Bridge Girder Crack Assessment Using Faster RCNN Inception V2 and Infrared Thermography

Murad Al Qurishee1, Weidong Wu2, Babatunde Atolagbe3, Joseph Owino4, Ignatius Fomunung4, Said El Said1, Sayed Mohammad Tareq2

1Tennessee Department of Transportation, Nashville, Tennessee, USA
2University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA
3Maryland Department of Transportation, Baltimore, Maryland, USA
4Department of Civil & Chemical Engineering, The University of Tennessee at Chattanooga, Chattanooga, Tennessee, USA

Abstract

Manual inspections of infrastructures such as highway bridge, pavement, dam, and multistoried garage ceiling are time consuming, sometimes can be life threatening, and costly. An automated computerized system can reduce time, faulty inspection, and cost of inspection. In this study, we developed a computer model using deep learning Convolution Neural Network (CNN), which can be used to automatically detect the crack and non-crack type structure. The goal of this research is to allow application of state-of-the-art deep neural network and Unmanned Aerial Vehicle (UAV) technologies for highway bridge girder inspection. As a pilot study of implementing deep learning in Bridge Girder, we study the recognition, length, and location of crack in the structure of the UTC campus old garage concrete ceiling slab. A total of 2086 images of crack and non-crack were taken from UTC Old Library parking garage ceiling using handheld mobile phone and drone. After training the model shows 98% accuracy with crack and non-crack types of structures.

Keywords

Bridge Girder, Convolution Neural Network, Crack Detection, Structural Health Monitoring, Infrared Thermography

1. Introduction

Bridges constitute a sizable number of infrastructures in our environment and
are often quite expensive to construct and maintain. Their integrity, safety, sustainability, reliability and maintenance are as important as the initial construction. These factors, however, are often impeded by deteriorating effects due to age and long-term service and exposure to harsh environmental conditions such as wind and earthquakes. In order to mitigate this fast deteriorating effect of bridges, the science of health monitoring emerged.

It is not a gainsaying that the role of infrastructure is critical, and serves as a significant index, in measuring the development of a nation. However, modern challenges in infrastructure development transcend merely building of roads, bridges, and other social facilities, but rather exploring means to mitigate the deterioration. Certainly, the rate of deterioration of US infrastructure, for example, has been a subject of concern amongst politicians, engineers and the public at large [1]. According to the ASCE 2017 report card, the US infrastructure received a cumulative grade of D+ (i.e. fair condition) with the bridges averaging a grade of C+ (i.e. good) [1] [2].

Collapse of critical infrastructures such as bridges can have far reaching implications on the safety and economy of a nation. The failure of the Silver Bridge in Ohio in 1967 which led to the death of 46 people, the collapse of the I-35 W Bridge in Minnesota in 2007 which killed 13 people, injured 145 and resulted in a direct economic loss of US$17 million and US$43 million in 2007 and 2008, respectively [3] [4], the recent Florida International University (FIU) pedestrian bridge disaster in 2018 and the Tennessee Bridge Railing Collapse in 2019 are relatable experiences showing the implications of bridge collapse.

Bridges, just like other infrastructure, are very costly to construct, and their failure cannot be allowed. To prevent sudden failure, bridges are now routinely inspected for member or connection failures and their performances are measured or predicted regularly, with necessary actions taken to reduce their rate of deterioration. Bridge condition assessment techniques that have been explored include to not limited to: non-destructive techniques e.g. ultrasonic pulse velocity method, impact-echo/impulse-response method, acoustic emission method, radiographic method, eddy current method, eddy current method, and infrared thermographic methods; and dynamic characteristics-based methods e.g. use of natural frequency changes, use of modal damping changes, use of FRF changes, use of mode shape curvature changes, use of modal strain energy changes, and flexibility changes [5]. More recently, the use of techniques, such as Digital Imaging Correlation and various classes of Deep Learning such as Convolutional Neural Network (CNN), Region-CNN (RCNN) and Faster-RCNN, are becoming increasingly popular in identification and classification of defects in concrete structures due to advances in computing as researchers are exploring means to automate structural health assessment.

Deep learning is a machine learning technique that utilizes the deep neural network [6]. It allows computational models of multiple processing layers to learn representations of data with multiple levels of abstraction [7]. By harness-
ing this feature of machine learning, it is possible to use a well-trained neural network to detect, and classify, defects in concrete structures such as bridges thereby aiding engineering judgements of the conditions of bridges. The Objectives of this paper are to develop a framework that can be used for automation of bridge inspection, train a network that can be used for concrete crack classification, develop an algorithm to obtain info. of crack size and location in structure and build a 3D crack visualization model to assist maintenance engineer to determine whether the crack needs immediate attention.

2. Literature Review

Many researchers developed different image processing algorithms to detect infrastructural health monitoring. But most of the algorithms are limited to analyze a single picture instead of real-time video analysis. Tsai, Y. C. et al. (2009) analyzed Sobel model and Otsu method for pavement crack detection [8]. After comparing Sobel and Otsu method with the segmentation based on fractal method, they found fractal method is better than Sobel and Otsu method to analyze pavement crack surface. They analyzed 1280 × 1024 size pavement crack surface image to compare the results. The Sobel method includes edges due to differential operations and noises. The Otsu method produces error segmentation due to uneven grey-level distribution on background and lower contrast of grey level between crack and background. Some of the existing algorithms with their objectives are discussed in Table 1.

We used Sobel, Canny edge, Roberts, Prewitt, and Laplacian of Gaussian filter detection algorithms to detect crack from a hand mobile phone capture image. But all the algorithm failed to produce satisfactory result to identify the crack as shown in Figure 1. With the same image, we used different corner detection algorithm such as BRISK Feature Corner Detection, Harris Feature Corner Detection, FAST Features Corner Detection, Min Eigen Features Corner Detection, MSER Features Corner Detection, SURF Features Corner Detection, Extract Corner Features, HOG Features Extraction around corner point. Although some of the corner detection algorithm shows better detection compared to Figure 1 algorithms, some of the algorithm’s detections are over-exceeded as shown in Figure 2.

In summary, the available image processing techniques are impractical to use crack detection because of their lack of identification capabilities. Moreover, most of the techniques are limited by a single image while in practice the technique should perform on a sequence of images rather than a single image. In this research a state-of-art technique named as a convolution neural network to produce useful detection information.

3. Data Acquisition

In this work, we used UAV (i.e. drone) to get access to bridge or anywhere we need to collect images from. With the drone a FLIR 6.8 mm f/1.3 thermal imaging camera were incorporated to capture the thermal images, damages inside
Table 1. Comparison of image processing models [1].

| Researcher                | Methods                        | Objectives of the methods                                                                 |
|---------------------------|--------------------------------|------------------------------------------------------------------------------------------|
| Ying and Salari (2010)    | Beamlet Transforms (BT)        | Extraction of curvilinear features of crack from pavement images.                        |
| Hough (1962)              | Hough Transform (HT)           | Recognize complex pattern in a picture.                                                   |
| Leymarie and Levine (1992)| Grassfire Transform (GT)       | Describe planer object shape including region and boundary features.                     |
| Canny (1986)              | Laplacian of Gaussian (LOG)    | For edge detection.                                                                      |
| Witkin (1984)             | Scale-space filtering          | Precisely localize a large-scale structure using the gaussian convolution.                |
| Sobel [14]                | Sobel operator                 | The pixel-level edge detection algorithm.                                                |
| Pallotta (2014)           | Zernike moments operator       | Compute all the edge parameters for subpixel detection.                                  |
| Ying-Dong et al. (2005)   | Sobel-Zernike moments operator | A fast subpixel edge detection method.                                                    |
| Dong, W (2008)            | Prewitt method                 | Wall crack assessment from image.                                                        |
| Jianzhuang et al. (1991)  | Otsu method                    | Classify the image into background and foreground and find the threshold value to minimize the variances. |
| Farhidzadeh et al. (2013) | fractal method                 | Reinforcement concrete damage assessment.                                                |
| Massaro, E., & Bagnoli, F. (2014) | Percolation method            | Large-size concrete surface crack detection method [21]-[32].                           |
| Pudney (1998)             | Distance-ordered homotopic thinning (DOHT) | Distance-ordered homotopic thinning (DOHT) is effective and efficient Skeletonization Algorithm to verify a real and synthetic 3D image. |
| Fortune, S. (1987)        | Voronoi diagram                | Exact shape analysis and accurate skeleton approximation.                                |
| Choi et al. (2003)        | Euclidean skeleton             | Efficient and accurate skeletonization method.                                           |
| Jia and Tang (2003)       | ND tensor voting               | Can apply roughly homogenous and periodic textures images.                               |

Figure 1. Crack detection using (b) Sobel, (c) Canny edge, (d) Roberts, (e) Prewitt, and (f) Laplacian of Gaussian filter detection.
Figure 2. (a) BRISK Feature Corner Detection, (b) Harris Feature Corner Detection, (c) FAST Features Corner Detection, (d) Min Eigen Features Corner Detection, (e) MSER Features Corner Detection, (f) SURF Features Corner Detection, (g) Extract Corner Features, (h) HOG Features Extraction around corner point.

Structural components can be identified, classified and evaluated. A short tutorial was made for the demonstration of data acquisition in this YouTube video link (https://www.youtube.com/watch?v=q4ZDH7MDh9E). A total of 2088 images of crack and not-crack were collected [37]. All the images were split into two groups of images: one for training and the other for testing or validation purpose. In the next step, a neural network model was developed and compared it with a commercially available deep convolutional Neural Network by using the labeled training images. Finally, the trained network applied to classify unknown images. Figure 3 shows the conceptual framework of data acquisition and model training. Figure 4 and Figure 5 show the DJI Panthom 4 pro Drone and the infrared thermal camera respectively. The difference between the infrared image and the naked eye image were shown in Figure 6.

4. Data Pre-Processing

All the images were labeled and annotated using Labellmg Software. The images
Figure 3. A conceptual framework of data acquisition and model training.
Figure 4. DJI Panthom 4 pro Drone.

Figure 5. 6.8 mm f/1.3 thermal camera.

Figure 6. Comparison of the Infrared image and the naked eye image.
were annotated as crack and non-crack image types. Additionally, PIX resizer software was used to reduce the image dimension to $838 \times 600$. Image data augmentation technique was implemented to enlarge the training and validation data sets. Figure 7 displays the sprite image of all 2086 images (before augmentation) used in this study.

5. Model Development

A 22-layer (20 hidden layer, one input layer and one output layer) convolution model was developed named as Visual Geometry Group (VGG 22). During the training process, different ratios of train vs. test images were implemented. One example is 1668 images used for training the model and 418 images used for validating purpose. The input size of images was maintained as $224 \times 224 \times 3$ (height $\times$ width $\times$ channel). In Figure 8, the VGG 22 model architecture is described.

6. Model Training

To train the model, a high-performance supercomputer in UTC SimCenter was used to run the deep learning simulations. Python is the computational tool used to implement the deep learning for damage assessment. During the VGG-22 model training, we tried multiple combinations of epoch and batch sizes. Ten
Figure 8. VGG-22 model.
percent of all images were used for the testing/validation and the rest images were used for training purpose. Data augmentation techniques were deployed to enlarge the training image dataset. Figure 9 represents a feature map of the crack image in hidden layer 3 after the application of 32 filters. Stochastic Gradient Descent (SGD) optimizer had been used to update the network. We get to a validation accuracy more than 78% (Figure 10), which is pretty good considering that the images are noisy and limited number of images are used for training the network. The gap between the training and validation accuracies because of overfitting. It is observed from Figure 11 that the loss tends to be relatively stable after 20 epochs. Figure 12 demonstrates the confusion matrix which is erroneous detection of the model. It measures the performance of an algorithm to detect the accurate class.

Although the VGG-22 model has 78% validation accuracy, faster RCNN model was chosen to increase the validation accuracy. The architecture of Faster RCNN has two network modules: first module is region proposal network (RPN) and the second module is used to generate bounding box for the detected objects [38]. With Faster RCNN, Inception v2 was used as a feature extractor.

Figure 9. VGG-22 feature map of hidden layer 3.
The feature extractor defines as the pattern recognition image processing algorithm. The input image size was taken as $838 \times 600$. The model was trained up to 60 k global steps. After the training, the model reached almost zero loss as shown in Figure 13. In Figure 14, the localization loss and in Figure 15 localization loss were shown compared to number of global steps.

7. Crack Location

Google Earth and Geotag software were used to retrieve the GPS coordinates.
Figure 12. VGG-22 confusion matrix.

Figure 13. Faster RCNN total loss vs. global steps.

Figure 14. Faster RCNN classification loss vs. global steps.
Figure 15. Faster RCNN localization loss vs. global steps.

Figure 16 shows the Google Earth view and the crack location in the UTC old garage building. To find the GPS coordinate of crack images, an open-source software, named as Geotag, was used to extract the image latitude, longitude, altitude, county name, city name, and province. The software can generate GPS coordinate based on the image metadata. In Figure 17, the red marked box shows the sample of the GPS coordinate found from our collected image data.

8. Crack Length

3D Drone mapping was used to determine the area and the length of cracks. Pix4D software was used to process the aerial photogrammetry images into 3D modelling. Using Pix4D software dimension tool the cracks measurement was completed as shown in Figure 18.

9. Results

The VGG-22 model shows 80.35% validation accuracy and 70.03% training accuracy at 19th global steps. The faster RCNN with Inception V2 shows 99% confident at 60 k global steps. A test image was provided in the faster RCNN and the model successfully detected the crack and non-crack structure with an average confident 99% as demonstrated in Figure 19. The 99% confidence means the trained model is 99% confident to classify the crack and non-crack image which is the best confident model percentage.

10. Conclusions

The bridge girder health assessment using deep learning with infrared thermography and drone is pushing the technology of traditional health monitoring. This method can save time, cost and make the structural health assessment more robust and efficient.
In this article, first we developed a 22-layer model and then used an open source model called faster RCNN with Inception V2. The VGG-22 model shows 80.35% validation accuracy and 70.03% training accuracy at 19th global
Figure 19. Crack detection using infrared thermography.

steps. The faster RCNN with Inception V2 shows 99% confidence level at 60k global steps. The faster RCNN with Inception V2 model can successfully detect the crack and non-crack image based on the training output. Moreover, infrared thermography images overcome the limitations of normal images especially dark place or inside of a structure. Using infrared thermography with convolution neural network and drone for the structural health assessment is a beneficial and cost saving technology.

Acknowledgements

Supercomputer resources used for deep learning simulations provided by Dr. Ethan Hereth and UTC SimCenter is greatly appreciated. We are also grateful to Dr. Yu Liang at UTC Department of Computer Science for his valuable suggestions and comments.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Wu, W., Qurishee, M.A., Owino, J., Fomunung, I., Onyango, M. and Atolagbe, B. (2018) Coupling Deep Learning and UAV for Infrastructure Condition Assessment Automation. 2018 IEEE International Smart Cities Conference, Kansas City, MO,
[2] ASCE (2017) Infrastructure Report Card for Pavements. 
https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Roads-Final.pdf

[3] Xie, F. and Levinson, D. (2011) Evaluating the Effects of the I-35W Bridge Collapse on Road-Users in the Twin Cities Metropolitan Region. Transportation Planning and Technology, 34, 691-703. https://doi.org/10.1080/03081060.2011.602850

[4] Deng, L., Wang, W. and Yu, Y. (2015) State-of-the-Art Review on the Causes and Mechanisms of Bridge Collapse. Journal of Performance of Constructed Facilities, 30. Article ID: 04015005. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000731

[5] Xu, Y.-L. and He, J. (2017) Smart Civil Structures. CRC Press, London. https://doi.org/10.1201/9781315368917

[6] Kim, P. (2017) MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence. https://doi.org/10.1007/978-1-4842-2845-6

[7] LeCun, Y., Bengio, Y. and Hinton, G.J.N. (2015) Deep Learning. Nature, 521, 436-444. https://doi.org/10.1038/nature14539

[8] Tsai, Y.-C., Kaul, V. and Mersereau, R.M. (2009) Critical Assessment of Pavement Distress Segmentation Methods. Journal of Transportation Engineering, 136, 11-19. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000051

[9] Ying, L. and Salari, E. (2010) Beamlet Transform-Based Technique for Pavement Crack Detection and Classification. Computer-Aided Civil and Infrastructure Engineering, 25, 572-580. https://doi.org/10.1111/j.1467-8667.2010.00674.x

[10] Xu, L. and Oja, E. (2009) Randomized Hough Transform. In: Encyclopedia of Artificial Intelligence, IGI Global, London, 1343-1350. https://doi.org/10.4018/978-1-59904-849-9.ch197

[11] Leymarie, F. and Levine, M.D. (1992) Simulating the Grassfire Transform Using an Active Contour Model. IEEE Transactions on Pattern Analysis & Machine Intelligence, 14, 56-75. https://doi.org/10.1109/34.107013

[12] Zhang, L., He, Y., Xie, X. and Chen, W. (2009) Laplacian Lines for Real-Time Shape Illustration. In: Proceedings of the 2009 symposium on Interactive 3D graphics and games, ACM, New York, 129-136. https://doi.org/10.1145/1507149.1507170

[13] Witkin, A.P. (1987) Scale-Space Filtering. in Readings in Computer Vision. Elsevier, New York, 329-332. https://doi.org/10.1016/B978-0-08-051581-6.50036-2

[14] Kanopoulos, N., Vasanthavada, N. and Baker, R.L. (1988) Design of an Image Edge Detection Filter Using the Sobel Operator. IEEE Journal of Solid-State Circuits, 23, 358-367. https://doi.org/10.1109/4.996

[15] Pallotta, L., Clemente, C., De Maio, A., Soraghan, J.J. and Farina, A. (2014) Pseudo-Zernike Moments Based Radar Micro-Doppler Classification. 2014 IEEE Radar Conference, Cincinnati, OH, 19-23 May 2014, 850-854. https://doi.org/10.1109/RADAR.2014.6875709

[16] Qu, Y.-D., Cui, C.-S., Chen, S.-B. and Li, J.-Q. (2005) A Fast Subpixel Edge Detection Method Using Sobel-Zernike Moments Operator. Image and Vision Computing, 23, 11-17. https://doi.org/10.1016/j.imavis.2004.07.003

[17] Dong, W. and Zhou, S.S. (2008) Color Image Recognition Method Based on the Prewitt Operator. 2008 International Conference on Computer Science and Software Engineering. Hubei, 12-14 December 2008, 170-173. https://doi.org/10.1109/CSSE.2008.567
[18] Liu, J.Z., Li, W.Q. and Tian, Y.P. (1991) Automatic Thresholding of Gray-Level Pictures Using Two-Dimension Otsu Method. China., 1991 International Conference on Circuits and Systems, Shenzhen, 16-17 June 1991, 325-327.

[19] Farhizadeh, A., Dehghan-Niri, E., Moustafa, A., Salamone, S. and Whittaker A. (2013) Damage Assessment of Reinforced Concrete Structures Using Fractal Analysis of Residual Crack Patterns. Experimental Mechanics, 53, 1607-1619. https://doi.org/10.1007/s11340-013-9769-7

[20] Massaro, E. and Bagnoli, F. (2014) Epidemic Spreading and Risk Perception in Multiplex Networks: A Self-Organized Percolation Method. Physical Review E, 90, Article ID: 052817. https://doi.org/10.1103/PhysRevE.90.052817

[21] Al Qurishee, M., Wu, W., Atolagbe, B., El Said, S. and Ghasemi, A. (2019) Non-Destructive Test Application in Civil Infrastructure.

[22] Qurishee, M., Iqbal, I., Islam, M. and Islam, M. (2016) Use of Slag as Coarse Aggregate and Its Effect on Mechanical Properties of Concrete. In: Proceedings of International Conference on Advances in Civil Engineering, 475-479.

[23] Qurishee, M.A. (2019) Low-Cost Deep Learning UAV and Raspberry Pi Solution to Real Time Pavement Condition Assessment.

[24] Al Qurishee, M. (2017) Application of Geosynthetics in Pavement Design.

[25] Hasmat, A., Qurishee, M., Iqbal, I., Zaman, M. and Wahid, M. (2018) Effectiveness of Using Slag as Coarse Aggregate and Study of Its Impact on Mechanical Properties of Concrete.

[26] Al Qurishee, M. and Fomunung, I. (2017) Smart Materials in Smart Structural Systems.

[27] Islam, M.A., Sisiopiku, V.P., Ramadan, O.E. and Hadi, M. (2019) A Framework for Performance-Based Traffic Operations Using Connected Vehicle Data. Simulation, 6, No. 8.

[28] Atolagbe, B. (2019) Automatic Mesh Representation of Urban Environments.

[29] Islam, M.A. (2019) A Literature Review on Freeway Traffic Incidents and Their Impact on Traffic Operations. Journal of Transportation Technologies, 9, 504-516. https://doi.org/10.4236/jtts.2019.94032

[30] Al Qurishee, M., Wu, W., Atolagbe, B., El Said, S., Ghasemi, A. and Tareq, S.M. (2008) Wireless Sensor Network and Its Application in Civil Infrastructure.

[31] Islam, M.A. (2018) Intergrading Connected Vehicle Data into the Transportation Performance Measurement Process. The University of Alabama at Birmingham, Birmingham.

[32] Islam, M.A. (2019) Requirements and Challenges Associated with Deployment of Connected Vehicles.

[33] Pudney, C. (1998) Distance-Ordered Homotopic Thinning: A Skeletonization Algorithm for 3D Digital Images. Computer Vision and Image Understanding, 72, 404-413. https://doi.org/10.1006/cviu.1998.0680

[34] Fortune, S. (1987) A Sweepline Algorithm for Voronoi Diagrams. Algorithmica, 2, 153. https://doi.org/10.1007/BF01840357

[35] Choi, W.-P., Lam, K.-M. and Siu, W.-C. (2003) Extraction of the Euclidean Skeleton Based on a Connectivity Criterion. Pattern Recognition, 36, 721-729. https://doi.org/10.1016/S0031-3203(02)00098-5

[36] Jia, J. and Tang, C.-K. (2003) Image Repairing: Robust Image Synthesis by Adaptive nd Tensor Voting. 2003 IEEE Computer Society Conference on Computer Vision
[37] Qurishee, M., Wu, W., Atolagbe, B., Owino, J., Fomunung, I. and Onyango, M. (2020) Creating a Dataset to Boost Civil Engineering Deep Learning Research and Application. *Engineering*, 12, 151-165. [https://doi.org/10.4236/eng.2020.123013](https://doi.org/10.4236/eng.2020.123013)

[38] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2015) Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38, 142-158. [https://doi.org/10.1109/TPAMI.2015.2437384](https://doi.org/10.1109/TPAMI.2015.2437384)