

AUTOMATIC DAMAGE DETECTION OF STONE CULTURAL PROPERTY BASED ON DEEP LEARNING ALGORITHM

Dohyung Kwon¹, Jeongmin Yu*

¹ Dept. of Cultural Heritage Industry, Korea National University of Cultural Heritage – (ehgud1, jmyu)nuch.ac.kr

Commission II, WG II/8

KEY WORDS: Damage detection, Deep learning, Disaster management, Damage classification.

ABSTRACT:

Outdoor stone cultural properties are continuously affected by the external environment such as wind, rain, and earthquakes. These cause damage to the cultural properties by not only threatening structural stability but also damaging the aesthetic value. Quick detection of these damages is important to enable appropriate preservation treatment in terms of cultural property conservation management. Even though conventional manual damage detection methods are widely used, they are limited by manpower, cost, and other external conditions. In this paper, we propose a system that automatically detects and classifies damage occurring in cultural properties using deep-learning technique to settle these drawbacks. In detail, the damages are classified into four types (i.e., crack, loss, detachment, biological colonization) based on Faster region-based convolutional neural network (R-CNN) algorithm. In addition, we construct an image dataset of stone damage, which is collected by the regular report of the National Designated Cultural Property in 2017 conducted by the Cultural Heritage Administration of S. Korea, and augment its dataset to enhance damage detection performance. From the experiment conducted, we achieved an average confidence score of 94.6% or more on the 20 test images.

1. INTRODUCTION

Outdoor stone cultural properties frequently lose their original appearance due to physical, chemical, and biological weathering (ICOMOS-ISCS, 2008). In addition, natural disasters (e.g., typhoons, earthquakes, floods) and artificial disasters (e.g., arson, damage, graffiti, and environmental degradation) are threats to cultural properties.

The damage caused by these factors does not only threaten the structural stability but also develops into blistering, peeling, fragmentation, bursting, etc., thus damaging the prototype and aesthetic value of the cultural property. Therefore, it is important to effectively detect damages to cultural properties through continuous monitoring, analyzing the causes, and enabling appropriate conservation treatment.

Recently, the visual inspections, such as photogrammetry and laser scanning (A. Oliveira, 2012) (A. Manzo, 2019), infrared thermal imaging techniques, and ground-penetrating radar (D. Angelis et al, 2017) (M. Manataki et al, 2018) (B. Johnston et al, 2018) are widely used to detect damage to cultural properties.

In particular, after acquiring image and point cloud data of cultural property, the damage detection by photogrammetry and laser scanning is analyzed by a 3D-textured mesh model. Some disadvantages of this method are as follows: it takes a long time to perform the modeling, and it is difficult to identify the damage if the accuracy of the constructed model is lowered.

The infrared thermal technique detects damage or cracks by expressing the temperature distribution obtained by measuring the infrared ray emitted from the target as a color image. The active/passive thermal infrared technique can be greatly distinguished based on the heat source. Both techniques can detect damage effectively, but there is a great deal of noise from the outside environment, such as the weather and seasons, which causes damage by injecting heat sources into the surface. The ground-penetrating radar radiates electromagnetic waves in the range of 60 MHz–8 GHz to the cultural property and analyzes the

returning waves to grasp the internal state of the stratum or cultural property. Although other damage detection methods are fast and accurate, this method requires a lot of experience and huge data to receive a lot of external environment and reliable analysis.

These existing damage detection methods are time consuming and costly, and there is a limitation in that they should be carried out by expert in field. In addition, the results of inspection are affected by external conditions such as weather and seasons.

To deal with this limitation, we propose a damage detection system using deep-learning technique and data augmentation method. The main contributions of this paper can be summarized as follows. Firstly, we adopt deep-learning technique for automatically detecting damage of cultural heritage. Furthermore, we construct an image database which is composed of the deterioration patterns that can often occur in cultural properties into four categories and augment the dataset to improve the performance of damage detection and test image taken in various environments.

The rest of this paper is organized as follows. Section 2 briefly reviews deep-learning methods for damage and object detection. In Section 3, a deep-learning system for automatic damage detection is proposed. Section 4 describes the implementation of the proposed method, and finally, Section 5 presents conclusions and future works.

2. RELATED WORKS

2.1 Damage Detection Methods

Damage detection *via* deep learning is mainly used to detect defects in a concrete structure, road pavement damage, and various industrial facilities (Cha et al, 2017). They applied the sliding window technique, and it is easy to scan images larger than 256×256 pixels; the trained network showed an accuracy of about 98% in the verification experiment. Also, it showed

* Corresponding author

robust detection performance without being influenced by lighting condition, shadow, image quality, camera specifications, and distance. (Z. Lei et al. 2016) carried out quantitative evaluation of cracks 3264 × 2448 pixels image datasets, which showed detection with an accuracy of about 90% (Chen et al, 2018). This greatly reduced the amount of time previously spent manually and robustness in a complex environment.

2.2 Deep Learning for Object Detection

Deep learning for object detection frequently uses region-based convolution neural network (R-CNN) series (R. Girshick et al, 2014). R-CNN uses selective search to generate region proposals. The detection performance is greatly improved by using the CNN feature based on region. Furthermore, SPPNet in (Kaiming He et al, 2015) is input regardless of the size of the image, and after the convolutional layers have been passed through the entire image, they are input to fully connected layers through a spatial pyramid pooling process. As a result, performance is improved. Afterward, Fast R-CNN (R. Girshick et al, 2015), which does not use additional disk space for feature caching, but has higher accuracy than R-CNN or SPPNet, was developed to update the results learned in all layers, improving the shortcoming of previous models.

3. PROPOSED METHOD

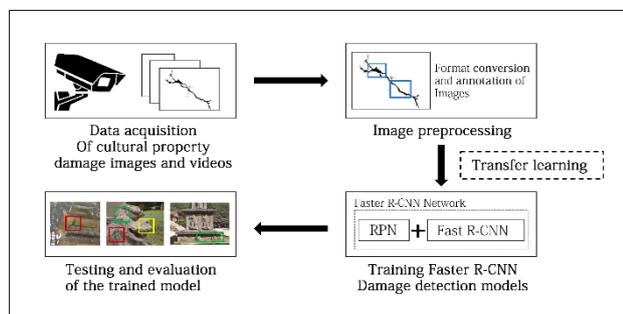


Figure 1. The workflow of the proposed method.

This paper proposes a deep-learning-based damage detection system to detect the damage area from the image of outdoor stone cultural property. Recently, various deep-learning algorithms are being developed for object detection. Faster R-CNN (S. Ren et al, 2017), which is a region-based detection model and somewhat slower to detect than recently developed methods but has high accuracy, has been adopted. The process of detecting the damage is shown in Figure 1. First, after acquiring an image captured by CCTV or camera, a training set is generated by performing format conversion and annotation of an image file to train a damage detection model. Then, a new learning model is created by transfer learning the Faster R-CNN model of the pre-trained inception v2 structure. Lastly, we detect the damage area and classify damage type through the learning model and display it as boundary box and score.

3.1 Database Construction

Deep-learning models require various training and test datasets for accurate feature extraction and recognition. However, since there are no datasets suitable for cultural property, we built a cultural property damage dataset for the experiment.

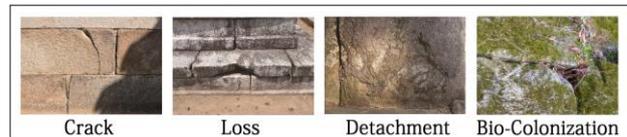


Figure 2. Damage type of dataset.

We categorized the four types of damage as follows: crack, loss, detachment, and biological colonization with reference to (ICOMOS-ISCS, 2008). The images were obtained from the regular report of the National Designated Cultural Property in 2017 conducted by the Cultural Heritage Administration of S. Korea. A total of 3,335 images were obtained, and 100 pieces were extracted for each type of damage.

3.2 Dataset Augmentation

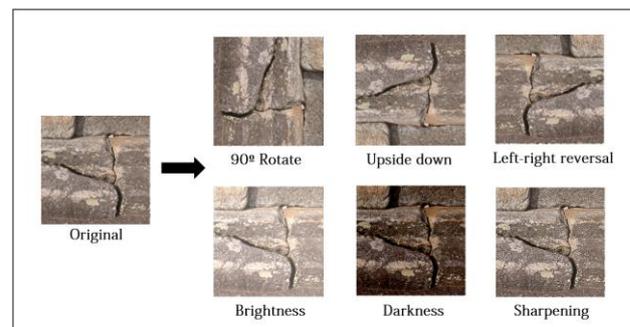


Figure 3. Augmentation training dataset.

One of the important factors that can improve the performance of the model by learning a deep-learning algorithm is the amount of training data. Therefore, we apply image dataset augmentation to learn the model for superior performance using a small amount of data.

The Faster R-CNN model does not care about the size of the input image, so we did not use the image segmentation, only the image transformation. One hundred existing images of each type were transformed into a total of six methods, as shown in Figure 3. The dataset was implemented with 90° rotation, upside down, and left and right reversal to prepare for various types of damage. Brightness of 50% and darkness of 50% were applied, considering the external weather and ambient noise at the time of image acquisition. Finally, image sharpening was performed.

3.3 Damage Detection via Faster R-CNN

Faster R-CNN is a form of Fast R-CNN with region proposal network added. The first feature is extracted through CNN to generate a feature map. The generated feature map is shared between region proposal network (RPN) and region of interest (ROI) pooling layer. RPN is a network that suggests a candidate region where an object may exist on an image. The extracted feature map is used to make a sliding window using nine anchor boxes with three scales and three ratios different from each other. Through this process, the probability of the object and the coordinates for the bounding box are generated for k -candidates, and candidate regions are created based on the generated probability and transmitted to the ROI pooling layer.

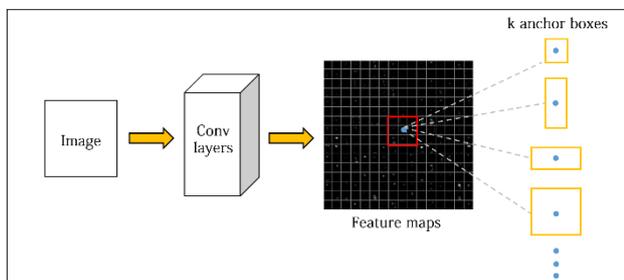


Figure 4. Illustration a process of region proposal network.

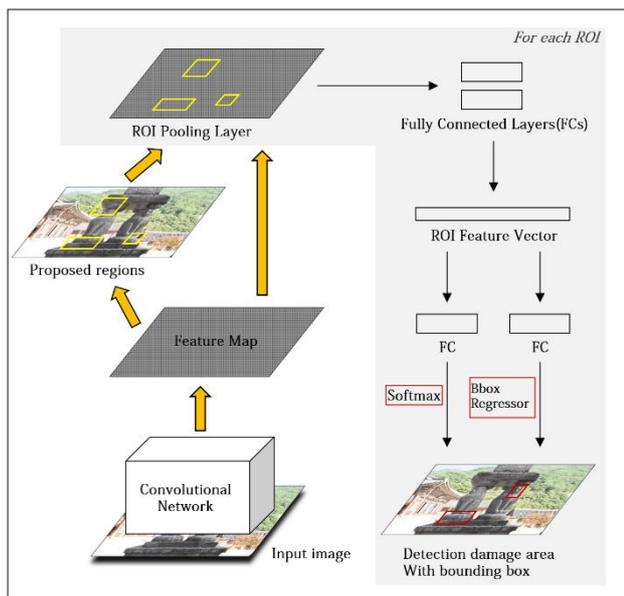


Figure 5. Process a faster R-CNN for damage detection.

During an RPN run, all anchors must be separated by foreground (positive anchor) and background (non-positive anchor) for the learning area. When the classification value of each anchor is P^* , the formula is given as follows:

$$P^* = \begin{cases} 1 \text{ (positive)}, & IoU > 0.7 \\ 0 \text{ (non-positive)}, & IoU < 0.3 \end{cases} \quad (1)$$

Intersection over union (IoU) means the width of the intersection area of the two areas divided by the value of the total area, and ground truth box means the actual box coordinates. The formula is given as follows:

$$IoU = \frac{\text{anchor} \cap \text{ground truth box}}{\text{anchor} \cup \text{ground truth box}} \quad (2)$$

In Faster R-CNN, we can summarize the loss function by adding the class classification loss function and the bounding box loss function with the above definition.

$$L(p_i, t_i) = \frac{1}{N_{cls}} \sum L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum L_{reg}(t_i, t_i^*) \quad (3)$$

where i = index of anchor,
 p_i = The predicted class probability of anchors,
 p_i^* = Ground truth label recall precision,
 t_i = Predicted bounding box,
 t_i^* = Ground truth box,
 λ = Balancing parameter.

i is the index of anchor, p_i is the probability that the object is detected by anchor, and p_i^* is the ground truth label, 1 for object, 0 for non-object. t_i is the coordinate vector of the predicted bounding box and t_i^* is the parameter of the ground truth box. λ is the default, typically 10. L_{cls} is the log loss for predicting the object or background, and L_{reg} is the location prediction loss. Also, N_{cls} is the same normalization value as the mini batch size, and N_{reg} is the same normalization value as the number of other locations.

4. EXPERIMENTAL RESULT

In this paper, we test various types of damage to a cultural property using the existing model. For the experiment, we conducted 100 train image datasets annotated by each type and composed of 20 test datasets. In addition, to improve the performance of the learning model, a comparative study will be conducted by augmenting the crack trainset among cultural property damage types.

4.1 Experimental Settings

Devices	Specification
CPU	Intel i9-7900X
GPU	GeForce GTX 1080 Ti
OS	Window 10 64x
Toolkit	TensorFlow 1.5, cuDNN 7.0.5, CUDA 9.0, Open CV 3.0
Language	Python

Table 1. System configurations for the experiment.

In this paper, the experimental environment for generating and evaluating a new detection model through Faster R-CNN based transfer learning is shown in Table 1. Learning and experiments were conducted using CUDA 9.0, cuDNN 7.0.5 and Open CV 3.0 in window based TensorFlow environment.

4.2 Image Pre-processing

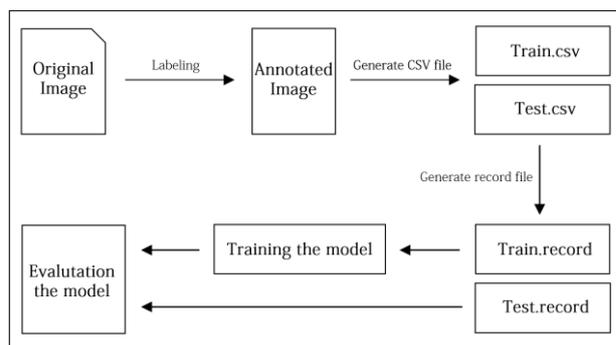


Figure 6. Workflow of image pre-processing.

The image pre-processing is required to use the RGB images and the image collected by the camera for training model. The image pre-processing procedure is shown in Figure 6. The images are labeled with an object on the image using the graphic image annotation tool, i.e., LabelImg, which is shown in Figure 7. Then, train and test comma-separated value (CSV) file are generated and converted into TFRecord file, which is a file format suitable for TensorFlow streaming, to perform detection model learning.

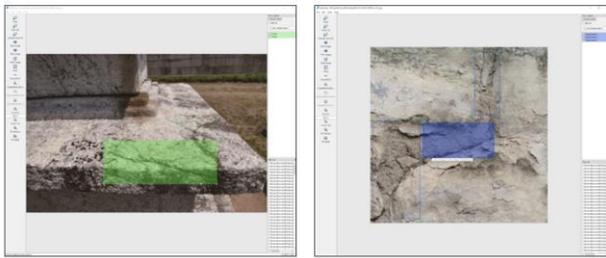


Figure 7. Image annotation using LabelImg.

4.3 Experimental Result

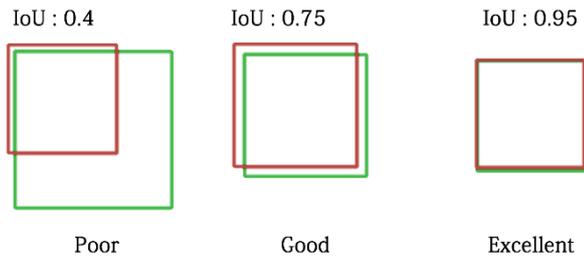


Figure 8. Illustration of IoU measure.

We conducted an experiment to detect four types of damage (i.e., crack, loss, detachment, biological colonization) occurring in cultural property and to perform image augmentation of cracks in damage type and compare the result before and after image processing. We conduct the quantitative evaluation as an evaluation measure which is illustrated in Figure 8. The cultural property damage detection model receives the input image, displays the bounding box in the part where the object is expected to be present, and displays the confidence score together with the classified damage class. For example, IoU is used to evaluate the overlap between predicted bounding boxes and ground truth bounding boxes, and they are displayed as scores. In this experiment, we set the threshold for IoU to 0.5. That is, if the ratio of the overlapping region between the prediction bounding box and the grounding bounding box exceeds 50%, the two bounding boxes are matched and the reliability score of the prediction bounding box is evaluated. The evaluation results are as follows.

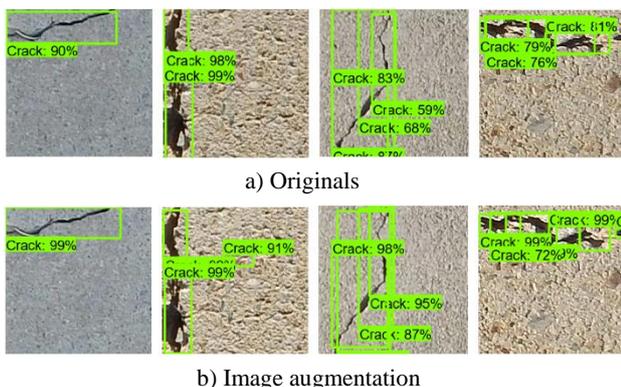


Figure 9. Detection results of crack.

As a result of quantitative evaluation of the crack detection using the bounding box and the score, it was confirmed that the image augmentation improved the score by an average of 17.5% when the amount of training dataset was increased by seven times



Figure 10. Detection results of loss detection.

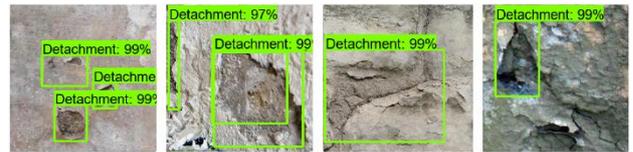


Figure 11. Detection results of detachment.

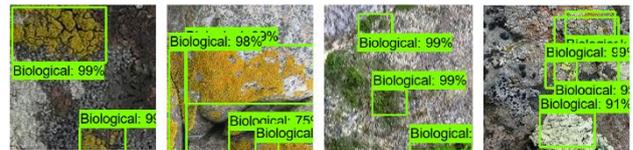


Figure 12. Detection results of biological colonization.

We confirm that the confidence score is higher in the other three types of experiments. In loss of damage type, where the damaged part is noticeable in the image, the score is detected up to 99% and the bounding box is drawn. However, if the edge of the loss part or other foreign matters are mixed, the score is detected as 50% and the damage is not detected in relatively small area. Also, detachment confirmed the detection of the extent of damage in the wide area of the damaged area, or the extent of the difference between the peripheral part and the shape, with an average of 98.67%.

5. CONCLUSION AND FUTURE WORKS

In this paper, we propose an automatic detection method using deep learning. Based on the highly accurate Faster R-CNN algorithm among multiple object detection networks, four types of damage dataset were constructed to detect damage location and extent and automatically classify the damage. To improve the detection performance, we used the image augmentation technique and the bounding box score increases by 17.5% on average. This study can improve the limitations of the existing methods in the field of damage detection of cultural properties. In addition, it is important to introduce a new conservation management method that can automatically detect early damage and enable preservation treatment in a short period of time, thereby continuing the value and beauty of cultural properties. In addition, since the damage area is small or boundary edges are not clear, future studies will be conducted to improve the accuracy of the damage detection, a study to detect other damages will be conducted, and also will conduct damage detection studies using images that are difficult to detect.

ACKNOWLEDGEMENTS

This work was supported by "2019 Graduate R&D Support Project at Korea National University of Cultural Heritage".

REFERENCES

Angelis, D., Tsourlos, P., Tsokas, G., Vargemezis, G., Zacharopoulou, G., 2017. Accessing a historic wall structure using GPR. The case of Heptapyrgion fortress Thessaloniki Greece. *2017 9th International Workshop on Advanced Ground Penetrating Radar (IWAGPR)*, pp. 1-6.

Chen, J., Liu, Z., Wang, H., Nunez, A., Han, Z., 2018. Automatic Defect Detection of Fasteners on the Catenary Support Device Using Deep Convolutional Neural Network. *IEEE Transactions on Instrumentation and Measurement*, 67(2), pp. 257-269.

Girshick, R., 2015. Fast R-CNN. *IEEE International Conference on Computer Vision (ICCV)*, pp. 1440-1448.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2016. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(1), pp. 142–158.

Vergès-Belmin, V., 2008. Illustrated glossary on stone deterioration pattern. ICOMOS-ISCS, pp. 1-41.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.37, pp. 1904-1916.

Johnston, B., Ruffell, A., Mckinley, J., Warke, P., 2018. Detecting voids within a historical building facade: A comparative study of three high frequency GPR antenna. *Journal of Cultural Heritage*, 32, pp. 117-123.

Lei, Z., Fan, Y., Yimin, D, Z., Ying, J, Z., 2016. Road crack detection using deep convolutional neural network. *IEEE International Conference on Image Processing*, pp. 3708-3712.

Manataki, M., Sarris, A., Oikonomou, D., Simirdanis, K., Strapazon, G., Trapero Fernandez, P., 2018. Contribution of GPR method in monitoring and evaluating the conservation state of Fortezza, Rethymno, Greece. *2018 17th International Conference on Ground Penetrating Radar (GPR)*, doi.org/10.1109/icgpr.2018.8441632.

Manzo, A., Chesi, C., 2019. Building Techniques and Structural Damage of Historical Constructions Detected Through the Point Cloud Survey. *Structural Analysis of Historical Construction*, pp. 135-143.

Oliveira, A., Oliveira, J, F., Pereira, J,M., Araujo, B, R., Boavida, J., 2012. 3D modelling of laser scanned and photogrammetric data for digital documentation: the Mosteiro da Batalha case study. *Journal of Real-time Image Processing*, 9(4) pp. 673-688

Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), pp. 1137–1149.

Y. J. Cha., W. Choi., Büyüköztürk, O., 2017. Deep Learning Based Crack Detection Using Convolutional Neural Networks. *Computer-Aided Civil and Infrastructure Engineering*, vol,32, no.5, pp. 361-378.