

도로 노면 파손 탐지를 위한 배경 객체 인식 기반의 지도 학습을 활용한 성능 향상 알고리즘

Performance Enhancement Algorithm using Supervised Learning based on Background Object Detection for Road Surface Damage Detection

심 승 보* · 전 찬 준** · 류 승 기***

* 주저자 : 한국건설기술연구원 전임연구원

** 공저자 : 한국건설기술연구원 전임연구원

*** 교신저자 : 한국건설기술연구원 선임연구위원

Seungbo Shim* · Chanjun Chun* · Seung-Ki Ryu*

* Korea Institute of Civil Engineering and Building Technology

† Corresponding author : Seung-Ki Ryu, skryu@kict.re.kr

Vol.18 No.3(2019)

June, 2019

pp.95~105

pISSN 1738-0774

eISSN 2384-1729

<https://doi.org/10.12815/kits.2019.18.3.95>

2019.18.3.95

요 약

최근 들어 도로 노면 파손의 위치 정보를 수집하기 위한 영상 처리 기술에 대한 연구가 활발히 진행되고 있다. 대표적으로 차량에 탑재가 가능한 스마트폰이나 블랙박스를 통해 영상을 얻고 이를 영상처리 알고리즘을 사용하여 인식하는 기술이 주로 사용된다. GPS 모듈과 연계하여 실제 파손 위치를 파악할 때 가장 중요한 기술은 영상 처리 알고리즘인데, 근래에는 대부분 인공지능을 통한 알고리즘이 연구 주제로 주목받고 있다. 이와 같은 맥락에서 본 연구에서도 영역 기반의 합성곱 방식 계열의 객체인식 (Object Detection) 방법을 사용한 인공지능 영상 처리 알고리즘에 대하여 논의하고자 한다. 도로 노면 파손 객체 인식 성능을 향상시키기 위하여 도로 노면 파손 영상 600여 장과 일반적인 도로 주행 영상 1500여 장으로 학습 데이터베이스를 구성하였다. 또한 배경 객체 인식 방법을 적용한 지도 학습을 수행하여 도로 노면 파손의 오탐을 감소시켰다. 그 결과 동일한 테스트용 데이터베이스를 통해 알고리즘의 인식 성능을 mAP 평균값 기준 9.44%만큼 향상시킨 새로운 방법을 소개하고자 한다.

핵심어 : 도로 노면 파손, 심층 신경망, 유지보수, 영역 기반 영상 분류, 배경 객체 인식

ABSTRACT

In recent years, image processing techniques for detecting road surface damaged spot have been actively researched. Especially, it is mainly used to acquire images through a smart phone or a black box that can be mounted in a vehicle and recognize the road surface damaged region in the image using several algorithms. In addition, in conjunction with the GPS module, the exact damaged location can be obtained. The most important technology is image processing algorithm. Recently, algorithms based on artificial intelligence have been attracting attention as research topics. In this paper, we will also discuss artificial intelligence image processing algorithms. Among them, an object detection method based on a region-based convolution neural networks method is used. To improve the recognition performance of road surface damage objects, 600 road surface damaged images and 1500 general road driving images are added to the learning database. Also, supervised learning using background object recognition method is performed to reduce false alarm and missing rate in road surface damage detection. As a result, we introduce a new method that improves the recognition

Received 8 April 2019

Revised 26 April 2019

Accepted 29 April 2019

© 2019, The Korea Institute of Intelligent Transport Systems. All rights reserved.

performance of the algorithm to 8.66% based on average value of mAP through the same test database.

Key words : Road surface damage, Deep neural network, Road maintenance, Region based convolutional neural networks, Background object detection

I. 서 론

1. 연구배경

아스팔트 포장 도로가 매년 증가함과 동시에 노후화된 포장도로도 지속적으로 증가하고 있다. 이렇게 도로 노면 파손으로 인한 교통사고는 증가 추세에 있다. 시간의 경과에 따라 노후화된 도로는 꾸준히 증가될 것으로 예상되지만, 이를 수리 보수할 인구는 점차 감소될 것으로 전망된다. 따라서 도로의 포장 품질 역시 지속적으로 개선되기 어려워지는 실정이다.

이 같은 이유로 포장된 도로의 손상 및 파손은 적절한 보수 시기를 놓쳐 안전사고를 야기시킨다. 그 중 가장 심각한 문제는 도로 포장 파손 영역이 운전자의 주행에 영향을 미쳐 교통사고를 발생시킨다는 점인데, 이는 인명 피해로 연결되어 더욱 악화되는 양상으로 나타난다. 정부에 보고된 바에 따르면 약 5년간 4,000여건이 넘는 사고가 발생하였는데, 이는 기존 대비 5배가 증가한 수치이다. 또한 서울시 도로 파손 (포트홀) 건수는 25,600건으로 이전 년도에 비해 4배 이상 급증하는 실정이 되었다. 이와 같은 수치는 우리나라 도로의 노후화가 급속도로 증가 추세에 있음을 의미한다. 이런 상황을 개선하기 위해서는 도로 포장 관리를 효율적으로 수행하기 위한 유지관리 자동화 핵심기술이 필수적이라 할 수 있다.

여러 핵심 기술 가운데서도 가장 중요한 것은 도로 노면의 파손 위치를 정확하게 수집하는 기술이라고 볼 수 있다. 이 중에서도 가장 주목받는 기술은 영상 처리를 활용한 기술이다. 고전적인 영상 처리 방식과 차별화된 Deep Learning을 이용한 방식이 세계적으로 그 성능을 인정받기 시작하였고, 이는 여러 분야에서 활발하게 적용되고 있다. 이러한 시대적 흐름에 따라 도로 노면 유지 관리 분야에서도 이 방식은 적용되고 있으며, 이에 대한 연구도 활발하게 이루어지고 있다.

본 연구에서는 Deep Learning 기반의 인공지능을 활용한 연구에 대하여 논의하고자 하며, 그와 관련된 핵심 내용은 아래와 같다. 첫째, 도로 노면 파손을 탐지하기 위한 기존의 영상 처리 연구에 대하여 기술한다. 둘째, 본 연구에서 제안하는 배경 객체 인식을 위한 학습용 데이터베이스를 확보하는 방식에 대하여 설명한다. 셋째, 도로 노면 객체 인식을 위한 새로운 심층 신경망과 그에 따른 학습 방법에 대하여 언급한다. 넷째, 성능 향상을 검증하기 위한 실험과 그와 관련된 분석 결과를 보여준다. 마지막으로, 결론 및 향후 연구 방향에 대하여 제시하고자 한다.

2. 관련 연구

차량에 탑재되어 도로 노면 파손을 탐지하기 위한 영상처리 기술이 실제로 적용되는 형태는 하드웨어 및 알고리즘의 형태에 따라 다양하다. 예를 들어, Jo et al.의 논문에서는 블랙박스로부터 영상을 제공받아서 포트홀을 탐지하는 시스템을 제안하였다 (Jo et al., 2015). 우선 포트홀이라고 판단되는 후보 영역을 추출한 후 여러 단계로 이루어진 필터를 거쳐 최종적으로 포트홀 영역을 선정하는 방식의 알고리즘이다. Zhang et al.은 3차원의 아스팔트 표면으로부터 포장 균열을 추출하는 심층 신경망을 제안하였는데(Zhang et al., 2017), 이는

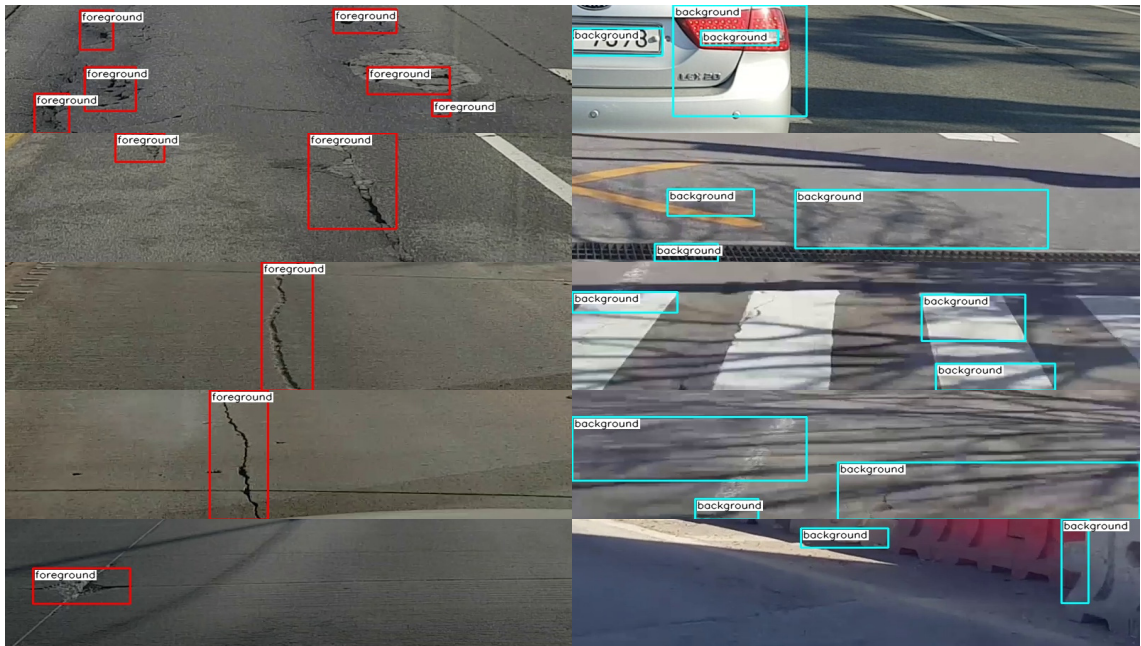
독창적인 라인 필터를 활용해 특징 지도를 생성하고 이를 심층 신경망의 입력으로 활용하는 것을 특징으로 한다. 그 결과 90.13%의 정확도와 87.63%의 재현율을 보여주었다. Maeda et al.의 연구는 스마트폰으로 촬영한 영상을 활용해 분류(Classification)와 위치 인식(Localization)을 동시에 수행하는 심층 신경망을 제안하였다 (Maeda et al., 2018). 이 연구를 수행하기 위해 약 9,000장에 가까운 영상을 수집하였고, 15,000개의 객체를 확보하였다. 또한 도로 노면의 파손 형태에 따라 가로 균열과 세로 균열로 구분하였고, 보다 구체적인 노면의 상태에 따라 위험 순위를 구분할 수 있도록 하였다. 여기서 사용한 심층 신경망은 Inception v2와 MobileNet 이고, 자체적으로 학습한 모델을 스마트폰에 탑재하여 실험까지 완료되었다. 총 8종류의 도로 노면 파손을 분류하고 그 위치까지 탐지할 수 있도록 알고리즘을 개발하였다. Yang et al.의 연구는 완전 합성곱을 활용한 심층 신경망을 활용하여 화소 레벨 단계에서 정확한 의미론적 분할을 할 수 있는 알고리즘을 제안하였다 (Yang et al., 2018). 이 알고리즘은 비대칭형 자기부호화 심층 신경망 형태로 설계되었고 800장 이상의 영상을 통해 학습되었으며, 결과적으로 97.96%의 정확도로 균열 영역을 탐지할 수 있다. Chun et al.의 연구는 오로지 합성곱 신경망과 역합성곱 신경망을 사용하여 자기부호화 형태의 심층 신경망을 제안하였다 (Chun et al., 2018). 그리고 신경망 학습을 위해 실제 도로 주행 영상을 확보하였고, 그 영상에 도로 노면 파손 영역을 표시하여 학습할 수 있는 데이터 세트를 구축하였다. 그리고 데이터 증강 작업을 바탕으로 데이터 수를 확장하여 인식 성능을 향상시킬 수 있도록 하였고, 그 결과 70.12%의 정확도를 가진 알고리즘을 제안하였다.

이와 유사한 맥락으로 본 논문에서도 도로 노면 파손 탐지를 위해 Girshick가 제안한 Fast R-CNN 기반의 인공지능 알고리즘을 사용하고자 한다 (Girshick, 2015). 이 같은 Fast R-CNN 기법을 이용하여 학습하기 위해서는 입력 영상 내에 탐지하고자 하는 전경 객체 (Foreground Object)가 있어야 한다. 그리고 전경 객체의 경계 상자를 기준으로 중첩 비율 (Intersection Of Union)에 따라 학습을 위한 전경 객체 경계 상자 후보군 (Foreground Object Bounding Box Candidates)과 배경 객체 경계 상자 후보군 (Background Object Bounding Box Candidates)으로 구성한다. 따라서 배경 객체 경계 상자 후보군은 항상 전경 객체가 있는 영상 내에서만 얻을 수 있다. 이와 같은 이유로 다양한 배경 객체 경계 상자 후보군으로 학습하기 위해서는 중첩 비율 기준을 조정하거나 전경 객체가 포함된 영상을 더 확보해야 한다. 중첩 비율 기준을 조정하는 경우는 배경 객체 경계 상자 후보군의 다양성을 확보하기는 불충분하고, 전경 객체가 포함된 영상을 확보하는 경우는 도로 노면 파손 영상을 직접 촬영하는 것으로 다수의 데이터를 확보하는데 어려움이 있다. 따라서 본 논문에서는 다수의 다양한 배경 객체 경계 상자 후보군을 확보하기 위한 새로운 기법을 제안하고자 한다.

II. 학습용 DB 확장을 위한 배경 객체 인식 기법

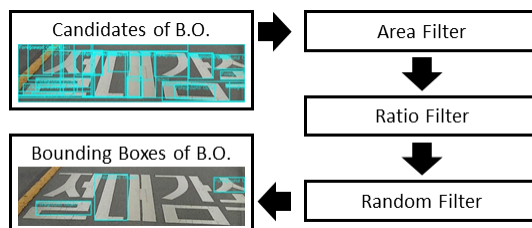
1. 배경 객체 인식 알고리즘

많은 배경 객체 경계 상자 후보군을 확보하기 위하여 우선적으로 일반 도로 주행을 통해 얻은 영상을 약 1500장 추가 확보하였다. <Fig. 1>과 같이 이 영상들은 도로 노면 파손 영역이 포함되지 않은 정상적인 도로 노면 영상이지만, 특이한 객체를 포함하고 있다. 이 영상으로 Fast R-CNN 기법에서 사용하는 Selective Search 알고리즘을 OpenCV에서 제공하는 파라미터의 초기값을 바탕으로 수행할 경우, 다수의 객체 후보군이 나오게 된다 (Uijlings et al., 2013). 이 객체 후보군들은 주로 차량, 도로 마커, 과속방지턱 등이 될 수 있다. 다음으로 이 후보군을 각종 필터에 통과시켜 <Fig. 1>과 같이 적절한 수의 배경 객체 (Background Object, B.O.)가 추출될 수 있도록 하였다.



<Fig. 1> (Left Column) Road Surface Images WITH the Damaged Region and its Bounding Boxes of the Foreground Object like Pothole, Crack, etc. (Right Column) Road Surface Images WITHOUT the Damaged Region and its Bounding Boxes of the Background Object like cars, road marker, etc.

구체적으로 배경 객체를 결정하는 절차는 <Fig. 2>와 같이 총 3단계로 구성되어 있다. 각 단계마다 필터 파라미터에 대한 임의의 가정을 사용한다. 우선적으로 Selective Search 알고리즘을 통해 얻은 배경 객체 경계 상자의 후보군은 면적 필터 (Area Filter)를 거친다. 본 논문에서는 경계 상자의 면적이 2,000화소 미만이거나 80,000화소 초과인 경우에 작거나 큰 객체로 가정한다. 따라서 예를 들어 경계 상자의 크기가 20×20인 경우에는 면적이 400화소가 되어 작은 객체로 판단하고 150×600인 경우에는 면적이 90,000화소가 되어 큰 객체로 판단하여 후보군에서 제외한다. 다음으로 비율 필터 (Ratio Filter)를 적용한다. 경계 상자의 너비가 높이의 5배보다 크거나 혹은 경계 상자의 높이가 너비의 5배보다 큰 객체를 제외한다. 끝으로 임의의 필터 (Random Filter)를 적용한다. 최종적으로 남은 후보군에서 임의로 3개를 선택하여 배경 객체를 결정하는데 그 결과는 <Fig. 2>와 같다. 결론적으로 ‘도로 마커의 일부, 차선의 일부, 과속방지턱의 일부, 차량의 일부’와 같이 도로 노면 영상에서 특징 있는 객체들이 최종 선별된다.

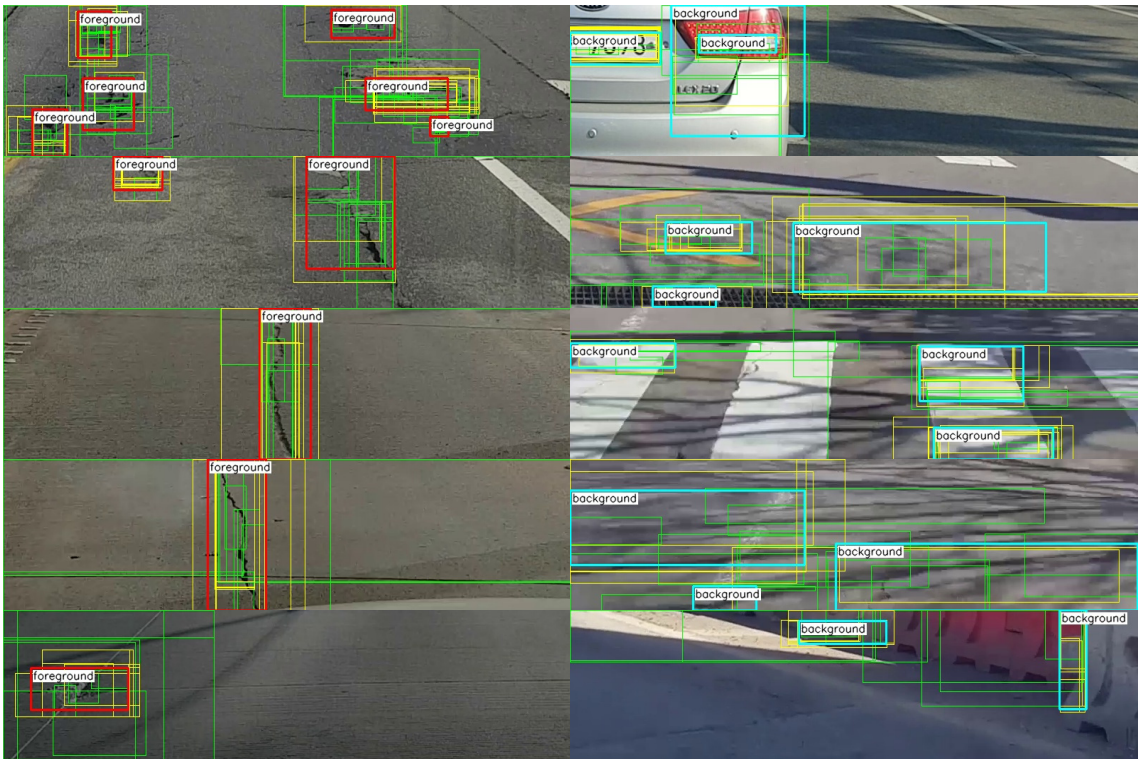


<Fig. 2> Flow Chart for Decision of the Background Object Bounding Boxes

2. 학습을 위한 영상 데이터 세트 구성

학습을 위한 영상 데이터 세트는 <Fig. 1>과 같이 도로 노면 파손 영역을 포함하고 있는 영상 약 600장과 파손 영역이 없는 영상 약 1,500장을 사용한다. Girshick가 제안한 Fast R-CNN 기법을 통해 인공지능을 학습하기 위해서는 한 장의 영상에서 다수의 후보 경계 상자가 필요하다 (Girshick, 2015). 우선 도로 노면 파손 영역을 포함하고 있는 영상의 경우에 대해서 설명해보면, 한 장의 영상에 평균적으로 2,000개의 후보 경계 상자들이 만들어진다. 이 중 학습을 위해서 사용하는 총 후보 경계 상자의 수는 Girshick가 사용한 방식과 동일하게 128개로 정하였다. 128개 중 32개는 전경 객체의 경계 상자와의 중첩 비율이 50% 이상인 후보 경계 상자들로 한다. 나머지 96개는 전경 객체의 경계 상자와의 중첩 비율이 10% 이상이고 50% 미만인 후보 경계 상자들로 선정한다.

다음은 도로 노면 파손 객체가 없고 3개의 배경 객체가 포함된 영상이다. 앞서 언급한 전경 객체가 포함되어 있는 영상에 적용했던 방법과 동일하다. 먼저 2,000개의 후보 경계 상자들 중 배경 객체의 경계 상자와의 중첩 비율이 50% 이상인 경계 상자 32개를 추출한다. 그리고 나머지 96개는 중첩 비율이 10% 이상이고 50% 미만인 것들로 구성하여 총 128개의 경계 상자를 학습에 사용한다. 이와 같은 방법으로 구성한 경계 상자들 중에 일부만 임의 선정하여 표시할 경우 <Fig. 3>와 같이 나타난다.

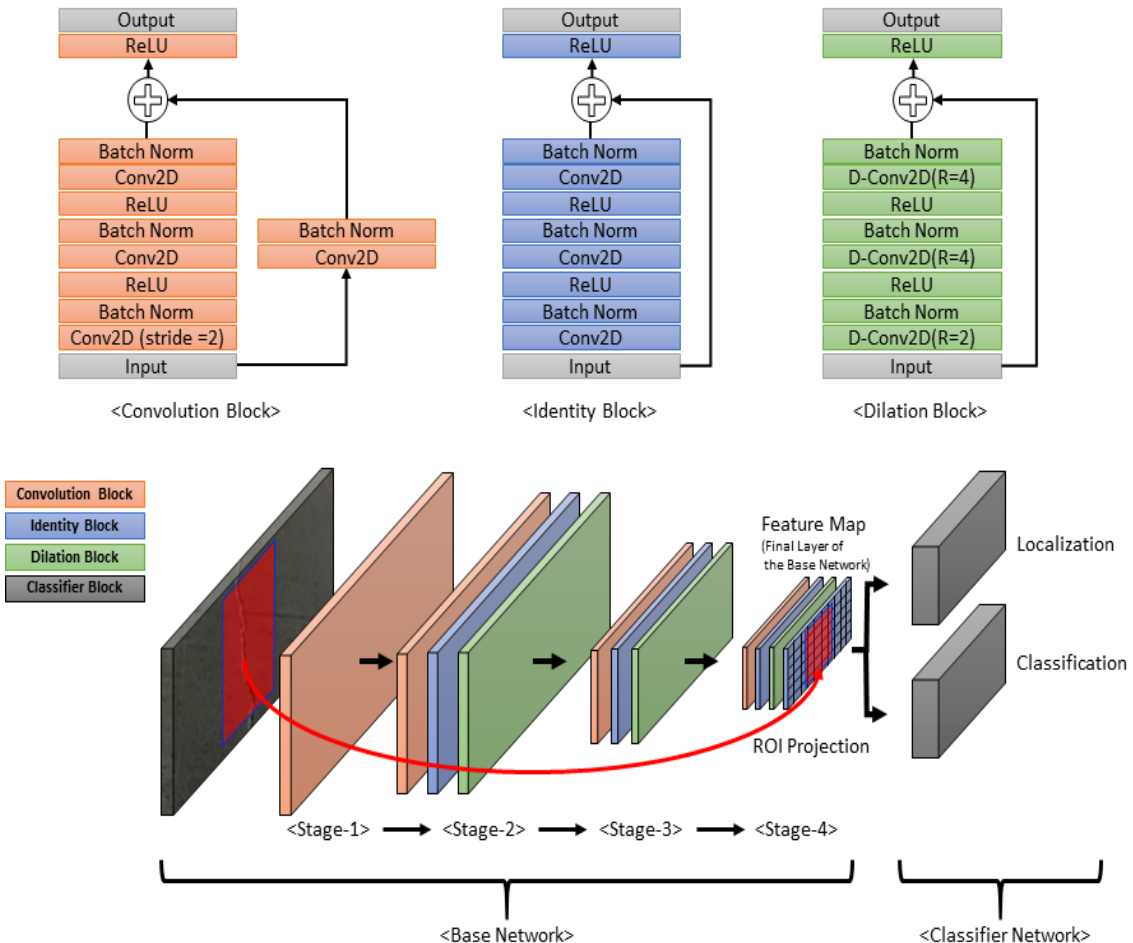


<Fig. 3> Road Surface Images with the Bounding Boxes. Red Bounding Boxes are Foreground Objects and Cyan Bounding Boxes are Background Objects. Yellow Bounding Boxes have IoU Overlap with a Red or Cyan Bounding Box of at least 50% and Green Bounding Boxes have IoU Overlap with a Red or Cyan Bounding Box between 10% and 50%.

III. 인식 성능 향상을 위한 심층 신경망 설계와 학습

1. 도로노면 파손 인식을 위한 심층 신경망 설계

Fast R-CNN 기반의 심층 신경망은 백본 신경망 (Backbone Network)과 분류 신경망 (Classifier Network)으로 구성된다 (Girshick, 2015). 본 논문에서는 알고리즘의 인식 성능을 향상시키기 위하여 새로운 심층 신경망을 <Fig. 4>와 같이 제안한다. 이 심층 신경망은 3가지 주요 블록으로 구성되는데, 우선 Convolution 블록은 모두 2차원 합성곱과 배치 정규화 (Batch Normalization) 그리고 활성화함수 (Activation Function)인 ReLU 함수를 사용한다 (Goodfellow, 2016; Ioffe et al., 2015; Nair et al., 2010). 총 3번에 걸쳐서 순차적으로 연산이 수행되는데, 마지막으로 입력에 별도의 2차원 합성곱 연산을 수행한 결과를 합하여 출력으로 내보내는 것이 특징이다. 다음으로 Identity 블록은 Convolution 블록과 동일한 구성이지만, 입력에 별도의 2차원 합성곱 연산이 없는 점이 유일한 차이점이다. 끝으로 Dilation 블록의 연산 순서는 Identity 블록과 동일하지만 2차원 합성곱 연산을 대신해 2차원 팽창형 합성곱 (Dilated Convolution) 연산을 사용했다는 점이다 (Yu et al., 2015).



<Fig. 4> Proposed Deep Neural Network Structure

3개의 신경망 블록을 활용하여 도로 노면 파손 인식을 위한 백본 신경망을 제안하였다. 이 신경망은 총 4 단계로 구성되고 각 단계마다 출력 채널수를 계속해서 증가시켰다. 첫 번째 단계에는 Convolution 블록을 사용하였고 출력 채널을 64개로 설정하였다. 두 번째 단계에서는 Convolution 블록, Identity 블록 그리고 Dilation 블록을 순서대로 연산하도록 배치하였고 출력 채널은 256개로 설정하였다. 세 번째 단계는 두 번째 단계와 구성은 동일하고 출력 채널만 512로 설정하였다. 네 번째 단계는 마지막에 Identity 블록을 추가하였고 출력 채널을 1024로 설정하였다.

마지막으로 분류 신경망은 객체 인식 (Object Detection)을 위한 영상 내 위치 인식(Localization)과 객체의 종류 분류 (Classification) 기능을 수행한다. 단계에서는 Girshick이 제안한 방식과 유사하게 ROI Pooling 연산과 여러 층의 신경망으로 구성하였다 (Girshick, 2015). ROI Pooling을 위한 신경망의 크기는 24×24 로 하였다. 이는 학습을 위해 보유한 영상에 포함된 노면 파손 영역의 평균적인 가로와 세로 비율을 1:1로 나타냈다. 이 특성을 직관적으로 반영하고자 가로와 세로의 비율을 동일하게 하였고, 학습할 때 GPU의 메모리를 고려하여 이와 같은 크기로 설정하였다.

2. 다단계(Multi-stage) 학습 방법

총 2,100여 장의 영상을 이용해 학습하기 위하여 3단계를 거쳐 학습을 수행한다. 첫 번째 학습 단계는 도로 노면 파손이 포함되어있는 약 600장의 영상만을 이용하여 학습을 수행한다. 이때 가중치의 초기값은 Xavier가 제안한 방식을 사용했고, 최적화 기법은 Adaptive Moment Estimation (Adam)을 사용하였으며 총 반복 횟수는 500회 수행하였다 (Glorot and Bengio, 2010; Kingma and Ba, 2015). 두 번째 학습 단계는 앞서 사용한 도로 노면 파손이 포함된 600여 장의 영상과 그렇지 않은 영상 600여 장을 임의 선별하여 총 1,200장을 이용해 학습을 수행하였다. 이 때 가중치의 초기값은 앞서 학습된 신경망 모델의 것을 사용하였다. 학습하는 방법은 동일하고 총 500회를 수행하였다. 그리고 마지막 단계에서 보유하고 있는 모든 영상 2,100여 장을 활용하여 동일한 학습 방법으로 1,000회 학습을 수행하였다.

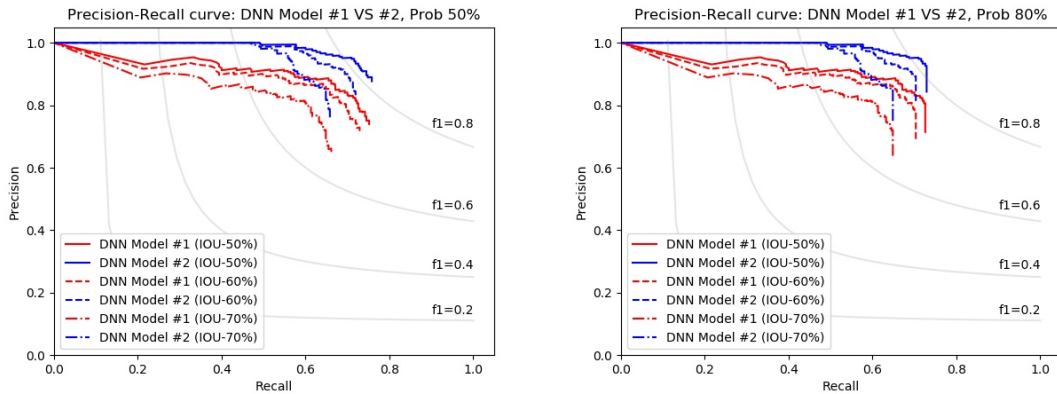
IV. 실험 결과

1. 실험 방법

본 논문에서 제안하는 방식의 성능을 분석하고자 두 개의 인공지능 모델을 비교한다. 두 개의 인공지능 모델의 공통점은 본 논문에서 제안한 심층 신경망을 사용한다는 점이다. 차이점은 첫 번째 심층신경망 모델 (DNN Model #1)은 도로노면 파손 영역이 포함된 영상 600여 장만을 이용해 학습을 수행한 것이고 두 번째 심층신경망 모델 (DNN Model #2)은 파손 영역이 없는 영상까지 포함하여 2,100여 장을 3단계를 걸쳐 학습한 인공지능 모델이다. 그리고 두 인공지능 모델을 분석하기 위해 동일한 테스트용 영상 데이터 세트를 활용한다. 총 200여 장으로 구성되어 있고, 모두 도로 노면 파손 영역을 포함하고 있다. 비교하는 방법으로는 Precision and Recall Curve (P-R Curve)와 mAP 방식을 사용한다. P-R Curve를 사용하면 재현율 대비 알고리즘의 성능 파악이 가능하다. 그리고 mAP를 사용하면 알고리즘의 전반적인 인식 성능을 정량적으로 파악하여 성능 비교가 가능하다.

2. 실험 분석

테스트용 영상 데이터 세트를 이용하여 두 알고리즘을 수행한 결과의 P-R Curve는 <Fig. 5>와 같다. 본 그래프에서는 중첩 비율(Intersection Of Union, IOU)의 비율에 따라 인식 성공 여부를 구분하였다. 즉, 영상 내에서 도로 노면 파손이 있는 참 영역과 알고리즘에 의해 인식한 영역 간의 중첩 비율 기준을 50%, 60%, 70%로 임의로 구분하여 인식 성능을 관찰하였다. 또한 객체 인식 확률(Probability)을 기준으로 50% 이상인 결과를 성공으로 보는 경우와 80% 이상인 결과를 성공으로 보는 경우로 나뉘어서 관찰하였다. 그 결과는 DNN Model #2의 성능이 전체적으로 우수한 것을 확인할 수 있었고, 중첩 비율이 낮을수록 그리고 인식 확률 기준이 낮을수록 인식 성능은 더 높게 나타났다.



<Fig. 5> Precision-Recall Curve of DNN Model #1 and DNN Model #2

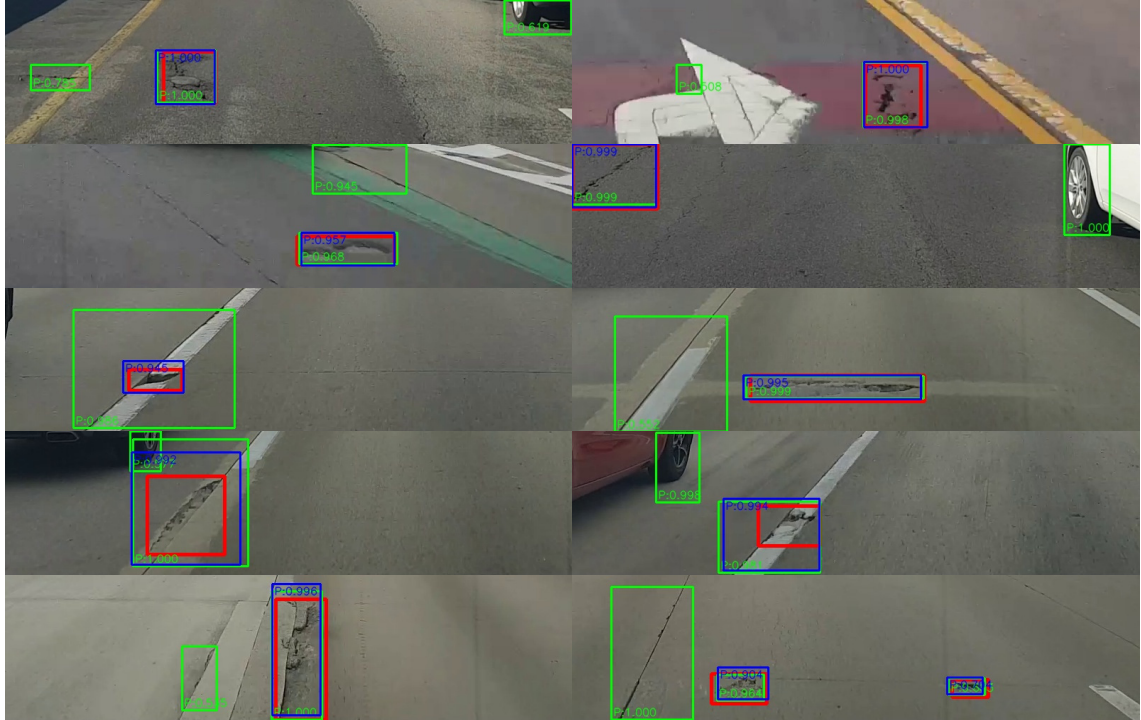
<Table 1> mAPs of the DNN Model #1 Algorithm and DNN Model #2 Algorithm

Condition of Threshold	DNN Model #1				DNN Model #2			
	50% IOU	60% IOU	70% IOU	Avg.	50% IOU	60% IOU	70% IOU	Avg.
50% Probability	88.12	86.30	81.92	85.44	96.78	95.34	92.75	94.96
80% Probability	87.73	85.92	81.76	85.14	96.10	94.93	92.49	94.51
Avg.	87.93	86.11	81.84	85.29	96.44	95.14	92.62	94.73

이와 같은 현상은 <Table 1>에서도 나타난다. 우선 알고리즘의 전반적인 성능을 나타내는 mAP의 평균 값을 비교해 볼 때, 동일하게 DNN Model #2의 성능이 다른 것에 비해 평균적으로 9.44% 높은 것으로 나타났다. 또한 중첩 비율 인식 기준 파라미터의 변화에 대한 인식 성능 변화도 더 크게 나타났다. DNN Model #1의 경우 객체 인식 성공 확률 기준을 80%에서 50%로 감소시켰을 경우 평균적으로 0.31% 증가하였다. 또 DNN Model #2의 경우는 0.45% 증가하였다. 반면에 객체 인식 성공 중첩 비율 기준을 70%에서 50%로 감소시켰을 경우 평균적으로 DNN Model #1은 6.09%, DNN Model #2는 3.82%가 증가하였다. 이를 통해 두 심층 신경망 모두 중첩 비율이 인식 성능에 더 민감하게 작용하는 것을 확인할 수 있다. 이는 두 심층 신경망이 동일하지만 학습용으로 사용하는 데이터만 다르므로 인식 성능에만 차이가 있을 뿐이고, 다른 경향에서는 유사한 특성을 가지고 있는 것으로 보인다.

다음으로 두 심층 신경망으로 테스트용 영상 데이터 세트를 수행한 결과를 살펴보면 <Fig. 6>와 같다. 붉

은색 경계 상자는 영상 내 도로 노면 파손 영역을 표시한 참 영역이다. 그리고 초록색 경계 상자와 파란색 경계 상자는 각각 DNN Model #1과 DNN Model #2의 알고리즘 수행 결과다. 배경 객체가 포함된 영상으로 학습한 DNN Model #2의 결과가 전경 객체만 포함된 영상으로 학습한 DNN Model #1보다 오탐이 비교적 적은 것으로 나타났다. 특히 차량과 도로 이음새 등과 같은 곳에서의 오탐이 감소하였고, 차선과 로드 마커를 인식하는 경우도 감소하였다.



<Fig. 6> Detection Results of Deep Neural Networks(DNN) Algorithms, Red Boxes are Ground Truth of Road Surface Damage Region, Green Boxes are Results of DNN Model #1 and Blue Boxes are Results of DNN Model #2.

V. 결 론

본 논문에서는 도로 노면 파손 영역 탐지 성능 향상을 위하여 일반 도로 주행 영상을 활용하는 방법을 제안하였다. 이를 위하여 우선 도로 주행 영상을 확보하였고, 이를 도로 노면 파손 영역이 있는 영상 데이터 세트와 그렇지 않은 일반 영상 데이터 세트로 구성하였다. 그리고 일반 영상 데이터 세트를 활용하여 도로 노면 파손 객체의 인식 성능을 향상시키는 방법을 개발하였다. 일반 영상에서 Selective Search 알고리즘으로 수행하여 다수의 배경 객체의 후보군을 생성하고 이 후보군에 각종 필터를 적용하여 3개의 배경 객체를 결정한다. 이 같은 방법으로 총 배경 객체가 포함된 1,500장의 영상을 추가하여 학습용 데이터 세트를 구성하였다. 그리고 도로 노면 파손을 인식하기 위한 새로운 심층 신경망을 제안하였고, 총 3단계에 걸쳐 다단계 학습을 진행하였다. 그리고 인식 성능 향상을 확인하기 위해서 학습 방법을 달리한 두 가지의 심층 신경망

모델을 만들어 실험하였다. 그 결과 배경 객체가 포함된 영상 데이터 세트를 추가로 학습한 모델이 mAP의 평균값이 9.44% 향상된 것으로 드러났고, 인식 결과에서도 차량, 차선 그리고 도로 이음새의 오인식이 크게 감소한 것으로 나타났다.

이 같은 방법은 심층 신경망을 활용한 기존의 도로 노면 파손 영역 탐지 알고리즘에도 적용이 가능하다. 이는 인식 성능을 높이기 위해 탐지하고자 하는 대상을 포함한 영상을 다수 확보하는 기존의 방법과는 달리 탐지 대상이 없는 영상을 이용한 방법이다. 앞서 개발된 알고리즘에도 이와 같은 방법을 사용한다면 인식 성능이 개선될 것으로 기대된다.

향후 연구로는 연산 속도의 향상을 위한 Region Proposal Network과 경계 상자의 정확도 향상을 위한 ROI Align의 적용이 필요하다고 보인다. 그리고 이를 바탕으로 Faster R-CNN 기법을 적용하여 실시간에 가까운 성능을 가지는 알고리즘 구현이 뒤따를 것이다. 끝으로 향후에 임베디드 시스템에 적용하여 실제 주행 차량에 탑재하여 그 성능을 확인하고자 한다.

ACKNOWLEDGEMENTS

본 논문은 국토교통과학기술진흥원의 “국토교통기술사업화지원 사업 [인공지능 기반의 도로 파손 정보 제공 기술 개발]”지원으로 수행하였습니다.

REFERENCES

- Chun C., Shim S., Kang S. and Ryu S.(2018), “Development and evaluation of automatic pothole detection using fully convolutional neural networks,” *Journal of Korea Institute of Intelligent Transport System*, vol. 17, no. 5, pp.55-64.
- Girshick R.(2015), “Fast r-cnn,” in *Proc. The IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp.1440-1448.
- Glorot X. and Bengio Y.(2010), “Understanding the difficulty of training deep feedforward neural networks,” in *Proc. the thirteenth international conference on artificial intelligence and statistics*, Sardinia, Italy, pp.249-256.
- Goodfellow I., Bengio Y. and Courville A.(2016), *Deep Learning*, MIT Press, Cambridge, MA.
- Ioffe S. and Szegedy C.(2015), “Batch normalization: accelerating deep network training by reducing internal covariate shift,” in *Proc. the 32nd International Conference on Machine Learning (ICML)*, Lille, France, pp.448-456.
- Jo Y. and Ryu S.(2015), “Pothole detection system using black-box camera,” *Sensors*, vol. 15, no. 11, pp.29316-29331.
- Kingma D. P. and Ba J. L.(2015), “ADAM: a method for stochastic optimization,” in *Proc. third International Conference on Learning Representations (ICLR)*, San Diego, CA, pp.1-15.
- Maeda H., Sekimoto Y., Seto T., Kashiya T. and Omata H.(2018), “Road damage detection and classification using deep learning neural networks with smartphone images,” *Computer-Aided*

- Civil and Infrastructure Engineering*, vol. 33, no. 12, pp.1127-1141.
- Nair V. and Hinton G. E.(2010), "Rectified linear units improve restricted boltzmann machines," in *Proc. the 27th International Conference on Machine Learning (ICML)*, Haifa, Israel, pp.807-814.
- Uijlings J. R., Van De Sande K. E., Gevers T. and Smeulders A. W.(2013), "Selective search for object recognition," *International Journal of Computer Vision*, vol. 104, no. 2, pp.154-171.
- Yang X., Li H., Yu Y., Luo X., Huang X. and Yang X.(2018), "Automatic pixel-level crack detection and measurement using fully convolutional networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp.1090-1109.
- Yu F. and Koltun V.(2015), "Multi-scale context aggregation by dilated convolutions," *arXiv preprint arXiv:1511.07122*.
- Zhang A., Wang K. C., Li B., Yang E., Dai X., Peng Y., Fei Y., Liu Y., Li J. Q. and Chen C.(2017), "Automated pixel-level pavement crack detection on 3Dasphalt surfaces using a deep-learning network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 10, pp.805-819.