

YOLO 모델의 전이학습을 적용한 엣지디바이스의 실시간 객체탐지 성능 분석

이영주

한국폴리텍대학 서울정수캠퍼스 인공지능소프트웨어과
e-mail:young2017@kopo.ac.kr

Real-time Object Detection Performance Analysis on Edge Devices using YOLO Model Transfer Learning

Youngju Lee

Dept. of Artificial Intelligence and Software, Seoul Jungsu Campus of Korea Polytechnics

요약

본 논문에서는 불법어촌 활동 감시 시스템 구축을 위한 사전단계로 엣지디바이스를 사용하여 실시간 객체 탐지 성능을 검증하였다. 현장데이터 부족으로 교내에 유사한 환경을 구축하고 실시간 객체 탐지에 적합한 YOLOv8 모델의 전이학습을 적용하여 엣지 디바이스에 적합한 모델을 검증하였다. Jetson Nano 엣지디바이스에서 x-large 모델을 적용했을 때 상대적으로 높은 성능을 얻었으나 모델의 처리 능력을 고려하여 nano 모델을 적용하여 성능을 분석하고 평가하였다.

1. 서론

최근 자율주행 시스템, 해상에서의 소형 인식 및 재난 예측 경보 등에서 실시간 객체 검출에 대한 연구가 수행되고 있다 [1-2]. 본 연구에서는 해안에서 발생가능한 불법 해루질 및 안전에 대한 경보시스템을 구축하기 위한 선행 단계로 엣지 디바이스를 이용한 실시간 객체탐지 알고리즘을 검증하고자 한다.

엣지디바이스의 실시간 객체 탐지 능력을 검증하기 위해 최근 연구되고 있는 실시간 객체 알고리즘인 YOLO를 도입하여 데이터의 개수별, 세대수 별로 학습된 모델을 사람 객체 클래스를 적용하였다[3]. 1000장의 데이터셋을 라벨링 후 전이학습을 적용하여 딥러닝 기반 소형객체 검출 지표를 검증하였다.

2. 본론

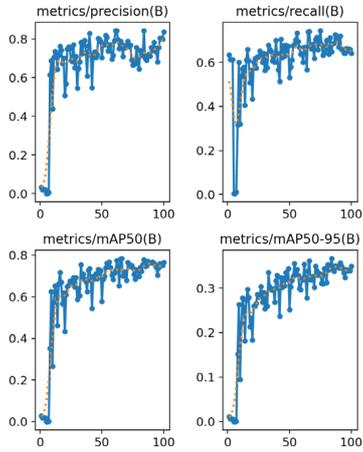
2.1 데이터셋 확보와 전이학습

데이터셋의 부족은 모델의 성능을 제한하는 주요 요인 중 하나입니다. 이를 극복하기 위해 국내 서해안을 중심으로 사람 객체를 포함한 약 100장의 해변 사진을 촬영하여 데이터셋을 확보하였습니다. 소량의 데이터셋을 라벨링 후 학습 결과 20% 미만의 정확도를 보여 다양한 환경에서 적용하기에는 부족하다 판단하였다. 따라서 많은 데이터 셋을 확보하기 위

해 교내 설치된 IP Camers 데이터를 활용하여 유사한 환경을 구축하고 1000장의 데이터셋을 확보하였다. 전이학습을 통해 기존의 YOLO 모델을 해양 환경에 특화된 모델로 변환하였습니다. 전이학습은 검증된 데이터 셋으로 사전 훈련된 모델을 이용하여 적은 양의 데이터 셋으로 훈련함으로써 모델의 성능을 향상시키는 방법이다. YOLOv8은 80개의 클래스를 포함하고 있는 COCO dataset으로 학습되었다. 교내 정문으로 출입하는 방문객을 기준으로 person 클래스를 생성하고 1000여장의 이미지를 라벨링 후 train, validation, test를 각각 600, 200, 200장으로 구분하여 전이학습을 시행하였다. YOLO는 nano, small, medium, large, x-large 모델로 사전 학습된 모델을 제공한다. 표 1은 사전 학습된 모델을 이용하여 교내 데이터셋을 적용한 전이학습 결과이다.

[표 1] 사전학습 모델을 적용한 전이학습 결과

| 사전학습 모델 | metrics/precision (B) | metrics/recall(B) | metrics/mAP50 | metrics/mAP50-95 |
|---------|-----------------------|-------------------|---------------|------------------|
| nano | 0.78204 | 0.53284 | 0.60003 | 0.23945 |
| small | 0.79739 | 0.56272 | 0.66481 | 0.28243 |
| medium | 0.81729 | 0.56847 | 0.6599 | 0.29909 |
| x-large | 0.84398 | 0.74443 | 0.78458 | 0.36678 |



[그림 1] x-large 전이학습모델의 정밀도, 재현율, mAP50, mAP50-95

2.2 모델 성능 평가

교내 CCTV 데이터를 활용하여 학습한 모델을 평가하기 위해 정밀도, 재현율, mAP50, mAP50-95 등의 지표를 사용하였으며 평가 결과는 그림 1과 같다. 엡지디바이스는 Nvidia사의 Jetson Nano를 적용하였다. Jetson Nano는 128개의 GPU 코어를 갖고 있어 지능형 영상처리 엡지디바이스로 사용되고 있다. validation 결과로는 x-large 모델을 적용시 정밀도 0.84, 재현율 0.74로 상대적으로 높은 값을 얻었다. 표 1의 결과에 따르면 large 모델부터 초당 프레임수가 0.6으로 엡지디바이스인 Jetson Nano에서 처리가 불가능함을 보이고 있어 프레임 처리 능력을 고려하면 실제 사용가능한 영역은 nano 모델이 적합할 것으로 판단된다. 그림 2는 nano 모델의 전이학습을 적용하여 실시간 측정된 영상이다. 대부분의 person 객체를 탐지하고 있다.

[표 2] 전이학습모델 적용 엡지디바이스 영상처리초당 프레임수

| 사전학습모델 | FPS(초당프레임수) |
|--------|-------------|
| nano | 6.4 |
| medium | 1.5 |
| large | 0.6 |



[그림 2] 교내 정문 방문객 탐지 결과

3. 결론

본 연구에서는 해안에서 발생 가능한 불법 해루질 및 안전에 대한 경보시스템을 구축하기 위한 선행 단계로 엡지디바이스의 실시간 객체 탐지 성능을 검증하였다. Jetson Nano 모델을 적용 시 YOLOv8 사전 학습 모델 중 가장 적은 nano 모델을 적용하는 것이 적합할 것으로 판단된다. 향후 다양한 데이터셋을 확보하여 엡지디바이스에 적합한 아키텍처와 하이퍼파라미터를 더욱 세밀하게 조정 및 성능 향상에 대한 연구가 필요할 것으로 판단된다.

참고문헌

- [1] 김하선, 김아란, 라해운, 김선영 “전이학습을 적용한 YOLOv8 기반 객체 검출 성능 분석”, 대한기계학회춘추 학술대회, Vol.2023. No.6, 2023년.
- [2] 김기관,노천명,이수봉,이순섭,이재철“해상에서 소형객체 인식을 위한 데이터 전처리 방안 연구”, 한국CDE학회는 문지, Vol.26. No.4, pp. 366-375, 12월, 2021년.
- [3] <https://docs.ultralytics.com/>