

ResNet 알고리즘 기반 폐색영역 학습 데이터셋 생성 및 비교 분석

이정민*, 배경호*, 윤중훈*

*㈜신한항업

e-mail:qpandora@daum.net

Occlusion area learning data set creation and comparative analysis based on ResNet algorithm

Jeong-Min Lee*, Kyoung-Ho Bae*, Joong-Hoon Yoon*

*Research Institute, Shinhan Aerial Survey CO.,LTD

요약

본 논문에서는 3D 모델링 텍스처에 가로수, 간판, 차량 등으로 가려지는 폐색영역을 검출하기 위해 학습 영상 개수별 학습 데이터셋 생성하였다. 차량, 간판, 도로표지판 대상으로 ResNet 알고리즘 기반 폐색영역 학습 데이터셋을 생성하였으며 그 결과, 학습 영상 개수가 많아질수록 정확도가 높아지는 경향이 나타났으며 정형화된 객체일수록 학습률이 일정하게 높은 것을 확인할 수 있었다.

검출하기 위해 학습 영상을 수집하여 라벨링 데이터를 만들고 학습 영상과 라벨링 데이터셋으로 ResNet 기반 학습데이터셋을 구축하였다.

차량, 간판, 도로표지판의 학습 영상을 수집하기 위해 객체당 60, 120, 240, 600장의 영상을 구글 스트리트뷰에서 캡처하였으며 총 4세트의 데이터셋을 구성하였다.

[표 1] 객체별 학습데이터셋 영상 개수

구분	영상 개수			
학습 영상 개수	50	100	200	500
검증 영상 개수	5	10	20	50
테스트 영상 개수	5	10	20	50
합계	60	120	240	600

1. 서론

컴퓨터 성능 및 기술의 발전으로 3D 공간정보 구축사업이 활성화되고 있다. 3D 공간정보는 건물, 도로 등의 요소로 이루어져 있으며 주로 시설물관리, 도시관리, 도시환경관리 등에 활용된다. 건물 3D 모델링 구축할 시 텍스처링 데이터에 가로수, 간판, 차량 등의 여러 가지 요소로 건물의 텍스처가 가려지는 폐색영역이 발생한다. 폐색영역을 빠르게 처리하기 위해 학습기반 폐색영역 검출 및 복원 기술이 필요하며 폐색영역 검출을 위한 학습 데이터셋 생성이 우선적으로 필요하다.

주로 영상처리에 사용되는 학습기반 알고리즘은 CNN을 사용한다. CNN은 일반적으로 영상 내의 특징을 추출하기 위한 알고리즘이며, CNN에서 Residual Block 네트워크를 추가하여 빠르고 효율적인 특징 추출 처리가 가능한 ResNet은 특정 객체에 대해 탐지할 수 있는 최적의 알고리즘이다[1].

본 실험에서는 ResNet 기반의 알고리즘을 사용하여 차량, 간판, 도로표지판 학습 데이터셋을 생성하고 영상개수별 데이터셋 테스트 결과를 비교 분석하였다.

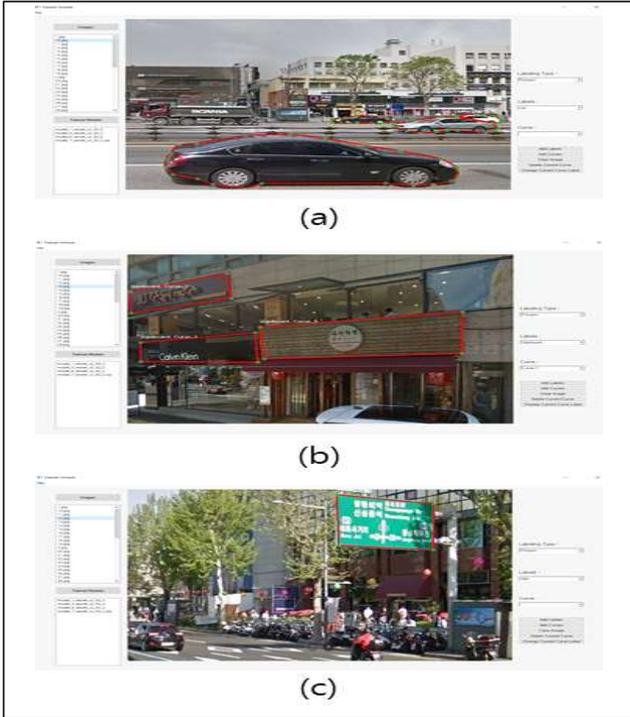
2. 연구방법

본 실험은 폐색영역으로 선정한 차량, 간판, 도로표지판을

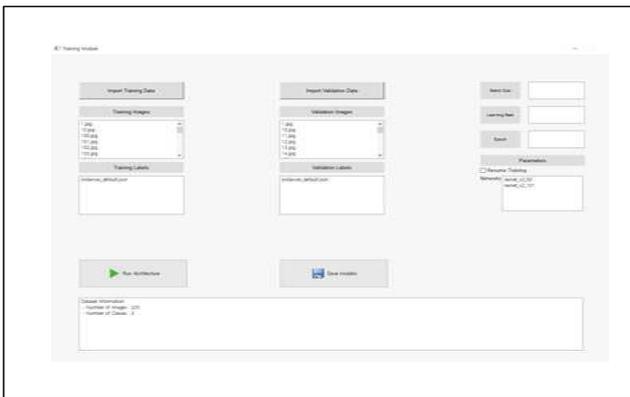
학습시키기 위한 데이터셋을 생성하기 위해 선차적으로 수집한 영상에서 폐색영역으로 해당하는 객체의 레이블을 추가하고 레이블에 해당되는 영역을 폴리곤 형태로 지정하여 주석 파일을 생성해야한다. 데이터셋은 수집한 JPG 포맷의 영상과 폴리곤 형태의 레이블링한 주석 내용으로 이루어진 JSON 포맷의 파일로 생성되며, 본 실험에서는 수집한 구글 스트리트뷰 영상에서 각각 차량, 간판, 도로표지판의 레이블링과 객체별 형태로 직접 폴리곤 형태로 지정하여 데이터셋을 생성하였다.

객체별 JPG 영상과 JSON 파일로 구성된 데이터셋과 검증

데이터셋을 입력데이터로 사용하였고 검증 데이터셋은 데이터셋을 학습할 때마다 학습 데이터가 오버피팅 되는 것을 방지하는 역할을 수행하며 ResNet 기반 객체별 학습데이터셋을 생성하였다.



[그림 1] 객체별 레이블링 데이터셋 구축 (a) 차량 레이블링 데이터셋 구축, (b) 간판 레이블링 데이터셋 구축, (c) 도로표지판 레이블링 데이터셋 구축

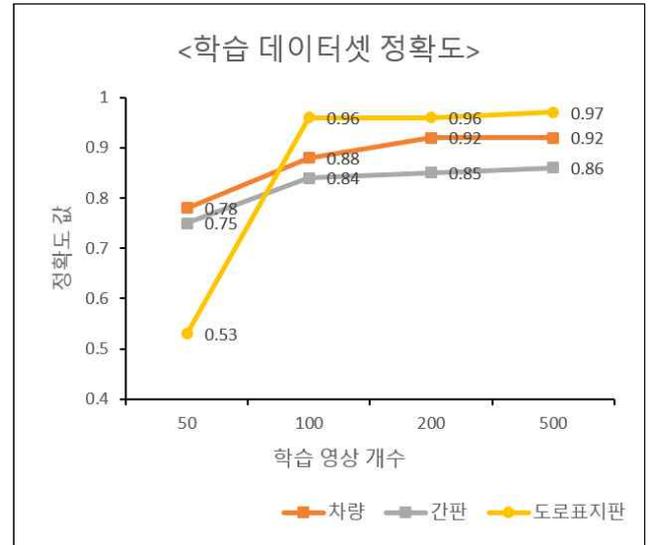


[그림 2] 객체별 학습데이터셋 구축

3. 연구내용 및 분석

실험 결과, 차량, 간판, 도로표지판 학습데이터셋 모두 영상 개수가 많아질수록 정확도가 높아지는 것을 확인할 수 있었다. 도로표지판의 경우 0.97로 가장 높은 정확도를 나타냈으며 이는 도로표지판의 색상 및 형태가 정형화되어있어 학습

률이 일정한 수준으로 높게 나타나는 특성을 확인할 수 있었다. 간판의 경우 가장 낮은 정확도를 도출하였는데 간판의 특성상 여러 색상과 다양한 크기의 형태로 인해 보여지며, 차량의 경우 색상은 다양하지만 정형화된 형태로 간판보다 일정 수준의 높은 정확도를 나타내고 있다.



[그림 3] 학습 데이터셋 객체별 정확도

4. 결론

영상 개수별 폐색영역 학습 데이터셋 구축 정확도를 알아보고자 차량, 간판, 도로표지판 대상으로 학습 데이터셋을 구축하여 비교 분석하였다. 학습 영상이 많아질수록 정확도가 높아졌으며 정형화된 객체일수록 정확한 학습 데이터셋을 구축할 수 있는 것을 확인하였다. 이는 건물 3D 모델링 텍스처에 발생하는 폐색영역을 처리하기 위해 ResNet 알고리즘을 이용하여 수많은 학습 영상을 수집이 필요하고 정형화된 객체뿐만 아니라 비정형 객체의 학습데이터셋을 구축하여 다양한 학습데이터셋 구축 연구가 필요한 것을 알 수 있었다.

감사의글

본 연구는 국토교통부 공간정보 기반의 실감형 콘텐츠 융복합 및 혼합현실 제공기술 개발 연구개발사업의 연구비지원(21DRMS-B147287-04)에 의해 수행되었습니다.

참고문헌

[1] Targ, S., Almeida, D., & Lyman, K., "Resnet in resnet: Generalizing residual architectures". arXiv preprint arXiv:1603.08029, 2016.