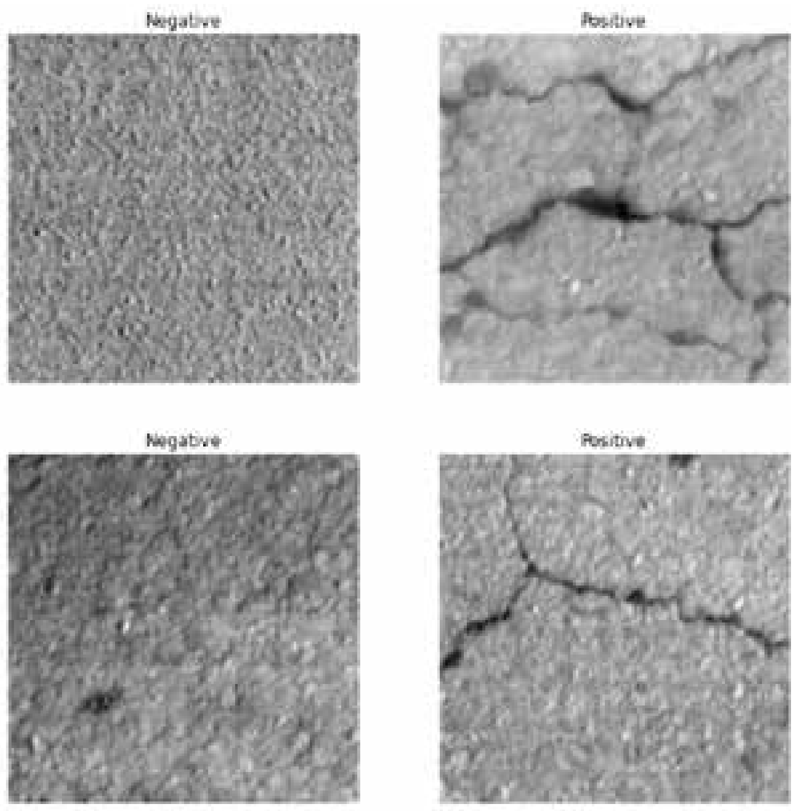

Figure 1: Inputs from different groups loaded with CNN model

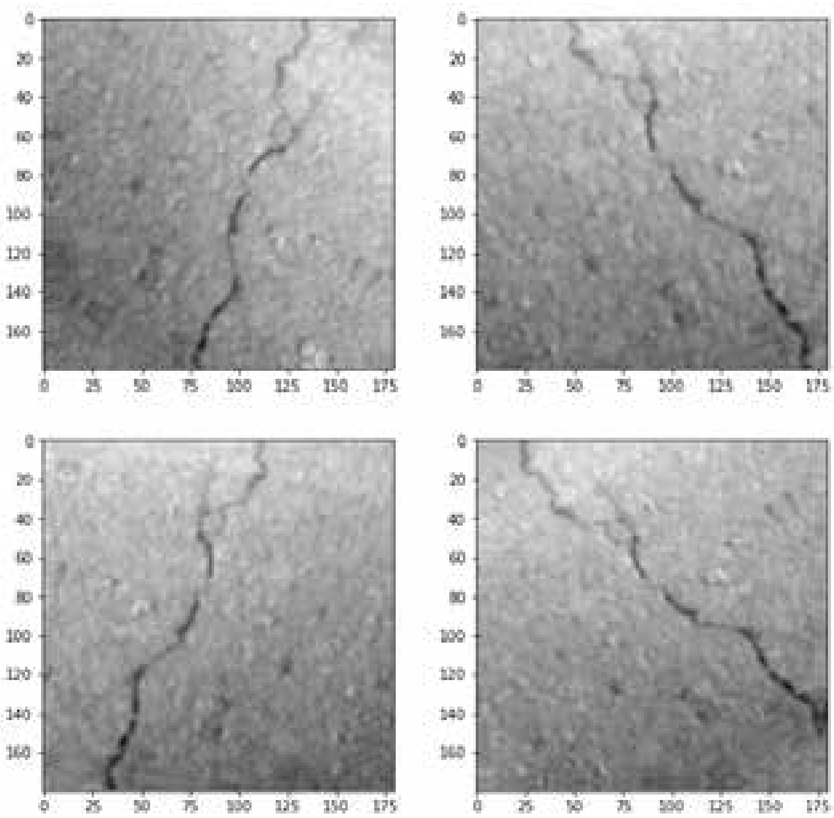

Figure 2: An example of image argumentation

roadways using a camera module is mounted 2.13 m above the ground on a vehicle. The images were captured while the vehicle was running at a speed of less than 9.83m/s. Detailed data collecting processes can be found in [6]. The raw data obtained from the field were then pre-processed to remove the unwanted effects of surface variations, scanning noises, and non-crack patterns. The final data set were generated using the sliding window technique [16], classified and documented in the "crack" and "non-crack" folders. A total of 4000 images are generated for each type of image data. Figure 1 presents some images from the positive/crack and negative/non-crack groups loaded by the proposed CNN model.

2.2. Convolutional Neural Network

The structure of a standard CNN model consists of an input layer, convolutional layers, pooling layers, and fully connected layers with an activation function to produce the output. The role of the convolutional layer is to apply the convolution to the raw input data and pass the results to the next layer. The pooling layer is extracted the dominant features from the input, usually using the maximum pooling or average pooling technique. The fully connected layers convert the two-dimensional features obtained from the previous layers into a one-dimensional vector and feed it into a softmax function to generate the outputs. Figure 3 illustrates a structure of a typical CNN model.

The CNN model used in this study has one node in the input layer and two nodes (i.e., crack or non-crack) in the output layer. To train the CNN model, the cracks and non-cracks images were loaded from the two separate folders and preprocessed to reduce the size of the pictures to 180 by 180 pixels before feeding to the proposed CNN model. In this study, about 90% of the entire inputs were used to train and validate the model, and 10% of the database was used for testing the accuracy of the trained CNN model.

3. Results and discussion

As previously mentioned, the trained CNN model was used to classify the crack and non-crack images in the test set. The following section present the prediction capabilities of the proposed model for the 330 unseen images in the testing dataset were presented. The performance of models would be evaluated through various performance metrics including Accuracy, Precision, Recall, and $F_{1-score}$. The confusion matrix and AR indicator would also be briefly discussed.

3.1. Model performance metrics

Accuracy refers to the ratio of the number of correctly predicted crack and non-crack images to the total number of input images.

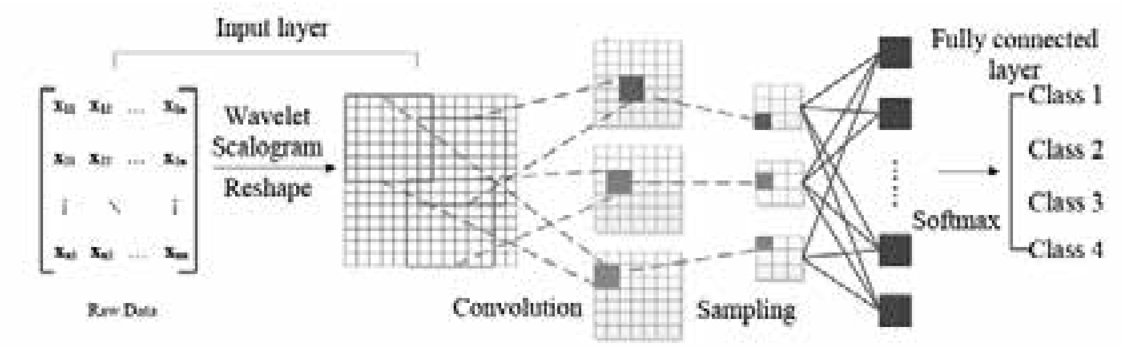

Figure 3: Architecture of CNN model

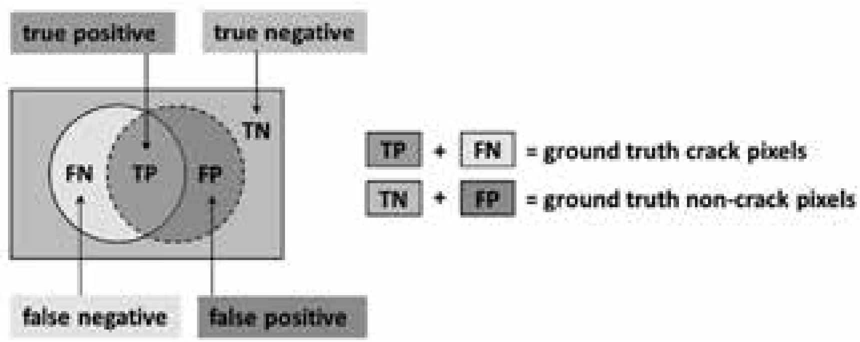

Figure 4: The schematic diagram for the performance metrics

Precision can be understood as the number of correctly predicted crack images divided by the number of crack images predicted by the classifier. The recall is the percentage of the number of correctly predicted crack images to the total number of cracked images. $F_{1-score}$ is the harmonic mean of precision and recall. Accuracy, Precision, Recall, and $F_{1-score}$ can be calculated through equations (1a-1d) using true-positive (TP), true negative (TN), false-positive (FP), and false-negative (FN), as illustrated in Figure 4.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1a)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1b)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (1c)$$

$$F_{1-score} = 2 \times \frac{\text{Precsion} \times \text{Rcall}}{\text{Precision} + \text{Recall}} \quad (1d)$$

An alternative way to present the performance results is using a confusion matrix. The columns of a confusion matrix represent the true value, and the rows show the predicted values assigned by the predictive model. The element aij (i is the row, and j is the column) indicates that the model assigned the value as i while the true value as in the database is I. The elements in the diagonal of the confusion matrix (aii in the light green cells) are the components correctly classified by the model. Additionally, an accuracy rate (AR) indicator (i.e., the percentage of predicted images that accurately matched the actual one) is also calculated for each group in the entire test set.

3.2. Model performance evaluation

A total of 330 images (i.e., 165 crack images and 165 non-crack images) in the training dataset were employed to test the trained CNN model. Performance results of the CNN model for the testing dataset are listed in Table 1. As can be seen, the trained CNN model can classify the crack/non-crack images at a high level of accuracy with an F1−score value of 93.0%. It is worth noting that high accuracy, precision, and recall values indicate a high positive detection rate, low false-positive rate, and low false-negative rate, respectively.

Table 1: Statistic on the performance metrics of CNN model

Accuracy (%)	Precision (%)	Recall (%)	$F_{1\text{-score}}$ (%)
93.0	98.8	88.6	93.4

Table 2: Confusion matrix for testing performance of CNN model

	Actual		
Prediction	Crack	Non-crack	Sum
Crack	163	21	184
Non-crack	2	144	146
Sum	165	165	330
AR (%)	98.8	87.3	93.0

The performance of the CNN model in terms of a confusion matrix for the testing dataset is presented in Table 2. It is interesting to note is that CNN showed a high level of accuracy in classifying the crack images with an AR value of 98.8%. The model, however, produced some misclassification for the non-crack group. To be specific, misidentified 21 non-crack images as crack ones.

4. Conclusions and recommendations

In this paper, a novel method using the Deep Learning approach to detect cracks on concrete surface is presented and discussed. The CNN model was developed using the available aerial images obtained from past publications. The trained CNN models were then utilized to categorize 330 images in the testing dataset. In terms of model performance, the CNN model demonstrated a high precision in detecting concrete cracks. In future studies, the proposed CNN application is recommended to integrate with a computer and a camera system mounted on a vehicle to test the crack recognition capability of the software for the concrete road in Vietnam./.

References

1. *Jong-Woo, K.; Sung-Bae, K.; Jeong-Cheon, P.; Jin-Won, N. Development of Crack Detection System with Unmanned Aerial Vehicles and Digital Image Processing. Proceedings of the World Congress on Advances in Structural Engineering and Mechanics (ASEM15), Incheon, Korea, 25–29 August 2015; pp. 1–11.*
2. *Aldea, E.; Hégarat-Mascle, S.L. Robust crack detection for unmanned aerial vehicles inspection in a contrary decision framework. J. Electr. Imaging 2015, 24, 1–16, doi: 10.1117/1.JEI.24.6.061119.*
3. *Jinhong Chen, Haoting Liu, Jingchen Zheng, Ming Lv, Beibei Yan, Xin Hu, Yun Gao, Damage Degree Evaluation of Earthquake Area Using UAV Aerial Image. International Journal of Aerospace Engineering, Vol. 2016, Article ID 2052603, 10 pages, 2016. https://doi.org/10.1155/2016/2052603.*
4. *Peng, X., Zhong, X., Zhao, C., Chen, Y. F., Zhang, T. (2020). The Feasibility Assessment Study of Bridge Crack Width Recognition in Images Based on Special Inspection UAV. Advances in Civil Engineering, 2020, 8811649. https://doi.org/10.1155/2020/8811649.*
5. *S. Li and X. Zhao. Image-based concrete crack detection using convolutional neural network and exhaustive search technique. Advances in Civil Engineering, Vol. 2019.*
6. *Zhou, S., & Song, W. (2021). Deep learning-based roadway crack classification with heterogeneous image data fusion. Structural Health Monitoring, 20(3), 1274–1293. https://doi.org/10.1177/1475921720948434.*
7. *K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. 2014, https://arxiv.org/abs/1409.1556.*
8. *C. Szegedy. Going deeper with convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9, Boston, MA, USA, June 2015.*
9. *K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, Boston, MA, USA, June 2016.*
10. *L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu. Road crack detection using deep convolutional neural network. Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. 3708–3712, IEEE, Phoenix, AZ, USA, September 2016.*
11. *H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata. Road damage detection using deep neural networks with images captured through a smartphone. 2018, https://arxiv.org/abs/1801.09454.*
12. *Z. Tong, J. Gao, and H. Zhang. Recognition, location, measurement, and 3D reconstruction of concealed cracks using convolutional neural networks. Construction and Building Materials, Vol. 146, pp. 775–787, 2017.*
13. *Nguyen, T. T., Ngoc, L. T., Vu, H. H., & Thanh, T. P. (2021). Machine learning-based model for predicting concrete compressive strength. International Journal of Geomate, 20(77), 197-204.*
14. *Nguyen, T.T., Pham, D. H., Pham, T.T., and Vu, H.H. (2020). Compressive strength evaluation of Fiber-Reinforced High Strength Self-Compacting Concrete with artificial intelligence. Advances in Civil Engineering, 2020, 3012139. https://doi.org/10.1155/2020/3012139.*
15. *Pham, T. T., Nguyen, T. T., Nguyen, L. N., Nguyen, P. V. (2020). A neural network approach for predicting hardened property of Geopolymer concrete. International Journal of GEOMATE, Oct. 2020, Vol.19, Issue 74, pp.193–201. ISSN: 2186-2982 (P), 2186-2990 (O), Japan, https://doi.org/10.21660/2020.74.72565.*
16. *Cha YJ, Choi W and Buyukozturk O. Deep learning-based crack damage detection using convolutional neural networks. Comput Aided Civ Inf 2017, 32: 361–378.*
17. *Dung CV. Autonomous concrete crack detection using deep fully convolutional neural network. Automat Constr 2019, 99: 52–58.*
18. *Srivastava N, Hinton G, Krizhevsky A. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014, 15: 1929–1958.*

Strength reduction of mudstone embankment...

(tiếp theo trang 30)

due to increased water content is inevitable. For the reason, it is important to develop design and quality control methods for weak rock embankments that take into account the reduction in strength due to water absorption and retention in order to improve durability and long-term performance./.

References

1. *M. Takagi, S. Yokota, K. Suga, S. Yasuda and H. Ota. The actual situation of the slope of earthfill that collapsed by an earthquake disaster in the Tomei Expressway Makinohara district. 55th Geotechnical Engineering Symposium, 193-196, 2010 (in Japanese).*
2. *H. Nakamura. Study on Design and Construction Method of High Performance Embankment on Expressway. Doctoral Dissertation. Department of Environmental Field Engineering, Graduate School of Engineering, Hokkaido University, 2014 (in Japanese).*
3. *H. Shima and S. Imagawa. Compressive settlement of slaking materials (weak rock) and its countermeasures. Tsuchi-to-Kiso. Japan Geotechnical Society, Vol.28, No.7, 45-52, 1980 (in Japanese).*
4. *NEXCO Test Methods, Test method for slaking of rocks, NEXCO Test Method 110, 2015 (in Japanese).*
5. *NEXCO Test Methods, Test method of crushing of rocks, NEXCO Test Method 109, 2015 (in Japanese).*
6. *Japanese Industrial Standards, Method for unconfined compression test of soils, JIS A 1216, 2020.*
7. *Japanese Industrial Standards, Test method for soil compaction using a rammer, JIS A 1210, 2020.*