

CBAM을 적용한 YOLOv7 기반 선박 분류 모델 연구

성은산¹, 박진호¹, 권태현², 김상우³, 조희정³, 윤주상*

동의대학교 인공지능학과¹, 동의대학교 IT융합학과², (주)서안에스앤씨³, *동의대학교 인공지능학과
ses0137112@naver.com, jhpark3679@gmail, dhxk9513@gmail.com, woogigi@hanmail.net,
shineeej@naver.com, *jsyoun@deu.ac.kr

A Study on the Improved YOLOv7 Model based on CBAM for Ship Classification

Seong Eun San, Jinhyo Park, Taehyeon Kweon, Kim Sang Woo, Cho Hee Jeong, Joosang
Youn*

Dept. of Artificial Intelligence, Graduate School, Dong-eui Univ. ¹, Dept. of IT Convergence,
Graduate School, Dong-eui Univ. ², SEOAHN S&C Co., Ltd.³, *Dept. of Artificial Intelligence,
Dong-eui Univ.

요약

본 논문은 YOLOv7 기반 선박 분류 모델의 배경에서 선박과 비슷한 특성을 가진 물체를 선박으로 인식하는 문제점을 개선한 선박 분류 모델을 제안한다. 이러한 문제점을 해결하기 위하여 YOLOv7에 어텐션 모듈을 적용했다. 실험 결과 어텐션 모듈을 적용한 YOLOv7 기반 선박 분류 모델은 선박 분류 정확도가 일부 향상되었을 뿐만 아니라 배경에서 선박과 비슷한 특성을 가진 물체를 선박으로 인식하는 문제점도 5% 개선되었다.

I. 서론

YOLOv7은 최근 많이 사용되는 객체 인식 모델 중 하나이다[1]. 이러한 객체 인식 모델은 선박 분류 모델로 사용될 수 있다. 선박 분류 모델은 항만에서 수집한 동영상으로 선박 분류를 위한 데이터셋을 구축하고 YOLOv7을 사용하여 학습한다. YOLOv7 기반의 선박 분류 모델을 항만 CCTV 동영상에 적용했을 때 배경에서 선박과 비슷한 특성을 가진 물체를 선박으로 인식하는 문제점이 발생했다[2].

이러한 문제를 해결하기 위하여 본 논문에서는 CNN의 성능향상 기법의 하나인 CBAM(Convolutional Block Attention Module)을 활용하여 YOLOv7 기반 선박 분류 모델의 성능을 향상하는 방법을 제안한다.

II. CBAM을 적용한 YOLOv7 모델

가. YOLOv7

객체 인식 모델은 크게 두 가지로 나뉜다. 영상에서 객체의 위치를 찾는 Localization과 객체가 어느 클래스 인지 분류하는 Classification을 두 단계에 걸쳐서 수행하는 Two-stage 방식과 이를 동시에 수행하는 One-Stage 방식이다. YOLOv7은 One-stage 객체 인식 모델 중 하나로 영상을 그리드로 분할하고 각 그리드에 대한 경계 상자와 클래스 확률을 예측한다. 이렇게 예측된 경계 상자 중에서 신뢰도가 높은 경계 상자를 선택하는 방식으로 객체를 인식해서 처리 속도가 빠르고 정확도가 높다.

나. YOLOv7 기반 선박 분류 모델

YOLOv7 기반 선박 분류 모델은 항만 CCTV 동영상으로 구축한 선박 분류를 위한 데이터셋으로 학습했다. 선박 분류를 위한 데이터셋은 선박이 바지선, 예인선, 급유선, 기타 총 4개의 클래스로 분류되어 있다. 선박 분류 데이터셋으로 학습한 YOLOv7 기반 선박 분류 모델은 클래스별로 바지선은 1.0, 예인선은 0.96, 급유선은 0.98, 기타 선박은 0.96의 정확도를 보였다[2].

하지만 그림 1과 같이 데이터셋에 사용되지 않은 항만 CCTV 동영상에 적용했을 때 배경에서 선박과 비슷한 특성을 가진 물체를 선박으로 인식하는 문제점이 발생했다.

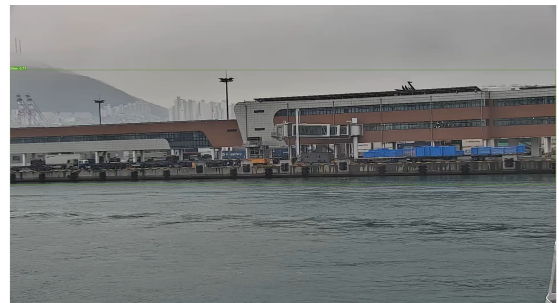


그림 1. YOLOv7 기반 선박 분류 모델이 배경에서 기타 선박을 검출한 결과 이는 성능지표의 Background FP(False Positive)[3]로도 확인되어 본 논문에서는 이러한 현상을 해결하기 위해 데이터셋을 증강하고 YOLOv7 백본에 어텐션 모듈을 적용했다.

다. CBAM의 구조

CBAM(Convolutional Block Attention Module)은 그림 1과 같이 채널 어텐션 모듈과 공간 어텐션 모듈로 구성되어 있다[4].

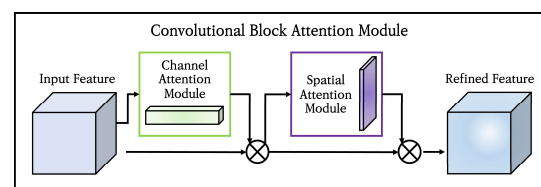


그림 2. CBAM의 구조

어텐션 모듈은 채널과 공간 각각에서 어텐션 맵을 생성해서 모델이 무엇을, 어디를 봐야 하는지 알려주어 불필요한 정보는 억제하고 중요한 정보

는 강조한다.

라. 제안하는 방법

YOLOv7의 백본은 다양한 크기를 가진 객체의 특징을 추출하기 위해 원본영상에서 CBS(Convolution+Batch N+SiLU)를 사용한다. 영상의 크기 별로 합성곱 연산을 하고 ELAN(Efficient Layer Aggregation Network) 구조로 추출된 특징을 합하여 Max Pooling으로 추출된 특징을 강조하고 해상도를 줄여 연산 비용을 감소시킨다. 본 논문에서는 추출된 선박의 특징을 강조하고 배경의 특징은 억제하기 위해 그림 3과 같이 CBAM을 추가하여 모델이 배경과 객체 중 어디에 집중해야 하는지 알려주어 신뢰도가 낮은 배경의 특징이 강조되지 않는 구조를 제안한다.

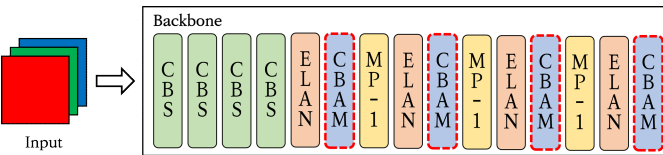


그림 3. CBAM을 적용한 YOLOv7의 백본

III. 실험

가. 데이터셋

배경에서 선박과 비슷한 특징을 가진 물체를 선박으로 인식하는 문제점을 개선하기 위해 기존 데이터셋의 반대쪽에서 수집된 동영상으로 데이터셋을 추가로 5,387장 구축했다. 이를 기존 7,438장인 데이터셋에 통합해 총 12,825장으로 증가했다. 통합된 데이터셋은 CBAM을 적용한 YOLOv7 모델과 기존 YOLOv7 모델 학습에 사용 하기 위해 train은 0.8, validation은 0.1, test는 0.1의 비율로 분할했다.

하이퍼파라미터는 기존 YOLOv7 모델과 CBAM을 적용한 YOLOv7 모델 모두 동일하게 learning rate는 0.001, batch size는 32, epoch는 300, image size는 640으로 설정했다.

나. 실험결과

통합된 데이터셋을 기존 YOLOv7으로 학습한 결과 여전히 배경에서 선박이 검출되었다. 이에 앞서 제안한 CBAM을 적용한 YOLOv7 기반 선박 분류 모델을 통합된 데이터셋으로 학습했고 결과는 표 1과 같다.

box_loss	0.012	precision	0.9876
obj_loss	0.004844	recall	0.9565
cls_loss	0.0004027	map@.5	0.9736
total_loss	0.01725	map@.5-95	0.7367

표 1. CBAM을 적용한 YOLOv7 기반 선박 분류 모델 학습 결과

Precision, Recall, map@.5, map@.5-95 모두 통합된 데이터셋으로 학습한 YOLOv7 기반 선박 분류 모델과 비슷한 결과를 보인다. 이를 정확도를 기준으로 비교한 것은 표 2와 같다.

	YOLOv7	YOLOv7+CBAM
바지선	0.91	0.94
예인선	0.99	0.96
급유선	0.99	0.99
기타	1.0	1.0
급유선 Background FP	0.75	0.70

표 2. YOLOv7과 CBAM을 적용한 YOLOv7의 정확도 및 Background FP 비교

선박 분류 정확도는 표 2와 같이 전체적으로 유지되었고 그림 4의 (좌)와 같이 선박이 잘 분류 되었다.



(좌)선박 분류 결과, (우)배경에서 선박이 검출되지 않은 결과

그림 4. CBAM을 적용한 YOLOv7 모델의 추론 결과

YOLOv7 기반 선박 분류 모델의 문제점이었던 급유선의 Background FP는 통합된 데이터셋으로 CBAM을 적용한 YOLOv7 모델을 학습한 결과 5% 개선되었다. 또한 그림 1과 같은 시점의 영상을 CBAM을 적용한 YOLOv7 모델로 추론했을 때 그림 4의 (우)와 같이 배경에서 선박이 검출되지 않았다.

IV. 결론

본 논문에서는 YOLOv7에 CBAM을 적용해 선박 분류 정확도를 개선하는 방법을 제안했다. 실험 결과 급유선의 Background FP를 5% 개선했고 배경에서 선박이 더 이상 검출 되지 않았다. 추후 Background FP 개선을 위해 데이터셋에 선박이 없는 배경 영상을 추가하고 클래스 불균형을 개선해 제안한 방법을 적용할 예정이다. 또한 YOLOv9에도 CBAM을 적용하여 성능을 비교할 예정이다.

ACKNOWLEDGMENT

본 논문은 과학기술정보통신부 및 정보통신기획평가원의 정보통신방송표준개발지원사업(RS-2024-00397768)과 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구결과임.

참 고 문 헌

[1] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023.

[2] 성은산, 윤주상. 항만 출입 선박 관리 자동화를 위한 딥러닝 기반 선박 분류 모델 연구. 한국산업정보학회, 2024.

[3] Jocher, Glenn. Ultralytics YOLOv5. Version 7.0, 2020, <https://github.com/ultralytics/yolov5>. Accessed 10.5281/zenodo.3908559.

[4] Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." Proceedings of the European conference on computer vision (ECCV). 2018.