

3D 모델 합성 데이터셋을 이용한 항공 영상 객체 탐지 연구

이도현, 김성용, 조규태

LIG 넥스원

dohyun.lee1@lignex1.com, sungyong.kim@lignex1.com, kyutae.cho2@lignex1.com

A Study on Aerial Image Object Detection Using a 3D Model Synthetic Dataset

Lee Do Hyun, Kim Sung Yong, Cho Kyu Tae

LIG NEX1

요약

본 연구에서는 실데이터 확보와 라벨링의 어려움을 해결하기 위해 3D 기반 이미지 합성 기법을 활용한 객체 탐지 학습 데이터 자동 생성 방법을 제안한다. Blender로 실제 드론 촬영과 유사한 시점에서 렌더링한 객체와 해당 마스크를 이용해 배경 이미지에 자연스럽게 합성하고, 실제 객체 크기와 유사한 범위 내에서 크기와 위치를 무작위로 설정하여 강건성을 확보하였다. 합성과 동시에 바운딩 박스를 자동 기록하여 대규모 학습 데이터를 효율적으로 생성할 수 있으며, 합성 데이터만으로 학습한 YOLOX-s 모델이 테스트셋 기준 mAP@[0.50:0.95] 0.563을 달성하였다. 향후 실제 데이터도 3D 모델로 제작해 동일 방식으로 학습하면 성능 향상을 기대할 수 있다.

I. 서론

최근 항공 영상 기반의 객체 탐지 기술은 국방, 재난 대응, 경찰 등의 분야에서 높은 활용 가치를 보이며 활발히 연구되고 있다 [1]. 특히 드론을 활용한 고공 촬영 영상은 넓은 지역을 신속하게 확보할 수 있는 장점이 있어, 다양한 인공지능 기반 컴퓨터 비전 알고리즘의 학습에 효과적으로 활용되고 있다. 그러나 실제 항공 영상 데이터를 수집하고 학습에 적합한 형태로 라벨링하는 과정은 많은 비용과 시간이 요구되며, 다양한 환경과 시나리오를 충분히 반영하기 어렵다는 한계가 존재한다.

기존의 항공 시점 객체 데이터는 대부분 고정된 객체를 기준으로 드론이 일정한 경로를 따라 이동하며 촬영한 이미지로 구성되며, 배경이나 조명, 위치 조건의 다양성이 부족한 경우가 많다. 이로 인해 실제 환경에서 발생할 수 있는 복잡한 상황을 충분히 반영하지 못하며, 탐지 모델의 일반화 성능에도 제약을 준다.

이에 본 연구에서는 이러한 한계를 극복하기 위해 3D 모델 기반의 합성 데이터셋 제작 방식을 활용하였다. 학습 데이터는 관심 객체의 3D 모델을 Blender[2]에서 드론 촬영 시점과 유사하게 렌더링한 뒤, VBS(Virtual Battlespace) 시뮬레이션 환경에서 추출한 항공 배경 이미지에 무작위로 합성하였다. 이를 통해 다양한 각도, 크기, 배경 조건이 반영된 학습용 데이터를 구성하였다.

반면, test 데이터는 VBS 시뮬레이션 환경 내에 3D 객체를 직접 배치하고, 드론 시점에서 실제로 촬영하듯 데이터를 수집함으로써, 실제 운용 환경에 가까운 구성을 실현하였다.

본 연구에서 생성한 데이터셋만을 활용하여 객체 탐지 모델을 학습한 결과, 테스트셋 기준으로 mAP@[0.50:0.95] 0.563을 달성하였다. 이는 실제 데이터 수집 없이도 자동 라벨링과 다양한 시나리오가 반영된 고품질 학습 데이터를 통해 높은 수준의 성능을 확보할 수 있음을 정량적으로 입증하는 결과이며, 향후 실존 객체를 기반으로 한 pretraining이나 도메인 확장 연구에도 활용 가능성이 크다.

II. 제안 방법

본 연구에서는 실제 항공 데이터를 수집하지 않고도 다양한 시나리오를 반영한 객체 탐지용 학습 데이터를 구축하기 위해, 3D 모델 렌더링과 시뮬레이션 환경 기반의 이미지 합성 파이프라인을 설계하였다. 이 파이프라인은 ① VBS 배경 이미지 확보, ② 3D 객체 렌더링 및 마스크 이미지 생성, ③ 이미지 합성과 자동 라벨링의 세 단계로 구성된다.

VBS 배경 이미지 확보

객체 합성에 사용되는 배경 이미지는 VBS 시뮬레이션 환경에서 수집한 항공 시점 이미지들 중 무작위로 선택하여 활용하였다. VBS는 다양한 지형과 조명 조건, 시점 구성이 가능하여 실제 드론 촬영을 대체할 수 있는 현실적인 배경을 제공한다. 본 연구에서는 객체가 배치될 환경적 다양성을 확보하기 위해 수집된 배경 이미지 중 랜덤하게 선택하여 합성에 사용하였다.

3D 객체 렌더링 및 마스크 생성

학습용 객체 이미지는 Blender를 이용해 렌더링하였다. 3D 모델은 일정 거리에서 시작하여 위쪽으로 각도를 변화시키며, 실제 드론이 상공에서 다양한 시점으로 촬영한 것처럼 보이도록 이미지를 생성하였다. 이때 객체와 배경을 분리하기 위한 마스크 이미지도 그림1과 같이 생성된다. 마스크 이미지는 이후 합성 과정에서 활용되며, 배경을 제거하고 객체만 배경 이미지 위에 삽입되도록 하는 역할을 수행한다. 이를 통해 객체 외의 배경 요소는 완전히 제거된 상태로 합성이 이루어지며, 자연스러운 외관 표현이 가능하다. 또한, 객체 이미지는 경계에 맞게 자동으로 크롭되어 저장되므로, 불필요한 여백 없이 정확한 위치 정렬이 가능하다.

이미지 합성과 자동 라벨링

합성 단계에서는 렌더링된 객체 이미지와 해당 객체의 마스크를 활용하

여, 객체만을 배경 이미지 위에 합성한다(그림 2). 이때, 마스크를 이용하여 객체 외곽의 불필요한 영역을 제거하고, 배경과의 경계가 자연스럽게 어우러지도록 합성을 수행한다. 또한, 실제 객체 크기와 유사한 범위 내에서 객체의 크기를 무작위로 설정하고, 배경 이미지 내 위치 역시 무작위로 배치하여 다양한 시나리오에 대응할 수 있는 강건성을 높였다. 더불어, 합성과 동시에 객체의 바운딩 박스를 자동으로 계산하여 라벨링 정보를 기록한다. 이러한 과정은 대규모 학습 데이터를 수작업 라벨링 없이 생성할 수 있도록 하며, 실제 데이터 수집 없이도 효율적인 객체 탐지 모델 학습을 가능하게 한다.

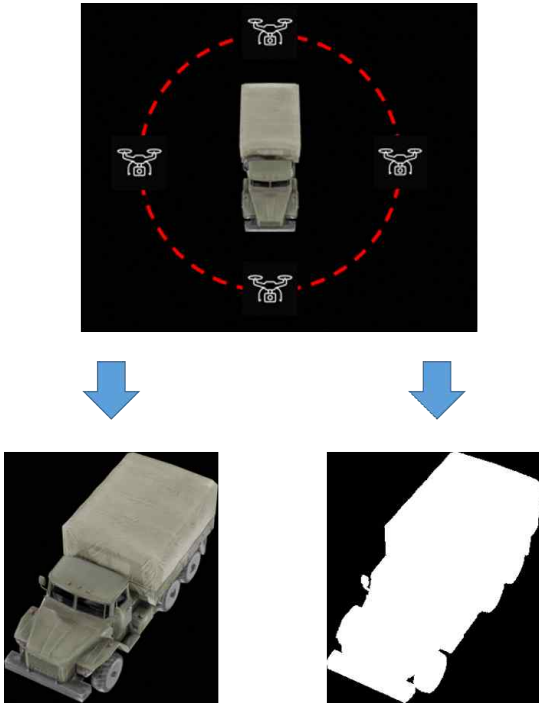


그림 1 3D 객체 렌더링 및 마스크 생성 이미지

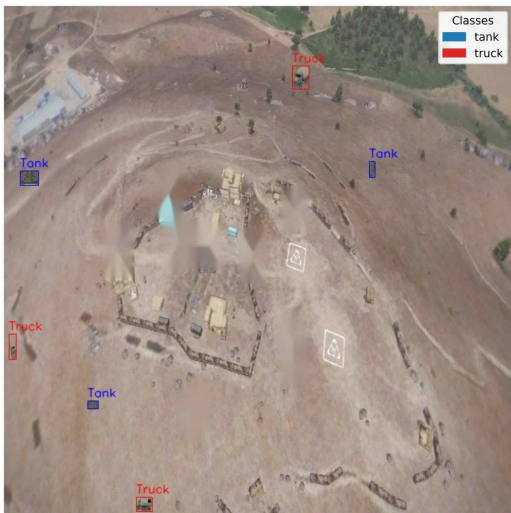


그림 2 합성 이미지 예시

III. 실험 결과

본 연구에서는 구축한 합성 데이터셋을 활용하여 객체 탐지 모델의 학

습과 평가를 수행하였다. 학습에는 매 epoch마다 새롭게 생성된 10,000장의 합성 이미지와 자동 생성된 라벨을 사용하였으며, 테스트는 VBS 시뮬레이션 환경에서 3D 객체를 직접 배치한 뒤 드론 시점으로 촬영한 총 2,000장으로 구성하였다. 이 중 400장은 검증(Validation)용으로 분할해 학습 중 체크포인트 선택에 사용하고, 나머지 1,600장으로 최종 테스트를 수행하였다. 객체 탐지기는 YOLOX-s(Megvii-BaseDetection, commit 6ddff48; 0.3.0 + 38 commits, 2025-06-08), 입력 640×640을 사용하였다 [3]. 합성 데이터만으로 학습한 결과, 테스트셋 기준 $mAP@[0.50:0.95] = 0.563$ 을 기록하였다. 학습은 최대 300 epochs로 진행했으며, 검증 성능이 가장 높은 체크포인트를 테스트에 사용하였다 (표 1). 이는 실제 촬영 데이터를 사용하지 않고도 시뮬레이션 배경과 드론 시점 렌더링 기반의 3D 객체 합성만으로도 의미 있는 탐지 성능을 달성할 수 있음을 보여준다. 추후 실제 영상 수집과 실데이터 기반 3D 객체 제작이 이루어지면, 제안 방법을 재평가할 계획이다.

표 1 데이터 통계 및 학습 설정 요약

항목	내용
학습 데이터	합성 10,000장/epoch
테스트 데이터	1,600장
검증 데이터	400장
입력 해상도	640x640
모델	YOLOX-s
학습 epoch	300 (best-val ckpt 활용)
평가 지표	$mAP@[0.50:0.95]$

IV. 결론

본 연구는 실제 항공 데이터 수집의 제약을 극복하고 다양한 시나리오를 반영한 객체 탐지 학습 데이터 구축을 위해 3D 기반 이미지 합성 방식을 제안하였다. Blender 렌더링과 VBS 시뮬레이션 배경을 결합하여 자동 라벨링이 가능한 고품질 합성 데이터를 생성하였으며, 합성 데이터만으로도 $mAP@[0.50:0.95] 0.563$ 의 성능을 달성하였다. 비록 합성 데이터는 실제 데이터와의 도메인 차이로 인해 실환경 적용 시 한계가 존재하지만, 본 연구의 방법은 실제 데이터를 기반으로 정밀한 3D 모델을 제작하여 동일한 합성 절차를 적용함으로써 이러한 한계를 완화하고 성능을 향상시킬 수 있는 잠재력을 지닌다. 마스크 기반의 배경 제거와 자동 라벨링 구조는 다양한 객체·환경에 적용 가능하며, 향후 정밀 3D 모델링, 합성·실제 데이터의 혼합 학습, domain adaptation, 다양한 기상·시점 변화를 반영한 시뮬레이션 고도화를 통해 성능 향상을 기대할 수 있다. 종합적으로, 본 연구는 자동화된 합성 데이터셋 구축이 객체 탐지 분야에서 데이터 확보와 모델의 일반화 성능 향상의 실질적 대안이 될 수 있음을 제시한다.

참 고 문 헌

- [1] Zhuang, S., Hou, Y., & Wang, D. (2025). Towards Efficient Object Detection in Large-Scale UAV Aerial Imagery via Multi-Task Classification. *Drones*, 9(1), 29.
- [2] Blender Online Community. (2024). Blender - a 3D modelling and rendering package. Stichting Blender Foundation. Retrieved from <https://www.blender.org>
- [3] Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv:2107.08430*.