



• 특집 • 자율제조(Autonomous Manufacturing)

모방 학습 기반 매니플레이터 Peg-in-hole 공정 자동화 프레임워크 개발

Development of an Imitation Learning-based Manipulator Framework for Peg-in-hole process Automation

이병현¹, 오기용^{1,2,#}
Byeong Hyun Lee¹ and Ki-Yong Oh^{1,2,#}

¹ 한양대학교 대학원 융합기계공학과 (Department of Mechanical Convergence Engineering, Graduate School, Hanyang University)

² 한양대학교 기계공학부 (School of Mechanical Engineering, Hanyang University)

Corresponding Author / E-mail: kiyongoh@hanyang.ac.kr, TEL: +82-2-2220-3116

ORCID: 0000-0003-2895-4749

KEYWORDS: Imitation learning (모방 학습), Manipulator (로봇 팔), Compliance control (순응 제어), Domain randomization (도메인 랜덤화)

This paper presents an advanced robotic automation framework that combines an impedance-based compliance controller with an imitation learning network for high-precision peg-in-hole assembly. The framework is characterized by three key features. First, it employs an impedance-based compliance controller to ensure stable contact. This approach enables the robot to adapt flexibly to external contact forces, functioning like a spring-damper system to prevent potential damage. Second, domain randomization is applied to both geometric and visual properties within a high-fidelity simulation environment. This strategy effectively narrows the reality gap, enhancing robustness against environmental uncertainties and visual disturbances. Third, the framework utilizes an action-chunking-transformer (ACT) network to predict precise action sequences based on multimodal data, reducing compounding errors in trajectory generation and improving assembly success rates. Each feature is supported by specific advancements, such as real-time force feedback integration, diverse simulation scenario generation, and multimodal sensor fusion. Extensive experiments conducted in various unseen environments demonstrate the framework's effectiveness, confirming its suitability for complex assembly tasks that require high adaptability and precision under diverse conditions.

Manuscript received: December 17, 2025 / Revised: January 28, 2026 / Accepted: February 9, 2026

NOMENCLATURE

u = Joint Torque Input
 J^T = Transpose of Jacobian Matrix
 F_{cmd} = Commanded Force in Task Space
 u_{null} = Null Space Control Input
 K_P = Proportional Gain of Spatial PID Controller
 K_D = Derivative Gain of Spatial PID Controller
 K_I = Integral Gain of Spatial PID Controller
 u_{task} = Task Space Error Vector

K_{stiff} = Virtual Stiffness Coefficient
 x = Current Position
 F_{target} = Target Contact Force
 F_s = Raw Sensor Measured Force
 M_d = Desired Mass Matrix
 D_d = Desired Damping Matrix
 K_d = Desired Stiffness Matrix
 \bar{J} = Dynamically Consistent Inverse of Jacobian
 k_v = Damping Coefficient for Null Space Control
 \dot{q} = Damping Coefficient for Null Space Control

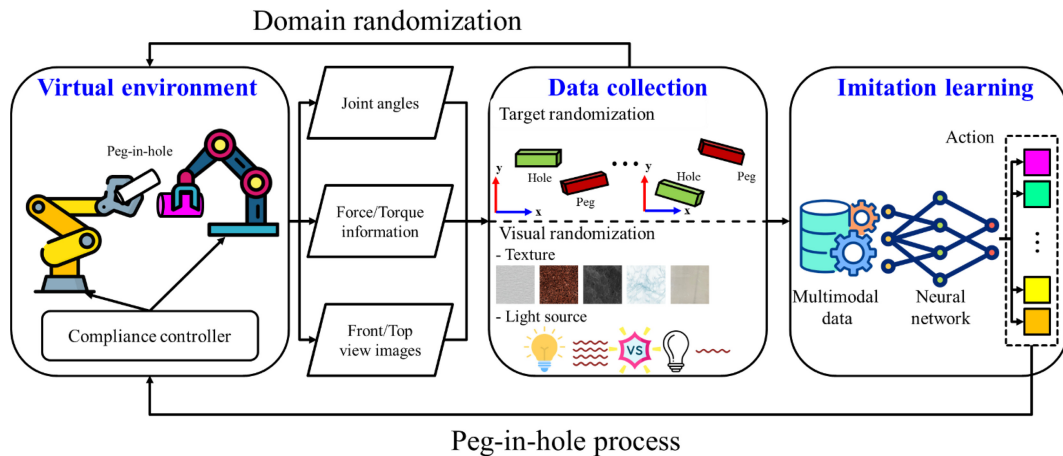


Fig. 1 Pipeline of the proposed framework for peg-in-hole process automation

1. 서론

최근 제조 지능화와 함께 스마트 팩토리 내에서 작업 성공률을 높이고 효율성을 극대화하기 위한 매니플레이터(Manipulator) 제어 기법에 대한 연구가 활발히 수행되고 있다. 특히 매니플레이터의 경로 계획(Path Planning) 관련 연구는 작업의 효율성과 안정성을 동시에 확보하기 위해 지속적으로 진행되어 왔으나[1], 사전에 정의된 환경 모델에 의존하기 때문에 비정상적인 환경 변화나 예측하지 못한 외란에 적응하지 못하고 여전히 비효율적인 움직임을 보인다는 단점이 존재한다.

최근 컴퓨터 연산 속도의 향상 및 데이터 처리 기술의 발전은 인공지능 기반의 로봇 제어 연구를 가속화하였다. 구체적으로, 보상 함수를 설정하여 이를 극대화하는 방향으로 로봇의 행동을 최적화하는 강화학습(Reinforcement Learning)이 다양한 제어 문제에 적용되고 있다[2]. 그러나 강화학습은 수렴을 위해 방대한 횟수의 시행착오가 필요하여 샘플(Sample) 효율성이 낮고, 학습 과정에서 발생할 수 있는 안전 문제로 인해 실제 제조 현장에 바로 적용하기에는 신뢰성이 부족하다는 한계가 있다[3,4].

이에 대한 대안으로 전문가의 시연 데이터를 기반으로 로봇이 작업을 학습하는 모방 학습(Imitation Learning)이 주목받고 있다 [5]. 모방 학습은 복잡한 보상 함수 설계 없이도 직관적인 교사가 가능하다는 장점이 있으나, 학습에 사용된 데이터 분포에 성능이 지나치게 의존적이라는 단점이 존재한다. 즉, 전문가 시연 데이터에 포함되지 않은 상황이 발생할 경우 로봇의 대응 능력이 현저히 떨어지기 때문에, 다양한 환경 변수를 포괄하는 고품질의 데이터를 구축하는 것이 필수적이다.

따라서 본 연구에서는 외력에 의한 접촉이 빈번한 펙인홀(Peg-in-hole) 공정에 효과적인 순응 제어기(Compliance Controller)를 가상 환경에 도입하여 높은 현실성을 갖는 학습 데이터를 취득하는 방법을 제안한다. 또한, 가상 환경과 실제 환경의 시각적 차이를 극복하기 위해 도메인 랜덤화(Domain Randomization) 기법을 적용하여 조명이나 질감 변화와 같은 광학적 외란에 취약한 이미지 기반 제어의 단점을 보완하였다[6]. 이를 바탕으로

본 연구는 Action-chunking-transformer (ACT) 네트워크[7]를 네트워크를 도입한다. ACT는 단일 시점의 행동이 아닌 일련의 행동 시퀀스를 한 번에 예측함으로써, 기존 모방 학습의 고질적인 문제인 오차 누적(Compounding Error)을 억제하고 정밀한 연속 동작 생성이 가능하기 때문에 본 공정에 적합하다. 최종적으로 본 연구는 힘/토크 센서 데이터, 매니플레이터의 관절 각도, 그리고 다중 시선(Multi-view) 이미지를 통합적으로 활용하는 ACT 네트워크를 구축하여, 환경 변화에 강건하면서도 높은 작업 성공률을 달성하는 자동화 프레임워크를 제공하고자 한다.

펙인홀 공정은 고전적인 연구에 따르면 전체 조립 작업의 약 40%를 차지할 만큼 제조 산업 현장에서 가장 빈번하게 수행되는 공정으로 알려져 있다[8]. 특히 최근에는 불확실성이 높은 환경에서의 정밀한 힘 제어와 위치 제어가 동시에 요구되는 접촉 기반 작업의 특성으로 인해, 지능형 로봇의 조작 성능을 검증하는 핵심 벤치마크로 그 중요성이 재조명 받고 있다[9,10]. 본 연구에서 제시하는 펙인홀 공정 자동화 프레임워크는 다양한 환경적 외란에 대해 강건한 제어 성능을 보이기 때문에, 스마트 팩토리 및 공정 자동화 등 다양한 정밀공학 분야에 활용이 가능할 것으로 사료된다.

2. 모방학습 기반 공정 자동화 프레임워크

본 장에서는 외란이 빈번하고 고정밀 제어가 요구되는 펙인홀 조립 공정을 성공적으로 수행하기 위해, 제어 이론과 데이터 기반 학습 기법이 결합된 통합 자동화 프레임워크를 제안한다. Fig. 1에 도시된 바와 같이 제안하는 프레임워크는 물리적 접촉에 유연하게 대응하는 하위 제어기(Low-level Controller)를 가상 환경에 탑재하였으며, 환경적 특성을 다양하게 포함하고 있는 데이터 취득 방법론, 해당 데이터를 기반으로 전문가의 작업 기술을 모방하는 상위 정책 네트워크(High-level Policy Network) 파이프라인으로 구성된다.

2.1 외란 강건성을 위한 순응 제어기

팩인홀 공정과 같이 로봇과 환경 간의 접촉이 빈번하게 발생하는 작업에서는 위치 제어만으로는 미세한 위치 오차나 외부 충격에 유연하게 대처하기 어렵다. 따라서 본 연구에서는 팩인홀 공정과 같이 외력이 빈번한 작업에서 로봇이 유연하게 대처할 수 있도록 임피던스(Impedance) 기반의 컴플라이언스 제어를 적용하였다. 본 제어기의 최종 목적은 작업 공간(Task Space)에서 계산된 필요 힘을 관절 공간(Joint Space)의 토크로 변환하여 로봇을 구동하는 것이다. 따라서 제어기의 최종 출력인 관절 토크 입력 u 는 식(1)과 같이 정의된다[11].

$$u = J^T F_{cmd} + u_{null} \tag{1}$$

여기서 로봇의 자코비안(Jacobian) 전치 행렬 J^T 는 작업 공간 힘 F_{cmd} 을 관절 토크로 맵핑(Mapping)하는 역할을 수행한다. 핵심이 되는 작업 공간 제어 힘 F_{cmd} 는 로봇이 목표 위치와 힘을 추종하면서 동시에 외력에 순응하도록 설계되어야 한다. 이를 위해 본 연구는 식(2)와 같은 공간 PID 제어(Spatial PID Control) 구조를 채택하여 F_{cmd} 를 산출한다.

$$F_{cmd} = K_p u_{task} + K_D \frac{d}{dt} u_{task} + K_I \int u_{task} d\tau \tag{2}$$

식(2)에서 u_{task} 는 제어기의 목표 값과 현재 값 사이의 오차를 정의하는 항이다. 본 연구에서는 위치 오차와 힘 오차를 결합하여 식(3)과 같이 가상의 오차 벡터 u_{task} 를 정의하였다.

$$u_{task} = K_{stiff}(x_d - x) + F_{target} - F_s \tag{3}$$

여기서 K_{stiff} 는 가상 강성 계수, x_d 와 x 는 각각 목표 위치와 현재 위치를 의미한다. F_{target} 은 목표 접촉 힘이며, F_s 는 필터링 된 센서 측정 외력이다. 식(3)과 같이 제어 입력을 설계한 물리적 근거는 로봇이 식(4)와 같은 임피던스 모델을 추종하도록 유도하기 위함이다[12].

$$F_s = M_d(\ddot{x}_d - \ddot{x}) + D_d(\dot{x}_d - \dot{x}) - K_d(x_d - x) \tag{4}$$

즉, 본 연구의 제어기 식(3)은 물리적으로 식(4)의 강성 항 $K_d(x_d - x)$ 과 외력 항 F_s 의 관계를 제어 입력으로 재해석한 것이다. 로봇은 외력이 목표치 F_{target} 보다 클 경우, 식(3)에 의해 그 힘의 차이를 줄이는 방향으로 가상의 위치 변위를 생성하게 되며, 이는 식(4)의 임피던스 모델에 따라 외부 환경에 순응하는 동작을 유도한다. 마지막으로, 식(1)의 영공간 제어 입력 u_{null} 은 주 작업인 F_{cmd} 에 영향을 주지 않으면서 관절의 에너지를 최소화하거나 불필요한 움직임을 억제하기 위해 식(5)과 같이 산출된다.

$$u_{null} = (I - J^T J^T)(-k_v \dot{q}) \tag{5}$$

이러한 제어 구조를 통해 로봇은 외력이 없는 자유 공간에서는 높은 위치 정밀도를 유지하고, 접촉이 발생하는 구속 공간에서는 외력에 순응하여 부품의 파손을 방지하고 조립 성공률을 높일 수 있다. 한편, 순응 제어기의 성능을 결정하는 핵심 파라미터인

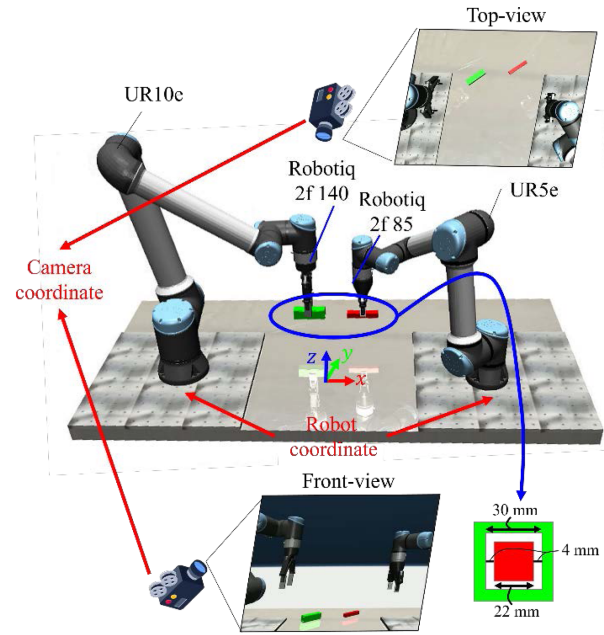


Fig. 2 Virtual environment of Mujoco simulator for the peg-in-hole task. This environment comprises two manipulators with two grippers

가상 강성(K_{stiff})과 감쇠 계수(D_d)는 예비 실험을 통해 선정되었다. 구체적으로, 외력이 인가되었을 때 시스템의 최종 수렴 오차가 2 mm 이하가 되도록 계수를 튜닝하여 과도한 위치 편차 없이 안정적인 접촉을 유지하도록 설정하였다. 선정된 파라미터 값들은 모든 학습 에피소드 및 테스트 과정에서 고정된 상수로 적용되어, 딥러닝 모델이 제어기의 일관된 동적 특성을 학습할 수 있도록 하였다.

2.2 학습 데이터 취득

본 연구에서는 실제 물리 현상을 높은 재현성으로 모사할 수 있는 MuJoCo 물리 엔진을 기반으로 학습 데이터 취득 환경을 구축하였다. 구축된 가상 환경의 전체적인 구성은 Fig. 2와 같다. 팩인홀 공정을 수행하기 위해 Universal Robotics사의 UR10e와 UR5e 매니퓰레이터로 구성된 듀얼 암(Dual-arm) 작업대를 구현하였으며, 각 매니퓰레이터의 제어에는 2.1절에서 제안한 컴플라이언스 제어를 적용하여 접촉 안정성을 확보하였다. 학습을 위해 수집되는 데이터는 정면 및 상면 시점의 RGB 이미지, 각 매니퓰레이터 말단 장치(End-effector)에 부착된 6축 힘/토크 센서(Force/Torque Sensor) 값, 그리고 로봇의 관절 각도 및 그리퍼 상태 정보로 구성된다.

실험에 사용된 작업 대상물인 팩과 홀은 각각 한 변의 길이가 22와 30 mm인 정사각형 단면을 갖는다(Fig. 2). 기구학적으로 이는 중심축을 기준으로 편차 4 mm의 조립 여유를 의미한다. 그러나 이 수치는 팩과 홀의 자세(Orientation)가 완벽하게 정렬된 이상적인 조건을 전제로 한다. 실제 공정에서는 로봇의 파지 과정에서 미세한 회전 오차가 필연적으로 발생하며, 팩이

기울어진 상태로 진입할 경우 기하학적 간섭이 발생하는 유효 범위는 4 mm보다 훨씬 좁아지게 된다. 즉, 단순한 위치 오차뿐만 아니라 자세 오차까지 복합적으로 작용하는 상황에서는 물리적 충돌이 불가피하므로, 다양한 외란이 존재하는 환경에서도 정밀한 제어 성능을 유지할 수 있는 강건한 학습 전략이 요구된다.

특히, 시뮬레이션 환경에서 학습된 모델이 실제 환경의 불확실성에 강건하게 대응할 수 있도록 도메인 랜덤화 기법을 적용하였다. 본 연구에서는 랜덤화 전략을 기하학적 위치 변화(Geometric Randomization)와 시각적 다양성(Visual Randomization)의 두 가지 주요 측면으로 구분하여 수행하였다. 한편, 마찰 계수나 링크 질량 등의 물리 인자 파라미터는 본 실험에서 고정된 값을 사용하였다. 이는 팩인홀 공정에서 발생하는 접촉 역학적 불확실성(Contact dynamics Uncertainty)을 2.1절의 임피던스 제어가 실시간으로 흡수 및 보상하도록 설계되었기 때문이다. 따라서 본 학습 프레임워크는 물리적 파라미터의 변동보다는, 제어가 대응하기 어려운 시각적 인지 오차와 기하학적 위치 오차를 극복하는 데 초점을 맞추었다.

첫째, 기하학적 다양성을 확보하기 위해 작업 대상물인 펌(Peg)과 홀(Hole)은 6자유도 자유 관절로 모델링 되었으며, 매 에피소드가 시작될 때마다 초기 위치와 자세를 무작위로 초기화하여 데이터의 다양성을 확보하였다. 구체적으로 펌과 홀의 기준 위치를 중심으로 X축 방향으로 ± 50 mm, Y축 방향으로 ± 100 mm범위 내에서 균등 분포를 따르는 위치 오차를 인가하였다. 자세의 경우, Z축 회전(Yaw)에 대해 $\pm \pi/4$ 범위의 무작위 회전 변위를 적용하여 다양한 진입 각도에 대한 대응 능력을 학습하도록 설계하였다.

둘째, 시각적 복잡도에 대한 네트워크의 인식 성능을 향상시키기 위해 환경의 광학적 특성을 무작위로 변동시켰다. 구체적으로 조도와 활성화되는 광원의 개수, 바닥 및 배경의 질감, 그리고 펌과 홀의 RGB 색상 값을 매 에피소드마다 변화시켜 특정 색상이나 조명 조건에 과적합(Overfitting)되는 것을 방지하였다.

데이터 수집 과정에서 두 매니퓰레이터가 수행하는 팩인홀 공정은 명확한 단계 구분을 위해 3개의 세부 동작으로 분할된다. 이는 대상물을 파지하고(Pick), 파지 후 조립 위치 상단까지 들어 올린 후(Lift), 최종적으로 펌을 홀에 삽입하여 조립(Peg-in-hole)을 완료 단계로 구성된다. 도메인 랜덤화를 통해 생성된 데이터셋의 시각적 다양성과 정의된 3단계 세부 동작은 Fig. 3에서 확인할 수 있다.

2.3 모방 학습 네트워크

복잡한 팩인홀 공정을 자동화하기 위해 본 연구에서는 전문가의 시연 데이터를 효과적으로 학습할 수 있는 ACT 아키텍처를 도입하였다[6]. ACT는 기존 행동 복제(Behavioral Cloning) 기법의 한계인 복리 오차 문제를 해결하기 위해, 단일 시점이 아닌 사전 정의한 길이의 행동 시퀀스(Action Chunk)를 예측하는 방식을 채택한다. 본 연구에서 구축한 네트워크는 크게 학습 단계에서의 잠재 변수 생성을 위한 CVAE (Conditional-variational-

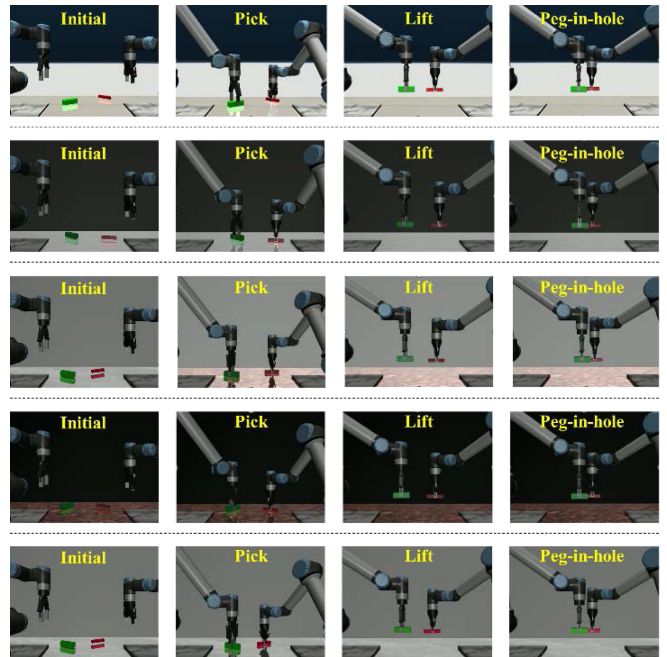


Fig. 3 Five representative dataset of the peg-in-hole task divided into three sub-tasks (Pick, Lift, Peg-in-hole). These front-view images denote that the domain randomization causes visual diversity

Autoencoder) 인코더와, 실제 행동을 생성하는 트랜스포머 기반의 정책 네트워크로 구성된다.

첫째, CVAE는 다양한 행동의 확률적 분포를 학습하기 위해 설계되었다[13]. 제공된 코드의 구현에 따라, CVAE는 BERT (Bidirectional Encoder Representations from Transformers) 스타일의 구조를 따르며[14], CLS 토큰(Special Classification Token), 현재의 로봇 관절 상태, 그리고 미래의 행동 시퀀스를 입력으로 받는다. 이 입력 토큰들은 고정된 정현파 위치 임베딩(Sinusoidal Positional Embedding)과 결합되어 인코더를 통과하며, 출력된 CLS 토큰은 선형 투영을 거쳐 잠재 공간(Latent Space)의 평균과 로그 분산을 생성한다. 학습 시에는 재매개변수화 트릭(Reparameterization Trick)을 통해 잠재 벡터를 샘플링하며, 이는 행동의 스타일(Style) 정보를 내포하게 된다.

둘째, 특징 추출 및 트랜스포머 인코더(Transformer Encoder)는 멀티모달(Multimodal) 데이터를 융합하는 역할을 수행한다. 시각 정보 처리를 위해 사전 학습된 ResNet-18을 백본 네트워크로 사용하였으며, 추출된 특징 맵은 컨볼루션 연산을 거쳐 차원이 축소된다[15]. 트랜스포머 인코더의 입력은 스타일 정보, 로봇의 관절 상태, 그리고 다중 카메라로부터 획득한 이미지 특징 맵으로 구성된다. 이때, 이미지 특징 맵은 공간적 정보를 보존하기 위해 2D 정현파 위치 임베딩(Embedding)이 더해지며, 1차원 데이터인 스타일일과 관절 정보에는 1D 위치 임베딩이 적용된다. 이러한 구조는 네트워크가 시각 정보와 기구학적 정보를 통합적으로 이해하여, 작업 환경 내에서 펌과 홀의 상대적 위치 관계를 정밀하게 파악하도록 돕는다.

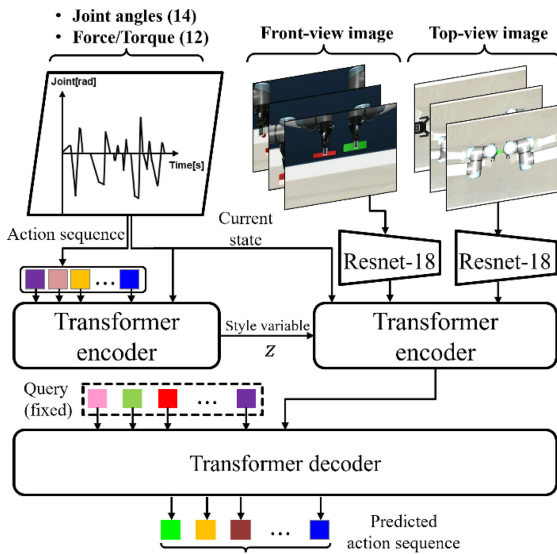


Fig. 4 Action chunking transformer network for the sophisticated contact-rich task using multi-modal information.

셋째, 트랜스포머 디코더(Transformer Decoder)는 인코딩된 정보를 바탕으로 최종 행동 시퀀스를 생성한다. 디코더는 k 개의 학습 가능한 위치 임베딩을 쿼리(Query)로 사용하여 인코더의 출력과 교차 어텐션(Cross-attention)을 수행한다. 디코더의 출력은 선행 레이어를 통과하여 최종적으로 k 시간 스텝에 해당하는 관절 위치 및 그리퍼 상태를 예측한다. 결과적으로 이러한 네트워크 구조는 시각적 인식과 행동 생성을 유기적으로 결합함으로써, 로봇이 불확실한 환경에서도 떨림 없이 부드럽고 일관된 궤적을 생성하여 정밀 조립 작업을 성공적으로 수행할 수 있는 기반을 제공한다.

3. 실험

3.1 네트워크 학습

본 연구에서 제안하는 펙인홀 공정 자동화 프레임워크의 성능을 검증하기 위해 ACT 기반의 모방 학습을 수행하였다. 모든 학습 및 실험은 Intel Core i9-11900K CPU와 NVIDIA GeForce RTX 3090Ti GPU가 탑재된 컴퓨팅 환경에서 진행되었으며, 학습 데이터셋은 숙련된 전문가의 시연 데이터 250개(Episodes)로 구성된다. 모델의 학습 효율성과 수렴성을 고려하여 배치 크기(Batch Size)는 8, 총 학습 스텝(Training Steps)은 300,000회로 모든 실험 조건에서 동일하게 설정하였다.

특히, 2.3절에서 제안한 도메인 랜덤화 기법과 멀티모달 정보의 유효성을 정량적으로 분석하기 위해, 학습 조건을 세 가지 케이스(Case)로 세분화하여 비교 실험을 설계하였다.

첫째, Case 1은 베이스라인(Baseline) 모델로서, 시각적 변화 없이 펙과 홀의 위치만 무작위로 변경하는 타겟 랜덤화(Target Randomization)만을 적용하였다. 입력 데이터로는 힘/토크 정보 없이 매니플레이터의 관절 각도 정보만을 사용하였다. 둘째,

Table 1 Definition of experimental cases for ablation study

Case	Case 1	Case 2	Case 3
Target randomization	O	O	O
Visual randomization	X	O	O
Input modality (Force/Torque)	X	X	O
Input modality (Joint angle)	O	O	O

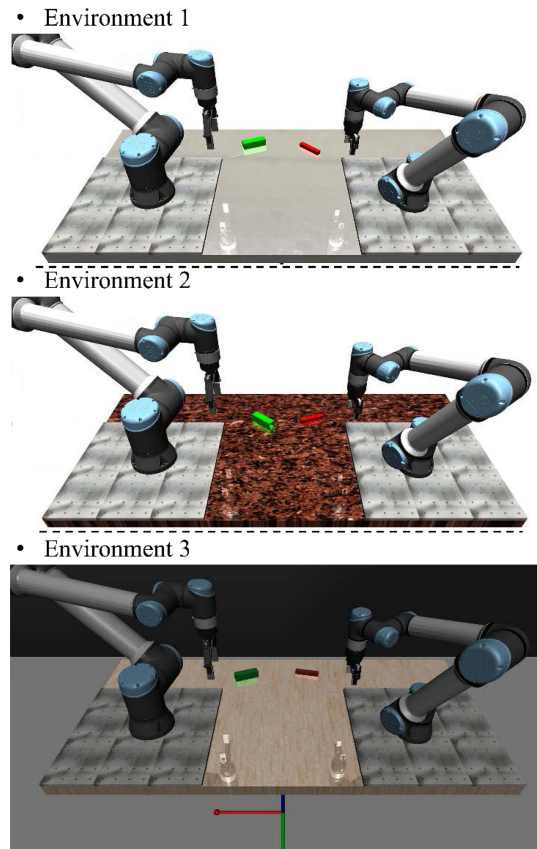


Fig. 5 Three test environment for verifying the robustness of the proposed framework

Case 2는 시각적 강건성을 평가하기 위한 모델로, 조명 및 텍스처 등 모든 종류의 도메인 랜덤화를 수행하여 시각적 다양성을 확보하였다. 그러나 입력 데이터는 Case 1과 동일하게 관절 정보만을 사용하여, 힘/토크 센서의 부재가 미치는 영향을 분석하고자 하였다. 셋째, Case 3은 본 연구에서 제안하는 최종 프레임워크이다. 도메인 랜덤화를 통해 구축된 다양한 시각적 환경 하에서 관절 정보뿐만 아니라 힘/토크 정보를 함께 학습함으로써, 접촉이 빈번한 조립 공정에서의 작업 성공률과 안정성을 극대화하고자 하였다. 각 실험 케이스별 상세한 학습 조건 및 입력 데이터 구성은 Table 1에 요약된 바와 같다.

3.2 실험 환경

제안된 프레임워크의 일반화 성능과 환경 강건성을 정량적으로 평가하기 위해, Fig. 5와 같이 학습에 사용된 환경과 전혀 다른

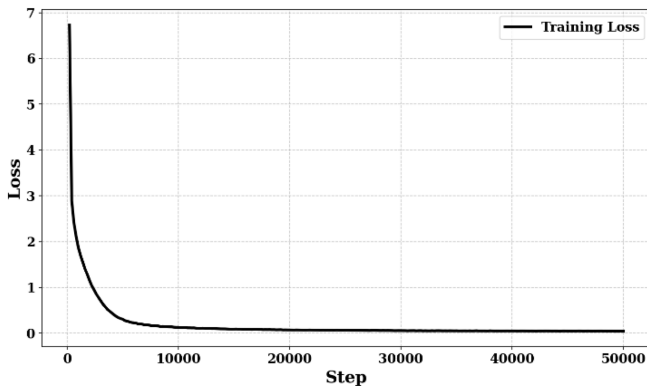


Fig. 6 Total loss of ACT network across 50,000 steps with respect to case 2

시각적 특성을 갖는 환경을 포함하여 총 세 가지 시나리오에서 실험을 진행하였다.

첫 번째 환경은 학습 데이터 수집에 주로 활용된 기본 환경으로, 시각적 랜덤화가 적용되지 않은 표준 상태이다. 이는 모델의 기본적인 학습 성취도를 확인하기 위한 대조군으로 설정되었다. 두 번째와 세 번째 환경은 학습 과정에서 모델이 한 번도 접하지 못한 새로운 시각적 환경이다. 구체적으로 조명의 위치와 세기, 작업대의 텍스처, 배경 색상 등을 학습 데이터의 분포와 상이하게 설정하여, 제안 모델이 시각적 외란에도 불구하고 강건하게 작업을 수행할 수 있는지 검증하고자 하였다.

각 환경마다 픽업 작업을 25회씩 반복 수행하여 작업 성공률을 측정하였다. 이때 단순한 최종 성공 여부뿐만 아니라 공정 단계별 실패 요인을 세밀하게 분석하기 위해, 2.3절에서 정의한 세 가지 하위 작업인 파지(Pick), 상승(Lift), 조립(Peg-in-hole, Pih) 단계별로 성공률을 산출하였다. 각 단계별 성공 여부는 다음과 같은 기준으로 판정하였다. 우선 파지 단계는 매니퓰레이터가 픽을 정확한 위치에서 안정적으로 파지하였는지 평가하며, 상승 단계는 파지 상태를 유지하며 조립 준비 위치까지 픽을 이탈 없이 이송하였는지를 확인한다. 마지막으로 조립 단계는 픽을 홀에 정밀하게 정렬시키고 삽입을 수행하여 최종적으로 공정을 완료하였는지 여부를 기준으로 성공률을 산출하였다. 이러한 단계별 평가는 특정 구간에서 발생하는 병목 현상을 파악하고, 멀티모달 정보 및 도메인 랜덤화가 각 동작 단계에 미치는 영향을 구체적으로 분석하는 데 활용된다.

4. 결과분석 및 고찰

본 절에서는 3.1절에서 정의한 세 가지 실험 케이스(Case 1, 2, 3)에 대한 학습 양상을 분석하고, 3.2절에서 제시한 세 가지 환경에서의 작업 성공률을 비교 평가한다.

우선, 학습 과정에서의 네트워크 수렴성을 확인하기 위해 대표적으로 도메인 랜덤화가 적용된 Case 2의 50,000 스텝까지의 손실 함수 변화를 Fig. 6에 도시하였다. Fig. 6은 Case 2의 결과

Table 2 Success rate for each case with respect to the sub-tasks (Pick, Lift, Peg-in-hole)

Success rate [%]	Case 1			Case 2			Case 3		
	Pi.	Li.	Pih	Pi.	Li.	Pih	Pih	Li.	Pih
Env. 1	80	80	64	84	84	64	100	100	88
Env. 2	8	8	0	88	88	48	88	88	76
Env. 3	68	68	44	80	80	64	88	88	68

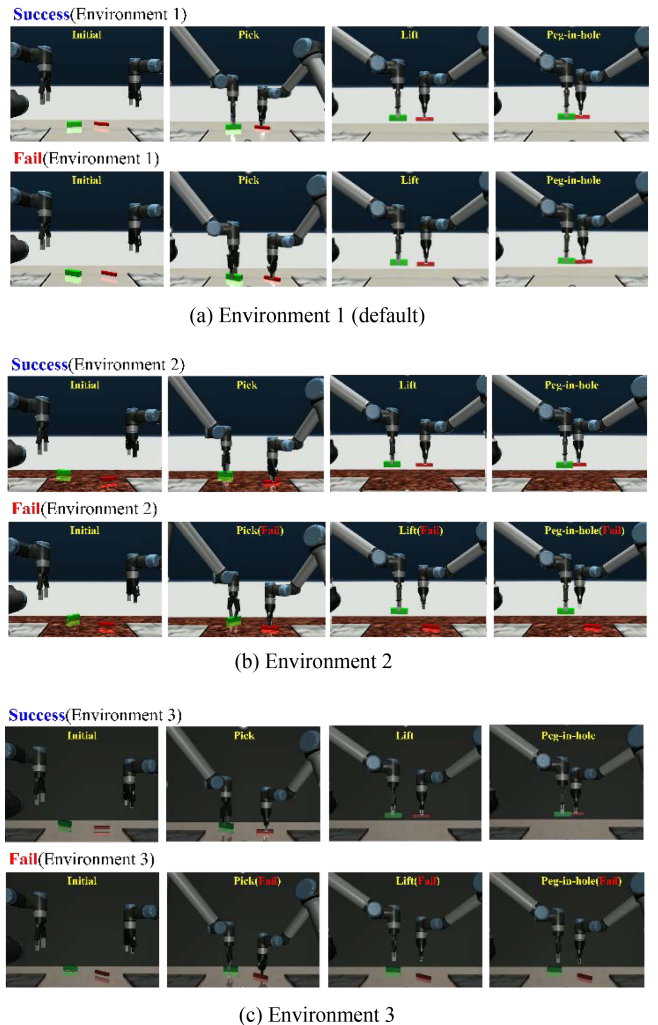


Fig. 7 Execution of Peg-in-hole task by using Imitation-learning for each environment. Comparison of the success/failure task from each case 1 and case 3

만을 나타내고 있으나, 데이터 로그 분석 결과 모든 케이스에서 학습이 진행됨에 따라 손실 값이 점차 감소하며 안정적으로 수렴하는 경향을 보였다. 총 300,000 스텝 학습을 완료한 시점에서의 최종 손실 값은 Case 1, Case 2, Case 3 각각 0.01670, 0.01414, 0.01856으로 확인되었다. 여기서 힘/토크 센서 정보를 추가로 활용한 Case 3의 손실 값이 상대적으로 높게 나타난 것은, 학습해야 할 데이터의 차원과 모달리티(Modality)가 증가함에 따라 최적화 난이도가 상승했기 때문인 것으로 사료된다.

학습된 각 학습 모델(Case 1, 2, 3)의 환경별 픽인홀 공정 작업 성공률은 Table 2와 같다. 본 절에서는 해당 결과를 바탕으로 환경 변화에 따른 모델의 강건성과 센서 데이터 융합의 효과를 분석한다.

첫째, 학습 데이터와 가장 유사한 기본 환경에서의 실험 결과이다. 해당 환경은 학습 데이터 분포 내에서 가장 높은 비중을 차지하므로, 모든 케이스에서 전반적으로 높은 성공률을 보였다. 특히 파지 및 상승 단계에서는 실패가 거의 발생하지 않았으나, Fig. 7(a)와 같이 최종 조립 단계에서 펌과 홀 간의 미세한 정렬 오차로 인한 충돌이 주요 실패 요인으로 작용했다. 특히 힘/토크 정보 없이 시각 정보에만 의존한 Case 1과 Case 2는 픽이 홀 입구에 걸렸을 때 이를 극복하지 못하고 멈추거나 과도한 힘을 주어 미션 실패로 이어지는 경향을 보였다. 주목할 점은 힘/토크 센서 데이터를 학습에 활용한 Case 3의 경우, 파지 단계 성공률이 100%를 기록하여 시각 정보만으로는 판단하기 어려운 정확한 파지 타이밍을 힘 정보를 통해 인지함을 확인하였다. 또한, 조립 단계에서 충돌이 발생하는 상황에서도 접촉 힘을 감지하고 궤적을 미세하게 수정하여 88%의 높은 성공률로 조립을 완료하였다. 이는 학습 과정에서 모델이 2.1절의 순응 제어기가 보인 접촉 적응 특성을 성공적으로 모방했음을 시사하며, 픽인홀과 같이 접촉이 빈번한 공정에서는 힘/토크 데이터의 활용이 필수적임을 입증한다.

두 번째 환경은 작업대의 재질이 대리석과 같이 복잡한 패턴을 가지며 작업물과 유사한 색상을 띠는 환경이다. 도메인 랜덤화를 적용하지 않은 Case 1의 경우, 학습 데이터에 존재하지 않았던 복잡한 배경 텍스처를 물체의 특징으로 오인하는 과적합 현상이 발생하였다. 이러한 시각적 혼동은 초기 물체 인식 단계에서부터 치명적인 오류를 야기하여, 파지 단계 성공률이 8%에 그치는 등 수행 능력이 현저히 저하되었으며 결과적으로 최종 조립까지 성공한 사례는 전무하였다(Fig. 7(b)). 반면, 다양한 텍스처와 색상 변화를 학습 과정에 포함시킨 Case 2와 Case 3는 배경과 객체를 효과적으로 분리해 내며 80%대의 안정적인 파지 성공률을 유지하였다. 이는 시각적 도메인 랜덤화가 네트워크로 하여금 배경의 텍스처가 아닌 물체의 고유한 형상 정보에 집중하도록 유도했음을 보여준다.

세 번째 환경은 복잡한 패턴은 부재하나 조도가 매우 낮아, 객체의 윤곽선이나 특징점을 식별하기 어려운 조건이다. Case 1은 텍스처 간섭이 심했던 두 번째 환경보다는 상대적으로 물체의 윤곽을 구분하기 용이하여 68%의 파지 성공률을 기록하였으며, 이에 따라 최종 조립 성공률 또한 44%로 유의미한 수치를 기록하였다. 하지만 어두운 환경으로 인해 정확한 깊이(Depth) 추정이나 미세 위치 파악에는 한계를 보여 절반 이상의 시도에서 실패하였다. 이에 반해, 다양한 조명 변동성을 포함한 데이터로 학습된 Case 2와 Case 3는 조도 변화에 강건한 특징을 학습하여 각각 64%와 68%의 더 높은 조립 성공률을 달성하였다. 이는 조도가 급격히 저하되거나 시각적 정보가 제한적인 상황에서도, 제안하는 도메인 랜덤화 기법과 멀티모달 학습 전략이 강건한 작업 성능 확보에 기여함을 확인하였다.

결론적으로, 단순히 실제 사용 환경과 동일한 데이터만을 사용하여 학습하는 것이 작업 성공률을 보장하지 않음을 확인하였다. 오히려 도메인 랜덤화를 통해 다양한 환경 데이터를 학습함으로써 네트워크가 환경 불변(Invariant)의 핵심 특징을 네트워크가 이해하도록 유도하는 것이 성능 향상에 기여함을 알 수 있다. 또한, 이러한 방식은 전이 학습(Transfer Learning) 측면에서도 효과적이며 처음 접하는 환경에서도 우수한 적응력을 보였다. 특히 힘/토크 센서의 상태 정보를 융합하여 학습한 경우(Case 3), 물체를 파지해야 할 정확한 순간을 포착하고 접촉 상황에 유연하게 대처함으로써 가장 높은 조립 완료율을 달성하였다.

5. 결론

본 연구에서는 임피던스 기반 순응 제어와 ACT 모방 학습을 결합하여, 외란에 강건한 픽인홀 공정 자동화 프레임워크를 제안하였다. MuJoCo 시뮬레이터를 활용한 멀티모달 데이터 학습과 도메인 랜덤화 기법을 통해 데이터 부족 문제를 해결하고 환경 적응력을 극대화하였다. 실험 결과, 제안 모델은 학습되지 않은 시각적 환경과 물리적 접촉 상황에서도 높은 성공률을 보이며 우수한 작업 강건성을 입증하였다. 특히, 다양한 시각적 및 물리적 변수가 적용된 테스트 환경에서 달성한 높은 성공률은, 본 프레임워크가 실제 환경이 내포한 불확실성과 노이즈에도 충분히 대응 가능함을 시사한다. 향후 연구에서는 이러한 강건성을 바탕으로 시뮬레이션과 실제 환경 간의 격차를 최소화하는 도메인 적응(Domain Adaptation) 기법을 적용하여, 향후 연구에서는 가상 환경의 학습 결과를 실제 로봇 시스템으로 확장하고, 다양한 형상의 작업물에 적용하여 프레임워크의 범용성을 검증할 계획이다.

ACKNOWLEDGEMENT

본 연구는 대한민국 행정안전부(No. RS-2024-00408982, 물류시설 특성을 반영한 지능형 화재탐지 및 스프링클러설비 기술개발)와 소방청 전기기반 모빌리티 관련 시설 및 부품 화재 대응 기술 개발 사업(No. RS-2024-00408270, 재사용 배터리의 운송·보관 단계에서의 화재 위험성 분석 및 화재안전기준(안 개발)의 연구비 지원으로 수행되었습니다.

REFERENCES

1. LaValle, S. M., (1998), Rapidly-exploring random trees: A new tool for path planning, The Annual Research Report.
2. Kober, J., Bagnell, J. A., Peters, J., (2013), Reinforcement learning in robotics: A survey, The International Journal of Robotics Research, 32(11), 1238-1274.

3. Brunke, L., Greeff, M., Hall, A. W., Yuan, Z., Zhou, S., Panerati, J., Schoellig, A. P., (2022), Safe learning in robotics: From learning-based control to safe reinforcement learning, *Annual Review of Control, Robotics, and Autonomous Systems*, 5(1), 411-444.
4. Zhao, W., Queraltá, J. P., Westerlund, T., (2020), Sim-to-real transfer in deep reinforcement learning for robotics: A survey, *Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI)*, 737-744.
5. Hussein, A., Gaber, M. M., Elyan, E., Jayne, C., (2017), Imitation learning: A survey of learning methods, *ACM Computing Surveys*, 50(2), 1-35.
6. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., Abbeel, P., (2017), Domain randomization for transferring deep neural networks from simulation to the real world, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 23-30.
7. Zhao, T. Z., Kumar, V., Levine, S., Finn, C., (2023), Learning fine-grained bimanual manipulation with low-cost hardware, *arXiv preprint arXiv:2304.13705*.
8. Nevins, J. L., Whitney, D. E., (1978), Computer-controlled assembly, *Scientific American*, 238(2), 62-75.
9. Lee, M. A., Zhu, Y., Srinivasan, K., Shah, P., Savarese, S., Fei-Fei, L., Garg, A., Bohg, J., (2019), Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks, *Proceedings of the International Conference on Robotics and Automation*, 8943-8950.
10. Schoettler, G., Nair, A., Luo, J., Bahl, S., Ojea, J. A., Solowjow, E., Levine, S., (2020), Deep reinforcement learning for industrial insertion tasks with visual inputs and natural rewards, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 5548-5555.
11. Khatib, O., (1987), A unified approach for motion and force control of robot manipulators: The operational space formulation, *IEEE Journal on Robotics and Automation*, 3(1), 43-53.
12. Hogan, N., (1985), Impedance control: An approach to manipulation: Part I—theory, *ASME Journal of Dynamic Systems, Measurement, and Control*, 107(1), 1-7.
13. Sohn, K., Lee, H., Yan, X., (2015), Learning structured output representation using deep conditional generative models, *Advances in Neural Information Processing Systems*, 28, 3483-3491.
14. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., (2019), BERT: Pre-training of deep bidirectional transformers for language understanding, *Proceedings of the North American Chapter of the Association for Computational Linguistics*, 4171-4186
15. He, K., Zhang, X., Ren, S., Sun, J., (2016), Deep residual learning for image recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-778.



Byeong Hyun Lee

M.S. in the Department of Mechanical Convergence Engineering, Hanyang University. His research interest is robust control of robot manipulators.

E-mail: byeonghyun98@hanyang.ac.kr



Ki-Yong Oh

Professor in the Department of Mechanical Engineering, Hanyang University. His teaching and research interests include applied dynamics and prognostics and health management in the field of robotic systems.

E-mail: kiyongoh@hanyang.ac.kr