Enhancing Dam Inspection with Pixel-Level CNN-FCN Approach Via 3D Texture Mapping

KRISADA CHAIYASARN, APICHAT BUATIK, VISHAL JANGID, SIRISILP KONGSILP and NAVID KHADEMI

ABSTRACT

Inspecting damage through visual inspection is an outdated and inefficient method, especially for large concrete dam structures. Furthermore, inspecting for damage to dam structures is costly, time-consuming, and hazardous. This research proposes a novel method for detecting and measuring cracks in the concrete structure of dams by utilizing a combination of the Convolutional Neural Network (CNN)-Fully Convolutional Network (FCN) algorithm and image-based 3D modeling. The method enables the identification of cracks on large structures at a pixel level, facilitating detection over a wide area and with high accuracy. Specifically, the approach detects cracks on the three-dimensional surface of the dam, providing a comprehensive and effective means of identifying potential structural issues. To accurately determine the size of the detected damage, the local thickness (LT) algorithm is introduced. This algorithm utilizes the pixel-precise detection capabilities of crack detection to provide highly precise crack size measurements. The CNN-FCN system segments cracks at the pixel level on the texture space obtained from a 3D model created through photogrammetry techniques. Then the trained CNN is employed to detect crack patches, which are then imported into a trained FCN system for pixel-level segmentation. The predicted crack pixels are converted to a physical scale by the LT algorithm to quantitatively measure the morphological features of cracks. The resulting crack map is then projected onto a 3D model, providing a comprehensive visual representation of the cracks. The proposed crack detection method can achieve an accuracy of over 90%. The study revealed that the presented system could efficiently and precisely detect, locate, and measure the size of cracks in large dam structures, highlighting the potential of the proposed method to enhance the efficiency and effectiveness of crack inspection. Furthermore, the visualization of cracks on 3D models can help increase inspectors' inspection efficiency.

Krisada Chaiyasarn, Associate Professor, Email: ckrisada@engr.tu.ac.th. Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani, Thailand

Apichat Buatik, PhD Candidate, Corresponding author e-mail: apichat.buat@dome.tu.ac.th. Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani, Thailand

Vishal Jangid, Master's Degree, Email: vishal.jang@dome.tu.ac.th. Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani, Thailand

Sirisilp Kongsilp, PhD, Email: sirisilp@engr.tu.ac.th. Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani, Thailand

Navid Khademi, Associate Professor, Email: navid.khademi@ut.ac.ir. (i) School of Civil Engineering, College of Engineering, University of Tehran, Tehran, Iran ;(ii) Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani, Thailand

INTRODUCTION

Concrete dams (Figure 1) are essential infrastructure components that must undergo routine inspection to guarantee their structural integrity and safety. Particularly when working with huge constructions, the conventional approach of visual examination is time-consuming, costly, and risky. The Convolutional Neural Network-Fully Convolutional Network (CNN-FCN) algorithm and image-based 3D modeling are combined in this work to present a unique approach for identifying and quantifying fractures on concrete dams. Various crack detection techniques have been developed for concrete structures, including visual inspection, ultrasonic testing, and digital image processing [1]. Visual inspection is the most common approach, but it is subjective and can miss small cracks not visible to the naked eye. Ultrasonic testing [2] effectively detects internal cracks, but they are expensive and require specialized equipment and expertise. Digital image processing [3] is a promising technique used to detect cracks on the surface of concrete structures. However, most existing methods rely on 2D images and do not provide accurate crack size measurements. A potential method for identifying and quantifying fractures in concrete buildings is image-based 3D modeling [4]. Creating high-resolution 3D models from multiple 2D images provides a comprehensive visual representation of the structure and enables the detection of cracks over a wide area. However, existing methods often rely on manual analysis of the 3D models, which can be time-consuming and error-prone. Convolutional Neural Networks (CNNs) and Fully Convolutional Networks (FCNs) are effective in detecting and segmenting objects in images [5]. In order to identify and quantify cracks in concrete dams, this study suggests a novel method that combines the CNN-FCN algorithm with image-based 3D modeling. CNN is a kind of classification machine learning technique that assigns the images patches to 'crack' and 'non-crack' categories. Besides, the FCN algorithm segments the cracks at the pixel level on the texture space obtained from a 3D model created through photogrammetry techniques. Stated more precisely, the trained CNN detects crack patches, which are then imported into a trained FCN system for pixel-level segmentation. The resulting 'crack map' is then projected onto a 3D model, providing a comprehensive visual representation of the cracks. To accurately determine the size of the detected damage, the local thickness (LT) algorithm is introduced [6]. This algorithm utilizes the pixel-precise detection capabilities of crack detection to provide exact crack size measurements. The LT algorithm converts the predicted crack pixels to a physical scale, enabling the quantitative measurement of the morphological features of cracks. In brief, a thorough and efficient technique for identifying possible structural problems in concrete dams—crack detection using a combination of the CNN-FCN algorithm and image-based 3D modeling—is suggested. Cracks may be found across a large area thanks to 3D modeling, and the CNN-FCN algorithm offers precise pixel-level segmentation and detection. The local thickness algorithm allows for accurate crack size estimations, and the 3D representation of cracks can increase inspection productivity. The study highlights the system's potential to enhance the efficacy and efficiency of crack inspection by demonstrating its capacity to quickly and accurately identify, locate, and measure the size of fractures in substantial dam structures.

Figure 1. Example pictures of a dam structure.

METHODOLOGY

The CNN-FCN crack detection method is used in this paper, and the findings are provided on a 3D concrete foundation. A dam construction evaluated in this research is depicted in Figure 1. The system starts with the technique for acquiring the texture space data and the images utilized to construct 3D models. The raw picture dataset from another dam structure is then used in the CNN and FCN training procedure. During training, picture patches from public crack datasets and texture spaces were incorporated to avoid overfitting and limit false detections caused by any distortions in the texture space. Lastly, a "crack map"—the output of the CNN-FCN to detect cracks in the texture space—can be produced.

Data Collection and 3D Modeling

The acquisition of crack image data is made using two different techniques. As a starting point, the 2D (two-dimensional) grid path method is used to collect general information about the dam. When data gathering is necessary for a 3D model, the 3D grid path approach is used instead. Both use a purposeful sweep of a curving motion by the digital single lens reflex (DSLR) camera to gather data (dimensions: $4,896$ x $3,264$ pixel^2) from the entire scene. The image-based 3D reconstruction approach reveals parts of the dam's relative size, form, and texture. To produce a 3D model, this method combines characteristics from Image Processing, Computer Vision, and Machine Learning. The 3D modeling procedure for this work is carried out using the Agisoft Metashape software.

Texture Data Reading

Theoretical groundwork based on applying a pinhole camera model serves as the foundation for the current paper [7]. Notably, this model calls for creating a scene view by projecting 3D points onto the picture plane via a perspective transformation; those interested in a more thorough explanation may find more information on this procedure in [3]. The present study has implemented a camera registration process that involves the simulation of a virtual camera on a 3D model, with the position of said virtual camera determined by that of the initial camera. Upon completing the camera registration

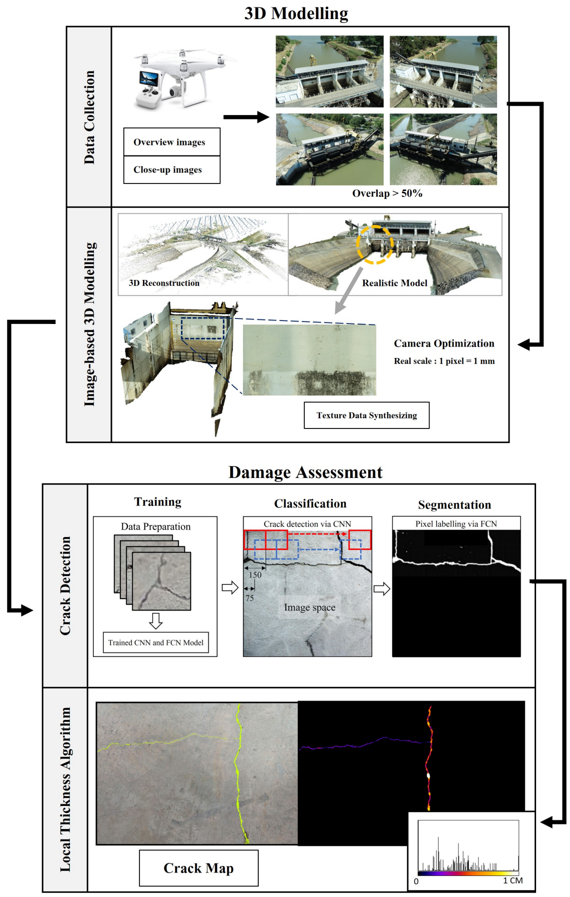

Figure 2. System overview and contribution of this paper.

process, the virtual camera's position is then displayed on the 3D model. To provide a visual representation of this process, Figure 2 has been included, which displays an example of the texture space data obtained from the camera registration process

Training Dataset Preparation

A mixture of $30,000$ patches from an open-source dataset makes up the training picture dataset. (for more detail, refer to [3]). The total number of $30,000$ image patches includes both cracked and uncracked versions of the photos, totaling $15,000$ patches in each case. Out of the $15,000$ crack-labeled photos in the dataset, $4,056$ were randomly chosen for image segmentation. There are $3,245$ and 811 image patches for picture segmentation in the training and validation datasets, respectively.

Damage Assessment and Crack Detection System

As shown in Figure 2, this work's suggested crack detection method is a combination of image classification and segmentation approaches [3]. The system is initialized through the training of CNN and FCN architectures. The CNN model, VGG19, is then used to detect possible crack-containing areas by removing crack patches using the sliding window technique. VGG19 is a machine learning classification algorithm that classifies image patches into "crack" and "non-crack" categories. It is used to detect probable areas of cracks to filter out crack patches. On the other hand, the FCN is then used to segment the crack locations that VGG19 identified. The two primary parts of the FCN are down- and up-sampling. The down-sampling component consists of convolutional layers [8], pooling layers [8], and dropout layers [9], while the up-sampling component consists of deconvolutional layers [10]. The FCN architecture used in this research was developed based on the U-Net architecture [11] and the work of Yang et al. [10], drawing upon their methodology for semantic segmentation. The FCN is the pixel-to-pixel convolutional network tasked with semantic segmentation, converting predictions from several classes to a semantic segmentation image for crack segmentation. Stated more precisely, FCN has been integrated to identify fractures at the pixel level, while CNN aids in improving the segmentation accuracy of cracks in FCN by deleting non-crack areas. The suggested system's main benefits include a large increase in fracture detection accuracy at the pixel level and an extension of the surface area that may be examined.

Local Thickness Algorithm

The local thickness (LT) method [12] is utilized to precisely analyze the damages and quantify the extent of the discovered damage (Figure 2). This algorithm utilizes the pixel-precise detection capabilities of crack detection to provide highly precise crack size measurements. Let $\omega \in R^3$ be the collection of every point in the crack being studied and $\underline{p} \in \omega$ an arbitrary point in it. As the local thickness $\tau(\underline{p})$ we define the diameter of the largest sphere which contains the point $(\underline{p})$ and which is completely inside the crack:

$$\tau(\underline{p}) = 2 \cdot max(r|\underline{p} \in sph(\underline{x}, r) \subseteq \omega, x \in \omega) \tag{1}$$

where $sph(\underline{x}, r)$ is the set of points inside a sphere with center x and radius r. This equation determines the average crack thickness. It is model-independent and enables fracture thickness to be calculated from 3D image data.

RESULTS AND DISCUSSION

Table I lists the outcomes of a concrete foundation structure's 3D modeling procedure. The projection values show the number of feature matching points with a root mean square (RMS) reprojection error. The RMS reprojection error indicates that the 3D modeling method's scanning error from 414 photos has an average error of 0.559 pixels in each image, leading to the assumption that the procedure has minimal faults and the conclusion that the model is highly accurate. The 3D model possesses a surface mesh of $2,282,086$ faces and a texture of $4,408 \times 2,343$ (texture size/ count), boasting high-resolution quality.

In the current experiment, a 3D model of dam construction was made using photographs captured by drones and DSLR cameras. In particular, the drones were used to take aerial pictures of the dam, while the DSLR camera was used to take close-up pictures of the structure to record its finer aspects. It should be emphasized that gathering picture data to build a 3D model is a difficult and complex task. Creating a precise and accurate 3D model is necessary for collecting high-quality texture data, which in turn

TABLE I. Reconstruction parameters.

Parameter Name	Value
Sparse point cloud	$1,247,792$ points
RMS reprojection error	0.559 pixels
Max reprojection error	1.832 pixels
Dense point cloud	$102,914,600$ points
Surface mesh	$2,282,086$ faces
Texture data (Region of Interest)	$4,408$ x $2,343$ pixels (1 mm/pixel)

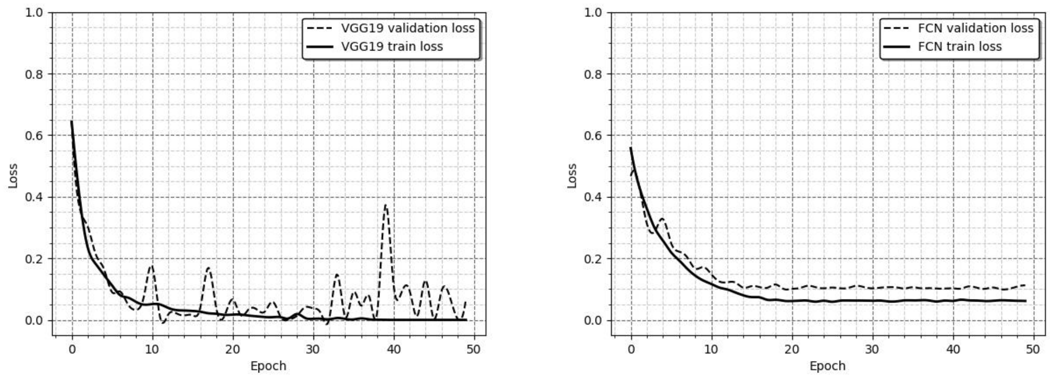

Figure 3. Loss of training and validation processes for (Left) image classification using VGG19 and (right) image segmentation using FCN.

depends on the knowledge and abilities of the drone operators. Additional challenges to this work include the manual development and selection of a virtual camera perspective, which must be properly handled to successfully create an accurate 3D model. Figure 3 in this investigation depicts the training and validation loss for image classification using VGG19 and image segmentation using FCN. The classifiers were trained over 50 epochs, as illustrated in the figure. Throughout the training process, all pre-trained networks achieved high levels of accuracy, with nearly 0.99 accuracies and 0.01 loss being attained. Figure 4 presents the results of crack segmentation on the dam structure. The figure includes (a) the original texture data, (b) the ground truth, (c) crack classification using VGG19, (d) the proposed crack segmentation using CNN-FCN, and (e) measurements of the crack size using the local thickness algorithm. Table II displays the results of the crack segmentation task on the dam using the proposed CNN-FCN system. The results indicate that this system achieved high accuracy, precision, recall, and F1 scores, with values of 0.999, 0.804, 0.803, and 0.817, respectively. Overall, this study's results demonstrate the proposed approach's efficacy in accurately detecting and segmenting cracks in dam structures.

TABLE II. Summary results of crack segmentation on a dam structure using CNN-FCN (proposed system).

Parameter Name	**Value**
True Positive (TP)	$2,702$ pixels
True Negative (TN)	$10,324,032$ pixels
False Positive (FP)	657 pixels
False Negative (FN)	553 pixels
Accuracy	0.999 pixels
Precision	0.804 pixels
Recall	0.830 pixels
F1 Score	0.817 pixels

Note: Crack pixel are $3,255$ pixels. Non-Crack pixel are $10,324,689$ pixels.

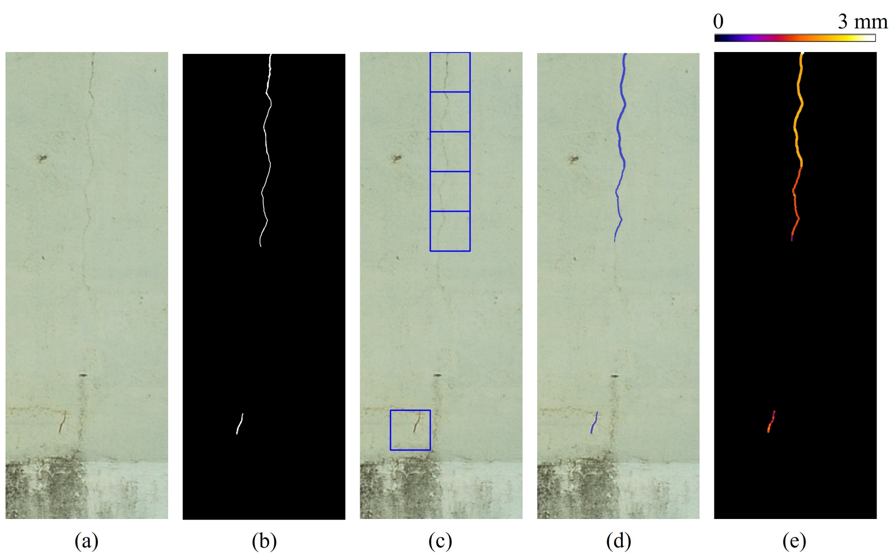

Figure 4. The results of crack segmentation on a dam structure are presented, including: (a) the original texture data, (b) the ground truth, (c) crack classification using VGG19, (d) crack segmentation using CNN-FCN (the proposed system), and (e) measurements of the crack size using the local thickness algorithm.

CONCLUSION

In summary, this study introduced a novel approach for detecting cracks in dam structures by utilizing a texture space derived from a 3D model. The proposed method offers a larger inspection area and enables scrutiny at the pixel level through effective texture space data. This paper's integrated CNN-FCN crack detection system demonstrated promising results, achieving coherent crack detection at the pixel level with high accuracy. The practical significance of this system lies in its ability to enhance the efficiency of inspectors while reducing time constraints. However, future studies are necessary to expand the coverage of large-scale structures, and additional image data for FCN training should be incorporated to continuously improve the accuracy of crack segmentation. The prospect of performing end-to-end learning in a multi-task setup with a shared encoder and two separate decoders for classification and segmentation is also a potential avenue for future research. Overall, this investigation serves as a stepping stone for further research in developing and enhancing a tool that can measure the width and length of cracks on a 3D model, ultimately contributing to the safety and maintenance of dams. The proposed method has significant potential for practical applications in structural engineering, particularly in inspecting and maintaining critical infrastructure such as dams.

ACKNOWLEDGMENT

The authors would like to thank the Research Unit, Thammasat School of Engineering (TSE) and Thammasat University, Thailand for providing scholarships and support for the research project.

REFERENCES

1. Anitha, M., R. Hemalatha, and S. Radha. 2021. “A survey on crack detection algorithms for concrete structures,” *Advances in Smart System Technologies: Select Proceedings of ICFSST 2019*:639–654.
2. Chakraborty, J., A. Katunin, P. Klikowicz, and M. Salamak. 2019. “Early crack detection of reinforced concrete structure using embedded sensors,” *Sensors*, 19(18):3879.
3. Chaiyasarn, K., A. Buatik, H. Mohamad, M. Zhou, S. Kongsilp, and N. Poovarodom. 2022. “Integrated pixel-level CNN-FCN crack detection via photogrammetric 3D texture mapping of concrete structures,” *Automation in Construction*, 140:104388.
4. Golding, V. P., Z. Gharineiat, H. S. Munawar, and F. Ullah. 2022. “Crack Detection in Concrete Structures Using Deep Learning,” *Sustainability*, 14(13):8117.
5. Ali, R., J. H. Chuah, M. S. A. Talip, N. Mokhtar, and M. A. Shoaib. 2022. “Structural crack detection using deep convolutional neural networks,” *Automation in Construction*, 133:103989.
6. Shakoorioskooie, M., M. Griffa, A. Leemann, R. Zboray, and P. Lura. 2022. “Quantitative analysis of the evolution of ASR products and crack networks in the context of the concrete mesostructure,” *Cement and Concrete Research*, 162:106992.
7. Szeliski, R. 2022. *Computer vision: algorithms and applications*, Springer Nature.
8. Cha, Y.-J., W. Choi, and O. Büyüköztürk. 2017. “Deep learning-based crack damage detection using convolutional neural networks,” *Computer-Aided Civil and Infrastructure Engineering*, 32(5):361–378.
9. Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. 2014. “Dropout: a simple way to prevent neural networks from overfitting,” *The journal of machine learning research*, 15(1):1929–1958.
10. Yang, X., H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang. 2018. “Automatic pixel-level crack detection and measurement using fully convolutional network,” *Computer-Aided Civil and Infrastructure Engineering*, 33(12):1090–1109.
11. Liu, Z., Y. Cao, Y. Wang, and W. Wang. 2019. “Computer vision-based concrete crack detection using U-net fully convolutional networks,” *Automation in Construction*, 104:129–139.

12. Hildebrand, T. and P. Rüegsegger. 1997. "A new method for the model-independent assessment of thickness in three-dimensional images," *Journal of microscopy*, 185(1):67–75.