

# 건설 현장을 위한 자율주행 로봇의 실시간 3D 객체 인지 알고리즘 구현

최지예<sup>1</sup>, 최인구<sup>2</sup>, 홍형근<sup>3</sup>, 전재욱<sup>3</sup>

<sup>1</sup>성균관대학교 바이오메카트로닉스학과

<sup>2</sup>HL 만도 SW Campus

<sup>3</sup>성균관대학교 전자전기컴퓨터공학과

jiyea0217@g.skku.edu, ingu.choi@hlcompany.com, whaihong@g.skku.edu, jwjeon@skku.edu

## Real-time 3D Object Perception Algorithm Implementation for Autonomous Driving Robots at Construction Sites

Ji-Ye Choi<sup>1</sup>, In-Gu Choi<sup>2</sup>, Hyeong-Keun Hong<sup>3</sup>, Jae-Wook Jeon<sup>3</sup>

<sup>1</sup>Department of Bio-Mechatronics, Sungkyunkwan University

<sup>2</sup>SW Campus, HL Mando

<sup>3</sup>Department of Electrical and Computer Engineering, Sungkyunkwan University

### 요 약

건설 현장에서 자율주행 로봇의 안전한 주행을 위해 동적 장애물의 정확한 인식 및 추적이 중요하다. 본 논문에서는 실시간 3D 객체 인식 및 추적을 위한 방법을 제안한다. Complex-YOLOv4 모델을 이용한 객체 인식, SORT 알고리즘 확장을 통한 객체 추적을 구현하였다. Jetson AGX Orin 보드의 ROS2 환경에서 시스템을 구축하여, 실시간 3D 객체 인식 및 추적이 가능함을 확인하였다.

### 1. 서론

건설 현장에서 자율주행 로봇의 안전한 주행을 위해서는 동적 장애물을 정확하게 인지하는 과정이 필수적이다. 특히, 사람과 차량, 로봇, 중장비의 정확한 인식과 추적이 중요하다. 자율주행 로봇의 동적 장애물 회피 경로 생성을 위해서는 우선적으로 임베디드 환경과 딥러닝 모델을 고려한 실시간성이 요구된다. LiDAR 와 카메라를 융합한 방법은 실시간성 보장이 어렵고, 카메라를 단독으로 사용하는 방법은 시각적 변화에 취약하다. 따라서, 본 논문에서는 LiDAR 기반의 실시간 객체 인식 및 추적을 제안한다. 객체 인식은 Complex-YOLOv4[1]을 활용하고, 객체 추적은 SORT[2]를 3D 로 확장한다. 또한, 로봇 환경에서 실행 되도록 Jetson AGX Orin 보드에서 시스템을 구축한다.

### 2. 선행 연구

#### 2.1 Complex-YOLOv4

Complex-YOLOv4 는 LiDAR 를 기반으로 실시간 3D 객체 인식을 수행하는 딥러닝 모델이다[1]. Complex-YOLOv4 는 세 단계로 이루어진다. 첫 번째는 LiDAR 데이터 전처리 과정이다. 관심영역에 대해 BEV(Bird-Eye-View) RGB map 으로 변환한다. 두 번째는

Complex-YOLOv4 를 활용한 추론 과정이다. 2D 객체 인식 모델인 YOLO 와의 차이점은 E-RPN 네트워크를 도입하여 Bounding Box 의 각도에 대한 Regression 으로 Heading 방향을 추정할 수 있다는 것이다. 결과적으로 Bounding Box 의 x, y, w, h, angle 을 얻을 수 있다. 세 번째는 3D Bounding Box Reconversion 이다. BEV 이미지에서 예측한 2D Bounding Box 를 3D 로 변환하는 과정이다. 2D Bounding Box 좌표를 LiDAR 기준의 3차원 좌표계로 변환한다. z 값은 Class 마다 사전에 정의한 높이를 반영한다. 카메라 이미지에 시각화하기 위해 LiDAR 기준 좌표계를 카메라 좌표계로 변환한다.

#### 2.2 SORT

SORT(Simple Online Real-time Tracking)는 실시간 2D 다중 객체 추적 알고리즘이다[2]. SORT 는 두 단계로 이루어진다. 첫 번째는 객체 인식 결과에 Kalman Filter 를 적용하여 다음 프레임에서의 객체 위치를 예측하는 과정이다. 두 번째는 Hungarian 알고리즘을 활용한 객체 매칭 과정이다. Kalman Filter 예측 결과와 현재 프레임의 객체 인식 결과를 IoU(Intersection over Union)를 기반으로 매칭한다. 객체에 ID 를 부여하고, 객체 업데이트, 등록, 삭제를 통해 추적이 이루어진다.

### 3. 시스템 구성

#### 3.1 자율주행 로봇

16 채널 LiDAR, 스테레오 카메라가 장착된 Scout mini 모델이며, Jetson AGX Orin 보드로 시스템이 구현된다. 로봇 구성은 다음 (그림 1)과 같다.



(그림 1) 자율주행 로봇의 센서 구성

#### 3.2 3D 객체 인식

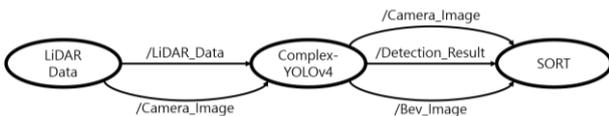
Complex-YOLOv4 를 자체 데이터셋으로 학습하여 LiDAR 센서만으로 실시간 3D 객체인식을 구현하였다. Rosbag 형식의 원본 데이터를 전처리하여 카메라 및 LiDAR 데이터를 추출하고, roLabelImg 라벨링 툴을 활용하여 Label 파일을 구성하였다. Complex-YOLOv4 에서 구성한 Class(사람, 자전거, 자동차)를 로봇에서 인식하려는 Class(사람, 로봇, 자동차, 중장비)로 변경하였다. epoch 은 4, batch 는 200, weight decay 는 0.0005, 그리고 momentum 은 0.9 로 설정하여 700 장의 데이터셋으로 학습하였다.

#### 3.3 3D 객체 추적

동적 객체 정보를 정확히 파악할 수 있도록 2D 객체 추적 알고리즘인 SORT 를 3D 로 확장하였다. 객체 인식으로 얻은 Bounding Box 의 정보를 모두 반영하여 정확한 추적이 이루어지도록 수정하였다. 객체 간 매칭 과정에서 Bounding Box 의 헤딩 방향을 반영하여 IoU 를 계산하도록 하였다. 또한, BEV 이미지에서의 Detection 결과를 바탕으로 추적한 후, 3D Reconversion 과정을 거치도록 수정하여 3 차원으로 확장하였다.

#### 3.4 Jetson AGX Orin 환경에서의 시스템 구축

로봇에는 Jetson AGX Orin 보드를 탑재하였다. 보드에는 ROS2 환경으로 (그림 2)와 같은 시스템을 구성하여 실시간으로 객체를 인식하고 추적하도록 하였다.



(그림 2) Jetson AGX Orin 환경 시스템 노드 그래프

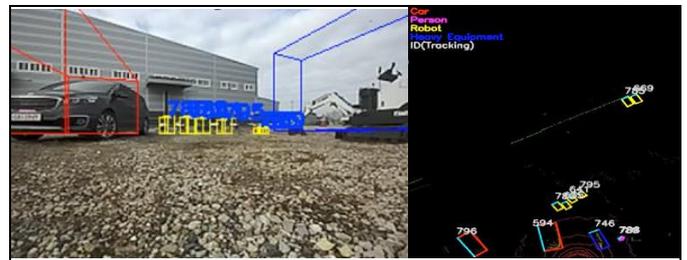
### 4. 시스템 구현 결과

자체 데이터셋으로 학습을 진행한 객체 인식 모델의 정확도를 평가한 결과 mAP 는 0.7449, inference time 은 39.56fps 이다. 평가 결과는 <표 1> 과 같다.

<표 1> 학습 후 모델 평가 결과

Class	Precision	Recall	AP	F1
Car	0.8265	0.895	0.8757	0.8594
Person	0.596	0.8912	0.7479	0.7143
Robot	0.6066	0.7569	0.5784	0.6735
Heavy Equipment	0.7	0.8235	0.7776	0.7568

자체 데이터셋 샘플에 대해 3D 객체 인식과 추적을 수행한 결과는 (그림 3)와 같다. 객체 인식 결과인 Bounding Box 를 카메라 이미지와 LiDAR BEV 이미지에 나타내었고, 객체 추적 과정에서 객체마다 부여한 ID 를 Bounding Box 의 상단부에 나타냈다. Jetson AGX Orin 에서 시스템을 구현한 결과, 실시간으로 3D 객체 인식 및 추적이 수행되는 것을 확인하였다.



(그림 3) 3D Object Detection 및 Tracking 수행 결과

### 5. 결론

본 연구는 건설 현장에서 실외 자율주행 로봇의 주행 환경 인지를 위한 3D 객체 인식 및 추적 방법을 제안하였다. 3D 객체 인식에는 Complex-YOLOv4 모델을 로봇 자체 데이터셋을 구축하여 학습을 다시 진행하였고, 3D 객체 추적은 SORT 알고리즘을 3 차원으로 확장하였다. Jetson AGX Orin 의 ROS2 환경으로 시스템을 구축하여 로봇에서 실시간 3D 객체 인식 및 추적이 가능함을 확인하였다. 하지만, 객체 인식에서 높이를 추론할 수 없고, 객체 인식 정확도가 추적 정확도에 영향을 미친다는 점에서 추가 연구가 필요하다.

### ACKNOWLEDGMENT

이 논문은 정부(교육부-산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0022098, 2024 년 미래형자동차 기술융합혁신인재양성사업)

### 참고문헌

[1] Simon, M., Milz, S., Amende, K., & Gross, H.-M., "Complex-YOLO: An Euler-Region-Proposal for Real-time 3D Object Detection on Point Clouds", ECCV, 2018, 197-209.

[2] Bewley, Alex and Ge, Zongyuan, Ott, Lionel and Ramos, Fabio, Upcroft, Ben, "Simple Online and Realtime Tracking", ICIP, 2016, 3464-3468.