

실시간 인스턴스 세그멘테이션 객체 탐지 및 적응형 배치 알고리즘을 활용한 저비용 빈피킹 시스템

Real-time Instance Segmentation-based Object Detection and Adaptive Placing Algorithm for Low Cost Bin-picking System

김기석¹, 신현표^{1,*}Ki-Suk Kim¹ and Hyun-Pyo Shin^{1,*}¹ 동양미래대학교 로봇자동화공학부 (School of Robot and Automation Engineering, Dongyang Mirae University)

Corresponding Author / E-mail: hpshin@dongyang.ac.kr, TEL: +82-2-2610-1816

ORCID: 0000-0002-1779-0577

KEYWORDS: Real-time object detection (실시간 객체 탐지), Instance segmentation (인스턴스 세그멘테이션), Adaptive placing algorithm (적응형 플레이싱 알고리즘), Bin-picking process (빈피킹 공정), Inverse kinematics based on optimization (최적화 기반 역기구학), Empty hole detection (빈 홀 탐지)

Robots are increasingly utilized in manufacturing and logistics, where bin-picking has become crucial for managing randomly placed objects. However, traditional methods often rely on expensive 3D vision systems, have limited adaptability to unstructured environments, and primarily focus on the picking process, neglecting the placing tasks. To address these challenges, this study presents a cost-effective system that combines a depth camera, YOLO-based instance segmentation, and optimization-based inverse kinematics for real-time object detection and stable manipulation. In the placing stage, an adaptive algorithm detects empty tray holes and generates grid patterns, ensuring reliable placement even in the presence of tray misalignments, occupied slots, or partial occlusions. Experimental validation revealed a 91% success rate in mixed-object environments during picking tasks and a 94% success rate for placing tasks, even with tray displacement and occlusion conditions. The results demonstrate that the system maintains stable performance across both picking and placing processes while minimizing reliance on expensive hardware and complex initial setups. By enhancing flexibility and scalability, the proposed approach offers a practical solution for intelligent automation and can serve as a foundation for broader applications in assembly, logistics, and service robotics.

Manuscript received: September 22, 2025 / Revised: November 25, 2025 / Accepted: December 17, 2025

1. 서론

로봇을 활용한 제조업과 물류 분야에서의 자동화가 빠르게 진행되고 있다[1-3]. 특히 다양한 형상과 자세의 부품을 자동으로 인식하고 이송, 조립하는 로봇 비전 기술의 중요성이 높아지고 있으며, 이는 생산성과 품질을 높이는데 핵심적인 역할을 한다[4-6]. 최근에는 실제 산업 현장에서 무작위로 놓인 부품을 집어 내는 빈피킹(Bin-picking) 기술이 자동차 조립, 전자부품 포장, 물류 분류 등 다양한 분야에 도입되고 있다[7-9].

그러나 지금까지 상용화된 빈피킹 시스템(Pickit3D, Ensenso

N, Photoneo PhoXi 등)은 비교적 고가의 3D 비전 시스템을 적용하여 대당 적게는 수백 만원에서 수천 만원에 이르는 구축 비용을 필요로 하여 중소기업에서는 기술 도입이 쉽지 않으며, 물체의 형상이나 자세 인식이 어려운 경우에 대해 사용자가 유연성 있게 대응하기 어려운 한계도 있다[10-12].

또한 기존의 빈피킹 시스템에 대한 연구는 물체를 집어 올리는 피킹(Picking) 단계에만 중점을 두고 있는 경우가 대부분이며 피킹한 물체를 다양한 환경에서 정확하고 안정적으로 위치 또는 배치시키는 플레이싱(Placing) 단계에 대한 연구는 부족하다[13]. 피킹이 이루어진 후에는 물체를 일관성 있게 정렬하여 위치시키기 위한 트레이로의 이동이 필요하다. 그러나, 빈피킹

공정을 수행하는 중에는 플레이싱 단계에서 예상치 못한 상황이 발생할 수 있다. 예를 들어, 물체를 위치시키기 위한 트레이의 자세가 사용자의 부주의나 실수로 틀어짐이 발생할 수 있고, 트레이 상에 물체를 위치시킬 공간이 이미 다른 물체에 의해 채워져 있을 수 있는데 이러한 상황에 실시간으로 적절히 대응하지 않으면 충돌과 파손이 일어날 수 있다. 그러므로 원활한 전 공정의 자동화 및 효율화를 위해서는 피킹 단계뿐 아니라, 플레이싱 단계의 자동화 또한 필요하다.

최근에는 비교적 저렴한 카메라와 인공지능 기반의 객체 인식 기술을 결합한 연구들이 활발히 진행되고 있다[14]. 특히 물체의 윤곽 정보를 추출해내는 인스턴스 세그멘테이션 기술은 물체 인식의 정확도를 높이고 다양한 형상의 물체를 구분하는 데에 효과적인 방법으로 주목받고 있다. 그러나, 특이점과 관절의 운동 범위 한계를 고려한 실시간 6-자유도 포즈 데이터(Pose Data) 계산 및 플레이싱에 대한 기술적인 과제가 남아있다. 6-자유도 직렬 로봇은 동일한 말단 자세에 대해 어깨-팔꿈치-손목의 방향 조합에 따라 복수의 해(Left/Right, Up/Down, Flip/No-Flip)가 존재하므로, 이들 중 실제 작업 환경에서 관절 한계·충돌 회피·자세 안정성을 만족하는 해를 일반화하여 선택하기 어렵다. 또한 특이점 근처에서는 수렴이 불안정하여 빈피킹 작업에서 안정성 저하로 연결될 수 있다. 따라서 관절 한계 및 특이점을 고려하여 신뢰성 있는 해를 선택할 수 있는 최적화 기반 역기구학의 적용이 필요하다.

본 연구에서는 전용 3D 비전 시스템을 대체할 수 있는 백만원 이하의 저비용 깊이 인식 카메라(Depth Camera)와 범용 YOLO 인공지능 모델을 활용한 피킹 및 플레이싱 공정 자동화 시스템을 제안한다. 깊이 인식 카메라와 인공지능 객체 인식을 통해 물체의 위치와 방향을 실시간으로 파악하고, 대칭 물체의 방향을 잘못 인식하는 문제를 슬라이싱 기반의 분석 기법으로 해결한다. 또한, 물체를 놓는 트레이의 자세가 달라지더라도 자동으로 이를 감지하고 보정할 수 있도록 적응형 플레이싱 알고리즘(Adaptive Placing Algorithm)을 결합하여 플레이싱을 안정적으로 수행한다. 이러한 방식은 고가의 장비나 복잡한 초기 설정 없이도 다양한 환경에 빠르게 적용할 수 있으며, 설치 및 유지 비용을 낮추고 작업 환경의 변화에도 유연하게 대응할 수 있다는 장점이 있다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 제안하는 공정에 사용된 로봇, 깊이 인식 카메라 등 하드웨어 시스템 구성에 대해 기술하였고, 3장은 실시간 인스턴스 세그멘테이션 객체 탐지 알고리즘을, 4장은 적응형 배치 알고리즘을 기술하였다. 5장에서는 협동로봇에 적용한 실험을 통해 제안하는 알고리즘을 검증하였다. 마지막으로 결론을 제시하였다.

2. 공정 시나리오 및 하드웨어

2.1 공정 시나리오

본 연구의 빈피킹 공정은 깊이 인식 카메라와 YOLO 인공지능 모델(YOLOv8m-seg)을 통해 물체를 인식하여 물체의 중심점을

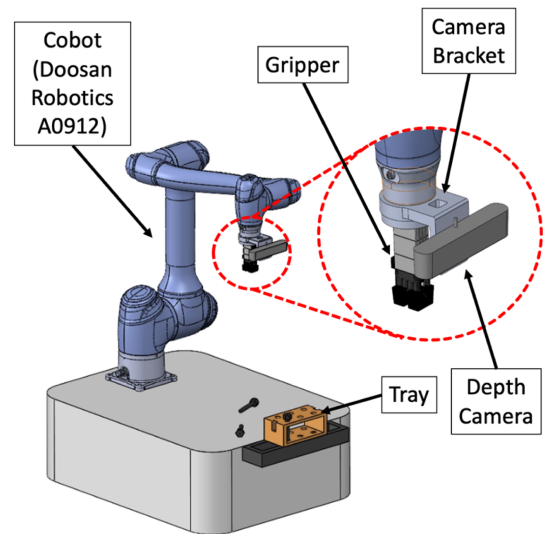


Fig. 1 Hardware system for the bin-picking and adaptive placing algorithm

도출하면 그 중심점의 3차원 좌표를 바탕으로 관절 값(Joint Angle)을 계산한 후 말단 장치가 물체를 잡으면서 이루어진다. 물체를 잡은 후에는 적응형 플레이싱 알고리즘을 통해 트레이의 자세를 고려한 빈 홀을 찾는 과정을 거치게 된다. 빈 홀이 존재하지 않는 경우 트레이를 교체하도록 신호를 보내고, 존재하는 경우 빈 홀에 물체를 배치하면서 공정이 진행되도록 한다.

2.2 하드웨어 구성

본 연구를 위한 주요한 하드웨어의 구성은 Fig. 1과 같다. 전체 시스템은 PC와 제어기, 협동로봇, 협동로봇 베이스, 깊이 인식 카메라(ZED2, Stereolabs), 카메라 브라켓, 그리퍼 그리고 트레이로 구성된다. 사용된 협동로봇은 두산로보틱스 A0912 모델이다. 말단부에는 그리퍼가 장착되어 물체를 파지할 수 있도록 하였다. 깊이 인식 카메라는 3D 프린터로 제작된 카메라 브라켓을 통해 로봇의 플랜지 하부에 고정되어 로봇의 위치 변화에 따른 실시간 시각 정보를 획득한다. 트레이는 실제 산업 환경의 부품 적재 공간을 모사한 요소로 일정한 간격의 홀을 포함한다.

3. 실시간 인스턴스 세그멘테이션 객체 탐지 빈피킹 시스템

3.1 빈피킹 시나리오

빈피킹 공정 순서도를 Fig. 2에 나타내었다. 먼저 YOLO 인공지능 모델을 이용하여 객체를 탐지한 후 물체의 형상을 나타내는 2D 마스크를 추출하여 3D 포인트 클라우드(Point Cloud)로 변환하고 이를 바탕으로 물체의 중심점을 구한다. 이후 PCA (Principal Component Analysis)를 활용하여 구한 축을 기준으로 물체의 이미지를 슬라이싱하여 가장 넓은 단면을 찾아 헤드 방향을 판정하고, 카메라를 기준으로 한 좌표계를 로봇 좌표계로

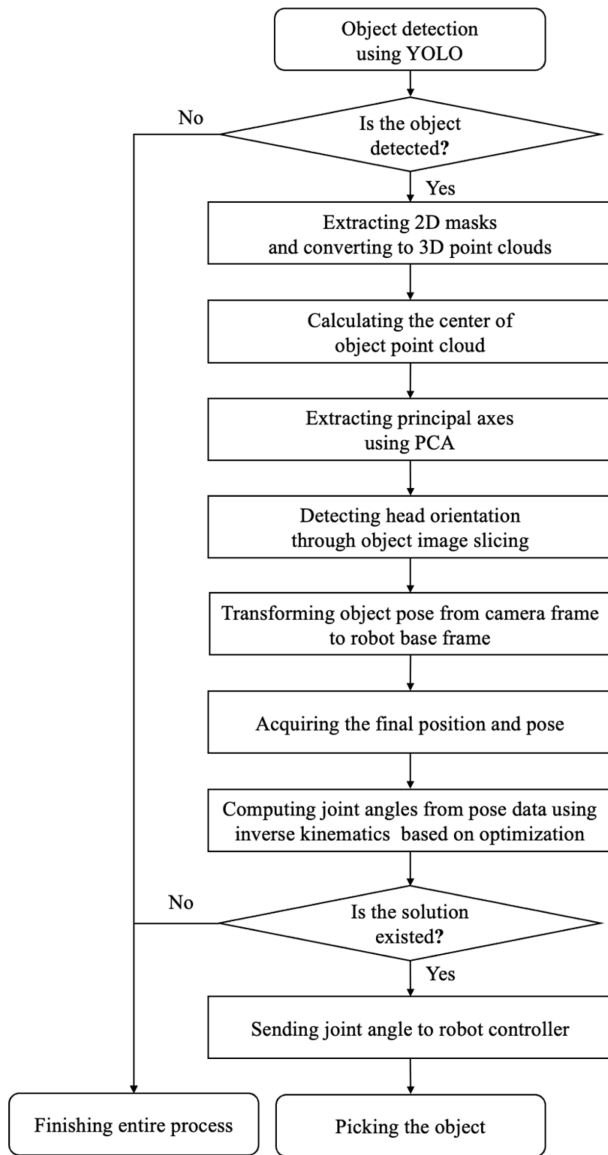


Fig. 2 Bin-picking process flow chart (before adaptive placing process)

변환하여 최종 위치와 자세를 도출한다. 로봇의 빈피킹 동작을 보다 원활하게 수행하기 위해 최적화 기반의 역기구학 계산을 적용한다. 로봇의 동작이 실제 작업 환경에서 안정적으로 수행되기 위해서는 다양한 제약 조건을 고려하는 것이 필요하다. 관절 가동 범위를 벗어나지 않도록 제한하고 로봇 팔과 주변 환경 간의 충돌을 방지하기 위한 회피 조건을 반영하는 것, 그리고 불연속적인 동작을 최소화하여 전체 경로를 부드럽게 움직이는 것이 필요하다. 이를 위해 최적화 기법을 활용하여 역기구학 문제를 해석하고, 각 관절의 값을 산출한다.

3.2 인스턴스 세그멘테이션 기반 객체 탐지

다양한 형상과 방향으로 놓인 물체들을 정확하게 피킹하기 위해서는 물체를 실시간으로 탐지하는 것이 필수적이다. 이를 위해 YOLO 기반 인공지능 모델을 활용하여 인스턴스 세그멘테이션

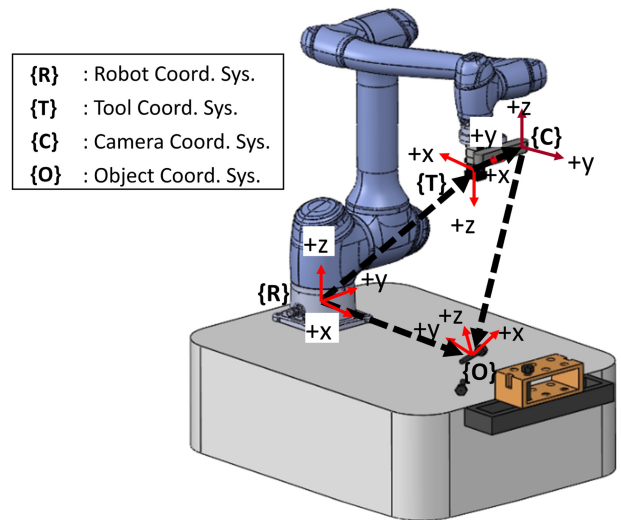


Fig. 3 Coordinate system designation

방식의 실시간 객체 검출 알고리즘을 구현한다. 이 알고리즘에서는 실시간으로 영상을 취득하여 물체에 대한 픽셀 마스크를 추출하고, 이에 대응하는 깊이 정보를 바탕으로 마스크 내부 픽셀들의 3차원 좌표를 수집하여 포인트 클라우드를 구성한다.

수집된 포인트 클라우드에서 무한대 값이나 'NaN' 값과 같은 이상 데이터를 제거한 뒤, PCA를 통해 물체의 방향성을 나타내는 주성분 축을 계산한다. PCA는 분산이 가장 큰 방향을 제1 주성분으로 정의하여 물체의 길이 축을 자동으로 추출하며, 나머지 축은 직교 관계를 통해 구성된다.

이렇게 구성된 길이 축을 기준으로 물체의 형상을 축 후보 방향으로 얇게 슬라이싱한 뒤, 각 슬라이스의 가로 폭을 비교하여 가장 넓은 지점을 찾고, 그 위치를 물체의 헤드 방향으로 판단한다. 이 과정은 볼트처럼 상대적으로 축 방향 길이가 긴 대칭적인 물체에서 자주 발생하는 방향 뒤바뀔 문제를 방지하는 데 효과적이다.

남은 두 축은 길이 축과의 외적을 통해 서로 직교하도록 만들고, 전체 축을 정리하여 오른손 좌표계로 구성한다. 이때, 카메라 좌표계의 Z-축이 반대로 설정되는 경우가 있기 때문에, 이를 점검하고 필요시 다시 보정하여 정확한 좌표계를 유지하도록 한다. Fig. 3에서 나타난 것처럼 카메라 좌표계에서 얻은 물체의 포즈 정보를 로봇의 베이스 좌표계로 변환하기 위해, 현재 관절각을 통해 계산한 정기구학 (베이스-툴 좌표 변환)과 툴-카메라 좌표 변환을 결합하여 로봇-카메라 좌표 변환을 구성한다. 카메라-툴 회전 변환을 통해 물체 포즈를 베이스 좌표계로 변환한다. 최종 회전은 로봇 제어기와 호환을 위해 Z-Y-Z 오일러 각으로 표현하고, 실제 피킹 동작 시 충돌 여유와 툴 오프셋(Tool Offset)을 반영한 좌표 값들이 다음 절의 역기구학에 전달된다.

3.3 최적화 기반 역기구학

빈피킹 공정에서는 작업을 수행하는 동안 비전 시스템이 매 순간 목표 위치와 자세를 실시간으로 제공하므로 로봇의 관절

Table 1 D-H Parameters of the Doosan Robotics A0912

D-H	Theta(δ) [Degree]	a [mm]	d [mm]	Alpha(α) [Degree]
Joint 1	θ_1	0	195	0
Link 1	$\theta_2 - 90$	0	39	-90
Joint 2	$\theta_3 + 90$	560	-	0
Link 2	θ_4	0	516	90
Joint 3	θ_5	0	-	-90
End	θ_6	0	124	90

한계나 작업 제약 조건을 고려하여 신속하게 적절한 피킹 자세를 도출할 수 있어야 한다.

전통적인 닫힌 해(Closed-form) 기반 역기구학은 주로 구형 손목 구조나 단순화된 로봇 모델을 전제로 한다. 이로 인해 툴 오프셋이 발생하거나 로봇의 구조가 조금만 달라져도 해가 존재하지 않거나 계산이 불가능한 경우가 생겨나기도 한다. 6-자유도 직렬 로봇은 동일한 말단 자세에 대해서도 어깨-팔꿈치-손목의 방향 조합에 따라 복수의 해가 도출된다. 그러나 이 중에서 어떤 해가 실제 작업 상황에 더 적합지를 일반화하여 선택하기가 어렵다. 더 나아가 관절 제한, 충돌 회피 여유, 자세 안정성과 같은 조건을 수식에 직접 포함하는 것도 매우 제한적이다. 자코비안(Jacobian) 기반의 수치적 역기구학 역시 초기 추정값에 크게 의존하며, 특이점 근처에서는 수렴이 불안정하다. 또한 최소 이동량, 중심 관절값 기준 구동 등 추가 목적을 동시에 반영하기 어렵다는 한계가 있다. 이러한 제약을 극복하기 위해 본 연구에서는 다양한 목적과 제약 조건을 통합적으로 고려할 수 있는 최적화 기반 역기구학을 적용한다.

로봇의 정기구학은 Table 1의 D-H 파라미터를 바탕으로 구성되며, 관절 각도 벡터 $q = [q_1, \dots, q_6]^T$ 에 대해 말단 포즈 행렬은 식(1)과 같이 나타낸다.

$$T(q) = \begin{bmatrix} R(q)p(q) \\ 0 & 1 \end{bmatrix} \quad (1)$$

$R(q)$ 는 말단의 방향을 나타내는 3×3 회전 행렬이며, $p(q)$ 는 3×1 말단 위치 벡터이다. 목표 포즈 T^* 는 깊이 인식 카메라로부터 얻은 목표 위치 $p^* = [x, y, z]^T$ 와 목표 방향 R^* 을 나타내는 Z-Y-Z 오일러 각 (ϕ, θ, ψ)를 사용하여 식(2)와 같이 구성된다.

$$R^* = R_z(\phi)R_y(\theta)R_z(\psi), \quad T^* = \begin{bmatrix} R^* p^* \\ 0 & 1 \end{bmatrix} \quad (2)$$

이때 목표와의 차이를 나타내는 포즈 오차는 두 가지로 정의된다. 첫째, 위치 오차는 말단 위치와 목표 위치 사이의 유클리디안 거리로 $e_R(q) = \|p(q) - p^*\|_2$ 로 계산된다. 둘째, 회전 오차는 회전 행렬의 대각합을 이용한 거리 척도로 식(3)과 같이 나타낼 수 있다.

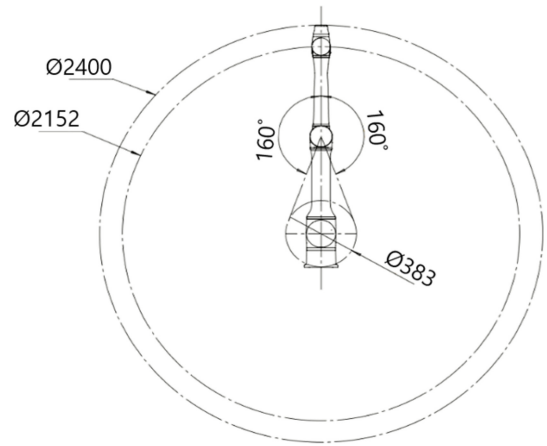


Fig. 4 Work range of the Doosan Robotics A0912

$$e_R(q) = \arccos\left(\frac{\text{tr}(R(q)R^{*T}) - 1}{2}\right) \in [0, \pi] \quad (3)$$

이 식을 통해 두 회전 행렬 $R(q)$ 와 R^{*T} 의 상대적인 차이를 계산하여 회전 각도(θ)를 추출한다. 구체적으로, $R(q)R^{*T}$ 는 목표 방향과 현재 방향 사이의 상대 회전을 나타내는 행렬이며, 이 행렬의 대각합을 이용하면 두 방향 사이의 회전 각도를 계산할 수 있다. 결과적으로 $e_R(q)$ 는 로봇 말단의 방향이 목표 방향으로 얼마나 회전되어 있는지를 나타내는 척도이며, 수치적으로 안전한 계산을 위해 $\arccos$ 의 입력은 $[-1, 1]$ 로 제한한다. 서로 다른 단위에 대해 다음의 목적 함수를 적용했다.

$$\min_q J(q) = \frac{e_p(q)}{2100} + \frac{e_R(q)}{\pi} \quad (4)$$

subject to $q^{\min} \leq q \leq q^{\max}$

여기에서 $J(q)$ 는 최소화하고자 하는 목적 함수이며, $e_p(q)$ 는 위치 오차를 의미하고 분모 2100은 Fig. 4에 나타난 작업 공간의 크기, π 는 최대 회전 각도로, 위치와 회전 오차 간의 단위 차이를 보정한다. 또한 $q^{\min}$, $q^{\max}$ 는 각 관절의 물리적 회전 한계를 나타낸다. 이 최적화 과정에서 제약 조건을 만족하는 순차 이차 계획법(Sequential Least Squares Programming, SLSQP)을 이용해 결과를 도출한다. SLSQP는 목적함수가 비선형이며 경계조건만 존재하는 문제를 효율적으로 해결할 수 있는 알고리즘으로, 로봇 관절 제한을 직접 포함할 수 있다. SLSQP 최적화의 반복 조건은 실시간 제어 요구사항과 수렴 안정성 사이의 균형을 고려하여 설정하였다. 우선 최대 반복 횟수는 500으로 설정하였으며, 이는 6-자유도 로봇의 비선형 역기구학 문제에서 수십 밀리초 수준의 실시간 계산 시간을 유지하면서도 충분한 오차 감소를 확보할 수 있는 상한 값이다. 또한 수렴 허용 오차는 1×10^{-4} 로 설정하였다. 이 값은 최적화 과정에서 위치 변화가 약 0.1 mm 이하, 회전 변화가 약 0.1° 이하로 안정되는 시점에서 알고리즘이 종료되도록 설정한 것이다. 6-자유도 로봇의 해 공간을 모두

포함하는 총 8가지의 어깨-팔꿈치-손목 방향 조합과 함께 무작위 초기 자세를 생성하여 각기 다른 시작점에서 시작하여 최적화를 반복 수행하고 그 중 비용 $J(q)$ 가 가장 작은 해를 선택한다. 이후 선택된 최적 관절 값 q^* 에 대해 정기구학식을 다시 적용하여 $T(q^*)$ 와 T^* 사이의 오차를 수치적으로 검증함으로써 정확성을 확보하는 과정을 거친다.

이와 같은 최적화 기반 접근은 목표 자세가 실시간으로 변하고 작업 제약을 복합적으로 적용해야 하는 빈피킹 환경에서 전통적인 닫힌 해 기반 역기구학 방식보다 더 높은 안정성과 유연성, 그리고 유지보수의 용이성을 제공한다.

4. 적응형 플레이싱 알고리즘

기존의 빈피킹 관련 연구들은 물체를 위치시킬 트레이의 위치 및 자세가 사전에 정의된 곳에 정확하게 정렬되어 있다고 가정하고 피킹 단계에 집중하는 경우가 많았다. 본 연구에서 제안하는 적응형 플레이싱 알고리즘은 물체를 잡은 이후 위치시키는 과정에서 사용자의 실수나 예상치 못한 환경 변화에 대응할 수 있도록 고안되었다. 사전 정의된 위치 및 자세대로 놓이지 못한 트레이의 위치 및 자세 오차, 물체를 규칙적으로 배치하기 위한 홀을 지닌 트레이의 경우 이미 일부 홀에 무작위로 물체가 놓여져 있어서 남은 홀을 찾아서 물체를 플레이싱 해야 하는 상황, 그리고 트레이가 작업 중 다른 장비나 물체로 인해 일부 가려지는 상황에도 대응할 수 있도록 함으로써 빈피킹 이후 공정까지의 자동화를 구현할 수 있다.

본 연구에서는 사전에 정의된 패턴이나 고정된 배열 정보에 의존하지 않고, 실시간으로 검출된 홀을 기반으로 트레이의 패턴을 유추하는 새로운 적응형 플레이싱 알고리즘을 고안하였다. 먼저 HSV (Hue, Saturation, Value) 임계화를 통해 트레이 영상에서 빈 홀 후보를 추출하고, 원형성 지표(Circularity)와 반경 기반 이상치 제거를 수행하여 신뢰성 있는 홀 집합을 구성한다. 이후 중앙값과 평균 절대 편차로 Y축 방향 이상치를 배제한 뒤, K-Means와 실루엣 점수(Silhouette Score)를 활용해 홀들을 행·열 패턴으로 군집화하여 실제 트레이의 격자를 복원한다. 이 과정을 통해 트레이의 위치·자세 오차, 부분 가림, 일부 홀 점유 상황에서도 배치 가능한 빈 홀을 안정적으로 도출한다.

4.1 적응형 플레이싱 시나리오

제안하는 방식은 빈 홀의 검출과 패턴 추출 방식을 통해 위에서 언급한 다양한 상황에서도 더 정확하고 안정적인 플레이싱 작업을 수행할 수 있게 한다. 본 절에서는 제안된 절차를 ROI (Region of Interest) 추출 및 빈 홀 탐색, 그리고 패턴 추출의 두 단계로 나누어 설명한다.

적응형 플레이싱 알고리즘에 대한 순서도를 Fig. 5에 나타내었다. 로봇 제어기로부터 트레이 감지 신호를 받게 되면 카메라를 이용하여 트레이 이미지의 유무를 파악하고 트레이 이미지에서

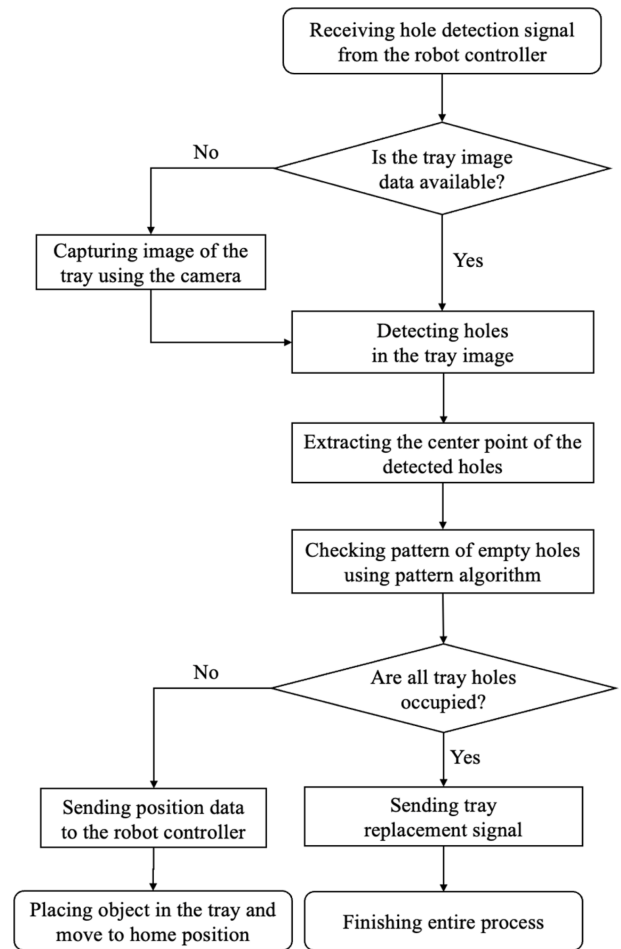


Fig. 5 Adaptive placing process flow chart

빈 홀을 찾는다. 이때 빈 홀이 감지된다면 해당하는 중심 좌표를 추출하며, 추출된 중심 좌표들을 로봇 좌표계로 변환하는 동시에 중심 좌표들을 유무를 0과 1로 구분하여 행렬을 구성한다. 1로 구분된 중심 좌표들은 기구학 계산을 진행한 이후 로봇 제어기로 전송하여 적응형 플레이싱 과정을 진행할 수 있도록 한다.

4.2 ROI 추출 및 빈 홀 탐색

트레이 상의 빈 홀을 탐지하기 위해 먼저 입력 영상을 색상, 채도, 명도를 나타내는 HSV 색 공간으로 변환한다. 이후 미리 설정된 색상, 채도, 명도 임계 값 범위에 해당하는 영역을 추출하여 각 픽셀 좌표를 나타내는 $M(x, y)$ 를 생성한다.

$$M(x, y) = \begin{cases} 1 & \begin{aligned} H_{min} &\leq H(x, y) \leq H_{max} \\ S_{min} &\leq S(x, y) \leq S_{max} \\ V_{min} &\leq V(x, y) \leq V_{max} \end{aligned} \\ 0 & otherwise \end{cases} \quad (5)$$

마스크 결과에 대해 필터를 적용하여 노이즈를 제거한 후 피킹한 물체의 플레이싱을 진행할 빈 홀 후보를 분리한다. 각 후보 영역은 식(6)과 같이 원형성 지표 C 를 통해 정의된다.

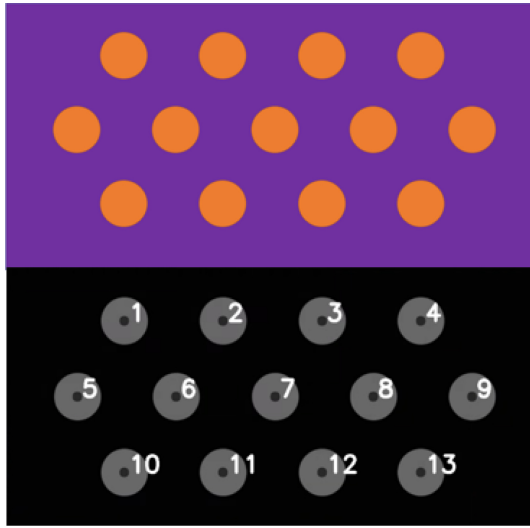


Fig. 6 Detected empty holes using the pattern generation algorithm

$$C = \frac{4\pi A}{P^2} \quad (6)$$

여기에서 A 는 빈 홀 후보 영역의 면적, P 는 둘레를 각각 나타낸다. $C \approx 1$ 에 가까울수록 원형성이 높아 실제 빈 홀일 가능성이 크다. 또한 반경 이상 데이터를 제거하기 위해 평균 반경 $\bar{r}$ 과 표준편차 σ_r 를 이용하여 다음 조건을 적용한다.

$$|r_i - \bar{r}| > k\sigma_r \quad (7)$$

이때 k 는 이상 데이터 판단을 위한 임계 계수로, 보통 2-3 정도로 설정되며, 평균값에서 너무 벗어나는 크기의 조건을 가진 홀은 제거된다. 이를 통해 트레이 상에서 조명이나 가림에 영향을 받지 않는 신뢰도 높은 빈 홀 후보 집합을 확보할 수 있다.

4.3 패턴 생성

ROI 추출을 통해 찾아낸 빈 홀 정보를 바탕으로 트레이의 격자 패턴을 만들어낸다. 단순히 홀의 좌표만을 이용하는 경우, 일부 홀이 인식되지 않으면 전체 격자 구조가 적절하게 생성되지 않는 문제가 있다. 이를 해결하기 위해 본 연구에서는 중앙값과 평균 절대 편차(Mean Absolute Deviation, MAD)를 이용하여 홀을 가로줄(행) 단위로 나누는 방식을 적용한다.

$$MAD = \text{med}(|y_i - \text{med}(Y)|) \quad (8)$$

여기에서 $Y = \{y_1, y_2, \dots, y_n\}$ 는 검출된 홀들의 Y -축 좌표 집합이다.

$$|y_i - \text{med}(Y)| > \lambda \cdot MAD \quad (9)$$

이상 데이터 제거 조건은 식(9)로 정의되며, λ 는 이상 데이터 제거의 민감도를 나타낸다. 이러한 과정을 통해 잡음에 해당하는 홀 좌표는 제외된다. 이후 데이터를 미리 정한 개수의 그룹으로 나눈 뒤, 각 그룹 내의 데이터가 중심점과 가까운 정도를 분류하는

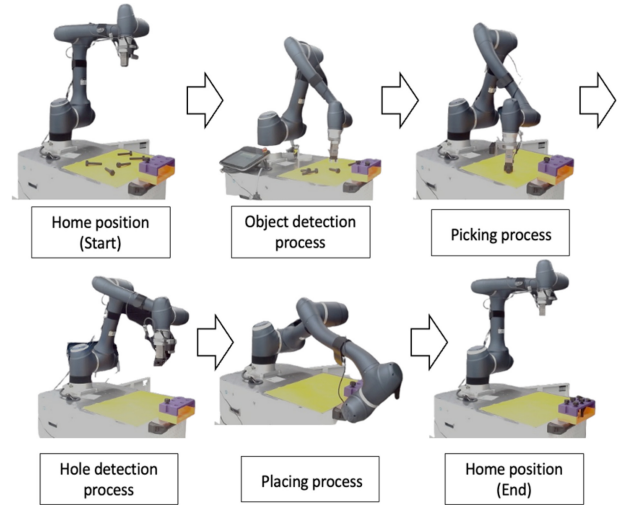


Fig. 7 Sequence of the entire bin-picking and adaptive placing algorithm

K-Means 군집화를 적용하여 홀을 행렬 형태로 정렬하였으며, 최적의 군집 개수는 각 점이 자기 군집 내에서의 일치와 차이를 나타내는 실루엣 점수를 통해 결정한다. 실루엣 점수는 식(10)과 같이 정의된다.

$$S = \frac{b-a}{\max(a,b)} \quad (10)$$

여기서 a 는 해당 점과 같은 군집 내 다른 점들 사이의 평균 거리, b 는 가장 가까운 다른 군집과의 평균 거리이다. 점수 S 가 클수록 더 좋은 군집화 결과임을 의미한다. Fig. 6은 제안된 방식에 따라 빈 홀이 검출되고, 최종적으로 규칙적인 격자 패턴이 생성되는 과정을 보여준다.

5. 실험을 통한 검증

제안한 실시간 인스턴스 세그멘테이션 기반 빈피킹 및 적응형 배치 시스템의 성능을 검증하기 위해 실제 실험 환경을 구축하고 검증 절차를 수행하였다. Fig. 7은 실제 실험 환경에서의 하드웨어 세팅과 빈피킹 및 플레이싱 과정을 보여준다.

5.1 객체 검출 성능

알고리즘의 유효성을 검증하기 위해 사용된 객체는 볼트이다. 볼트 검출 단계에서는 YOLO 인스턴스 세그멘테이션을 활용하여 영상 데이터에서 볼트에 해당하는 마스크를 실시간으로 추출하였다. 추출된 마스크 내부 픽셀은 깊이 정보와 결합하여 3차원 포인트 클라우드로 변환되었으며 이를 바탕으로 PCA 분석을 통해 길이 방향 축 및 헤드를 판별하였다. 약 9,000장의 테스트 이미지를 대상으로 학습된 모델을 적용한 결과, Fig. 8에 나타난 그래프와 같이 mAP@0.5는 약 93%, mAP@0.5-0.95는 약 87%를

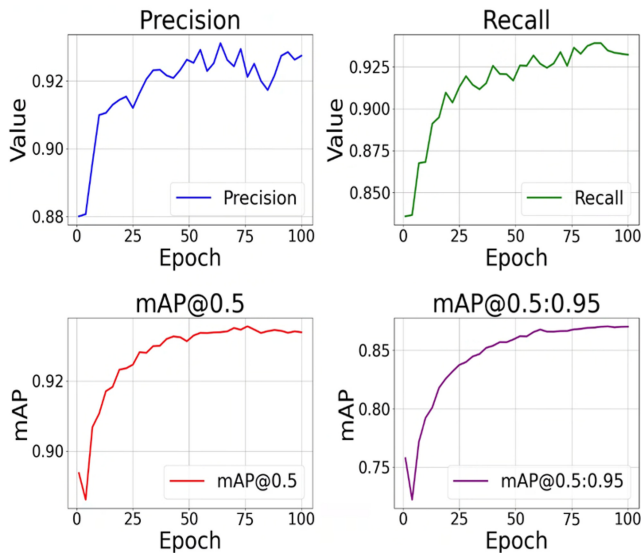


Fig. 8 mAP value using YOLO custom detect model

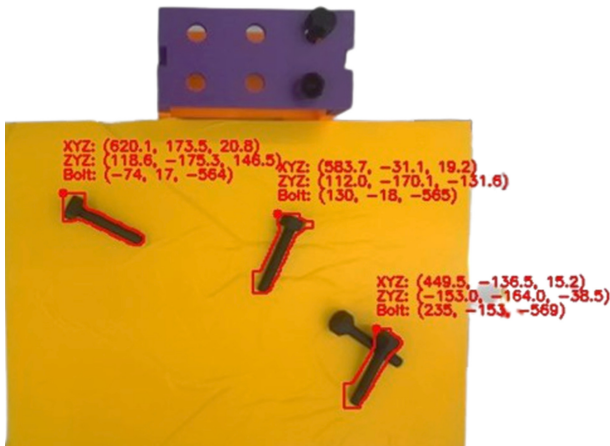


Fig. 9 Bolt detection using the real-time instance segmentation algorithm

기록하였다. Fig. 9에서 확인할 수 있듯이 볼트를 안정적으로 인식하였으며 마스크 기반 포인트 클라우드 추출 후 PCA를 통해 객체의 길이 방향 축 및 헤드를 안정적으로 판별할 수 있었다. 특히 슬라이싱 기반 폭 최대 구간 분석 기법을 적용함으로써 대칭 구조를 가진 볼트에서 발생할 수 있는 방향 뒤바뀔 문제를 효과적으로 해결할 수 있음을 확인할 수 있었다.

5.2 트레이 빈 홀 검출 성능

트레이 상의 빈 홀 검출은 HSV 임계화와 노이즈 제거를 통해 ROI를 추출한 뒤, 원형성 지표 및 반지름 이상 데이터 제거 기법으로 수행한다. 이후 중앙값 + MAD 기반 Y축 간격 분할 및 K-Means실루엣 점수 기반 보정을 통해 행과 열의 정렬을 진행하였다.

100회의 반복 실험에서 홀 검출 성공률은 97%로 나타났다. 또한 트레이 일부가 가려진 환경과 작업자의 실수 등으로 생기는

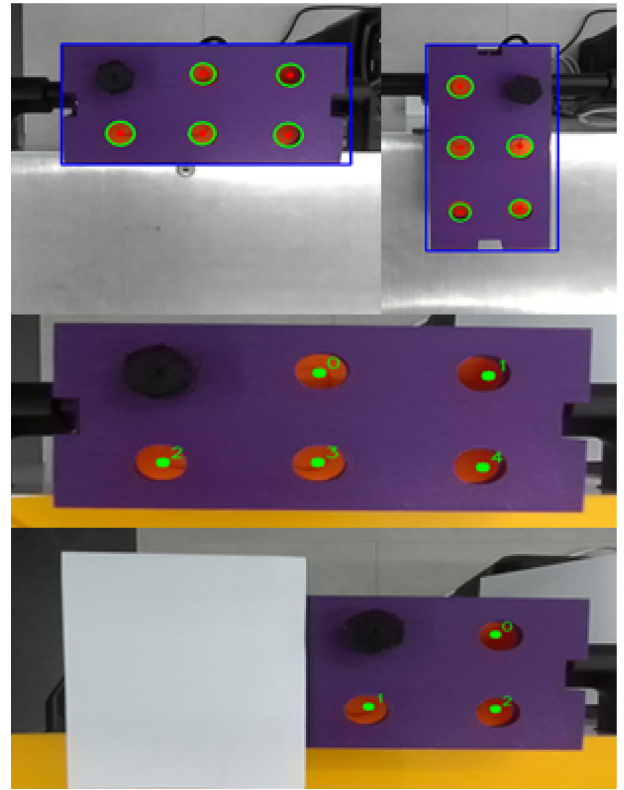


Fig. 10 Experiment on the adaptive placing algorithm (tray with aligned-pattern)

트레이의 변위가 바뀌는 환경(0-90°, 5° 간격)에 대한 실험에서도 각각 96%, 93%의 성공률을 기록하였다. Fig. 10은 가로세로 방향으로 홀이 정렬된 패턴의 트레이에서 홀을 검출하는 장면을 나타내며 제안된 알고리즘이 다양한 자세 및 부분 가림 상황에서도 홀 패턴을 비교적 안정적으로 복원하는 결과를 보여준다.

또한, 지그재그 패턴의 홀 배치 구조를 가진 트레이를 적용하여 실험을 진행하였다. Fig. 11은 트레이가 모두 비어 있는 상황, 일부 홀이 부품으로 점유되어 있는 상황에서의 검출 장면을 나타낸다. 총 100회의 실험 중 빈 홀 검출 성공률은 97%로 측정되었다. 실패 사례는 ROI 검출 실패(약 2%), 패턴 추출 실패(약 1%)로 나타났다.

5.3 피킹 및 플레이싱 성공률

Table 2에 제시된 것처럼 100회에 걸친 볼트 피킹 실험을 수행한 결과 볼트만 사용한 환경에서의 무작위 자세 및 겹침의 상황에서 전체 성공률은 95%로 나타났다. 실패 사례는 주로 포인트 클라우드 노이즈에 의한 자세 추정 오차(약 4%)와 로봇 그리퍼의 미끄러짐(약 1%)에서 발생하였다. 플레이싱의 성능은 피킹이 성공한 총 95회의 배치 실험으로 검증했다. 그 결과, 배치 성공률은 96%, 위치 오차는 평균 0.5 mm, 표준편차는 ± 0.4 mm로 나타났다.

제안한 빈피킹 및 적응형 플레이싱 알고리즘의 성능을 추가적으로

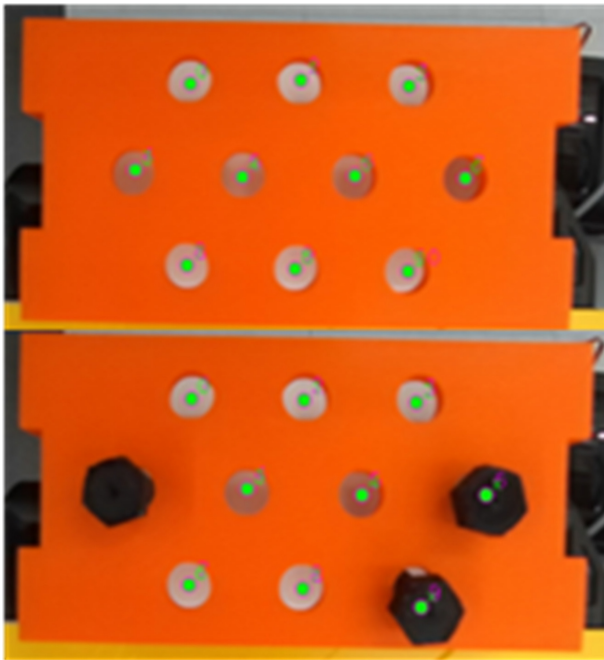


Fig. 11 Experiment on the adaptive placing algorithm (tray with zigzag-pattern)

Table 2 Experiment result of picking and placing

Experimental conditions	Number of experiment	Success rate [%]	Mean error [mm]	Standard deviation [mm]
Picking (only bolts exist)	100	95	-	-
Placing (only bolts exist)	95	96	0.5	± 0.4
Picking (various objects exist)	100	91	-	-
Placing (various objects exist)	91	94	0.84	± 0.47

검증하기 위하여 피킹하고자 하는 대상 물체가 다양한 크기 및 형상의 물체와 섞여서 무작위로 배치된 환경(크기와 길이가 다른 볼트, 너트, 축)에서도 피킹 및 플레이싱 실험을 진행하였다. 구체적인 실험 장면이 Fig. 12에 나타나 있다. 총 100회의 실험 중 볼트를 인식하지 못한 미검출은 6회, 볼트가 아닌 다른 물체를 인식한 오검출은 3회 발생하였다. 피킹 성공률은 91%, 플레이싱 성공률은 피킹이 성공한 총 91회의 배치 실험을 기준으로 약 94%로 나타났으며, 평균 위치 오차는 0.84 mm, 표준편차는 ± 0.47 mm로 나타났다. 볼트만 포함된 환경의 실험 대비 약간 증가한 편차를 보였으나 전체적인 편차는 1 mm 이하의 수준을 유지하고 있다.

이러한 결과는 빈피킹 뿐 아니라 플레이싱 공정의 자동화를 통해 공정 효율을 높일 수 있고 플레이싱 과정에서의 사고나 파손을 미연에 방지하는 것이 가능함을 보여준다.



Fig. 12 Picking and placing experiments in situations where objects of various shapes exist

6. 결론

본 연구에서는 저비용 깊이 인식 카메라와 YOLO 기반 인공 지능 모델을 활용하여 협동로봇 빈피킹 및 플레이싱 공정을 자동화하는 시스템을 제안하고 실험으로 검증하였다. 인스턴스 세그멘테이션을 통해 다양한 형상과 자세의 물체를 실시간으로 인식하였고, 최적화 기반 역기구학을 적용하여 안정적이고 유연한 로봇 동작을 구현하였다. 또한, 트레이 위치 및 자세 오차, 홀 점유, 부분 가림 등 실제 산업 현장의 불확실성에 대응할 수 있는 적응형 플레이싱 알고리즘을 개발하였다. 피킹 대상 물체가 다양한 형상의 물체와 함께 무작위로 배치된 환경에서 91%의 피킹 성공률을 보였고, 플레이싱은 94%의 성공률을 보임으로써 대칭 형상의 물체에서 발생하는 방향 오류와 임의 자세의 트레이와 같은 복잡한 상황에서도 제안하는 알고리즘을 효과적으로 적용할 수 있음을 보였다. 이러한 결과는 고가의 3D 비전 시스템에 비해 설치와 유지 비용을 크게 절감하면서도 중소 제조 환경에 적용 가능한 실용적 대안을 제시하며, 향후 조립, 물류, 서비스 로봇 등 다양한 산업 분야에서 지능형 자동화 공정의 확산과 효율성 향상에 기여할 것으로 기대된다.

REFERENCES

- Zhuang, C., Li, S., Ding, H., (2023), Instance segmentation based 6d pose estimation of industrial objects using point clouds for robotic bin-picking, *Robotics and Computer-Integrated Manufacturing*, 82, 102541.
- Kwon, D. H., Ha, T. M., Lee, J. S., Jeong, Y. S., Kim, Y. G., Kim, H. G., Song, S. J., Dae Gil, O., Lee, G., Jeong, J. W., (2024), The study of barista robots utilizing collaborative robotics and AI technology, *Journal of Drive and Control*, 21(3), 36-45.

3. Cho, J., Kang, S. S., Kim, K. K., (2019), Object recognition and pose estimation based on deep learning for visual servoing, The Journal of Korea Robotics Society, 14(1), 1-7.
4. Duc Hanh, L., Luat, N. V., Bich, L. N., (2022), Combining 3D matching and image moment based visual servoing for bin picking application, International Journal on Interactive Design and Manufacturing (IJIDeM), 16(4), 1695-1703.
5. Ojer, M., Lin, X., Tammara, A., Sanchez, J. R., (2022), Pickingdk: A framework for industrial bin-picking applications, Applied Sciences, 12(18), 9200.
6. Kwon, Y., Lee, D., Kang, H., Choi, J., Lee, I., (2024), Object pose estimation and motion planning for service automation system, The Journal of Korea Robotics Society, 19(2), 176-187.
7. Zhao, Y., Yang, J., Wang, S., Li, X., (2023), Towards one shot & pick all: 3D-OAS, an end-to-end framework for vision guided top-down parcel bin-picking using 3D-overlapping-aware instance segmentation and GNN, Robotics and Autonomous Systems, 167, 104491.
8. Dolezel, P., Stursa, D., Kopecky, D., (2024), Memory efficient deep learning-based grasping point detection of nontrivial objects for robotic bin picking, Journal of Intelligent & Robotic Systems, 110(3), 110.
9. Kozák, V., Sushkov, R., Kulich, M., Přeučil, L., (2021), Data-driven object pose estimation in a practical bin-picking application, Sensors, 21(18), 6093.
10. Boschetti, G., Sinico, T., Trevisani, A., (2023), Improving robotic bin-picking performances through human-robot collaboration, Applied Sciences, 13(9), 5429.
11. Fu, K., Dang, X., Zhang, Q., Peng, J., (2024), Fast uois: Unseen object instance segmentation with adaptive clustering for industrial robotic grasping, Actuators, 305.
12. Dias, J., Simões, P., Soares, N., Costa, C. M., Petry, M. R., Veiga, G., Rocha, L. F., (2023), Comparison of 3D sensors for automating bolt-tightening operations in the automotive industry, Sensors, 23(9), 4310.
13. Torres, P., Arents, J., Marques, H., Marques, P., (2022), Bin-picking solution for randomly placed automotive connectors based on machine learning techniques, Electronics, 11(3), 476.
14. Chen, Q., Wan, L., Pan, Y.-J., (2023), Robotic pick-and-handover maneuvers with camera-based intelligent object detection and impedance control, Transactions of the Canadian Society for Mechanical Engineering, 47(4), 486-496.



Ki-Suk Kim

B.S. candidate in the School of Robot and Automation Engineering, Dongyang Mirae University. His current research interests include quadruped walking robot based on SLAM and robot vision using artificial intelligence.
E-mail: kimkisuk@dongyang.ac.kr



Hyun-Pyo Shin

Associate Professor in the School of Robot and Automation Engineering, Dongyang Mirae University. His current research interests include the robot control, design of ultra-precision positioning stages based on flexure hinges and redundantly actuated parallel mechanisms.
E-mail: hpshin@dongyang.ac.kr