

콘크리트 구조체 균열 탐지에 대한 Mask R-CNN 알고리즘 적용성 평가

Application of Mask R-CNN Algorithm to Detect Cracks in Concrete Structure

배 병 규¹ Bae, Byongkyu 최 용 진² Choi, Yongjin
윤 강 호³ Yun, Kangho 안 재 훈⁴ Ahn, Jaehun

Abstract

Inspecting cracks to determine a structure's condition is crucial for accurate safety diagnosis. However, visual crack inspection methods can be subjective and are dependent on field conditions, thereby resulting in low reliability. To address this issue, this study automates the detection of concrete cracks in image data using ResNet, FPN, and the Mask R-CNN components as the backbone, neck, and head of a convolutional neural network. The performance of the proposed model is analyzed using the intersection over the union (IoU). The experimental dataset contained 1,203 images divided into training (70%), validation (20%), and testing (10%) sets. The model achieved an IoU value of 95.83% for testing, and there were no cases where the crack was not detected. These findings demonstrate that the proposed model realized highly accurate detection of concrete cracks in image data.

요 지

구조물의 상태를 파악하기 위한 균열조사는 정밀안전 진단에 필수적인 검사 항목이다. 그러나 육안으로 이루어지는 균열조사 방식은 현장 상황의 변화에 따라 주관적으로 수행될 수 있다. 이러한 육안검사의 한계를 극복하기 위해 본 연구에서는, ResNet, FPN, Mask R-CNN을 백본(Backbone), 넥(Neck), 헤드(head)로 구성된 합성곱 신경망을 바탕으로, 이미지 데이터에서의 콘크리트 균열 탐지를 자동화하고 그 성능을 IoU 값을 바탕으로 분석하였다. 해석에 사용된 데이터는 총 1,203개의 이미지 데이터로 구성하였으며, 이 중 70%를 훈련(Training)에, 20%를 검증(Validation)에, 그리고 10%의 데이터를 시험(Testing)에 사용하였다. 시험 결과의 평균 IoU값은 95.83%로 산정되었고, 또한 이미지 내 균열이 전혀 탐지되지 않는 경우는 존재하지 않아, 본 연구에 가정한 모델이 콘크리트의 균열 탐지를 성공적으로 수행하는 것을 확인하였다.

Keywords : Convolutional neural network, Crack detection, Deep learning model, Intersection over union, Mask R-CNN

1 정회원, 부산대학교 사회환경시스템공학과 석사과정 (Member, Graduate Student, Dept. of Civil and Environmental Engrg., Pusan National Univ.)
2 정회원, 조지아 공과대학 토목환경공학부 박사후연구원 (Member, Postdoctoral Fellow, School of Civil and Environmental Engrg., Georgia Institute of Technology)
3 비회원, 부산대학교 토목공학과 학사졸업 (Graduate Student, Dept. of Civil Engrg., Pusan National Univ.)
4 정회원, 부산대학교 사회환경시스템공학과 정교수 (Member, Prof., Dept. of Civil and Environmental Engrg., Pusan National Univ., Tel: +82-51-510-7627, Fax: +82-51-510-9596, jahn@pusan.ac.kr, Corresponding author, 교신저자)

* 본 논문에 대한 토의를 원하는 회원은 2024년 12월 31일까지 그 내용을 학회로 보내주시기 바랍니다. 저자의 검토 내용과 함께 논문집에 게재하여 드립니다.

1. 서론

국내 기반시설물 중 노후 시설의 비율은 빠른 속도로 증가하고 있다. 2020년 기준 국토교통부의 통계에 따르면, 도로, 터널, 교량과 같은 중대형 기반시설의 36.8%가 30년 이상의 노후 시설이었으며, 20년 후에는 중대형 기반시설물의 78.9%가 노후 시설물이 될 것으로 판단된다(NARS, 2015). 콘크리트는 노후화에 따라 시설물의 하중, 환경적인 요인 등에 의해 균열이 생기는 문제가 발생하게 된다. 구조물에 생긴 균열은 콘크리트의 수밀성과 기밀성을 저하시켜 콘크리트 내부의 철근의 부식, 시설물의 구조적 손상 등의 성능저하를 발생시키기 때문에, 콘크리트 균열을 탐지하여 유지보수 하는 것은 매우 중요한 사항이다.

콘크리트 구조물에 발생한 균열의 탐지는 구조물의 정밀안전 진단에 필수적인 항목이다. 현장에서의 균열 조사는 균열자, 균열 현미경 등의 균열 측정기를 활용하여 육안으로 관찰하여 찾아내는 방식으로 이루어진다(KALIS, 2022). 기존 방식의 균열조사는 기술자의 시간과 노력이 소요되며, 그 결과가 주관적 판단에 기인할 수 있다는 한계를 가지고 있다.

육안검사의 한계를 극복하기 위해, 합성곱 신경망(Convolutional Neural Network, CNN)을 활용하여 콘크리트 균열을 탐지하여 객관성과 신뢰성을 확보하기 위한 연구가 활발히 진행되고 있다. Dung(2019)은 이미지 세분화(Image segmentation) 작업에 널리 사용되는 FCN(Fully convolutional network)을 활용하여 도로 균열을 픽셀단위로 탐지하였으며, 균열 탐지의 정확도가 약 90%가 되는 것을 확인하였다. Golding et al.(2022)은, 원본 이미지에 회색조(Grayscale), 윤곽선 검출(Edge detection), 임계처리(Thresholding) 기법을 적용하여, CNN 알고리즘의 균열 탐지 성능을 향상시킬 수 있음을 보고하였다. 이외에도, CNN 알고리즘을 바탕으로 콘크리트의 균열 탐지를 자동화하기 위한 시도로서, 열화상 이미지와 환경변수를 이용한 콘크리트 균열 깊이 예측 머신 러닝 분석(Kim et al., 2021), 딥러닝과 영상처리기법을 이용한 콘크리트 지반 구조물 균열 탐지(Kim et al., 2018), CNN을 기반으로 한 1단계 객체탐지 알고리즘 YOLOv2를 활용한 콘크리트 균열 탐지(Jung et al., 2019) 등의 연구가 수행되었다. 한편, CNN 기반의 이미지 분석 알고리즘은 빠른 속도로 발전을 거듭하며 개발되고 있으며, 2단계 객체 탐지 알고리즘인 Mask R-CNN(He et al.,

2017)의 경우, 2018년 3월 20일에 발표된 Mask R-CNN 2.1(Abdulla, 2018)이 가장 최근에 개발되었고 타분야에 성공적으로 수행된 사례가 있다.

본 연구의 목적은, Mask R-CNN 2.1의 객체 탐지 성능을 향상시키기 위해 ResNet(He et al., 2016), FPN(Lin et al., 2017) 알고리즘과 결합하여 모델을 생성하고, 이미지 데이터를 바탕으로 한 균열탐지에 적용하여 그 성능을 분석하고 사용성을 평가하는 것에 있다.

2. 객체탐지 모델 구성

2.1 모델 구성

본 연구에 사용된 2단계 객체탐지 모델의 구조를 Fig. 1에 도식화하였다. 모델은 3가지 단계(Backbone, Neck, Head)를 통해 콘크리트 균열의 탐지를 수행한다. 첫번째 요소인 백본(Backbone)은 이미지를 입력으로 받아 여러가지 해상도의 고차원 특징맵(Feature map)들을 생성한다. 넥(Neck)은 백본에서 생성된 특징맵들을 통합하여 하나의 특징맵을 생성한다. 마지막으로 헤드(Head)는 생성된 특징맵(Feature map)을 기반으로 객체의 위치, 크기 등을 예측한다. 본 연구에서는 백본, 넥, 헤드의 요소로서 각각 ResNet, FPN(Feature pyramid network), 인스턴스 세분화(Instance segmentation)을 적용하여 Mask R-CNN 알고리즘을 구성하고 이를 적용하였다.

2.2 ResNet

백본(Backbone)으로 사용된 ResNet 알고리즘은, 단순히 신경망층을 쌓아 단계적으로 특징맵(Feature map)을 추출해내는 Simonyan and Zisserman(2014)가 발표한 VGGNet 알고리즘의 구조와는 달리, 입력층, 중간층의 출력을 직접 출력층으로 전달하는 단축연결(Shortcut connection) 기법을 활용하여 그라디언트 소실문제를 해결하고, 속도와 성능향상을 이루었다. 또한 ResNet 알고리즘은 VGGNet 알고리즘과 비교하였을 때 8배 깊은 신경망을 가지면서 모델의 복잡성은 높지 않아 정밀한 특징맵을 만들어낼 수 있다(He et al., 2017).

2.3 FPN (Feature Pyramid Network)

넥(Neck)으로 사용된 FPN(Feature pyramid network)

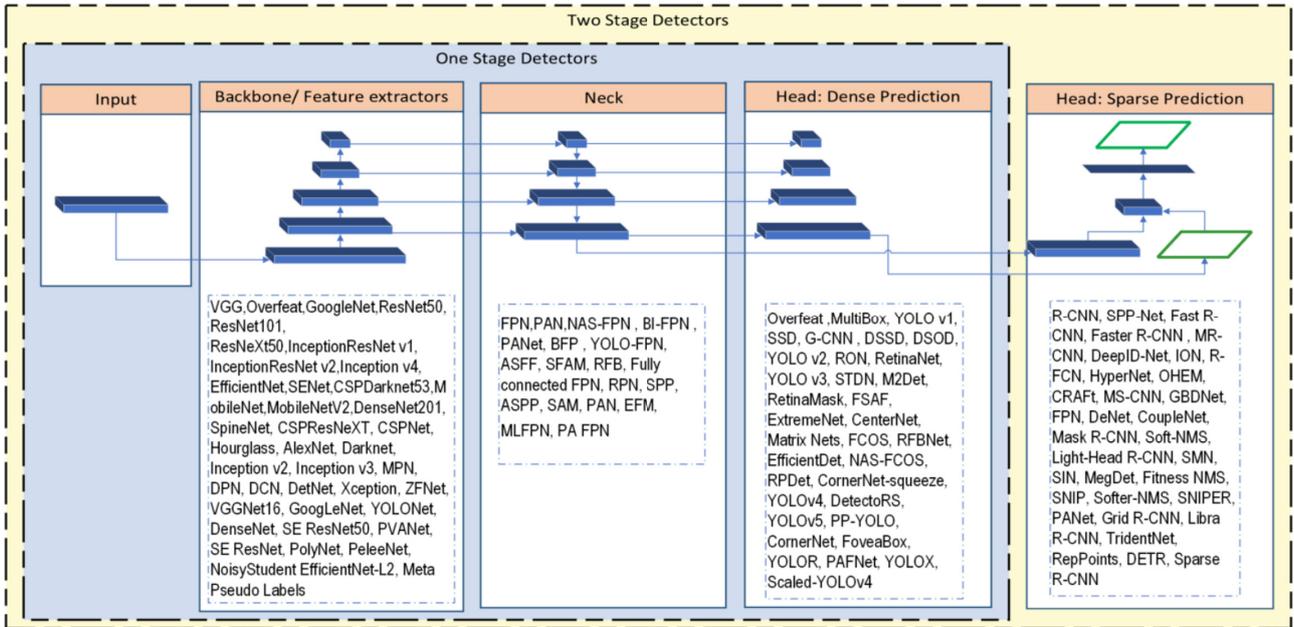


Fig. 1. Structure of the two-stage detection model (Bouraya and Belangour, 2021)

은 컨볼루션 레이어(Convolutional layer)를 피라미드 구조로 쌓은 형태로, ResNet에서 얻은 다양한 해상도의 특징맵(Feature map)들을 업샘플링(Up-sampling)을 통해 저해상도 특징맵의 특징을 고해상도 특징맵에 포함시켜 하나의 특징맵을 만든다. 이를 통해 객체의 세부적인 정보부터 경계조건과 같은 포괄적인 정보를 통합할 수 있으며, 객체의 형상이 다르거나 변형도 높은 정확도의 탐지가 가능하다(Lin et al., 2017).

2.4 Mask R-CNN

Mask R-CNN 알고리즘의 기반이 되는 CNN은 계층적 구조를 통해 이미지의 복잡한 패턴을 효과적으로 인식하고 처리하는 딥러닝 알고리즘으로, 이미지 처리 분야에서 광범위하게 활용된다. CNN은 Convolution 과정을 통해 이미지의 특징을 추출하고, Pooling 과정을 통해 레이어 사이즈를 줄여주는 Convolutional Layer와 Pooling Layer, 추출해낸 특징을 집약하고 목적에 맞게 출력력을 조정하는 Fully Connected Layer로 구성된다(LeCun et al., 1998; Park et al., 2023).

Mask R-CNN은 CNN을 기반으로 발전한 2단계 객체 인식 알고리즘으로, 기존의 Faster R-CNN 알고리즘의 단점을 보완하기 위해 만들어졌다. Faster R-CNN에서는 RoI(Region of interest) Pooling 기법을 사용할 때, 관심 영역을 고정된 크기의 그리드 형태로 변환하는 양자

화(Quantize) 과정에서 일부 위치 정보가 손실되는 문제가 발생하게 된다. 그러나 Mask R-CNN에서는 RoI Align 기법을 사용하여 양자화 과정 없이 관심 영역 좌표를 소수점 단위로 정확하게 계산하고, 해당 좌표의 특성 값을 선형 보간(Linear interpolation) 방식으로 추출하여 위치 정보가 손실되는 것을 해결하였다(He et al., 2017). 또한, Mask R-CNN은 인스턴스 세분화(Instance segmentation)를 통해, 탐지하고자 하는 객체를 픽셀 단위로 구분하고, 각 객체의 경계를 정확하게 파악하여 객체를 탐지의 정확도를 크게 향상시켰다.

본 연구에 사용된 Mask R-CNN의 버전은 2.1으로, 불필요한 드롭아웃 레이어(Dropout layer)의 제거, 객체 탐지 후보영역을 세밀화, 미니 배치(Mini batch)의 크기를 증가시키는 방식으로 이전버전의 Mask R-CNN 보다 높은 정확도를 보인다(Abdulla, 2018).

3. 해석방법

총 1,203장의 콘크리트 균열 이미지 데이터로 구성된 데이터 세트를 사용하였다. 데이터 세트는 훈련 세트(Train set), 검증 세트(Validation set), 시험 세트(Test set)로 분할하였으며, 각각의 비율은 70%, 20%, 10%이다.

모델 학습 시에는 과적합(Overfitting)을 방지하기 위해 Early stopping 기능을 사용하였다. Early stopping은 훈련 오차(Train loss)가 감소할 때 검증 오차(Validation



Fig. 2. Concrete crack images (Özgenel, 2018)

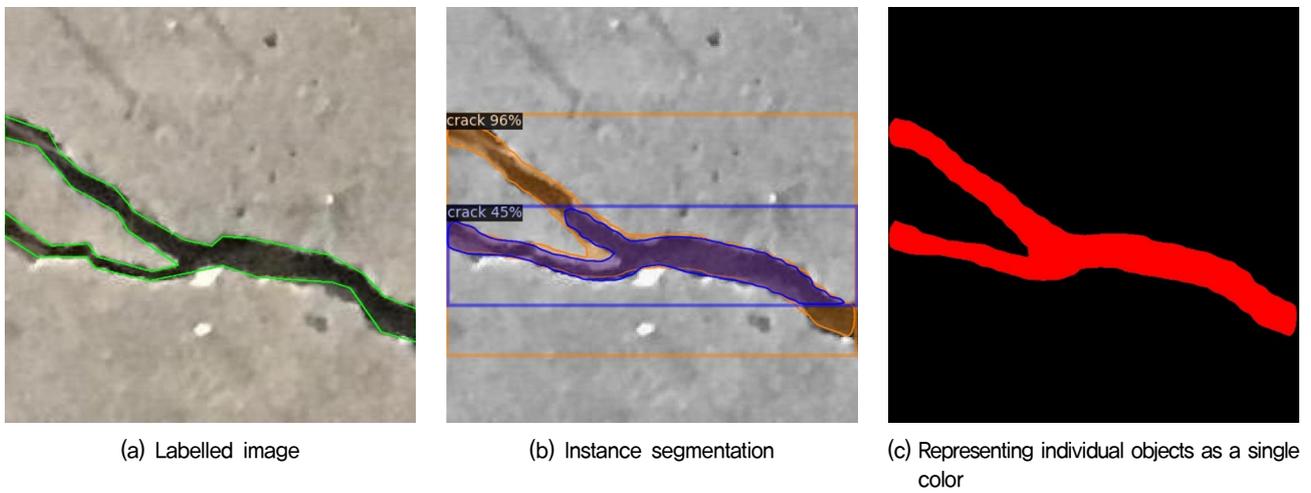


Fig. 3. Processing image data to calculate IoU

loss)가 증가하면 학습을 종료하는 기능이다. 이를 통해 모델이 지나치게 훈련되어 성능이 저하되는 것을 방지할 수 있다.

실제 콘크리트 균열이 포함된 영역(Ground truth)은 Fig. 3(a)의 콘크리트 균열을 따라 라벨링(Labeling)된 영역을 좌표로 계산한 넓이를 사용하였다. Mask R-CNN의 객체 탐지 결과는 인스턴스 세분화(Instance segmentation)에 의해 Fig. 3(b)와 같이 각각의 콘크리트 균열이 개별적인 객체로 인식되어 서로 다른 색으로 예측이 표시된다. 예측 영역(Predict union)의 넓이를 계산하기 위해서는 개별로 인식된 객체들의 RGB값을 Fig. 3(c)와 같이 [255, 0, 0]로 변환하여 통합된 영역의 면적을 계산하였다.

모델의 정량적 성능 평가는 IoU(Intersection over union)를 사용하였다(Cheng et al., 2021; Kim et al., 2018)(Fig. 4). IoU는 실제 콘크리트 균열이 포함된 영역(Ground truth)

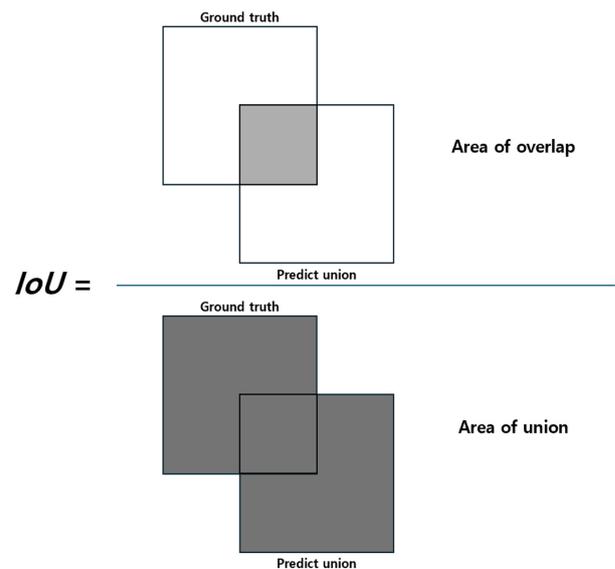


Fig. 4. Methodology for calculating the IoU

과 Mask R-CNN 알고리즘으로 예측해낸 콘크리트 균열 영역(Predict union)이 겹치는 영역 대비 전체 영역의 비율을 나타낸다.

4. 콘크리트 균열 이미지 분석 결과

Fig. 5는 본 연구에 사용된 모델의 콘크리트 균열 탐지 결과 예시이다. 모델은 전반적으로 Fig. 5(a)에 보인 것처럼 콘크리트 균열을 높은 정확도로 탐지하였다. 그러나 일부 시험 데이터에서는 Fig. 5(b)와 같이 균열 뿐만 아니라 균열이 아닌 영역도 균열로 판단하는 현상이

나타나기도 하였다.

모델의 예측성능을 정량적으로 평가하기 위해 사용된 시험 세트(Test set)에 포함된 122개의 콘크리트 균열 예측 이미지의 IoU를 측정하였다. 그 결과, 시험 세트(Test set)의 평균 IoU 값은 95.83%로 나타났다(Fig. 6). 122개의 시험 이미지 데이터 중 80%이상의 IoU값을 보인 예측 결과는 모두 118개로, 96.72%를 차지하였으며, IoU의 값이 80% 이하로 나온 예측 결과는 모두 4장으로 3.27%로 나타났다. IoU가 80% 이하인 경우의 예가 Fig. 5(b)에 제시되어 있다. 이미지 데이터 내의 균열을 전혀 탐지하지 못한 경우는 찾을 수 없었다.

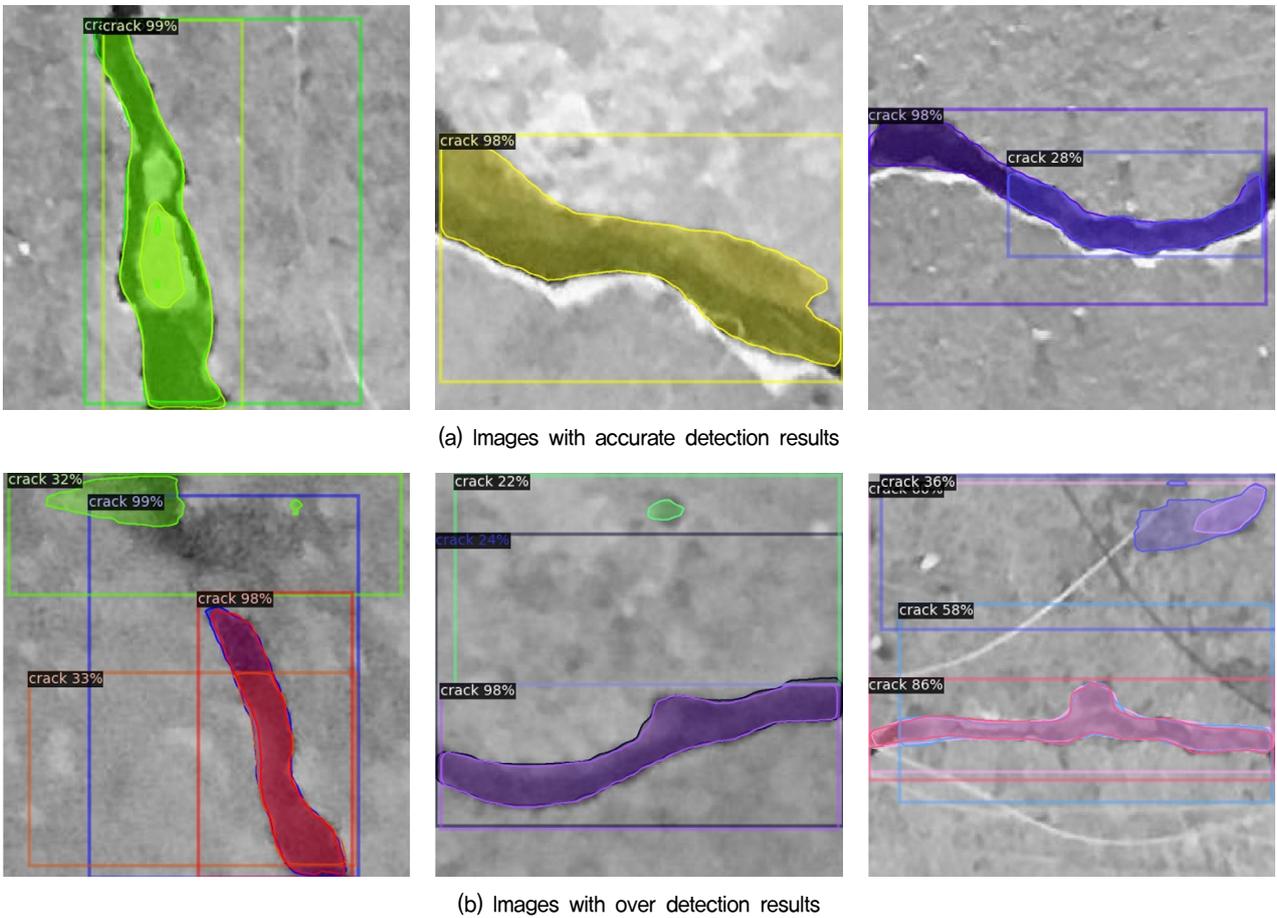


Fig. 5. Detection results from the concrete crack images

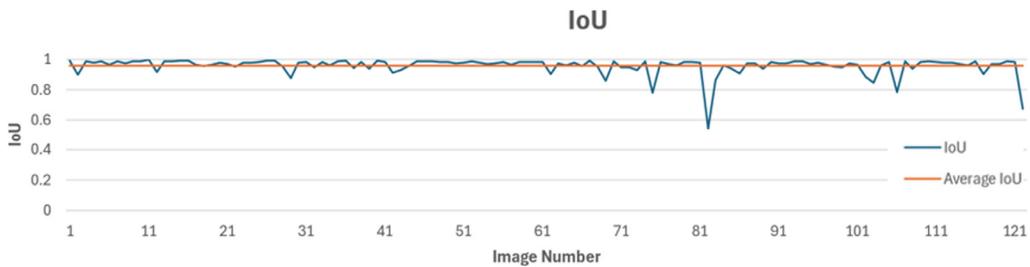


Fig. 6. IoU for test set images

객체 탐지 모델의 성능을 평가할 때, 찾고자 하는 객체를 올바르게 탐지하였다고 판단하는 기준이 되는 IoU의 임계값으로, 타분야에서는 50%를 적용하는 경우가 있다(Everingham et al., 2010). 그러나 본 연구와 같은 예측 모델을 활용한 균열 탐지에서, 탐지 누락은 구조물의 안정성 평가에 악영향을 미칠 수 있으므로, 높은 수준의 예측 일치성을 바탕으로 균열 등의 구조물 손상을 탐지할 수 있어야 할 것으로 판단한다.

5. 결론

본 연구에서는 콘크리트 균열 탐지를 자동화하기 위해, 콘크리트 균열 이미지 데이터를 사용하여 합성곱 신경망 기반 모델을 생성하였다. 해당 모델은 ResNet, FPN, Mask R-CNN을 백본(Backbone), 넥(Neck), 헤드(head)로 구성하였으며, 그 성능을 IoU 값을 바탕으로 분석하였다. 총 1,203개의 이미지 데이터 세트 중, 70%는 모델 훈련 세트(Train set), 20%는 검증 세트(Validation set), 10%는 시험 세트(Test set)로 사용하였다. 해석 결과에 따른 결론은 다음과 같다.

- (1) 실제 콘크리트 균열이 포함된 영역(Ground truth)과 Mask R-CNN 알고리즘으로 예측해낸 콘크리트 균열 영역(Predict union)이 겹치는 영역 대비 전체 영역의 비율을 측정하는 IoU값의 평균은 95.83%로 나타났다.
- (2) 시험 세트(Test set)에 포함된 122개의 콘크리트 균열 예측 이미지의 IoU에서 80%이상의 IoU값을 보인 예측 결과는 모두 118개로, 96.72%를 차지하였고, IoU의 값이 80% 이하로 나온 예측 결과는 모두 4장으로, 3.27%로 나타났다.
- (3) 이미지 데이터 내의 균열을 전혀 탐지하지 못한 경우는 찾을 수 없었으며, 또한 IoU 값이 80%이하로 예측되는 경우는 콘크리트 균열이 아닌 부분을 균열로 과도하게 탐지하기 때문인 것으로 분석되었다.

가정된 합성곱 신경망 모델을 활용한 콘크리트 균열 탐지는 전반적으로 우수한 성능을 보였다. 추후, 더 많은 수의 데이터 구축과 학습을 통해 탐지 정확도를 증가시켜, 해당 모델을 콘크리트 균열 이미지 분석의 자동화, 실용화에 활용할 수 있을 것으로 판단한다.

감사의 글

이 과제는 부산대학교 기본연구지원사업(2년)에 의하여 연구되었음.

참고문헌 (References)

1. Abdulla, W. (2018), Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow (v2.1) [Computer software]. GitHub. https://github.com/matterport/Mask_RCNN/releases
2. Bouraya, S. and Belangour, A. (2021), "Deep Learning based Neck Models for Object Detection: A Review and a Benchmarking Study", *International Journal of Advanced Computer Science and Applications*, Vol.12, No.11, pp.161-167.
3. Cheng, B., Girshick, R., Dollár, P., Berg, A. C., and Kirillov, A. (2021), "Boundary IoU: Improving Object-centric Image Segmentation Evaluation", In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp.15334-15342).
4. Dung, C. V. (2019), "Autonomous Concrete Crack Detection Using Deep Fully Convolutional Neural Network", *Automation in Construction*, Vol.99, pp.52-58.
5. Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010), "The Pascal Visual Object Classes (voc) Challenge", *International Journal of Computer Vision*, Vol.88, pp. 303-338.
6. Golding, V. P., Gharineiat, Z., Munawar, H. S., and Ullah, F. (2022), "Crack Detection in Concrete Structures Using Deep Learning", *Sustainability*, Vol.14, No.13, p.8117.
7. He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017), "Mask r-cnn", In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2961-2969).
8. He, K., Zhang, X., Ren, S., and Sun, J. (2016), "Deep Residual Learning for Image Recognition", In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp.770-778).
9. Jung, S. Y., Lee, S. K., Park, C. I., Cho, S. Y., and Yu, J. H. (2019), A method for detecting concrete cracks using deep-learning and image processing. *Journal of the Architectural Institute of Korea Structure & Construction*, Vol.35, No.11, pp.163-170.
10. KALIS (2022), Detailed Guidelines for Implementing Safety and Maintenance of Facilities (Safety Inspection and Diagnosis) Report, Korea.
11. Kim, A. R., Kim, D., Byun, Y. S., and Lee, S. W. (2018), "Crack Detection of Concrete Structure Using Deep Learning and Image Processing Method in Geotechnical Engineering", *Journal of the Korean Geotechnical Society*, Vol.34, No.12, pp.145-154.
12. Kim, A. R., Kim, D., Byun, Y. S., and Lee, S. W. (2018), "Crack Detection of Concrete Structure Using Deep Learning and Image Processing Method in Geotechnical Engineering", *Journal of the Korean Geotechnical Society*, Vol.34, No.12, pp.145-154.
13. Kim, J., Jang, A., Park, M. J., and Ju, Y. K. (2021), "Comparison Analysis of Machine Learning for Concrete Crack Depths Prediction Using Thermal Image and Environmental Parameters", *Journal of Korean Association for Spatial Structures*, Vol.21, No.2, pp.99-110.
14. LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998), "Gradient-

- based Learning Applied to Document Recognition”, *Proceedings of the IEEE*, Vol.86, No.11, pp.2278-2324.
15. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017), “Feature Pyramid Networks for Object Detection”, In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp.2117-2125).
16. NARS (2015), Basic Research for the Development of Comprehensive Fiscal Statistics System for Infrastructure Report, Korea.
17. Özgenel, Ç. F. (2018), “Concrete Crack Images for Classification”, Mendeley Data, v1 <http://dx.doi.org/10.17632/5y9wdsg2zt.1>
18. Park, J., Kim, K. Y., and Yun, T. S. (2023), “Morphological Analysis of Hydraulically Stimulated Fractures by Deep-Learning Segmentation Method”, *Journal of the Korean Geotechnical Society*, Vol.39, No.8, pp.17-28.
19. Simonyan, K. and Zisserman, A. (2014), Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556*.

Received : April 9th, 2024

Revised : May 27th, 2024

Accepted : May 28th, 2024