

영상 폐색영역 검출 및 해결을 위한 딥러닝 알고리즘 적용 가능성 연구

배경호¹, 박흥기^{2*}

¹(주)신한항공 연구소, ²가천대학교 토목환경공학과

A Study on the Applicability of Deep Learning Algorithm for Detection and Resolving of Occlusion Area

Kyoung-Ho Bae¹, Hong-Gi Park^{2*}

¹Research Institute, Shinhan Aerial Survey CO.,LTD

²Department of Civil & Environmental Engineering, Gachon University

요약 최근 드론을 이용한 공간정보 구축이 활성화되면서 공간정보 산업발전에 많은 기여를 하고 있다. 하지만 드론 공간정보는 카메라의 중심투영에 의한 발생하는 폐색영역 뿐 아니라 가로수, 보행자, 현수막과 같은 적치물에 의한 폐색영역이 필연적으로 발생한다. 이러한 폐색영역을 효율적으로 해결하기 위한 다양한 방안이 연구되고 있다. 본 연구에서는 폐색영역 해결을 위해 원초적인 재촬영이 아닌 딥러닝 알고리즘을 적용하기 위한 다양한 알고리즘별 조사 및 비교 연구를 수행하였다. 그 결과, 객체 검출 알고리즘인 HOG부터 기계학습 방법인 SVM, 딥러닝 방식인 DNN, CNN, RNN 까지 다양한 모델들이 개발 및 적용되고 있으며, 이 중 영상의 분류, 검출에 가장 보편적이고 효율적인 알고리즘은 CNN 기법임을 확인하였다. 향후 AI 기반의 자동 객체 탐지와 분류는 공간정보 분야에서 각광받는 최신 과학기술이다. 이를 위해 다양한 알고리즘에 대한 검토와 적용은 중요하다. 따라서, 본 연구에서 제시하는 알고리즘별 적용 가능성은 자동으로 드론 영상의 폐색영역을 탐지하고 해결할 수 있어 공간정보 구축의 시간, 비용, 인력에 대한 효율성 향상에 기여할 것으로 판단된다.

Abstract Recently, spatial information is being constructed actively based on the images obtained by drones. Because occlusion areas occur due to buildings as well as many obstacles, such as trees, pedestrians, and banners in the urban areas, an efficient way to resolve the problem is necessary. Instead of the traditional way, which replaces the occlusion area with other images obtained at different positions, various models based on deep learning were examined and compared. A comparison of a type of feature descriptor, HOG, to the machine learning-based SVM, deep learning-based DNN, CNN, and RNN showed that the CNN is used broadly to detect and classify objects. Until now, many studies have focused on the development and application of models so that it is impossible to select an optimal model. On the other hand, the upgrade of a deep learning-based detection and classification technique is expected because many researchers have attempted to upgrade the accuracy of the model as well as reduce the computation time. In that case, the procedures for generating spatial information will be changed to detect the occlusion area and replace it with simulated images automatically, and the efficiency of time, cost, and workforce will also be improved.

Keywords : Occlusion Area, Deep Learning, Drone, Image, Detection

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

*Corresponding Author : Hong-Gi Park(Gachon Univ.)

email: hgpark@gachon.ac.kr

Received October 4, 2019

Revised October 31, 2019

Accepted November 1, 2019

Published November 30, 2019

1. 서론

영상 기반의 공간정보 구축은 수치표고모델, 토지피복도 등으로 활용되며 변화탐지, 재해 모니터링 연구 등을 활용되어 핵심적인 공간정보 구축 방법이다. 하지만 이러한 영상은 카메라의 중심투영과 다양한 적치물로 인해 필연적으로 건물의 옆면이나 지붕에 의해 가려지거나 그림자, 가로수 등으로 인해 지표면에 화소값이 부여되지 않는 폐색영역이 발생한다는 점이다. 이러한 폐색영역은 영상에서 방사방향으로 멀어질수록 그 영향이 커지며, 고층건물이나 장애물이 다수 분포하는 도심지역에서 빈번히 발생한다. 최근 좁은 지역에서는 국가기본도, 대축척 수치지도 등 제작에 사용되는 전통적인 항공사진측량 방식의 시간적, 비용적 비효율성 때문에 드론, UAV, MMS 등을 이용한 측량이 도입 및 적용되고 있는데, 작업 상 많은 장점에도 불구하고 도심지역에서의 장애물에 의한 폐색영역 발생 문제가 여전히 해결되어야 하는 문제로 남아있다.

기존 연구에서 폐색영역을 추출 및 보정하기 위한 방법은 대체로 항공 및 위성영상을 대상으로 실시되었다. 기본적인 원리는 수치표고자료와 카메라의 외부표정요소를 계산해 건물 또는 장애물에 의한 폐색영역을 계산한 후 여러 방향에서 촬영한 영상을 사용해 해당 영역의 정보를 추출, 모자이킹하여 채우는 방식이다[1-5]. 그러나 다른 시점에서 촬영한 영상이라도 장애물이나 인접 건물에 2차로 가려져 적절한 대체 영역을 추출할 수 없어 재촬영을 하거나 대상 영역이 복잡해 수작업을 통해 세밀한 공정을 진행해야 하는 경우가 있어 시간, 인력, 비용적인 측면에서 효율성이 떨어진다는 단점이 있다.

특히, 최근에는 모바일매핑시스템, 드론 등을 이용한 실감형 3차원 공간정보 구축이 증가하기 있기 때문에 폐색영역 해결 방안이 필요한 시점이다.

이에 본 연구에서는 최신 기술인 딥러닝을 이용한 객체 학습화 및 패턴화를 실시하여 가로수, 현수막, 차량 등의 폐색을 유발하는 객체를 자동으로 탐지하고 이를 해결하기 위한 딥러닝 알고리즘 현황 분석[6-9]을 실시하였으며 최적의 폐색영역 해결 알고리즘을 제시하고자 하였다.

2. 영상 내 폐색영역 탐지 및 복원기법

2.1 전통적인 탐지 및 복원기법

전통적으로 폐색영역은 카메라의 외부표정요소와 수치표고모델을 이용한 가시권 분석을 통해 검출한다. 각 화소에 대해 방향을 정해 촬영위치로부터 각도를 계산하고, 수치표고 모델로부터 폐색이 발생할 것으로 판단되는 각도 한계를 계산해 그 각도를 벗어나면 폐색지역으로 판단하는 방식이다. 이후 폐색 지역에 대해서는 1차로 화소값을 배정하지 않고, 인접 위치에서 중복도를 고려해 촬영한 부가영상에서 해당 위치가 보인다면 색조정합을 통해 채우는 방식이다. Fig. 1은 전통적인 폐색영역 복원 기법을 나타낸 것으로 기준시점이 되는 L1 위치에서 촬영한 경우 건물과 절벽으로 인해 발생하는 폐색영역 A와 B를 L2 위치에서 촬영한 부가영상으로부터 추출해 채우는 방법이다.

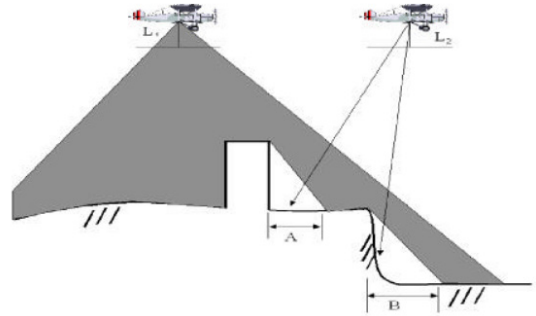


Fig. 1. Occlusion area at the point L1 and its recovery by images obtained by point L2

반면, 지상에서 드론, UAV, MMS 등에 의해 획득된 영상의 경우에는 중심투영에 따른 폐색 뿐 아니라 Fig. 2와 같이 가로수, 신호등, 플랜카드, 차량, 보행자 등 다양한 요소들이 건물의 외벽을 가리는 장애물에 의한 폐색도 발생한다. 이러한 요인들에 따른 폐색 영역은 가려진 영역의 화소를 제거한 후 인접 영상에서 촬영한 텍스처를 추출해 복원하게 된다.



Fig. 2. Occlusion area occurred in downtown area

그러나 앞서 언급한 바와 같이 전통적인 방법에 따른 폐색영역 보원방법은 다른 시점에서 촬영된 영상에서 대체할 수 있는 영역이 있다는 가정 하에 공정과정 중 수작업이 진행되는 부분들이 많아 시간, 인력, 비용적으로 비효율적이다.

2.2 학습기반 탐지 및 복원기법

전통적인 방법의 한계를 극복하기 위해 최근 연구되고 있는 분야는 사람, 차량, 건물 등과 같이 특정한 클래스에 속하는 객체를 영상에서 자동으로 찾는 컴퓨터 비전 기술과 영상처리 기술이 결합된 딥러닝 기반 알고리즘을 적용하는 것이다. 다수의 영상 내에서 검출하고자 하는 특징을 학습하고, 검출기 모델 내 여러 탐지 클래스를 구분해 자동으로 검출하도록 하는 방법으로 효율성이 높다. 특히, 보행자 인식, 제스처 등에 대한 식별 시에도 딥러닝 기술이 사용되는 만큼 형태가 불규칙한 경우에 학습기반의 탐지 기술을 적용할 때 효용성이 더 높을 것으로 기대되고 있다[10,11].

학습 기반으로 객체를 탐지하기 위한 방법은 영상 내 영역 간 밝기차이를 이용해 객체의 특징을 추출 하는 분류기 방식과 이들을 결합해 인공신경망 형태로 엮은 딥러닝 방식으로 나눌 수 있다.

2.2.1 HOG(Histogram of Oriented Gradients)

대상 영역을 일정 크기의 셀로 분할하고, 각 셀마다 변화의 크기가 일정값 이상인 픽셀들의 방향에 대한 히스토그램을 생성하여 히스토그램의 도수를 일렬로 연결한 벡터를 의미한다. 즉, 셀에서 변화가 나타나는 지점을 추출하는 특징 추출 알고리즘이다[12]. 지역적인 변화가 나타나는 경우에는 안정적이라고 평가되고 있으나 기하학적인 정보를 이용하는 것이 아니라 분포에 대한 정보만을 이용하기 때문에 잘못된 대상과 매칭이 되는 경우도 빈번히 발생한다. 특히, 물체가 회전되거나 형태 변화가 심하게 발생하는 경우에는 객체 검출 성공률을 급격히 떨어지는 것으로 평가되어 형태변화가 심하지 않고 패턴이 단순한 경우에 윤곽선을 추출하기 위한 용도로 적합하다.

실제 200장의 영상으로 학습 데이터베이스를 구축한 후 6명으로 20장씩 구성된 테스트 영상을 입력해 가장 유사한 대상자를 찾는 실험을 진행했을 때, 정면이 아니거나 얼굴 좌우대칭이 평행하지 않을 때 검출되지 않는다는 결과가 발표된 바 있다[13].

2.2.2 AdaBoost

적응형(adaptive)와 가속(boosting)이라는 의미의 조합으로 생성된 AdaBoost 알고리즘은 영상처리 분류에 사용되는 기계학습 알고리즘으로 Yoav Freund와 Robert Schapire에 의해 개발되었다. 영상에서 탐지하고자 하는 객체와 닮은 형상적 특징을 학습해 반복적으로 이용하는 방식으로 동일한 샘플로 반복 학습을 통해 가중치를 수정함으로써 분류 정확도를 높인다. 즉, 특징 선택과 관측값에 해당하는 오차율을 산출하고, 가중치를 수정하는 단계로 진행되며, 반복 시 추출되는 새로운 특징인 약한 분류기를 선형적으로 결합함으로써 강한 분류기를 구성하게 된다[14]. 여러 개의 약한 분류기에서 특징별 비교가 가능하기 때문에 이를 결합하게 되면 안정성이 높아진다는 장점이 있다.

Adaboost 방법을 적용해 야간에 운행하는 차량에 대한 검출 실험을 진행하였을 때 거리가 멀수록 검출 정확도는 떨어지나 100m 이내이면 91.3%, 거리가 200m까지 늘어나면 81.7%로 나타나 80% 이상의 검출률이 확보된 것으로 발표된 바 있다[15].

2.2.3 SVM(Support Vector Machine)

SVM은 기계학습 알고리즘으로 패턴 인식, 자료 분석을 위한 모델이며, 분류와 회귀 분석에서 많이 사용된다. 확보된 데이터 군을 기반으로 새로운 자료가 어느 카테고리에 속할지 판단하는 선형분류모델을 생성하는 방식으로 카테고리가 다른 자료들을 가장 큰 마진(margin)으로 분리해내는 초평면(hyperplane)을 찾아 분류한다 [16,17]. 문제는 카테고리가 분명하지 않은 자료들 또는 이상치들이 포함되었을 때 분리 정확도가 떨어진다는 점이다. 이러한 문제점을 해결하기 위해 고안된 방법 중 대표적인 것이 RBF 커널 SVM으로 Fig. 3에 표현된 바와 같이 자료들을 고차원 특징 공간으로 사상한 후 선형으로 분류하는 방식을 의미한다[18]. 특히, SVM 방식은 자료의 양이 적을 때 우수한 성능을 보이기 때문에 현재도 꾸준히 사용되고 있다.

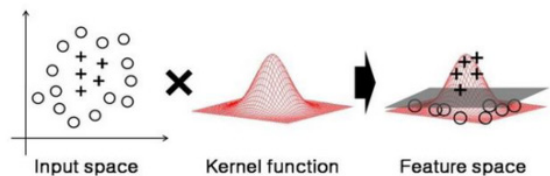


Fig. 3. Example of RBF SVM

임명 등 (2019)은 시뮬레이션을 통한 학습 후 가려진 유무에 따른 인식 정확도 분석결과, 물체가 일부 가려진 상황에서도 87.5%의 인식률을 나타내었다고 평가했으며, Kuo *et al.*(2013)는 형상의 수가 190개 이상일 때 RBF 필터를 사용한 경우와 유사성 비교 알고리즘만 사용한 경우를 비교했을 때 정확도가 0.935와 0.929로 RBF 사용 시 성능이 향상되었다고 발표했다[19,20].

2.2.4 ANN(Artificial Neural Network)

SVM은 처리할 자료의 양이 늘어나거나 그 형태가 불규칙한 경우에 성능이 저하된다. 이에 입력, 은닉, 출력 계층으로 구성된 인공신경망 기술을 구현해 많은 양의 자료를 다룰 수 있는 기술이 개발되었으며, 여기서 신경망이란 학습된 자료로부터 모델을 찾아내는 기법을 의미한다[21]. 다양한 인공신경망 기술 중 ANN은 초기에 개발된 모델로 Fig. 4에 도시된 것과 같이 은닉 계층이 하나이다. 일반적으로 ANN이라는 용어 자체가 인공신경망 자체를 의미하므로 후속 개발된 기술들과의 차별성을 위해 얕은(shallow) 신경망으로 부르기도 한다. 신경망의 성능은 은닉층의 수에 의존하는데, 얕은 신경망으로 불리는 초기 ANN의 경우에는 은닉 계층이 1개이기 때문에 훈련에 소요되는 시간이 오래 걸리며, 수렴 속도 역시 떨어진다 하는 한계가 있다. 이에 최근에는 은닉층의 개수를 늘려 결합한 심층 신경망이 이용되고 있다.

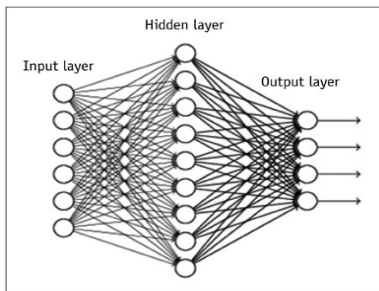


Fig. 4. Networks of ANN

2.2.5 DNN(Deep Neural Network)

심층신경망은 은닉 계층을 쌓아서 만든 기술로 수십개에서 수백개의 은닉 계층으로 구성되기도 한다 (Fig. 5). 이에 따라 은닉 계층이 하나인 ANN에 비해 성능이 우수하며, 적용할 수 있는 분야도 더 넓은 것으로 알려져 있다. ANN에서는 은닉 계층이 1개이므로 은닉계층에서 산출된 결과가 바로 출력되는 형태였으나 심층신경망에서

는 이전 은닉 계층의 결과가 제 2의 은닉계층으로 순차적으로 들어가는 형태로 구성되며, 추가되는 계층들은 점진적으로 하위 계층들의 특징들을 규합할 수 있어 더 정교한 처리가 가능하다. 이에 ANN에 비해 적은 수의 자료 만으로 복잡한 모델링이 가능하다는 장점이 있다 [23,24]. 참고로, ANN이 기존의 기계학습 방법인 반면 은닉층의 개수가 많은 심층신경망부터는 딥러닝이라고 그 용어를 구분한다.

DNN의 경우 은닉층이 많은 대신 계산 시간이 늘어나며, 신경망에 대한 가중치가 늘어나 학습결과가 과적합되는 경우가 나타난다. 이를 해결하기 위해 드롭아웃(drop-out)이나 복합임무 학습(multi-task learning)과 같은 정규화 기법을 사용한다. 드롭아웃이란 다양한 조합의 DNN을 이용함에 있어 하나의 DNN에 과도하게 학습되지 못하도록 네트워크의 일부를 생략하는 방식이며, 복합임무 학습이란 추정하고자 하는 임무 외 상관성이 있는 다른 임무를 추가적으로 학습시켜 원래 추정하고자 했던 임무의 성능을 향상시키는 방법을 의미한다.

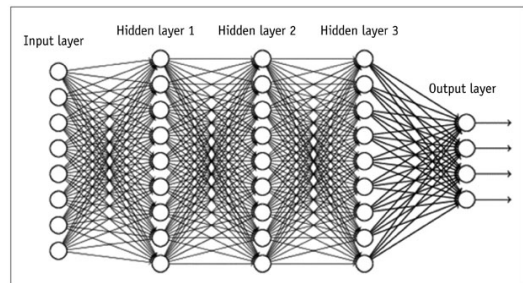


Fig. 5. Networks of DNN

2.2.6 CNN(Convolution Neural Network)

CNN은 하나 또는 여러 개의 컨볼루션 레이어와 인공 신경망 레이어들의 조합으로 이루어진 모델로 Y. LeCun에 의해 처음 소개되었다[25]. 최소한의 전처리를 실시하도록 설계된 것으로 가중치와 통합 계층들을 추가로 활용할 수 있기 때문에 2차원 구조의 입력자료를 활용하기에도 적합하다. 특히, 다중 은닉계층 기반의 심층신경망에서 나타내던 학습 시간, 변수의 개수 등의 비효율성을 개선할 수 있다는 장점이 있다[26]. 영상 분류 분야의 경우, 심층신경망을 사용하면 입력 노드 각각에 화소값을 넣어 학습을 위한 계산이 복잡해지고 커져 비효율적이며, 결과의 신뢰도가 낮아진다는 문제점이 있어 최근 영상과 관련된 분야에서는 이러한 문제점이 보완된 CNN이 가장 널리 사용되고 있다.

CNN 중 영상에 대한 탐지 및 분류에 사용되는 모델은 영역을 기반으로 하는 R-CNN(Region with CNN)이며, R-CNN이 발표된 이후 Fast R-CNN, Faster R-CNN, Mask R-CNN 과 같이 다양한 형태로 개발 또는 개선되어 왔다. R-CNN은 영상에서 객체를 탐지할 후보 영역을 추출한 뒤 CNN 연산을 실시해 획득한 특징에 대해 HOG나 SVM, 선형 회귀방법을 적용해 객체를 탐지하는 방법으로 2013년 FAIR에서 개발하였다[27]. 기존의 HOG, SVM만을 적용한 경우에 비해 안정적인 결과를 산출하지만 학습하는데 시간이 오래 걸리며 저장공간이 대용량으로 요구된다는 문제점이 있다. 또한, 검출 시에도 영상 1장을 처리하는데 수십 초가 소요될 만큼 속도가 느리다는 단점이 있다[28]. 이러한 비효율성은 객체 후보 영역 전체에 대해 개별적인 연산을 하기 때문인데, CNN중 초기 모델이다보니 성능에 초점을 둔데 따른 영향으로 판단된다.

Fast R-CNN은 R-CNN의 개별 연산에 따른 속도 저하 문제를 보완해 2015년 Microsoft에서 발표한 모델이다. 영상에서 객체가 있을 만한 영역만 추출하는 RIO Pooling(Region Of Interest Pooling)이라는 계층을 도입해 정규화된 특징을 추출함으로써 최소한의 관심영역에서 객체 탐지를 수행, 정확성과 속도 문제를 개선한 것이다. Girshick (2015)의 연구 결과에 따르면 R-CNN과 Fast R-CNN을 비교한 결과, 학습에 소요되는 시간이 각각 84초와 9.5초로 R-CNN의 시간이 짧으며, 검출 또한 47초와 0.32초로 큰 차이를 보인다[29]. 정확도 측면에서도 66.0%와 66.9%로 미소하게 개선된 결과를 보여 활용성이 높은 편이다. 추가로, 선택적 탐색 과정을 개선해 속도를 획기적으로 향상시킨 Faster R-CNN도 발표된 바 있는데, 이는 RPN(Region Proposal Network)를 통해 객체가 있을 것으로 예상되는 영역을 제안한 후 Fast R-CNN을 수행하는 방식이다. 단, 크기가 다른 객체를 동시에 학습하는 경우 작은 객체에 대한 검출 성능이 떨어지는 것으로 알려져 있다[30]. Faster R-CNN은 현재 카카오톡 로드뷰에서 사용되는 기법이다[31].

Mask R-CNN은 Fast R-CNN 적용 후 객체에 해당 하는지 여부를 마스킹하는 네트워크를 추가한 모델로 2017년 Facebook AI팀에서 개발했다. 즉, 정확한 픽셀 위치를 추출하기 위해 CNN을 통과하면서 관심 위치 영역 위치에 생기는 소수점 오차를 2차원 선형 보간을 통해 감소시킴으로써 정확도를 높인다[32,33]. 국내에서는 건물 벽 폐색영역 추출 실험을 수행함에 있어 Fast

R-CNN과 Mask R-CNN 간의 성능 분석을 실시한 바 있는데, 검출률은 차량과 사람에서 높게 나타났고, 정확도는 신호등, 차량, 사람, 가로수 등에서 80% 이상으로 높게 나타남을 확인했다. 또한, 이러한 결과를 바탕으로 폐색영역 탐지 및 자동적 폐색영역 처리가 가능하다는 결론을 내린 바 있다[34].

2.2.7 RNN(Recursive Neural Network)

RNN은 기본적인 ANN 구조에서 이전 시간의 은닉층의 출력값을 다음 시간의 은닉층의 입력값으로 넣는 경로가 추가된 형태로, 이전 시간의 결과가 다음 시간에 영향을 미치는 과정이 끊임없이 반복되는 모델이다. 시간에 의한 영향을 분석하기 위해 활용되는 모델로 동일한 지역에 대해 기존 학습 자료 외 신규 자료가 확보되는 경우와 같이 시간차가 존재하는 자료를 함께 다룰 때 일반적으로 사용한다. 단, 이전 시간에 의한 영향이 시간이 지날수록 작아져 어느 정도 지나면 없어지기 때문에 실제로는 현재 시간에 가까운 시간 자료만 고려해 예측하게 된다는 문제가 있으며, 이 경우 시간차이가 많이 나는 자료를 이용한 예측에서는 성능이 떨어지는 것으로 알려져 있다 [35,36]. 이에 장시간 그 영향을 보존해 기억할 수 있도록 하는 기술 개발 측면에서 많은 연구가 진행 중이다.

이상의 다양한 학습기반 객체 탐지 알고리즘 중 영상의 학습을 위한 딥러닝 모델로 가장 널리 활용되는 것은 CNN이다. CNN은 기본적으로 광학영상을 학습시켜 딥러닝을 수행하기 위해 개발되었기 때문에 드론, UAV와 같은 다양한 항체에 탑재된 센서로부터 취득된 영상을 학습데이터로 사용한다면 도심지에서 폐색지역을 탐지 및 보완하는데 있어 활용할 수 있을 것으로 판단된다. 특히, 가장 최신 모델인 Mask R-CNN은 픽셀 레벨의 위치 정확도로 객체를 구분할 수 있는 수준이다.

단, 이러한 딥러닝 기반의 분류 기술을 적용함에 있어 최종 탐색 및 분류 성과에 영향을 미치는 것은 모델 학습에 이용된 영상의 해상도나 정확도, 왜곡 여부이다. 모델 학습에 사용된 영상과 다른 영상을 이용해 분류를 실시하거나 해상도가 다른 경우, 촬영 시기가 차이가 많이 나는 경우에는 현저히 정확도가 떨어지게 된다. 또한, 기하학적인 변화나 오차에도 큰 영향을 받는다고 알려져 있다. 즉, 대상 영역 전체에 대해 일반화된 조건이나 모델을 찾을 수 없어 딥러닝을 기술을 적용하는데 있어서는 여전히 많은 양의 영상을 학습시켜야 한다는 한계가 존재하며, 이로 인해 아직까지는 새로 발표된 기술의 적용 가능성을 점검하거나 적용한 초기 결과를 비교한 정확도

분석을 수행하는 단계에 머물러 있다.

3. 딥러닝 기술을 이용한 드론 영상의 폐색영역 해결 방안

실감형 공간정보 구축을 위해서는 대상 객체의 폐색 혹은 적치물이 없는 상태의 벽면 객체 정보가 필요하다. 이러한 장애요인은 주로 가로수, 차량, 사람, 현수막과 같은 장애물에 의해 발생된다. 이러한 장애물은 대부분 정형데이터로 객체를 패턴화하여 딥러닝 기술로 학습데이터셋으로 지정한다면 폐색을 유발하는 객체의 자동 검출 및 해결이 가능하다.

학습데이터셋의 정확도에 따라 대상 객체의 추출율과 정확도는 직접적으로 영향을 받기 때문에 다양한 데이터를 이용하여 학습데이터셋을 만들어야 한다. 이렇게 패턴화된 데이터셋을 이용하여 드론 영상의 폐색을 유발하는 객체를 자동으로 탐지하고 해결할 수 있다.

Fig. 6에 나타난 객체 검출 작업의 흐름도에 나타난다. 드론 혹은 모바일매핑시스템 등에서 구축된 영상을 대상으로 폐색영역 해결을 실시한다. 이를 위해 객체별 데이터셋을 사전에 제작하여 폐색을 유발하는 객체에 대한 검출 및 영역을 자동으로 지정하는 단계를 수행한다. 폐색을 유발하는 객체는 차량, 현수막, 사람 등과 같은 정형데이터와 구름, 그림자 등의 비정형데이터로 구분하여 학습데이터셋을 제작하여 DB로 관리한다. 검출된 폐색영역을 대상으로 Inpainting 기술을 적용하여 폐색영역을 해결하고 보정한다.

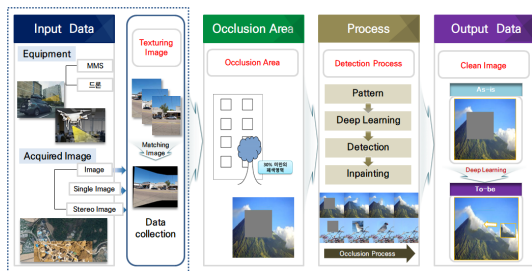


Fig. 6. Process of object detection

세부적으로 모의영상을 생성해 폐색영역을 해결함에 있어 딥러닝 기술의 적용 방안은 다음과 같다. 먼저, 다양한 형상으로부터 건물 가로수, 차, 신호등, 현수막과 같은 객체에 대한 학습을 실시한다. 이후 드론에서 획득한 영

상에서 앞서 학습한 데이터를 사용해 건물을 비롯한 객체를 검출한 뒤 추출한 객체에 대한 폐색영역의 면적 및 비율을 계산해 모의영상을 생성한 뒤 마스킹하는 것이다. 이 때, 객체 탐지를 위한 모델은 미리 학습된 COCO 모델이 있으므로 해당 모델에 포함되지 않은 객체들(예 : 현수막)을 선정해 추가 학습 후 활용하여야 한다. Fig. 7은 드론 기반 영상 촬영 시 발생할 수 있는 폐색영역을 지형에 의한 경우와 지물에 의한 경우로 나눈 뒤, 딥러닝 기술을 기반으로 이를 검출하고 마스킹하기 위한 방법의 개념도 형식으로 정리한 것이다.

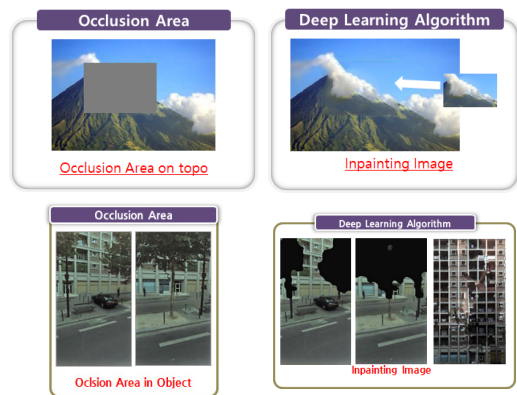


Fig. 7. Concept of occlusion area detection and its recovery for the images obtained by drone

위와 같이 딥러닝 기술을 기반으로 모의영상을 생성하여 채우게 된다면 기존의 수많은 수작업을 대체하는 것이 가능하다. 이 경우, 인력적, 시간적, 비용적인 효율성을 극대화할 수 있어 작업 방법의 개선 측면에서도 의의가 있으며, 나아가 드론 영상을 활용한 3차원 공간정보 구축 분야의 활성화에도 기여할 수 있으리라 판단된다.

4. 결론

본 연구에서는 실감형 공간정보 구축에서 문제가 되는 폐색영역을 최소화할 수 있는 방안으로 최신 기술인 딥러닝 기술의 적용 가능성을 분석하였다.

먼저, 영상으로부터 객체를 검출하기 위한 알고리즘은 분류기 기반 탐지 알고리즘인 HOG, AdaBoost부터 기계학습 모델인 SVM, ANN, 은닉층의 개수를 늘린 딥러닝 방식의 DNN, CNN, RNN까지 다양하게 개발되고 있다. 이 중 가장 대표적인 것은 하나 또는 여러개의 컨볼

루션 레이어와 인공신경망을 CNN으로 영상 부분에 최적화된 모델이다. R-CNN, Fast R-CNN, Faster R-CNN, Mask R-CNN에 이르기까지 단점을 보완하여 효율적인 처리가 가능하도록 다양한 모델로 고도화되었으며, 현재도 영상 부분에서 많은 연구가 이루어져 그 효용성이 가장 높다고 할 수 있다.

영상의 폐색영역을 해결하기 위한 가장 적합한 알고리즘은 영상의 여러 개층으로 구분하여 학습시키고 패턴화하는 CNN 기법이 가장 적합한 것으로 분석된다. 특히 R-CNN, Mask R-CNN은 영상 탐지 및 분류에 효율적이며 처리 속도 및 정확도 측면에서 더욱 효율적이다.

컴퓨터 비전기술의 발달로 실감형 3차원 공간정보에 대한 수요는 지속적으로 증가할 것이다. 하지만 실감형 공간정보를 구축하기 위해서는 폐색영역과 같은 문제를 최단시간에 해결하고 이를 서비스할 수 있는 방안이 필요하다. 따라서 재촬영 등의 원초적인 방법이 아닌 폐색을 유발하는 객체를 자동으로 검출하고 이를 해결할 수 있는 딥러닝 기술은 영상의 폐색영역을 해결하기 위한 최적의 기술이다.

향후 딥러닝 알고리즘 개발과 다양한 학습데이터셋이 구축된다면 폐색영역을 해결하기 위한 시간, 인력, 비용적인 측면에서의 효율성을 높일 수 있으며, 나아가 드론 영상의 활용도 향상도 기대할 수 있다.

References

- [1] O. Kwon, *Detection and Restoring the Occlusion Area for Generating Digital Orthoimage*, Master's thesis, Seoul National University, Seoul, Korea, pp.13-16, 2000.
- [2] J. Yom., D. Lee, D. Kim, "Automatic 3D building reconstruction by integration of digital map and stereo imagery for urban area", *KSCE Journal of Civil Engineering*, Vol.8, No.4, pp.443-449, July. 2004
DOI: <https://doi.org/10.1007/bf02829168>
- [3] M. Seo, D. Y. Han, B. K. Lee, Y. I. Kim, "Detecting and restoring the occlusion area for generating the true orthoimage using IKONOS image". *Korean Journal of Remote Sensing*, Vol.22, No.2, pp131-139, Apr. 2006.
- [4] J. Youn, G. H. Kim, 2008, "Visible height based occlusion area detection in true orthophoto generation", *Journal of the Korean Society of Civil Engineers D*, Vol.28, No.3D, pp.417-422, May. 2008.
- [5] E. J. Yoo, D. Lee, "Detection and recovery of occlusion areas caused by building sidewalls and aerial photos", *Proceedings of Korean Society of Surveying, Geodesy, Photogrammetry, and Cartography 2016*, KSGPC, Suwon, Korea, pp.156-158, Apr. 2016.
- [6] S. Shim, C. Chun, S. Choi, S. Ruy, "Road Surface Damage Detection based on Object Recognition using Fast R-CNN", *The Korean Institute of communications and Information Sciences*, Vol.18. No.2, pp.104-113, Apr. 2019.
DOI: <https://doi.org/10.12815/kits.2019.18.2.104>
- [7] H. C. Song, M. Kang, T. Kim, "Object Detection based on Mask R-CNN from Infrared Camera", *Journal of Digital Contents Society*, Vol.19, No.6, pp.1213-1218, June 2018.
DOI: <https://doi.org/10.9728/dcs.2018.19.6.1213>
- [8] I. Choi, J. Seo, H. Park, "Object Recognition of Low Resolution Images based on Deep Learning", *Proceeding of Korea Computer Congress 2017*, KIISE, Jeju, Korea, pp.782-784, June 2017.
DOI: <https://doi.org/10.1109/access.2019.2941005>
- [9] D. S. Jeong, H. Kim, J. Shin, J. Paik, "Deep Learning-Based Person Re-identification Using Semantic Segmentation", *Proceeding of the Institute of Electronics and Information Engineers*, IEIE, Incheon, Korea, pp.392-394, Nov. 2018.
- [10] A. Tang, K. Lu, Y. Wang, J. Huang, H. Li, "A real-time hand posture recognition system using deep neural networks". *ACM Transactions on Intelligent Systems and Technology*, Vol.6, No.2, p21, 2015.
DOI: <https://doi.org/10.1145/2735952>
- [11] W. Ouyang, X., Wang, "Joint deep learning for pedestrian detection", *Proceedings of the IEEE International Conference on Computer Vision*. IEEE, NSW, Australia, pp.2056-2063, Dec. 2013.
DOI: <https://doi.org/10.1109/iccv.2013.257>
- [12] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, CA, USA, pp.1-9, June 2005.
DOI: <https://doi.org/10.1109/cvpr.2005.177>
- [13] B. Kim, S. Oh, J. Kim, "Design of RBFNNs Pattern Classifier-based Two-dimensional Face Recognition System Using HOG and Adaboost Algorithm", *Proceeding of the Korean Institute of Electrical Engineers*, KIEE, Kosung, Korea, pp.77-78, Apr. 2015,
DOI: <https://doi.org/10.5370/kiee.2014.63.6.797>
- [14] Y. Freund, R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of Computer and System Sciences*, Vol.55, No.1, pp.119-139, 1997.
DOI: <https://doi.org/10.1006/jcss.1997.1504>
- [15] S. Kim, J. Park, J. Lee, "Monocular Image and AdaBoost Learning Based Nighttime Preceding Vehicle Detection for ADAS and Intelligent Headlamp System", *Journal of Institute of Control, Robotics and Systems*, Vol.23, No.10, pp.886-893, Oct. 2017.
DOI: <https://doi.org/10.5302/jicross.2017.17.0134>

- [16] C. Cortes V, Vapnik, "Support-vector networks". *Machine Learning*, Vol.20, No.3, pp.273-297, July 1995.
- [17] V. Vapnik, The nature of statistical learning theory, p.313. New York: Springer-Verlag New York, 2000, pp.123-167.
DOI: <https://doi.org/10.1007/978-1-4757-2440-0>
- [18] X. Zhang, L. Li, D. Pi, "Toward Optimization of SVM Learning with RBF Kernel", *Proceedings of International Technical Conference on Circuits Systems, Computers and Communications, ITC-CSCC*, pp.437-440, July 2006.
- [19] M. Im, D. K. Yoon, B. Kim, "A study on Object Recognition Algorithm based on SVM Machine Learning Algorithm", *Proceeding of Institute of Control, Robotics and Systems, ICROS, Kyeong-ju, Korea*, pp.272-273, May 2019.
- [20] B. C. Kuo, H. H. Ho, C. Li, C. Jung, J. S. Taur, "A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol.7, No.1, pp.317-326, 2013.
DOI: <https://doi.org/10.1109/istars.2013.2262926>
- [21] P. P. Plehiers, S. H. Symoens, I. Amghizar, G. B. Marin, C. V. Stevens, K. M. Van, "Artificial Intelligence in Steam Cracking Modeling: A Deep Learning Algorithm for Detailed Effluent Prediction", Preprint version, July 2019.
DOI: <https://doi.org/10.1016/j.eng.2019.02.013>
- [22] Neural Networks and Deep Learning. Available From: <http://neuralnetworksanddeeplearning.com/chap5.html> (accessed Sep. 27, 2019)
- [23] G. E. Hinton, R. R. Salakhutdinov. "Reducing the Dimensionality of Data with Neural Networks", *Science*, Vol.313, No.5786, PP.504-507, July 2006
DOI: <https://doi.org/10.1126/science.1127647>
- [24] M. Oh, H. Choi, S. Kim, J. Jang, J. Jin, M. Cheon (2017). Analysis of social welfare and estimation model based on machine learning, Technical Report, Korea Institute for Health and Social Affairs, Korea, pp.54-84.
- [25] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition", *Neural Computation*, Vol.1, No.4, pp.541-551, Mar. 1989.
DOI: <https://doi.org/10.1162/neco.1989.1.4.541>
- [26] A. Krizhevsky, I. Sutskever, G. E. Hinton. "Imagenet classification with deep convolutional neural networks". *Proceedings of Advances in neural information processing systems, NIPS, NV, USA*, pp.1097-1105, Dec. 2012.
DOI: <https://doi.org/10.1145/3065386>
- [27] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation". *Proceedings of the IEEE conference on computer vision and pattern recognition*, IEEE, OH, USA, pp. 580-587. June 2014.
DOI: <https://doi.org/10.1109/cvpr.2014.81>
- [28] J. R. Uijlings, K. E. Van De Sande, T. Gevers, A. W. Smeulders. "Selective search for object recognition". *International journal of computer vision*, Vol.104, No.2, pp.154-171, Apr. 2013.
DOI: <https://doi.org/10.1007/s11263-013-0620-5>
- [29] R. Girshick, "Fast r-cnn", *Proceedings of the IEEE international conference on computer vision*, IEEE, Santiago, Chile, pp. 1440-1448, Dec. 2015.
DOI: <https://doi.org/10.1109/iccv.2015.169>
- [30] S. Ren, K. He, R. Girshick, J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks". *Proceedings of Advances in neural information processing systems, NeurIPS, Montreal, Canada*, pp.91-99, Dec. 2015.
DOI: <https://doi.org/10.1109/tpami.2016.2577031>
- [31] Kakao, Available From: <https://brunch.co.kr/@kakao-it/66> (accessed Sep. 27, 2019)
- [32] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask r-cnn", *Proceedings of the IEEE international conference on computer vision*, IEEE, Venice, Italy, pp.2980-2988, Oct. 2017.
DOI: <https://doi.org/10.1109/iccv.2017.322>
- [33] K. Zhao, J. Kang, J. Jung, G. Sohn. "Building Extraction from Satellite Images Using Mask R-CNN with Building Boundary Regularization", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, IEEE, CA, USA, pp.247-251, June 2018.
DOI: <https://doi.org/10.1109/cvprw.2018.00045>
- [34] H. J. Kim, J. M. Lee, K. H. Bae, Y. D. Eo, "Application Research on Obstruction Area Detection of Building Wall using R-CNN Technique", *Journal of Cadastre & Land InformatiX*, Vol.48, No.2, pp.213-225, Dec. 2016.
DOI: <https://doi.org/10.22640/LXSIRI.2018.48.2.213>
- [35] A. Sperduti, A. Starita, "Supervised neural networks for the classification of structures", *IEEE Transactions on Neural Networks*, Vol.8, No.3, pp.714-735, May 1997.
DOI: <https://doi.org/10.1109/72.572108>
- [36] P. Frasconi, M. Gori, A. Sperduti, "A general framework for adaptive processing of data structures", *IEEE Transactions on Neural Networks*, Vol.9, No.5, pp.768-786, Sep. 1998.
DOI: <https://doi.org/10.1109/72.712151>

배 경 호(Kyoung-Ho Bae)

[정회원]



- 2006년 2월 : 동아대학교 토목공학과 (공학박사)
- 2006년 4월 ~ 2017년 7월 : 공간정보산업협회 책임연구원
- 2017년 7월 ~ 현재 : (주)신한항업 연구소장

<관심분야>

공간정보, 지도제작, 정보통신

박 흥 기(Hong-Gi Park)

[정회원]



- 1988년 2월 : 연세대학교 토목공학과 (공학박사)
- 1997년 1월 ~ 1997년 12월 : 국립지리원 전문직공무원
- 1990년 3월 ~ 현재 : 가천대학교 토목환경공학과 교수

<관심분야>

측량정책, 공사측량, 수치지도/공간정보