

영상품질별 학습기반 알고리즘 폐색영역 객체 검출 능력 분석*

이정민^{1*} · 함건우¹ · 배경호² · 박흥기³

Detection Ability of Occlusion Object in Deep Learning Algorithm depending on Image Qualities*

Jeong-Min LEE^{1*} · Geon-Woo HAM² · Kyoung-Ho BAE² · Hong-Ki PARK³

요 약

정보화 사회로 진입하면서 공간정보의 중요성은 급격하게 부각되고 있다. 특히 스마트시티, 디지털트윈과 같은 Real World Object의 3차원 공간정보 구축 및 모델링은 중요한 핵심기술로 자리매김하고 있다. 구축된 3차원 공간정보는 국토관리, 경관분석, 환경 및 복지 서비스 등 다양한 분야에서 활용된다. 영상기반의 3차원 모델링은 객체 벽면에 대한 텍스처링을 생성하여 객체의 가시성과 현실성을 높이고 있다. 하지만 이러한 텍스처링은 영상 취득 당시의 가로수, 인접 객체, 차량, 현수막 등의 물리적 적치물에 의해 필연적으로 폐색영역이 발생한다. 이러한 폐색영역은 구축된 3차원 모델링의 현실성과 정확성 저하의 주요원인이다. 폐색영역 해결을 위한 다양한 연구가 수행되고 있으며, 딥러닝을 이용한 폐색영역 검출 및 해결방안에 대한 연구가 수행되고 있다. 딥러닝 알고리즘 적용한 폐색영역 검출 및 해결을 위해서는 충분한 학습 데이터가 필요하며, 수집된 학습 데이터 품질은 딥러닝의 성능 및 결과에 직접적인 영향을 미친다. 따라서 본 연구에서는 이러한 학습 데이터의 품질에 따라 딥러닝의 성능 및 결과를 확인하기 위하여 다양한 영상 품질을 이용하여 영상의 폐색영역 검출 능력을 분석하였다. 폐색을 유발하는 객체가 포함된 영상을 인위적이고 정량화된 영상 품질별로 생성하여 구현된 딥러닝 알고리즘에 적용하였다. 연구결과, 밝기값 조절 영상 품질은 밝은 영상일수록 0.56 검출비율로 낮게 나타났고 픽셀크기와 인위적 노이즈 조절 영상 품질은 원본영상에서 중간단계의 비율로 조절된 영상부터 결과 검출비율이 급격히 낮아지는 것을 확인할 수 있었다. F-measure 성능평가 방법에서 노이즈 조절한 영상 품질 변화가 0.53으로 가장 높게 나타났다. 연구결과로 획득된 영상 품질별에 따른 폐색영역 검출 능력은 향후 딥러닝을 실제 적용을 위한 귀중한 기준으로 활용될 것이다. 영상 취득 단계에서 일정 수준의 영상 취득과 노이즈, 밝기값, 픽셀크기 등에 대한 기준을 마련함으로써 딥러닝을 실질적인 적용에 많은 기여가 예상된다.

2019년 09월 23일 접수 Received on September 23, 2019 / 2019년 09월 26일 수정 Revised on September 26, 2019 / 2019년 09월 26일 심사완료 Accepted on September 26, 2019

* 본 연구는 국토교통부 공간정보 기반의 실감형 콘텐츠 융복합 및 혼합현실 제공기술 개발 연구개발사업의 연구비지원(19DRMS-B147287-02)에 의해 수행되었습니다.

1 (주)신한항공 기업부설연구소 연구원 Research Institute, Shinhan Aerial Survey CO.,LTD(Corresponding author: insomnie08@gmail.com)

2 (주)신한항공 기업부설연구소 연구소장 Research Institute, Shinhan Aerial Survey CO.,LTD

3 가천대학교 토목환경공학과 교수 Professor of Gachon university(hgpark@gachon.ac.kr)

* Corresponding Author E-mail : insomnie08@gmail.com

주요어 : 딥러닝, 3차원 모델링, 텍스처링, 영상품질, 폐색영역

ABSTRACT

The importance of spatial information is rapidly rising. In particular, 3D spatial information construction and modeling for Real World Objects, such as smart cities and digital twins, has become an important core technology. The constructed 3D spatial information is used in various fields such as land management, landscape analysis, environment and welfare service. Three-dimensional modeling with image has the high visibility and reality of objects by generating texturing. However, some texturing might have occlusion area inevitably generated due to physical deposits such as roadside trees, adjacent objects, vehicles, banners, etc. at the time of acquiring image. Such occlusion area is a major cause of the deterioration of reality and accuracy of the constructed 3D modeling. Various studies have been conducted to solve the occlusion area. Recently the researches of deep learning algorithm have been conducted for detecting and resolving the occlusion area. For deep learning algorithm, sufficient training data is required, and the collected training data quality directly affects the performance and the result of the deep learning. Therefore, this study analyzed the ability of detecting the occlusion area of the image using various image quality to verify the performance and the result of deep learning according to the quality of the learning data. An image containing an object that causes occlusion is generated for each artificial and quantified image quality and applied to the implemented deep learning algorithm. The study found that the image quality for adjusting brightness was lower at 0.56 detection ratio for brighter images and that the image quality for pixel size and artificial noise control decreased rapidly from images adjusted from the main image to the middle level. In the F-measure performance evaluation method, the change in noise-controlled image resolution was the highest at 0.53 points. The ability to detect occlusion zones by image quality will be used as a valuable criterion for actual application of deep learning in the future. In the acquiring image, it is expected to contribute a lot to the practical application of deep learning by providing a certain level of image acquisition.

KEYWORDS : *Deep Learning, 3D Modeling, Texturing, Image Quality, Occlusion Area*

서론

스마트시티, 디지털트윈 등의 Real World Object의 3차원 공간정보 구축 및 모델링 기술을 지속적으로 발전하고 있으며, 고정밀의 3차원 모델링 데이터는 다양한 분야에서 광범위하게 활용되고 있다. 실세계의 다양한 객체에 대한 정확한 3차원 모델링은 국토관리, 시설물 관리, 환경 및 복지 분야에서 중요한 기반인프라

로 사용되고 있다. 또한 Virtual Singapore과 같이 도심 전체에 대한 실감형 3차원 모델링 구축을 지속적으로 발생할 것이다. 현재 3차원 모델링은 주로 항공사진, 드론촬영, MMS(Mobile Mapping System) 등을 이용한 영상기반으로 구축되고 있다. 이러한 영상은 모델링 후 텍스처링 처리를 함으로써 실감형 3차원 모델링으로 구현된다.

하지만, 영상 취득 당시의 가로수, 차량, 현수막 등의 물리적 적치물에 의해 필연적으로 폐색

영역이 발생하며, 발생된 폐색영역은 3차원 모델링의 현실성과 가시성을 저하시키며 나아가 3차원 모델링 성과 및 품질을 저하시킨다. 따라서 3차원 모델링의 폐색영역 해결은 중요한 핵심기술이다.

이를 위해 최근 딥러닝 알고리즘을 적용하여 폐색영역을 탐지하고 해결하고자 하는 다양한 연구가 시도되고 있다. 객체 검출은 영상에서 찾고자 하는 객체의 위치와 크기를 찾아내는 것이며, 컴퓨터 비전에서 꾸준히 연구가 진행되어 왔다. 최근에는 딥러닝 기술의 발전으로 객체 검출 성능이 매우 향상된 것을 확인할 수 있다 (Hariharan *et al.*, 2014; LeCun *et al.*, 2015).

하지만 기존의 연구는 주로 딥러닝 알고리즘을 이용한 폐색영역 탐지 및 해결 가능성 중심으로 연구를 수행하였으며, 학습 데이터의 품질 및 영향에 대한 연구는 미비한 실정이었다. 딥러닝 알고리즘 적용한 폐색영역 검출 및 해결을 위해서는 입력되는 학습 데이터의 품질이 중요하다. 충분한 학습 데이터를 이용한 데이터셋을 구축하여 딥러닝 알고리즘에 적용하여야 한다. 또한, 이때 일정 수준 이상의 영상품질이 학습되어야 우수한 딥러닝의 성과와 결과에 도출할 수 있다.

이에 본 연구는 텐서플로우 기반으로 구축된 딥러닝 알고리즘을 대상으로 폐색을 유발하는 객체가 포함된 영상을 인위적이고 정량화된 영상품질로 생성하였다. 연구의 목표는 노이즈, 밝기값, 픽셀크기 등 다양한 영상품질에 따른 딥러닝 알고리즘의 성과와 결과를 분석하여 향후 폐색영역 검출 및 해결을 위한 영상품질의 기준을 제시하고자 한다.

영상 처리를 위한 딥러닝 알고리즘 종류

1. 딥러닝 알고리즘

딥러닝 알고리즘을 이용한 객체 탐지는 주어진 영상에서 특징점을 이용하여 특정 객체를 확인하고 그 크기와 위치를 탐지하는 것을 말한다. 폐색영역에서 딥러닝 알고리즘을 적용한 장점은 불규칙하게 폐색을 유발하는 가로수, 차량, 현수막 등과 같은 특정 객체를 탐지하여 위치를 자동으로 지정할 수 있는 장점이 있다. 따라서 딥러닝을 이용한 폐색영역 해결은 컴퓨터 비전 기술 발전과 더불어 공간정보 분야에서 중요하게 활용되고 있다.

딥러닝은 기본적으로 학습 모델에 대한 추론 과정과 학습과정을 가진다. 추론과정은 학습된 내용을 기반으로 새로운 입력에 대해 답을 얻어내는 과정이며, 학습은 주어진 학습 데이터를 기반으로 최적의 추론을 수행하기 위해 추론 구조 또는 학습 파라미터(가중치)들을 설정하는데, 간단히 말하면 귀납적으로 배우는 단계로 볼 수 있다.

2. CNN 및 R-CNN Algorithm

영상기반 딥러닝의 기초는 CNN(Convolutional Neural Network) 알고리즘이다. 인공신경망은 'Layer' 를 연속적으로 쌓은 구조를 통해 인간의 계층적 인지구조를 모방한다. 그림 1과 같이 CNN은 일반적으로 Convolution, Batch Normalization, Activation 및 Pooling Layer

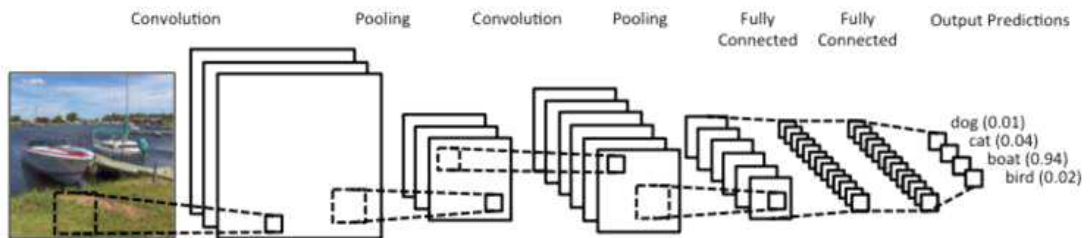


FIGURE 1. Concept of CNN

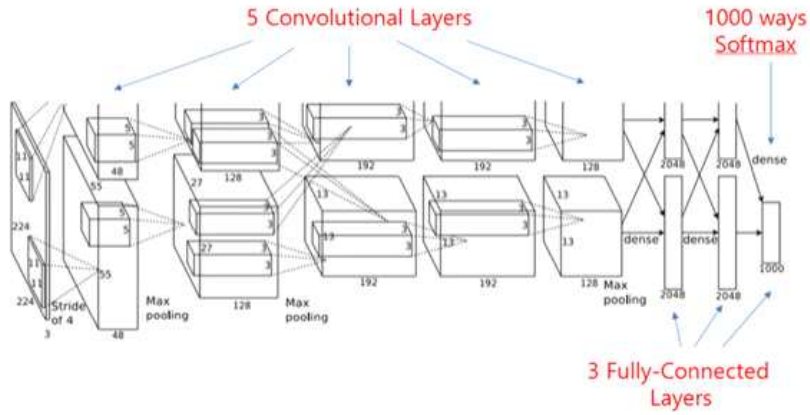


FIGURE 2. Convolutional layers of CNN

로 구성된다. Convolution 연산은 특정 크기의 커널(Kernel)을 입력 영상의 일부분과 겹쳐 Matrix Multiplication을 수행한다. 보통 하나의 Layer에서 수십~수백개 커널이 사용된다. 따라서 Convolution의 결과로 구한 Feature Map은 커널의 개수와 같은 채널수를 가지는 다채널 영상의 형태를 가진다. 인공신경망 이전에도 ‘Hand-designed Feature’를 사용하는 기법들이 있었다(Zhou *et al.*, 2014).

그림 2는 5개의 Convolution Layer들을 거쳐 3개의 Fully-Connected Layer를 통해 1000가지의 Softmax가 나오는 CNN 상세 구성을 나타낸다.

CNN 알고리즘을 개량한 R-CNN(Region-based Convolutional Neural Network)은 영

상내의 물체가 있을 법한 영역 후보들을 먼저 인식하고 이들의 스코어를 매겨 물체를 인식하는 방법이다. 선택적 영역에 대해서 CNN 연산을 수행하여 얻어진 특징으로 SVM(Support Vector Machine)과 Linear Regression으로 객체 여부를 판단하고 위치를 결정하게 된다. 영상검출 분야에서 주로 적용하여 SIFT(Scale-Invariant Feature Transform)나 HOG(Histogram of Oriented Gradient) 기반의 컴퓨터 비전 알고리즘보다 우수한 탐지능력을 가지고 있다.

3. Faster R-CNN Algorithm

Faster R-CNN은 기본적으로 검출 정확도가 R-CNN보다 좋으며, 학습의 결과를 연결 신경

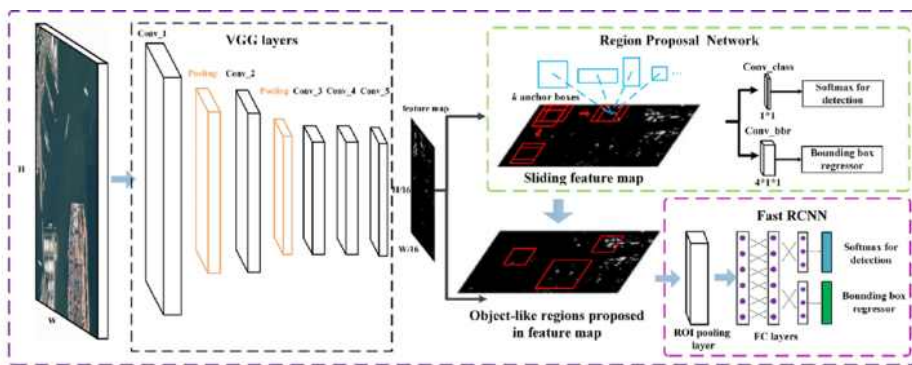


FIGURE 3. Structure of faster R-CNN

망에 있는 모든 레이어에 갱신할 수 있고 Feature Caching을 위해 별도의 디스크 공간이 필요 없는 장점이 있다. 그림 3은 Faster R-CNN의 구조를 나타낸다. ROI(Region Of Interest) Pooling이라는 레이어 개념을 도입하여 CNN에서 얻어진 특징 지도의 일부 영역으로부터 정규화된 특징을 추출한다. Faster R-CNN은 각 후보 영역에 CNN을 반복적으로 적용하며 입력 영상에 CNN을 한 번만 적용하고 ROI Pooling으로 객체를 판별하기 위한 특징을 추출한다. 이 방법으로 Faster R-CNN은 CNN에 비해 특징점 추출 시간을 줄이는 장점이 있다. 또한, 객체 판별이 더 우수하므로 일반적으로 큰 객체에 대해 검출이 가능하다. 따라서 영상기반의 객체 탐지 및 검출에 가장 많이 사용되고 있는 알고리즘이다.

4. Mask R-CNN Algorithm

Mask R-CNN은 그림 4와 같이 Faster R-CNN에서 영상 내의 분할된 바운딩 박스에서 객체 마스크링 네트워크를 추가하여 성능을 향상시킨 알고리즘이다. 영상에서 여러 ROI 후보를 지정하여 해당 위치의 특징점을 RoIAlign 방식으로 추출하는 기법이다. 추출된 특징점으로 객체를 분류하며 해당 객체의 위치와 크기를 탐지한다. 특히, Faster R-CNN은 큰 객체에 탐지 능력이 우수한 반면 Mask R-CNN은 사람과 자동차의 탐지에 우수한 성능을 나타내고 있다.

이러한 장점은 작은 객체에 대한 마스크 레이어를 추가함으로써 처리 속도와 시스템 크기에 효율적으로 적용될 수 있다.

학습 데이터셋 및 영상품질별 테스트 영상 생성

본 논문에서 제안하는 연구 흐름도는 그림 5와 같다. 테스트 영상품질 조절 후 Mask R-CNN 적용 결과를 이용하여 검출 영상품질 기준을 제안한다. 건물 텍스처에 다양한 형태의 폐색을 일으키는 가로수를 검출 객체로 지정하여 학습 데이터셋을 생성하고 밝기, 픽셀 크기, 노이즈와 같은 품질 조절한 영상을 테스트 영상으로 사용하여 Mask R-CNN 기법에 적용한 가로수 검출 결과를 비교 분석한다. 즉, 학습 데이터셋 검출 오류를 줄일 수 있는 검출 영상품질 기준을 세우고자 조절한 검출 영상별 폐색영역 검출 탐지율 분석 수행하였다.

학습 데이터셋 생성 및 Mask R-CNN 검출 실험 환경은 Intel core i7-7700 @4.20GHz, 메모리 16GB, GPU NVIDIA GeForce GTX 1070 8휴, Python 3.6, Linux Ubuntu 16.04 64bit 운영체제에서 시행하였다. 실험데이터는 학습 데이터셋 생성하기 위한 영상과 품질조절을 위한 테스트 영상을 사용한다.

1. 학습 데이터셋 생성방법

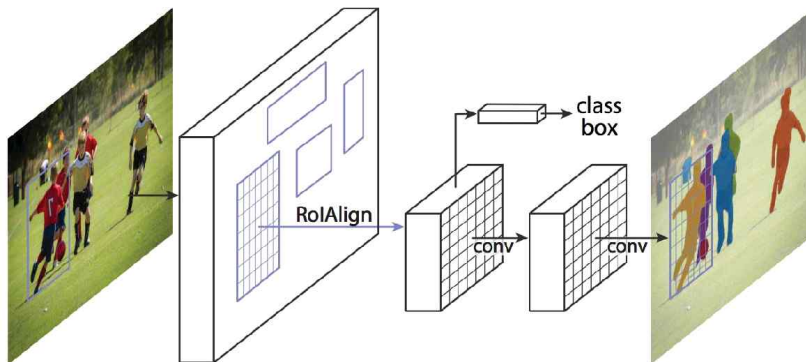


FIGURE 4. Concept of mask R-CNN

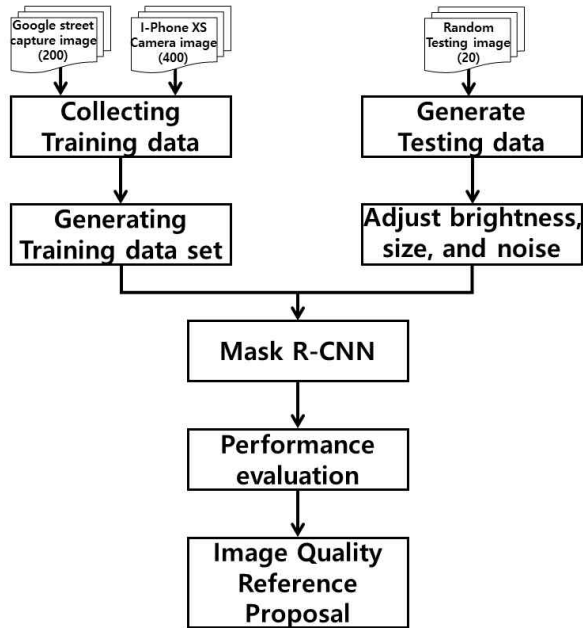


FIGURE 5. Flow of study

본 실험에 사용된 학습 영상은 총 600장이며 구글 스트리트뷰(<https://www.instant-streereview.com>)에서 캡처한 200장과 아이폰 XS로 촬영한 400장의 영상으로 구성되어있다. 그림 6은

옥스퍼드 대학교에서 개발한 무료소프트웨어인 VIA(VGG Image Annotator) Tool로 영상, 오디오 및 비디오에서 수동 레이블링 작업 가능하며 Json 포맷 파일로 제공한다. Json 포맷 파

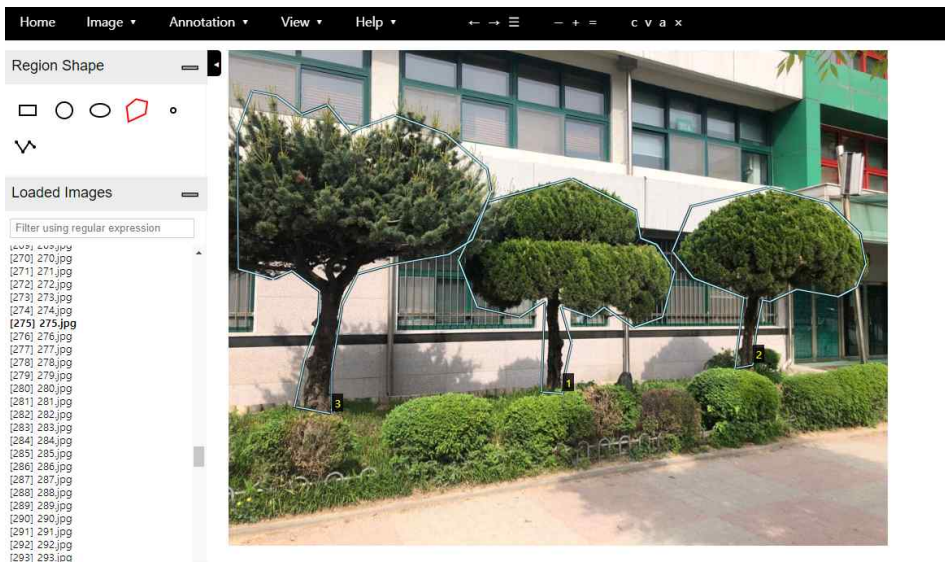


FIGURE 6. Tree labeling using via tool

일은 영상 내에 관심 영역으로 지정한 레이블링 데이터가 포함되어있다. 수집한 학습 영상에서 VIA Tool을 이용하여 관심 영역인 가로수를 지정하고 레이블링 작업을 수행하였다.

학습 데이터셋 생성방법은 GirHub에서 제공하는 matterport 페이지 (https://github.com/matterport/Mask_RCNN/releases)의 모델로 총 30번의 학습을 수행했다. 가로수 학습 데이터셋을 생성하기 위해 600장의 학습 영상과 Via Tool로 레이블링 작업한 데이터를 사용하여 진행했다. 위 학습 데이터셋을 생성하기 위한 시간은 3.5일이 소요되었다.

2. 영상품질별 테스트 영상 생성방법

영상품질에 따라 검출 결과를 알아보기 위해 영상에서 밝기, 픽셀크기, 노이즈, 등 다양한 특성을 인위적인 변형을 주었다. 테스트 영상은 무작위로 총 20장의 영상을 선별하였으며, Python을 이용하여 밝기, 픽셀크기, 노이즈를 원본 영상 기준으로 5가지의 범위로 나누어 처리하였다. 1을 원본 영상으로 두었으며 밝기값은 최저 비율 0.4, 최대 비율 2.0, 픽셀크기와 노이즈는 최저 비율 0.1, 최대 비율 1로 설정하였다. 밝기 특성은 PIL(Python Imaging Library)라이브러리를 사용하여 0.4, 0.7, 1.5,

TABLE 1. Brightness images by ratio

Artificial brightness case	
0.4	0.7
	
1.0	1.5
	
2.0	
	

TABLE 2. Pixel size images by ratio

Artificial pixel size cases	
1.0	0.7
	
	
	

2.0 비율별 부여하여 생성하였고, 픽셀크기는 OpenCV(Open Source Computer Vision) 라이브러리로 0.1, 0.3, 0.5, 0.7 비율의 리사이즈로 생성하였으며 마지막 노이즈 특성은 Numpy(Numerical Python) 라이브러리로 영상 값에 노이즈가 전체적으로 퍼지는 가우스 노이즈를 적용하고 0.1, 0.3, 0.5, 0.7 비율을 곱하여 생성하였다. 표 1~3은 각각 밝기, 픽셀크기, 노이즈의 품질 요소를 총 5가지의 범위로 나눈 테스트 영상을 나타낸다. 밝기값의 영상 변화를 준 테스트 영상은 시각적으로 보았을 때 차이가 도드라지게 생성되었으며, 픽셀크기와 노이즈에서 변화를 준 테스트 영상은 객체가 확

인될 수 있을 정도의 영상으로 생성되었다. 즉, 표 3에서 비율이 1에 가까울수록 좋은 품질이고 0.1에 가까울수록 나쁜 품질을 의미한다.

3. 학습모델 적용 Mask R-CNN 테스트

Mask R-CNN 검출 진행 과정은 다음과 같다. 컨볼루션 레이어를 지나 특징맵을 산출하고 학습 데이터셋 정보를 이용하여 영상에서의 픽셀이 검출 대상에 해당되는 것을 분류하는 마스크 네트워크에 적용하면 검출 결과가 나타난다. Mask R-CNN 결과 영상은 검출 객체를 사각형 모양으로 이루어진 바운딩 박스와 바운딩 박

TABLE 3. Noise images by ratio



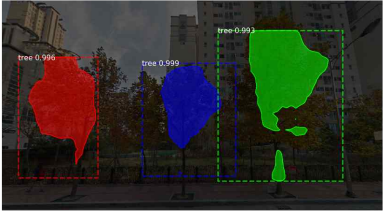
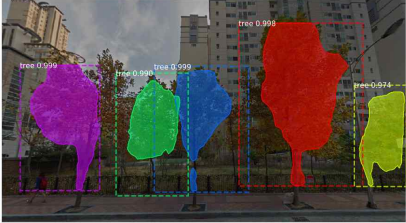
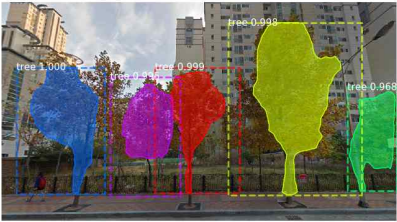
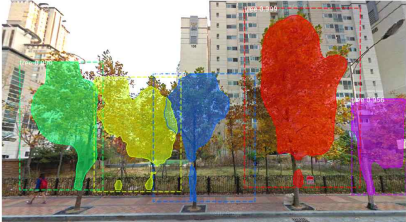

스안에 검출 객체의 영역을 마스킹한 형태로 결과가 나타난다. 영상품질에 따라 생성한 테스트 영상을 학습 데이터셋과 Mask R-CNN을 이용하여 객체 검출 수행하였고 영상 한 장마다 가로수 평균 검출 소요시간은 45초로 나왔다.

실험결과 및 성능평가 분석

본 실험에서 학습 데이터셋과 Mask R-CNN으로 테스트 영상을 객체 검출 수행하고 그 결과 성능평가 분석하였다. 표 4~6은 Mask R-CNN으로 품질별 테스트 영상에 따른 가로수 객체 검출 결과를 나타낸다. 영상품질별 결

과를 동일하게 비교할 수 있도록 테스트 결과 영상 중 하나의 대표 영상을 선정하였다. 밝기 값에서 검출된 결과는 원본영상과 가까운 값으로 조절한 영상에서 전체 7개 가로수 중 5개 검출 결과로 큰 차이는 없었고 가장 높은 밝기 값과 가장 낮은 밝기 값을 주어진 영상에서 전체 7개 가로수 중 3개의 적은 비율의 검출 결과가 나타났다. 픽셀크기를 조절한 영상에서는 원본 영상 결과가 전체 7개 가로수 중 5개로 가장 좋은 것으로 나타났으며 0.1의 비율은 준 영상이 7개 가로수 중 1개의 가로수만이 검출되었다. 노이즈를 조절한 영상에서는 원본 영상 결과 전체 7개 가로수 중 5개의 검출이 가장 좋

TABLE 4. Results of mask R-CNN detection in brightness images

Artificial brightness case	
0.4	0.7
	
1.0	1.5
	
2.0	
	

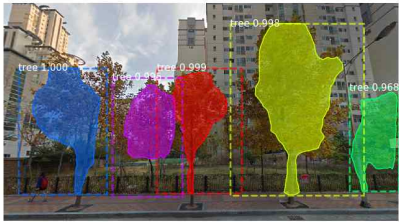
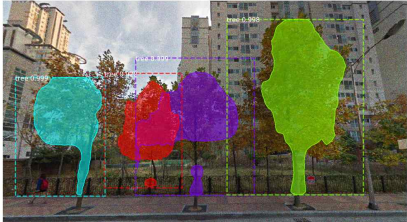
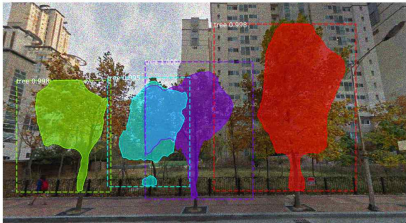
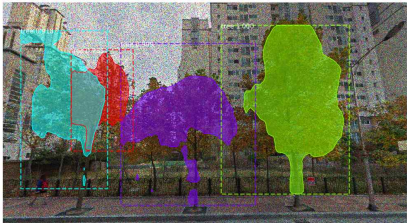

은 것으로 나타났고 최저 비율로 주어진 0.1의 값에서 검출된 가로수는 없는 것으로 나타났다.

또한, 테스트 영상에서 학습된 검출기를 적용한 결과 영상을 판별하였을 때 가로수에 가깝게 나타나는 비율은 전반적으로 전체 검출 면적에서 알맞게 검출된 비율이 높게 나타났지만, 이 파리 없이 나뭇가지로 양상한 가로수는 상대적으로 검출이 낮게 나타났다. 이는 영상품질이

아닌 학습된 검출기에서 나뭇가지로 이루어진 학습 영상 데이터가 부족한 것으로 보여진다.

본 실험의 학습모델 검출 성능평가를 위해 Precision-Recall 지표를 사용하였다. Precision-Recall은 일반적으로 기계학습에 사용되는 학습모델 평가 방법이며 정밀도인 Precision과 재현율인 Recall을 식 1~2로 산출할 수 있다.

TABLE 5. Results of mask R-CNN detection in noise images

Artificial noise case	
1.0	0.7
	
0.5	0.3
	
0.1	X
	

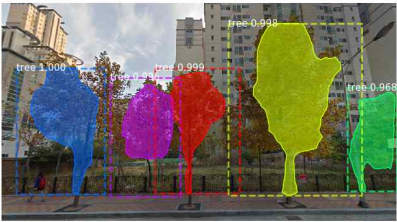
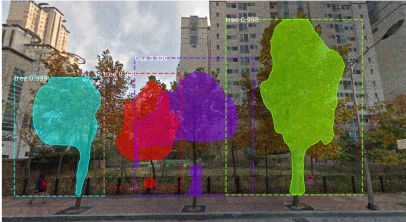
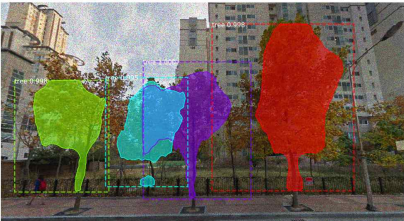
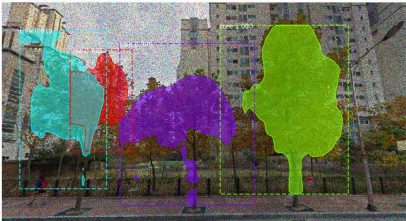
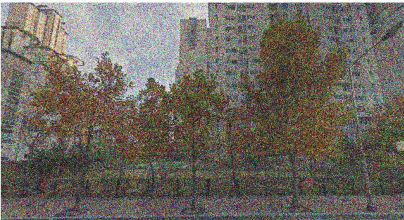
$$prec = TP / (TP + FP) \quad (1)$$

$$rec = TP / (TP + FN) \quad (2)$$

Precision은 검출된 데이터에서 바르게 검출된 비율로써 얼마나 정확하게 검출되었는지 나

타내고 Recall은 검출해야 할 데이터에서 바르게 검출된 비율을 나타낸다. 예를 들어 하나의 영상에서 7개의 가로수와 2대의 자동차가 존재한다고 가정하고 가로수를 검출하고자 한다. 여기서 가로수 검출 결과 총 5개의 객체검출이 발생하였고 4개는 가로수, 1개는 자동차로 나타났을 때 Precision은 5개의 검출 결과에서 4개의

TABLE 6. Results of mask R-CNN detection in noise images

Artificial noise case	
1.0	0.7
	
0.5	0.3
	
0.1	
	

가로수 검출된 결과를 나타내고 Recall은 7개의 검출해야할 가로수에서 4개의 가로수 검출 결과를 나타낸다. Precision-Recall 계산하기 위한 인자인 TP(True Positive)는 검출 결과에서 알맞은 객체로 검출된 개수, FP(False Positive)는 검출 결과에서 잘못된 객체로 검출된 개수, FN(False Negative)은 검출 되어 할

객체가 검출 되지 않은 개수를 나타낸 것으로 표 7과 같다.

표 8은 영상품질별 Precision-Recall로 산출한 학습모델 성능평가 결과이다. 25개의 테스트 영상에서 총 100개의 가로수가 있으며, 표 8과 같이 Precision-Recall 비율을 확인할 수 있다. 밝기와 픽셀크기 영상특성에서 Precision은 대

TABLE 7. Factor for calculating Precision–Recall

	Predicted \oplus	Predicted \ominus
Actual \oplus	TP (True positive)	FN (False negative)
Actual \ominus	FP (False positive)	TN (True negative)

TABLE 8. Precision–Recall results by image quality type

Brightness			Pixel size			Noise		
Ratio	Precision	Recall	Ratio	Precision	Recall	Ratio	Precision	Recall
0.4	1.00	0.65	1.0	1.00	0.71	1.0	1.00	0.71
0.7	1.00	0.70	0.7	0.97	0.61	0.7	0.97	0.65
1.0	1.00	0.71	0.5	0.97	0.66	0.5	0.95	0.65
1.5	1.00	0.69	0.3	1.00	0.57	0.3	0.98	0.54
2.0	0.97	0.56	0.1	0.97	0.30	0.1	0.65	0.20

부분 가로수에 해당되는 비율로 나타났고 노이즈에서 0.1의 비율을 주어진 영상에서 0.65로 다른 객체로 검출된 결과가 높은 것으로 나타났다. Recall을 비교했을 때 밝기에서 최대값은 0.71, 최소값은 0.56이고 픽셀크기에서 최대값은 0.71, 최소값은 0.30이며, 노이즈에서 최대값은 0.71, 최소값은 0.20이 나타났다. 이는 노이즈를 준 영상에서 검출해야 할 가로수가 제대로 검출이 되지 않는 성능이 나타났다.

그림 7~9는 영상품질별 Precision–Recall

결과 값을 이용한 그래프를 나타내주었다. 밝기 특성 조절한 그래프는 차이가 크게 없이 나타났고 픽셀크기와 노이즈 특성 조절한 그래프에서 0.5의 비율로 조절한 영상부터 급격하게 떨어지는 것을 확인할 수 있다.

F-measure(F1-score)는 Precision과 Recall 값으로 조화평균 계산된 분류 평가 지수이며 다음 식 3과 같다.

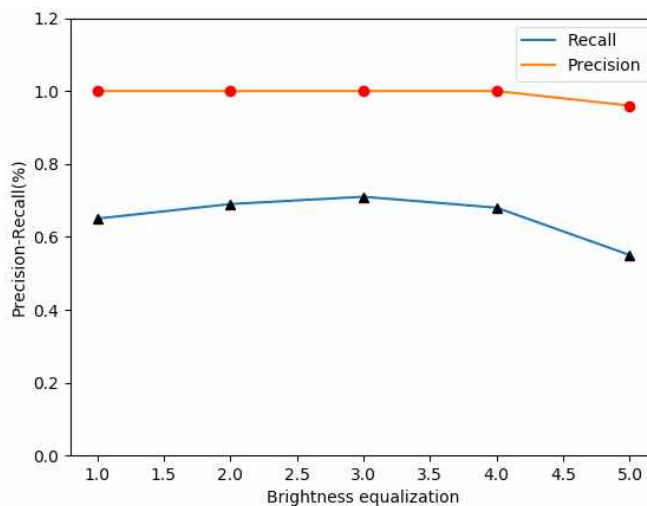


FIGURE 7. Precision–Recall based brightness images performance curves

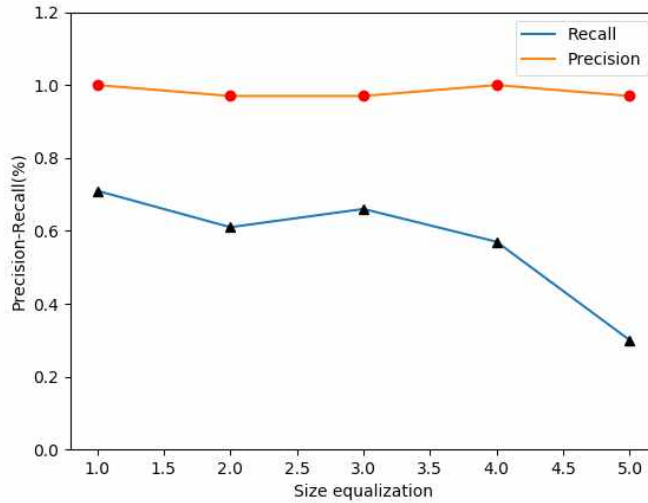


FIGURE 8. Precision-Recall based pixel size images performance curves

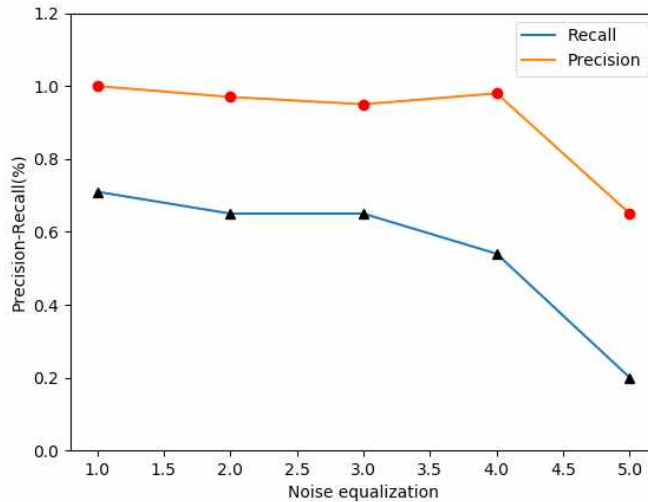


FIGURE 9. Precision-Recall based noise images performance curves

$$F1 \triangleq \frac{2}{\frac{1}{prec} + \frac{1}{rec}} = \frac{2 \cdot prec \cdot rec}{prec + rec} \quad (3)$$

$$= \frac{TP}{TP + (FP + FN) / 2}$$

Precision과 Recall 둘 다 고려하여 높은값을

나타내는 것이 학습 데이터셋 성능이 좋다고 평가할 수 있으며 Precision과 Recall은 서로 반비례하는 경향이 나타나기 때문에 균형점을 맞춰주는 F-measure를 사용하는 것이 합리적인 성능평가 비교가 가능하다. 또한, F-measure는 Precision과 Recall에서 낮은 값에 가중치를

TABLE 9. F-measure results by images quality type

Division	F-measure		
	Brightness	Pixel size	Noise
1	0.79	0.83	0.83
2	0.82	0.75	0.78
3	0.83	0.79	0.77
4	0.81	0.72	0.69
5	0.71	0.46	0.30

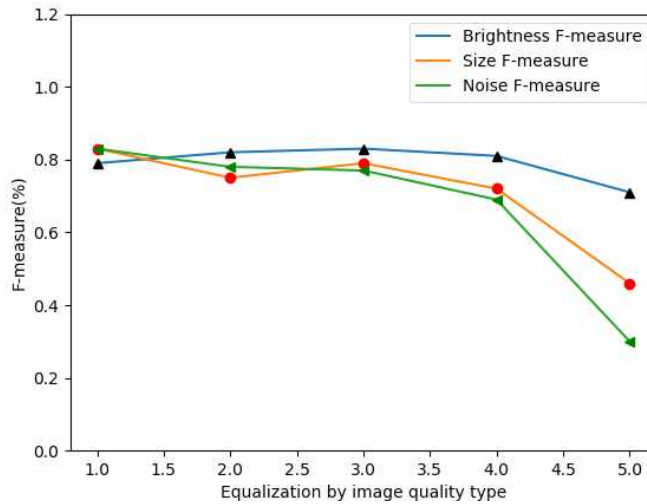


FIGURE 10. Results of F-measure

주는 평균으로 Precision과 Recall 정량적인 수치값을 나타내주어 정확한 비교가 가능하기 때문에 객체 검출분야에 사용된다. 표 9는 영상품질별 F-measure로 산출한 학습모델 성능평가 결과 값이다. 밝기에서 최대값은 0.83, 최소값은 0.71이고 픽셀크기에서 최대값은 0.83, 최소값은 0.46이며, 노이즈에서 최대값은 0.83, 최소값은 0.30이 나타났다. 밝기값을 조절한 테스트 영상 결과는 영상의 검출능력이 평이하게 나타났다. 픽셀사이즈를 조절한 테스트 영상 결과는 5단계에서 낮게 떨어지는 경향을 보였으며 노이즈를 조절한 테스트 영상 결과는 4단계부터 5단계까지 낮은 성능검출 결과가 나타났다.

그림 10은 표 9의 결과를 그래프로 나타내주었으며, Precision과 Recall 사이에서 낮은 값에 가까운 곳으로 F-measure 결과 값이 나타

났고, 노이즈 값을 조절한 테스트 영상이 차이가 가장 많이 발생한 것을 보여준다.

결론

본 논문은 딥러닝 알고리즘을 적용한 폐색영역 검출 및 해결을 위한 학습 데이터의 품질을 결정하기 위한 연구이다. 이를 위해 영상처리에 주로 사용되는 R-CNN 딥러닝 알고리즘을 적용하여 폐색영역 검출 및 해결을 모색하였다. 실험을 위해 600장의 영상을 이용하여 학습 데이터셋을 구축하였으며 100개의 가로수를 대상으로 실험하였다. 구축된 영상에서 밝기값 조절, 픽셀 크기 조절, 인위적 노이즈 추가에 대한 인위적 영상품질 항목을 설정하여 영상품질별 폐색영역 검출 능력을 분석하였다. 분석방법은 폐색영역 객체의 검출의 정확성(Precision)과 지

정객체의 검출비율(Recall)을 산출하였으며, 이와 더불어 영상의 검출능력을 나타내는 F-measure 방식을 적용하여 정확성과 검출비율에 대한 종합적인 분석을 실시하였다. 영상품질별 딥러닝 알고리즘 적용 폐색영역 검출 능력 분석 결과는 다음과 같다.

첫째, 영상 밝기값 조절을 통한 영상품질은 어두운 영상보다 밝은 영상에서 정확성과 검출비율이 떨어짐을 확인하였다. 원본영상을 기준으로 인위적인 밝기값을 조절한 결과 2.0에서 검출비율이 0.56으로 떨어져 가장 낮은 검출비율이 산출되었다. 이는 딥러닝 처리에 필요한 특징점 추출이 밝기값이 너무 강한 영상에서는 특징점 추출이 어렵기 때문이다.

둘째, 픽셀 크기 조절을 통한 영상품질은 원본영상에서 0.5 이하일 경우에 급격히 낮은 검출비율을 보이고 있다. 또한 특이한 점은 픽셀 크기에 따른 딥러닝 처리 결과는 가로수에 대한 검출 정확도는 비교적 우수하나, 가로수의 지정 객체 검출 비율은 낮은 현상을 보이고 있다. 이는 픽셀 크기에 따라 가로수라는 검출은 가능하지만 지정된 가로수의 검출능력은 떨어짐을 의미한다.

셋째, 인위적 노이즈 추가를 통한 영상품질은 원본영상의 0.5 이하에서 급격하게 정확도와 검출비율이 떨어짐을 확인하였다. 이는 정확도와 검출비율이 같은 양상을 보이고 있으며, 노이즈 때문에 가로수 검출 능력과 학습된 지정 객체인 가로수의 검출비율도 떨어짐을 의미한다.

딥러닝 폐색영역 검출에 가장 많은 영향을 미치는 요소는 노이즈로 검출능력의 정확성과 검출비율에 동시에 영향을 미치는 영향 요소임을 확인하였다. 또한, 픽셀크기는 검출비율에는 직접적인 영향을 미치지않지만, 상대적으로 검출 정확도에는 영향을 적게 미쳤다. 마지막으로 밝기값은 가장 적은 영향을 미치는 요소임을 확인하였다. 본 연구는 제한된 영상과 한정적인 객체를 이용하여 영상품질별 검출능력을 분석하였다. 영상기반의 딥러닝에 관한 기초자료 및 학습 데이터의 기준이 부족한 시점에서 향후 딥러닝 학습데이터의 기준을 제시하였다. 향후 다양한 학

습 데이터셋과 딥러닝 알고리즘을 적용한다면 공간정보 기술개발에 많은 기여가 예상된다.

KAGIS

REFERENCES

- Choi, H.S. and E.M. Kim. 2015. Detection of road signs region and recognition of directional information. Proceeding of Korean Society for Geospatial Information Science 197-198 (최한승, 김의명. 2015. 도로표지영역 검출 및 방향정보 인식. 한국지형공간정보학회 추계학술대회:197-198).
- Girshick, R. 2015. Fast R-CNN. Proceedings of the IEEE international conference on computer vision 1440-1448.
- Girshick, R., J. Donahue, T. Darrell and J. Malik. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition 580-587.
- Hariharan, B., P. Arbelaez, R. Girshick and J. Malik. 2014. Simultaneous detection and segmentation. In European Conference on Computer Vision 297-312..
- Hinton, G.E. and R.R. Salakhutdinov. 2006. Reducing the Dimensionality of Data with Neural Networks. science 313(5786):504-507.
- Kim, J.K. and J.W. Choi. 2017. Object Detection from Camera, Lidar Fusion based on Deep Learning, Proceeding of Communications and Networks 270-271 (김재겸, 최준운. 2017. 딥러닝 기반 카메라, 라이다 센서 융합 인지기술. 한국통신학회 하계종합학술발표대회:270-271).

- Krizhevsky, A., I. Sutskever and G.E. Hinton. 2012. ImageNet Classification with Deep Convolutional Neural Networks. In Advances in neural information processing systems 1097–1105.
- Kwon, S.I. and E.M. Kim. 2019. Recognition of Flat Type Signboard using Deep Learning. Journal of The Korean Society of Surveying, Geodesy, Photogrammetry, and Cartography 37(4):219–231 (권상일, 김의명. 2019. 딥러닝을 이용한 판류형 간판의 인식. 한국측량학회 37(4):219–231).
- LeCun, Y., Y. Bengio and G. Hinton. 2015. Deep learning. nature, 521(7553):436.
- Lowe, D.G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. International journal of computer vision, 60(2):91–110.
- Mees, O., A. Eitel and W. Burgard. 2016. Choosing Smartly: Adaptive Multimodal Fusion for Object Detection in Changing Enviroments. 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems 151–156.
- Park, S.J., S.M. Choi, h.j. Lee and j.b. Kim. 2016. Spatial analysis using R based Deep Learnin. Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology. 6(4):1–8 (박세정, 최성묵, 이흥재, 김종배. 2017. R 기반의 딥 러닝을 이용한 공간 정보 분석. 예 술 인 문 사 회 융 합 멀 티 미 디 어 논 문 지 . 6(4):1–8).
- Ren, S., K. He, R. Girshick and J. Sun. 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems 91–99.
- Sa Gong, H.S. and S.Y. Lim. 2018. Digitla Twin Spatial Construction Strategy Leading 4th industry. KRIHS POLICY BRIEF (661):1–6 (사공호상, 임시영. 2018. 4차 산업혁명을 견인하는 디지털 트윈 공간 (DTS) 구축 전략, 국토정책Brief (661):1–6).
- Yoo, E.J. and D.C. Lee. 2016. Detection and Recovery of Occlusion Areas caused by Building Sidewalls in Aerial Photos. Proceeding of The Korean Society of Surveying, Geodesy, Photogrammetry, and Cartography 2016(4):156–158 (유은진, 이동천. 2016. 건물벽면에 의해 항공영상에 발생하는 폐색지역 탐색과 복원. 한국측량학회 학술대회 (4):156–158).
- Zhou, B., A. Lapedriza, J. Xiao, A. Torralba and A. Oliva. 2014. Learning deep features for scene recognition using places database. In Advances in neural information processing systems 487–495. **KAGIS**