

# OODIS: 산업 안전을 위한 전방위 객체 탐지 시스템

## OODIS: Omnidirectional Object Detection for Industrial Safety

윤지일<sup>1</sup> · 노초환<sup>1</sup> · 신초희<sup>1</sup> · 전제환<sup>1</sup> · 신명지<sup>1</sup> · 박재일<sup>1</sup> · 이태용<sup>2</sup> · 복형준<sup>2</sup> · 장지운<sup>2</sup> ·  
이지윤<sup>3</sup> · 이민호<sup>3</sup> · 조영근<sup>†</sup>

Jiil Yun<sup>1</sup>, Chohwan Ro<sup>1</sup>, Chohee Shin<sup>1</sup>, Jehwan Jeon<sup>1</sup>, MyungJi Shin<sup>1</sup>, Jaeil Park<sup>1</sup>, Taeyong Lee<sup>2</sup>, Hyeongjun Bok<sup>2</sup>, Jiwoon Jang<sup>2</sup>, Jiyun Lee<sup>3</sup>, Minhoo Lee<sup>3</sup>,  
Younggun Cho<sup>†</sup>

**Abstract:** In industrial environments, vehicle structures restrict the operator's visibility, increasing the risk of collisions with nearby workers. Existing multi-sensor fusion systems attempt to mitigate this risk but require sensor-dependent mounting and struggle to integrate multiple sensors with limited, non-overlapping fields of view. To address these limitations, we introduce OODIS, a platform-independent system combining a 360° camera and a 360° LiDAR for comprehensive omnidirectional perception. We deploy OODIS on both a forklift and a mobile robot to evaluate its object detection and tracking capabilities across real-world factory, indoor, and outdoor environments. Using mAP and MOTA metrics, the results demonstrate consistent and robust performance across these diverse settings.

**Keywords:** Omnidirectional object detection, Platform-Independent sensor mount, Real-World factory

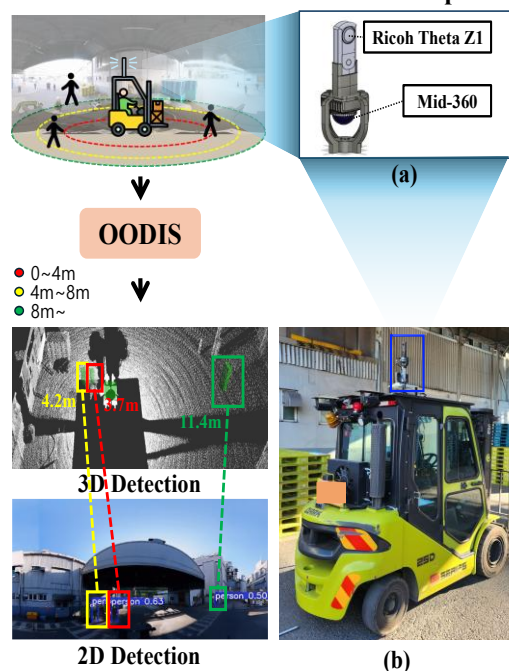
### 1. 서론

산업 현장에서 지게차 운전자는 운전석 위치와 차체 구조로 인해 시야가 제한되며, 이는 작업자를 제때 인지하지 못해 충돌 사고로 이어진다. 이러한 문제를 완화하기 위해 다중 센서를 결합해 관측 영역을 확장하는 접근이 제안된다. 그러나 센서 간 상호 가림을 방지하기 위해 각 센서를 별도의 위치에 장착해야 한다는 구조적 제약이 존재하며, 이는 실제 산업 환경에서의 적용성을 제한하는 요인으로 작용한다.

이러한 한계를 해소하기 위해, 최근에는 360° 카메라와 라이다 등 전방위 감지 센서를 활용해 넓은 관측 범위를 확보하는 접근이 주목받는다. 본 논문에서는 360° 카메라와 라이다를 기반으로 다중 센서 구성의 제약을 제거한 전방위 객체 탐지 시스템(OODIS)을 제안한다. 제안된 시스템은 CJ 공장 및 학교의 실내·외 등 다양한 실제 환경

에서 지게차와 로봇을 주행하며 성능을 평가한다. 또한 객체까지의 거리에 기반한 위험도 시각화 기능을 통해 OODIS의 범용성과 실용성을 검증한다.

### Omnidirectional Sensor Platform Independent



[Fig. 1] Overview of the Proposed System (OODIS)

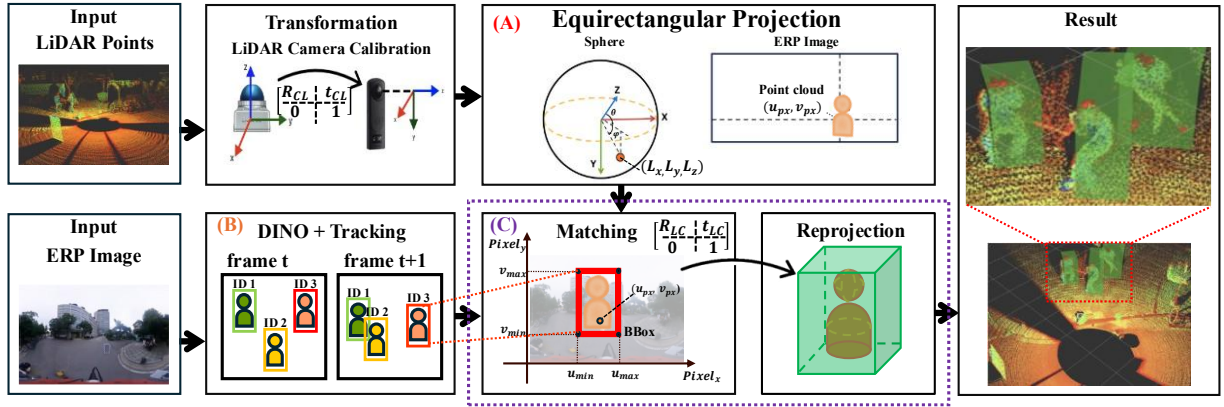
※ This project was funded by Korea Robotics Society (KROS), and is currently supported by the publication grant

1. Undergraduate Student, Inha University, Incheon, Korea,  
wldlf2819@naver.com, chghks430@inha.edu,  
chohui7876@naver.com, smgmsg12@gmail.com  
12200519@inha.edu, jaejae5220@inha.edu

2. Incheon Plant, CJ CheilJedang, Incheon, Korea,  
taeyong.lee@cj.net, hj.bok@cj.net, jiwoon.jang1@cj.net

3. Master Student, Inha University, Incheon, Korea,  
jiwoo6315@gmail.com, mino@inha.edu

† Advisor, Inha University, Incheon, Korea (yg.cho@inha.ac.kr)



[Fig. 2] Pipeline of the proposed system

## 2. 본 론

### 2.1 센서 마운트

본 연구에서 사용한 센서 마운트는 360° 카메라와 라이다를 통합하여 다중 센서 구성에서 발생할 수 있는 제약을 최소화한다. 360° 센서는 자체적으로 전방위 관측이 가능하므로, 별도의 센서 배치를 추가적으로 고려할 필요가 없으며 구조적으로 간결한 형태로 구현한다. [Fig. 1-(a)]는 제안 시스템의 센서 마운트 구성을, [Fig. 1-(b)]는 실제 지게차에 장착된 형태를, [Fig. 3]는 실제 SCOUT 로봇에 장착된 형태를 나타낸다.

### 2.2 3차원 전방위 객체 탐지 시스템

본 시스템은 Equirectangular Projection(ERP)<sup>[1]</sup>를 통해 3차원 라이다 포인트를 2차원으로 변환하여 연산 복잡도를 줄이고, 왜곡이 심한 전방위 영상에서 안정적인 탐지를 수행하기 위해 DINO<sup>[2]</sup>와 OC-SORT<sup>[3]</sup>를 활용한다. 이후 2차원 평면의 기하학적 경계조건을 기반으로 포인트와 객체에 대응 관계를 확립한다.

시스템의 입력은 라이다 포인트와 카메라 이미지이며, 라이다 포인트들은 ERP 투영을 위해 라이다 좌표계에서 카메라 좌표계로 변환된다. [Fig. 2-(A)] ERP는 3차원 좌표( $L_x, L_y, L_z$ )를 구면 좌표계로 변환한 후, 2차원 이미지 좌표( $u_{px}, v_{px}$ )로 매핑하는 투영방식이다. 투영 관계는 다음과 같이 정의되며,  $W$ 와  $H$ 는 투영 이미지의 폭과 높이를 의미한다.

$$u_{px} = \frac{W}{2\pi} \arctan\left(\frac{L_y}{L_x}\right) + \frac{W}{2} \quad (1)$$

$$v_{px} = \frac{H}{\pi} \arctan\left(\frac{L_z}{\sqrt{(L_x)^2 + (L_y)^2}}\right) + \frac{H}{2} \quad (2)$$

이를 통해 3차원 라이다 포인트를 2차원 이미지 공간으로 변환하여 연산 복잡도를 감소시킨다.

[Fig. 2-(B)] DINO는 self-attention 매커니즘을 통해 전역

적 맥락 정보를 활용하여 왜곡이 큰 ERP 이미지에서도 강건한 객체 특징을 추출한다. OC-SORT는 프레임 간 객체의 연속성을 유지하여 일관된 bounding box(bbox)를 제공한다. 각 bbox는  $(u_{min}, v_{min}), (u_{max}, v_{max})$ 로 정의되며, 이후 라이다 포인트와의 매칭 단계에서 기하학적 제약 조건으로 활용된다.

[Fig. 2-(C)] Matching은 투영된 포인트 좌표( $u_{px}, v_{px}$ )가 각 bbox의 경계조건  $u_{min} \leq u_{px} \leq u_{max}$ 와  $v_{min} \leq v_{px} \leq v_{max}$ 을 만족하는지를 검사하여 포인트와 객체 간 대응 관계를 결정한다. 이러한 경계조건 기반 Matching은 복잡한 3차원 탐색을 수행하지 않고도 2차원과 3차원의 결합과정을 단순화하는 장점을 갖는다. 대응된 포인트들은 라이다 좌표계로 복원되어, 각 객체까지의 실제 거리 산출에 활용된다.

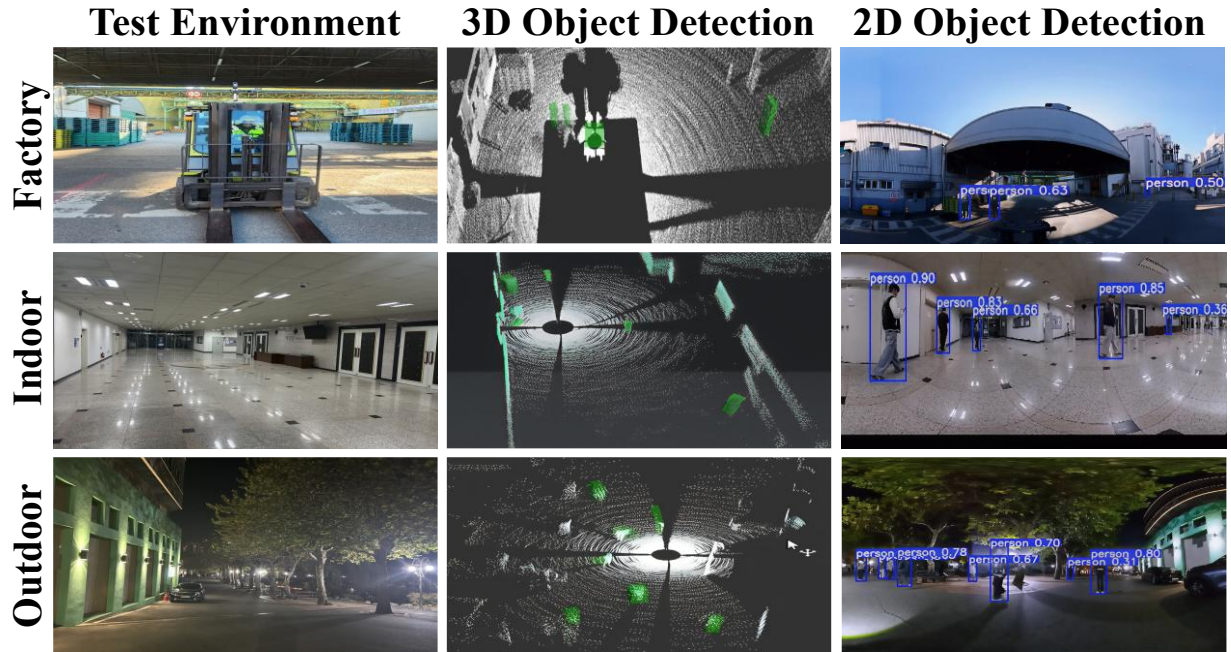


[Fig. 3] Sensor mount on SCOUT Robot

## 3. 실험

### 3.1 실험 환경

본 실험에서는 360° 카메라(Ricoh Theta Z1)와 라이다(Mid-360)를 사용하여 다양한 환경에서 데이터를 취득한다. [Fig. 4]의 첫 열은 공장, 실내, 실외 환경의 전경을



[Fig. 4] OODIS's results for three test environments: The first column presents results from the factory, the second from the school indoor environment, and the third from the outdoor environment at night

나타내며, 공장 데이터는 실제 CJ 제조 공장에서 수집된 것으로 산업 현장의 조건을 직접 반영한다. 공장 환경에서는 지게차를, 실내·외 환경에서는 SCOUT 로봇을 실제 주행시켜 데이터를 수집한다. 수집된 영상은 10초 단위로 분할한 뒤, 15프레임으로 샘플링하여 평가에 활용한다.

AP(Average Precision)<sup>[4]</sup>와 CD(Centroid Distance)산출을 위해 모든 프레임의 사람에 대해 bbox를 수동으로 라벨링하고 이를 정답 데이터셋으로 활용한다. IDF1(Identification F1-Score)와 MOTA(Multiple Object Tracking Accuracy)<sup>[5]</sup>의 산출을 위해 동일 인물에 대해 프레임 간 ID가 일관되도록 수동으로 라벨링을 수행한다.

### 3.2 실험 결과 및 평가

본 연구에서는 제안된 시스템의 성능을 다각적으로 분석하기 위해 정량적 실험과 정성적 실험을 병행하여 평가한다. 정량적 실험은 객체 탐지 및 추적 성능을 수치적으로 검증하기 위한 것이며, 정성적 실험은 다양한 환경에서의 실제 주행 상황을 통해 시스템의 동작 특성과 실용성을 확인하기 위한 목적을 가진다.

#### 3.2.1 정량적 평가

[Table 1] Detection Performance Across Environments

	Factory	Indoor	Outdoor
AP	0.588	0.700	0.095
CD (pixel)	24	27	30

[Table 2] Tracking Performance Across Environments

	Factory	Indoor	Outdoor
MOTA	89.1	76.7	53.4
IDF1	83.5	84.9	67.2

[Table 1]에서 탐지 성능을 보면, 실내 환경의 AP가 0.700으로 가장 높고, 공장 환경은 0.588 정도의 수준을 유지한다. 실외 환경은 배경의 조도가 낮은 이유로 AP가 0.095까지 낮아지지만, CD는 세 환경 모두 24-30px 범위로 유사하여 탐지 위치의 일관성이 유지됨을 확인한다.

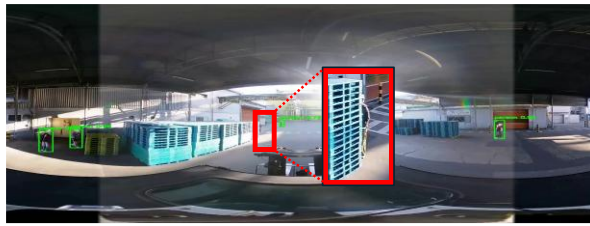
[Table 2]의 추적 성능에서는 공장 환경의 MOTA가 89.1%로 가장 높으며, 실내는 76.7%, 실외는 53.4%를 기록한다. IDF1 역시 실내 84.9%와 공장 83.5%에서 높게 나타나며, 실외 환경에서만 감소한다. 이는 제안된 시스템이 산업 현장 및 실내 환경에서는 일관된 탐지 및 추적 성능이 유지됨을



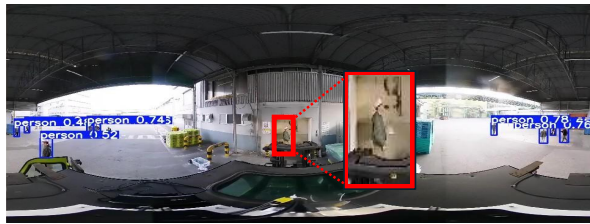
보여준다.

### 3.2.2 정성적 평가

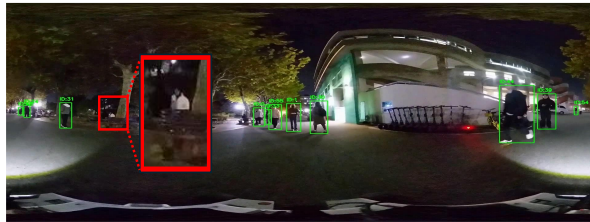
[Fig. 4]는 제안된 OODIS 시스템이 세 환경에서 실제 주행을 통해 획득한 결과를 시각적으로 나타낸다. 첫 번째 열은 각 실험 환경의 전경을, 두 번째 열은 3차원 공간상에서 객체 탐지 결과를 제시한다. 세 번째 열은 ERP기반 2차원 투영면에서의 객체 탐지 결과를 보여준다. 이를 통해 제안된 시스템의 환경 적응성을 시각적으로 보여준다.



(a)



(b)



(c)

[Fig. 5] The Limitation of Proposed System (failure case): (a) The person is vertically blocked (b) The person is horizontally blocked (c) The person is in low light conditions

### 3.2.3 한계점

[Table 1]의 CD는 환경 간 안정적인 위치 산출을 보인다. AP는 실내, 공장 등에서 높으나 실외에서는 크게 감소한다. 이는 실내, 공장 환경이 일정한 조도와 단순한 배경을 갖는 반면, 실외 환경은 조도 변화와 복잡한 배경, 저조도 노이즈 등의 요인으로 탐지 난이도가 증가하기 때문이다.

[Table 2]의 추적 성능은 [Table 1]의 탐지 성능과 유사한 환경별 변화를 보인다. MOTA는 AP의 경

향과 일치하며, IDF1 역시 실내와 공장에서 높게 유지되지만 실외에서는 크게 감소한다. 이는 실외 환경에서 탐지 정확도가 저하되면서 추적 성능 또한 함께 감소한 결과로 해석된다. 실외의 복잡한 배경과 조도 변화는 객체 연관 과정에서 ID 스위칭과 추적 단절을 유발해 전체 추적 성능 저하로 이어진다.

공장 환경에서는 장비와 구조물에 의한 가림으로 탐지가 불안정하며, 이는 [Fig. 5-(a)]의 세로 가림과 [Fig. 5-(b)]의 가로 가림 사례에서 확인된다. 또한 야간 조건에서는 조도 저하로 인해 [Fig. 5-(c)]와 같이 탐지 및 추적 성능이 감소한다.

## 4. 결 론

본 논문에서는 ERP를 활용 및 DINO와 OC-SORT 모듈을 적용한 전방위 객체 탐지 시스템(OODIS)을 제안한다. 공장과 실내 환경에서 수행한 실험에서는 안정적인 탐지 및 추적 성능이 확인되나 가림 및 저조도 상황에서는 성능 저하가 발생한다. 향후 연구에서는 가림 및 저조도 환경에서의 성능 저하 문제를 개선하기 위해 가림 인지 기반 탐지 기법과 저조도 영상 처리 알고리즘을 도입하고 다양한 환경 조건을 포함한 학습 데이터 확장과 검증 체계 고도화를 통해 시스템의 강건성과 실환경 적응성을 더욱 향상시키고자 한다.

## 5. References

- [1] Lin, Bo-Hong, et al. "360° map establishment and real-time simultaneous localization and mapping based on equirectangular projection for autonomous driving vehicles." *Sensors* 23.12 (2023): 5560.
- [2] Zhang, Hao, et al. "Dino: Detr with improved denoising anchor boxes for end-to-end object detection." *arXiv preprint arXiv:2203.03605* (2022).
- [3] Maggolino, Gerard, et al. "Deep oc-sort: Multi-pedestrian tracking by adaptive re-identification." *2023 IEEE International conference on image processing (ICIP)*. IEEE, 2023.
- [4] Padilla, Rafael, Sergio L. Netto, and Eduardo AB Da Silva. "A survey on performance metrics for object-detection algorithms." *2020 international conference on systems, signals and image processing (IWSSIP)*. IEEE, 2020.
- [5] Bernardin, Keni, and Rainer Stiefelhagen. "Evaluating multiple object tracking performance: the clear mot metrics." *EURASIP Journal on Image and Video Processing* 2008.1 (2008): 246309.