

무작위로 적재된 종이박스 재사용을 위한 YOLO-OBB 기반의 최상단 박스 인식 및 회전각 추출
 이신우, 박재한, 신수용*
 국립금오공과대학교
 nakgongs@kumoh.ac.kr, rimitpark@kumoh.ac.kr, *wdragon@ku.oh.ac.kr

Top Box Recognition and Rotation Angle Extraction Based on YOLO-OBB for Reusing Randomly Stacked Cardboard Boxes

Shin Woo Lee, Jae Han Park, Soo Young Shin*

Kumoh National Institute of Technology.

요약

최근 물류 산업에서 탄소 저감과 자원 순환을 위해 1회용 종이박스를 재사용하는 시스템의 필요성이 대두되고 있다. 본 논문에서는 무작위로 적재된 박스 더미에서 로봇이 박스를 파지하기 위한 OBB(Oriented Bounding Box) 기반의 비전 인식 시스템을 제안한다. 제안된 시스템은 YOLOv11-OBB를 적용하여 기존 AABB 방식의 한계인 회전 오차를 극복하고, 박스의 위치(x,y)와 회전 각도(θ)를 정밀하게 추출한다. 실험 결과, 제안 모델은 최상단 박스를 효과적으로 구분하고 로봇 그리퍼 제어에 필수적인 회전 정보를 제공함을 확인하였다.

I. 서론

전자상거래 확산으로 국내 골판지 상자 시장은 연간 약 53조 원 규모로 성장했으나, 단순 재활용 과정에서 막대한 탄소 배출 문제가 발생하고 있다. 이를 해결하기 위해 자원 효율을 극대화하는 폐쇄형 순환 공급망 구축이 필수적으로 논의되고 있다 [1]. 특히 종이박스의 재사용은 지속 가능한 물류의 핵심이나 [2], 무작위 적재 환경에서 오염되거나 변형된 박스를 탐지하는 데는 기술적 난관이 존재하여 엣지 컴퓨팅 기반의 탐지 연구가 활발하다 [3]. 기존의 비전 기술은 복잡한 환경에서 로봇 파지에 필요한 정밀 좌표와 회전 정보를 제공하는 데 한계가 있어, 비마커 기반의 정밀 인식 기술이 요구된다 [4][5]. 특히 AABB(Axis-Aligned Bounding Box) 방식은 객체를 수평 직사각형으로만 인식하여 회전된 박스 탐지 시 오차를 유발한다. 따라서 본 논문에서는 무작위 적재 환경에서도 박스의 회전 정보를 포함한 정밀 데이터를 추출할 수 있는 YOLOv11-OBB 기반의 비전 시스템을 제안한다.

II. 본론

1. 시스템 구성 및 OBB 알고리즘

본 연구의 시스템은 비전 센서, AI 연산 장치, 로봇 매니퓰레이터로 구성되며, 이는 지능형 자동 분류 시스템의 전형적인 아키텍처를 따른다 [6]. 시스템은 입력 이미지에서 최상단 박스를 선별하고 중심 좌표(x,y)와 회전 각도(θ)를 로봇에 전송한다. 기존 AABB 방식은 회전된 박스를 탐지할 때 불필요한 배경을 포함하는 오차를 발생시킨다 [7]. 이를 해결하기 위해 적용된 OBB는 객체의 회전각을 직접 학습하여 정밀한 경계 영역을 제공한다 [8]. 특히 최신 YOLOv11 구조는 복잡한 객체 인식에서 탁월한 성능을 입증하고 있어 본 시스템의 핵심 알고리즘으로 채택하였다 [9][10][11].

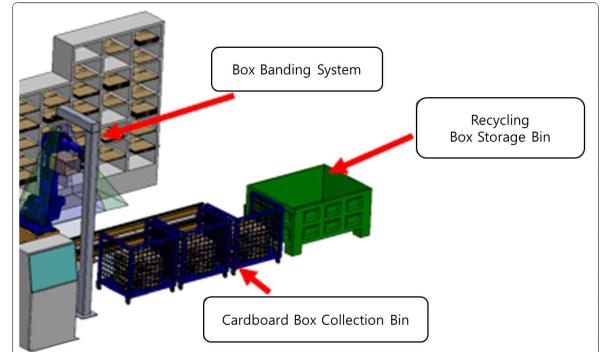


Fig. 1. Overview of the Proposed AI Robot Sorting System
 그림 1. 제안하는 AI 로봇 분류 시스템의 개요도

2. 데이터셋 및 전처리

데이터 수집을 위해 알루미늄 프로파일 카트에 박스를 무작위 적재하고, 정밀 깊이 정보를 제공하는 Intel RealSense D435i를 활용하여 수직 뷰에서 이미지를 취득하였다 [12][13]. 데이터 라벨링은 LabelImg-OBB 툴을 사용하였다 [14]. 이미지의 너비와 높이를 각각 W 와 H 라 할 때, 네 모서리의 좌표 집합 $P = \{(x_i, y_i)\}_{i=1}^4$ 에 대하여 클리핑 및 정규화를 수행하여 정규화된 좌표 \hat{x}_i, \hat{y}_i 를 산출한다, 이는 식 (1)과 같다.

$$\hat{x}_i = \frac{\text{clip}(x_i, 0, W)}{W}, \quad \hat{y}_i = \frac{\text{clip}(y_i, 0, H)}{H} \quad (1)$$

여기서 $\text{clip}(v, \min, \max)$ 함수는 값을 $[\min, \max]$ 범위 내로 제한하는 연산이다. 전처리 과정에서는 네 모서리 좌표를 이미지 경계 내로 클리핑한 후 0.0~1.0 사이로 정규화하여 식 (1)의 결과를 출력하여 모델 입력 효율을 높였다.

3. 실험 환경 및 학습

실험은 실제 물류 현장을 모사한 120cm×120cm 크기의 카트 환경에서 진행되었다. 데이터셋은 Train과 Validation 세트로 8:2 분할하였으며, YOLOv11n-OBB 모델을 사용하여 Epochs 100회로 학습을 수행하였다.

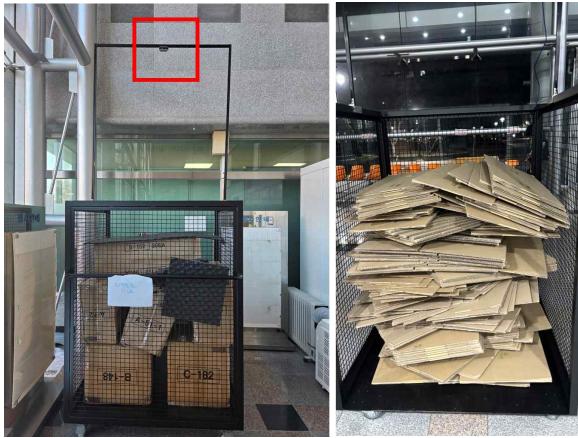


Fig. 2. Experimental Setup with Camera Mount
그림 2. 데이터 수집을 위한 실험 환경 및 카메라 마운트

2. 결과 분석

학습된 모델은 식 (2)과 같이 클래스, 신뢰도, 좌표, 크기, 회전 각도를 포함한 벡터 P_{box} 를 출력한다.

$$P_{box} = (class, conf, cx, cy, w, h, \theta) \quad (2)$$

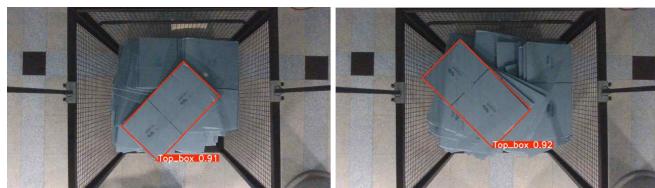


Fig. 3. Inference Results and Data Visualization
그림 3.OBB 모델 추론 결과 및 데이터 시각화

실험 결과, 모델은 무작위로 적재된 박스 더미에서 최상단 박스를 90% 이상의 높은 신뢰도로 탐지하였다. 특히 추출된 θ 값은 박스의 기울기를 정밀하게 반영하여, 로봇 그리퍼의 제 6축을 박스 면에 수직 정렬시키는 핵심 제어 변수로 활용됨을 확인하였다.

III. 결론

본 논문에서는 물류 박스 재사용을 위한 OBB 기반 비전 인식 시스템을 구현하였다. YOLOv11-OBB 모델을 통해 무작위 적재 박스의 2차원 위치(x, y)와 회전 각도(θ)를 성공적으로 추출하였다. 현재 시스템은 2D 이미지 기반으로, 향후 연구에서는 Depth 정보를 결합하여 Z축 높이 차이를 인식하고 그리퍼 제어의 안정성을 높이는 3차원 과정 시스템으로 확장할 계획이다.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (RS-2025-00553810, 50%)

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2025-RS-2024-00437190, 50%) supervised by the IITP(Institute for Information & Communications

Technology Planning & Evaluation)

참 고 문 헌

- [1] Difrancesco, Rita Maria, and Arnd Huchzermeier. "Closed-loop supply chains: a guide to theory and practice." *International Journal of Logistics Research and Applications* 19.5 (2016): 443–464.
- [2] Govindan, Kannan, Hamed Soleimani, and Devika Kannan. "Reverse logistics and closed-loop supply chain: A comprehensive review to explore the future." *European journal of operational research* 240.3 (2015): 603–626.
- [3] Yang, Xi, et al. "Detecting defects with support vector machine in logistics packaging boxes for edge computing." *IEEE Access* 8 (2020): 64002–64010.
- [4] Medjram, Sofiane, Jean-François Brethe, and Khairidine Benali. "Markerless vision-based one cardboard box grasping using dual arm robot." *Multimedia Tools and Applications* 79.31 (2020): 22617–22633.
- [5] Fragapane, Giuseppe, et al. "Planning and control of autonomous mobile robots for intralogistics: Literature review and research agenda." *European journal of operational research* 294.2 (2021): 405–426.
- [6] Vaidya, Saurabh, Prashant Ambad, and Santosh Bhosle. "Industry 4.0 - a glimpse." *Procedia manufacturing* 20 (2018): 233–238.
- [7] Ding, Jian, et al. "Learning RoI transformer for oriented object detection in aerial images." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.
- [8] 김현규, et al. "HBB 와 OBB 기반 모발 객체 검출 라벨링 기법 비교분석." *Journal of the Korea Institute of Information & Communication Engineering* 28.3 (2024).
- [9] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [10] He, Lu-hao, et al. "Research on object detection and recognition in remote sensing images based on YOLOv11." *Scientific Reports* 15.1 (2025): 14032.
- [11] Kotthapalli, Manikanta, Deepika Ravipati, and Reshma Bhatia. "YOLOv1 to YOLOv11: A comprehensive survey of real-time object detection innovations and challenges." *arXiv preprint arXiv:2508.02067* (2025).
- [12] Mirota, D., and J. Scaife. "Intel RealSense Depth Camera D435i IMU Calibration." (2019).
- [13] Servi, Michaela, et al. "Comparative evaluation of Intel RealSense D415, D435i, D455 and Microsoft Azure Kinect DK sensors for 3D vision applications." *IEEE Access* (2024).
- [14] Asghari Ilani, Mohsen, and Yaser Mike Banad. "LabelImg: CNN-Based Surface Defect Detection." *arXiv e-prints* (2025): arXiv-2509.