

임베디드 보드에서의 YOLO 기반 드론 탐지

유병호, 박한빈, 김민성, 최해철

한밭대학교 정보통신공학과

really4725@naver.com, phb2303@naver.com, rudwls7271@naver.com, choihc@hanbat.ac.kr

YOLO based Drone detection on Embedded Board

ByeungHo Yu, HanBin Park, MinSung Kim and Haechul Choi

Dept. of Information and Communications Engineering, HANBAT NATIONAL UNIVERSITY

요약

최근 드론의 용도는 취미, 공연, 농업, 안전, 군사, 연구, 물자수송 등 다양한 분야와 목적으로 활용되고 있다. 더불어 드론의 불법적 활용으로 인한 안전 및 법적 문제 또한 빈번히 발생하고 있어, 이런 문제들을 예방하기 위한 드론의 탐지 기술이 활발히 연구되고 있다. 본 논문은 카메라로 촬영된 영상에서 조류와 같은 다른 객체와 구별하여 드론을 탐지하는 기술과 상공에서 바라본 객체들을 탐지하는 기술을 구현한다. 제안 방법은 딥러닝 기반의 YOLOv4를 사용하였다. UAV_123 데이터셋으로 학습한 실험 결과, mAP는 85%, Recall은 85%, Precision은 81%의 정확도를 보였다. 제안 방법은 인명 구조, 배송, 건축 뿐만 아니라 안티 드론 시장에서도 효과적으로 활용될 수 있을 것으로 기대된다.

1. 서론

군사용으로 개발된 무인항공기, '드론(Drone)'은 4차 산업혁명 시대에 등장하여 사용 분야가 다양한 분야로 확장되며, 전문가뿐 아니라 일반인까지 널리 사용되고 있다.

이에 따라 드론의 활용영역이 다양화되고, 넓어질수록 역설적으로 사생활 침해 범죄 또한 급증하고 있다. 드론으로 인한 사생활 침해가 사회문제로 떠오르면서 드론을 탐지하여 이와 같은 문제에 대처하려 한다.

특정 영상들을 입력값으로 넣어주어 학습시켜 생성된 모델로 드론과 일반 객체를 구분하여 탐지한다. 딥러닝 기반의 YOLOv4 탐지 알고리즘을 이용하여 드론과 객체를 탐지한다. 객체 탐지에 성공 후, 적절한 대처를 통해 드론을 활용한 물류 관리, 시설 관리, 산림 관리, 실종자 수색 등에 활용될 수 있다. 또한, 드론을 불법적으로 이용하는 사생활 침해, 중요시설 촬영 등에도 대처할 수 있다.

2. 본론

2.1 Data Set & Learning

본 논문의 목적인 드론 탐지를 위해 약 252,409장의 데이터를 확보하였다.

첫 번째 데이터 세트는 다양한 장소에서 촬영할 수 있으며, 고화질의 영상취득이 용이한 UAVs(Unmanned Aerial Vehicles) 데이터 세트를 활용하였다. 이 중 90,039장을 training data로 22,510장을 test data로 사용하였으며, 해당 데이터 세트는 상공에서 지상에 있는 객체들

을 찍은 사람, 차, 자전거, 새 등 외에도 10개의 클래스를 가진 데이터 세트로 이루어져 있다. 이는 다양한 클래스에 대해 구별하여 탐지할 수 있도록 데이터 클래스들을 구성하였다. 또한, 해당 데이터 세트의 좌표 format과 Yolo format이 달라 Yolo format으로 변환하는 작업을 거쳤다. 그리고, 학습하는 과정에서 사용한 Weight 파일은 기존에 학습되어 있는 모델로부터 업데이트하는 방법인 Fine Tuning 하였다.

두 번째 데이터 세트로 학습한 웨이트 파일을 받아 테스트하였다. 이는 직접 촬영한 드론 이미지 44,986장과 drone-vs-bird 챌린지[1]의 94,874장의 드론 이미지를 사용하여 109,360개의 training과 30,500의 test로 랜덤하게 나누어 학습되었다.

2.2 YOLOv4

YOLO는 You Only Look Once의 약자이며, 처음으로 One-Stage-detection 방법을 고안해 객체 탐지를 가능하게 만들었다. 현재 버전은 YOLOv5까지 나왔으며, 본 논문에서는 YOLOv4를 사용하였다.

YOLOv4는 기존 모델에서 Bag of Freebies, Bag of Specials 기법들을 도입했다. BOF는 inference cost의 변화 없이 성능 향상을 꾀하는 기법들이다. 적대적 이미지를 일부러 생성하여 학습시키는 SAT, 서로 관련 없는 분할된 네 개의 이미지를 합쳐서 학습시키는 Mosaic, DropOut처럼 독립적으로 랜덤하게 drop 하는 것이 아니라 feature의 일정 범위를 함께 드랍하는 Drop Block 기법 등이 이에 해당한다. BOS는 Architecture 관점에서의 기법들이 주를 이루고, Post Processing

도 포함되어 있으며, inference cost를 증가시켜 정확도를 높이는 기법들을 의미한다.

또한, YOLO의 고질적인 문제인 small object에 취약하다는 점을 극복하기 위해 Input network resolution을 512 x 512로 늘려 사용했다. 하나의 이미지에서 여러 크기의 오브젝트를 찾는 네트워크의 표현력을 높이기 위해 Parameters를 키워줬다.

아래 사진은 YOLOv4의 구조를 나타낸다.

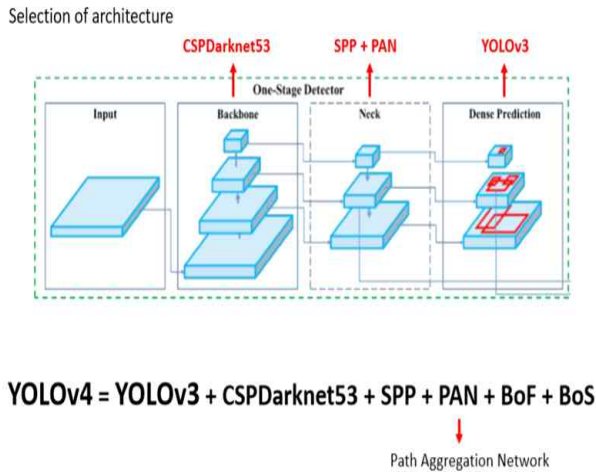


그림 1. YOLOv4의 구조 [2]

2.3 Jetson Tx2

NVIDIA사에서 개발한 임베디드 보드인 Jetson Tx2는 Deep Learning에 특화된 기기이다. 영상 처리나 딥러닝 등을 구현할 수 있을 수준의 고성능 GPU를 탑재하고 있으며, Linux 계열의 OS를 지원하며, 여러 라이브러리와 프레임워크 등의 많은 소프트웨어와의 호환성을 보이는 AI Computing 장치이다.

그리고, NVIDIA에서 제공하는 JetPack은 AI 응용 프로그램을 구축할 때 사용하는 포괄적인 솔루션이다. cuDNN, CUDA 툴킷, TensorRT 플랫폼 소프트웨어를 번들로 제공하여 학습 환경을 바로 사용할 수 있다.

또한, NVIDIA에서 만든 Deep Learning Inference Optimization Library인 Tensor RT는 학습된 딥러닝 모델을 최적화하여 NVIDIA GPU 상에서의 높은 처리량을 제공하는 모델 최적화 엔진이다.

3. 결과

UAV_123 데이터 세트를 사용하여 학습된 네트워크의 모델 성능을 파악했다. 데이터 세트의 20%를 검증 데이터로 사용하여 테스트를 진행하였는데, 해당 검증 데이터로 모델을 평가하기 위한 성능척도는 mAP와 GIOU를 사용하였다. mAP는 85%, Recall은 85%, Precision은 81%의 정확도를 보였고, GIOU는 학습 결과에서 10 epoch에서 0.02에 근접하고 20 epoch에서 0에 수렴하는 수치를 보여준다.

먼저, GIOU란 IoU를 Loss에 활용하기 위해 정의된 Loss function으로 0에 가까울수록 손실이 없다는 것을 의미한다.

mAP는 각 class마다 한 AP를 갖게 되는데, 모든 class의 AP에 대해 평균값을 낸 것을 mAP라고 한다. 여기서 AP는 Precision-Recall의 면적을 의미한다. Precision은 모델이 검출한 것들 중에 옳게 검출한 것을 Precision이라고 하며, Recall은 실제 옳게 검출되어야 할 결과물들 중에서 옳다고 예측한 것의 비율을 의미한다.

데스크탑과 Tx2에서 객체를 탐지한 결과를 비교해보았을 때, 임베디드 보드에서 실시간으로 탐지하는 데 있어 무리가 있는 결과가 나왔다. 데스크탑에서의 GPU 환경은 GTX1060 6G로 평균 FPS 23.83이 나왔지만 임베디드 보드에서의 GPU 환경으로 NVIDIA Pascal, 256 CUDA cores로 평균 FPS 2.84로 실시간 탐지하기 위한 부족한 수치를 보여주었다.

결과적으로 정확한 탐지는 가능했지만 탐지하는 시간적인 부분에서 현저히 떨어지는 모습을 보여주었다. Tx2의 GPU 기술의 고도화와 객체 탐지 알고리즘의 최적화가 개발된다면 실시간 객체 탐지를 기대할 수 있을 것이다.

아래 그림은 Tx2에서 객체 탐지하는 순간의 GPU 사용량을 보여주는 그래프이다.

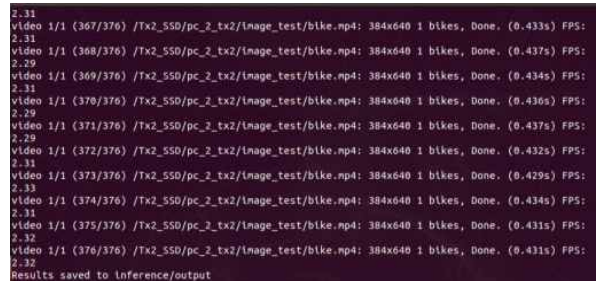


그림 2. Tx2에서 객체 탐지 시 FPS

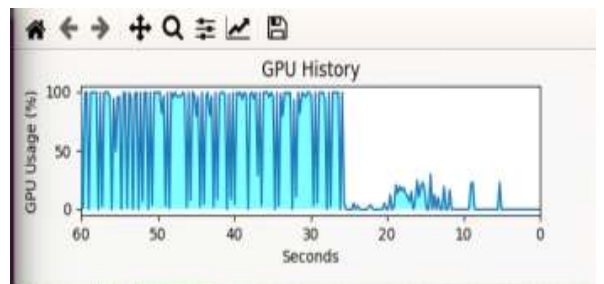


그림 3. Tx2에서 객체 탐지 시 GPU 사용량

하단 그림은 차례로 UAV_123 데이터 세트를 이용하여 학습한 결과 그래프, 해당 모델로 탐지한 결과와 drone-vs-bird 챌린지 데이터 세트를 이용하여 학습된 모델 탐지 결과이다.

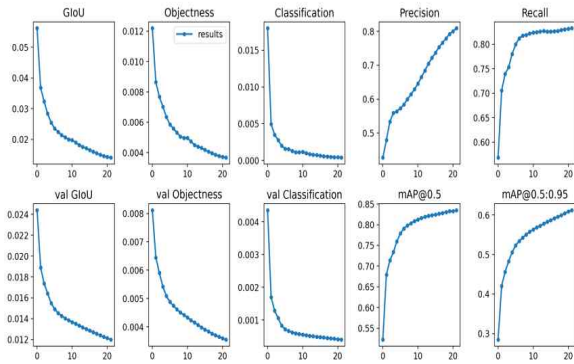


그림 4. UAV_123 데이터 세트를 이용하여 학습한 결과



그림 5. UAV_123 데이터 세트를 이용하여 학습된 모델 탐지 결과



그림 6. drone-vs-bird 챌린지 데이터 세트를 이용하여 학습된 모델 탐지 결과

4. 기대효과

본 작품에서는 상공에서 지상을 찍은 1번 데이터 세트와 드론과 다른 비행체들을 구분할 수 있는 2번 데이터 세트로 구성되어 있으며, 활용 방안도 무궁무진해질 수 있다.

현재 드론은 다양한 분야에서 쓰이고 있으며, 드론의 효용성이 날이 갈수록 증가하고 있다. 예를 들어, 실종자 수색, 교통·수사·범죄 예방, 택배 수송, 산림 관리 등 사실상 활용 영역은 무한대로 넓어지고 있으며, 1번 데이터 세트로 해당 분야들에 활용될 수 있다.

하지만, 장점만 있는 것은 아니다. 이로 인한 위협과 부작용도 함께 증가하고 있는 것 또한 사실이다. 불법적인 촬영으로 인한 사생활 침해, 불법 비행, 중요시설 촬영 등 문제가 끊임없이 터져 나오고 있다.

이에 따른 안티 드론 관련 시장이 급격히 확대되고 있다. 선행되어야 할 필수적인 부분은 해당 드론을 감지할 수 있어야 하며, 2번 데이터 세

트로 이러한 부작용을 해결하는데 활용할 수 있을 것이다.

비록 결과는 임베디드 보드에서의 실시간 탐지는 어렵다고 판단되었지만, Tx2 보드의 실시간 탐지가 가능한 수준으로 성능이 향상되고 검증이 된다면 추후 실시간까지의 가능성을 염두에 뒤 볼 만하다. 본 작품을 토대로 드론 및 객체 탐지에 그치는 게 아니라 보안작업을 거친다면 실시간으로 객체를 탐지하는 기술을 앞에서 언급한 다양한 시장 및 산업 분야에서 활용할 수 있을 것이라 기대한다.

참고문헌

- [1] Angelo Coluccia, Alessio Fascista, Arne Schumann, Lars Sommer, Marian Ghenescu, Tomas Piatrik, Geert De Cubber et al., "Drone-vs-Bird Detection Challenge at IEEE AVSS2019", 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 25 November 2019
- [2] Alexey Bochkovskiy, Chien-Yao Wang and Hong-Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection", In arXiv.org, 23 Apr 2020