

CNN을 사용한 노면 크랙의 위치 검출과 분류

이동규¹¹신한대학교 IT융합공학부 교수

Position Detection and Classification of Road Surface Crack using Convolutional Neural Network

Dong-Gyu Lee¹¹Professor, School of IT Convergence Engineering, Shinhan University¹Corresponding author: dglee@shinhan.ac.kr

Received June 16, 2020; Revised June 16, 2020; Accepted June 21, 2020

ABSTRACT

본 논문에서는 CNN을 사용하여 도로의 노면에 나타나는 크랙을 촬영한 영상으로부터 크랙의 위치와 모양을 검출하고 크랙의 종류를 분류하는 방법을 제안한다. CNN 기법을 물체의 세그멘테이션을 위해 사용한 의미론적 세그멘테이션 방법을 크랙의 검출을 위해 사용하였고, 검출된 결과 영상으로부터 크랙의 종류를 구분할 수 있는 신경망을 사용하였다. 신경망의 학습을 위해 지도학습 방법을 사용하였으며, 주행 중인 차량에 설치된 카메라를 통해 촬영한 도로 영상과 이 영상의 크랙 위치를 세그멘테이션한 결과, 크랙 분류 결과를 학습데이터로 사용하였다. 실제 도로 영상에 적용해 본 결과, 크랙의 위치를 효과적으로 검출하였고 크랙의 위치와 모양으로부터 크랙의 종류를 분류할 수 있었음을 확인하였다. 이 방법은 도로의 노면 관리에 효율적으로 사용될 수 있을 것으로 생각된다.

In this paper, we propose a method using CNN for detecting the location and shape of cracks and classifying the types of cracks from the crack images appearing on the road surface. The semantic segmentation method using CNN technique for object segmentation was used for the detection of cracks and a neural network capable of distinguishing the type of crack was applied to the detected crack image. Supervised learning method was used for training neural networks and a road image taken through a camera installed in a driving vehicle, the image of segmented area corresponding to crack position and the crack classification result was used as training data. As a result of deploying it to the real road image, the location of the crack was effectively detected and it was confirmed that the type of crack could be classified based on the location and shape of the crack. It is thought that this method can be used for efficient management of road surfaces.

Keywords: Road surface crack, Pothole, Image classification, Convolutional neural network, Semantic segmentation

1. 서론

도로의 노면은 도로의 노후화와 지반의 침하 등으로 인해 포장 면에 균열이 발생하여 크랙이 생기기도 하고 일부가 떨어져 나가거나 침식작용으로 인해 구멍이 생기는 포트홀이 발생하기도 한다. 이러한 노면 상태의 불균일은 주행 중인 운전자에게 큰 위험이 되기도 하고 사고로 이어지기도 한다. 그러므로 이러한 위험 상황을 조기에 발견하고 신속한 유지 보수가 이루어지도록 하기 위해서 크랙의 유무를 검출하고 도로 현장의 정보를 신속히 전달하는 관리체계가 필요하다.

노면의 크랙을 검출하기 위해서 여러 가지 방법이 사용될 수 있으나 도로의 주행 중에 차량에 설치된 카메라를 통해 촬영된 영



상으로부터 크랙을 자동으로 검출하는 시스템이 효율적이다¹⁻³. 도로 영상에서 크랙의 위치를 검출하기 위해서 기존의 여러 가지 영상인식 기법들이 사용되었으나 다양한 영상인식 분야에 딥러닝 기법의 하나인 CNN (Convolutional Neural Network)을 적용함으로써 기존의 방법에 비해 뛰어난 성능을 얻을 수 있었고 크랙의 검출에도 우수한 성능을 얻을 수 있었다⁴⁻⁷. 영상의 분류와 위치 검출을 위해 R-CNN (Region-CNN) 기법 등이 제안되었다⁸. R-CNN은 입력 영상을 작은 크기의 관심 영역으로 나누어 다수의 중첩된 관심 영역을 설정하고, 각각의 영역을 CNN의 입력으로 사용하여 특징 맵(feature map)이 생성되도록 한다. 이러한 방법에 의해 영상의 분류와 위치 검출을 동시에 할 수 있도록 하였으나, 지나친 반복 계산으로 인해 막대한 계산량과 처리 시간이 필요하다⁹.

CNN을 레이블링에 적용한 의미론적 세그멘테이션(Semantic segmentation) 알고리즘에서는 fully-CNN이라는 신경망 구조를 사용하여 픽셀 단위로 세그멘테이션을 할 수 있었다^{10,11}. fully-CNN은 학습시킨 영상과 유사한 영상이 나타나는 영역의 세밀한 정보를 출력하지 않고 해당하는 픽셀의 화소값을 강조하는 형태로 작동하기 때문에 픽셀의 위치를 찾아내는 데 사용할 수 있다. 이러한 특성은 도로 영상에서 크랙의 위치와 모양을 검출하고 이 모양으로부터 크랙의 종류를 구분하는데 효율적일 수 있다.

본 논문에서는 CNN 기법을 사용한 의미론적 세그멘테이션 알고리즘을 픽셀 단위의 크랙 검출에 사용하여 입력 영상으로부터 크랙의 위치와 모양을 검출하고 어떤 종류의 크랙인지를 분류할 수 있는 시스템을 제안한다.

2. 제안한 신경망 구조

본 논문에서 사용한 신경망은 크랙의 위치를 검출하기 위한 신경망과 검출된 크랙의 위치와 모양으로부터 크랙의 종류를 구분하기 위한 신경망으로 구성되어 있다. 카메라에서 입력된 영상에서 크랙이 나타날 수 있는 영역을 정해서 고정된 크기의 영상을 신경망의 입력으로 사용한다. 크랙의 위치와 모양을 검출한 크랙위치영상은 입력 영상과 동일한 크기의 영상이 출력되도록 하며, 크랙위치영상을 입력으로 하여 크랙의 종류를 구분한 결과를 출력하도록 하는 2단계 구조로 되어 있다. Fig. 1은 전체 신경망의 구조를 나타낸다.

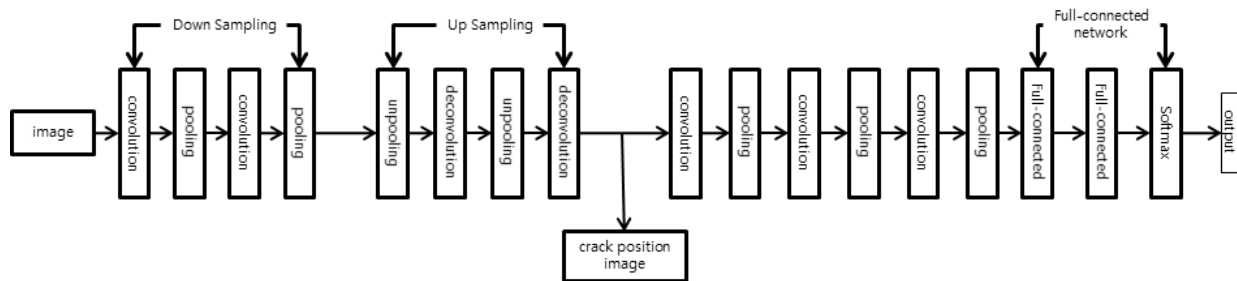


Fig. 1. Proposed CNN structure for detection of crack position and classification of crack type

2.1 크랙 위치 검출을 위한 CNN

의미론적 세그멘테이션에서 사용한 방법을 적용하여 픽셀 단위의 크랙 위치를 검출할 수 있도록 한다. 도로의 노면 영상을 입력으로 사용하고, 입력 영상과 같은 크기이고 픽셀 단위의 크랙 검출 결과가 계산된 크랙위치영상이 출력되도록 한다. CNN에서 사용되는 여러 층의 기본 요소(합성곱 층, 풀링 층, 활성화 함수)의 조합과 반대기능을 하는 역 합성곱 층, 역 풀링 층으로 구성되어 있다.

2.1.1 합성곱 층

입력 영상에 대해 합성곱 연산의 결과를 구한다. 합성곱 층에서는 전 연결 신경망(full connected network)과 달리 필터 크기에 해당하는 영역의 픽셀만 다음 단계로 전달되고 픽셀 사이의 평면적인 위치 분포가 반영된 연산 구조로 되어있다. 합성곱을 계산하는데 사용하는 필터 k 가 $H \times H$ 크기이고, 이 필터로부터 구해진 특징 맵이 z_k 라고 하면 좌퓯값(i, j)에서의 계산 결과는 식(1)과 같이 표현된다.

$$z_k(i, j) = f_k \otimes x(i, j) = \sum_{p=-\frac{H-1}{2}}^{\frac{H-1}{2}} \sum_{q=-\frac{H-1}{2}}^{\frac{H-1}{2}} f_k(p, q) x(i + p, j + q) + bias \quad (1)$$

여기서 \otimes 는 합성곱 연산을 의미하며 $f(p, q)$ 는 필터의 계수를 나타내고, 필터의 크기 H 는 일반적으로 홀수를 사용한다.

합성곱 연산의 결과는 활성화 함수를 통과하는데 일반적으로 ReLu 함수를 사용한다. 필터를 일정한 간격으로 이동시키면서 반복적인 연산을 수행하는데, 이때의 이동 간격을 스트라이드(stride)라고 하며 이에 따라 결과 영상의 크기가 줄어들게 된다.

2.1.2 풀링 층

풀링 층에서는 정해진 영역으로부터 하나의 대퓯값을 선택함으로써 픽셀의 화소값과 위치 변화에 대한 민감도를 떨어뜨리고 두드러진 특징 만이 추출되도록 한다. 최대 풀링이나 평균 풀링 등이 사용되며 최대 풀링은 식(2)와 같이 표현된다.

$$p_g = \max_{i \in g} h_i \quad (2)$$

여기서 g 는 풀링이 이루어지는 영역을 나타내며 이 영역의 크기에 따라 영상의 세밀한 특징을 제거하고 명확한 특징만이 남도록 해주며 객체의 이동에 둔감하도록 만들어 준다. 합성곱 연산과 마찬가지로 일정한 간격으로 이동하면서 풀링이 이루어지는 영역을 선택하는데 스트라이드의 크기에 따라서 출력 영상의 크기가 줄어든다.

2.1.3 역 풀링(unpooling) 층

역 풀링은 풀링의 반대 과정으로 하나의 대퓯값을 해당 영역의 적절한 위치에 저장하고, 그 외의 나머지 위치의 값을 결정해 주어야 한다. switch라는 저장 위치를 이용해 풀링 과정에서 선택된 위치를 저장하고 역 풀링에서 switch에 저장된 위치를 참고하여 픽셀값을 복원하기도 한다^[11]. 본 논문에서는 화소의 밝기를 복원할 필요가 없고 레이블링과 같이 크랙이 존재하는 픽셀들의 위치를 검출하는 것이 목적이므로 switch 메모리를 사용하지 않고 Fig. 2와 같이 해당 영역에 동일한 값들을 복사함으로써 업 샘플링의 결과를 얻을 수 있다. 이때 스트라이드 값을 풀링 영역의 크기와 같아지도록 설정하여 풀링 영역이 겹치지 않도록 하였다. 이는 역 풀링 시 동일한 위치에 대응되는 화소값이 여러 개가 발생하는 것을 방지하기 위함이다.

역 풀링은 원래의 화소값을 복원할 수가 없어 세밀한 화소값은 손실되지만 합성곱 과정에서 추출된 특징이 강조된 영상을 얻을 수 있도록 해준다. g_j 영역에 대한 역풀링 결과 h_k 는 식(3)과 같이 표현할 수 있다.

$$h_k(i) = w_k(i) p_k(j) \quad \forall i \in g_j \quad (3)$$

여기서 p_k 의 j 번째 요소는 주변에 있는 g_j 의 영역에 영향을 받으므로 가중치 $w_k(i)$ 에 의해 복원된 값이 결정된다. 가중치 $w_k(i)$ 의 분포에 따라 역 풀링의 픽셀 복원 방법이 달라진다. switch 메모리를 사용하는 경우 풀링시 저장된 값이 역 풀링의 가중치 $w_k(i)$ 에 해당하고, 동일한 값을 복사하는 형태를 사용할 경우 모든 가중치가 1로 설정된 것과 같다. Fig. 2는 최대 풀링을 사용했을 때의 결과와 이로부터 역 풀링을 수행한 복원 영상을 나타낸다.

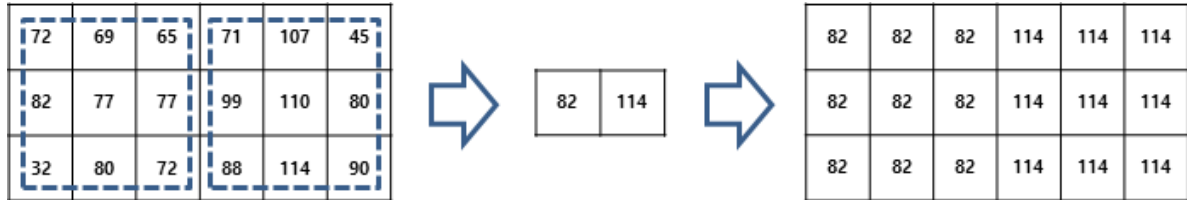


Fig. 2. Maxpooling and unpooling for pixel extraction and restoration

2.1.4 역 합성곱(Deconvolutional Network) 층

역 합성곱 층에서는 합성곱 층에서와 같은 연산이 진행되지만, 영상을 복원하기 위한 필터가 사용된다. 역 풀링에 의해 크기가 늘어난 영상을 입력으로 하여 출력 영상을 생성하기 위한 필터 F_j 와 합성곱 연산이 계산된다. 이는 합성곱 층에서 특징을 추출하는 과정의 역 과정에 해당한다.

$$\hat{x} = F_j \otimes \hat{z} \quad (4)$$

여기서 식 (4)는 합성곱 연산과 동일한 연산을 수행하지만, 이때 사용되는 F_j 의 필터 계수는 지도학습을 위해 제시된 크랙의 레이블링 영상과 유사한 출력을 만들어 낼 수 있도록 학습 과정에서 반복적인 계수 갱신을 통해 적합한 필터 계수를 찾아나가게 된다.

2.1.5 다운 샘플링 과정과 업 샘플링 과정

입력 영상은 여러 개의 합성곱 층과 풀링 층을 통과하면서 영상의 특징이 추출되고, 스트라이드와 풀링을 통해 출력 영상의 크기가 줄어든다. 이 과정에서 입력 영상이 다운 샘플링되는 효과가 발생한다. 또한, 추출된 특징 맵은 역 합성곱 층과 역 풀링 층을 통과하면서 입력 영상의 크기와 같아지도록 출력 영상의 크기를 늘려 주면서, 업 샘플링이 이루어지고 결과 영상의 크기가 늘어난다. 다운 샘플링 과정에서 사용된 은닉층의 순서는 업 샘플링 과정에서 역순으로 사용되고, 또한 다운 샘플링 과정에서 은닉층의 설정 파라미터들, 즉 합성곱의 필터 크기, 스트라이드와 풀링 연산의 크기는 업 샘플링 과정에서도 동일하게 적용되어야 입력 영상과 같은 크기의 출력 영상을 복원해 낼 수 있다.

2.2 크랙의 종류 구분을 위한 CNN

크랙위치영상은 크랙이 존재하는 픽셀 위치에 흰색의 화소값이 출력되도록 하였으므로 출력 화소값의 분포와 모양으로부터 크랙의 종류를 구분하도록 한다. 크랙위치영상을 입력으로 하고, 합성곱 층과 풀링 층에 의해 결과 영상의 크기를 줄여나가면서 특징을 추출한다. 전 연결 신경망(full-connected network)에 추출된 특징을 입력함으로써 크랙의 위치와 모양으로부터 종류를 구분할 수 있도록 한다. 마지막 출력층의 활성화 함수는 소프트맥스 함수를 사용하여 각 크랙의 종류에 포함될 확률값이 출력되도록 한다.

3. 학습데이터 생성

크랙의 위치와 모양을 검출하고 크랙의 모양에 따른 종류의 구분을 위해 학습데이터를 작성한다. 학습데이터는 주행 중인 차량에 설치된 카메라로부터 영상을 취득하여 학습데이터를 수집하였다. 각 영상에서 도로의 크랙 위치가 보일 수 있는 부분을 정해서 800x200 크기의 영역을 추출하여 영상을 저장하고, 크랙의 위치와 모양을 저장하기 위해 입력 영상과 같은 크기의 크랙위치영상을 생성하였다. 크랙이 없는 영상과 크랙이 발견된 영상을 선택하고 크랙의 종류를 구분하여 학습데이터를 작성한다. 크랙위치영상은 크랙의 위치를 흰색으로 표시하여 작성하고, 육안으로 크랙 종류를 분류한 결과를 저장하였다. 크랙의 모양과 형태를 구분하기 위해 크랙의 종류를 5가지로 정의하였고, 크랙없음(non-crack), 수직크랙(vcrack), 수평크랙(hcrack), 포트홀(pothole), 지역크랙(local crack)으로 분류하였다. 크랙의 종류별 모양과 분류 결과의 예시는 Fig. 3과 같다.

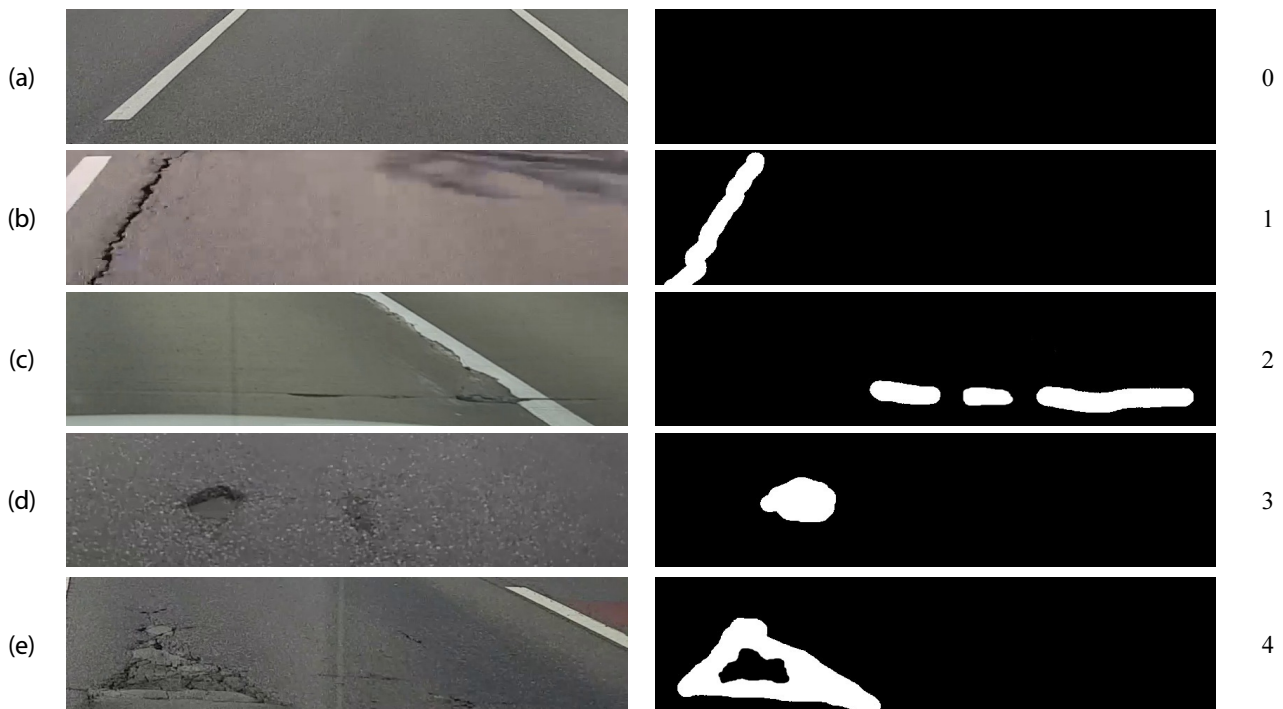


Fig. 3. Training Data for input image, crack position image and crack type (a) non-crack (b) vertical crack (c) horizontal crack (d) pothole (e) local crack

4. 실험결과

학습(training)과 적용(deployment) 단계에서 사용한 입력 영상은 800x200의 칼라 영상을 사용하였으며, 크랙위치영상은 800x200의 흑백 영상이 출력되도록 하였다. 크랙의 종류를 분류한 결과는 5가지의 결과를 출력할 수 있도록 5개의 출력 노드를 사용하였다. 크랙의 위치와 모양에 따라 입력 영상, 크랙위치영상, 크랙의 종류에 해당하는 학습데이터를 작성하였으며 총 800장의 영상을 학습데이터로 사용하였다. 적용 단계에서는 학습된 신경망의 성능평가를 위해 실제 도로를 주행한 차량에서 촬영된 동영상으로부터 입력 영상을 추출하고 신경망으로부터 계산된 크랙위치영상과 분류 결과를 평가하였다.

Fig. 4는 신경망 모델에서 출력된 크랙위치영상과 분류 결과의 예를 나타내며, 크랙이 있는 부분에서 검출 결과가 표시되고 있고 적절한 분류 결과를 보이고 있다.



Fig. 4. Crack position and classification result of real road image

Table 1은 크랙의 종류 별 학습데이터의 수와 적용한 영상의 수, 크랙의 위치 검출율, 분류정확도를 나타낸다. 적용 과정에서 사용한 영상은 동영상의 모든 프레임을 입력으로 사용하였기 때문에 주행 위치에 따라 크랙의 위치와 크기가 변할 수 있으므로 화면의 적절한 위치에 도달했을 때의 결과를 사용하였다. 위치 검출률은 육안으로 판정한 크랙 영역이 크랙 영상에 나타나는 비율을 나타내고, 분류 정확도는 육안으로 판정한 크랙의 종류와 신경망에서 분류된 결과를 비교하였다.

실험 결과 94.6%의 높은 검출율과 91.0%의 정확도를 얻을 수 있었다. 일반적으로 크랙의 형태가 수직 크랙이 많았으며 수평 크랙은 적게 나타났다. 수평과 수직 크랙이 모두 포함된 복합 크랙은 지역 크랙으로 분류하였으나 포트홀과의 구분이 모호한 부분이 있어 분류 결과도 다소 낮은 정확도를 보였다. 또한, 포트홀은 내부의 굴곡이 영상으로 구분할 수 있어야 하지만 조명의 영향등으로 포트홀의 내부가 불명확하거나 원형 크랙의 형태로 되어 있는 경우 지역 크랙으로 분류되기도 하였다. 제한한 방법은 크랙의 위치를 검출하는 데 효과적이었으며 크랙의 종류를 구분할 수 있었으나, 크랙의 종류를 나누기 위한 명확한 정의가 필요할 것으로 보인다.

Table 1. Performance evaluation result about crack position and classification

크랙종류	학습 데이터의 수	검사 영상	위치 검출율(%)	분류 정확도(%)
크랙없음	300	350	98	97
수직크랙	180	134	92	96
수평크랙	80	45	97	94
포트홀	120	51	94	82
지역크랙	120	70	92	86
합계	800	650	94.6	91.0

5. 결론

본 논문에서는 CNN 기법을 도로의 크랙 영상에 적용함으로써 크랙의 위치 검출과 크랙의 종류를 구분 할 수 있는 방법을 제안하였다. 크랙의 위치 검출을 위해 CNN 기반의 레이블링 방법인 fully- CNN을 변형하여 크랙의 위치를 나타내는 크랙 영상을 생성하였고 이로부터 크랙의 모양과 위치를 기준으로 크랙의 종류를 구분하도록 하였다. 도로에서 나타나는 다양한 크랙 영상을 학습데이터로 사용하였고, 지도학습의 형태를 사용하기 위해 입력 영상의 크랙 위치를 표시한 영상과 육안으로 분류한 크랙의 종류를 제시하여 신경망을 학습시켰다. 학습된 CNN에 실제 도로 영상을 적용해 본 결과 크랙의 위치와 모양을 검출하는데 효과적인 방법임을 확인하였고 크랙의 종류를 자동 분류할 수 있었다. 향후, 크랙의 종류와 모양에 대한 명확한 기준이 정립되어야 할 것으로 판단되며, 본 연구의 결과를 바탕으로 도로 영상에서 크랙의 위치를 검출하고 현장 상황을 기록, 관리할 수 있는 기능을 자동화함으로써 효율적인 도로 관리에 활용할 수 있는 노면 관리시스템으로 사용할 수 있음을 확인하였다.

Acknowledgement

※ 본 논문은 2019년도 신한대학교 학술연구비 지원으로 연구되었음.

References

1. 김정주, 강병호, 최수일, “2차원 라이다 기반 3차원 포트홀 검출 시스템”, 한국디지털콘텐츠학회 논문지, 제18권, 제5호, pp. 989-994, 2017.8.
2. 김영로, 조영태, 류승기, “명암과 움직임 정보를 이용한 포트홀 검출”, 전자공학회 논문지, 제 52권, 제 11호, pp. 137-146, 2015.
3. 김정민, 현세권, 채정환, 도명식, “드론영상을 이용한 물체탐지알고리즘 기반 도로균열탐지”, 한국ITS학회 논문지, 제18권, 제 6호, pp.155-163, 2019.12.
4. G. M. Jog, C. Koch, M. Golparvar-Fard, and I. Brilakis, “Pothole properties measurement through visual 2D recognition and 3D reconstruction”, International Conference on Computing in Civil Engineering, pp. 553-560, 2012.
5. E. Buza, S. Omanovic, and A. Huseinovic, “A pothole detection with image processing and spectral clustering”, Proc. the 2nd International Conference on Information Technology and Computer Networks, Antalya, Turkey, pp. 48-53, 2013.
6. A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, “Real time pothole detection using Android smartphones with accelerometers,” Proc. IEEE International Conference on Distributed Computing in Sensor Systems and Workshops, Barcelona, Spain. 2011.
7. T. Kim, and S.-K. Ryu, “Review and analysis of pothole detection methods”, Journal of Emerging Trends in Computing and Information Sciences, Vol. 5, No. 8, pp. 603-608. 2014.
8. R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 580-587, 2015.
9. 심승보, 전찬준, 류승기, “Fast R-CNN을 이용한 객체 인식 기반의 도로 노면 파손 탐지 기법”, 한국ITS학회 논문지, 제18권, 제2호, pp. 104-113, 2019.
10. J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431-3440, 2015.
11. M. D. Zeiler, R. Fergus, “Visualizing and Understanding Convolutional Networks”, European conference on computer vision (ECCV), pp. 818-833, 2014.