

파레트가 다단 적재된 실내 환경에서의 지게차 포킹 자동화 알고리즘

이순교¹, 박주형², 우효원³, 박유현⁴, 이창현⁴, 이석⁵, 이기백^{6*}

^{1,5}한국철도기술연구원 미래교통물류연구소

^{2,3,4,6}광운대학교 전기공학과

Forklift forking automation algorithm in an indoor environment where pallets are stacked in multiple stages

Soon-Kyo Lee¹, JooHyeong Park², Hyowon Woo³, YuHyun Park⁴, Chang-Hyun Lee⁴, Suk Lee⁵,
Ki-Baek Lee^{6*}

^{1,5}Korea Railroad Research Institute, Future Transport Logistics Research Center

^{2,3,4,6}Department of Electrical Engineering, Kwangwoon University

As the logistics industry grows, small warehouses face inefficiencies due to manual labor and budget constraints. Forklifts are vital for large tasks, but precise actions like inserting forks into pallet slots still require human input. This thesis presents an algorithm to automate forklift guidance and forking in multi-stage pallet stacking using an RGB-D camera and YOLOv5. Our approach allows real-time pallet detection and precise selection from stacked arrays. By calculating focal length relationships, the system measures the distance for accurate fork insertion and adjusts movement for both floor-level and multi-stage stacking. Tests show improved efficiency, offering a cost-effective automation solution for small warehouses.

Keywords: Forklift Automation, RGB-D Camera, YOLOv5, Multi-Stage Pallet Stacking, Autonomous Forking

논문접수일 : 2024.10.14. 논문수정일 : 2024.11.22. 게재확정일 : 2024.12.07.

본 연구는 국토교통부/국토교통과학기술진흥원의 지원(과제번호RS-2021-KA163201), 그리고 2024년도 광운대학교 우수연구자 지원 사업으로 수행되었음

1. 한국철도기술연구원 첨단물류시스템연구실 선임연구원

2. 광운대학교 전기공학과 학사과정

3. 광운대학교 전기공학과 석사과정

4. 광운대학교 전기공학과 석박사통합과정

5. 한국철도기술연구원 첨단물류시스템연구실 책임연구원

6*. 광운대학교 전기공학과 교수, Corresponding Author: kblee@kw.ac.kr

1. 서론

1.1 연구배경 및 목적

물류 산업의 급성장으로 인해 산업 내 효율성과 비용 절감이 중요한 과제로 떠오르며, 이에 대규모 물류 창고들은 자동화 시스템을 도입하고 있다. 자동화 물류 시스템(Molka-Danielsen et al., 2018)은 기존 인력을 대체하여 물류 작업의 효율성을 높이는 장점이 있으나, 산업 현장에서는 공간과 예산 문제로 인해 이를 적용하는 데 한계가 있다. 산업 현장 중 많은 소규모 물류 창고들이 여전히 수작업에 의존하면서 작업 효율이 낮아지고, 작업시간과 인건비, 운영비가 증가하고 있다. 특히 중소규모 창고나 군부대는 물류 인프라 부족으로 인해 상하차 작업을 수작업으로 처리하게 되어 어려움을 겪고 있으며, 영세 업체들도 비슷한 문제에 직면하고 있다. 적재 공간의 제약을 해결하기 위해 기존 장비에 자동화 기술을 적용하려는 연구가 주목받고 있다. 그중 AGV(Automated Guided Vehicle), 로봇팔, 지게차 등 다양한 형태로 물류 자동화에 대한 연구가 활발히 진행되고 있다. 먼저 AGV는 분산 제어 시스템(De Ryck et al., 2018), 동적 슬롯팅 기법(Tekinerdogan & Catal, 2021), 그리고 자율성 향상 기술(Zhang et al., 2023)을 통해 작업 속도를 높이는 방향으로 발전하고 있다. 로봇팔의 경우에는 협업과 증강 현실 기술(De Pace et al., 2020)을 도입하여 보다 복잡한 작업을 안전하게 수행할 수 있도록 연구되고 있다. 이러한 기술을 통해 작업자와 로봇팔이 같은 공간에서 효율적으로 협업할 수 있게 하며, 작업의 정확성과 안정성을 향상시킨다. 또한, 증강 현실을 활용하여 로봇팔의 동작을 시각화하고 조작을 간편하게 하는 방안도 연구되고 있다.

그중 특히 지게차는 다른 장비들에 비해 비교적 대규모의 수작업을 보완하는 역할을 하고 있다. 적재 화물 운반과 팔레트 홈에 포크를 정확히 삽입하는 작업은 여전히 세밀한 조작이 요구되어 인력의 개입이 불가피하다. 따라서 안정성과 정확성을 확보하기 위해 상당한 인력이 필요하다. 불과 몇 년 전에는 사람이 직접 조작해야만 하는 분야를 AI(Artificial Intelligence)로 대체하기 어려웠지만, 현재의 AI 기술은 많은 발전을 이루고 있으며, 지게차에도 AI를 도입한 자동화에 관한 연구가 진행되고 있다. 머신러닝과 RFID를 기반으로 UHF-RFID(Ultra High Frequency-Radio Frequency Identification) 게이트를 활용하여 지게차의 위치 및 동작을 모니터링하고, 팔레트와 지게차의 진입 및 출입 방향을 인식하는 등 지게차와 팔레트 간의 작업을 자동화(Motroni et al., 2021)하고 있다. 또한 RGB-D 카메라와 딥러닝을 이용하여 DDRNet(Deep Dual-Resolution Networks)을 사용해 2차원 이미지에서 팔레트를 분할하는 방법(Meng et al., 2024)을 제시하고 있다.

이와 같은 기존의 연구들은 복잡한 알고리즘을 사용하거나 다수의 센서를 결합하여 팔레트 인식을 수행해왔다. 다량의 센서를 활용하여 팔레트를 인식하고 자율주행 기술을 구현했으며, 화물 운반 과정에서 지게차의 포크를 팔레트 홈에 자동으로 삽입하는 기술이 개발되었다(Kelly et al., 2007; Shao et al., 2023; Meng et al., 2023). 하지만 기존의 자율주행을 넘어 자율 운반을 구현하기 위해서는 다단으로 적재된 팔레트 중 원하는 특정 팔레트를 확실히 인식하는 과정이 필요하다. 현재 기술로는 2차원 레이저 스캐너와 Faster R-CNN(Ren et al., 2016)을 활용하여 팔레트의 위치를 탐지하고 추적(Mohamed et al., 2020)하며, 이는 여러 팔레트가 존재하는 경우에도 작동한다. 현 시점에서 딥러닝 기법과 소량의 센서를 활용한 팔레트 인식 및 포킹 기술이 가장 효율적인 방안으로 평가된다. 다만, 식별 시 여러 팔레트를 동시에 인식하면서 개별 팔레트를 정확히 특정하지 못해 원하는 팔레트를 선택하는 데 어려움이 있다. 우리는 이러한 문제를 극복하기 위해 실시간 검출이 가능한 딥러닝 기법인 YOLOv5 모델(Zhang et al., 2022)과 RGB-D 카메라를 사용하여 다양한 환경에서도 적은 비용으로 다단 팔레트 인식 문제를 해결하는 방법을 제안한다.

먼저, RGB-D 카메라를 통해 2차원 영상 데이터를 수집한 후, YOLOv5 모델을 사용하여 팔레트를 실시간으로 탐지하고 위치를 추적한다. YOLOv5를 본 연구에서 선택한 이유는 경량화된 구조 덕분에 적은 하드웨어 자원으로도 높은 성능을 발휘하며, 실시간으로 객체를 탐지하는 데 필요한 빠른 처리 속도와 높은

정확도를 동시에 제공하기 때문이다. 팔레트가 인식되면 초점 거리(Focal length)를 이용한 높이 추정 관계식을 적용하여 지게차와 팔레트 간의 거리를 계산하고, 지게차가 포크를 정확히 팔레트 홈에 삽입할 수 있도록 이동 경로를 조정한다. 이를 통해 딥러닝 모델이 안정적인 영상을 바탕으로 물체 인식을 수행할 수 있으며, 포킹 작업의 정확도와 신뢰성을 높인다. 마지막으로, 모델의 성능을 검증하기 위해 다양한 시나리오에서 데이터를 수집한다.

본 연구는 선행연구(Lee et al., 2024)를 확장하여 진행되었다. 선행연구에서는 YOLOv3(Mao et al., 2019) 객체 탐지 모델을 통해 실시간 카메라 피드에서 팔레트를 인식하고, 이를 반시계 방향 회전(CCW), 시계 방향 회전(CW), 정렬(Aligned)의 세 가지 class로 분류하였다. 또한, 연구에서는 팔레트의 위치를 정확하게 감지한 후 유도 알고리즘을 수행하였다. 해당 알고리즘은 팔레트와 올바르게 정렬하기 위해 지게차의 특정 방향 회전이나 이동을 계산하였다. 이 과정은 전제 또는 하역을 위해 팔레트와의 상호작용을 위한 조정으로 이어졌다.

선행연구를 바탕으로 상·하역을 위한 자동화 지게차 구현을 위해 다단 적재된 팔레트를 가정하고 팔레트의 선택, 높이 추정 및 포킹까지의 과정을 수행할 수 있는 시스템을 제안하며 유효성을 평가한다.

1.2 관련 연구

1) YOLO를 이용한 탐지

YOLO(You Only Look Once)는 실시간 객체 탐지를 위한 딥러닝 모델로, 이미지에서 물체를 빠르고 정확하게 탐지하는 데 사용된다(Zhao et al., 2019). 기존의 객체 탐지 방식은 이미지에서 여러 영역을 슬라이딩 윈도우 방식으로 탐색한 후, 각 영역에 대해 객체 여부를 판단하는 복잡한 과정을 거친다. 하지만 YOLO는 이러한 방식과 달리 이미지를 한 번에 전체적으로 분석하여 객체의 위치와 클래스 정보를 동시에 예측하는 단일 네트워크 구조를 채택하고 있다.

YOLO는 이미지 전체를 한 번에 처리하는 단일 신경망을 사용하여, 주어진 이미지의 여러 객체를 동시에 탐지할 수 있다(Terven et al., 2023). 이 과정에서 YOLO는 이미지를 그리드로 나누고, 각 그리드 셀에서 객체의 바운딩 박스와 클래스 확률을 예측한다. 이를 통해 YOLO는 높은 처리 속도와 상대적으로 높은 정확도를 유지하면서도 실시간으로 객체 탐지가 가능하다. 이러한 특성 덕분에 YOLO는 실시간 탐지가 필요한 다양한 응용 분야에서 매우 효과적이며, 특히 지게차의 팔레트 인식과 같은 산업용 애플리케이션에 적합하다.

지게차가 팔레트를 인식하는 데 YOLO를 사용하면, 실시간으로 팔레트의 정확한 위치를 탐지하고, 팔레트에 대한 정보(크기, 모양, 위치)를 실시간으로 파악하는 데 매우 효과적이다. 지게차는 자동화된 창고 환경에서 팔레트를 안전하고 정확하게 처리해야 하며, 팔레트를 감지하는 작업은 매우 중요하다. YOLO의 빠른 처리 속도 덕분에 지게차는 움직이는 동안에도 실시간으로 팔레트를 인식할 수 있어 작업 효율성이 크게 향상된다.

YOLO는 첫 버전인 YOLOv1(Terven et al., 2023)에서 시작해, YOLOv2(Sang et al., 2018)와 YOLOv3에서는 성능을 개선하며 더 빠른 속도와 정확도를 달성한다. YOLOv4(Wu et al., 2020)는 모델 최적화와 데이터 증강 기법을 도입하여 정확도와 속도 측면에서 균형을 이룬다. YOLOv5는 모델 경량화와 최적화된 구조 덕분에 실시간 응용 분야에서 뛰어난 성능을 보인다. YOLOv6(Norkobil Saydirasulovich et al., 2023)는 정확도와 추론 속도를 더욱 높이고, YOLOv7(Gallo et al., 2023)은 다양한 크기의 객체를 탐지하는 성능을 향상시키며 작은 모델에서도 매우 높은 정확도를 제공하며 최근에는 학습과 추론의 효율성을 강화한 YOLOv8(Vijayakumar et al., 2024)이 등장했다.

다만, 본 논문에서는 가장 최신 버전의 YOLOv8이 아닌 YOLOv5를 사용했다. 단순히 객체 탐지에 중점을 둔 연구이기에 YOLOv5 사용해도 될뿐더러 부가적 기능(세그멘테이션, 키포인트 검출 등)을 사용하지

얇아 경량화에 있어 유리하며 지게차에 부담이 되지 않을 소형화된 하드웨어에 YOLOv8과 같은 고성능을 요구하는 모델보단 YOLOv5를 통해 확장성의 이득을 취할 수 있었다.

2) 초점 거리를 이용한 거리 추정

사물의 거리를 추정하는 데 있어 초점 거리는 중요한 역할을 하며, 영상에서 물체의 크기 및 위치를 추정할 때 필수적인 요소로 작용한다. 카메라 렌즈의 초점 거리를 활용해 영상의 픽셀 크기를 바탕으로 사물과 카메라 간의 실제 거리를 계산할 수 있다. 이를 통해 다양한 영상 처리 및 물체 인식 작업에서 거리 정보를 얻을 수 있다.

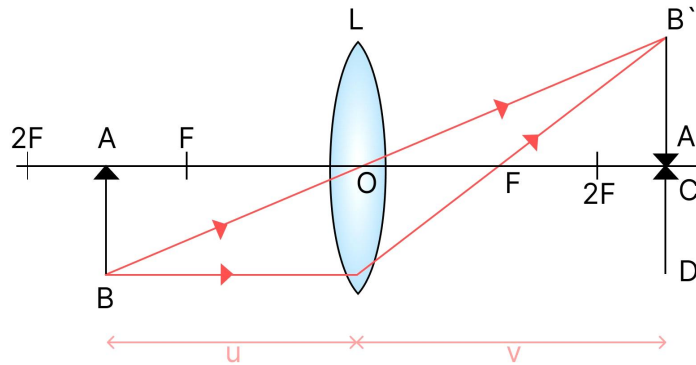


Figure 1. Focal length

$$A : A' = B : B' \quad (1)$$

Figure 1. 은 초점 거리를 이용한 실제 높이 값의 관계를 보여주는 그림이다. 관계식(1)처럼 비례관계를 카메라 렌즈를 기준으로 이미지상의 거리를 이용하여 실제 높이를 추정할 수 있게 된다. A' 와 B' 가 각각 렌즈를 기준으로의 실제거리와 지면으로부터 실제 높이라고 하였을 때 렌즈를 투과하여 물체가 이미지로 보일 때의 세로 픽셀이 A, B 가 Focal length이다. 이 4개의 요소가 비례관계로서 A', A, B 이 3개의 변수를 수집한다면 실제 높이를 추정할 수 있다. 이와 같이 렌즈의 굴곡으로 인한 실제 사물과 렌즈를 투과한 빛에 의해 투영된 사물의 크기 비율은 초점 거리와 렌즈의 세로 길이에 의존되고 있음을 알 수 있다. 초점 거리는 카메라 렌즈의 중심에서 이미지 센서까지의 거리를 의미한다. 초점 거리는 카메라의 시야각과 물체를 확대하거나 축소하는 정도를 결정하는 중요한 파라미터이다. 초점 거리가 길수록 시야각은 좁아지고 물체는 더 크게 보이며, 초점 거리가 짧을수록 시야각이 넓어지고 물체는 작게 보인다.

거리 추정을 위해 얇은 렌즈 방정식(Thin Lens Equation)에 기반한 관계식을 사용한다. 기본적으로, 물체의 실제 크기와 이미지 상의 물체 크기 간의 관계를 사용하여 거리를 추정한다. 이때 거리 추정을 위한 초점 거리의 관계식(2)은 다음과 같다. 이때 f_p 는 픽셀(pixel)단위로 환산한 초점 거리, f_{mm} 는 초점 거리로 카메라의 고유 상수이며, H_p 는 카메라를 통해 출력되는 해상도의 세로 픽셀 개수, H_s 는 카메라 렌즈의 높이 값으로 고유 상수이다.

$$f_p = \frac{f_{mm} \times H_p}{H_s} \quad (2)$$

렌즈의 초점 거리를 이용한 거리 추정은 다양한 응용 분야에서 사용된다. 특히 지게차가 팔레트를 인식하는 경우, 팔레트와 지게차 간의 거리를 정확히 계산하여 팔레트를 안전하고 정확하게 들어 올릴 수 있다

록 돕는다. 또한, 이 기술은 자동 운전, 로봇 내비게이션, AR(Augmented Reality) 등 다양한 분야에서 유용하게 사용된다. 이러한 초점 거리는 카메라를 이용한 거리 추정에서 중요한 역할을 하며, 물체의 픽셀 크기와 실제 크기 간의 관계를 통해 정확한 거리 추정을 가능하게 한다. 이를 통해 지게차와 같은 자동화 기계가 실시간으로 물체의 위치를 인식하고 조정할 수 있다.

2. 연구방법론 및 테스트

우리는 바닥 적재와 다단 적재 모두에 대응 가능한 자동 포킹을 실현하기 위해 지게차의 포크가 정확히 팔레트 홈에 도달할 수 있는 거리를 계산한다. 먼저 RGB-D 카메라와 YOLOv5를 사용하여 2차원 영상에서 여러 팔레트 중 단순히 가장 아래 또는 가장 위 팔레트뿐만이 아닌 팔레트 그 사이에 위치한 팔레트 중 어디에 위치하더라도 원하는 팔레트 하나를 특정한 후, 인식된 팔레트에 초점 거리 관계식을 적용하여 카메라에서 팔레트까지의 거리를 추정한다. 이후, 추정된 거리를 바탕으로 지게차가 얼마나 이동해야 포크가 특정 팔레트 홈에 맞춰질 수 있는지를 계산한다. 결과적으로, 우리의 방식이 적용된 지게차는 2차원 영상 데이터를 통해 팔레트의 3차원적 위치를 추정하고, 팔레트의 적재 상태에 상관없이 포크 삽입 작업을 자동으로 수행할 수 있다.

기존 방식의 경우 UHF-RFID 태그와 안테나를 사용하여 부착된 태그를 인식하고, 신호의 위상을 분석하고 판별하는 방식(Motroni et al., 2021)으로 자동화를 구현하였지만 신호의 영향을 줄 수 있는 산업 환경 요인으로 인한 통신 장애, 데이터를 수집하기 위한 장비나 기술적 비용, 특정 환경에 맞춘 시스템에 의한 환경적 제약 등으로 실제 산업 환경에서 사용하기 어려운 부분이 있다.

우리는 대형 물류 창고에서의 다양한 적재 상황을 가정하여, 바닥에 놓인 팔레트뿐만 아니라 여러 층에 적재된 팔레트까지도 자동 포킹이 가능하도록 하는 방식을 제안한다. 이를 위해, 지게차의 포크가 정확히 팔레트 홈에 도달할 수 있도록 거리를 계산한다. 먼저, RGB-D 카메라와 YOLOv5를 사용하여 2차원 영상에서 여러 팔레트를 실시간으로 인식하고, 팔레트의 높이와 위치까지 정확하게 추정한다. 이때, 단순히 가장 아래나 위에 있는 팔레트뿐만 아니라, 중간 레벨에 위치한 팔레트도 정확하게 선택할 수 있도록 시스템을 설계한다. 선택된 팔레트에 대해 초점 거리 관계식을 적용하여 카메라에서 팔레트까지의 거리를 추정하고, 이 거리를 바탕으로 지게차가 포크를 정확히 특정 팔레트 홈에 삽입할 수 있도록 이동 경로를 계산한다. 결과적으로, 제안된 방식은 지게차가 2차원 영상 데이터를 활용하여 팔레트의 3차원적 위치를 정확하게 추정하고, 팔레트의 적재 상태와 상관없이 포크 삽입 작업을 자동으로 수행할 수 있게 한다. 이로써 다단 팔레트의 선택적 포킹이 가능해지며, 물류 작업의 효율성을 크게 향상시킬 수 있다.

2.1 팔레트 높이 추정

1) YOLOv5 데이터 세트 준비 및 학습

YOLOv5는 이미지에서 객체를 빠르고 정확하게 탐지하는 실시간 객체 탐지 모델로, 여러 단계의 CNN(합성곱 신경망)을 사용해 물체의 특징을 추출하고 예측하는 구조로 되어 있다. YOLOv5는 YOLOv3에서 발전된 형태로, CSPDarknet(Li et al., 2024)과 PANet(Xu et al., 2024)을 백본과 넥에 적용하여 피쳐 맵을 더욱 효율적으로 결합하고 탐지 성능을 향상시켰다. YOLOv5는 모델 경량화와 효율성 개선에 집중하여, 더 작은 크기의 모델에서도 높은 성능을 유지하면서 빠른 탐지가 가능하게 해준다. 이러한 특징 덕분에 YOLOv5는 YOLOv3에 비해 더 빠른 속도와 더 높은 정확도를 제공한다.

이 모델을 사용하여 팔레트를 식별하고 위치를 예측하는 작업을 진행하며, 데이터셋에는 단일 적재된 팔레트 학습 이미지 1800장과 테스트 이미지 300장, 다단 적재된 팔레트 학습 이미지 2100장과 테스트 이미지 530장이 포함되어 있다. 이 데이터셋은 팔레트가 다양한 각도와 조건에서 촬영된 이미지를 포함하고 있으며, 이를 통해 YOLOv5 모델이 다양한 환경에서도 정확하게 팔레트를 인식할 수 있도록 훈련했다. 특히, 이 모델은 pre-trained된 YOLOv5s 모델을 기반으로 하여, 팔레트 이미지를 커스텀 데이터셋으로 학습시켰

다.

모델 학습에 사용되는 데이터셋에는 각 이미지에 대해 레이블 정보가 포함되어 있다. 각 레이블은 객체의 클래스와 위치 정보를 담고 있으며, 이를 바탕으로 모델은 객체를 정확히 탐지하고 분류했다. 팔레트의 상태를 인식하기 위해 3가지 클래스를 설정했다: 반시계 방향 회전, 시계 방향 회전, 정렬. 이 클래스는 팔레트가 어느 방향으로 회전된 상태인지를 나타내며, 지게차가 팔레트를 포킹하기 전에 정확히 정렬될 수 있도록 돕기 위한 중요한 정보를 제공했다.

YOLOv5 모델은 각 이미지를 입력으로 받아, 그 이미지 내에서 탐지된 객체(팔레트)의 클래스와 위치를 출력한다. 출력 값은 객체의 바운딩 박스 좌표(좌상단과 우하단 좌표)와 해당 객체의 클래스 정보이다. 이를 통해 팔레트의 정확한 위치와 상태를 예측할 수 있으며, 다양한 환경에서 안정적으로 팔레트를 인식하는 데 중요한 역할을 한다.

2) 높이 추정

RGB-D 카메라를 사용하여 실시간으로 영상 데이터를 수집하고, 해당 영상에 YOLOv5를 적용해 팔레트를 탐지한다. 팔레트가 인식된 후, RGB-D 카메라로 수집된 2차원 영상의 픽셀 값을 기반으로 초점 거리 관계식을 이용하여 팔레트까지의 거리를 추정한다.

① YOLOv5를 이용한 모든 팔레트 탐지 : YOLOv5 모델을 사용하여 입력된 이미지에서 모든 팔레트를 탐지한다. 모델은 이미지를 분석하여 팔레트의 경계 상자를 반환하며, 각 팔레트의 위치 정보를 획득한다. 이를 통해 탐지된 팔레트들의 좌표와 크기가 계산된다.

② 팔레트 선택: 다단으로 적재된 팔레트 중에서 특정 팔레트를 선택하는 과정이다. 사용자는 탐지된 팔레트들 중에서 포킹을 원하는 팔레트를 선택할 수 있다. 선택된 팔레트는 화면에서 강조 표시되며, 이후 팔레트의 높이와 위치를 계산하여 지게차가 정확한 포킹 작업을 수행하도록 지원한다.

③ 경계 상자와 픽셀 크기 계산 : 탐지된 팔레트의 경계 상자 좌표를 바탕으로 팔레트의 중심 좌표와 픽셀 높이를 계산한다. 이를 통해 팔레트의 크기를 추정하고, 높이 정보를 추출할 수 있는 기초 데이터를 제공한다.

④ 깊이(Depth) 값 추출 : 깊이 이미지를 사용하여 탐지된 팔레트의 중심부에 해당하는 깊이 값을 추출한다. 이 값은 팔레트가 실제로 얼마나 떨어져 있는지를 측정하는 데 사용되며, 이후 실제 높이를 계산하는 데 중요한 역할을 한다.

⑤ 깊이 보정 값 : 깊이 센서의 오차를 줄이기 위해 보정 값을 적용한다. 특히 팔레트가 가까운 거리에 있을 때 발생하는 깊이 데이터의 왜곡을 줄이기 위해 보정 값을 사용(Yang et al., 2020)해야 하며, 실험을 진행하며 보정 값을 사용하여 더 정확한 높이 추정이 가능해진다. 보정상수를 K 라 할 때 높이를 구할 때 사용하는 값으로 실제 거리와 측정된 거리의 오차를 보정하는 역할로 1.6으로 세팅하였다.

⑥ 실제 높이 추정 : 추출된 깊이 값과 계산된 픽셀 크기를 바탕으로 팔레트의 실제 높이를 추정한다. 이 과정은 지게차가 팔레트를 정확하게 들어 올릴 수 있도록 팔레트의 위치와 높이를 파악하는 데 사용된다. 함수식(3)은 초점거리 계산식을 토대로 본 연구에 적용한 것으로 h_{real} 은 지상으로부터 팔레트 중심까지의 실제 높이, y_z 는 이미지 상의 지상과 팔레트 경계면의 y 좌표, y_t 는 적재된 팔레트 중에서 목표가 되는 팔레트의 중심 y 좌표, D_m 은 목표 팔레트의 깊이 측정값이다. 이는 Figure 2. 와 같이 비례식을 이용한 높이 추정을 본 연구에 맞게 변형한 것이다.

$$h_{real} = \frac{(y_z - y_t) \times K \times D_m}{f_p} \quad (3)$$

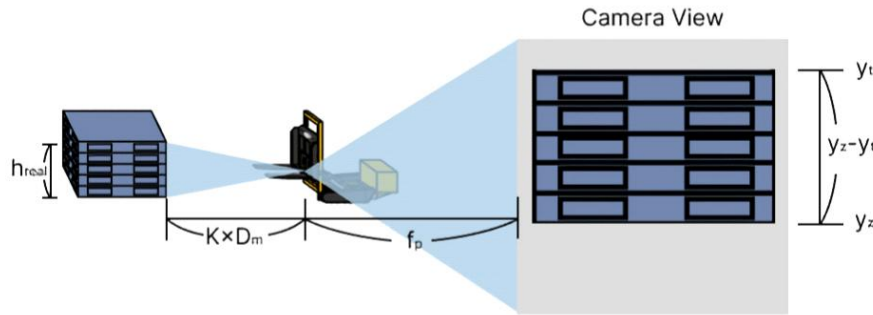


Figure 2. Height estimation using proportional formula

3) 전체 과정

팔레트 높이 추정 및 지게차의 정확한 포킹 작업을 위한 전체적인 과정은 영상 데이터 수집, 특정 팔레트 인식 및 선택, 팔레트 높이 추정의 순서로 진행된다.

먼저, RGB-D 카메라를 사용해 실시간으로 팔레트가 적재된 환경의 영상 데이터를 수집한다. 수집된 영상은 YOLOv5 모델에 입력되어 팔레트를 탐지하고, 모델은 다단 적재된 팔레트들에 대해 경계 상자 (Bounding Box)를 생성한다. 이때 각 팔레트의 중심 좌표와 크기를 계산하여 탐지된 여러 팔레트의 정보를 기록한다.

그다음, 사용자는 특정 팔레트를 선택할 수 있다. 이 코드를 통해 다단 적재된 팔레트 중에서 원하는 순번의 팔레트를 정확하게 선택한다. 선택된 팔레트는 YOLOv5 모델이 설정한 경계 상자를 통해 팔레트의 중심 좌표 및 크기를 계산한 후, 특정 팔레트를 대상으로 이후 작업이 진행된다.

팔레트가 선택되면, 초점 거리 관계식을 적용해 선택된 팔레트의 높이와 위치를 추정하는 과정이 시작된다. RGB-D 카메라로 수집된 2차원 영상에서 픽셀 크기를 계산하고, f_{mm} (1.88mm)을 이미지 해상도와 렌즈 높이 값을 이용해 f_p 로 환산 후 K 와 함께 높이 추정 함수식에 적용하여 d_m 을 보정하여 계산한다. 이러한 보정 과정을 통해 가까운 거리에서 발생할 수 있는 오차를 줄이며, 선택된 팔레트의 정확한 높이와 위치를 추정할 수 있게 된다. Figure 3. 은 이러한 과정의 흐름을 다이어그램으로 나타낸 것이다.

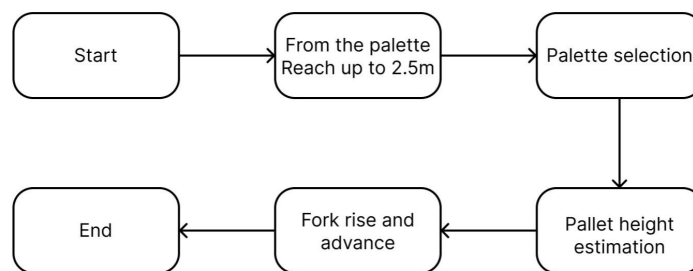


Figure 3 Pallet height estimation algorithm implementation flowchart

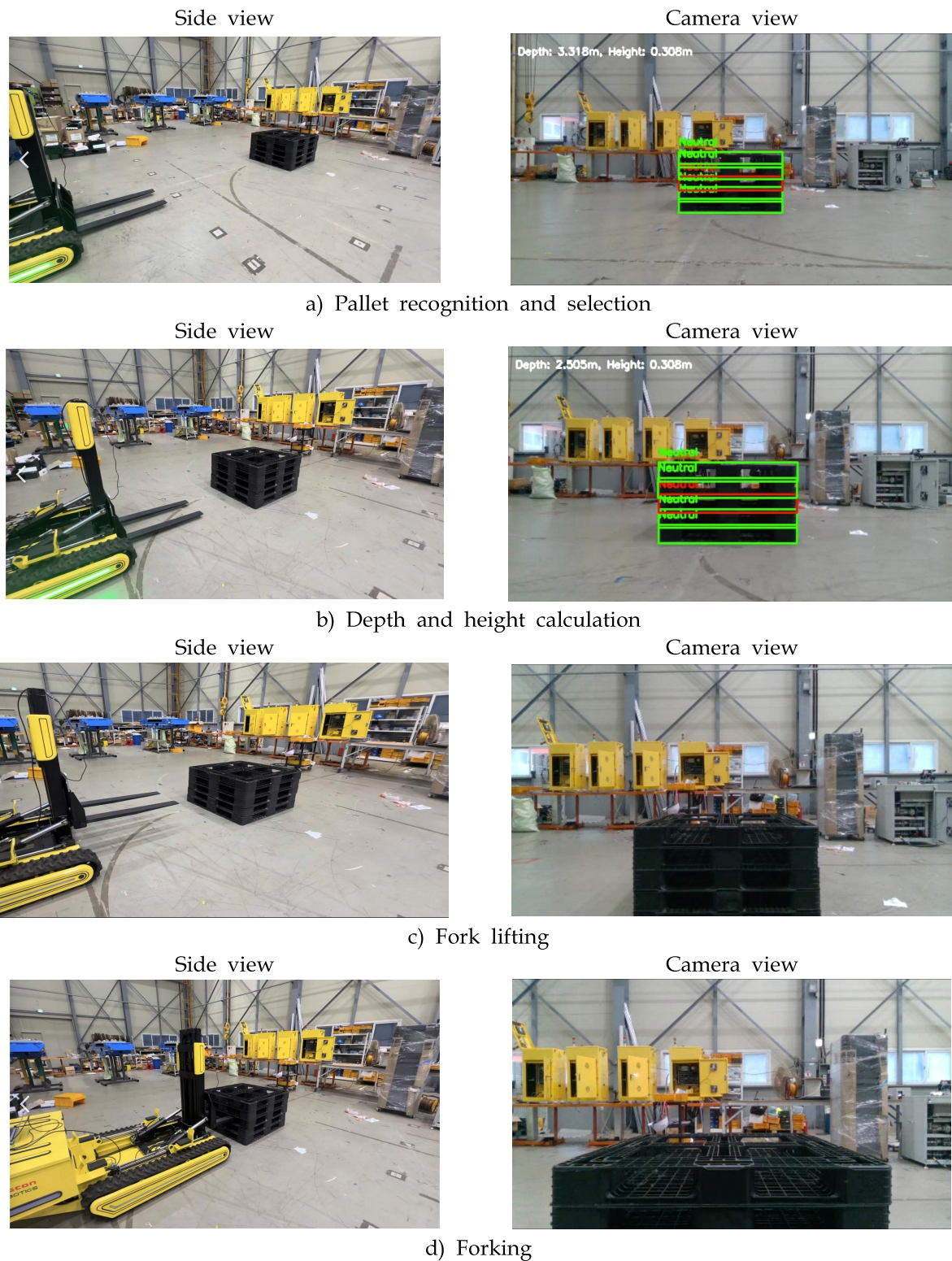


Figure 4. Forklift operation & Visualization of deep learning model output pallet height

4) 실제 방법론에 따른 구동 테스트

Figure 4. 는 적재된 팔레트를 인식 후 특정 팔레트의 위치를 파악하여 포킹하는 과정까지의 현장 시험 이미지와 카메라 영상 화면이다. 현장 테스트 이미지와 카메라 비디오 화면은 적재된 팔레트 더미에서 특

정 팔레트를 인식, 선택 및 포크하는 전체 프로세스를 보여준다. 그림은 실제 장면을 나타내는 측면 뷰와 팔레트에 대한 필 수 높이 및 깊이 정보를 제공하는 딥러닝 모델 출력을 시각화하는 카메라 뷰 등 다양한 뷰를 보여주고 있다. a의 경우 모든 팔레트를 인식하며 그 중 목표로 하는 층의 팔레트만 빨간색 bounding box로 표시하게 된다. b의 경우 팔레트 인식 완료 후 지게차가 팔레트에 접근하게 된다. 충분히 접근 후 목표로 하는 팔레트의 높이를 계산한다. c의 경우 계산된 높이만큼 포크를 상승 또는 하강 시킨다. d의 경우 포크만 앞으로 전진하여 팔레트 홈의 포크를 정확히 끼워 넣는다.

2.2 시나리오 테스트 결과

1) 팔레트 인식 테스트

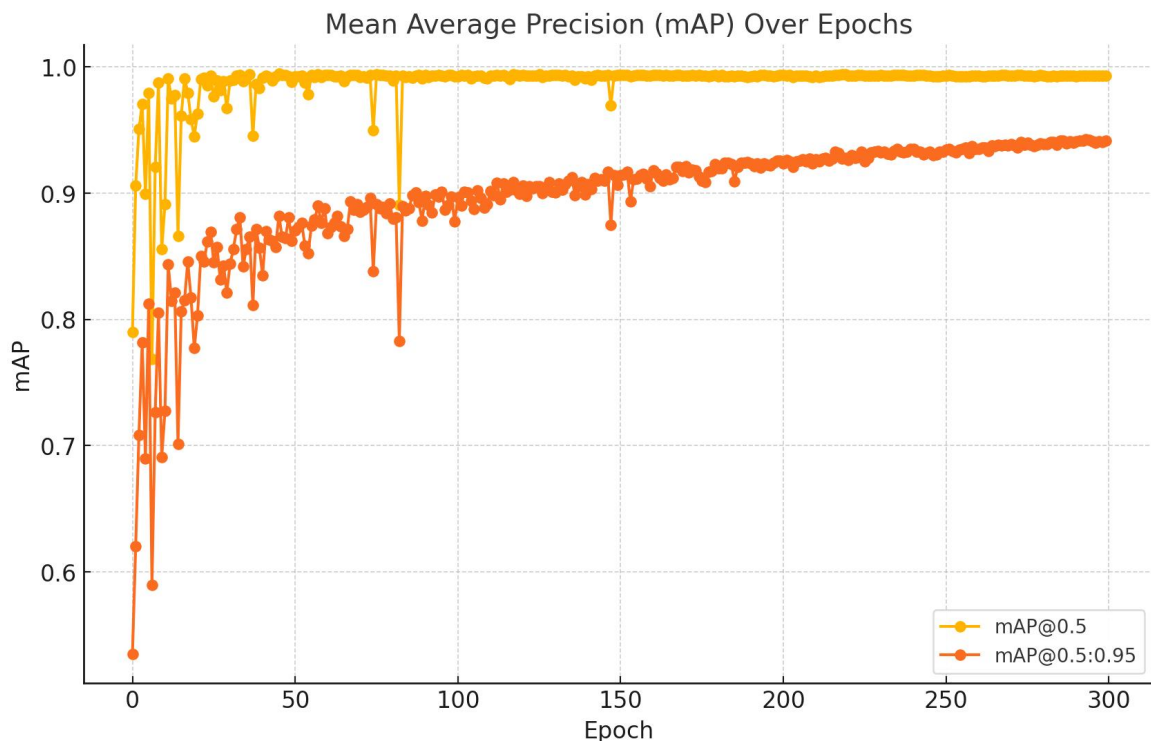


Figure 5. Training Performance of the Pallet Detection Model: mAP Progression Over Epochs

Figure 5. 는 팔레트 인식 모델의 학습 과정을 보여주며, mAP(Mean Average Precision) 값이 에포크(epoch)마다 개선되고 높은 수준에 도달했음을 나타낸다. mAP@0.5(노란색)는 임계값 0.5에서 모델의 정확도를 나타내며, 학습 초기에 변동이 있었지만 빠르게 수렴하여 약 99% 이상의 높은 정확도를 기록한다. 이는 팔레트 탐지와 위치 인식에서 매우 높은 신뢰성을 제공한다. mAP@0.5:0.95(주황색)는 다양한 IoU(Intersection over Union) 임계값에서 평균적으로 계산된 값으로, 더 까다로운 조건을 반영하지만 95%에 근접하는 안정적인 성능을 보여준다.

팔레트 인식의 오류는 학습 초기 단계에서 변동이 있었으나 학습이 진행되면서 감소했고, 대부분의 경우 1% 미만의 낮은 오류율을 기록했다. 이러한 오류는 물류 환경에서 무시할 수 있는 수준으로, 탐지 작업의 정확도와 효율성을 크게 저해하지 않는다. mAP@0.5:0.95가 mAP@0.5보다 낮은 이유는 더 엄격한 기준을 적용했기 때문이며, 이는 모델의 성능이 뛰어나다는 점을 강조한다.

물류 공간에서는 이러한 오류가 작업 효율성에 미치는 영향이 미미하다. 높은 정확도를 유지하며, 환경 조건이 일정하고 단순한 물류 작업의 특성상 탐지 오류가 발생해도 반복적인 탐지나 보정 작업을 통해 쉽

게 해결할 수 있다.

결론적으로, 팔레트 인식 모델은 높은 수준의 성능과 안정성을 달성했으며, 물류 환경에서의 오류는 대체로 무시할 수 있는 수준이다. 팔레트 탐지 작업에서 충분히 신뢰할 수 있는 성능을 제공한다.

2) 거리 추정 실험 및 결과

팔레트의 높이를 추정하기 전 팔레트 높이 함수식(3)에 대입할 변수 중 깊이를 YOLOv5로 감지된 팔레트를 RGB-D 카메라로 깊이를 추출하여 대입하게 된다. 이때 팔레트의 높이 추정에 대한 오차를 줄이기 위해서 깊이의 정확도가 곧 높이 추정식의 정확도가 된다. 우리가 제안하는 높이 추정식은 깊이의 의존도가 높은 계산식이기에 깊이의 오차가 적어야 한다. 이러한 오차를 줄이기 위해선 본 실험에서 사용하는 RGB-D 카메라인 Intel RealSense D415 카메라의 데이터시트에 표기에 따라 2m이내에선 $\pm 2\%$ 이하의 오차가 발생하는 점, 그 이후로는 오차가 급격히 증가한다는 점을 토대로 깊이를 추출할 수 있는 최소 범위부터 3m까지 카메라를 이동시켜 실제거리와 깊이 비교 실험을 진행한다.

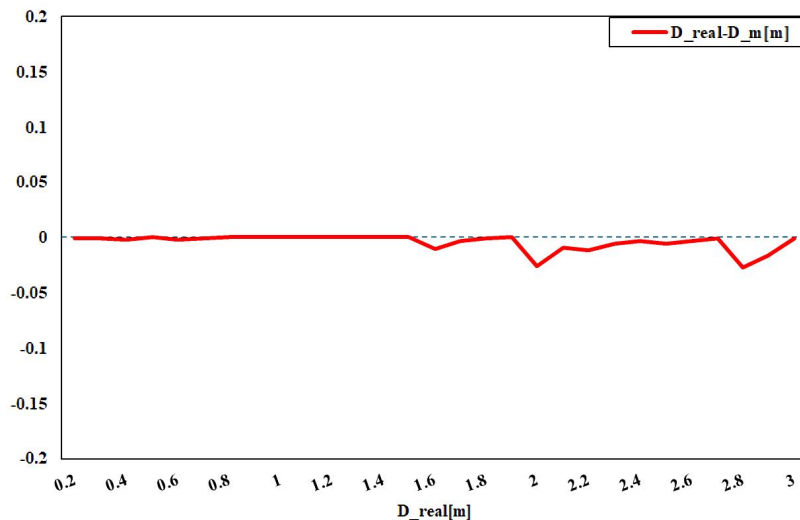


Figure 6. Depth estimation experiment results

Figure 6. 은 실제 거리 D_{real} 과 D_m 의 관계 및 차이를 나타낸 것이다. 깊이 값의 경우 D_{real} 이 증가하더라도 D_m 과의 오차가 대체적으로 변화가 크지 않음을 보인다. 이것으로 3m 이내에는 깊이 값의 큰 오차가 발생하지 않아 충분히 높이 값을 신뢰할 수 있다. 다만 팔레트의 홈에 정확히 지게차의 포크가 안착되어야 하기 때문에 포킹 알고리즘이 수행하기 안정적인 거리를 찾아야 한다. 0.8m에서 2m 사이의 거리에서는 다른 구간에 비해 상대적으로 안정적인 성능을 보였다. 그러나 2m를 넘는 거리에서는 약간의 흔들림이 발생하였다.

안정적인 성능을 위해서는 2m 이내에서 알고리즘을 실행하는 것이 이상적이다. 그러나 작업 범위에 영향을 미치기 때문에 최적의 위치를 결정할 때 포크의 길이를 고려해야하기 때문에 1.5m, 2m, 2.5m, 3m 구간에서 높이 추정 실험을 진행한다.

3) 높이 추정 실험 및 결과

거리 추정을 통해 충분히 팔레트에 접근한 후 포킹 알고리즘을 진행해야 하기 때문에 1.5m, 2m, 2.5m, 3m 각각에 접근했다고 가정하고 높이 추정 알고리즘을 실행하기에 적절한 위치를 찾는다.

Figure 7. 은 1.5m, 2m, 2.5m, 3m 각각의 거리에서 팔레트를 1단계, 2단계, 3단계, 4단계, 5단계에 두고

진행한 높이 추정 실험 결과를 보여주고 있다. a)에서 깊이 오차는 2단계에서 정점에 달하지만 그 이상 레벨에서 감소하고 높이 오차는 미미하며 5단계에서 약간만 증가했다. b)에서는 깊이 오차가 변동하여 2단계 레벨에서 정점을 찍고 이후 감소하며 높이 오차는 작은 변동을 보이지만 전반적으로 낮은 수준을 유지했다. c)에서는 깊이 오차가 커서 2단계에서 정점을 찍고 더 높은 레벨에서 낮아지는 반면, 높이 오차는 1단계와 2단계에서 가장 높고 상위 레벨로 갈수록 감소하는 반면, 높이 오차는 5단계에서 약간 정점을 찍으며 미미한 수준을 유지한다.

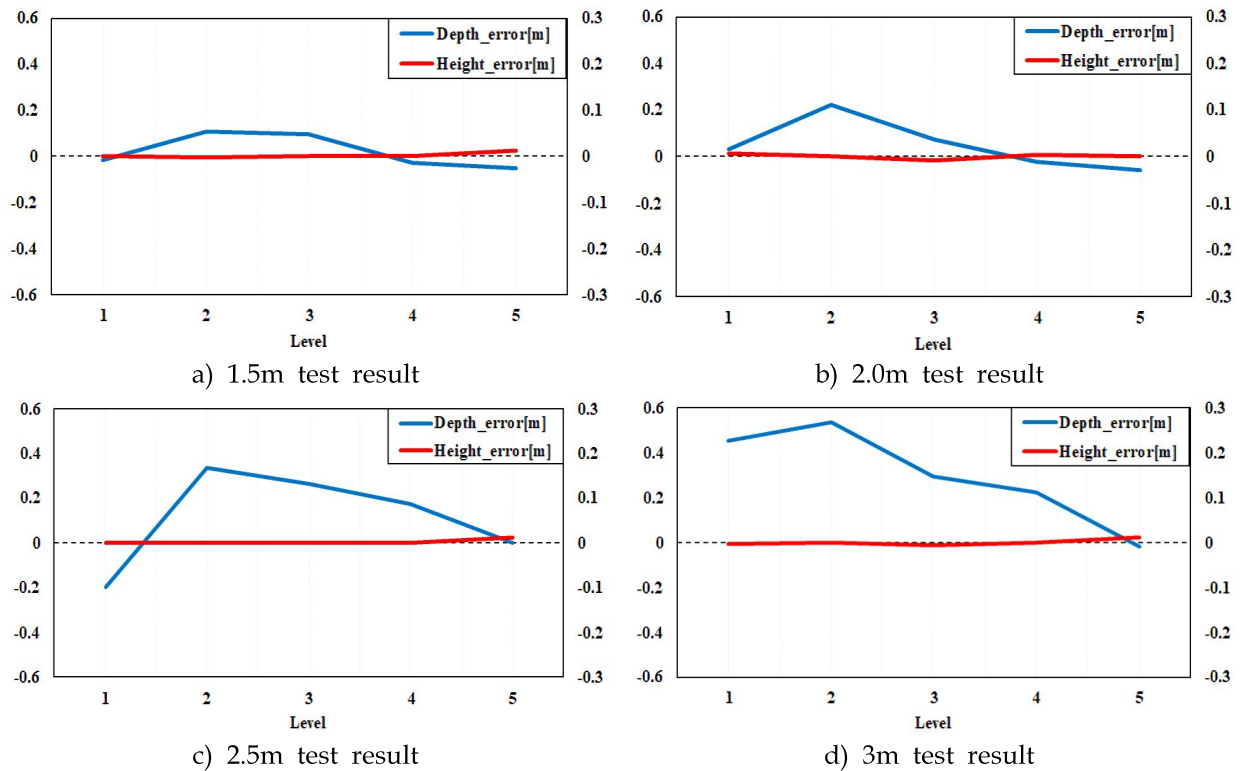


Figure 7. Error in height according to depth by distance

거리가 멀어질수록 깊이 오차는 점진적으로 증가하는 경향을 보였으며, 특히 대부분 중간단계에서 정점을 찍은 후 점점 줄어드는 경향이 보인다. 반면 높이 오차는 모든 거리에서 상대적으로 미미한 수준을 유지했으며, 2m를 제외한 나머지는 5단계에서 변동이 있었다.

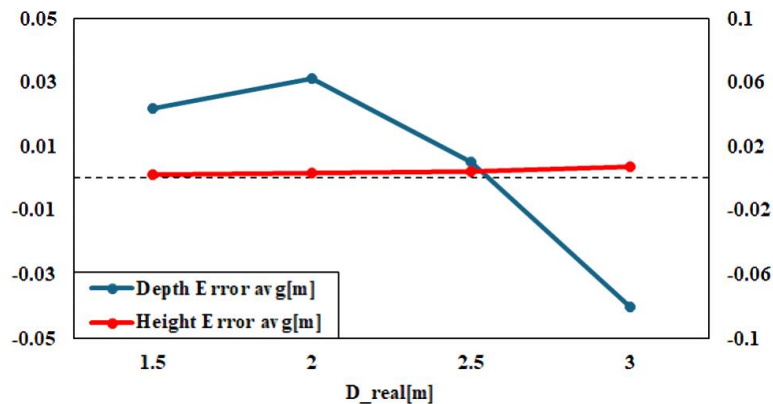


Figure 8. Changes in overall depth and height according to distance

Figure 8. 은 높이 추정 실험을 토대로 실제 거리마다 깊이와 높이의 총 오차 평균을 내어 거리별 알고리즘의 안정성을 확인할 수 있다. 1.5m일 때는 포크의 길이(1.061m)로 인해 포킹을 수행하기에 안전하지 않아 적합하지 않다. 2m에서의 높이 오차는 상대적으로 작은 반면, 깊이 오차는 불균형적으로 커서 시스템에 비현실적인이다. 3m에서는 깊이 오류가 모두 상당히 높다. 하지만 2.5m는 깊이와 높이 오류 모두 균형이 잡혀 있고 팔레트의 홈의 너비(260m×90m)를 고려하여 수용 가능한 수준이므로 포킹 알고리즘을 효율적으로 실행하는 데 가장 적합한 거리이다.

3. 결론

본 연구에서는 RGB-D 카메라와 YOLOv5를 활용하여 다단으로 적재된 팔레트를 실시간으로 인식하고, 이를 기반으로 팔레트의 위치와 높이를 정확하게 추정하여 지게차의 자동 포킹 작업을 가능하게 하는 기술을 제안하였다. 기존 연구들이 복잡한 알고리즘과 다수의 센서를 결합하여 자율 주행과 물체 인식을 해결한 반면, 본 연구는 상대적으로 적은 비용과 간단한 장비를 사용하여 팔레트 인식 및 포킹 문제를 해결하였다. 이는 특히 소규모 물류 창고나 군부대와 같은 제한된 예산과 공간을 가진 환경에서도 적용 가능하다는 점에서 실용적 가치를 지닌다.

연구의 핵심은 다음과 같은 과정을 통해 이루어졌다. RGB-D 카메라를 이용하여 2차원 영상을 수집한 후, YOLOv5를 통해 실시간으로 팔레트를 탐지하고, 각 팔레트의 위치와 높이를 정확하게 계산하였다. 팔레트가 인식된 후 초점 거리 관계식을 활용해 팔레트의 거리를 추정하고, 이를 기반으로 지게차가 포크를 정확히 팔레트 홈에 삽입할 수 있는 경로를 산출하였다. 또한, 실험 과정에서 발견된 깊이 인식 오차를 해결하기 위해 보정 값을 추가하여 가까운 거리와 먼 거리에서 모두 안정적인 높이 추정이 가능하도록 하였다.

본 연구의 결과로, 기존 방식보다 더 간단하고 효율적인 시스템을 통해 다단으로 적재된 팔레트 중 특정 팔레트를 선택하고, 그 팔레트의 높이와 위치를 정확하게 추정하여 지게차의 자동 포킹 작업을 성공적으로 수행할 수 있었다. 이는 물류 자동화의 효율성을 크게 향상시킬 수 있으며, 소규모 창고에서도 저비용의 자동화 시스템을 적용할 수 있는 가능성을 제시한다.

향후 연구에서는 더욱 복잡한 물류 환경에서 다양한 장애물과 조건에 대응할 수 있는 모델의 확장성을 검토하고, 실시간 처리 속도를 더욱 개선하여 대규모 물류 환경에도 적용 가능한 범용적 시스템으로 발전시키는 데 주력할 것이다.

본 연구의 결과로, 기존 방식보다 더 간단하고 효율적인 시스템을 통해 다단으로 적재된 팔레트 중 특정 팔레트를 선택하고, 그 팔레트의 높이와 위치를 정확하게 추정하여 지게차의 자동 포킹 작업을 성공적으로 수행할 수 있었다. 이는 물류 자동화의 효율성을 크게 향상시킬 수 있으며, 소규모 창고에서도 저비용의 자동화 시스템을 적용할 수 있는 가능성을 제시한다.

향후 연구에서는 더욱 복잡한 물류 환경에서 다양한 장애물과 조건에 대응할 수 있는 모델의 확장성을 검토하고, 실시간 처리 속도를 더욱 개선하여 대규모 물류 환경에도 적용 가능한 범용적 시스템으로 발전시키는 데 주력할 것이다.

참고문헌

- De Pace, F., Manuri, F., Sanna, A., & Fornaro, C. (2020). A systematic review of Augmented Reality interfaces for collaborative industrial robots. *Computers & Industrial Engineering*, 149, 106806.
- De Ryck, M., Versteyhe, M., & Debrouwere, F. (2020). Automated guided vehicle systems, state-of-the-art control algorithms and techniques. *Journal of Manufacturing Systems*, 54, 152-173.
- Gallo, I., Rehman, A. U., Dehkordi, R. H., Landro, N., La Grassa, R., & Boschetti, M. (2023). Deep object detection of crop weeds: Performance of YOLOv7 on a real case dataset from UAV images. *Remote Sensing*, 15(2), 539.
- Kelly, A., Nagy, B., Stager, D., & Unnikrishnan, R. (2007). Field and service applications-an infrastructure-free automated guided vehicle based on computer vision-an effort to make an industrial robot vehicle that can operate without supporting infrastructure. *IEEE Robotics & Automation Magazine*, 14(3), 24-34.
- Lee, S. K., Kim, S., Woo, H., Lee, S., & Lee, K. B. (2024). Design of Vehicle-mounted Loading and Unloading Equipment and Autonomous Control Method using Deep Learning Object Detection. *The Journal of Korea Robotics Society*, 19(1), 79-91.
- Li, F., Sun, T., Dong, P., Wang, Q., Li, Y., & Sun, C. (2024). MSF-CSPNet: A Specially Designed Backbone Network for Faster R-CNN. *IEEE Access*.
- Mao, Q. C., Sun, H. M., Liu, Y. B., & Jia, R. S. (2019). Mini-YOLOv3: real-time object detector for embedded applications. *Ieee Access*, 7, 133529-133538.
- Meng, D., Ren, Y., Yu, X., Yin, X., Wang, W., & Men, J. (2024). High-Precision Positioning and Rotation Angle Estimation for a Target Pallet Based on BeiDou Navigation Satellite System and Vision. *Sensors*, 24(16), 5330.
- Mohamed, I. S., Capitanelli, A., Mastrogiovanni, F., Rovetta, S., & Zaccaria, R. (2020). Detection, localisation and tracking of pallets using machine learning techniques and 2D range data. *Neural Computing and Applications*, 32, 8811-8828.
- Molka-Danielsen, J., Engelseth, P., & Wang, H. (2018). Large scale integration of wireless sensor network technologies for air quality monitoring at a logistics shipping base. *Journal of industrial information integration*, 10, 20-28.
- Motroni, A., Buffi, A., Nepa, P., Pesì, M., & Congi, A. (2021). An action classification method for forklift monitoring in industry 4.0 scenarios. *Sensors*, 21(15), 5183.
- Norkobil Saydirasulovich, S., Abdusalomov, A., Jamil, M. K., Nasimov, R., Kozhamzharova, D., & Cho, Y. I. (2023). A YOLOv6-based improved fire detection approach for smart city environments. *Sensors*, 23(6), 3161.
- Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6), 1137-1149.
- Sang, J., Wu, Z., Guo, P., Hu, H., Xiang, H., Zhang, Q., & Cai, B. (2018). An improved YOLOv2 for

- vehicle detection. *Sensors*, 18(12), 4272.
- Shao, Y., Fan, Z., Zhu, B., Lu, J., & Lang, Y. (2023). A Point Cloud Data-Driven Pallet Pose Estimation Method Using an Active Binocular Vision Sensor. *Sensors*, 23(3), 1217.
- Tekinerdogan, B., & Catal, C. (2021). Design of a reference architecture for developing smart warehouses in industry 4.0. *Computers in Industry*.
- Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*, 5(4), 1680-1716.
- Vijayakumar, A., & Vairavasundaram, S. (2024). Yolo-based object detection models: A review and its applications. *Multimedia Tools and Applications*, 1-40.
- Wu, D., Lv, S., Jiang, M., & Song, H. (2020). Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Computers and Electronics in Agriculture*, 178, 105742.
- Xu, L., Cui, Q., Xu, W., Chen, E., Tong, H., & Tang, Y. (2024). Walk in views: Multi-view path aggregation graph network for 3d shape analysis. *Information Fusion*, 103, 102131.
- Yang, L., Dryanovski, I., Valenti, R. G., Wolberg, G., & Xiao, J. (2020). RGB-D camera calibration and trajectory estimation for indoor mapping. *Autonomous Robots*, 44, 1485-1503.
- Zhang, J., Yang, X., Wang, W., Guan, J., Ding, L., & Lee, V. C. (2023). Automated guided vehicles and autonomous mobile robots for recognition and tracking in civil engineering. *Automation in Construction*, 146, 104699.
- Zhang, Y., Guo, Z., Wu, J., Tian, Y., Tang, H., & Guo, X. (2022). Real-time vehicle detection based on improved yolo v5. *Sustainability*, 14(19), 12274.
- Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11), 3212-3232.

요약문

물류 산업이 성장함에 따라 소형 창고는 수작업과 예산 제약으로 인해 비효율성에 직면하고 있다. 지게차는 대형 작업에서 필수적이지만, 팔레트 슬롯에 포크를 정확히 삽입하는 등의 정밀한 작업에는 여전히 사람의 개입이 필요하다. 이 논문은 RGB-D 카메라와 YOLOv5 모델을 사용하여 다단계 팔레트 적재에서 지게차 안내와 포크 작업을 자동화하는 알고리즘을 제안한다. 이 방법은 실시간으로 팔레트를 감지하고, 쌓여 있는 팔레트 배열 중에서 정확하게 선택할 수 있게 한다. 시스템은 초점 거리 관계를 계산하여 지게차와 팔레트 사이의 거리를 측정하고, 포크 삽입을 위한 정확한 위치를 계산한 후 지게차의 움직임을 바닥 수준과 다단계 적재 작업에 맞게 조정한다. 테스트 결과, 시스템은 효율성을 크게 향상시켰으며, 공간이 제한된 소형 창고에 적합한 비용 효율적인 자동화 솔루션을 제공한다.

주제어: 지게차 자동화, RGB-D 카메라, YOLOv5, 다단 팔레트 적재, 자율 포킹