

R-CNN 기법을 이용한 건물 벽 폐색영역 추출 적용 연구

Application Research on Obstruction Area Detection of Building Wall using R-CNN Technique

김혜진* · 이정민** · 배경호*** · 어양담****
Hye Jin Kim · Jeong Min Lee · Bae Kyoung Ho · Yang Dam Eo

Abstract

For constructing three-dimensional (3D) spatial information occlusion region problem arises in the process of taking the texture of the building. In order to solve this problem, it is necessary to investigate the automation method to automatically recognize the occlusion region, issue it, and automatically complement the texture. In fact there are occasions when it is possible to generate a very large number of structures and occlusion, so alternatives to overcome are being considered. In this study, we attempt to apply an approach to automatically create an occlusion region based on learning by patterning the blocked region using the recently emerging deep learning algorithm. Experiment to see the performance automatic detection of people, banners, vehicles, and traffic lights that cause occlusion in building walls using two advanced algorithms of Convolutional Neural Network (CNN) technique, Faster Region-based Convolutional Neural Network (R-CNN) and Mask R-CNN. And the results of the automatic detection by learning the banners in the pre-learned model of the Mask R-CNN method were found to be excellent.

Keywords: 3D Spatial information, Occlusion area, Deep-learning, R-CNN

1. 서 론

공간정보의 활용 극대화를 위해서는 최신 공간정보에 기반하여 다양한 수요처 및 사용자 요구에 부응하는 3차원 공간정보가 구축되어야 한다(조명흠 외

2011; 박지혜·서용철, 2017). 3차원 공간정보는 지적 정보·지하시설물도·건축물 대장 등 다른 정보들과 연계되어 다양한 활용 업무를 창출하고 국가공간정보인프라로 활용될 수 있다. 또한 디지털 트윈 국토 실현의 일환으로 영상정보와 3차원 국토를 재현함으로써 지

* 건국대학교 신기술융합학과 석사과정 Department of Advanced Technology Fusion, Konkuk University (first author: merrybingo@konkuk.ac.kr)

** (주)신한항공 연구소 연구원 Shinhan Aerial Survey Research Institute (insomnie08@gmail.com)

*** (주)신한항공 연구소 연구소장 Shinhan Aerial Survey Research Institute (qpandora@hanmail.net)

**** 건국대학교 신기술융합학과 교수 Department of Advanced Technology Fusion, Konkuk University (corresponding author: eoandrew@konkuk.ac.kr)



Figure 1. In case of a occlusion area (a) an aerial photograph (b) a ground photograph
Source: Naver Aerial View and Street View

적도, 생태자연도, 도로망도 등 다양한 주제의 공간정보를 구축, 관리하여 지적, 도시계획, 환경, 도로, 건설, 재난재해 등의 행정업무의 기반자료로 활용할 수 있다(Kim et al. 2010).

3차원 공간정보 중 빌딩의 경우 Mobile Mapping System (MMS), 항공사진 및 드론에 카메라를 탑재하여 다수의 영상을 취득하고, 취득된 영상의 기하학적 관계를 구현하여 3차원 모델링 방법으로 제작하고 있으나(Remondino et al. 2011; 이준상·이임진, 2018). 복잡한 시가지의 경우 벽면 텍스처를 추출하는데 폐색영역 문제가 발생하는 경우가 많아 이를 보정하는데 시간적 경제적 손실이 발생하고 있다.

이를 해결하기 위해선 폐색영역을 자동 인식하여 이를 검출하고 텍스처를 자동 보완하는 자동화기법 연구가 필요하다. 최근 대두되고 있는 딥러닝 기반의 알고리즘을 이용하는 경우, 고정밀 데이터를 이용한 폐색지역 패턴화하고, 학습기반의 폐색영역을 자동 검출하는 대안이 될 수 있다.

본 연구는 최근 영상 내 객체 추출에서 우수한 성과를 발표하는 Region-based Convolutional Neural Network (R-CNN) 기법의 향상된 알고리즘인 Faster R-CNN과 Mask R-CNN 2가지를 공간벽면 텍스처 추출이 가능한 지상영상에 적용하여 폐색영역 추출 성능을 분석하고 Mask R-CNN 방법의 미리 학

습된 모델에 현수막, 가로수를 추가로 학습시켜 폐색영역 추출 성능을 분석하였다.

2. 본 론

2.1. 지상사진의 폐색영역 검출

건물 등의 객체에 대한 3차원 공간정보 구축의 가장 일반적인 방법은 객체를 대상으로 영상정보를 획득하여 영상을 건물 벽면에 텍스처링하는 방법이다. 하지만 객체에 대한 3차원 공간정보를 구축 과정에서 다양한 원인으로 폐색영역이 발생한다(김성수·김병국, 2002; 신도경 외, 2011). 폐색영역은 그림 1(a)과 같이 항공사진의 경우 중심투영으로 촬영된 영상에서 객체가 멀어짐에 따라 발생하게 되고 그림 1(b)와 같이 지상사진 촬영 시 건물에 가려진 벽면, 기복변위로 인한 그림자, 가로수 등으로 인해 발생하게 되며 주로 밀집된 도심지에서 나타난다.

기존 폐색영역의 탐지와 복원관련 연구는 3D 모델링을 수행했을 때 건물의 벽면 등 폐색영역이 발생한 곳의 텍스처 취득이 어려우므로 여러 중첩 촬영한 영상을 모자이크 제작하여 폐색영역을 해결하는 방안을 연구하였다, 또한 지상에서 촬영한 영상을 사용하여 폐색영역 부분의 텍스처링을 진행하는 방안과 Unmanned Aerial Vehicle (UAV) 영상에서 획득한 텍스처 정보를 이용하여 3D 모델의 건물 형태를 텍스처링하는 방안 등에 대한 연구가 수행되어 왔다(정성혁·이재기, 2008; Xiao et al. 2009; Gruen et al. 2013). 또한 지상사진에서 폐색을 유발하는 객체 검출 관련 연구로 자율주행 자동차를 위한 데이터 셋인 KITTI 데이터 셋과 CNN 기법을 활용하여 운전 차량으로부터 각 방향으로 가장 가까운 장애물을 검출하는 연구가 수행되었다. 그림 2는 지상사진으로부터 폐색을 유발하는 객체를 탐지하고 이것의 영상 내 경계선을 빨간색 점으로 표시한 결과이다(Levi et al. 2015).



Figure 2. Obstacle detection results
Source: Levi et al. 2015

그런데, 이것은 기복이나 구조물 고도에 의해 생기는 지표상의 폐색영역을 추출해내는 것으로서 대부분 중복 촬영된 영상으로부터 폐색영역을 복원할 수 있다는 가정이 있다. 현실적으로 지상에서 촬영된 지상 사진의 경우 차량이나 가로수, 보행자, 플래카드 등에 의해서 건물의 벽면이 폐색될 수 있고, 이것은 건물과 매우 가까운 거리 폐색유발 요인이 존재함으로써 복원이 매우 어렵다. 폐색영역 검출은 정확도가 복원영역에 영향을 미치므로 3차원 공간정보 구축에 큰 비중을 차지하는 중요한 단계라 할 수 있다. 그림 3(a)는 최근 자동 폐색영역 탐지에서 다량의 장애물이 구하고자 하는 텍스처를 가리고 있으므로 장애물 객체를 먼저 인식하고 그림 3(b)와 같이 폐색영역 처리를 한 후 그림 3(c)와 같이 목표표 하는 구조물 벽면 텍스처가 폐색영역을 자동복원 처리하는 연구가 진행된 바 있다(Hammoudi et al. 2012).

2.2. 학습기반 객체탐지 기법

객체탐지란 제공된 영상에서 사람, 얼굴 등과 같이 찾고자 하는 객체를 확인하여 위치와 크기를 탐지하는 것이다(Papageorgiou et al. 1998). 주로 중심투영으로 인하여 발생하는 폐색영역을 찾기 위해서는 폐색요인을 정확하게 분석해야 하며, 불규칙한 형태로 나타나는 폐색영역을 탐지하기 위해서는 학습기반의 방법이 효과적으로 보인다(김혜진 외, 2009).

현재 학습기반 객체탐지를 수행하기 위하여 객체의

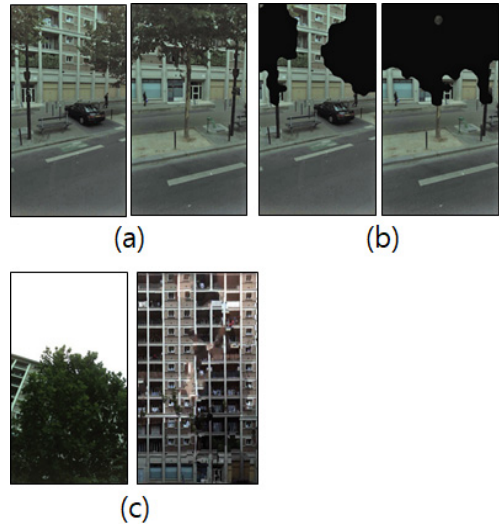


Figure 3. Occlusion area restoration process (a) Occlusion areas of building wall (b) Masking of occlusion areas (c) Restored building wall texture

Source: Hammoudi et al. 2012

특징을 추출하는 알고리즘으로 영상에서의 영역과 영역 간의 밝기차를 이용하는 Haar-like feature, 각 픽셀의 방향성에 대한 히스토그램을 이용한 HOG (Histogram of Oriented Gradients)와 객체를 분류시키는 알고리즘으로 약한 분류기(weak classifier)를 선형적으로 결합하여 강한 분류기(strong classifier)로 분류 능력을 향상시키는 AdaBoost, 선형 분류 모델을 만든 후에 이진 분류를 수행하여 문제를 해결하는 Support Vector Machine (SVM) 등이 사용되고 있다 (Ho et al. 2009; Sivaraman and Trivedi, 2013; Sivaraman and Trivedi, 2014).

기존의 방법은 객체 특징을 추출하기 위하여 입력 데이터 제작하기 위한 비용과 객체 분류시키기 위한 학습데이터 연산 비용이 많이 발생하는 단점이 있다. 이러한 비용을 줄이기 위해 인공지능명 중 하나인 CNN은 영상에서의 특징을 추출한 convolutional layer와 fully connected layer로 객체를 검출하고 영

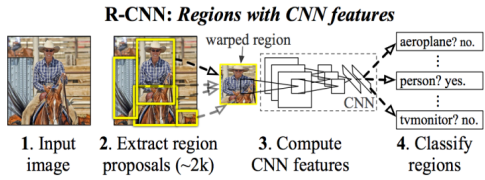


Figure 4. Object detection system overview of R-CNN algorithm

Source: Girshick et al. 2014

상을 대상으로 객체탐지에 향상된 성능을 보여준다 (Simard et al. 2003). R-CNN은 영상에서 후보 영역을 제안하고 이를 기반으로 학습하며 CNN 연산에서 얻은 특징으로 SVM과 Linear regression로 객체 검출을 한다. R-CNN은 입력 영상으로부터 Selective Search 방법으로 객체가 추정되는 곳에 약 2000개의 후보영역을 만들고 후보영역에서 영상의 색상이나 텍스처 등의 정보뿐만 아니라 영상 내의 계층 구조를 같이 활용하여 객체의 특징을 추출한다.

그림 4는 R-CNN 기법의 객체탐지 방법을 설명한 그림이다. Girshick et al. (2014)는 이러한 방식을 적용함에 따라 객체 검출에 있어 SIFT나 HOG 기반의 알고리즘보다 좋은 결과가 나타난다고 하였다.

본 연구에서는 Faster R-CNN과 Mask R-CNN 알고리즘을 도입하여 객체탐지를 수행하였다. Faster R-CNN은 영상에서 하나의 관심영역을 만들어 객체탐지를 수행하며, Region Proposal Network(RPN) 단계를 CNN안에 넣어서 R-CNN의 알고리즘의 정확성과 속도 문제를 개선하였다. (그림 5)

CNN을 통과한 feature map에서 sliding window를 이용하여 각 anchor마다 박스의 점수를 계산한다. (그림 6) Mask R-CNN은 Faster R-CNN을 적용 후 분할된 바운딩 박스에서 객체에 해당하는 것을 각 픽셀로 마스킹하는 네트워크를 추가한 것이다. (그림 7) 정확한 픽셀 위치를 추출하기 위해 선형보간법으로 CNN을 통과하여 생기는 오차를 감소시켰다(Ren et al. 2015; He et al. 2017).

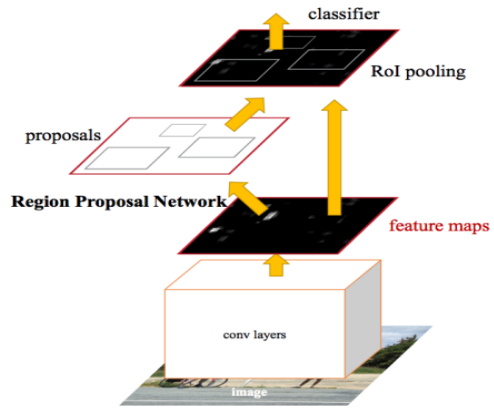


Figure 5. Object detection system overview of Faster R-CNN algorithm

Source: Ren et al. 2015

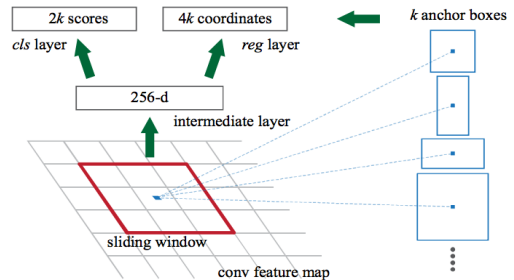


Figure 6. Region proposal network (RPN)

Source: Ren et al. 2015

3. 실험 및 분석

3.1. 실험영상과 실험방법

본 실험은 스트리트 뷰 영상에서 건물의 폐색지역을 탐지하기 위해 Faster R-CNN 방법과 Mask R-CNN 방법을 사용하고 비교 분석한다. 실험에 사용된 영상은 도심지의 신호등, 차량, 가로수, 현수막과 같은 장애물로 건물에 폐색지역이 발생한 곳을 선정하였고 구글 스트리트 뷰(<https://www.instant-streetview.com>)에서 캡처하였다.

faster R-CNN 방법의 소스코드는 tensorflow git

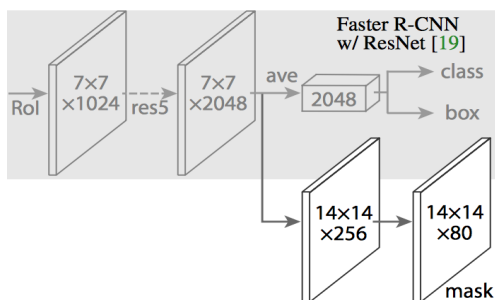


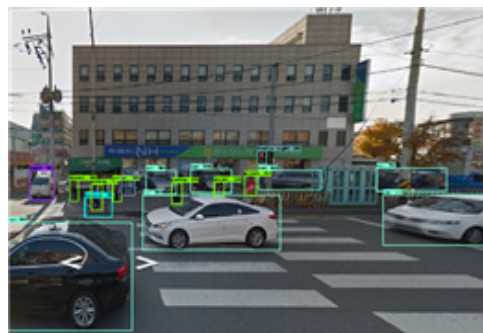
Figure 7. Head architecture of Mask R-CNN.
Source: He et al. 2017

hub 페이지(https://github.com/tensorflow/models/tree/master/research/object_detection)에서 다운받았고 본 연구를 위해 학습된 모델을 사용하고 탐지 결과를 저장하도록 수정하였다. mask R-CNN 방법은 Python3에서 Keras와 Tensorflow로 구현한 소스코드를 기반으로 (https://github.com/matterport/Mask_RCNN) input 영상 경로를 설정하고 탐지 결과를 저장하도록 수정 활용하였다.

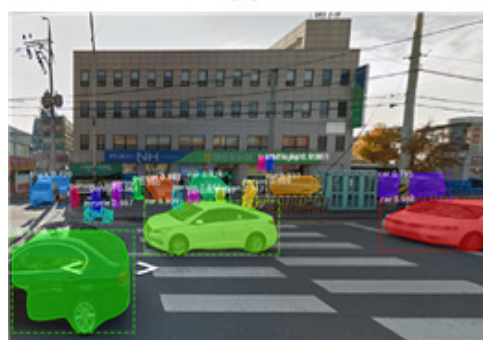
객체 탐지 모델은 미리 학습된 COCO 모델을 사용하였다. (https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md) COCO는 대규모 객체 탐지, 분할 및 캡션 데이터 셋으로 객체 세분화, 컨텍스트 인식, 330K 이미지 (> 200K 라벨), 80개의 객체 카테고리 등의 기능을 가지고 있다. faster R-CNN의 모델은 다른 모델에 비해 속도가 빠르고 정확도가 높은 faster_rcnn_nas_coco로 선정하였다. mask R-CNN의 모델은 matterport github 페이지(https://github.com/matterport/Mask_RCNN/releases)에서 제공되는 mask_rcnn_coco.h5를 사용하였다.

3.2. 폐색영역탐지 결과분석

미리 학습된 모델로 검출한 결과 신호등, 차량, 오토바이, 자전거, 사람 항목은 탐지 결과가 좋았다. 그러



(a)



(b)

Figure 8. Object detection results (a) Faster R-CNN (b) Mask R-CNN

나 가로수나 현수막은 대부분 탐지하지 못했다. 그림 8~11은 폐색영역을 검출한 결과이다.

그림 8과 그림 9의 검출결과를 보면 지상에서의 폐색을 발생시키는 신호등, 차량, 사람을 탐지한 경우이다. 그러나 그림 10과 그림 11의 검출결과를 보면 가로수나 현수막이 대부분의 폐색을 발생시키나 탐지하지 못하는 경우이다. coco 데이터 셋에는 'potted plant'라는 항목이 미리 학습되어있어 가로수가 화분으로 탐지 되었으나 20장 중 1장이 탐지되는 등 검출율이 좋지 않았다. 현수막은 미리 학습 되지 않은 항목으로 탐지되지 않았다. 가로수나 현수막 등 폐색을 발생시키나 기존 모델에 학습되어있지 않은 항목을 학습시킬 필요성을 느꼈다. 따라서 기존 모델에 현수막 데이터를 학습시키고 탐지하려고 하였다.

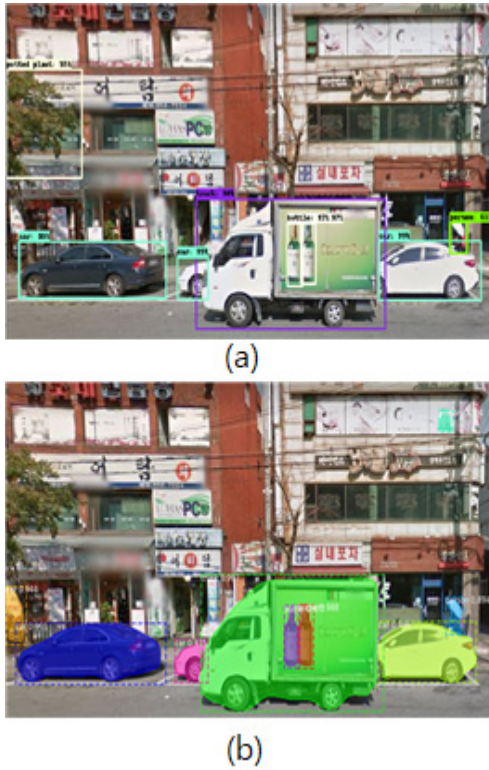


Figure 9. Object detection results (a) Faster R-CNN (b) Mask R-CNN

3.3. 학습영상과 학습방법

새로운 항목을 학습시키기 위해서는 두 가지 과정이 필요하다. 첫째로 학습시킬 충분한 영상 데이터를 구축하고 둘째로 구축된 영상 별 라벨링 작업을 수행해야 한다. coco 데이터 셋에는 현수막 항목이 포함되어 있지 않지만 다른 이미지 (~120k)가 많이 포함되어 있으므로 훈련된 가중치는 자연 이미지에서 흔히 볼 수 있는 많은 기능을 이미 배웠다. 따라서 작은 데이터 셋만으로 충분하다. 학습시킬 영상 데이터는 현수막이 포함된 영상 225장과 가로수가 포함된 영상 225장을 준비하였다. 학습 영상 개수 별 검출율과 정확도의 변화를 알아보기 위하여 학습 세트(training set)를

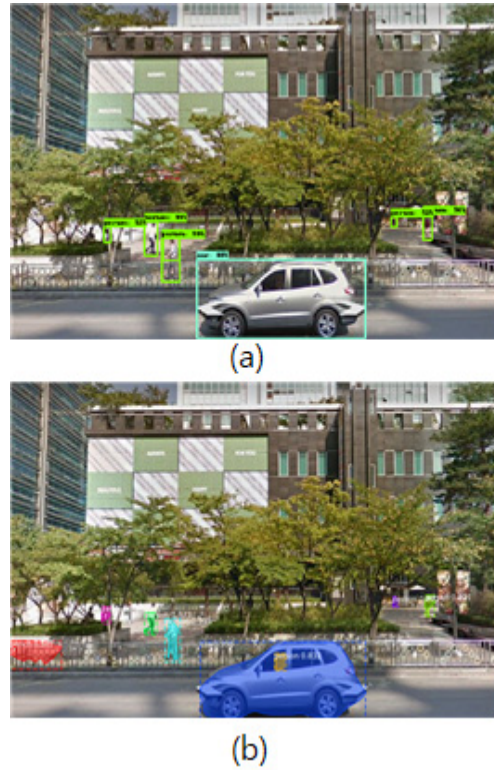


Figure 10. Object detection results (a) Faster R-CNN (b) Mask R-CNN

5세트(10장, 50장, 100장, 150장, 200장)로 나누었고 검증 세트(validation set)로 25장을 사용하였다. 검증 세트는 모델의 성능을 평가하기 위해서 필요하다. 정확도를 측정하고 모델을 튜닝하며, 오버피팅을 막는 역할을 한다. 다음으로 이미지에 주석을 달기 위한 라벨링 작업이 필요하다. 그림 12는 VIA(VGG Image Annotator) Tool을 사용하여 이미지에서 현수막 영역을 표시하는 과정의 캡처 화면이다. 영역은 Polygon 점들로 기록되며 Json 파일 형태로 저장된다.

학습시키는 코드는 위에서 받은 matterport github 소스코드에 포함되어있다. 실험 환경은 표 1과 같은 환경에서 이루어 졌으며, 준비한 데이터의 경로 설정 후 epoch값은 30으로 Steps per epoch값은 100으로



(a)



(b)

Figure 11. Object detection results (a) Faster R-CNN (b) Mask R-CNN

설정하고 학습을 수행했다. 한 세트의 학습시간은 평균 36시간이 소요됐다.

3.4. 학습모델 결과분석

그림 13~16은 Mask R-CNN 방법의 미리 학습된 coco 모델에 현수막과 가로수를 학습시켜 객체를 검출한 결과이다.

그림 13(b) 검출 결과를 보면 현수막 열한 개를, 그림 14(b) 검출 결과에서는 현수막 다섯 개를 정확히 탐지하는 것을 볼 수 있다. 그림 15(b) 검출 결과에서 가로수 여섯 개 중 다섯 개를 탐지하고 그림 16(b) 검출 결과에서는 가로수 네 개중 세 개의 가로수를 탐지

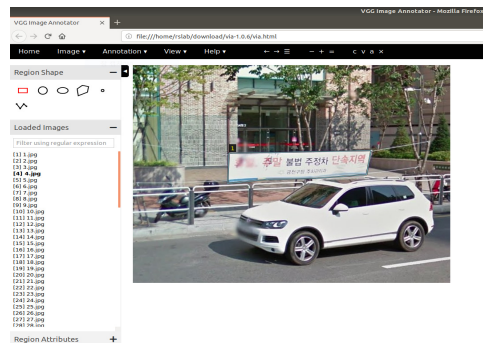


Figure 12. Image labeling using the VGG image annotator tool

Table 1. Experimental environment.

Operating System	Ubuntu 16.04 64bit
CPU	Intel core i7-7700 @3.60GHz
RAM	16GB
GPU	Geforce GTX 1050 Ti 4GB
Language	Python 3.6
Toolkit	Keras, Scikit-image, OpenCV2

했으나 탐지 영역이 정확하지 않는 것을 볼 수 있고 두 개를 하나의 객체로 탐지하는 것을 볼 수 있다. 객체의 형태가 일정한 현수막 경우에는 좋은 탐지 결과를 보였으나 객체의 형태가 일정하지 않고 여러 객체가 겹치는 가로수의 경우에는 탐지결과가 좋지 못했다.

학습한 모델의 성능평가를 위해 precision-recall을 측정하였다. precision은 검출된 결과가 얼마나 정확 한지를 나타낸다. 즉, 검출된 결과들 중 실제 물체가 얼마나 포함되어 있는지를 나타낸다. recall은 대상 물체들을 빠뜨리지 않고 잡아내는 정도를 나타내며 통계적 분류 분야에서 precision과 recall은 다음 식(1), (2)와 같다.

$$precision = \frac{tp}{tp + fp} \quad (1)$$



(a)



(b)

Figure 13. Banner detection results (a) Input Image (b) Output Image

$$recall = \frac{tp}{tp + fn} \quad (2)$$

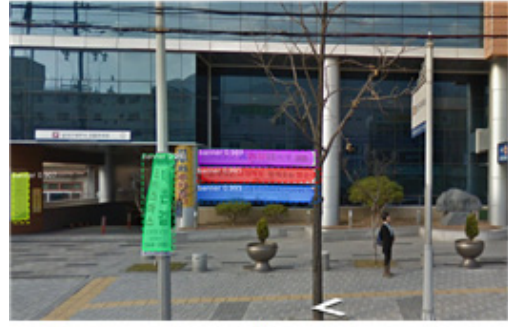
여기서 tp는 실제 정답이면서 실험 결과가 참으로 나온 경우이고, fn는 실제 정답이나 실험 결과가 거짓으로 나온 경우이다. 그리고 fp는 실제 정답이 아니나 실험 결과가 참으로 나온 경우를 말한다. 표 2는 위의 정의를 표로 나타낸 것이다.

Precision과 recall을 계산하기 위하여 tp는 실제 정답을 제대로 탐지한 경우를, fn은 실제 정답이나 탐지하지 못한 경우 그리고 fp는 다른 항목으로 오 탐지한 경우를 썼다.

그림 17은 현수막 학습 영상개수 별 검출율과 정확도를 측정하여 그래프로 나타낸 것이다. 현수막 탐지



(a)



(b)

Figure 14. Banner detection results (a) Input Image (b) Output Image

성능 평가를 위해서 학습과 검증 용도의 225장 이외의 20장의 영상을 사용하였고 총 67개의 현수막이 포함되어있다. 검출율을 높이기 위하여 학습 시킬 때 세로 모양의 현수막이나 다양한 크기의 영상을 학습시켰고 학습 영상 개수가 많아질수록 검출율이 증가되는 것을 볼 수 있었다. 현수막의 객체 탐지는 0.78~0.93 사이의 높은 정확도를 보였다. 100개 이상 학습 영상이 많아질수록 정확도가 떨어지는 것을 볼 수 있는데 이는 현수막과 형태가 비슷한 간판을 탐지하는 경우가 증가했기 때문이다. 현수막과 간판을 다른 클래스를 분리하고 같이 학습시켜야 현수막의 정확도가 높아질 것으로 보인다.



(a)



(b)

Figure 15. Tree detection results (a) Input Image (b) Output Image



(a)



(b)

Figure 16. Tree detection results (a) Input Image (b) Output Image

Table 2. Definition of precision and recall.

Source: Powers and David 2011

		true condition	
		positive	negative
predicted condition	positive	True Positive (tp)	False Positive (fp)
	negative	False Negative (fn)	True Negative (tn)

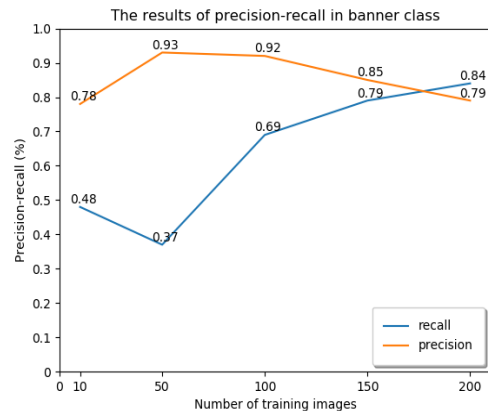


Figure 17. The results of precision-recall in banner class

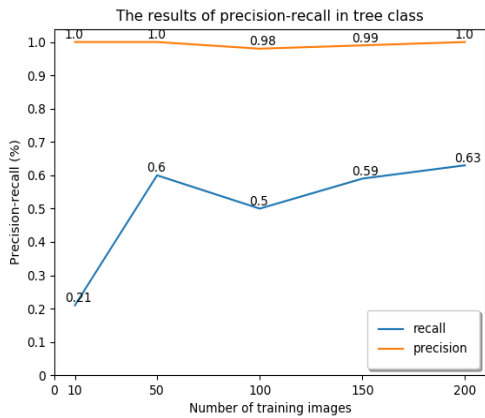


Figure 18. The results of precision-recall in Tree class

그림 18은 가로수 학습 영상개수 별 검출율과 정확도를 측정하여 그래프로 나타낸 것이다. 현수막과 마찬가지로 가로수 탐지 성능 평가를 위해서 학습과 검증 용도의 225장 이외의 20장의 영상을 사용하였고 총 112개의 가로수가 포함되어있다. 현수막 탐지 결과와는 다르게 가로수는 오탐지가 거의 발생하지 않아 정확도가 0.98~1 사이로 측정되었으나 탐지율이 최고 0.63으로 높지 않았다.

4. 결론

학습기반 폐색영역 자동 검출은 패턴화, 추출화를 거쳐서 모의영상 제작으로 이행된다.

본 실험에서는 Faster R-CNN방법과 Mask R-CNN방법으로 폐색영역 탐지를 수행했고, 현수막과 가로수를 추가로 학습시켜 건물의 폐색영역을 탐지하고 마스크를 생성함으로써 자동적으로 폐색영역을 처리할 수 있다는 것을 확인했다. 실제로 현수막과 가로수 이외에도 전봇대, 공중전화부스 등 지상에서 건물에 폐색을 유발하는 항목은 다양하다. Mask R-CNN 기법으로 형태가 일정한 현수막과 같은 객체탐지는 좋은 결과를 보인 반면 가로수와 같이 형태가 일정하

지 않고 여러 객체가 겹쳐있는 경우는 좋지 않는 결과를 보였다.

본 연구의 활용 방안으로 항구/해안 감시의 선박 탐지나 시각장애인을 위한 물체인식 등을 제안할 수 있다. 향후 연구에서는 고품질 공간정보 데이터 생성에 활용할 수 있도록 Semantic segmentation 기법을 활용하여 가로수에 의해 발생하는 폐색영역을 자동탐지하고 폐색영역으로 탐지된 영상을 자동으로 복원하는 부분폐색 복원 연구를 지속할 예정이다.

감사의 글

본 연구는 국토교통부 공간정보 기반의 실감형 콘텐츠 융복합 및 혼합현실 제공기술 연구개발사업의 연구비지원(18DRMS-B147287-01)에 의해 수행되었습니다. 또한 이 논문은 국토교통부의 스마트시티 석·박사과정 지원 사업으로 지원되었습니다.

참고문헌

References

- 김성수, 김병국. 2002. 3D 가상도시 구축을 위한 건물 텍스처 이미지의 왜곡보정. 한국지형공간정보학회지. 10(2):49-56.
- Kim SS, Kim BG. 2002. Adjustment of texture image for construction of a 3D virtual city. *Journal of the Korean Society for Geospatial Information Science*. 10(2):49-56.
- 김혜진, 한유경, 최재완, 김용일. 2009. 건물 그림자와 폐색 보정을 통한 고해상도 위성영상의 분류정확도 향상. 대한원격탐사학회. 대한원격탐사학회 2009년도 춘계학술대회 논문집(Mar):13-17.
- Kim HJ, Han YK, Choi JW, Kim YI. 2009. High resolution satellite image classification enhancement using restoration of building shadow

- and occlusion. The Korean Society of Remote Sensing, Proceedings of the KSRS Conference (Mar):13-17.
- 박지혜, 서용철. 2017. 3차원 공간정보 오픈 플랫폼을 활용한 일조 시간 산정. 한국지리정보학회지. 20(3):80-89.
- Park JH, Suh YC. 2017. Calculation of the Duration of Sunshine Using a Three-Dimensional Spatial Information Open Platform. *Journal of the Korean Association of Geographic Information Studies*. 20(3):80-89.
- 신도경, 박찬우, 박진욱, 김용민, 박기태, 문영식. 2011. 폐색영역 보정을 위한 변위 추정을 이용한 특징 기반 깊이 맵 생성 기법. 한국정보기술학회논문지. 9(5):217-230.
- Shin DK, Park CW, Park JW, Kim YM, Park KT, Moon YS. 2011. Feature based Depth Map Generation for Compensation of Occlusion Region using Disparity Estimation. *The Journal of Korean Institute of Information Technology*. 9(5):217-230.
- 이준상, 이임근. 2018. 드론 촬영 기반 사진 스캐닝 기술을 활용한 3D 모델링데이터 생성방법에 관한 연구. 한국정보통신학회논문지. 22(6):874-880.
- Lee JS, Lee IG. 2018. The 3D Modeling Data Production Method Using Drones Photographic Scanning Technology. *Journal of the Korea Institute of Information and Communication Engineering*. 22(6):874-880.
- 정성혁, 이재기. 2008. 디지털 영상자료를 이용한 3D GIS의 사실적 모델링 및 가시화. 한국측량학회. 26(1):73-83.
- Jung SH, Lee JK. 2008. Application of Photo-realistic Modeling and Visualization Using Digital Image Data in 3D GIS. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*. 26(1):73-83.
- 조명흠, 서정택, 이철용, 박영진. 2011. 3차원 공간정보 기반의 가상 시뮬레이션을 통한 피해정보 구현. 한국방재학회. 11(5):175-184.
- Cho MH, Seo JT, Lee CY, Par YJ. 2011. The Realization of Disaster Information using Virtual Simulation based on 3D Spatial Information. *Journal of the Korean Society of Hazard Mitigation*. 11(5):175-184.
- Girshick R, Donahue J, Darrell T, Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition. pp.580-587.
- Gruen A, Huang X, Qin R, Du T, Fang W, Boavida J, Oliveira A. 2013. Joint processing of UAV imagery and terrestrial mobile mapping system data for very high resolution city modeling. *Remote Sens Spatial Inf Sci*. pp.175-182.
- Hammoudi K, Dornaika F, Soheilian B, Vallet B, McDonald J, Paparoditis N. 2012. Recovering occlusion-free textured 3D maps of urban facades by a synergistic use of terrestrial images, 3D point clouds and area-based information. *Procedia Engineering*. 41(2012): 971-980.
- He K, Gkioxari G, Dollár P, Girshick R. 2017. Mask R-CNN. IEEE International Conference on Computer Vision (ICCV). pp.2980-2988.
- Ho WT, Lim HW, Tay YH. 2009. Two-stage license plate detection using gentle Adaboost and SIFT-SVM. *Intelligent Information and Database Systems*. pp.109-114.
- Kim HT, Kim SB, Go JS, Eo YD, Lee BK. 2010.

- Building 3D Geospatial Information using Airborne Multi-Looking Digital Camera System. *Journal of Convergence Information Technology*. 5(1):15-22.
- Levi D, Garnett N, Fetaya E. 2015. StixelNet: A Deep Convolutional Network for Obstacle Detection and Road Segmentation. British Machine Vision Conference 2015. pp.109.1-109.12.
- Papageorgiou CP, Oren M, Poggio T. 1998. A general framework for object detection. Sixth International Conference on Computer Vision. pp.555-562.
- Powers, David MW. 2011. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*. 2(1): 37 - 63.
- Remondino F, Barazzetti L, Nex F, Scaioni M, Sarazzi D. 2011. UAV photogrammetry for mapping and 3D modeling-Current status and future perspectives. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XXXVIII-1(C22): 25-31.
- Ren S, He K, Girshick R, Sun J. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. Neural Information Processing Systems (NIPS). pp.91-99.
- Simard PY, Steinkraus D, Platt JC. 2003. Best practices for convolutional neural networks applied to visual document analysis. Document Analysis and Recognition. p.958.
- Sivaraman S, Trivedi MM. 2013. A review of recent developments in vision-based vehicle detection. Intelligent Vehicles Symposium (IV). pp.310-315.
- Sivaraman S, Trivedi MM. 2014. Active learning for on-road vehicle detection: A comparative study. *Machine vision and applications*. 25(3): 599-611.
- Xiao J, Fang T, Zhao P, Lhuillier M, Quan L. 2009. Image-based street-side city modeling. *In ACM transactions on Graphics (TOG)*. 28(5):114.

2018년 10월 04일 원고접수(Received)

2018년 10월 08일 1차심사(1st Reviewed)

2018년 10월 29일 2차심사(2st Reviewed)

2018년 11월 22일 게재확정(Accepted)

초 록

3차원 공간정보 구축을 위해 건물 텍스처를 촬영하는 과정에서 폐색영역 문제가 발생한다. 이를 해결하기 위해선 폐색영역을 자동 인식하여 이를 검출하고 텍스처를 자동 보완하는 자동화 기법 연구가 필요하다. 현실적으로 매우 다양한 구조물 형상과 폐색을 발생시키는 경우가 있으므로 이를 극복하는 방안들이 고려되고 있다. 본 연구는 최근 대두되고 있는 딥러닝 기반의 알고리즘을 이용하여 폐색지역 패턴화하고, 학습기반 폐색영역 자동 검출하는 접근을 시도한다. 영상 내 객체 추출에서 우수한 성과를 발표하는 Convolutional Neural Network (CNN) 기법의 향상된 알고리즘인 Faster Region-based Convolutional Network (R-CNN)과 Mask R-CNN 2가지를 이용하여, 건물 벽면 촬영 시 폐색을 유발하는 사람, 현수막, 차량, 신호등에 대한 자동 탐지하는 성능을 알아보기 위해 실험하고, Mask R-CNN의 미리 학습된 모델에 현수막을 학습시켜 자동탐지하는 실험을 통해 적용이 높은 결과를 확인할 수 있었다.

주요어 : 3차원 공간정보, 폐색영역, 딥러닝, R-CNN