

# 모션블러 이미지에 대한 CNN 모델의 균열 검출 성능

이철희<sup>1</sup> · 김동구<sup>1</sup> · 김동규<sup>2\*</sup>

<sup>1</sup>정회원, 한국건설기술연구원 지반연구본부 수석연구원

<sup>2</sup>정회원, 한국건설기술연구원 지반연구본부 선임연구위원

## Crack detection performance of CNN models on motion blur images

Chulhee Lee<sup>1</sup> · Dongku Kim<sup>1</sup> · Donggyou Kim<sup>2\*</sup>

<sup>1</sup>Senior Researcher, Dept. of Geotechnical Engineering Research, Korea Institute of Civil Engineering and Building Technology

<sup>2</sup>Senior Research Fellow, Dept. of Geotechnical Engineering Research, Korea Institute of Civil Engineering and Building Technology

\*Corresponding Author : Donggyou Kim, dgkim2004@kict.re.kr

### Abstract

In this study, we analyzed the effect of motion blur on images used for detecting cracks in concrete tunnel linings on the performance of CNN models. Motion-blurred images with intensities ranging from 10 to 50 were generated on the Kaggle and KICT datasets. A semantic segmentation model with ResNet 18, ResNet 34, VGG 11, and Alex-Net as backbones for feature extraction was employed, all pre-trained on the U-Net architecture. The performance of these models in crack detection was then assessed. It was observed that detection accuracy decreased across all models as the intensity of motion blur increased for each dataset. Within the same model, the F1-score on the KICT dataset showed over 20% higher performance than on the Kaggle dataset. This study demonstrates that CNN-based crack detection performance is affected by the quality of the image data and that the crack detection accuracy of CNN models can vary depending on the quality of the dataset used in training.

**Keywords:** CNN (convolutional neural network), Motion blur, Dataset, Crack detection, F1 score

### 초 록

본 연구에서는 터널 콘크리트 라이닝에 발생하는 균열을 검출하는데 활용되는 이미지에 발생하는 모션블러가 CNN 모델들의 성능에 미치는 영향을 분석하였다. Kaggle과 KICT 데이터셋에 10~50 범위의 강도를 갖는 모션블러 이미지를 생성하고, U-Net 아키텍처에 사전 학습된 ResNet 18, ResNet 34, VGG 11 및 AlexNet를 특징추출을 위한

### OPEN ACCESS

Journal of Korean Tunnelling and  
Underground Space Association  
26(6)777-792(2024)  
<https://doi.org/10.9711/KTAJ.2024.26.6.777>

eISSN: 2287-4747

pISSN: 2233-8292

Received November 1, 2024

Revised November 20, 2024

Accepted November 20, 2024



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2024, Korean Tunnelling and Underground Space Association

백본으로 사용한 의미론적 분할 모델을 활용하여 균열 검출 성능을 분석하였다. 각각의 데이터셋에 모션블러가 증가함에 따라 모든 모델에서 검출 정확도가 감소하는 것으로 나타났다. 동일한 학습 모델에서 KICT 데이터셋의 F1-score가 Kaggle 데이터셋보다 20% 이상 높은 성능을 갖는 것으로 평가되었다. 본 연구의 결과로, CNN 기반 균열 검출 성능은 이미지 데이터의 품질에 따라 차이가 발생하며, 기 구축된 학습 데이터셋의 품질에 따라서도 CNN 모델의 균열 검출 성능이 상이할 수 있는 것으로 판단된다.

**주요어:** CNN (convolutional neural network), 모션블러, 데이터셋, 균열 검출, F1 score

## 1. 서론

터널의 콘크리트 라이닝에 발생하는 구조적 균열 정보는 터널 구조물의 안전성과 내구성을 평가하는 중요한 기초이므로 균열을 정확하게 검출하는 것은 유지관리 측면에서 큰 의미가 있다(Ni et al., 2019). 기존의 인력 기반 외관조사 결과의 낮은 신뢰도, 인건비 상승 그리고 위험한 조사 환경으로 인해 전통적인 수동 방법보다 안전하고 효율적인 비전(vision) 검사 방법이 요구되고 있다(Balaguer et al., 2014). 이러한 요구에 맞춰 건설 유지관리 분야에서는 향상된 점검 및 모니터링의 핵심 구성요소로 인식되고 있는 컴퓨터 비전(computer vision, CV) 기술이 활용되고 있다. 최근 몇 년 전부터는 카메라를 활용한 이동식 터널 스캐닝 시스템으로 취득한 영상 이미지로부터 기계학습(machine-learning, ML)의 한 분야인 딥러닝(deep-learning, DL)의 합성곱 신경망(convolutional neural network, CNN)을 활용하여 균열과 같은 손상을 검출하기 위한 다양한 연구가 수행되고 있다(Alidoost et al., 2022).

균열 검출을 위한 이미지 기반 CV는 분할(segmentation)방식, 이미지 분류(classification), 객체 탐지(object detection)로 구분된다. 이미지 수준, 객체 수준 그리고 픽셀 수준에서 시각적 정보를 제공한다. 분할 방식은 의미론적(semantic) 분할과 인스턴스(instance) 분할이라는 두 가지 하위작업으로 세분화된다(Guo et al., 2024). 이미지 분류 방식은 이미지 패치(patch)를 생성하고 각 패치에 균열이 있는지 여부를 진단하여 이미지 영역에서 균열의 위치를 파악하는 데 중점을 두고 있다. 콘크리트 표면 손상 검출을 위해 가장 광범위하게 사용되는 이미지 분류의 CNN 아키텍처(architecture)는 AlexNet (Krizhevsky et al., 2012), visual geometry group (VGG) networks (Simonyan and Zisserman, 2015), Inception networks (Szegedy et al., 2016), 및 ResNet (He et al., 2016) 등이 있다. 균열 검출을 위한 분류 네트워크의 성능을 비교한 연구에 따르면, 손상 유형의 다양성과 학습데이터의 품질로 인해 단일 네트워크가 지속적으로 다른 네트워크보다 우수한 성능을 발휘하지 못하는 것으로 파악되었다(Guo et al., 2020; Ali et al., 2021; Miao and Srimahachota, 2021; Shin et al., 2021). 이러한 연구 결과는 학습데이터의 품질 차이에 따라 균열 검출을 위한 최적의 네트워크를 선별하는 것에 한계가 있음을 보여주고 있다.

객체 탐지 방식은 CNN 기반 탐지기를 사용하여 균열과 다른 객체를 구별하는 것이 주된 목표이다. 객체 탐지 모델에는 두 가지 유형이 있으며, Faster Region-based Convolutional Neural Network (Faster RCNN)과 같은

지역 기반 탐지기(region-based detector)는 영역 제안(region proposal), 특징 추출기(feature extractor) 및 분류기(classifier) 등으로 사용된다. Single shot multibox detector (SSD)(Ren et al., 2017)와 you only look once (YOLO)(Redmon et al., 2016)와 같은 단일 단계 탐지기(one-stage detector)는 모든 경계 상자를 한 번에 예측하고 처리속도가 빨라서 모바일 장치에 적합한 경향을 띠고 있다. Kumar et al. (2020)은 Faster R-CNN, SSD, YOLO의 성능을 비교하였다. 그 결과 Faster R-CNN이 SSD와 YOLO보다 훨씬 더 정확하지만 처리 시간이 더 오래 걸리는 것으로 분석되었다. Xue and Li (2018)는 철도 터널 스캐닝 장비인 moving tunnel inspection (MTI-100)로 취득한 이미지에서 균열을 검출하기 위한 Faster RCNN을 제안하고 설계하였다. GoogLeNet (Szegedy et al., 2015), AlexNet 및 VGG networks과 비교하여 보다 개선된 정확도를 보여주었다. Li et al. (2021)은 높은 정밀도로 터널 표면 결함을 자동으로 탐지하기 위하여 Faster RCNN을 활용한 Metro Tunnel Surface Inspection System (MTSIS)를 개발하였다. 그러나 손상 검출의 경우 균열, 박락, 누수 위치 파악 및 분류에 대한 성능은 향상되었지만 상태평가가 가능한 수준에는 미흡하다고 평가하였다. 정량적으로 손상을 평가하기 위해서는 데이터 셋 수집 문제와 고속으로 수집된 이미지의 품질 문제는 향후 해결해야 할 문제로 언급하였다.

객체 탐지는 결함을 찾을 수 있지만 결함의 모양과 윤곽을 정확하게 그릴 수 없는 단점이 있다(Spencer Jr et al., 2019). 반면 의미론적 분할은 각 픽셀에 클래스를 레이블(label)하여 픽셀 수준의 결함 위치를 파악할 수 있기 때문에, 최근 연구에서는 표면 손상 탐지를 위하여 의미론적 분할을 적용하는 추세이다. 콘크리트 구조물 표면에서 균열을 픽셀 수준으로 식별하기 위해 fully convolutional network (FCN)(Long et al., 2015), DeepLab (Chen et al., 2018), U-Net (Liu et al., 2019), DenseNet (Mei et al., 2020), SegNet (Badrinarayanan et al., 2017) 등 의미론적 분할 네트워크가 제안되었다. 또한, 표면 손상 탐지의 새로운 방법인 인스턴스 분할은 개별 손상에 대한 자세한 정보를 제공하며 Mask R-CNN (Wei et al., 2019)과 segment objects by locations (SOLO)(Wang et al., 2020)가 있다.

다양한 CNN기반 알고리즘을 통해 균열을 성공적으로 검출할 수 있었지만 여전히 해결해야 할 몇 가지 중요한 기술적 문제가 있다(Bae et al., 2021). 실제로 CNN기반 균열 손상 검출 아키텍처의 성능은 다양한 조건에서 수집된 균열 이미지의 품질에 크게 좌우된다(Liu et al., 2020). 이동식 터널 스캐닝 시스템은 터널의 콘크리트 라이닝 표면을 이미지로 안전하게 수집할 수 있는 가장 효율적인 방법이다. 그러나 차량의 진동과 콘크리트 라이닝 표면에 근접 촬영이 어려워 수집된 이미지에서 모션블러(motion blur, MB)와 해상도 부족이 발생할 수 있다. 여기서 MB는 흔들리거나 이동하는 물체를 촬영할 때 발생하는 흐릿한 이미지 효과를 의미한다. 또한 디지털 영상은 획득, 보관, 가공, 전송 과정에서 품질 저하가 발생할 수 있다. 대용량의 학습데이터는 다수의 인력에 의해 이미지의 수집 및 가공이 이루어지기 때문에 작업자의 숙련도에 따라 다른 품질의 이미지가 데이터셋에 포함될 수 있다. 이로 인해 낮은 품질의 이미지가 데이터셋에 포함되고, 이미지 정보가 손실됨으로써 유지관리 실무에서 요구하는 0.3 mm 균열 폭과 같은 미세 균열을 검출하지 못하는 경우가 많다. 이러한 문제점들에도 불구하고 이미지 데이터 셋의 품질에 대한 체계적인 연구는 많이 부족한 상황이다(Lee et al., 2024).

본 연구에서는 터널의 콘크리트 라이닝에 발생하는 균열을 검출하기 위해서 사용되는 이미지 데이터의 품질이

CNN 모델들의 성능에 미치는 영향을 분석하였다. 이것은 기존 연구에서 서로 다른 촬영조건 및 해상도를 갖는 이미지 데이터로 구축된 학습데이터로 인해 CNN 모델들의 성능 비교가 어려운 점에 대해서 착안하였다. 사용된 데이터는 공개 데이터인 Kaggle 이미지 데이터셋(Kaggle, 2024)과 실제 이동식 터널 스캐닝 시스템으로 80 km/h의 속도로 운영 중인 터널의 콘크리트 라이닝 표면을 촬영한 이미지 데이터셋(이하 KICT 데이터셋)을 활용하였다. 이미지의 선명도에 가장 큰 영향을 미치는 MB에 대하여 모션블러 필터를 사용하여 수평 강도(intensity) 10~50까지 갖는 이미지들을 생성하였다. 학습모델로는 전이학습(transfer learning, TL)이 수행된 U-Net을 사용하였고, 특성추출을 위해서는 사전 학습된 AlexNet, VGG, ResNet 모델을 활용하였다. 생성된 MB 이미지를 활용하여 CNN기반 균열 검출 성능의 강건성(robustness)을 분석하였다.

## 2. 모션블러 이미지 데이터셋 구축

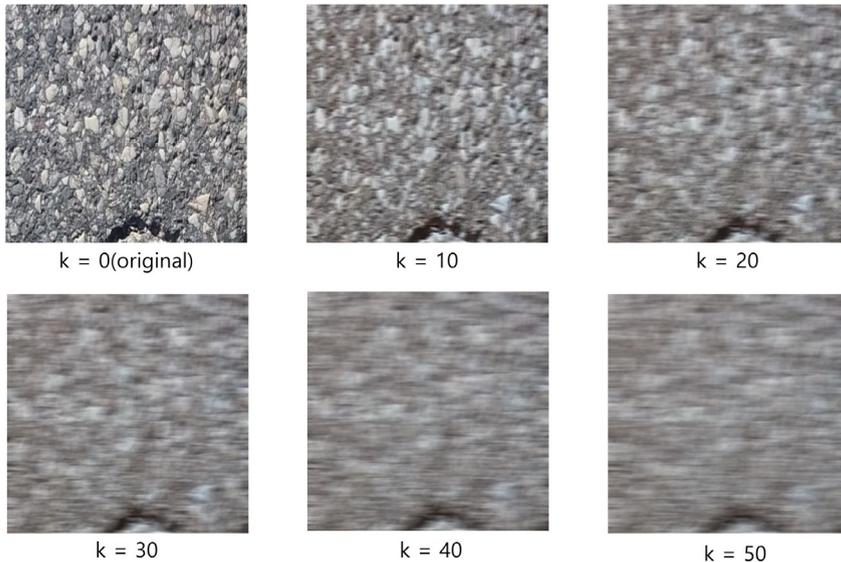
고속으로 이동하는 이동식 터널 스캐닝 시스템으로 터널의 콘크리트 라이닝 표면을 촬영할 경우, 도로 및 차량 진동에 의한 카메라의 떨림으로 이미지에 MB가 발생한다. 또한, 터널의 각기 다른 구조적, 환경적 특성으로 취득되는 이미지의 품질은 일관성을 유지하기 어렵다. 그러므로 CNN기반 균열 검출 측면에서 다양한 MB에 대한 강건성을 분석하는 것이 매우 필요하다. 이에 본 연구에서는 공개된 Kaggle 데이터셋과 KICT 데이터셋을 활용하여 MB 강도를 10~50까지 발생시킨 이미지를 생성하였다. 여기서 강도는 중심이 되는 픽셀 좌우로 블러 현상을 발생시키기 위해서 사용되는 주변 픽셀 수를 의미한다. 이미지에 MB 효과를 부여하기 위한 알고리즘(식 (1))과 생성 방법은 다음과 같다.

$$k = \frac{1}{3} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad (1)$$

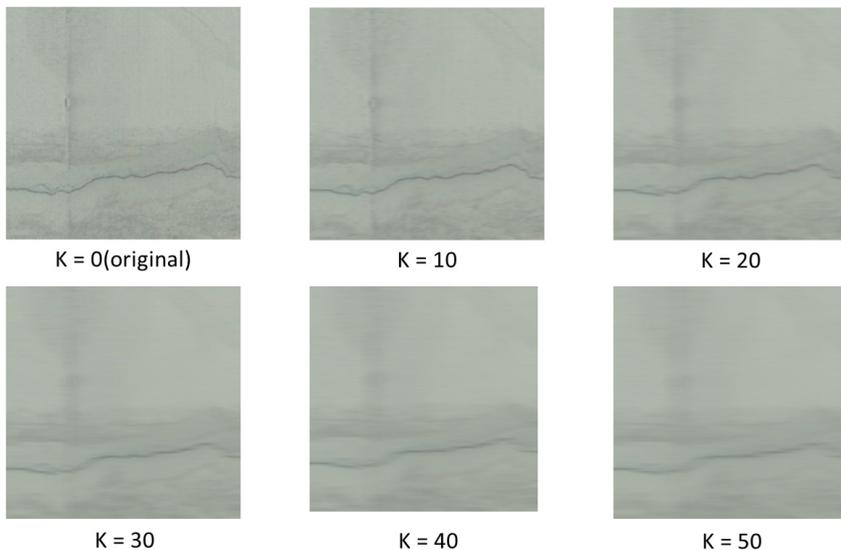
1.  $k \times k$ 의 정방향 0 행렬을 커널(kernel)로 생성한다.
2. 커널 행렬의 중앙행에 1을 부여한다(짝수인 경우  $k/2$ , 홀수인 경우  $k/(k+1)$ )
3. 커널 행렬의 전체 값을  $k$ 로 나누어 평균화한다.
4. 입력 이미지에 대해서  $k \times k$  크기의 커널을 합성곱으로 연산을 수행한다.
5. 중앙행에  $1/\text{size}$ 의 값이 분포되고 합성곱을 통해 수평방향으로 좌우 픽셀의 평균치를 도출하여 수평 블러 효과를 생성한다.

Kaggle 데이터셋은 11,298개의 이미지( $448 \times 448$ )이며, KICT 데이터셋은 27,736개의 이미지( $512 \times 512$ )로 구성되어 있다. 각각의 데이터셋에 커널 사이즈  $k$ 를 10, 20, 30, 40, 50으로 변화시켜 5종류의 MB 이미지를 Fig. 1과 같이 생성하였다. DL 학습시에 Kaggle 데이터셋은 학습 이미지 9,603개와 실험 이미지 1,695개로 분류하여

사용하였다. KICT 데이터셋은 전체 이미지에서 균열이 포함된 이미지의 비율이 5% 이하이기 때문에 무작위로 실험 데이터를 추출할 경우, 균열이 없는 다수의 이미지가 포함될 가능성이 높다. 그래서 mask 파일에서 균열이 포함된 픽셀수로 정렬한 후, 상위 1/3에서 100개 이미지, 중위 1/3에서 100개 이미지, 하위 1/3에서 100개 이미지를 무작위로 추출하여, 실험 데이터에서 균열의 분포가 다양하게 포함되도록 하였다. 이렇게 2가지 데이터셋에 대해서 원본 이미지와 5종류의 MB 데이터셋을 포함해서 총 12개의 데이터셋을 구축하였다.



(a) Kaggle dataset with generated 5 intensities of motion blur



(b) KICT dataset with generated 5 intensities of motion blur

Fig. 1. Generated motion blur image dataset

### 3. 전이학습 기반 CNN 훈련

CNN의 일반화 성능(generalisation capability)은 레이블이 지정된 대량의 학습 데이터에 크게 의존한다는 것은 잘 알려진 사실이다. 그러므로 학습 데이터 수가 줄어들면 분류 정확도(classification accuracy) 또한 떨어지며, 이로 인해 과적합(over-fitting) 문제가 발생할 수 있다. CNN의 뛰어난 이미지 인식 성능은 여전히 대규모 데이터셋의 학습을 전제하고 있다. 그러나 유지관리 분야에서 충분하고 효과적인 데이터를 수집하는 것은 여전히 어려운 일이다. 이러한 점에서 TL 기술은 학습데이터의 양에 대한 네트워크의 의존도를 줄일 수 있기 때문에, 새로 취득한 데이터나 기존의 데이터를 최대한 활용할 수 있는 강력한 도구로 인식되고 있다(Yu et al., 2022). 기존 ML 방법과 달리 TL은 데이터셋을 소스 도메인과 타겟 도메인의 두 그룹으로 구분한다. TL은 소스 도메인에서 학습한 지식을 관련 타겟 도메인에 적용하여 타겟 문제를 해결한다(Pan and Yang, 2010).

학습 모델은 의미론적 분할(semantic segmentation)의 U-Net 아키텍처를 활용하여 구축하였다. 특징 추출(feature extraction)을 담당하는 백본(backbone)으로 사전 학습된 ResNet 18, ResNet 34, VGG 11 및 AlexNet으로 총 4개의 네트워크를 사용하였다. 여기서 모델 학습에는 U-Net의 일종인 Dynamic UNet (Fig. 2)을 사용하였다(Dynamic UNet, 2024). Dynamic UNet은 다양한 인코더(encoder)를 사용하여 특징맵(feature map)을 추출하고 이와 대칭적인 디코더(decoder)를 구성하여 원본 이미지와 동일한 크기의 출력을 생성한다. 레이어의 입출력 U-Net의 인코더는 보통 사전 학습된 상태로 사용되며, ImageNet과 같은 대규모 데이터셋에서 학습된 가중치를 활용한다. 이를 통해 모델의 학습을 가속화하고, 일반화 성능을 향상시킬 수 있다.

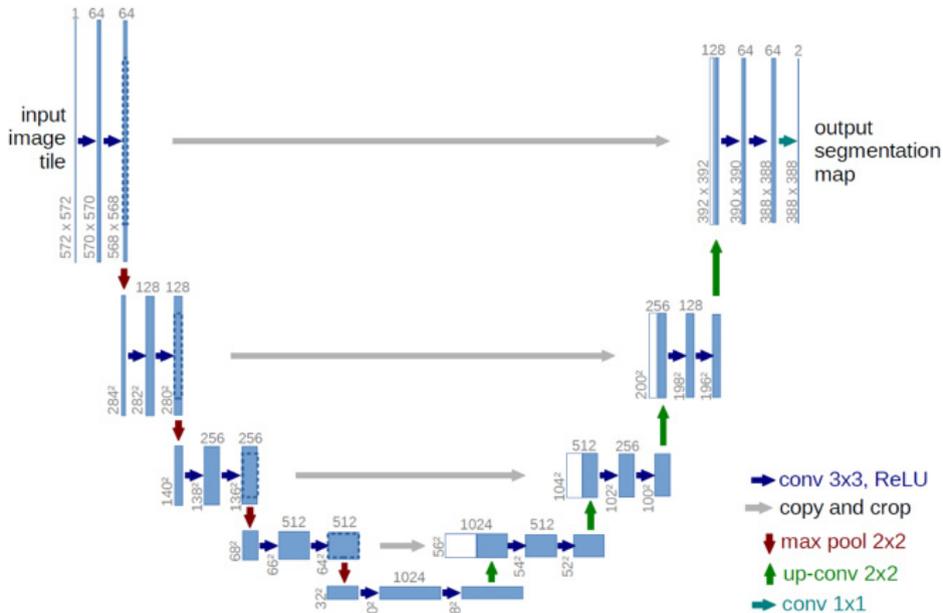


Fig. 2. Schematic diagram of Dynamic UNet by fastai

본 연구에서는 TL기반의 CNN모델을 활용한 균열 검출 성능 분석을 위해 파이썬(python)의 DL 라이브러리인 fastai를 활용하였다. Fastai는 pytorch를 기반으로 TL이 적용된 CNN 모델을 구축할 수 있도록 지원하고 있다 (Fastai, 2024). TL기반의 CNN 모델 구축시 Kaggle과 KICT 데이터셋은  $256 \times 256$  사이즈로 이미지 크기를 동일하게 조정하였다. Fig. 3과 같이 구축한 12종의 MB 데이터셋에서 분류한 학습데이터로 학습모델 Dynamic UNet으로 구축하고, 학습데이터에 사전 학습된 ResNet 18, ResNet 34, VGG 11 그리고 AlexNet을 특징추출용 백본으로 사용하여 균열 검출 결과를 도출하였다.

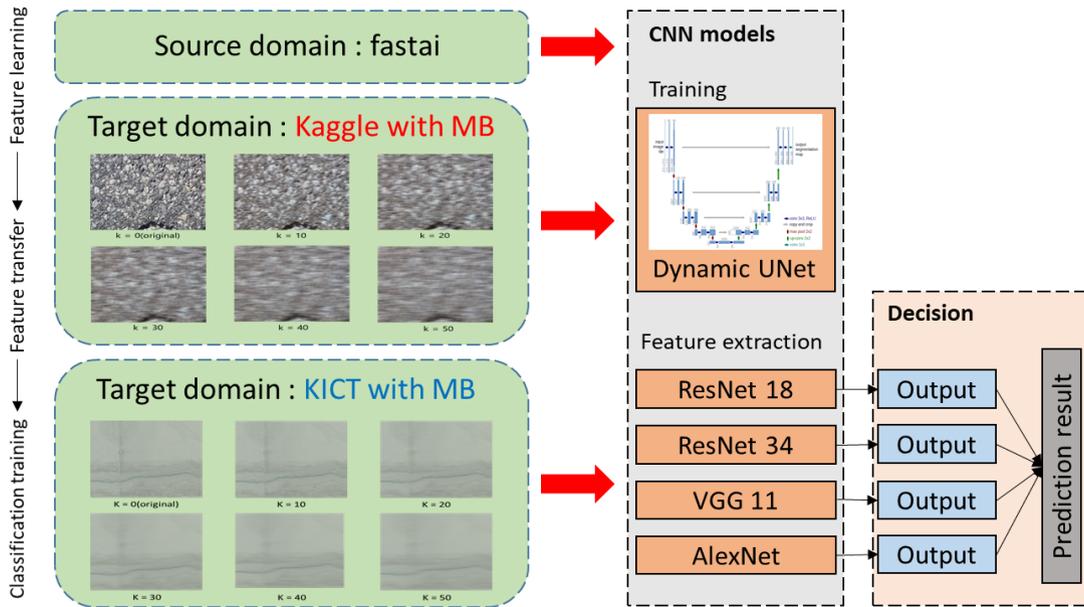


Fig. 3. Schematic of crack diagnosis about motion blur images using CNN

## 4. CNN기반 균열 검출 성능 분석

### 4.1 CNN 모델의 성능 평가 지표

각기 다른 CNN 모델들의 균열 검출 성능의 강건성을 분석하기 위해 재현율(recall), 선택성(selectivity), 정밀도(precision), 음성 예측도(negative predictive value, NPV), 정확도(accuracy)와 F1-score 등을 포함한 통계적 지표를 사용하였다. 재현율은 실제 긍정 케이스 중 모델이 올바르게 예측한 비율이다. 정확도는 전체 케이스 중에서 올바르게 예측된 케이스의 비율이며, 정밀도는 긍정으로 예측된 케이스 중 실제로 긍정인 케이스의 비율을 나타낸다. 선택성은 TN (true negative)을 식별하는 모델의 성능을 측정하는 비율이며, 음성 예측도는 모델이 부정으로 예측한 케이스 중 실제로 부정인 비율이다. F1-score는 재현율과 정밀도의 조화평균으로 두 지표의 균형을 고려한 평가지표이다. 평가 지표의 값이 높을수록 평가된 모델의 성능이 더 좋다는 것을 의미한다. 이러한 지표의

수학적 표현은 식 (2)~(7)과 같다. 이러한 지표들은 Table 1과 같이 confusion matrix에 기반하며, 분류 문제에서 CNN 모델의 성능을 평가하는데 활용된다.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Selectivity = \frac{TN}{TN + FP} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$NPV = \frac{TN}{TN + FN} \quad (5)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (7)$$

여기서, TP (true positive)은 모델이 긍정으로 올바르게 예측한 경우의 수, FP (false positive)는 모델이 긍정으로 잘못 예측한 경우의 수, TN은 모델이 부정으로 올바르게 예측한 경우의 수 그리고 FN (false negative)는 모델이 부정으로 잘못 예측한 경우의 수를 의미한다.

Table 1. Confusion matrix

TP	FN	Recall
FP	TN	Selectivity
Precision	NPV	Accuracy

## 4.2 Kaggle 데이터셋의 모션블러 영향 분석

Kaggle 데이터셋에 MB 강도 10~50을 부여하여 생성한 이미지를 활용하여 CNN 4개 모델에 대한 균열 검출 성능을 confusion matrix에 기반하여 분석하였다. Table 2는 평가된 재현율(recall), 선택성(selectivity), 정밀도(precision), 음성 예측도(negative predictive value, NPV), 정확도(accuracy)와 F1-score를 보여주고 있다.

**Table 2.** Evaluation indices of different models of backbone for crack diagnosis from Kaggle dataset with MB

Model - blur	Evaluation index (%)					
	Recall	Selectivity	Precision	NPV	Accuracy	F1 score
ResNet 18 - 0	64.96	99.38	79.28	98.73	98.17	71.41
ResNet 18 - 10	53.33	99.45	77.90	98.32	97.83	63.31
ResNet 18 - 20	35.03	99.68	79.77	97.68	97.41	48.68
ResNet 18 - 30	13.75	99.87	79.43	96.96	96.85	23.44
ResNet 18 - 40	3.40	99.98	83.55	96.61	96.59	6.53
ResNet 18 - 50	2.97	99.98	82.34	96.59	96.57	5.73
ResNet 34 - 0	64.32	99.41	79.86	98.71	98.18	71.25
ResNet 34 - 10	50.59	99.56	80.84	98.23	97.85	62.23
ResNet 34 - 20	35.35	99.70	81.30	97.70	97.45	49.27
ResNet 34 - 30	24.38	99.79	81.20	97.32	97.15	37.51
ResNet 34 - 40	16.67	99.87	82.88	97.06	96.96	27.75
ResNet 34 - 50	7.23	99.95	84.36	96.74	96.70	13.31
VGG 11 - 0	35.53	99.62	77.41	97.70	97.37	48.70
VGG 11 - 10	17.26	99.80	76.14	97.07	96.91	28.14
VGG 11 - 20	7.56	99.86	66.07	96.74	96.62	13.57
VGG 11 - 30	2.38	99.90	47.26	96.57	96.48	4.53
VGG 11 - 40	6.45	99.55	34.05	96.70	96.28	10.84
VGG 11 - 50	2.51	99.95	64.21	96.57	96.53	4.82
AlexNet - 0	37.93	99.54	74.85	97.78	97.38	50.34
AlexNet - 10	21.08	99.80	79.65	97.21	97.04	33.34
AlexNet - 20	12.83	99.89	80.65	96.92	96.83	22.13
AlexNet - 30	8.60	99.88	72.91	96.78	96.68	15.38
AlexNet - 40	0.70	99.99	83.97	96.51	96.51	1.38
AlexNet - 50	2.79	99.96	73.63	96.58	96.55	5.37

본 연구에서 구축한 MB 이미지 데이터셋은 불균형 데이터에 대한 CNN 모델에 대한 강건성을 분석하기 위한 것이다. 이러한 불균형 데이터셋에서 정밀도와 재현율은 일반적으로 사용되는 평가지표이다. 이상적으로 정밀도와 재현율이 둘다 높으면 좋지만 실제로 두 평가지표를 모두 높이는 것은 한계가 있다. 그러므로 정밀도와 재현율을 조화평균한 F1-score로 MB 이미지에 대한 CNN 모델의 강건성을 분석하였다.

Fig. 4는 Kaggle 데이터셋에 대한 4개의 CNN 모델의 균열 검출 성능을 F1-score로 나타낸 것이다. 원본 이미지에서 ResNet 18은 71.41%, ResNet 34는 71.25%, VGG 11는 48.7%, AlexNet은 50.34%로 나타났다. 4개 모델 중 ResNet 18이 가장 높은 균열 검출 성능을 갖는 것으로 나타났다. MB 강도가 10에서 50으로 증가할수록 4개 모델에서 전부 F1-score가 감소하여 균열 검출 성능이 저하되는 것으로 나타났다. 여기서 ResNet 34의 경우

MB 강도가 10일 때 62.23%, 20일 때 49.27%, 30일 때 37.51%, 40일 때 27.75% 그리고 50일 때 13.31%로 비교적 선형적 감소 경향을 보이고 있다. 다른 3개 모델에서는 MB 강도 30~50에서 F1-score가 불규칙한 결과로 나타났다. ResNet 34 모델이 MB의 변화에도 균열 검출에 있어서 비교적 높은 성능을 발휘하는 것으로 볼 수 있다.

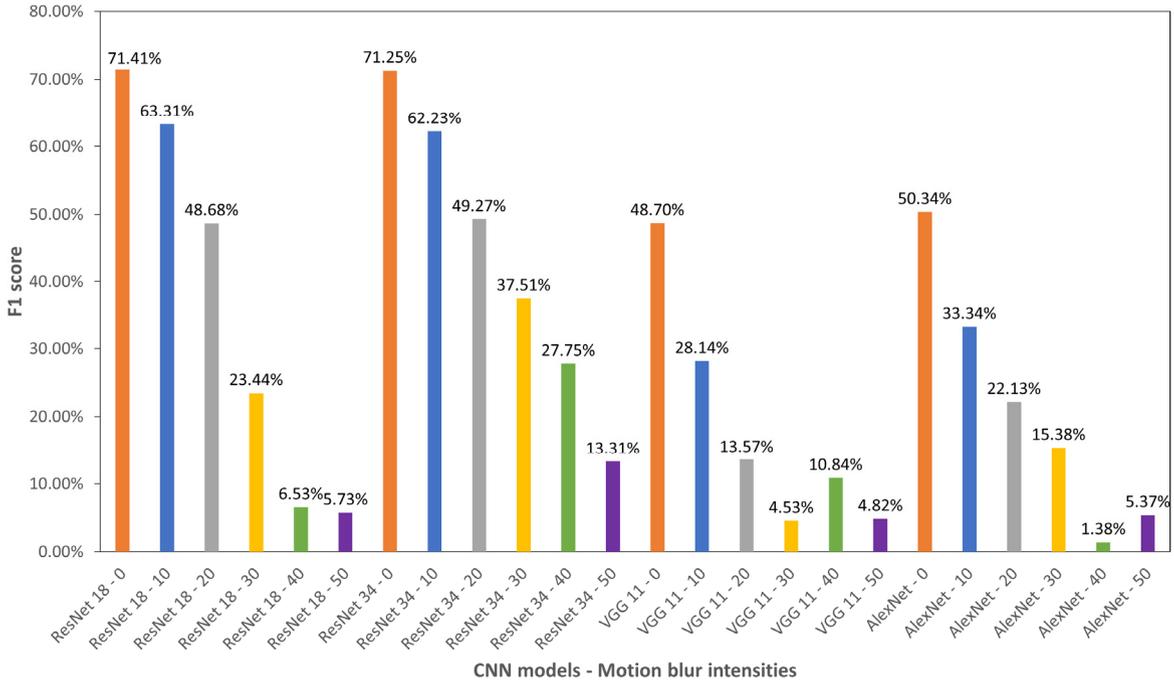


Fig. 4. Results of crack diagnosis by 4 CNN models of backbone about 5 motion blur intensities of Kaggle dataset

#### 4.2 KICT 데이터셋의 모션블러 영향 분석

KICT 데이터셋에 MB 강도 10~50으로 생성된 이미지를 활용하여 CNN 4개 모델에 대한 균열 검출 성능을 분석하였다. Table 3은 평가된 재현율(recall), 선택성(selectivity), 정밀도(precision), 음성 예측도(negative predictive value, NPV), 정확도(accuracy)와 F1-score를 보여주고 있다.

**Table 3.** Evaluation indices of different CNN models of backbone for crack diagnosis from KICT dataset with MB

Model - blur	Evaluation index (%)					
	Recall	Selectivity	Precision	NPV	Accuracy	F1 score
ResNet 18 - 0	92.46	99.19	85.74	99.60	98.85	88.97
ResNet 18 - 10	54.14	99.43	83.32	97.62	97.16	65.63
ResNet 18 - 20	29.59	99.66	82.05	96.40	96.14	43.49
ResNet 18 - 30	11.89	99.85	80.71	95.55	95.44	20.72
ResNet 18 - 40	5.50	99.91	76.66	95.24	95.17	10.27
ResNet 18 - 50	1.12	99.99	82.25	95.04	95.03	2.21
ResNet 34 - 0	93.41	99.18	85.77	99.65	98.89	89.43
ResNet 34 - 10	57.67	99.35	82.33	97.80	97.26	67.83
ResNet 34 - 20	24.31	99.74	82.93	96.15	95.95	37.59
ResNet 34 - 30	12.39	99.85	81.62	95.57	95.46	21.52
ResNet 34 - 40	5.39	99.92	77.93	95.24	95.18	10.08
ResNet 34 - 50	2.29	99.96	76.44	95.09	95.06	4.45
VGG 11 - 0	71.36	98.61	73.10	98.49	97.25	72.22
VGG 11 - 10	31.12	99.42	73.80	96.47	95.99	43.78
VGG 11 - 20	12.47	99.64	64.54	95.57	95.26	20.90
VGG 11 - 30	6.70	99.82	66.39	95.30	95.15	12.17
VGG 11 - 40	3.74	99.87	60.94	95.16	95.05	7.05
VGG 11 - 50	0.47	99.99	68.51	95.00	95.00	0.93
AlexNet - 0	47.05	99.10	73.37	97.25	96.49	57.33
AlexNet - 10	25.27	99.37	67.86	96.18	95.65	36.82
AlexNet - 20	8.40	99.87	77.16	95.38	95.28	15.15
AlexNet - 30	2.83	99.96	77.91	95.12	95.08	5.46
AlexNet - 40	1.96	99.94	63.49	95.07	95.02	3.80
AlexNet - 50	0.88	99.98	71.24	95.02	95.01	1.74

Fig. 5는 KICT 데이터셋을 활용한 4개의 CNN 모델의 균열 검출 성능을 F1-score로 나타낸 것이다. 원본 이미지에서 ResNet 34은 89.43%로 ResNet 18의 88.97%, VGG 11의 72.22%와 AlexNet의 57.33% 보다 높은 균열 검출 성능을 갖는 것으로 나타났다. Kaggle 데이터셋의 결과와 동일하게 MB 강도가 증가할수록 F1-score가 감소하는 경향을 나타내고 있다. 원본 이미지에서 MB 강도가 10정도 발생하면 모든 CNN 모델에서 약 20% 이상의 균열 검출 성능이 감소하는 결과를 보이고 있다. Kaggle 데이터셋인 Fig. 1(a)에서 아스팔트 표면과 같이 이미지에서 특징점이 많은 경우에는 MB 강도 10은 육안으로 미세한 차이가 인식되지만, KICT 데이터셋인 Fig. 1(b)와 같이 이미지에서 특징점이 잘 나타나지 않을 경우에는 육안으로도 식별이 어려울 수 있다. 이러한 시각적 오차는 CNN 기반 균열 검출의 성능 저하에 영향을 미치는 요인으로 볼 수 있다.

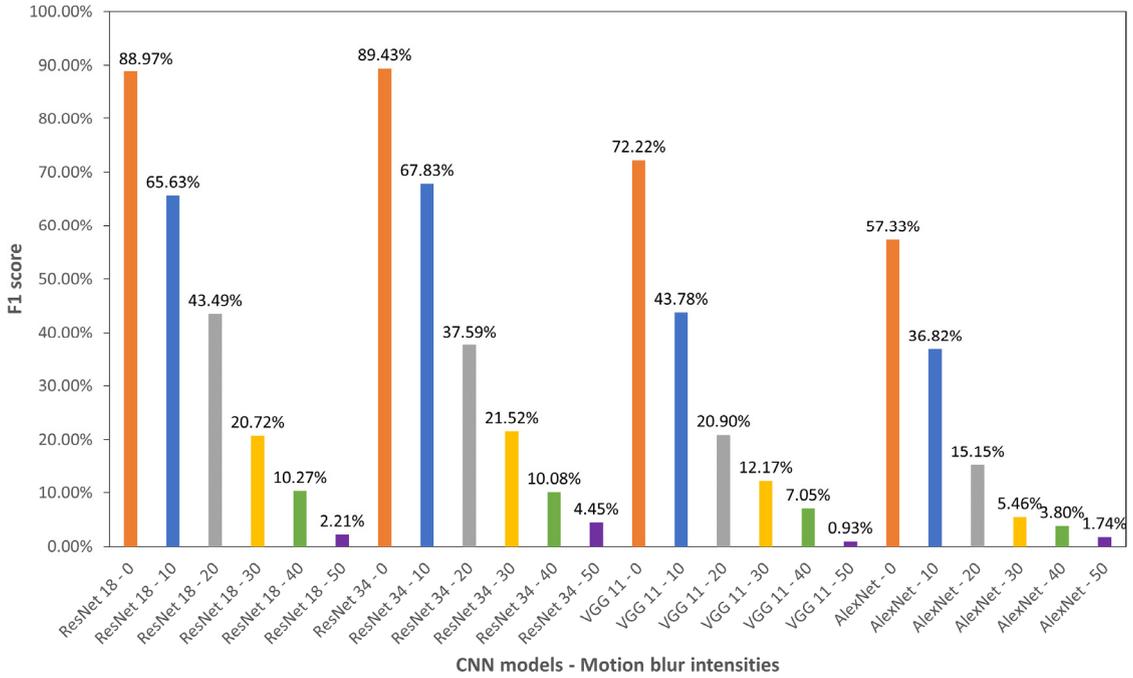


Fig. 5. Results of crack diagnosis by 4 CNN models of backbone about 5 motion blur intensities of KICT dataset

Fig. 6은 Kaggle 데이터셋과 KICT 데이터셋의 균열 검출 성능을 비교한 것이다. 원본 이미지에 대한 F1-score를 비교한 것이다. 동일한 학습 모델을 활용하였지만, 균열 검출 결과에서는 큰 차이를 보이고 있다. ResNet 18은 Kaggle 데이터셋에서 71.74%이고 KICT 데이터셋에서는 88.97%로 약 17%, ResNet 34은 약 18%, VGG 11은

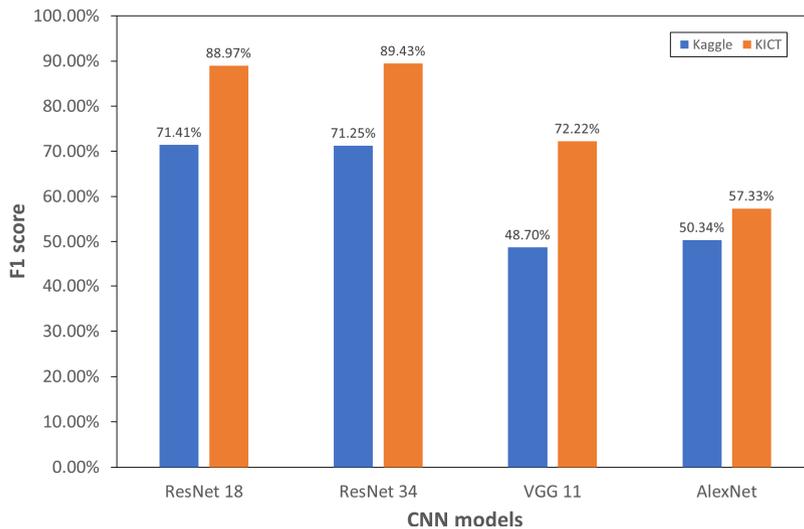


Fig. 6. Comparison of crack diagnosis by 4 CNN models of backbone between Kaggle and KICT dataset

약 23%, 그리고 AlexNet은 27%의 차이가 있는 것으로 나타났다. 또한, 미세한 차이지만 Kaggle 데이터셋에는 ResNet 18이 ResNet 34보다 높은 결과를 보였으나, KICT 데이터셋에서는 ResNet 34가 더 높은 결과로 나타났다. 그리고 KICT 데이터셋에서 VGG 11의 72.22%는 Kaggle 데이터셋의 ResNet 18의 71.41% 그리고 ResNet 34의 71.25%보다 높게 나타났다. 또한, Kaggle 데이터셋에서 비슷한 성능을 보였던 VGG 11과 AlexNet은 KICT 데이터 셋에서 VGG 11은 약 30% 이상 향상된 반면 AlexNet은 약 7% 정도로 작은 성능 개선 효과를 보임으로써 학습 데이터의 품질 차이에 의한 CNN 모델의 성능적 우위를 파악하기에는 어려운 것으로 볼 수 있다.

이러한 결과로부터 CNN기반 균열 검출 성능은 이미지 데이터의 품질에 따른 차이가 발생할 수 있으며, 구축되는 데이터셋의 품질에 따라서도 CNN 모델의 균열 검출 성능이 상이할 수 있는 것으로 판단된다. 그러므로 다양하게 개발되는 CNN 아키텍처의 발전과 함께 시설물 유지관리 측면에서는 데이터셋을 위한 취득단계부터 전처리 단계까지 이미지의 품질을 정량적으로 평가하는 것이 필요할 것으로 사료된다.

## 5. 결론

본 연구에서는 터널의 콘크리트 라이닝에 발생하는 균열을 검출하기 위해서 사용되는 이미지의 MB 강도 변화에 따른 CNN 모델의 균열 검출 성능을 분석하였다. 기존 연구에서 서로 다른 촬영조건 및 해상도를 갖는 이미지로 구축된 학습데이터로 인해 CNN 모델들의 성능 비교가 어렵다는 문제점을 가지고 있다. Kaggle 데이터셋과 KICT 데이터셋에서 10~50 범위의 MB강도를 가진 이미지를 생성하여 총 12개의 데이터셋을 구축하였다. U-Net 아키텍처 기반에 백본으로 ResNet 18, ResNet 34, VGG 11 그리고 AlexNet 등 4종류의 CNN 모델을 활용하여 균열 검출 성능을 F1-score를 통해 분석하여 도출된 결론은 다음과 같다.

1. Kaggle과 KICT 데이터셋에 대한 4개의 CNN 모델의 균열 검출 성능을 분석한 결과, 이미지의 MB 강도가 10에서 50으로 점차 증가할수록 F1-score는 낮아지며 CNN 성능 또한 감소하는 것으로 확인되었다. 이러한 결과를 통해 CNN을 활용한 균열 검출 시, 이미지의 MB는 CNN 모델의 성능을 저하시키는 요인으로 판단된다.
2. 동일한 학습모델을 활용한 경우, Kaggle과 KICT 데이터셋의 F1-score는 약 20% 이상의 큰 차이를 보였다. 해당 결과로부터 구축되는 학습 데이터셋의 품질에 따라서도 CNN 모델의 균열 검출 성능이 상이할 수 있는 것으로 판단된다.
3. 본 연구의 결과로 다양하게 개발되는 CNN 아키텍처와 함께 시설물 유지관리 측면에서 데이터셋을 위한 취득 단계부터 분류, 정제를 통한 전처리 단계까지 이미지의 품질을 정량적으로 평가하여 고품질의 일관성을 확보할 수 있는 데이터 관리 기술이 필요할 것으로 사료된다.

## 감사의 글

본 연구는 국토교통과학기술진흥원의 기반시설 첨단관리(total care) 기술개발사업(RS-2022-00142566)의 지원으로 수행되었습니다. 이에 감사드립니다.

## 저자 기여도

이철희는 연구 설계, 데이터 수집, 원고 작성을 하였고, 김동구와 김동규는 연구개념, 분석결과 및 원고 검토를 하였다.

## References

1. Ali, L., Alnajjar, F., Jassmi, H.A., Gocho, M., Khan, W., Serhani, M.A. (2021), "Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures", *Sensors*, Vol. 21, No. 5, 1688.
2. Alidoost, F., Austen, G., Hahn, M. (2022), "A multi-camera mobile system for tunnel inspection", *iCity. Transformative Research for the Livable, Intelligent, and Sustainable City*, Springer, pp. 211-224.
3. Badrinarayanan, V., Kendall, A., Cipolla, R. (2017), "SegNet: a deep convolutional encoder-decoder architecture for image segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481-2495.
4. Bae, H., Jang, K., An, Y.K. (2021), "Deep super resolution crack network (SrcNet) for improving computer vision-based automated crack detectability in in situ bridges", *Structural Health Monitoring*, Vol. 20, No. 4, pp. 1428-1442.
5. Balaguer, C., Montero, R., Victores, J.G., Martínez, S., Jardón, A. (2014), "Towards fully automated tunnel inspection: a survey and future trends", *Proceedings of the 31st ISARC, Sydney, Australia*, pp. 19-33.
6. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H. (2018), "Encoder-decoder with atrous separable convolution for semantic image segmentation", *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, pp. 801-818.
7. Dynamic UNet, Unet model using PixelShuffle ICNR upsampling that can be built on top of any pretrained architecture, <https://docs.fast.ai/vision.models.unet.html> (November 1, 2024).
8. Fastai, <https://docs.fast.ai/> (September 1, 2024).
9. Guo, J., Liu, P., Xiao, B., Deng, L., Wang, Q. (2024), "Surface defect detection of civil structures using images: Review from data perspective", *Automation in Construction*, Vol. 158, 105186.
10. Guo, J., Wang, Q., Li, Y., Liu, P. (2020), "Façade defects classification from imbalanced dataset using meta learning-based convolutional neural network", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 35, No. 12, pp. 1403-1418.
11. He, K., Zhang, X., Ren, S., Sun, J. (2016), "Deep residual learning for image recognition", *Proceedings of*

- the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770-778.
12. Kaggle, Crack segmentation dataset, <https://www.kaggle.com/datasets/lakshaymiddha/crack-segmentation-dataset> (September 1, 2024).
  13. Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012), "Imagenet classification with deep convolutional neural networks", Proceedings of the Advances in Neural Information Processing Systems 25, NIPS, Stateline, pp. 1097-105.
  14. Kumar, S.S., Wang, M., Abraham, D.M., Jahanshahi, M.R., Iseley, T., Cheng, J.C.P. (2020), "Deep learning-based automated detection of sewer defects in CCTV videos", Journal of Computing in Civil Engineering, Vol. 34, No. 1, 04019047.
  15. Lee, C., Kim, D., Kim, D. (2024), "Analysis of the application of image quality assessment method for mobile tunnel scanning system", Journal of Korean Tunnelling and Underground Space Association, Vol. 26, No. 4, pp. 365-384.
  16. Li, D., Xie, Q., Gong, X., Yu, Z., Xu, J., Sun, Y., Wang, J. (2021), "Automatic defect detection of metro tunnel surfaces using a vision-based inspection system", Advanced Engineering Informatics, Vol. 47, 101206.
  17. Liu, Y., Yeoh, J.K.W., Chua, D.K.H. (2020), "Deep learning-based enhancement of motion blurred UAV concrete crack images", Journal of Computing in Civil Engineering, Vol. 34, No. 5, 04020028.
  18. Liu, Z., Cao, Y., Wang, Y., Wang, W. (2019), "Computer vision-based concrete crack detection using U-Net fully convolutional networks", Automation in Construction, Vol. 104, pp. 129-139.
  19. Long, J., Shelhamer, E., Darrell, T. (2015), "Fully convolutional networks for semantic segmentation", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, pp. 3431-3440.
  20. Mei, Q., Gül, M., Azim, M.R. (2020), "Densely connected deep neural network considering connectivity of pixels for automatic crack detection", Automation in Construction, Vol. 110, 103018.
  21. Miao, P., Srimahachota, T. (2021), "Cost-effective system for detection and quantification of concrete surface cracks by combination of convolutional neural network and image processing techniques", Construction and Building Materials, Vol. 293, 123549.
  22. Ni, F., Zhang, J., Chen, Z. (2019), "Zernike-moment measurement of thin-crack width in images enabled by dual-scale deep learning", Computer-Aided Civil and Infrastructure Engineering, Vol. 34, No. 5, pp. 367-384.
  23. Pan, S.J., Yang, Q. (2010), "A survey on transfer learning", IEEE Transactions on Knowledge and Data Engineering, Vol. 22, No. 10, pp. 1345-1359.
  24. Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016), "You only look once: unified, real-time object detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 779-788.
  25. 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, Vol. 39,

No. 6, pp. 1137-1149.

26. Shin, H., Ahn, Y., Tae, S., Gil, H., Song, M., Lee, S. (2021), "Enhancement of multi-class structural defect recognition using generative adversarial network", *Sustainability*, Vol 13, No. 22, 12682.
27. Simonyan, K., Zisserman, A. (2015), "Very deep convolutional networks for large-scale image recognition", *ArXiv*, Vol. 1409, 1556.
28. Spencer Jr, B.F., Hoskere, V., Narazaki, Y. (2019), "Advances in computer vision-based civil infrastructure inspection and monitoring", *Engineering*, Vol. 5, No. 2, pp. 199-222.
29. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabino- vich, A. (2015), "Going deeper with convolutions", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, pp. 1-9.
30. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. (2016), "Rethinking the inception architecture for computer vision", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp. 2818-2826.
31. Wang, X., Kong, T., Shen, C., Jiang, Y., Li, L. (2020), "SOLO: segmenting objects by locations", *Proceedings of the Computer Vision - ECCV 2020: Part XVIII* 16, Glasgow, UK, pp. 649-665.
32. Wei, F., Yao, G., Yang, Y., Sun, Y. (2019), "Instance-level recognition and quantification for concrete surface bughole based on deep learning", *Automation in Construction*, Vol 107, 102920.
33. Xue, Y., Li, Y. (2018), "A fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 33, No. 8, pp. 638-654.
34. Yu, Y., Samali, B., Rashidi, M., Mohammadi, M., Nguyen, T.N., Zhang, G. (2022), "Vision-based concrete crack detection using a hybrid framework considering noise effect", *Journal of Building Engineering*, Vol. 61, 105246.