

## Defect-Detection Model for Underground Parking Lots Using Image Object-Detection Method

Hyun Kyu Shin<sup>1</sup>, Si Woon Lee<sup>2</sup>, Goo Pyo Hong<sup>3</sup>, Lee Sael<sup>2</sup>, Sang Hyo Lee<sup>4</sup> and Ha Young Kim<sup>5,\*</sup>

<sup>1</sup>Architectural Engineering, Hanyang University, ERICA, Ansan, 15588, Korea

<sup>2</sup>Department of Artificial Intelligence, Ajou University, Suwon, 16499, Korea

<sup>3</sup>Division of Architecture and Civil Engineering, Kangwon National University, Samcheok, 25913, Korea

<sup>4</sup>Division of Smart Convergence Engineering, Hanyang University, ERICA, Ansan, 15588, Korea

<sup>5</sup>Graduate School of Information, Yonsei University, Seoul, 03722, Korea

\*Corresponding Author: Ha Young Kim. Email: haimkgetup@gmail.com

Received: 03 September 2020; Accepted: 22 September 2020

**Abstract:** The demand for defect diagnoses is gradually gaining ground owing to the growing necessity to implement safe inspection methods to ensure the durability and quality of structures. However, conventional manpower-based inspection methods not only incur considerable cost and time, but also cause frequent disputes regarding defects owing to poor inspections. Therefore, the demand for an effective and efficient defect-diagnosis model for concrete structures is imminent, as the reduction in maintenance costs is significant from a long-term perspective. Thus, this paper proposes a deep learning-based image object-identification method to detect the defects of paint peeling, leakage peeling, and leakage traces that mostly occur in underground parking lots made of concrete structures. The deep learning-based object-detection method can replace conventional visual inspection methods. A faster region-based convolutional neural network (R-CNN) model was used with a training dataset of 6,281 images that utilized a region proposal network to objectively localize the regions of interest and detect the surface defects. The defects were classified according to their type, and the learning of each exclusive model was ensured through test sets obtained from real underground parking lots. As a result, average precision scores of 37.76%, 36.42%, and 61.29% were obtained for paint peeling, leakage peeling, and leakage trace defects, respectively. Thus, this study verified the performance of the faster RCNN-based defect-detection algorithm along with its applicability to underground parking lots.

**Keywords:** Faster R-CNN; deep learning; defect detection; concrete structures

### 1 Introduction

The quality control of civil structures has become an emerging public issue and has escalated the importance of defect diagnosis of structures. In underground parking lots made of concrete structures, defects are generated on the surface of the structure after a long period of repeated contraction and



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

expansion caused by the constant pressure from the surrounding water and earth and its own material characteristics. Therefore, frequent leakages occur in these parking lots owing to penetration of underground water into the surface defects. As leakage of underground structures is a factor that adversely affects their durability, it is important to prevent defects that cause leakages [1]. Waterproofing of underground parking lots is mostly carried out with liquid waterproofing materials; however, their utility is diminished when the surface is damaged by defects such as leakage and peeling on the surface of the structures, even for small damages. Thus, surface defects in underground parking lots are a critical problem, which undermines the structural safety and deteriorates the building quality [2]. Therefore, periodic inspections and maintenance are necessary for the early detection of such surface defects and ensuring building quality and durability through appropriate repair. However, in conventional methods, routine inspections are often neglected owing to time and cost issues, despite the fact that surface defects can be easily identified with the naked eye.

To solve such problems, an automated inspection technology has been recently developed, which can replace the traditional human inspection method. For instance, the defect or deformation of a structure can be detected using an ultrasonic transducer [3,4], or by receiving an abnormal signal from the installed sensors in real time [5–7]. Although these methods have the advantage of locating and monitoring defects, expensive equipment is required to transmit and receive signals that use a predefined threshold to determine the defect, thus making defect-detection very challenging for diverse defects. Therefore, research has been conducted to replace conventional methods with techniques that use the same principle as human visual inspection, i.e., based on real image data instead of electrical signals. For instance, Wang et al. [8] classified the damage of Masonry Historic Structures based on defect types using image data, and Cha et al. [9] used a deep learning-based image analysis model to detect surface crack damages of a concrete structure. Furthermore, Liang [10] proposed an image-based post-disaster inspection method for a reinforced concrete bridge system.

These studies experimentally proved that deep learning-based image processing models can exhibit excellent results in detecting and analyzing concrete defects with diverse patterns in various environments. This application of deep learning-based object detection has recently become popular. By recognizing and learning various objects in an image, it can automatically extract and analyze the representative features of the image. Thus, this method can be effectively applied to identify atypical features such as defects on the surface of a structure. Therefore, this study proposed a method to detect surface defects in underground parking lots using a deep learning-based object-detection method and experimentally verified its applicability.

## **2 Literature Review**

### ***2.1 Necessity of Managing Defects in Underground Parking Lots***

Leakage in underground parking lots is caused by three factors: design, construction, and maintenance [11]. The outer wall of a basement is subjected to continuous external forces owing to water pressure and earth pressure being exerted simultaneously; therefore, water leakage can easily occur even when a small waterproofing defect occurs during the building operation phase. In addition, water can penetrate and spread easily when a defect occurs on the surface of a structure because the concrete constituting the underground structure is a permeable material. Neglecting defects can lead to serious financial, functional, and structural problems [12]. Therefore, this paper proposes an efficient image-based detection method for defects caused by maintenance factors occurring during the operation of underground parking lots.

## **2.2 Related Research on Image-Based Defect Analysis**

Several studies have attempted to convert conventional manpower-based inspections to computer-aided inspections, such as by automation of safety inspection and defect diagnosis, to ensure efficient maintenance of structures. In particular, studies on automatically detecting and classifying defects on the surface of a structure based on image data exist [13–25].

Image processing techniques (IPTs) for the inspection and monitoring of structural safety have been extensively investigated in combination with various computing technologies. To identify the characteristics of superficial damages on building and infrastructure, previous studies have presented damage detection methods based on several IPTs such as Gabor filters and histogram of oriented gradient (HOG) and machine learning-based methods such as support-vector machine [13], histogram shape-based thresholding [14], and region-based segmentation [15].

Moreover, numerous researchers have focused on detecting and recognizing specific damage on the surface of structures. For instance, Yeum et al. [16] proposed a sliding-window technique to identify cracks near bolts on steel structures with various channel image features extracted using HOG. German et al. [15] introduced a local-entropy-based threshold to extract spalling damage maps. Koch et al. [14] proposed automated pothole detection in asphalt pavement images using a histogram shape-based threshold. Their results demonstrated that each method could detect and identify specific damages in structures such as asphalt pavement, bridges, and infrastructure with reasonable accuracy.

However, the morphological features of digital images extracted by IPT are parameter dependent [17]. Although Cheng et al. [18] proposed an effective method to determine the threshold value in real time, the reliability of this method is significantly affected by the presence of noise in the digital images [15,19,20].

Therefore, to overcome these challenges, a deep learning-based convolutional neural network (CNN) approach has recently been proposed to automatically attain the representative features of concrete-structure damage based on digital images [21]. The greatest advantage of the CNN is it can learn and automatically extract representative features from large-scale image data [21,22]. Thus, CNN-based image analysis has been applied as a method for extracting multi-dimensional features or solving complex problems in various fields, including face recognition, automatic driving, and object recognition tasks [23–25].

## **2.3 Related Research on Image-Based Defect Analysis**

Numerous studies have been conducted in the civil engineering field to detect damage on the surface of structures. For instance, Lin et al. [21] optimized a CNN technique for image analysis, which is resistant to noise and can automatically extract damage features from a low-resolution image to detect damage on structures. In addition, Maeda et al. [26] classified eight types of damage on road surfaces by applying CNN as an image analysis method. Moreover, Yang et al. [22] proposed a fully convolutional network (FCN) in their research, which can efficiently investigate and classify cracks into pixels using a computer image-based analysis method. Furthermore, Wang et al. [8] proposed a damage classification model that combined the CNN model with the sliding-window technique to replace manpower-based inspection of aged concrete brick buildings.

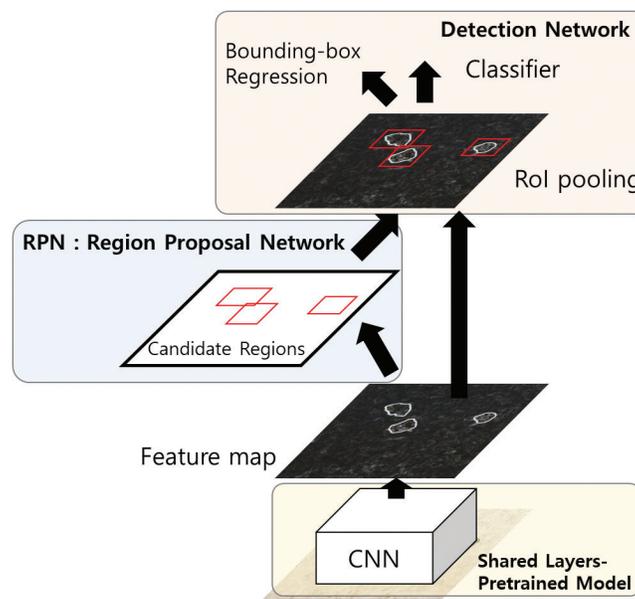
As described above, the analysis based on the CNN model can be optimized to extract various characteristics of the structure surface image data. In addition, several studies have been conducted to generalize the CNN model via customization of the neural network architecture according to the target dataset [8,22]. Interestingly, few studies have been conducted to detect leakage traces or leakage peeling. To address this gap, this study applied the faster region-based CNN (R-CNN) model—optimized for the object recognition method—to detect leakage traces and leakage peeling in underground parking lots.

### 3 Methodology

Various surface defects on the concrete structure, wall patterns of an underground parking lot, and indoor illumination intensity significantly influence the identification of leakage traces and leakage peeling through images. The deep learning-based CNN can solve the problems caused by various image conditions by learning images collected from various environments. Therefore, this study utilized a CNN-based image analysis to overcome the limitations of conventional IPTs and applied the faster R-CNN proposed by Ren et al. [27] for identifying specific objects in an image to detect leakage traces and leakage peeling on a concrete surface.

#### 3.1 Faster R-CNN

The faster R-CNN is an improved object recognition model first proposed by Girshick et al. [28]. It is used to identify the location of an object by proposing a Region of Interest (RoI). The advanced object detection model is advantageous in providing more sensitive damage-detection and reducing the computational burden in the region proposal stages by replacing the selective search approach with a convolution shared network to generate the region proposals [29]. The overall architecture consists of a region proposal network (RPN) and a detection network, as shown in Fig. 1.



**Figure 1:** Schematic operation model of faster R-CNN (adopted from Fang et al. [30])

In addition, Fig. 2 shows the RPN operation process that consists of an FCN for generating the object proposals, including localization information, to deliver better performance in detecting damage locations from the concrete structure images. The probabilities of candidate objects were calculated using region boundary box coordinates by applying the sliding window approach with the spatial window filter and anchor box parameters to extract the RoI maps [27,31]. The region proposals of the RoI pooling were fed into the second module—the detection network—to refine the object localization proposals. The detection network shared the same convolutional layer with the RPN, as shown in Fig. 3.

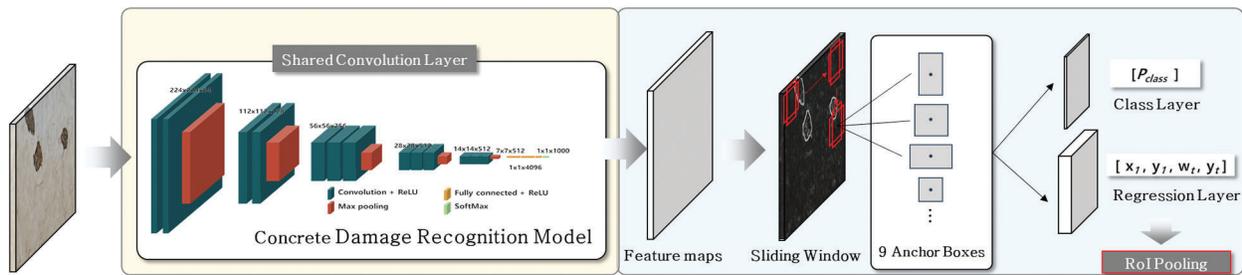


Figure 2: Region proposal network operation process

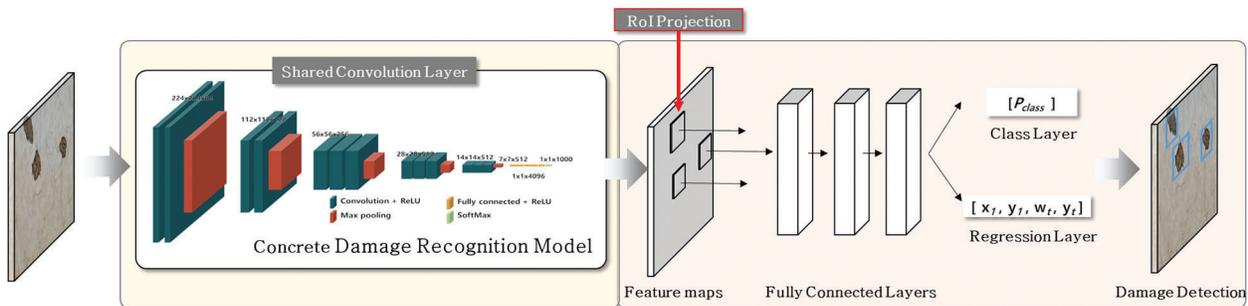


Figure 3: Detection network operation process

### 3.2 Development of Underground Parking Lot Defect Image Dataset

A total of 9,815 images and 14,842 annotations were collected using smartphones and digital cameras in an underground parking lot to constitute the training dataset. In this study, the data collected for paint peeling, leakage peeling, and leakage traces in the underground parking lot were used to optimize the concrete damage detection model as shown in Fig. 4; the results are shown in Tab. 1.

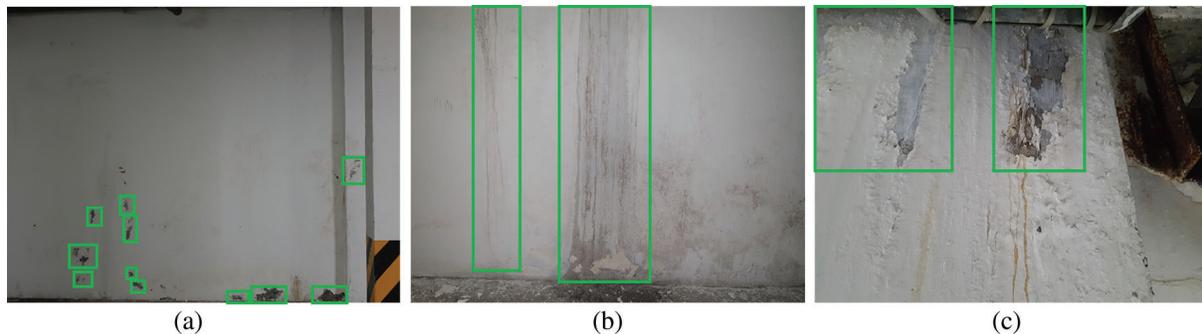


Figure 4: Examples of defect image dataset. (a) Class 1: Paint peeling, (b) Class 2: Leakage trace, (c) Class 3: Leakage peeling

Table 1: Data distribution per class

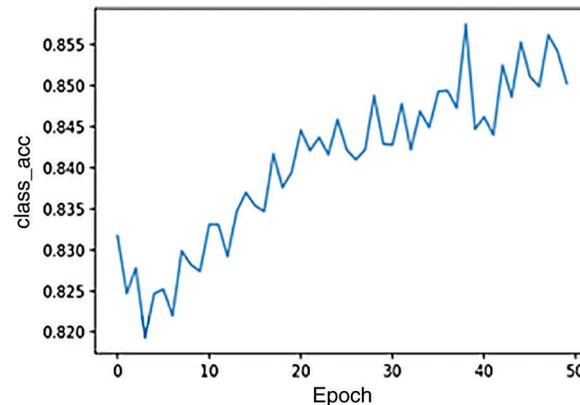
Category	Training data	Verification data	Test data	Total
Class 1 (paint peeling)	5299	1325	1656	8280
Class 2 (leakage traces)	164	41	51	256
Class 3 (leakage peeling)	818	205	256	1279
Class 4 (leakage traces)	164	41	51	256

However, owing to the small number of images collected, the data augmentation strategies used horizontal and vertical flipping with orthogonal rotation strategies in the image generation stages. As a result, the number of images in the dataset was expanded to 157,000 through data augmentations.

#### 4 Implementation and Results

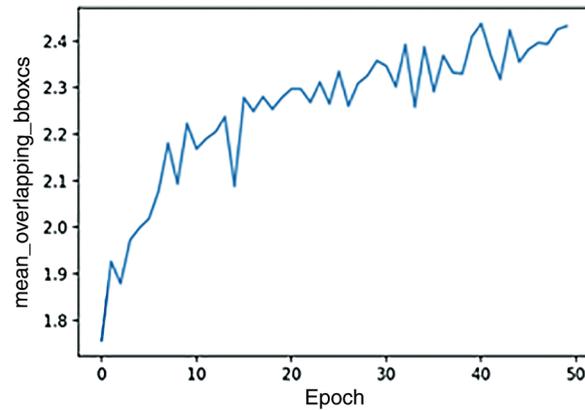
In this study, experiments were performed using the Keras platform on a workstation with a GPU (GeForce GTX 1080  $\times$  8) and a CPU (Intel Xeon E5-2630 v4 @ 2.2 GHz  $\times$  20). In this experiment, several key hyper-parameters were used in their default setting in the network, as established by Ren et al. [27]. Thus, the network was trained using the Adam Optimizer [32] with a learning rate of 0.0001 for 50 epochs. Moreover, nine anchor shapes were used in the RPN:  $128 \times 128$ ,  $256 \times 256$ , and  $512 \times 512$ , each in three aspect ratios (1:1, 1:2, and 2:1). Thereafter, the performance was evaluated with a test set and other raw images. The region proposals having IoUs greater than 0.5 were adopted for the detection of damages in the target.

Three individual defect-detection models were designed to test the effectivity in detecting each class of defects. Figs. 5–10 show the learning results per epoch unit for the individual models of Class 1 (peeling). Fig. 5 shows the ratio of class matching, i.e., accuracy. Fig. 6 shows the mean overlapping box. Figs. 7 and 8 show the class and RPN loss curves for the Cls layer of faster R-CNN, respectively. Figs. 9 and 10 show the class and RPN loss curves of the Regr layer of faster R-CNN, respectively. The success of learning is evident from these graphs.

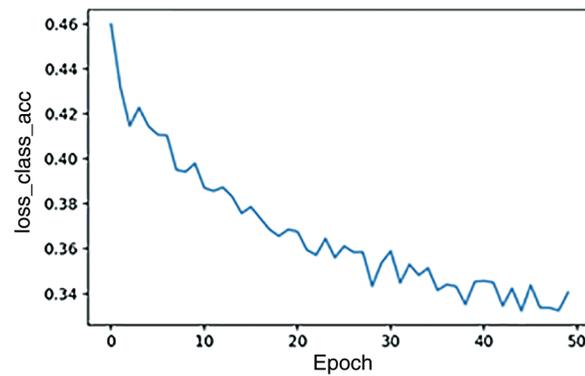


**Figure 5:** Accuracy of Class 1 (paint peeling)

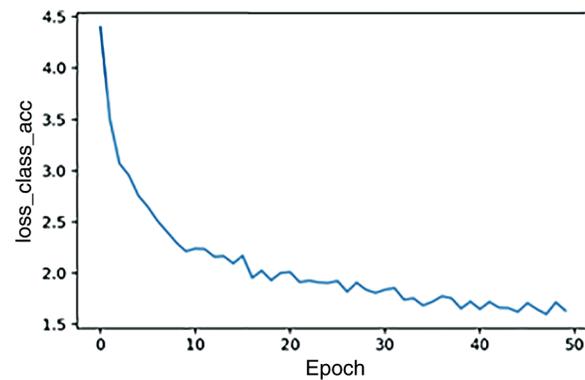
Similarly, Figs. 11–16 show the learning results per epoch unit for the individual models of Class 2 (leakage traces). Fig. 11 shows the accuracy of class matching. Fig. 12 shows the mean overlapping box. Figs. 13 and 14 show the class and RPN loss curves for the Cls layer of faster R-CNN. Figs. 15 and 16 show the class and RPN loss curves of the Regr layer of faster R-CNN. The success of learning was evidently confirmed from these graphs.



**Figure 6:** Mean overlapping box of Class 1 (paint peeling)

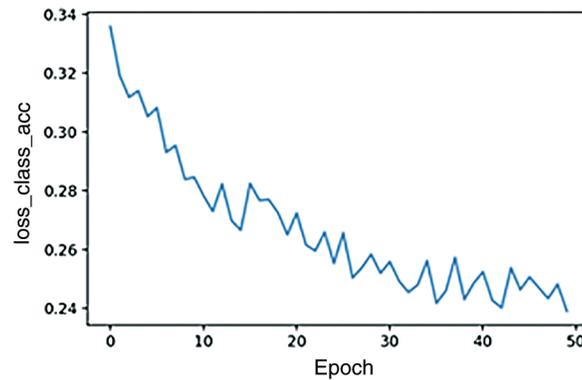


**Figure 7:** Cls layer class loss curve of Class 1 (paint peeling)

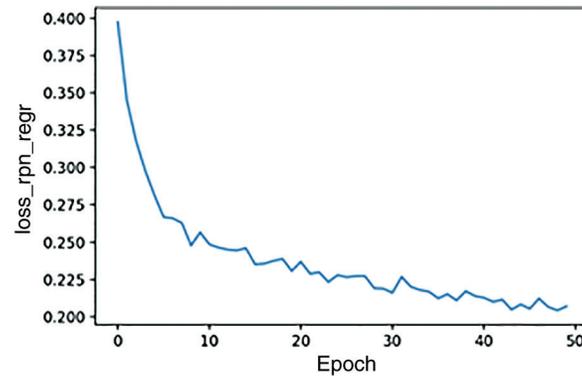


**Figure 8:** Cls layer RPN loss curve of Class 1 (paint peeling)

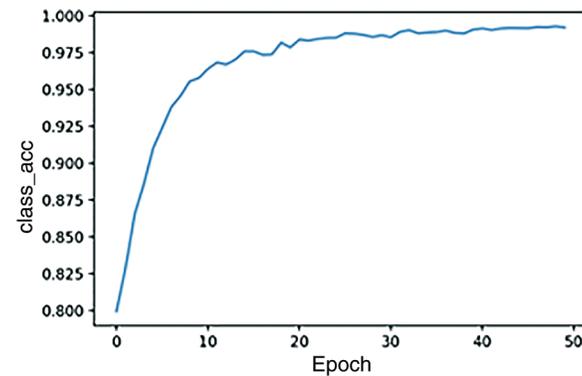
Furthermore, Figs. 17–22 show the learning results per epoch unit for the individual models of Class 3 (leakage peeling). Fig. 17 shows the accuracy of class matching. Fig. 18 shows the mean overlapping box. Figs. 19 and 20 show the class and RPN loss curves for the Cls layer of faster R-CNN, respectively. Figs. 21 and 22 show the class and RPN loss curves of the Regr layer of faster R-CNN, respectively. The learning success was confirmed from these graphs.



**Figure 9:** Regr layer class loss curve of Class 1 (paint peeling)



**Figure 10:** Regr layer RPN loss curve of Class 1 (paint peeling)

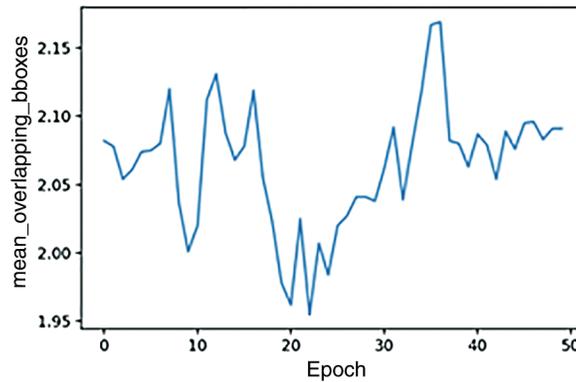


**Figure 11:** Accuracy of Class 2 (leakage trace)

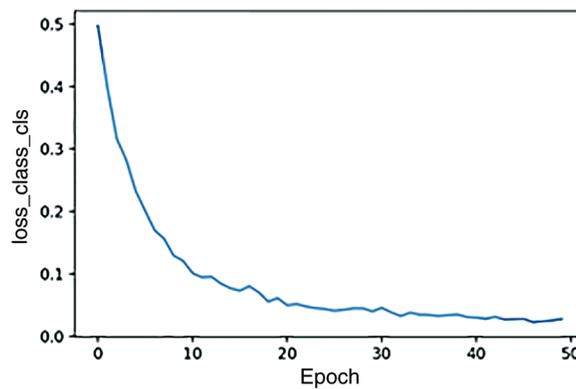
In addition, Figs. 23–25 show the results for the test set of the detection model. The blue box shows the prediction results for the location and type of a defect using faster R-CNN. As shown in the graph, notable results were obtained from certain images.

An examination of the detection performance using the average precision (AP) is presented in Tab. 2. This experiment demonstrated the feasibility of visual inspection of concrete damage in practical environments. In this study, the performance measurement was based on the AP, as individual defect models were developed for each class. The reference IoU was set to 50% for measuring the AP. The

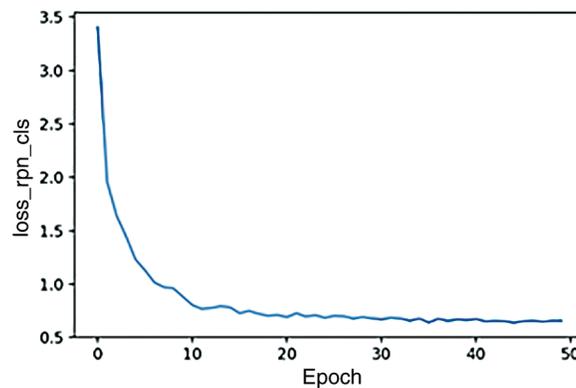
overall results are presented in [Tab. 2](#). Our proposed model achieved the AP scores of 37.76%, 61.29%, and 36.42% for the exclusive models of Class 1 (paint peeling), Class 2 (leakage traces), and Class 3 (leakage peeling), respectively. As reflected above, despite having a small learning dataset, the detection of patterns of Class 2 data was more accurate owing to their conspicuity when compared with other classes.



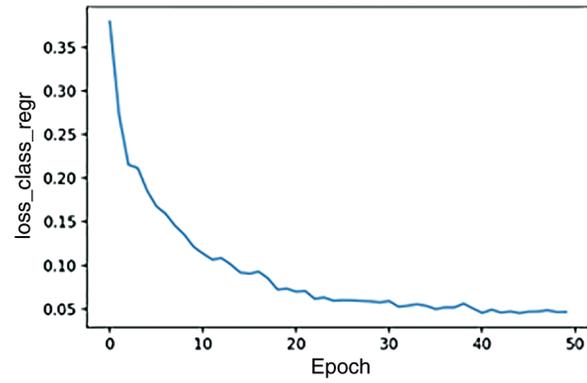
**Figure 12:** Mean overlapping box of Class 2 (leakage trace)



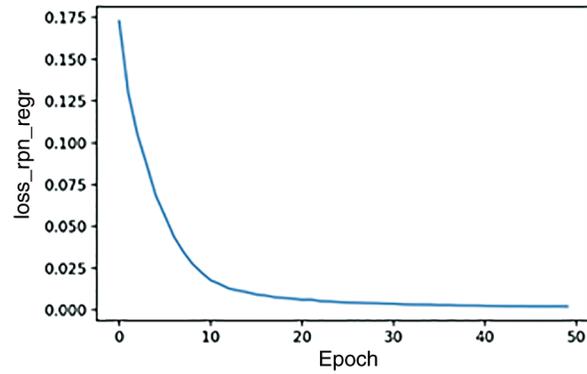
**Figure 13:** Cls layer class loss curve of Class 2 (leakage trace)



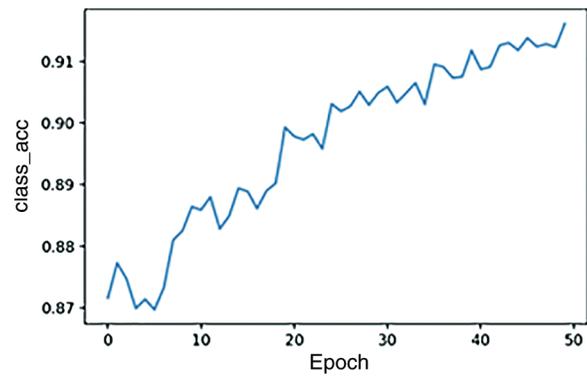
**Figure 14:** Cls layer RPN loss curve of Class 2 (leakage trace)



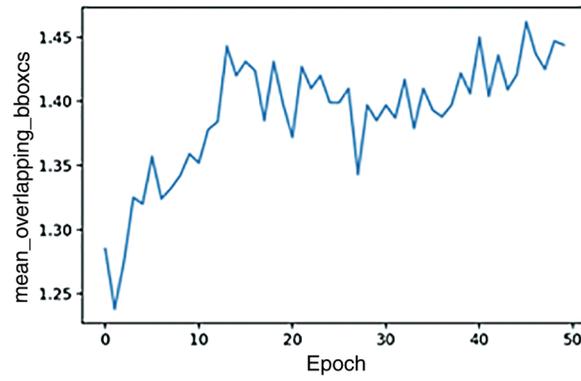
**Figure 15:** Regr layer class loss curve of Class 2 (leakage trace)



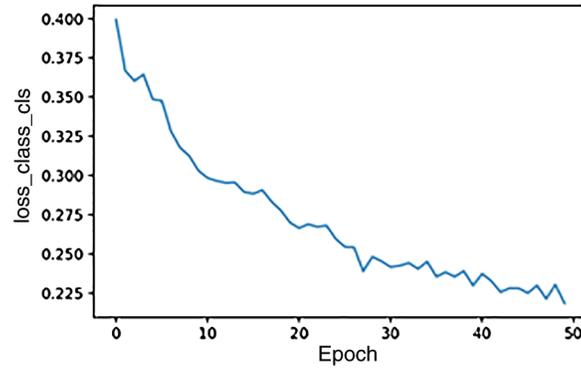
**Figure 16:** Regr layer RPN loss curve of Class 2 (leakage trace)



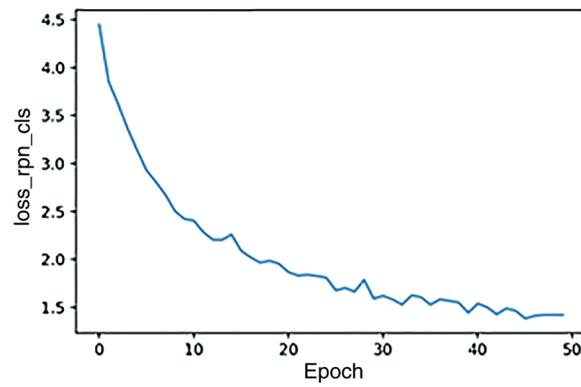
**Figure 17:** Accuracy of Class 3 (leakage peeling)



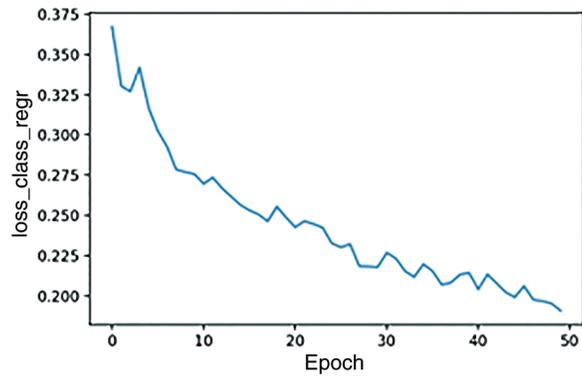
**Figure 18:** Mean overlapping box of Class 3 (leakage peeling)



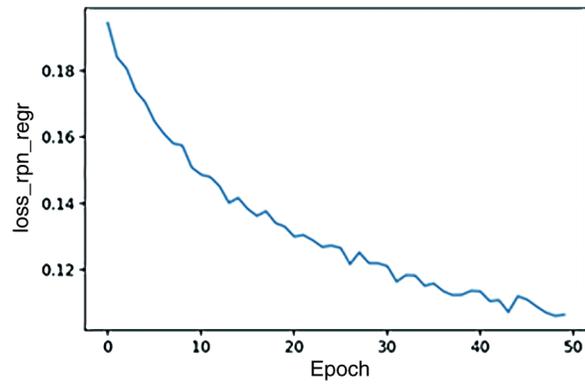
**Figure 19:** Cls layer class loss curve of Class 3 (leakage peeling)



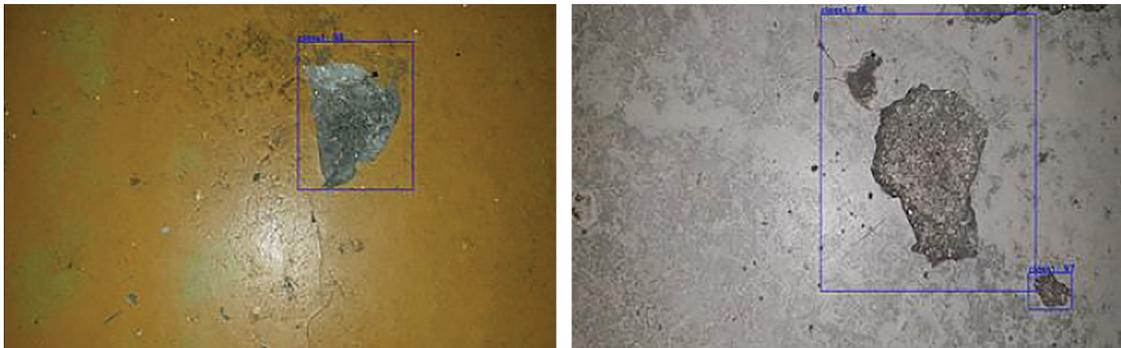
**Figure 20:** Cls layer RPN loss curve of Class 3 (leakage peeling)



**Figure 21:** Regr layer class loss curve of Class 3 (leakage peeling)



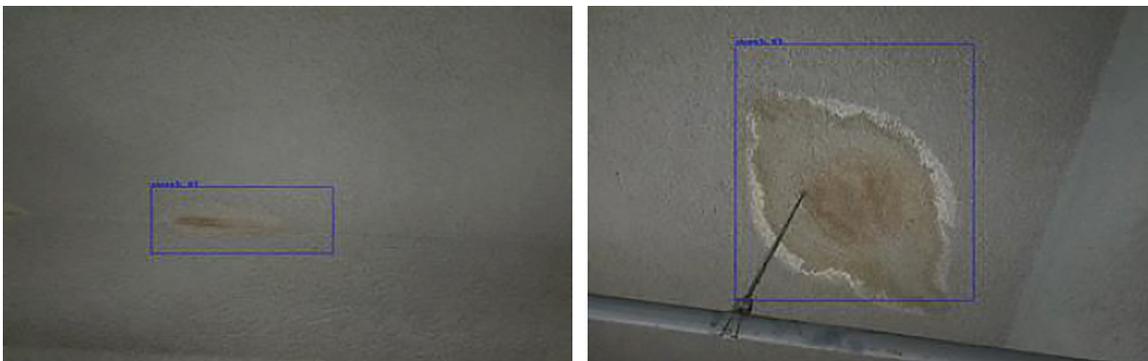
**Figure 22:** Regr layer RPN loss curve of Class 3 (leakage peeling)



**Figure 23:** Example of defect detection for Class 1 (paint peeling)



**Figure 24:** Example of defect detection for Class 2 (leakage peeling)



**Figure 25:** Example of defect detection for Class 3 (leakage trace)

**Table 2:** Average Precision (AP) measured per test data reference model

Model	AP (%)
Class 1 (paint peeling) exclusive defect detection model	37.76
Class 2 (leakage traces) exclusive defect detection model	61.29
Class 3 (leakage peeling) exclusive defect detection model	36.42

## 5 Discussion and Conclusion

This study proposed a model that can automatically detect various defects in underground parking lots using deep learning methodology and verified its applicability through experiments. The damage-detection model based on faster R-CNN was used in this experiment, and the observable surface defects in an underground parking lot were classified into three types: paint peeling, leakage peeling, and leakage trace.

Digital cameras and smart phones were used to collect 9,815 defect images with  $4,032 \times 1,960$  resolution from underground parking lots. In this study, 6,281 images were used as training dataset and 1,544 uncropped images were inputted for validation. Another dataset of 1,963 images was prepared to evaluate the performance of the proposed damage-detection model.

As a result, the exclusive detection models for Class 1 (paint peeling) and Class 3 (leakage peeling) showed AP scores of 37.76% and 36.42%, respectively. The best performance was exhibited by the

exclusive model for Class 2 (leakage trace) with an AP score of 61.29%. This proves that a wide range of defects in underground parking lots can be detected using image-based models.

The proposed deep learning-based defect-detection model solved the problems of poor inspection caused by limitations such as expensive diagnosis equipment, labor cost, and lack of manpower with respect to the conventional evaluations of safety and inspection for facility maintenance. In addition, an efficient and proactive facility maintenance system can be implemented using this method. The facility performance-evaluation tool proposed in this study can maximize efficiency and improve accuracy in terms of both time and cost.

**Acknowledgement:** The authors would like to thank the Ministry of Land, Infrastructure and Transport of the Korean government for funding this research project.

**Funding Statement:** This research was supported by a grant (19CTAP-C152020-01) from Technology Advancement Research Program (TARP) funded by the Ministry of Land, Infrastructure and Transport of the Korean government.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] P. Mendes, J. G. Lopes, J. D. Brito and J. Feiteira, "Waterproofing of concrete foundations," *Journal of Performance of Constructed Facilities*, vol. 28, no. 2, pp. 242–249, 2014.
- [2] M. Y. L. Chew and N. De Silva, "Benchmarks to minimize water leakages in basements," *Structural Survey*, vol. 21, no. 4, pp. 131–145, 2003.
- [3] Z. Liu, H. Yu, C. He and B. Wu, "Delamination detection in composite beams using pure Lamb mode generated by air-coupled ultrasonic transducer," *Journal of Intelligent Material Systems and Structures*, vol. 25, no. 5, pp. 541–550, 2014.
- [4] B. Park, H. Sohn, P. Malinowski and W. Ostachowicz, "Delamination localization in wind turbine blades based on adaptive time-of-flight analysis of noncontact laser ultrasonic signals," *Nondestructive Testing and Evaluation*, vol. 32, no. 1, pp. 1–20, 2017.
- [5] B. Spencer Jr, M. E. Ruiz-Sandoval and N. Kurata, "Smart sensing technology: Opportunities and challenges," *Structural Control and Health Monitoring*, vol. 11, no. 4, pp. 349–368, 2004.
- [6] J. Li, K. A. Mechitov, R. E. Kim and B. F. Spencer Jr, "Efficient time synchronization for structural health monitoring using wireless smart sensor networks," *Structural Control and Health Monitoring*, vol. 23, no. 3, pp. 470–486, 2016.
- [7] G. Morgenthal, N. Hallermann, J. Kersten, J. Taraben, P. Debus *et al.*, "Framework for automated UAS-based structural condition assessment of bridges," *Automation in Construction*, vol. 97, pp. 77–95, 2019.
- [8] N. Wang, Q. Zhao, S. Li, X. Zhao and P. Zhao, "Damage classification for masonry historic structures using convolutional neural networks based on still images," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1073–1089, 2018.
- [9] Y. J. Cha, W. Choi and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 361–378, 2017.
- [10] X. Liang, "Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization," *Computer-Aided Civil and Infrastructure Engineering*, vol. 34, no. 5, pp. 415–430, 2019.
- [11] B. H. Cho, B. H. Nam, S. Seo, J. Kim, J. An *et al.*, "Waterproofing performance of waterstop with adhesive bonding used at joints of underground concrete structures," *Construction and Building Materials*, vol. 221, pp. 491–500, 2019.

- [12] M. T. A. Chaudhary, "Structural condition assessment and functional appraisal of a large basement affected by groundwater leakage," *Engineering Structures*, vol. 52, pp. 384–393, 2013.
- [13] P. H. Chen, H. K. Shen, C. Y. Lei and L. M. Chang, "Support-vector-machine-based method for automated steel bridge rust assessment," *Automation in Construction*, vol. 23, pp. 9–19, 2012.
- [14] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Advanced Engineering Informatics*, vol. 25, no. 3, pp. 507–515, 2011.
- [15] S. German, I. Brilakis and R. DesRoches, "Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments," *Advanced Engineering Informatics*, vol. 26, no. 4, pp. 846–858, 2012.
- [16] C. M. Yeum and S. J. Dyke, "Vision-based automated crack detection for bridge inspection," *Computer-Aided Civil and Infrastructure Engineering*, vol. 30, no. 10, pp. 759–770, 2015.
- [17] E. Zalama, J. Gómez-García-Bermejo, R. Medina and J. Llamas, "Road crack detection using visual features extracted by Gabor filters," *Computer-Aided Civil and Infrastructure Engineering*, vol. 29, no. 5, pp. 342–358, 2014.
- [18] H. Cheng, X. Shi and C. Glazier, "Real-time image thresholding based on sample space reduction and interpolation approach," *Journal of Computing in Civil Engineering*, vol. 17, no. 4, pp. 264–272, 2003.
- [19] S. Behnke, "Face localization and tracking in the neural abstraction pyramid," *Neural Computing & Applications*, vol. 14, no. 2, pp. 97–103, 2005.
- [20] T. Yamaguchi and S. Hashimoto, "Fast crack detection method for large-size concrete surface images using percolation-based image processing," *Machine Vision and Applications*, vol. 21, no. 5, pp. 797–809, 2010.
- [21] Y. Z. Lin, Z. H. Nie and H. W. Ma, "Structural damage detection with automatic feature-extraction through deep learning," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 12, pp. 1025–1046, 2017.
- [22] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang *et al.*, "Automatic pixel-level crack detection and measurement using fully convolutional network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1090–1109, 2018.
- [23] D. C. Cireşan, U. Meier, L. M. Gambardella and J. Schmidhuber, "Deep, big, simple neural nets for handwritten digit recognition," *Neural Computation*, vol. 22, no. 12, pp. 3207–3220, 2010.
- [24] F. Eyben, F. Weninger, S. Squartini and B. Schuller, "Real-life voice activity detection with LSTM recurrent neural networks and an application to Hollywood movies," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, 2013.
- [25] C. Szegedy, A. Toshev and D. Erhan, "Deep neural networks for object detection," in *Advances in Neural Information Processing Systems Conf.*, Lake Tahoe, Nevada, USA, 2013.
- [26] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1127–1141, 2018.
- [27] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems*, Montreal, Canada, 2015.
- [28] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Columbus, Ohio, USA, 2014.
- [29] H. Jiang and E. Learned-Miller, "Face detection with the faster R-CNN," in *12th IEEE Int. Conf. on Automatic Face & Gesture Recognition*, Washington DC, USA, 2017.
- [30] Q. Fang, H. Li, X. Luo, L. Ding, H. Luo *et al.*, "Computer vision aided inspection on falling prevention measures for steeplejacks in an aerial environment," *Automation in Construction*, vol. 93, pp. 148–164, 2018.
- [31] Y. Xu, G. Yu, Y. Wang, X. Wu and Y. Ma, "Car detection from low-altitude UAV imagery with the faster R-CNN," *Journal of Advanced Transportation*, vol. 2017, pp. 1–10, 2017.
- [32] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization." *arXiv preprint, arXiv: 1412. 6980*, 2014.