

# 지하시설물 안전점검을 위한 딥러닝 기반 콘크리트 균열 검출

전익¹ · 이임평² · 김동규³\*

¹정회원, (주)이노팜 인공지능팀 책임연구원

²비회원, 서울시립대학교 공간정보공학과 교수

³정회원, 한국건설기술연구원 지반연구본부 선임연구위원

## Crack detection in concrete using deep learning for underground facility safety inspection

Eui-ik Jeon<sup>1</sup> · Impyeong Lee<sup>2</sup> · Donggyou Kim<sup>3\*</sup>

<sup>1</sup>Senior Researcher, AI Team, Innopam Co., Ltd.

<sup>2</sup>Professor, Dept. of Geoinformatics, University of Seoul

<sup>3</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

The cracks in the tunnel are currently determined through visual inspections conducted by inspectors based on images acquired using tunnel imaging acquisition systems. This labor-intensive approach, relying on inspectors, has inherent limitations as it is subject to their subjective judgments. Recently research efforts have actively explored the use of deep learning to automatically detect tunnel cracks. However, most studies utilize public datasets or lack sufficient objectivity in the analysis process, making it challenging to apply them effectively in practical operations. In this study, we selected test datasets consisting of images in the same format as those obtained from the actual inspection system to perform an objective evaluation of deep learning models. Additionally, we introduced ensemble techniques to complement the strengths and weaknesses of the deep learning models, thereby improving the accuracy of crack detection. As a result, we achieved high recall rates of 80%, 88%, and 89% for cracks with sizes of 0.2 mm, 0.3 mm, and 0.5 mm, respectively, in the test images. In addition, the crack detection result of deep learning included numerous cracks that the inspector could not find. If cracks are detected with sufficient accuracy in a more objective evaluation by selecting images from other tunnels that were not used in this study, it is judged that deep learning will be able to be introduced to facility safety inspection.

**Keywords:** Deep learning, Semantic segmentation, Concrete lining, Facility safety inspection, Micro crack

### OPEN ACCESS

Journal of Korean Tunnelling and Underground Space Association  
25(6)555-567(2023)  
<https://doi.org/10.9711/KTAJ.2023.25.6.555>

eISSN: 2287-4747  
pISSN: 2233-8292

Received October 31, 2023  
Revised November 13, 2023  
Accepted November 14, 2023



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://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 © 2023, Korean Tunnelling and Underground Space Association

## 초 록

현재 지하시설물의 균열을 영상 취득 시스템으로 취득한 경우 점검자가 취득된 영상에서 육안검사를 수행하여 미세균열을 판단한다. 점검자에 의존한 노동집약적인 방법은 점검자의 주관적인 판단에 영향을 받는 문제점을 가지고 있다. 최근에는 딥러닝을 활용하여 자동으로 콘크리트 균열을 탐지하기 위한 연구가 활발하게 수행되고 있다. 대부분의 연구에서는 공개 데이터셋을 활용하거나 분석과정의 객관성이 충분하지 못해 실제 업무에 적용하기 어려운 점이 있다. 본 연구는 실제 검사 시스템과 동일한 형태의 영상을 시험 데이터셋으로 선정하여 딥러닝 모델들을 평가하였다. 균열 탐지의 정확도를 향상시키기 위하여 딥러닝 모델들의 장단점을 상호 보완할 수 있는 앙상블 기법을 적용하였다. 시험 영상에서 폭 0.2 mm, 0.3 mm 및 0.5 mm의 균열들은 각각 80%, 88% 및 89%의 높은 재현율로 탐지되었다. 딥러닝을 적용한 균열 탐지 결과에서는 점검자의 육안 검사 과정에 찾지 못한 다수의 균열들을 포함하고 있었다. 향후 본 연구에서 사용하지 않은 다른 터널의 영상을 시험 영상으로 선정하여 보다 더 객관적인 평가에서 충분한 정확도로 균열을 탐지하게 된다면, 시설물 안전 점검 방식에 딥러닝의 도입이 가능할 것으로 판단된다.

**주요어:** 딥러닝, 의미론적 분할, 콘크리트 라이닝, 시설물 안전 점검, 미세균열

## 1. 서론

현재 터널과 같은 지하 시설물의 점검 및 진단시 이동식 차량에 영상 취득 시스템을 탑재하여 터널 내부손상 정보를 수집하는 경우가 증가하고 있다. 작은 손상, 특히 미세 균열의 경우 수집된 터널 내부 영상을 점검자가 육안으로 보면서 손상 정보를 파악하고 있다. 점검자의 육안에 의존한 전수검사 방식은 매우 노동집약적인 한계가 존재하고, 점검자의 주관적인 판단에 영향을 받는 한계점이 존재한다(Bae et al., 2022). 이러한 문제점을 개선하기 위하여 딥러닝을 활용하여 균열을 자동으로 검출하기 위한 연구들이 수행되고 있다.

Kim and Cho (2018)는 이상적인 조건이 아닌 실제 다양한 현장 환경의 콘크리트 표면의 균열을 AlexNet을 사용하여 탐지하였다. Kim et al. (2018)은 객체분류 모델을 이용하여 공개 자료의 영상에서 균열의 존재 유무를 판단하고, 균열이 존재하면 의미론적 분할을 수행하여 균열을 추출하는 연구를 수행하였다. 콘크리트 표면에서 균열 분할에 딥러닝이 충분히 활용 가능함에 따라, 딥러닝 모델의 성능 비교 및 개선하기 위한 연구들이 수행되었다. Liu et al. (2019)은 분류 모델인 VGG-16을 응용하여 의미론적 분할 모델인 DeepCrack을 제안하였다. 연구자들은 구조물의 균열 탐지에 의미론적 분할의 대표 모델인 UNet, DeepLabv3+, FCN-8s들의 성능을 비교하였다(Hadinata et al., 2021; Lee et al., 2023). 딥러닝 모델 비교 및 개선 이외에 균열의 폭을 확장하여 탐지 능력을 향상하는 방법을 제안하고, 이를 증명하기 위해 개선된 UNet 모델과 다양한 공개 균열 자료들에 적용하여 제안한 방법을 검증하였다(Pandey and Achara, 2022).

선행 연구들에서는 주로 공개 데이터셋을 사용하여 균열 탐지에 대한 딥러닝의 효용성 증명 및 성능 개선을 수행하였다. 균열 탐지를 위한 대표적인 공개 데이터셋은 Kaggle의 crack segmentation이 있다. Ham et al. (2022)은 터널 내 균열 탐지에서 공개 자료의 활용 가능성을 분석하였다. 터널 영상 취득 시스템(이하 시스템)과 유사한

공개 자료는 활용 가능하다고 증명하였다. Bae et al. (2022)은 공개 자료와 시스템의 영상을 혼합하여 공개 자료의 활용 가능성을 정량적으로 분석하였으나 공개 자료는 큰 의미가 없는 것으로 나타났다. Bae (2023)는 대략적으로 균열이 기록된 자료를 정밀하게 재구축한 소량의 고품질 자료가 정확도 향상에 더 도움된다고 제안하였다. 이 과정에서 CNN 기반의 DeepLabv3+와 어텐션 모듈을 활용한 DANet, 그리고 트랜스포머 기반의 SETR을 사용하였다. 하지만 고품질의 자료가 충분하지 않아 SETR의 성능이 낮게 나타났는데, 이는 현재 COCO와 같은 공개 자료 기반의 영상 분할 모델의 성능 평가에서 트랜스포머 기반 모델들이 최상위권을 차지하고 있는 점과 대치된다. 즉, 트랜스포머 모델을 활용하기 위해서는 충분한 자료 확보가 필요한 점과 현재 시스템의 영상에서 정확한 균열을 탐지하기 위해서는 공개 자료보다는 시스템 기반의 정밀한 학습 자료 구축이 필요한 것으로 판단할 수 있다.

본 연구에서는 4개의 터널에서 시스템으로 획득한 16개의 콘크리트 라이닝 스팬 영상과 기존 균열이 기록된 자료를 활용하여 정밀한 균열 자료를 제작하였다. 딥러닝 모델의 객관적인 성능 평가를 위해 16개의 영상 중 2개를 시험 자료로 선정하고, 나머지 14개는 딥러닝 모델의 원활한 학습을 위해 데이터 증강 기법을 적용하여 충분한 양의 학습 자료로 구축하였다. 딥러닝 모델은 정확한 균열 탐지를 위해 선행 연구에서 선정한 DeepLabv3+, 트랜스포머 기반의 최신 모델인 Mask2Former, 그리고 CNN 기반의 UPerNet을 선정하였다. 모델들의 성능 비교 및 평가를 통해 장점과 단점을 분석하고 앙상블 기법을 적용하여 균열 탐지 성능을 향상하고자 하였다.

## 2. 방법론

육안으로 구분하는 미세균열 검출방식을 개선하기 위하여 딥러닝을 활용한 일련의 연구들이 진행되었다(Bae et al., 2022; Ham et al., 2022; Bae, 2023). 선행 연구들에서 터널 영상 취득 시스템에서 수집된 단순 균열 자료와 추가 레이블 작업으로 생성된 정교한 균열 자료를 비교하여 고품질 데이터셋의 중요성을 강조하였다. Fig. 1에서 보여주듯이 본 연구에서는 터널 영상 취득 시스템으로 수집한 단순 균열 자료를 정교한 균열 자료로 제작하였다.

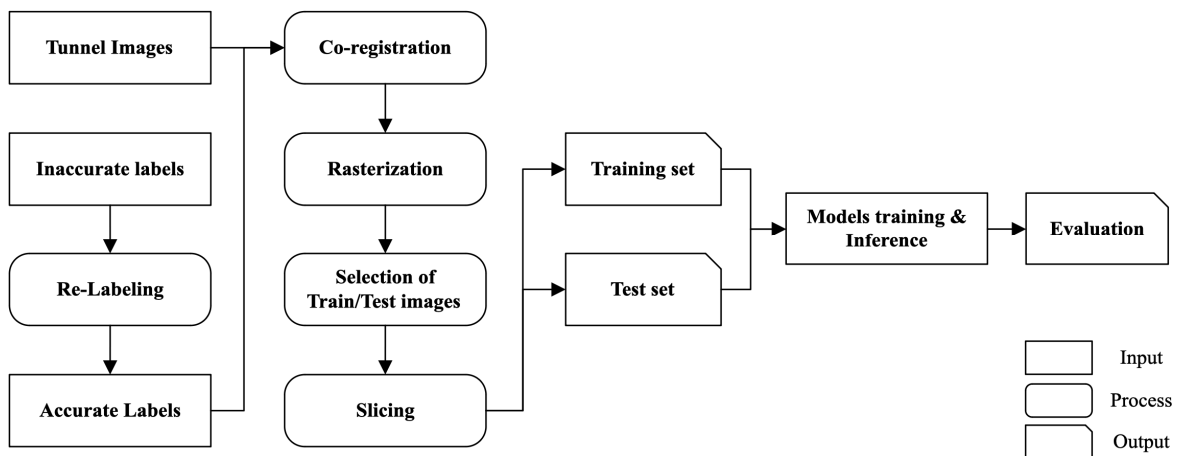


Fig. 1. Workflow of data preparation, processing, analysis in this study

제작된 정교한 균열 자료를 기반으로 이미지 증강 기법으로 딥러닝 모델의 학습이 충분히 이루어질 수 있는 양의 학습 데이터셋을 생성하였다. 본 연구에서는 선행 연구에서 우수한 성능을 보인 CNN 기반의 DeepLabv3+와 Vision transformer 기반의 Mask2former를 선정하고, 최근 공개된 CNN 기반의 InterImage를 기반 모델로 한 UPerNet을 추가로 선정하였다. 선정된 모델들은 Fig. 1에서 보여주는 학습 및 평가 과정을 거쳤다.

## 2.1 학습 데이터 구축

본 연구에서는 4개의 터널에서 영상시스템으로 수집한 기존 결과물들을 활용하였다. 해당 결과물은 약 10,000 × 23,000화소의 16개의 영상과 각각의 영상의 균열이 기록된 외관망도로 구성되어 있다. 다만, 외관망도는 Fig. 2와 같이 균열을 단순하게 표현하였기 때문에 학습 자료로 구축하기 위해 정교한 균열 자료로 다시 제작하는 과정을 수행하였다. 제작된 정교한 균열 자료는 QGIS 소프트웨어를 활용하여 영상과 상호정합하고 균열을 포함한 화소의 값은 1, 그 이외에 화소는 0으로 설정하여 래스터화하였다(Bae et al., 2022). 앞선 연구에서는 정교하게 제

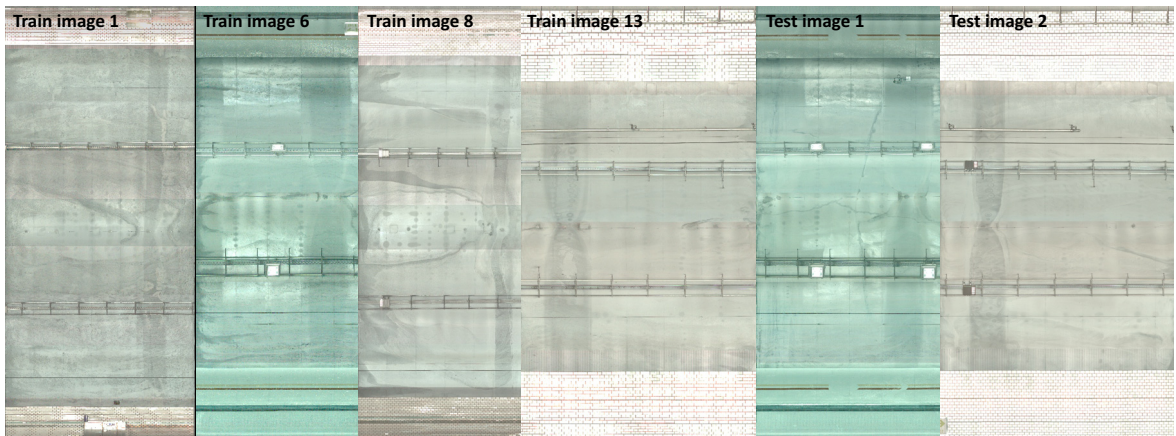


Fig. 2. Example of train and test images acquired by image data acquisition system for tunnel

Table 1. Summary of training and test datasets in this study

Image name	Tunnel	Pixels	Crack length	Image name	Tunnel	Pixels	Crack length
Train image 1	A	10,035 × 23,040	10.83 m	Train image 9	C	9,002 × 23,040	9.83 m
Train image 2	A	10,056 × 23,040	11.07 m	Train image 10	C	10,036 × 23,040	11.19 m
Train image 3	A	9,024 × 23,040	12.22 m	Train image 11	C	10,107 × 23,040	11.02 m
Train image 4	A	8,974 × 23,040	17.84 m	Train image 12	C	10,011 × 23,040	14.08 m
Train image 5	A	9,995 × 23,040	11.81 m	Train image 13	D	10,000 × 18,268	17.88 m
Train image 6	B	8,000 × 20,583	9.13 m	Train image 14	D	10,000 × 18,268	13.78 m
Train image 7	B	8,974 × 23,040	13.02 m	Test image 1	B	9,000 × 20,583	18.36 m
Train image 8	C	9,103 × 23,040	9.88 m	Test image 2	D	10,000 × 18,268	13.86 m

작된 2개의 영상을 타일링 후 영상 회전 등의 데이터를 증강하고 학습/검증/시험 데이터셋으로 재분류를 했다. 이와 같은 방식은 학습/검증/시험 데이터셋으로 분류할 때, 학습과 시험 데이터셋이 상당히 유사할 수 있어서 객관성이 낮다는 한계를 가지고 있다. 그래서 본 연구에서는 16개의 영상 중 14개의 영상을 학습에 사용하고, 학습에는 관여하지 않는 2개의 영상을 시험 자료로 하여 딥러닝 모델의 성능 평가의 객관성을 높일 수 있도록 하였다 (Table 1). 본 연구의 학습과 시험에 사용한 각각의 영상의 예시는 Fig. 2와 같다.

선별된 영상들은 딥러닝 모델의 학습에 사용할 수 있도록 특정 크기의 타일로 분할하는 과정이 필요하다. 타일로 분할하기 전의 균열 자료는 폴리라인으로 구성되어 있어 일정한 폭을 가진 균열을 정확히 반영하지 못하여, 모폴로지 연산 중 팽창의 반경을 10으로 적용하였다. 타일의 크기는 딥러닝 모델의 입력 크기에 따라 다르게 설정되는데, 본 연구에서 선정한 모델들의 입력 영상의 크기에 결정하였다. 3개의 모델들은 공통적으로 ADE20K 데이터셋에 DeepLabv3+, Mask2Former, UPerNet 순으로 입력 영상을  $512 \times 512$ ,  $640 \times 640$ ,  $896 \times 896$  크기로 설정하여 학습된 결과를 제공한다. 본 연구에서도 이와 동일하게 3개의 타일 크기로 각각 제작하였다.

분할된 타일은 75%의 중복도를 갖도록 하고 패치에 균열이 없으면 학습에서 제외하여 6,599쌍, 5,117쌍, 3,526쌍의 타일을 생성하였다. 생성된 타일은 영상 증강 기법 중 90도, 180도, 270도 회전을 사용하여 26,392쌍, 20,468쌍, 14,104쌍으로 증강하였다. 시험 데이터셋은 학습 데이터셋과 다르게 중복도 없이  $512 \times 512$ ,  $612 \times 612$ ,  $892 \times 892$  크기로 타일을 분할하고, 균열이 포함되지 않은 패치도 시험 데이터셋에 포함하였다. 2개의 시험 영상을 그대로 사용한 것과 동일하게 하기 위해 영상 증강 기법은 적용하지 않아 최종적으로 각각 1,346쌍, 837쌍, 399쌍의 타일이 생성되었다.

## 2.2 딥러닝 모델

본 연구에서는 미세 균열의 분할 성능 비교 및 평가에 DeepLabv3+ (Chen et al., 2018), Mask2Former (Cheng et al., 2022), UPerNet (Xiao et al., 2018)을 사용한다(Fig. 3). DeepLabv3+는 전통적인 의미론적 분할 모델로 인코더-디코더 구조와 피라미드 풀링 모듈을 결합을 특징으로 하며 ResNet50을 백본으로 선정하였다. Mask2Former는 masked 어텐션으로 구성된 트랜스포머를 디코더로 하는 특징을 가지고 있으며, Swin 트랜스포머를 백본으로 선정하여 높은 수준의 정확도를 보이고 있다(Liu et al., 2021). UPerNet은 피라미드 풀링 모듈을 기반으로 저수

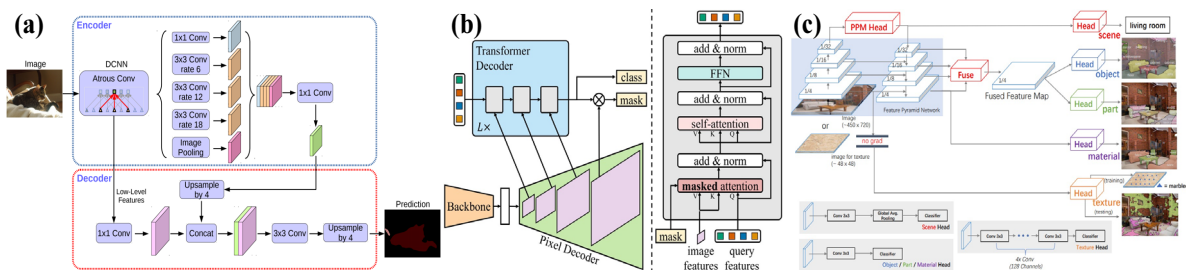


Fig. 3. Model's architecture: (a) DeepLabv3+, (b) Mask2Former, (c) UPerNet

준에서부터 고수준의 의미 정보를 융합하여 분할 정확도를 높인 모델로, 최근 공개된 CNN 기반의 기반 모델인 InternImage (Wang et al., 2023)를 백본으로 하였다. InternImage는 CNN로 만들어진 기반 모델로 트랜스포머의 기반 모델보다 Cityscapes, ADE20K 자료에 대한 좋은 성능을 보였다.

### 2.3 평가 지표

의미론적 분할의 평가 지표는 화소 정확도, 정밀도, 재현율, F1 값, 그리고 IoU를 사용하며 식 (1)~(4)와 같이 표현된다. 본 연구는 영상에서 균열과 그 이외 배경을 분류하는 이진 분류이므로 균열을 참, 나머지를 거짓으로 할 때, 실제 값과 분류결과에 따라 참양성(True Positive, TP), 거짓양성(False Positive, FP), 거짓음성(False Negative, FN), 참음성(True Negative, TN)으로 표현할 수 있다. 화소 정확도는 분할 결과에서 참양성과 참음성의 합의 비율(식 (1)), 정밀도는 참으로 분류한 결과 중에서 실제 참으로 판별한 비율(식 (2)), 재현율은 실제 참값에서 참값으로 판별한 비율을 의미한다(식 (3)). 화소 정확도는 객체들이 차지하는 화소 수의 비율이 불균등하면 분할 결과의 해석할 때 과소 또는 과대 해석의 가능성이 존재한다(Jeon et al., 2021). 분할 결과 평가에는 정밀도와 재현율 뿐만 아니라 정밀도와 재현율의 조화평균을 의미하는 F1 값을 함께 사용한다(식 (4)). 본 연구에서는 균열보다 배경의 화소 수가 월등히 많기 때문에 화소 정확도는 평가 지표에서 제외하고 정밀도, 재현율, F1 값을 평가 지표로 사용했다.

$$Pixel\ accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1\ score = \frac{2TP}{2TP + FN + FP} \quad (4)$$

## 3. 실험 결과

본 연구에서 딥러닝 모델의 학습에 사용한 하드웨어의 CPU, Memory, GPU 사양은 Intel(R) Xeon(R) CPU E5-2698, 256GB, NVIDIA사의 Tesla V100 32GB이다. 학습에 사용한 소프트웨어는 다양한 의미론적 분할 모델과 기능을 제공하는 Pytorch 1.11.0 기반의 OpenMMLab의 MMSegmentation을 사용하였다. 각 모델들의 학습

에서 최적화 모델은 공통적으로 AdamW를 사용하고, 학습률은 모델 순서대로 0.0001, 0.0001, 0.00002, Iteration은 80,000회로 하여 학습하도록 설정하였다. 본 연구의 선행 연구들에서 Kaggle의 균열 데이터셋을 활용하거나 1차 학습에 사용하였으나, 본 연구에서는 학습 자료가 충분한 것으로 판단하여 Kaggle 데이터셋은 사용하지 않고 ADE20K 데이터셋으로 사전 학습된 결과물을 이용하여 전이학습을 수행하였다.

### 3.1 딥러닝 모델들의 성능 비교

각 딥러닝 모델들의 학습 과정에서의 손실값, 정밀도, 재현율, F1 값의 변화를 Fig. 4에 도식화하였다. 먼저 손실값의 경향을 모델별로 살펴보면 DeepLabv3+과 UPerNet은 빠른 속도로 손실값이 감소하였다. 그러나 DeepLabv3+는 반복 학습 과정에서 손실값이 근소하게 증가하는 듯한 경향을 보이고 있으나 UPerNet은 손실값이 빠르게 감소 후 완만하게 일정한 값에 수렴하는 형태인 것을 알 수 있다. Mask2Former는 다른 모델보다 감소 속도가 느리지만 일정하게 손실값이 적어지는 형태를 나타냈다. 이런 경향은 평가 지표에서도 유사하게 반영되는 것을 알 수 있는데, DeepLabv3+의 재현율은 비교적 일정한 값을 유지하지만 정밀도는 크게 변화하면서 지속적으로 값이 낮아지는 경향을 보인다. 그에 반면 Mask2Former는 정밀도가 점차 증가하며 이는 Mask2Former의 학습 횟수가 증가하면 정밀도가 보다 더 증가할 수 있을 것으로 생각할 수 있다. 마지막으로 UPerNet은 손실 값이 빠르게 일정한 값으로 수렴한 것과 마찬가지로 평가지표들도 다른 모델과 다르게 빠르게 안정적인 형태의 그래프로 나타나는 것을 확인할 수 있다.

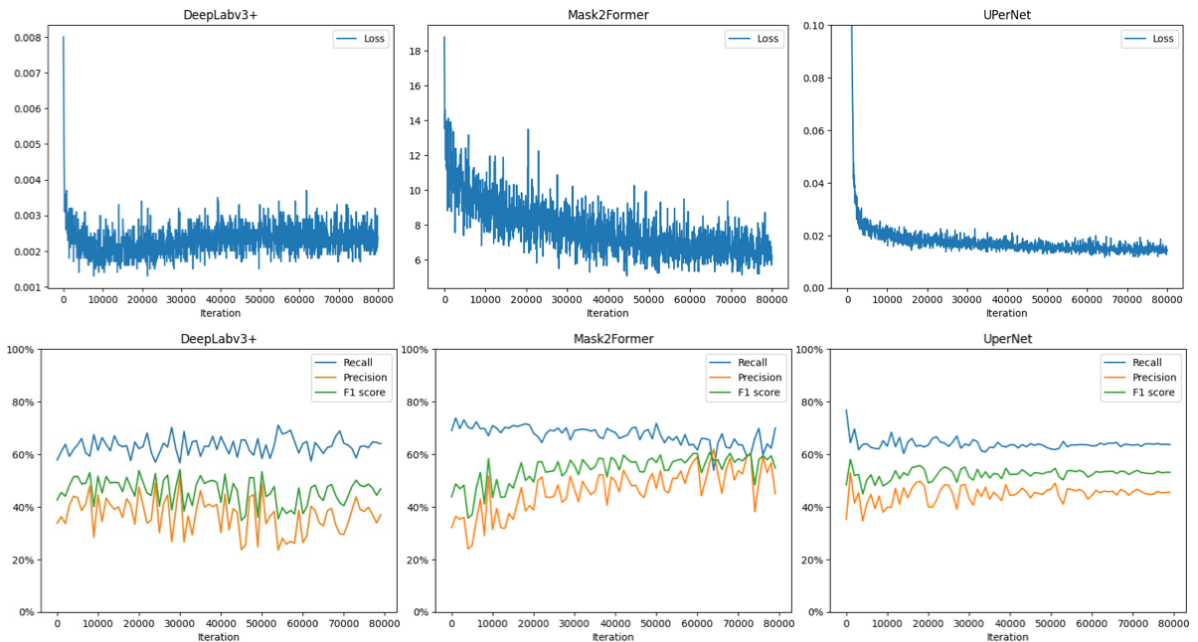


Fig. 4. Loss, precision, recall, F1 score curves of training

각 모델들의 학습 과정에서 평가 지표들의 변동이 발생하여 F1 값이 가장 높은 것을 기준으로 테스트 영상에 대한 추론 결과를 분석하였다. 이때 F1 값은 DeepLabv3+, Mask2Former, UPerNet 순으로 54%, 61%, 58%이다. 먼저, 각 딥러닝 모델의 성능을 정량적으로 평가하기 위해 두 개의 시험 영상에 대한 추론 결과의 정밀도, 재현율, F1 값을 Table 2에 정리하였다. Table 2에서 모델에 따라 F1 값을 기준으로 시험 영상 1에서는 DeepLabv3+, Mask2Former, UPerNet 순으로 58%, 72%, 69%로 Mask2Former가 다른 모델보다 3~14% 좋은 성능을 보였다. 이러한 편차의 주된 요인은 재현율은 같거나 크지 않은 차이를 보였지만, 정밀도에서는 17%의 큰 차이 때문이다.

**Table 2.** Results for each test image according to deep learning models

Model	Test image 1			Test image 2		
	Prec.	Recall	F1 score	Prec.	Recall	F1 score
DeepLabv3+	0.58	0.59	0.58	0.79	0.76	0.77
Mask2Former	<b>0.75</b>	<b>0.69</b>	<b>0.72</b>	<b>0.89</b>	<b>0.81</b>	<b>0.85</b>
UPerNet	0.70	<b>0.69</b>	0.69	0.83	0.78	0.80

시험 영상 1에 비해 0.3 mm, 0.5 mm의 폭을 가지는 균열이 주를 이루고 있는 시험 영상 2은 F1 값의 차이가 크지 않는 것을 알 수 있다. 이러한 경향은 GT와 함께 모든 모델의 결과를 도식화한 Fig. 5에서 확인할 수 있다. 시험 영상 2의 중심부에 있는 폭 0.5 mm의 균열은 모든 모델에서 완벽하게 탐지되었으며, 상단의 폭 0.3 mm 균열도 모델마다 근소한 차이가 있지만 균열 존재함을 충분히 인식할 수 있는 것으로 판단할 수 있다. 시험 영상 1의 중심부의 폭 0.3 mm 균열도 마찬가지로 높은 정확도로 탐지되었지만, 상단부의 폭 0.2 mm 균열에서는 차이가 있는 것을 알 수 있다. 가지적으로 비교를 하면 UPerNet, Mask2Former, DeepLabv3+ 순으로 탐지 결과가 좋지 않은 것을 확인할 수 있으며, 다른 균열과 다르게 UPerNet이 Mask2Former보다 0.2 mm 균열을 더 많이 탐지하는 것으로 판단할 수 있다. 이러한 결과는 모델들의 균열 크기 별 재현율을 정리한 Table 3에서 정량적으로 확인할 수 있다. 균열 크기가 폭 0.3 mm, 0.5 mm 균열에서는 재현율의 차이가 최대 7%였지만, 0.2 mm에서는 차이가 19%로 매우 크게 나타났다. 결과적으로 DeepLabv3+ 모델의 성능은 모든 균열에서 다른 모델보다 낮고, Mask2Former와 UPerNet은 균열 크기에 따라 탐지 성능이 달라짐에 따라 두 모델을 함께 이용하는 앙상블 기법을 활용하면 보다 높은 정확도로 균열을 탐지할 수 있을 것으로 판단된다. 그래서 DeepLabv3+를 제외한 Mask2Former와 UPerNet의 결과를 단순히 합하는 앙상블 기법을 적용한 결과를 다음 장에서 분석하였다.

**Table 3.** Results for each crack according to deep learning models

Crack size	0.2 mm	0.3 mm	0.5 mm
DeepLabv3+	52%	80%	82%
Mask2Former	65%	<b>85%</b>	<b>89%</b>
UPerNet	<b>71%</b>	78%	84%



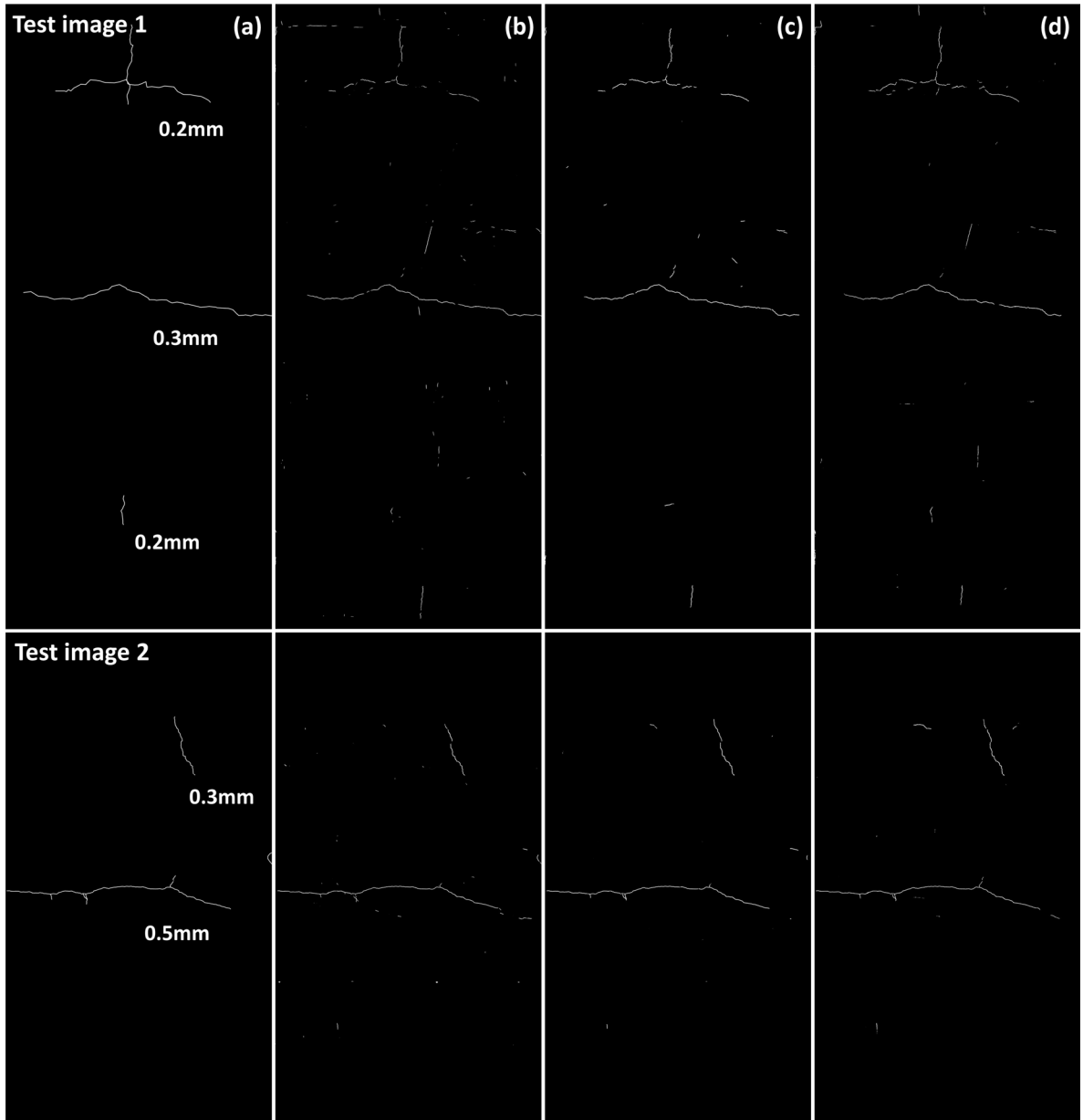


Fig. 5. Visual comparison of inference results: (a) GT. Results of (b) DeepLabv3+, (c) Mask2Former, (d) UPerNet

### 3.2 앙상블 기법 적용

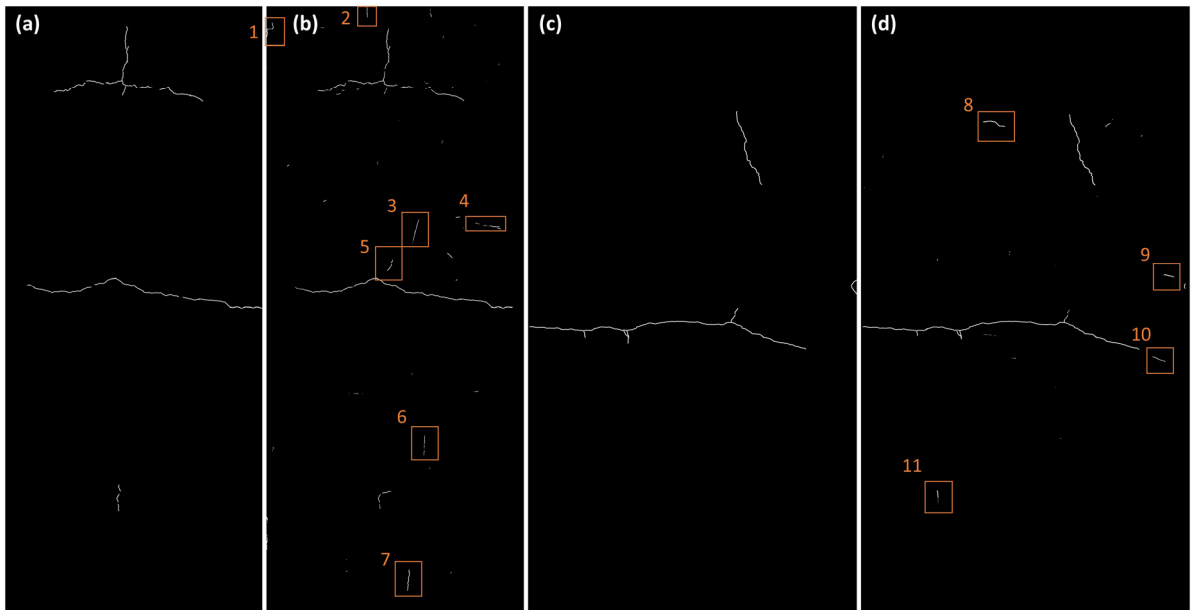
앙상블은 여러 개의 모델을 학습시켜 그 결과를 통합하는 기법을 의미하는 것으로, 일반적으로 개별 모델의 오차를 서로 상쇄시켜 모델의 성능을 개선하는 효과를 가지고 있다. 모델 추론 결과의 앙상블 기법은 대표적으로 Hard voting, Soft voting이 있다. Hard voting은 모델들이 가장 많이 선택한 클래스로 결정하는 방법, Soft voting

은 모델이 출력한 값들을 모두 더해서 가장 큰 값을 가진 클래스로 분류하는 방법을 의미한다(Kim and Kim, 2022). 이와 같이 모델의 결과를 통합하는 여러 앙상블 기법들이 존재하지만, 본 연구에서는 미세균열 탐지에서 앙상블 기법의 활용 가능성의 확인이 목적이기 때문에 각 모델의 추론 결과를 단순히 결합하는 간단한 방법을 선택하였다. 즉, 두 모델의 추론 결과에서 둘 다 또는 둘 중 하나의 모델이 균열이라고 판단했으면 그 화소는 균열로 결정되고, 그렇지 않으면 배경이라고 판단한 것이 된다.

이렇게 두 개의 모델의 추론 결과를 결합한 후 정확도 평가를 수행한 결과를 Table 4와 Fig. 6에 나타냈다. Table 4는 두 개의 시험 영상과 균열 폭 크기 별로 평가 지표 값들을 나타낸 것으로, 첫 번째 시험 영상에서 정밀도와 재현율은 65%, 79%로 Table 2의 최고값과 비교하여 정밀도는 낮아지고 재현율은 상승했다. 이는 두 번째 시험 영상에서도 동일하게 나타났는데, 재현율은 3개의 균열 크기에서 모두 80% 이상의 수치를 나타냈으며 특히

**Table 4.** Results for each test image of ensemble

	Precision	Recall	F1 score
Test image 1	65%	79%	71%
Test image 2	80%	84%	82%
Crack size: 0.2 mm	90%	80%	85%
Crack size: 0.3 mm	88%	88%	88%
Crack size: 0.5 mm	97%	89%	93%



**Fig. 6.** Visual comparison of ensemble results: (a) GT, (b) ensemble of test image 1, (c) GT, (d) ensemble of test image 2

폭 0.2 mm 균열에서 9% 이상 매우 높게 상승하였다. Fig. 5에서 UPerNet이 폭 0.2 mm 균열 탐지에 강인하고 폭 0.3 mm 균열 탐지에 강인한 Mask2Former의 결과가 상호 보완된 것으로 단순한 앙상블 기법을 적용하여 높은 정확도로 폭 0.2 mm 균열을 탐지한 것으로 판단할 수 있다.

단순한 결과의 결합 방법은 오탐지를 증가시키는 단점이 있는데, 이는 정밀도가 낮아지는 결과로 이어진다. Fig. 6에는 정밀도를 낮추는 균열 이외에 다수의 균열들이 탐지된 결과들을 보여주고 있다. 어떤 이유로 이런 오탐지가 발생하는지 가시적으로 확인하기 위해 11개의 주요 오탐지를 Fig. 7에 확대했다. Fig. 7에서 아주 미세하고 주위의 오염으로 인해 정확히 판단하기 어려운 5번과 10번을 제외한 나머지는 명확하게 균열로 판단된다. 이러한 결과는 현재 점검자의 육안에 의존한 전수검사 방식은 균열 기록에 누락이 발생할 수 있음을 알 수 있다.

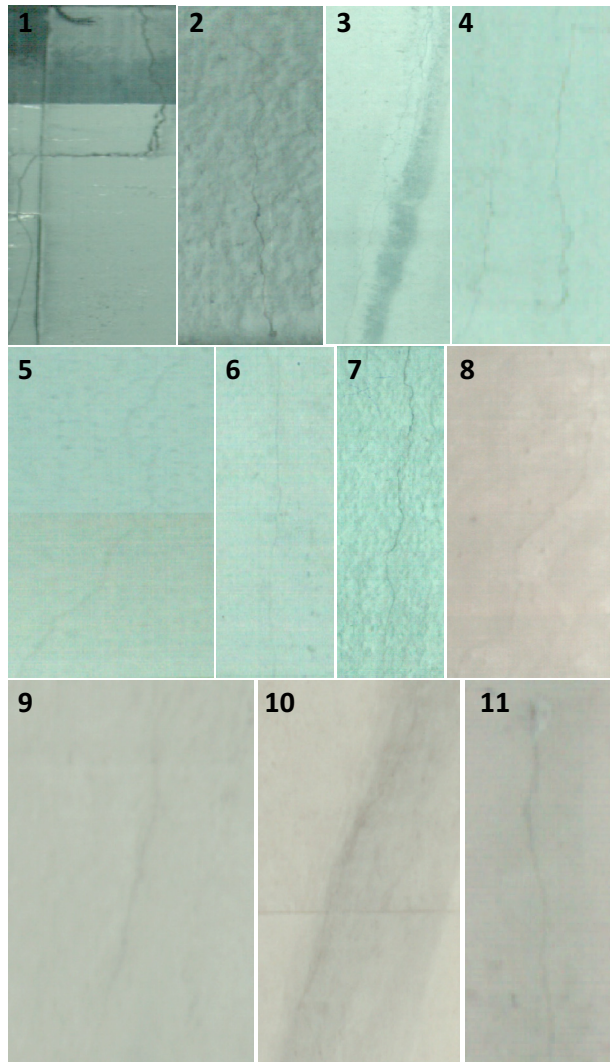


Fig. 7. Subset images identified as cracked

## 4. 결론

본 연구는 운영 중인 터널의 콘크리트 라이닝에서 발생한 균열들을 자동으로 인식하고 식별하기 위한 연구의 일환으로 수행되었다. 선행연구와의 주요 차이점으로는 고품질 학습 자료의 충분한 확보, 객관적인 평가를 위한 실제 검사시스템과 동일한 형태의 시험 영상 사용, 그리고 앙상블 기법을 활용한 균열 검출 정확도의 향상이 있다. 이러한 차이점을 기반으로 시험 영상에 대한 추론 결과에서 폭 0.2 mm, 0.3 mm, 0.5 mm의 균열들을 각각 80%, 88%, 89%의 높은 정확도로 검출하였다. 점검자의 직접 검수 과정에 찾지 못한 다수의 균열들을 딥러닝 모델의 결과에서 찾을 수 있었다. 본 연구의 결과에서 현재 점검자의 육안에 의존한 균열 전수 검사 방식에 딥러닝을 활용하여 주요 균열을 자동으로 탐지하여 그 영역을 제공하면 점검자가 탐지 결과를 확인하여 최종적으로 결정하는 방식으로의 전환이 가능할 것으로 판단된다. 다만, 본 연구에서 학습 자료는 4개의 터널에서 획득한 16개를 사용하여 그 중 2개의 터널의 2개 영상을 시험 자료로 선정하였다. 이는 학습 자료와 시험 자료의 유사성이 존재하는 한계가 있으므로, 향후 기 확보한 4개의 터널 이외에 새로운 터널에서 획득한 자료로 보다 객관적인 평가를 수행하여 실 활용이 가능할 것으로 판단된다. 또한, 본 연구에서 적용한 앙상블 기법은 추론 결과를 결합하는 단순한 방식으로, 전반적인 균열 검출에 강인한 Mask2Former와 폭 0.2 mm의 아주 미세한 균열 검출에 탁월한 UPerNet을 학습 및 결합 단계에서 다양한 앙상블 기법을 적용하면 보다 높은 정확도로 균열을 탐지할 수 있을 것으로 판단된다.

## 감사의 글

본 연구는 한국건설기술연구원 주요사업으로 지원을 받아 수행된 연구(인공지능을 활용한 대심도 지하 대공간의 스마트 복합 솔루션 개발)로 이에 감사합니다.

## 저자 기여도

전의익은 딥러닝 모델 학습, 결과분석 및 원고 작성을 하였다. 이임평은 연구 개념 설계 및 원고 검토를 하였고 김동규는 연구목적 및 내용 수립, 데이터 수집과 원고 검토를 하였다.

## References

1. Bae, S.H. (2023), Deep learning-based crack detection for facility safety inspection - focusing on tunnel cement concrete lining, Master Thesis, University of Seoul, pp. 35-46.
2. Bae, S.H., Ham, S.W., Lee, I.P., Lee, G.P., Kim, D.G. (2022), "Deep learning based crack detection from tunnel cement concrete lining", Journal of Korean Tunnelling and Underground Space Association, Vol. 24, No. 6, pp. 583-598.

3. 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.
4. Cheng, B., Misra, I., Schwing, A.G., Kirillov, A., Girdhar, R. (2022), "Masked-attention mask transformer for universal image segmentation", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, USA, pp. 1290-1299.
5. Hadinata, P.N., Simanta, D., Eddy, L., Nagai, K. (2021), "Crack detection on concrete surfaces using deep encoder-decoder convolutional neural network: a comparison study between U-Net and DeepLabV3+", Journal of the Civil Engineering Forum, Vol. 7, No. 3, pp. 323-334.
6. Ham, S.W., Bae, S.H., Lee, I.P., Lee, G.P., Kim, D.G. (2022), "An evaluation methodology for cement concrete lining crack segmentation deep learning model", Journal of Korean Tunnelling and Underground Space Association, Vol. 24, No. 6, pp. 513-524.
7. Jeon, E.I., Kim, S.H., Park, S.Y., Kwak, J.W., Choi, I.H. (2021), "Semantic segmentation of seagrass habitat from drone imagery based on deep learning: a comparative study", Ecological Informatics, Vol. 66, 101430.
8. Kim, A.R., Kim, D., Byun, Y.S., 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.
9. Kim, B.H., Cho, S.J. (2018), "Automated vision-based detection of cracks on concrete surfaces using a deep learning technique", Sensors, Vol. 18, No. 10, pp. 3452.
10. Kim, H.S., Kim, M.H. (2022), "A study on the image-based malware classification system that combines image preprocessing and ensemble techniques for high accuracy", KIPS Transactions on Computer and Communication Systems, Vol. 11, No. 7, pp. 225-232.
11. Lee, T.H., Kim, J.H., Lee, S.J., Ryu, S.K., Joo, B.C. (2023), "Improvement of concrete crack segmentation performance using stacking ensemble learning", Applied Sciences, Vol. 13, No. 4, 2367.
12. Liu, Y., Yao, J., Lu, X., Xie, R., Li, L. (2019), "DeepCrack: A deep hierarchical feature learning architecture for crack segmentation", Neurocomputing, Vol. 338, pp. 139-153.
13. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B. (2021), "Swin transformer: Hierarchical vision transformer using shifted windows", Proceedings of the IEEE/CVF International Conference on Computer Vision, Montrealers, Canada, pp. 10012-10022.
14. Pandey, R.K., Achara, A. (2022), CoreDeep: Improving crack detection algorithms using width stochasticity, arXiv preprint arXiv:2209.04648.
15. Wang, W., Dai, J., Chen, Z., Huang, Z., Li, Z., Zhu, X., Hu, X., Lu, T., Lu, L., Li, H., Wang, X., Qiao, Y. (2023), "InternImage: Exploring large-scale vision foundation models with deformable convolutions", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, pp. 14408-14419.
16. Xiao, T., Liu, Y., Zhou, B., Jiang, Y., Sun, J. (2018), "Unified perceptual parsing for scene understanding", Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, pp. 418-434.