

지하구조물 콘크리트 균열 탐지를 위한 semi-supervised 의미론적 분할 기반의 적대적 학습 기법 연구

심승보¹ · 최상일² · 공석민³ · 이성원^{4*}

¹정회원, 한국건설기술연구원 차세대 인프라연구센터 전임연구원

²비회원, 한국건설기술연구원 차세대 인프라연구센터 수석연구원

³정회원, 한국건설기술연구원 차세대 인프라연구센터 박사후 연구원

⁴정회원, 한국건설기술연구원 차세대 인프라연구센터 선임연구위원

Adversarial learning for underground structure concrete crack detection based on semi-supervised semantic segmentation

Seungbo Shim¹ · Sang-Il Choi² · Suk-Min Kong³ · Seong-Won Lee^{4*}

¹Research Specialist, Future Infrastructure Research Center, Korea Institute of Civil Engineering and Building Technology

²Senior Researcher, Future Infrastructure Research Center, Korea Institute of Civil Engineering and Building Technology

³Post-Doctoral Researcher, Future Infrastructure Research Center, Korea Institute of Civil Engineering and Building Technology

⁴Senior Research Fellow, Future Infrastructure Research Center, Korea Institute of Civil Engineering and Building Technology

*Corresponding Author : Seong-Won Lee, swlee@kict.re.kr

OPEN ACCESS

Journal of Korean Tunnelling and
Underground Space Association
22(5)515-528(2020)
<https://doi.org/10.9711/KTAJ.2020.22.5.515>

eISSN: 2287-4747

pISSN: 2233-8292

Received July 21, 2020

Revised August 4, 2020

Accepted August 4, 2020



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 © 2020, Korean Tunnelling and Underground
Space Association

Abstract

Underground concrete structures are usually designed to be used for decades, but in recent years, many of them are nearing their original life expectancy. As a result, it is necessary to promptly inspect and repair the structure, since it can cause lost of fundamental functions and bring unexpected problems. Therefore, personnel-based inspections and repairs have been underway for maintenance of underground structures, but nowadays, objective inspection technologies have been actively developed through the fusion of deep learning and image process. In particular, various researches have been conducted on developing a concrete crack detection algorithm based on supervised learning. Most of these studies requires a large amount of image data, especially, label images. In order to secure those images, it takes a lot of time and labor in reality. To resolve this problem, we introduce a method to increase the accuracy of crack area detection, improved by 0.25% on average by applying adversarial learning in this paper. The adversarial learning consists of a segmentation neural network and a discriminator neural network, and it is an algorithm that improves recognition performance by generating a virtual label image in a competitive

structure. In this study, an efficient deep neural network learning method was proposed using this method, and it is expected to be used for accurate crack detection in the future.

Keywords: Adversarial learning, Crack detection, Semantic segmentation, Health monitoring, Image processing

초 록

통상적으로 콘크리트 지하 구조물은 수십 년 이상 사용할 수 있도록 설계되지만 최근 들어 구조물 중 상당수가 당초의 기대 수명에 근접하고 있는 실정이다. 그 결과 구조물 고유의 기능이 상실되고 다양한 문제가 야기될 수 있어 신속한 점검과 보수가 요구되고 있다. 이를 위해 지금까지는 지하 구조물 유지관리를 위하여 인력 기반의 점검과 보수가 진행되었으나 최근에는 인공지능과 영상 기술의 융합을 통한 객관적인 점검 기술 개발이 활발하게 이루어지고 있다. 특히 딥러닝을 활용한 영상 인식 기술을 적용하여 지도학습 기반의 콘크리트 균열 탐지 알고리즘 개발에 관한 연구가 다양하게 진행되고 있다. 이러한 연구들은 대부분 지도학습 형태 영상 인식 기술로 많은 양의 데이터를 바탕으로 개발이 되는데, 그중에도 많은 수의 라벨 영상(Label image)이 요구된다. 이를 확보하기 위해서는 현실적으로 많은 시간과 노동력이 필요한 실정이다. 본 논문에서는 이와 같은 문제를 개선하고자 적대적 학습 기법을 적용하여 균열 영역 탐지 정확도를 평균적으로 0.25% 향상시키는 방법을 기술하고자 한다. 이 적대적 학습은 분할(Segmentation) 신경망과 판별자(Discriminator) 신경망으로 구성되어 있고, 가상의 라벨 영상을 경쟁적인 구조로 생성하여 인식 성능을 높이는 알고리즘이다. 본 논문에서는 이 같은 방법을 활용하여 효율적인 심층 신경망 학습 방법을 제시하였고, 향후에 정확한 균열 탐지에 활용될 것으로 기대한다.

주요어: 적대적 학습, 균열 탐지, 의미론적 분할, 상태 점검, 영상처리

1. 서 론

1.1 지하구조물 유지관리의 필요성

소규모 지반 콘크리트 구조물은 도로 터널, 철도 터널, 공동구 터널 등 다양한 형태로 구성된다. 이러한 일련의 콘크리트 구조물은 대부분 수십 년 이상을 사용할 수 있도록 설계된다. 하지만 최근 들어서 이러한 구조물 중 기대 수명에 임박한 구조물의 수가 증가하면서 노후화에 따른 콘크리트 손상이 지속적으로 발생하고 있다. 이 문제는 비단 국내에 국한된 문제는 아니며, 전 세계적으로 터널과 같은 지하 구조물의 노후화가 진행되어 가고 있다. 그 중에서도 미국의 경우를 살펴보면, 대부분의 도로와 철도 터널이 건설된 지 50년 이상이 되었다(FTA, 1997). 일본에도 이와 비슷한 실정으로 ICT 기술을 활용하여 유지보수를 하며 구조물의 기대 수명을 늘리기 위한 기술을 개발하고 있고 우리나라도 2029년이 되면 30년 이상 된 구조물의 비중이 33.7%까지 올라가게 되어 기존 구조물의 안전성에 대한 관심도가 높아지고 있다(Park and Lee, 2016). 아울러 현재의 검사 장비에 대한 불편함과 개선에 대한 호소가 높아지고 있는 가운데, 4차 산업 기술과 접목된 기술 개발에 대한 수요가 증가하고 있다(Hong et al., 2020).

지하 콘크리트 구조물에 발생하는 손상의 종류로는 균열, 박리, 박락, 재료분리 등으로 구분되는데 이런 현상의 원인은 누수, 동결 융해, 탄산화, 철근 부식, 팽창이다(Lee et al., 2018). 이러한 다양한 손상은 구조물의 성능을 저하시켜 사용자에게 안전 문제를 일으키는 요인으로 작용할 수 있어 신속한 점검과 보수가 요구된다. 이에 시설 안전 관리 주체는 정기 점검과 일상 점검으로 그 점검 항목을 구분하여 체계적으로 수행하고 있다(Kim et al., 2015). 하지만 이 같은 점검은 그 측정 도구가 균열자와 균열 현미경을 사용하는 육안점검이 주를 이룬다. 따라서 점검자의 주관적인 판단과 경험에 의존하여 손상 유무를 판단하는 것이므로 객관성과 신뢰성이 다소 떨어진다는 단점이 있다. 그러므로 객관적인 평가를 위하여 고해상도 영상 센서와 같은 장비를 이용해 정량화와 수치화를 수행하려는 연구가 진행되고 있다(Dawood et al., 2017).

1.2 지하구조물 상태 점검을 위한 선행 연구

콘크리트에서 발생할 수 있는 손상 중 균열을 탐지하기 위한 연구는 다른 어떤 연구 분야보다도 가장 활발하게 오랜 시간 동안 진행되어왔다. 콘크리트 구조물에서 허용되는 균열의 폭은 1 mm 미만으로 그 중에서도 0.3 mm의 균열을 집중적으로 검출하려고 한다(Kim and Cho, 2018). 이 같은 이유로 고해상도와 고분해능을 갖는 영상 센서를 활용한 연구가 활발히 진행되어왔다. 전통적인 연구들은 각종 필터를 사용하거나 균열의 고유 패턴을 사용하는 영상 처리 방식들이 대다수였다. Yu et al. (2007)이 제안한 방식은 히스토그램, 엣지 변환 등과 같은 방식을 사용하여 최종적으로 Neighbor region linking을 통해 균열 여부를 판단한다.

다음으로는 각종 필터 기반으로 만들어진 특징에 기계 학습 기법을 적용하는 방법들이 소개되었다. 이러한 알고리즘의 특징은 우선 입력 영상에 균열 정보를 가장 잘 드러나게 할 수 있는 특징 추출 알고리즘을 개발한다는 점에 있다. 그리고 이를 바탕으로 하여 최종적으로 균열 여부를 판단할 수 있는 기계 학습 모델을 분류 알고리즘으로 적용한다(O'Byrne et al., 2014; Zhang et al., 2014). 이들은 서로 다른 방식으로 균열의 특징을 추출하지만, Support vector machine은 기계 학습 모델을 공통적으로 사용하였다.

최근 들어서는 딥러닝을 활용한 영상 처리 기술로 콘크리트 표면의 손상을 탐지하는 연구가 활발하게 진행 중이다. Feng et al. (2020)은 Encoder와 Decoder 단계로 구성된 Auto-Encoder 방식을 바탕으로 하여 심층 신경망을 활용하여 콘크리트 댐 표면에 발생하는 균열을 탐지하는 알고리즘을 제안하였다. 그는 드론에 20 M급 카메라를 장착하여 18시간 동안 촬영한 영상을 활용하여 데이터를 구성하였다. 이를 바탕으로 심층 신경망을 설계하였고 Skeleton 영상까지 생성하여 균열의 길이를 측정하였다. 그 결과 80.45%의 Recall, 80.31%의 Precision, 79.16%의 F-measure를 갖는 알고리즘을 완성하였다. 다음으로 콘크리트 벽면에서 발생하는 균열을 비롯한 다른 손상을 탐지하기 위해 분류(Classification) 방식의 알고리즘도 있다(Cha et al., 2017). 그는 고해상도의 DSLR 카메라를 통해 촬영한 영상으로 데이터를 수집하였고 이를 256 × 256의 크기로 분할하여 학습용 DB를 확보하였다. 이것을 분류를 위한 심층 신경망 개발에 활용하였고, 90% 이상의 정확도를 얻었다. 이와 유사한 연구로 Kim and Cho (2018)가 제안한 방법이 있다. 그들은 인터넷 영상 자동 탐색 및 수집 프로그램(ScrapeBox)을 통해 인터넷 상에

서 균열, 조인트 등의 영상을 수집하였다. 그리고 이를 활용하여 심층 신경망 기반의 분류 알고리즘을 개발하였다. 그리고 이를 바탕으로 하여 고해상도 영상에서 Sliding Window를 띄워 영역 내에서 손상을 판단할 수 있도록 하였다. 그 다음 연구로 그는 Mask R-CNN 알고리즘을 활용하여 영상 내에서 균열의 영역까지 화소 단위로 분할 (Segmentation) 할 수 있는 기술을 개발하였다(He et al., 2017; Kim and Cho, 2019). 그 결과 IoU 50%를 기준으로 95%의 정확도를 보여주는 알고리즘을 최종적으로 제안하였다.

지금까지 소개된 인공지능 방식은 모두 지도학습의 형태를 이루고 있다. 이러한 지도학습은 다량의 데이터를 사용할 때 그 우수한 성능을 보장한다. 하지만 수많은 데이터를 확보하는 것은 현실적으로 많은 어려움이 있고 아울러 지도학습을 위한 라벨 영상 데이터를 확보하는 것조차 많은 시간이 소요된다. 이러한 단점을 보완하기 위해서 준지도 학습(Semi-supervised learning)을 바탕으로 한 연구가 진행되고 있다(Hung et al., 2018). 이는 기존의 Ground truth 데이터를 확보하기 위해서 수행했던 Labeling작업을 하지 않고도 균열 탐지 알고리즘의 인식 성능을 높이는 방법이다. 본 논문에서는 임의의 입력 영상을 대상으로 Confidence image를 생성하여 이를 다시 균열 탐지 알고리즘의 학습에 사용하는 적대적 학습(Adversarial learning)을 적용하여 인식 성능을 향상하는 방법을 적용하였다.

본 논문은 우선 콘크리트 균열 탐지를 위한 데이터 확보 방법과 구성에 대하여 설명하고 다음으로 판별자 신경망을 포함한 적대적 학습 구조에 대하여 기술한다. 마지막으로 기존의 지도 학습 방법과 적대적 학습 방법을 비교하여 학습 결과와 성능에 대하여 평가하는 방법을 서술한다.

2. 심층 신경망 학습을 위한 콘크리트 균열 영상

2.1 콘크리트 균열 영상의 확보

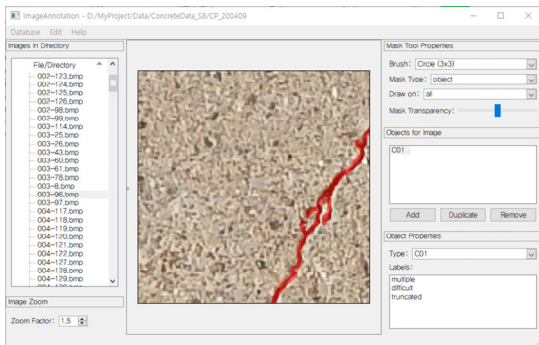
콘크리트 균열을 탐지하는 심층 신경망 알고리즘을 개발하기 위해서는 우선 영상 데이터가 필요하다. 균열이 있는 영상 데이터를 확보하기 위해서 기본적으로 촬영을 해야 하지만, 많은 시간과 인력이 필요한 관계로 선행 연구에서 이용한 데이터를 차용한다. 이들 대표적인 영상 데이터 세트는 Utah State University (USU)에서 제공하는 영상으로 16 mega-pixel의 니콘 카메라로 촬영한 데이터를 공개하였다(Dorafshan et al., 2018). 54개의 교량, 72개의 콘크리트 벽, 그리고 104개의 포장 면에서 발생할 수 있는 균열을 근접 촬영한 영상들이다. 공개된 영상은 크기는 256 × 256으로 되어 있고 JPEG 파일 포맷의 형태로 되어 있다. 이 영상 데이터에서 대표적인 영상을 선정하여 살펴보면 Fig. 1과 같다. USU에서 촬영한 영상은 교량을 비롯한 포장에서 생길 수 있는 균열을 포함하고 있어 다양성을 확보하였다고 볼 수 있다. 따라서 본 논문에서는 데이터의 다양성을 확보하고자 USU에서 제공하는 영상을 사용하고자 한다.



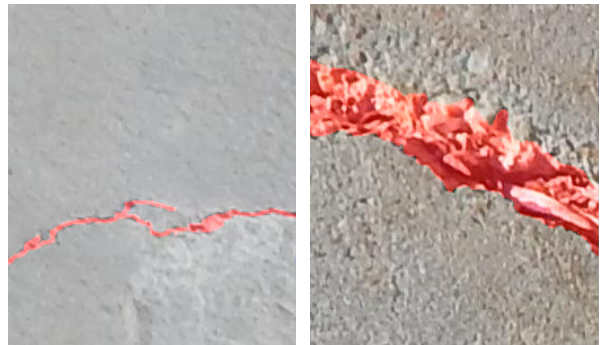
Fig. 1. Concrete crack images from USU

2.2 균열 영상의 데이터 세트 확보

화소 단위로 균열을 분할할 수 심층 신경망 알고리즘을 개발하기 위해서는 기존과 다른 데이터 세트가 필요하다. Dorafshan et al. (2018)이 제공하는 영상 데이터 세트는 분류(Classification) 알고리즘을 개발하기 위한 것이다. 하지만 본 논문에서는 분할 알고리즘을 개발하기 위한 것으로 새로운 데이터의 생성이 필요하다. 이를 위해서 본 논문의 저자들은 Marszalek and Schmid (2007)가 개발한 LEAR Image Annotation Tool을 활용하여 Fig. 2(a)와 같이 Labeling 작업을 수행했다. Labeling 작업을 수행한 영상의 수는 약 600여 장으로 그 구성은 교량 및 콘크리트 벽면을 촬영한 영상이 되겠다. 이 작업을 통해 영상 내에 균열 영역이 표시된 약 600여 장을 라벨 영상을 확보하였고, 그 결과는 Fig. 2(b)와 같다.



(a) Crack image label program



(b) Label images

Fig. 2. Image data annotation for segmentation

2.3 학습 성능 비교를 위한 데이터 세트의 구분

기존의 지도학습과 적대적 학습 결과를 비교하기 위해 본 논문에서는 K-fold cross validation 방식을 사용한다 (Kohavi, 1995). 이는 데이터 분포에 따른 성능 차이 여부를 확인하기 위한 실험이다. 이를 위해 영상 데이터 세트

를 Fig. 3과 같이 구성한다. 이 때 K를 4로 설정하여 전체 데이터를 4등분하고 각각을 G1, G2, G3, G4로 구분한다. K가 1일인 경우 전체 데이터의 75%인 G1, G2, G3을 학습용(Train) 데이터로 사용하고 나머지 25%인 G4를 검증용(Val) 데이터로 사용한다. 학습용 데이터를 절반으로 나눠서 균열 영상(Crack image)과 라벨 영상(Label image)을 모두 활용하여 지도학습(Supervised learning)을 수행한다. 그리고 학습용 데이터의 나머지 절반에 해당하는 균열 영상만을 추가로 사용하여 적대적 학습(Adversarial learning)을 수행한다. 그리고 이렇게 얻은 학습 결과를 동일한 검증용 데이터에 적용해 평가(Evaluation) 결과를 분석하도록 한다. 이는 학습용 데이터의 절반만을 사용하여 얻은 지도 학습의 결과와 라벨 영상이 없는 균열 영상을 추가적으로 사용하여 얻은 적대적 학습의 결과를 비교하기 위한 것이다. 끝으로 이 같은 실험을 K가 2, 3, 4일 때도 동일하게 적용하여 인식 성능을 확인한다.

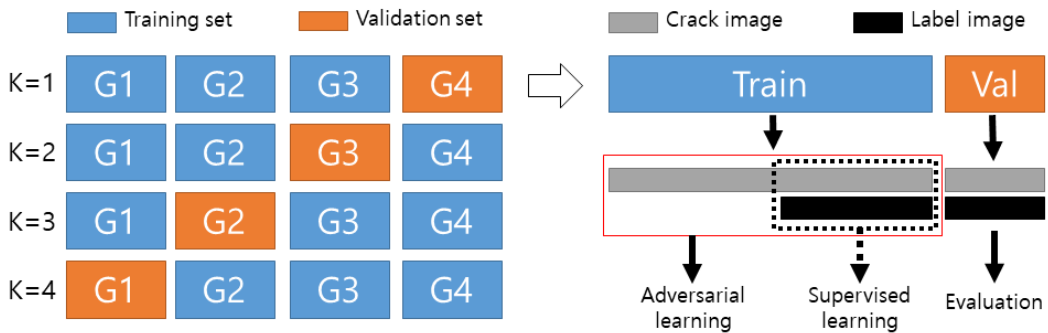


Fig. 3. Data set configuration for K-fold cross validation method

3. 균열 탐지를 위한 적대적 학습

3.1 적대적 학습을 위한 신경망의 구조

적대적 학습은 Fig. 4와 같이 두 가지의 신경망으로 이루어진다. 하나는 입력 영상이 들어왔을 때 균열의 위치를 정확히 추출하기 위한 분할 신경망(Segmentation network)이다. 두 번째는 라벨 영상의 진위 여부를 판단하는 판별자 신경망(Discriminator network)이다. 분할 신경망은 자기 부호화 형태를 차용하는데, 본 논문에서는 구체적으로 FRRN (Full-Resolution Residual Networks)을 사용한다(Pohlen et al., 2017). 이 신경망의 역할은 입력 영상에서 균열의 영역을 화소 단위로 분할하는 역할을 수행한다. 다음으로 판별자 신경망은 Convolution 연산을 사용하고 활성화 함수로는 Leaky Relu 함수를 사용하는 것이 특징이다(Xu et al., 2015). Convolution 연산이 수행될 때, 총 5단계로 진행되고 Stride는 모두 2로 설정하여 매 연산마다 크기가 절반으로 줄어들어 최종적으로 8×8 이 된다. 이를 원래의 크기인 256×256 으로 복원하기 위하여 Up-sample 함수를 사용하고 이렇게 도출된 최종 결과를 Confidence image라 칭한다. 이를 통해 입력 영상과 동일 크기의 가상 라벨 영상을 활용하여 준지도 학습을 기반으로 한 분할 신경망 학습을 수행한다.

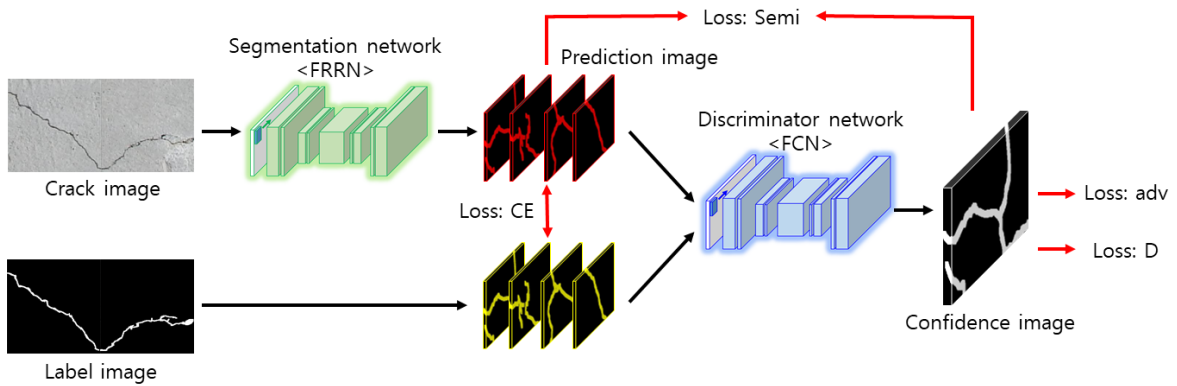


Fig. 4. Configuration of the adversarial learning

3.2 적대적 학습의 방법

적대적 학습을 진행하기 위해서는 3종류의 학습을 동시에 수행해야 한다. 그리고 학습용 데이터는 균열 영상과 라벨 영상이 짝을 이루고 있는 지도학습용 데이터 세트와 라벨 영상이 없는 준지도 학습용 데이터 세트가 필요하다. 우선적으로 판별자 신경망을 학습한다. 판별자 신경망은 지도학습 방식이 적용되고 입력은 예측 영상(Prediction image)과 라벨 영상이다. 예측 영상은 지도학습용 데이터 세트의 균열 영상과 준지도학습용 데이터 세트의 균열 영상을 분할 신경망에 입력으로 넣었을 때 얻게 되는 결과 영상이다. 이 두 개의 영상 세트를 판별자 신경망의 입력으로 학습을 수행하는데, 라벨 영상의 결과에 대해서는 True로 설정하고 예측 영상의 결과에 대해서는 False로 설정한다. 하지만 이 같은 정직한 학습으로는 새로운 영상을 생성하기에 어려움이 있으므로 속이는 학습을 한 차례 더 수행한다. 준지도학습용 데이터 세트의 균열 영상을 분할 신경망에 입력으로 넣어 얻게 된 결과 영상을 판별자 신경망의 입력으로 넣을 때, 그 결과를 True로 설정하여 적대적 학습이 되도록 구조를 설계한다. 이럴 경우 정직한 학습을 수행할 때 Loss:D 값은 0으로 수렴하게 되고 속이는 학습을 수행할 때 Loss:adv 값은 0으로 수렴하게 된다. 이는 예측 영상을 판별자 신경망의 입력으로 할 때, 한번은 가짜 라벨 영상(False)으로 간주하는 것이고, 다른 한번은 진짜 라벨 영상(True)으로 간주하는 것이다. 동일한 입력을 두고 서로 다른 방향으로 수렴하도록 학습을 수행하는 것으로 상호간의 경쟁적 학습이 된다. 다음으로 수행해야 할 것은 분할 신경망의 학습이다. 지도학습용 데이터 세트인 균열 영상과 라벨 영상을 활용하여 분할 신경망의 가중치를 업데이트한다. 끝으로 살펴볼 것은 준지도 학습이다. 이는 분할 신경망을 원래의 라벨 영상이 없는 균열 영상으로 학습하는 것이다. 앞서 경쟁적인 구조로 학습된 판별자 신경망의 출력인 Confidence image를 라벨 영상으로 간주하여 분할 신경망의 가중치를 업데이트 한다.

3.3 적대적 학습의 손실 함수

적대적 학습의 손실함수는 두 개의 심층 신경망에 대한 동시 학습을 목적으로 한다. 이들은 판별자 신경망의 손실함수와 분할 신경망의 손실함수의 합으로 구성된다. 최종 손실함수의 값은 L_{all} 이고 각각의 손실함수의 값은 L_D , L_{seg} 로 표시되며 식 (1)과 같이 정의한다.

$$L_{all} = L_D + L_{seg} \quad (1)$$

판별자 신경망을 학습하기 위해 사용한 손실함수는 식 (2)와 같다. X 는 입력 영상을 가리키고, Y 는 라벨 영상을 나타낸다. $S(\cdot)$ 는 분할 신경망의 결과로 예측 영상을 나타내고 $D(\cdot)$ 는 판별자 신경망의 결과로 Confidence image를 말한다. y 는 가중치 균형을 위한 값으로 예측 영상을 입력으로 하였을 때는 0이 되고, 라벨 영상을 입력으로 하였을 때는 1이 된다. 끝으로 i 는 영상 내에서 화소의 위치를 뜻하고 N 은 전체 화소의 수를 말한다.

$$L_D = - \sum_i^N [(1 - y_i) \log(1 - D(S(X))_i) + y_i \log(D(Y)_i)] \quad (2)$$

분할 신경망의 손실함수는 식 (3)과 같다. 이는 총 세 가지의 세부 손실함수 합으로 구성되어 있다. 첫 번째 손실함수인 L_{ce} 는 지도 학습을 수행할 때 사용하는 Cross-Entropy (CE) 함수로 식 (4)와 같다. 다음으로 L_{adv} 은 판별자 신경망의 손실함수로 예측 영상을 입력으로 하고 식 (5)와 같이 표현한다. 끝으로 L_{semi} 는 준지도 학습을 수행할 때 사용하는 손실함수로 식 (6)과 같이 정의한다. 이 식에서 T_{semi} 는 0.2로 설정하였고, I 는 Identity matrix로 0 또는 1로 구성된다. 그리고 \hat{Y} 는 분할 신경망에 의해 균열로 판단된 화소로 0 또는 1로 구성된다.

$$L_{seg} = L_{ce} + L_{adv} + L_{semi} \quad (3)$$

$$L_{ce} = - \sum_i^N Y_i \log(S(X)_i) \quad (4)$$

$$L_{adv} = - \sum_i^N \log(D(S(X))_i) \quad (5)$$

$$L_{semi} = \sum_i^N I(D(S(X))_i > T_{semi}) \hat{Y}_i \log(S(X)_i) \quad (6)$$

3.4 적대적 학습의 환경 조건

K-fold cross validation 방법으로 수행할 때는 최적화 함수는 분할 신경망은 SGD, 판별자 신경망은 ADAM을 사용하였다(Kingma and Ba, 2015; Ruder, 2016). SGD에서 사용한 매개변수로 Learning rate는 0.0002, Momentum은 0.9, Weight decay는 0.0005이고 ADAM에서 사용한 매개변수로 Learning rate는 0.001, beta-1은 0.9, beta-2는 0.999로 설정하였다. 아울러 분할 신경망에 전이학습 기법을 적용하였다. 적대적 학습을 수행하기에 앞서 분할 신경망을 학습용 데이터 세트의 50%만을 사용하여 학습을 수행하였다. 이때는 ADAM을 사용하였고 매개변수는 앞서 설명한 것과 동일하다. 총 1,500의 Epoch 동안 가장 인식 성능이 좋은 모델을 선정하여 전이학습 모델로 사용하도록 하였다. 그리고 적대적 학습을 수행할 때는 총 5,000회의 Iteration을 수행하였고, batch size는 4로 설정하였다. 알고리즘의 구현은 Windows 10 기반의 Pytorch를 사용하였고, 세 가지 실험은 모두 동일한 학습 환경에서 수행되었으며, 개발용 PC 사양은 i7-7700, 32GB RAM, NVIDIA-RTX 2080TI이다.

4. 실험 결과

4.1 성능 지표

분할 신경망의 정확도를 평가하기 위해 Long et al. (2015)이 사용한 지표를 차용한다. 그의 논문에 따르면 n_{ij} 는 Class j에 속해 있을 것으로 예측된 Class i의 모든 화소 수를 나타낸다. n_{cl} 는 Class의 개수를 가리키고, t_i 는 Class i에 속해 있는 모든 화소 수를 지칭한다. 예를 들어 0이 배경 영역이고 1이 균열 영역이라고 했을 때, n_{00} 는 라벨 영상에서 배경 영역을 예측 영상에서도 배경 영역이라고 나타내는 모든 화소의 수를 말한다. 또한 t_1 는 균열 영역에 해당하는 모든 화소의 수를 나타내는 기호다. 이 기호들의 의미를 고려하면 식 (7)과 식 (8)은 화소 단위에서의 정확도를 나타내는 지표이고 식 (9)와 식 (10)은 면적 단위에서의 정확도를 나타내는 지표를 말한다. 이 지표를 활용하여 지도학습만 사용하여 완성된 모델과 적대적 학습을 사용하여 얻은 모델 간의 정확도를 비교한다. 끝으로 식 (11)은 균열 영역만 고려하여 화소 단위의 탐지 정확도를 성능 지표로 추가한다.

$$\text{Pixel accuracy} : \sum_i n_{ii} / \sum_i t_i, \text{ where } t_i = \sum_j n_{ij} \quad (7)$$

$$\text{Mean accuracy} : (1/n_{cl}) \sum_i n_{ii} / t_i \quad (8)$$

$$\text{Mean IoU} : (1/n_{cl}) \sum_i n_{ii} / (t_i + \sum_j n_{ji} - n_{ii}) \quad (9)$$

$$\text{Frequency weighted IoU} : \left(\sum_k t_k \right)^{-1} \sum_i t_i n_{ii} / \left(t_i + \sum_j n_{ji} - n_{ii} \right) \quad (10)$$

$$\text{Crack IoU} : \sum_i n_{ii} / \left(t_i + \sum_j n_{ji} - n_{ii} \right), \text{ where } i = 1 \text{ (only crack region)} \quad (11)$$

4.2 인식 성능 비교

앞서 소개한 실험 방법과 지표를 이용하여 인식 성능을 비교한 결과는 Table 1과 같다. 데이터의 그룹에 따라 지도학습을 수행하여 얻은 인식 성능과 적대적 학습을 수행하여 얻은 인식 성능을 비교하였다. 총 4번의 실험을 수행한 결과 Mean IoU (Intersection of Union)와 Crack IoU의 지표에서는 모든 값이 지도학습보다 적대적 학습에서 높게 나온 것을 확인할 수 있었다. 하지만 다른 3가지 지표에서는 성능이 향상된 것과 반대로 감소한 것도 있다. 특히 K가 2인 경우에 Pixel accuracy, Mean accuracy, 그리고 Frequency weighted IoU 지표에서 지도학습의 결과가 적대적 학습의 결과보다 우수한 것으로 나타났다. 그럼에도 불구하고 다른 3번의 실험에서는 모두 적대적 학습의 결과가 지도학습의 결과보다 같거나 높은 것으로 나타나 준지도학습의 효과가 있는 것으로 볼 수 있다. 그리고 평균적인 값을 비교해보면 적대적 학습을 수행할 경우 지도 학습에 비하여 상대적으로 모든 지표에서 성능이 높은 것으로 나타났다. 특히 Crack IoU의 값을 살펴보면 지도 학습의 평균 결과는 71.34%이고 적대적 학습의 결과는 71.59%로 0.25% 상승한 것으로 나타났다. 비록 소량의 데이터를 활용한 비교이지만 두 개의 학습 성능에 대한 우위는 확인할 수 있다는 점에서 의의가 있다.



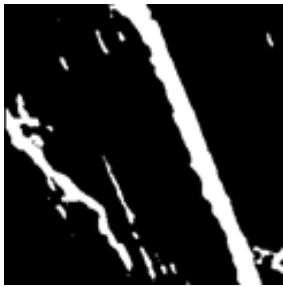
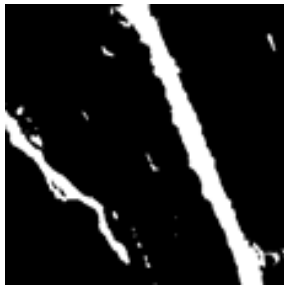



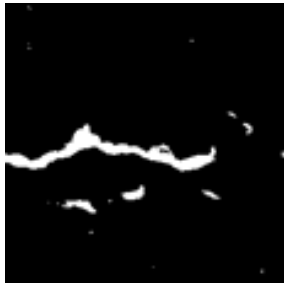




Table 1. Performance index of K-fold cross validation method (unit: %)

Learning type	K	Pixel accuracy	Mean accuracy	Mean IoU	Frequency weighted IoU	Crack IoU
Supervised learning	1	98.18	87.49	83.36	96.56	68.62
	2	98.54	88.64	84.58	97.23	70.67
	3	98.57	92.14	86.46	97.34	74.41
	4	98.45	89.48	85.02	97.08	71.66
	Avg.	98.44	89.44	84.86	97.05	71.34
Adversarial learning	1	98.19	87.42	83.41	96.57	68.70
	2	98.53	88.85	84.60	97.22	70.72
	3	98.65	91.82	86.91	97.45	75.23
	4	98.45	89.88	85.05	97.08	71.71
	Avg.	98.46	89.49	84.99	97.08	71.59

4.3 인식 결과 비교

두 학습 방식을 통해 얻은 분할 신경망을 활용하여 균열 탐지를 수행한 결과 영상은 Table 2와 같다. 첫 번째 열은 손상영역을 포함하고 있는 균열 영상이고 두 번째 열은 손상 영역을 표시한 라벨 영상이다. 세 번째와 네 번째 열은 각각 지도학습을 통해 얻은 심층 신경망의 결과 영상이고 적대적 학습을 통해서 얻은 심층 신경망의 결과 영상이다. 이들의 결과 영상들이 두 번째 열의 영상과 유사하게 나올수록 손상 탐지가 정확하게 이루어진 것으로 볼 수 있다. 이 결과 영상에서 볼 수 있듯이 지도학습의 결과 영상이 적대적 학습의 결과 영상보다 잡음 영역이 더 많은 것으로 나타나 적대적 학습을 통해서 얻은 심층 신경망의 성능이 더 나은 것으로 판단할 수 있다.

Table 2. Result images

Crack images	Label images	Supervised learning result images	Adversarial learning result images
			
			
			

5. 결론

본 논문에서는 콘크리트 지하 구조물에서 발생할 수 있는 손상 영역을 영상으로 탐지하기 위한 적대적 학습 기법의 효과에 대해서 규명하였다. 기존의 연구에서는 지도학습의 형태로 심층 신경망을 개발하여 손상영역을 탐지하고자 하였다. 하지만 이는 다수의 영상 데이터를 기반으로 하는 알고리즘으로 이를 위해 손상 영상을 확보하는 것뿐만 아니라 학습을 위한 라벨 영상을 생성하는 것에서부터 많은 시간과 노력이 요구된다. 이러한 문제를 개선하기 위해 라벨 영상을 생성하는 시간과 노력을 줄이고자 준지도 학습 기법의 한 종류인 적대적 학습을 적용하였다. 이를 통해서 새로운 손상 영상에서 Confidence image를 생성할 수 있는 방법을 확보하였고 이를 통해서 분할 신경망의 인식 성능을 향상시켰다. 다시 말해, 240장의 균열 영상과 라벨 영상을 활용하여 학습한 분할 신경망에 240장의 균열 영상만을 추가하여 정확도를 향상시켰다. 비록 심층 신경망 학습을 위한 실험으로는 소량의 영상 데이터를 사용하였지만, 결과적으로 탐지 영상에서 잡음 영역이 감소하는 것을 확인하였고, 평균적으로 0.25%만큼 Crack IoU 향상되어 균열 영역을 정확히 탐지하는 것으로 나타났다. 일반적으로 지도학습 기반의 알고리즘을 개발할 경우 인식 성능을 향상시키기 위해서 라벨 영상의 수를 늘리는 방법이 가장 보편적이다. 이는 시간과 노동력과 같은 추가 비용이 필요한 작업으로 본 논문에서 제안하는 방법으로 그만큼의 라벨 영상을 확보하는 비용을 줄이는데 기여하였다. 비록 작은 성능향상이지만 240장이라는 소량의 데이터를 고려한다면 대량의 데이터를 확보할 경우 더 큰 성능 향상을 기대할 수 있을 것으로 판단된다. 또한 콘크리트 구조물의 상태를 진단하는 경우 미소 균열에 의해서 상태 등급이 크게 좌우되기 때문에 이 같은 탐지 정확도 향상은 사실상 매우 중요하다. 이는 화소 하나의 영역에 해당하는 균열의 존재 여부에 따라서 구조물의 보강 및 보수 여부가 구별되기 때문이다. 이러한 점을 감안한다면 본 논문에서 제안한 영상 처리를 통한 잡음 영역의 감소와 균열 탐지 인식 성능은 구조물 진단의 정확성을 향상시키는 성과로 볼 수 있다.

본 논문에서는 적대적 학습 기법을 콘크리트 지하 구조물의 상태 점검을 위한 영상 처리 알고리즘 개발에 활용하였다. 또한, 심층 신경망의 학습을 위해 여러 종류의 손실함수를 사용하였다. 이 함수들 간의 가중치 균형 (Weight balance)이나 다른 형태의 손실 함수를 통해 인식 성능은 차이를 보일 수 있다. 동일한 데이터나 학습 환경이라도 가중치를 업데이트하는 방법에 따라 정확도가 향상될 수 있다. 이와 같은 방식으로 Arjovsky et al. (2017)은 기존의 Deep Convolutional Generative Adversarial Network에 새로운 손실 함수를 제안하여 성능을 개선한 바가 있다(Radford et al., 2015). 따라서 앞으로 콘크리트 균열의 정확한 탐지를 위하여 심층 신경망의 가중치를 업데이트하는 새로운 손실 함수에 대한 연구가 필요할 것으로 전망된다. 또한 본 논문에서 개발한 내용은 교량 또는 도로 등의 균열을 바탕으로 개발한 것으로 향후 지하 공간에 적용하기 위해서 많은 제약이 있을 것으로 예상된다. 특히 조도와 그림자와 같은 촬영 조건으로 인해 발생할 수 있는 문제에 대해서 향후 연구를 통해 해결하고자 한다.

감사의 글

본 논문은 국토교통부(국토교통과학기술진흥원) 국토교통기술촉진연구사업의 ‘자율형 무인 이동식 로봇암을 이용한 딥러닝 기반의 소단면 공동구 내부 구조물 결함 탐지 자동화 장비 핵심 기술 개발(20CTAP-C157234-01)’ 과제를 통해 수행되었습니다. 연구지원에 감사드립니다.

저자 기여도

심승보는 연구 개념 및 설계, 원고 작성을 하였고, 최상일과 공석민은 데이터 수집 및 분석을 하였으며, 이성원은 논문 지도 및 최종 검토를 수행하였다.

References

1. Arjovsky, M., Chintala, S., Bottou, L. (2017), “Wasserstein generative adversarial networks”, Proceedings of the 34th International Conference on Machine Learning, Sydney, pp. 214-223.
2. Cha, Y.J., Choi, W., Büyüköztürk, O. (2017), “Deep learning-based crack damage detection using convolutional neural networks”, Computer-Aided Civil and Infrastructure Engineering, Vol. 32, No. 5, pp. 361-378.
3. Dawood, T., Zhu, Z., Zayed, T. (2017), “Machine vision-based model for spalling detection and quantification in subway networks”, Automation in Construction, Vol. 81, pp. 149-160.
4. Dorafshan, S., Thomas, R.J., Maguire, M. (2018), “SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks”, Data in Brief, Vol. 21, pp. 1664-1668.
5. Feng, C., Zhang, H., Wang, H., Wang, S., Li, Y. (2020), “Automatic pixel-level crack detection on dam surface using deep convolutional network”, Sensors, Vol. 20, No. 7, pp. 2069.
6. FTA (Federal Transit Administration) (1997), “Inspection Policy and Procedures for Rail Transit Tunnels and Underground Structures”, TCRP Synthesis 23, Transportation Research Board National Research Council. [Online]. Available: <http://onlinepubs.trb.org/onlinepubs/tcrp/tsyn23.pdf>
7. He, K., Gkioxari, G., Dollár, P., Girshick, R. (2017), “Mask R-CNN”, Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, pp. 2961-2969.
8. Hong, S.H., Kim, J.G., Cho, J.Y., Kim, T.H. (2020), “A study on the necessity of verification and certification system of inspection and diagnostic equipment for infrastructure using advanced technologies”, Journal of the Society of Disaster Information, Vol. 16, No. 1, pp. 163-177.
9. Hung, W.C., Tsai, Y.H., Liou, Y.T., Lin, Y.Y., Yang, M.H. (2018), “Adversarial learning for semi-supervised semantic segmentation”, arXiv:1802.07934. [Online]. Available: <https://arxiv.org/abs/1802.07934>
10. Kim, B., Cho, S. (2018), “Automated vision-based detection of cracks on concrete surfaces using a deep learning technique”, Sensors, Vol. 18, No. 10, pp. 3452.

11. Kim, B., Cho, S. (2019), “Image-based concrete crack assessment using mask and region-based convolutional neural network”, *Structural Control and Health Monitoring*, Vol. 26, No. 8, pp. e2381.
12. Kim, M.K., Sohn, H., Chang, C.C. (2015), “Localization and quantification of concrete spalling defects using terrestrial laser scanning”, *Journal of Computing in Civil Engineering*, Vol. 29, No. 6, pp. 04014086.
13. Kingma, D.P., Ba, J. (2015), “ADAM: a method for stochastic optimization”, *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, San Diego, pp. 1-15.
14. Kohavi, R. (1995), “A study of cross-validation and bootstrap for accuracy estimation and model selection”, *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montreal, Vol. 14, No. 2, pp. 1137-1145.
15. Lee, Y.I., Kim, B., Cho, S. (2018), “Image-based spalling detection of concrete structure using deep learning”, *Journal of the Korea Concrete Institute*, Vol. 30, No. 1, pp. 91-99.
16. Long, J., Shelhamer, E., Darrell, T. (2015), “Fully convolutional networks for semantic segmentation”, *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, pp. 3431-3440.
17. Marszalek, M., Schmid, C. (2007), “Accurate object localization with shape masks”, *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, pp. 1-8.
18. O’Byrne, M., Ghosh, B., Schoefs, F., Pakrashi, V. (2014), “Regionally enhanced multiphase segmentation technique for damaged surfaces”, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 29, No. 9, pp. 644-658.
19. Park, C.H., Lee, H.I. (2016), *Future trend of capital investment for Korean transportation infrastructure*, Construction and Economy Research Institute of Korea, pp. 6-8.
20. Pohlen, T., Hermans, A., Mathias, M., Leibe, B. (2017), “Full-resolution residual networks for semantic segmentation in street scenes”, *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, pp. 4151-4160.
21. Radford, A., Metz, L., Chintala, S. (2015), “Unsupervised representation learning with deep convolutional generative adversarial networks”, arXiv:1511.06434, [Online], Available: <https://arxiv.org/abs/1511.06434>
22. Ruder, S. (2016), “An overview of gradient descent optimization algorithms”, arXiv:1609.04747. [Online]. Available: <https://arxiv.org/abs/1609.04747>
23. ScrapeBox—The Swiss Army Knife of SEO! Available online: <http://www.scrapebox.com/> (accessed on 29 June 2020).
24. Xu, B., Wang, N., Chen, T., Li, M. (2015), “Empirical evaluation of rectified activations in convolutional network”, arXiv:1505.00853, [Online], Available: <https://arxiv.org/abs/1505.00853>
25. Yu, S.N., Jang, J.H., Han, C.S. (2007), “Auto inspection system using a mobile robot for detecting concrete cracks in a tunnel”, *Automation in Construction*, Vol. 16, No. 3, pp. 255-261.
26. Zhang, W., Zhang, Z., Qi, D., Liu, Y. (2014), “Automatic crack detection and classification method for subway tunnel safety monitoring”, *Sensors*, Vol. 14, No. 10, pp. 19307-19328.