딥러닝과 드론을 이용한 동일 모양 건축 자재 수량 측정에 관한 실증적 연구

문지환*, 송누리*, 최재갑*, 박진호**, 김계영**
*숭실대학교 융합소프트웨어학과
**숭실대학교 소프트웨어학부
e-mail:gkrrydn_ji@naver.com, nuri@soongsil.ac.kr, kor_03@naver.com,
j.park@ssu.ac.kr, gykim@ssu.ac.kr

Empirical study for counting same shaped building material quantity using UAV and deep learning

Ji-Hwan Moon*, Nu-lee Song*, Jae-Gab Choi*,
Jin-Ho Park**, Gye-Young Kim**
*Dept of Convergence Software, Soong-Sil University
**School of Software, Soong-Sil University

요 약

최근에 카메라를 사용하여 물체를 인식하고 수량을 측정하는 연구가 진행되고 있다. 하지만 건축 자재와 같이 인접한 동일 물체의 경우, 인식과 측정이 어려우며, 이와 관련된 연구는 부족한 실정이다. 본 논문에서는 이러한 문제를 해결하기 위해 딥러닝과 드론을 사용하는 방법을 통하여 자재 더미의 윗면과 측면의 촬영, 드론 로그를 사용하한 드론 높이와 Yaw값 획득, 영상 내 자재 종류와 영역 정보 등 미리 학습된 딥러닝 모델을 사용한 획득방법, 드론 촬영 시의 Yaw값 차이를 이용해 윗면과 측면이 공통으로 가지는 직선 검색, 자재 더미의 가로, 세로, 높이의 비율 등을 통한 각 화소 비율과 자재 개수의 영역별 가상의 체적값 분석, 드론의 화각, 높이를 사용한 체적 사진측량값 획득, 가상 체적과 사진측량법 사이의 오차 최소화 비율 등을 연구하였다.

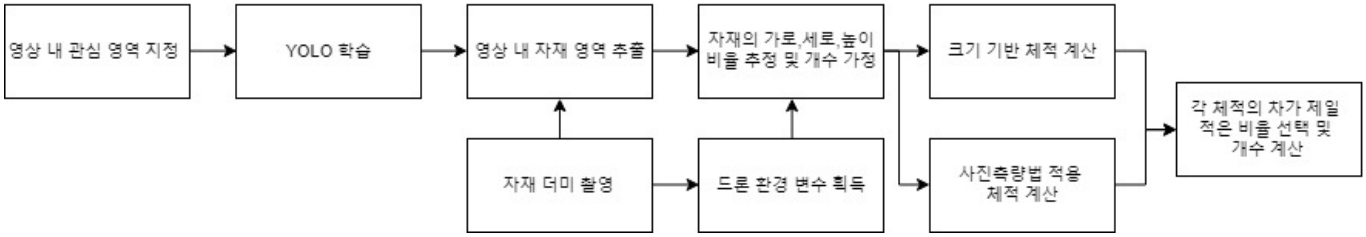
1. 서론

카메라를 사용한 영상 내 물체 인식 혹은 두 개 이상의 영상 내 동일 물체 인식에 관련된 연구가 활발히 진행되고 있다. 기존엔 색상이나 모양 같이 자재가 가지는 특징을 비교하여 물체를 인식 했지만[4], 관심 물체와 비슷한 모양이나 색상이 존재할 경우 정확도가 매우 낮은 결과를 보였다. 하지만 기술의 발전으로 인한 연산 처리 속도의 급격한 상승으로 딥러닝이 대두되면서 정확도는 향상되었고, 기존의 물체가 가지는 특징을 학습하여 관심 물체의 특징과 다른 물체의 특징 차이를 정밀하게 구분할 수 있게 되었다.[16] 물체 인식 기술의 발전으로 물체의 수량을 세는 연구도 활발히 진행되었으나[17], 물체의 체적을 알고 있는 상황에서는 물체가 가지는 특징 사이의 거리를 계산해 물체와 카메라 사이의 거리를 측정하거나, 물체를 특징 단위로 구분해 하나의 영역을 만들고 그 영역을 하나의 개체로 측정하는 등 오류가 많아 거리 측정 방법은 기준이 되는 물체의 특징이 유일하지 않는 건축 자재 더미에선 계산이 어려운 문제가 발생한다.

본 논문에서는 위와 같은 문제점을 가지는 건축 자재 더미의 수량을 측정하기 위해 딥러닝과 드론을 사용에 대하여 연구하였으며, 자재 더미 단일 자재의 기 체적값을 가진 상태에서, 동일한 모양의 자재가 쌓여있는 상태의 자재 더미 영역 추출을 위한 딥러닝 활용법과 수량을 계산할 때 드론과 사진측량법을 이용한 촬영 카메라의 화각, 높이, 촬영시의 Yaw값, 촬영된 영상의 가로, 세로 화소값이 사용된다. 학습 데이터는 인터넷 상의 임의의 자재 더미 영상 수집 등에 대하여 연구하였다.

2. 관련 연구

기존 연구로는 움직이는 물체를 차영상을 통해 물체 영역을 지정하고 BMA(Block Matching Algorithm)와 템플릿 매칭(Template Matching)을 사용해 움직이는 차량을 구분하거나[1] 물체가 가지는 색과 깊이 정보를 활용하여 물체의 색상 히스토그램과 STR(Steering Kernel Regression)을 구해 다른 물체와 구분하거나 추적하는 연구가 있다.[2,3,15] 하지만 이는 움직이지 않거나 비슷한 색상이 존재할 경우 낮은 정확도를 보인다. 또 다른 연구로 물체가 가지는 고유의 모양에서 특징을 얻어 물체를 구분하거나 추적한다. 이 방법은 SIFT[5] 같은 특징 기술자를 이용한다.[6,8] 여기서 성능을 높이기 위해 물체의 특징을 반 지도학습 분류기를 통해 학습시키고 물체를 인식하거나[7] 물체의 크기를 구하기 위해 카메라 정보와 특징



(그림 1) 전체적인 실험 흐름도

집합을 사용하여 3D AR(Augment Reality)을 구축하기도 한다.[9] 특징 기술자를 사용하면 물체가 가지는 특징을 비교하여 비교적 정확하게 물체를 찾아낼 수 있지만, 영상 내 비슷한 특징을 가지는 다른 물체가 있을 경우나 중요한 특징이 가려지거나 작게 표현되는 경우 전혀 다른 특징점이 연결되는 단점이 있다. 최근엔 그래픽카드 같은 하드웨어의 발전으로 딥러닝을 사용한 연구가 많다. 이 방법은 많은 학습 데이터 및 학습 시간이 필요하지만, 기존에 진행되었던 연구보다 높은 정확도를 가진다. 관련 연구로는 기존 연구에 사용했던 정보들을 엔트로피 기준(Entropy-based) 특징 추출자를 통해 얻고 이들을 베이시안 분류기를 통해 학습하여 물체를 더 정밀하게 검출하거나[10], PASCAL VOC, MS COCO 데이터 집합 영상에서 목표 물체가 가지는 영역을 학습시킨 후 새로운 영상에서 목표 물체의 영역을 찾는 연구[12, 13, 14]가 있다.

3. 실험 방법

본 실험은 크게 영상을 사용해 관심 영역을 설정하고 전처리하는 딥러닝 모델 학습, 수량 계산에 필요한 정보를 설정 및 추출하는 부분, 마지막으로 건축 자재 더미의 수량을 계산하는 부분으로 나뉜다. 이를 간단한 흐름도로 표현하면 (그림 1)과 같다.

3.1. 딥러닝 모델 학습

딥러닝 학습 모델은 R-CNN의 한 종류인 Yolo(You Look Only Once)[11]를 사용한다. Yolo를 학습시킬 영상은 자재 더미의 윗면과 측면이 나온 것으로 한다. 영상 내에선 자재 더미 영역을 관심 영역으로 지정한 후 영상 크기를 448x448로 조정한다. 학습 시 사용할 건축 자재는 벽돌, 시멘트, 연석 3가지로 하고, 모델 학습은 검증 시 발생하는 평균 오류율이 1.0 이하가 될 때 중단한다.

3.2. 드론의 영상, 환경 변수

이 연구에서 사용한 드론은 DJI Phantom4 pro plus이다. 이 드론은 내부적으로 84°의 화각과 360°기준의 Yaw, 드론이 동작하는 바닥에서 드론 촬영 위치까지의 높이 값을 가진다. 드론으로 촬영한 영상은 5472x3078로 자재의 윗면, 측면 모습을 담는다. 이 변수들은 사진측량법을 사용하여 자재 더미의 체적을 계산할 때 사용한다.

3.3. 수량 계산

먼저, 드론으로 촬영한 두 영상의 자재 더미 영역을 Yolo를 사용해 예측한다. 그리고 측면의 Yaw값과 윗면의 Yaw값을 비교해 자재의 가로, 세로 성분 중 두 영상에서 공통으로 찍힌 직선을 찾는다. 만약 윗면 촬영시의 Yaw값이 약 90°가량 차이가 난다면 체적을 구성하는 직선 중 세로가 공통으로 찍힌 것이고, 180° 혹은 별 차이가 없다면 가로가 공통으로 찍힌 것이다. 다음으로 자재 더미와 드론 사이의 거리 차로 인해 발생하는 공통 직선의 화소 차이를 없애고, 높이 화소 값으로 체적의 모든 직선들을 나눠 계산이 간단해 지도록 한다. 이는 가로, 세로, 높이를 W,H,D 라 가정하고 H_s,H_u를 두 영상에서 공통으로 촬영된 세로 값이라 한다면 다음과 같은 식(1)이 된다.

$$W\colon H_u,\ D\colon H_s \Rightarrow WH_s\colon H_uH_s,\ DH_u\colon H_sH_u \qquad (1)$$

$$\Rightarrow WH_s\colon H_uH_s\colon DH_u$$

$$\Rightarrow \frac{WH_s}{DH_u}\colon \frac{H_uH_s}{DH_u}\colon 1$$

이제 식(1)에서 얻은 비율에서 높이를 중심으로 자재가 존재한다고 가정한다. 즉, 높이 비율만큼 단일 건축 자재가 쌓여있는 층수를 가정하고, 해당 길이(h_a)를 기준으로 가로, 세로의 길이를 정한다. 이 가정된 상황에서 가로, 세로를 이용해 자재 더미 윗면의 대각선 길이를 구한다. 다음으로 사진측량법과 가정된 높이를 사용해 영상이 촬영할 수 있는 대각선 길이를 구한다. 이 대각선 길이를 구하는 식은 다음과 같으며, FOV는 화각, h_r는 드론의 높이를 의미한다.

$$diagonal = 2\times(h_d - h_a)\times\tan\left(\frac{FOV}{2}\right) \qquad (2)$$

이 식(2)를 통해 얻은 영상의 길이와 화소를 이용해 하나의 화소가 가지는 실제 길이를 구한다. 그리고 이를 Yolo를 통해 얻은 더미 영역의 화소에 적용시켜 드론에서 촬영할 수 있는 실제 더미의 대각선 길이를 구한다. 그 후, 가정한 상황에서 얻은 대각선 길이와 이 때 촬영할 수 있는 실제 대각선 길이의 차이를 비교하고, 이 과정을 가정

한 자재 더미의 높이가 드론을 넘지 않을 때 까지 반복한다. 그리고 가장 차이가 적은 상황의 자재 비율의 체적을 계산해 자재의 실제 수량을 알아낸다.

4. 실험 결과

실험 결과는 Yolo를 사용한 윗면, 측면 영상 내의 자재 종류 구별 및 자재 더미 영역 오차율에 따라 크게 달라진다.

먼저, Yolo 모델을 사용하여 학습에 사용되지 않은 임의의 건축 자재 영상을 종류별로 6개씩 사용하여 검증한 결과 약 78%의 정확도를 보였다. 자재 종류 인식을 못한 영상들은 벽돌을 연석이라 잘못 인식하거나, 자재 더미의 색상이 주변 환경과 비슷한 색일 경우 자재를 인식하지 못하였다.

(a) (b)

(그림 2) yolo를 사용하여 벽돌(a)과 연석(b)를 인식한 영상

다음으로, 학습 모델을 사용하여 얻은 건축 자재 더미 영역의 정확도다. 이것 또한 자재 더미 주변 환경에 따라 많은 차이를 보인다. (그림 3)과 같이 주변 색상이 비슷할 경우엔 원래 영역보다 더 적게 혹은 많게 인식하는 경우가 많고, 한 영상 내 같은 건축 자재라도 영역의 크기를 다르게 하는 경우가 있다.

(그림 3) 주변 환경과 색상이 비슷하여 자재 더미 영역이 축소된 영상

건축 자재 더미 수량은 Yolo 자재 종류 인식과 자재 더미 영역 정확도에 종속적인 결과를 보인다. (그림 2)와 같이 자재 더미 영역과 종류를 정확하게 인식한다면 약 94%의 정확도를 보이지만, (그림 3)과 같이 자재 종류는 정확하게 인식했지만 자재 더미 영역의 오차가 있을 경우 정확도는 80%까지 떨어진다.

5. 결론

딥러닝 모델을 학습시킬 땐 학습 목적으로 하는 물체와 물체 주변 환경의 다양성이 존재해야 한다. 그래야 학습 모델을 사용할 때 물체와 주변 환경을 충분히 고려해 정확한 값을 도출해내기 때문이다. 하지만 이번 연구에선 인터넷 상의 존재하는 소량의 건축 자재 자료를 바탕으로 학습하였다. 그래서 자재의 종류와 영역의 대한 인식 정확도가 떨어지고, 이에 종속적인 자재 수량 계산 정확도도 영향을 받았다. 또한 제안한 방법은 자재가 차지하는 사각형의 체적을 계산한다. 그래서 (그림 2)-(a)처럼 자재 더미의 일부분이 비어있는 경우엔 이를 인식하지 못해 수량 계산 오차가 발생한다.

앞서 설명된 문제점을 개선하기 위해 향후에는 직접 촬영 혹은 건축 자재 데이터 집합을 찾아 학습 데이터를 늘려 종류 및 영역 인식 정확도를 증가시키고, 앞서 설명했던 것처럼 자재 더미 영역의 공백 부분에 대한 예외 처리를 하여 자재 수량 계산 정확도를 증가시켜야 할 것이다.

ACKNOWLEDGMENT

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업의 연구결과로 수행되었음" (IITP-2018-2018-0-01419)

참고문헌

[1] P. S. Khude, S. S. Pawar, "Object detection tracking and counting using enhanced BMA on static background videos", IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2013.

[2] Camplani, M., Paiement, A., Mirmehdi, M., Damen, D., Hannuna, S., Tao, L., Burghardt, T.: Multiple human tracking from RGB-D data: a survey. IET Comput. Vis. (2017)

[3] Zoidli,"Visual object tracking based on local steering kernels and color histograms", Tefas circuit ans system for video technology IEEE Trans. Vol23 May 2013

[4] Pornpanomchai et al, "Object detection and counting system", Congress on Image and Signal Processing, vol. 2, pp. 61 - 65. IEEE (2008)

[5] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints" International Journal of Computer Vision, 60, 2 (2004), pp. 91-110.

[6] Belongie et al, "Shape matching and object

recognition using shape contexts", 2002 IEEE Trans. Pattern Anal. Mach. Intell. 24, 4, 509 - 522

[7] R. Fergus, P. Perona, A. Zisserman, "A Sparse Object Category Model for Efficient Learning and Exhaustive Recognition", Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2005.

[8] D. Lowe et al. "Scene modeling, recognition and tracking with invariant image features.", IEEE Int. Symp. on Mixed and Augmented Reality, 2004.

[9] Vincent Lepetit and Pascal Fua (2005), "Monocular Model-Based 3D Tracking of Rigid Objects: A Survey", Foundations and Trends® in Computer Graphics and Vision: Vol. 1: No. 1, pp 1-89.

[10] R. Fergus, P. Perona, A. Zisserman, "Object Class Recognition by Unsupervised Scale-Invariant Learning", Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2003.

[11] Joseph Redmon et al,"You Only Look Once:Unified, Real-Time Object Detection", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788

[12] Sean Bell et al, "net: Detecting objects in context with skip pooling and recurrent neural networks". CVPR 2016, abs/1512.04143, 2015

[13] Spyros Gidaris et al, "Object detection via a multi-region and semantic segmentation-aware cnn model", In The IEEE International Conference on Computer Vision (ICCV), December 2015.

[14] Wanli Ouyang et al, "Deepidnet: Deformable deep convolutional neural networks for object detection", In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.

[15] Barinova O et al, "On detection of multiple object instances using Hough transforms". IEEE Trans Pattern Anal Mach Intell 2012.

[16] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images", Proc. IEEE, vol. 101, no. 3, pp. 652-675, Mar. 2013.

[17] M.-R. Hsieh, Y.-L. Lin, and W. H. Hsu. Drone-based object counting by spatially regularized regional proposal networks. In The IEEE International Conference on Computer Vision (ICCV), 2017.