https://doi.org/10.7735/ksmte.2026.35.5.381

J. Korean Soc. Manuf. Technol. Eng.
ISSN 2508-5107(Online)

Check for updates

센서 노이즈에 강건한 합성 깊이 데이터 기반 금속 물체의 5자유도 로봇파지 기술 개발

오민수[a], 이재학[a,*]

Depth-Noise-Robust 5-DOF Robotic Grasping of Metallic Objects Using Synthetic Depth Data

Minsoo Oh[a], Jaehak Lee[a,*]

[a] *Autonomous Manufacturing & Process Research Division, Korea Institute of Industrial Technology (KITECH)*

ARTICLE INFO

Article history:
Received 3 July 2026
Revised 4 August 2026
Accepted 7 August 2026

Keywords:
Robotic grasping
Synthetic depth data
Sim-to-real
Metallic object
5-DOF grasping
GQCNN

ABSTRACT

Metallic objects with specular surfaces induce severe optical noise in depth sensors, degrading the reliability of vision-based robotic grasping. This paper proposes a five-degree-of-freedom (5-DOF) grasping framework that is robust to such optical noise, using synthetic depth data generated by a physical sensor noise model. Using the SAPIEN simulator, depth images are rendered through an active stereo sensor model combined with the Landau et al. noise model to reproduce reflection noise on metallic surfaces. A grasp-quality convolutional neural network is trained on this dataset, and the conventional top-down grasp is extended to 5-DOF by adding tilt about the grasp axis based on wrench space invariance. Real-world experiments on ten objects at three tilt angles demonstrate that the proposed method achieves an overall grasp success rate of 90.4%. Additionally, it reliably grasps wall-adjacent objects inaccessible to vertical approaches.

1. 서 론

로봇 파지(grasping)는 빈피킹(bin-picking), 조립, 물류 등 비정형 제조 자동화의 핵심 요소 기술이다. 특히 다품종 소량생산 환경에서는 사전에 정의된 위치·자세가 아니라 무작위로 적재된 물체를 인식하고 파지해야 하므로, 비전 센서로 획득한 데이터로부터 파지 자세를 추론하는 능력이 요구된다. 이에 따라 RGB-D 영상에 딥러닝을 적용하여 파지 위치를 직접 검출하는 학습 기반 파지가 제안된 이래[1], 데이터 기반 파지 작업에 관한 다양한 연구가 이루어져 왔다[2].

특히 합성 데이터와 해석적 파지 품질 지표를 결합한 Dexterity Network (Dex-Net) 계열의 연구가 활발히 진행되었다. Dex-Net 1.0은 대규모 3차원 객체·파지 데이터셋을 구축하였고[3], Dex-Net 2.0은 합성 메쉬로부터 파지 성공 확률을 예측하는 Grasp-Quality CNN (GQCNN)을 제안하였다[4]. 이후 연구는 빈피킹을 위한 정책 학습[5], 흡착식 파지[6], 이종 그리퍼를 아우르는 범용 파지로 확장되었다[7]. 한편 GG-CNN은 파지 후보의 이산 샘플링 없이 픽셀 단위로 파지 품질을 생성하여 경량·실시간 파지를 구현하였으며[8], PointNetGPD는 점군을 직접 입력으로 사용하여 파지를 평가하였다[9].

나아가 평면적 수직 파지를 넘어 클러터 환경에서 6자유도 파지

* Corresponding author. Tel.: +82-31-8040-6163
E-mail address: ljh1125@kitech.re.kr (Jaehak Lee).

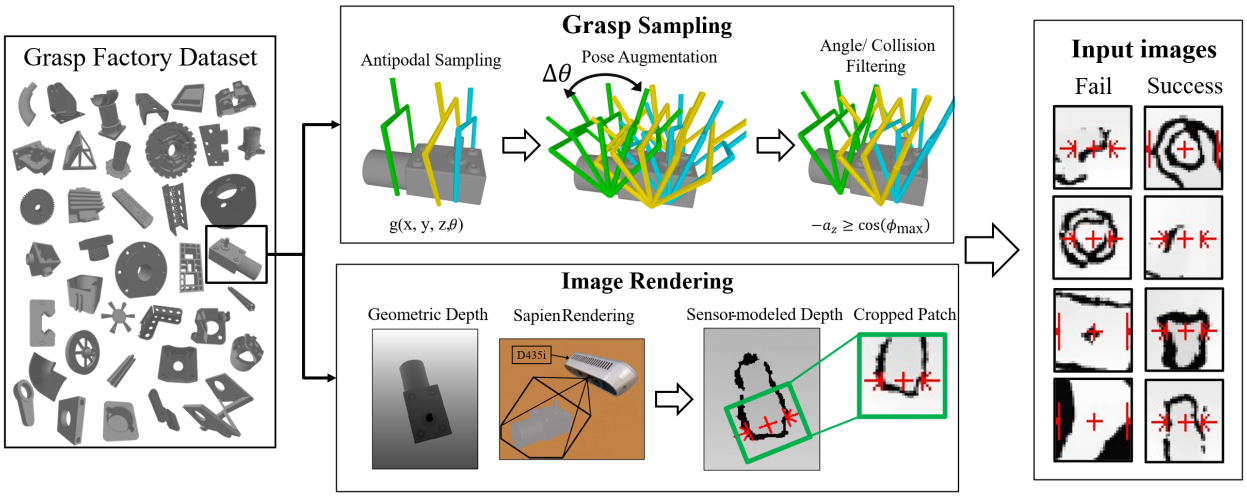


Fig. 1 Pipeline of the synthetic grasp dataset generation

를 생성하는 연구가 활발하다. Contact-GraspNet은 관측된 점군으로부터 6자유도 파지를 생성하였고[10], AnyGrasp은 공간·시간 영역에서 강건하고 효율적인 파지 인식을 달성하였으며[11], DexGraspNet 2.0은 대규모 합성 클러터 데이터셋과 생성 모델을 통해 다지(多指) 파지로 확장하였다[12].

그러나 이러한 깊이 기반 파지 기법은 입력 깊이 영상의 품질에 크게 의존한다는 공통된 한계를 갖는다. 실제 제조 현장에서 다루는 금속 부품은 표면의 강한 반사로 인해 능동형 깊이 센서가 투사하는 적외선 패턴을 교란시키며, 그 결과 깊이 영상에 홀(hole), 경계 왜곡, 표면 잡음과 같은 심각한 노이즈가 발생한다[13]. 이러한 노이즈는 파지 후보 영역의 형상 정보를 훼손하여 학습 기반 파지의 신뢰성을 저하시키는 주요 원인이 된다[4, 8].

또한 학습 기반 파지는 대량의 라벨링된 깊이 데이터를 필요로 하지만, 실제 환경에서 노이즈가 포함된 실측 데이터를 사람이 라벨링하여 대규모로 생성하는 것은 비용·시간 측면에서 비효율적이다. 이를 보완하기 위해 시뮬레이션 기반 합성 데이터가 널리 활용된다. 기존의 합성 데이터 기반 파지 연구는 이상적인 깊이 영상에 통계적 센서 노이즈를 주입하여 sim-to-real 격차를 줄이고자 하였다[4]. 그러나 이러한 노이즈 모델은 일반적인 깊이 센서의 특성을 단순화하여 모사한 것으로, 금속 표면의 반사로 인해 발생하는 홀과 심한 형상 왜곡까지 충실히 재현하기는 어렵다. 따라서 금속 부품을 안정적으로 파지하기 위해서는 실제 센서가 금속 표면에서 겪는 노이즈를 보다 정확하게 반영하는 물리 기반 노이즈 모델이 요구된다.

한편 다수의 기존 파지 방식은 고정된 카메라 광축에 정렬된 2차원 평면상의 수직 파지(top-down)를 가정하므로, 벽면에 인접하거나 기울여야만 그리퍼가 접근할 수 있는 물체에 대해서는 공간적 제약이 존재한다. 앞서 언급한 6자유도 파지는 이를 해소할 수 있으나 일반적으로 높은 계산 비용과 대규모 학습 데이터를 요구한다[10, 11].

본 연구에서는 금속 표면의 노이즈와 수직 파지의 공간적 제약이라는 두 가지 문제를 동시에 해결하기 위한 5자유도 파지 프레임워크를 제안한다. 구체적으로 (1) 기존 물리 시뮬레이터(SAPIEN)의 능동형 깊이 센서 모델과 Landau 등의 적외선·깊이 노이즈 모델을 활용하여 금속 표면의 반사 노이즈가 반영된 합성 깊이 데이터셋을 구성하고[14, 15], (2) 이를 GQCNN 학습에 활용하였으며[4], (3) clean depth, 일반 Gaussian noise, 제안한 물리 기반 노이즈의 세 학습 조건을 동일한 네트워크와 실제 로봇 실험 조건에서 비교하여 노이즈 모델의 독립적 기여를 정량적으로 평가하고, (4) 렌치 공간(wrench space)에 근거한 기울임(tilt)을 도입하여 기존 파지 데이터를 추가 비용 없이 5자유도(5-DOF) 파지 데이터로 확장하였다. 실제 금속·비금속 10종 물체에 대한 제안 방법의 전체 실험에서는 90.4%의 파지 성공률을 달성하였고, 수직 접근이 불가능한 벽면 인접 물체까지 안정적으로 파지하였다.

2. 합성 데이터 기반 5자유도 파지 알고리즘

2.1 합성 파지 데이터셋 생성

본 연구에서는 금속 부품 파지 학습을 위한 합성 깊이 데이터셋을 구축하였다. 객체 형상으로는 대규모 객체 중심 파지 데이터셋인 GraspFactory에서 제공하는 3차원 메쉬를 활용하였다[16]. Fig. 1은 데이터셋 생성의 전체 파이프라인을 나타낸다. 데이터셋에서 무작위로 선택한 총 750종의 3차원 모델을 대상으로 각 물체에 대해 파지 후보 샘플링과 깊이 영상 렌더링을 수행하여 (깊이 패치, 파지 라벨) 쌍을 생성하였으며, 최종적으로 약 800만 장 규모의 데이터셋을 확보하였다.

파지 후보는 antipodal sampling을 통해 생성된다. 물체 표면에서 서로 마주보는 한 쌍의 접촉점을 선택하고, 해당 두 점이 그리퍼 폭 이내에 들어오는 후보만 선별한다. 그렇게 선택된 후보를 파지 중심 위치와 접근 각도를 포함하는 g(x, y, z, θ)로 표현한다. 생성된 후보에 대해서는 파지축을 중심으로 회전시켜 다양한 기울임 각도를 부여한다. 이때 기울임 각도의 범위는 학습의 용이성과 실제 환경의 제약을 고려하여 본 연구에서는 최대 기울임 각도를 30°로 설정하였다. 이후 그리퍼 접근각($-a^z \geq \cos(\varphi_{max})$) 및 충돌 조건에 따른 필터링을 적용하여, 충돌 없이 접근 가능한 유효 후보만을 유지한다.

깊이 영상은 물체의 기하학적 깊이를 먼저 렌더링한 뒤, 2.2절에서 기술하는 SAPIEN 센서 모델을 통해 실제 RealSense D435i 센서의 특성을 반영한 센서 모델링 깊이로 변환한다. 금속 부품 파지를 대상으로 하므로 모든 객체를 금속 재질로 가정하여, 금속 표면 특유의 반사 특성에 따른 노이즈와 홀이 깊이 영상에 반영되도록 하였다. 이후 각 파지 후보를 중심으로 정렬하여 잘라낸 깊이 패치를 추출하고, 해당 파지의 성공/실패 라벨과 함께 학습 데이터로 저장한다. 이와 같이 노이즈가 반영된 패치 단위의 데이터를 사용함으로써, 네트워크가 금속 표면의 노이즈가 존재하는 조건에서도 안정적인 파지를 학습하도록 하였다.

2.2 물리 센서 노이즈 모델 기반 깊이 데이터

금속 부품의 사실적인 깊이 데이터를 획득하기 위해, SAPIEN 시뮬레이터에 내장된 능동형 깊이 센서(active depth sensor) 모델을 활용하였다. 해당 모듈은 좌·우 적외선 영상을 시뮬레이션한 뒤 스테레오 매칭을 통해 깊이를 복원하는 방식으로, 실제 RealSense D435i와 동일한 능동형 스테레오 원리를 모사한다. 여기에 Landau 등이 제안한 적외선·깊이 노이즈 모델을 함께 적용함으로써[14], 센서 고유의 양자화·정합 오차와 표면 반사 특성에 따른 노이즈가 깊이 영상에 자연스럽게 표현될 수 있다. 이를 통해 단순한 이상적 깊이 값이 아닌, 실제 금속 표면의 반사로 인해 생기는 물리적 노이즈가 합성 데이터에 반영된다.

Fig. 2는 동일한 금속 부품에 대한 깊이 영상을 네 가지 조건에서 비교한 것이다. (a)는 RGB 영상, (b)는 실제 RealSense D435i로 취득한 깊이 영상, (c)는 노이즈가 없는 기하학적 깊이 영상, (d)는 제안한 센서 노이즈 모델을 적용한 사실적 깊이 영상을 나타낸다. 실제 센서 영상 (b)에서는 금속 표면의 반사로 인해 물체 경계와 반사가 강한 영역의 깊이 값이 소실되고 표면이 불규칙하게 왜곡된다. 이상적인 기하학적 깊이 (c)는 이러한 열화를 포함하지 않는 반면, 제안 모델을 적용한 영상 (d)는 경계 부근의 홀과 표면 왜곡을 재현한다. 이러한 정성적 유사성은 합성 데이터와 실측 데이터 간의 분포 차이를 줄일 가능성을 보여주지만, 영상 비교만으로 실제 파지 강건성의 향상을 단정할 수는 없다. 따라서 물리 기반 노이즈 모델의 독립적 기여는 3.2절에서 clean depth 및 Gaussian noise 학습 모델과의 실제 파지 성공률 비교를 통해 정량적으로 검증한다.

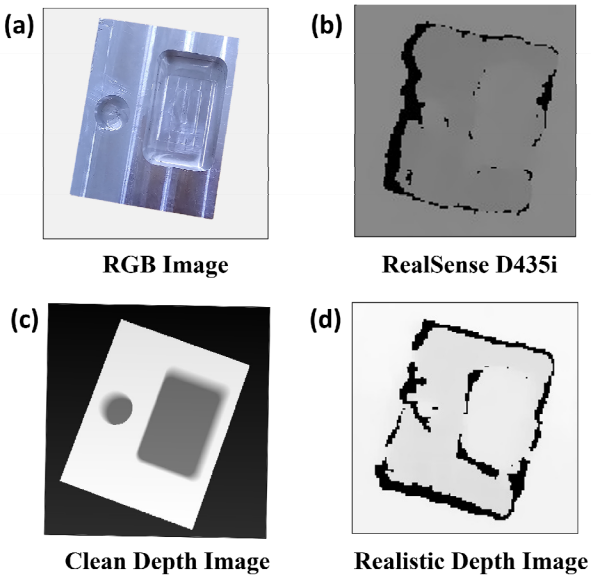


Fig. 2 Depth images under different rendering conditions

2.3 Wrench space 기반 기울임 증강과 5자유도 확장

기존의 수직 파지는 위치 (x, y, z)와 평면 회전(yaw)으로 정의되는 4자유도 파지에 해당한다. 이 경우 그리퍼 접근 방향이 한 방향으로 제한되기 때문에 기울여 접근해야 하는 물체나 벽면에 인접한 물체와 같이 수직 방향으로는 손가락이 충분히 진입할 수 없는 상황에서 파지가 제한된다. 본 연구에서는 이러한 제약을 완화하기 위해 파지축(grasp axis)을 중심으로 그리퍼를 기울이는 기울임(tilt)을 도입하여, 평면상의 위치와 회전에 더해 접근 각도까지 포함하는 5자유도로 파지 자유도를 확장하였다. 이를 통해 동일한 파지점에 대해서도 서로 다른 접근 각도를 선택할 수 있어, 그리퍼와 주변 환경 간의 충돌을 회피하면서 더 넓은 범위의 물체 자세에 대응할 수 있다. 얇은 띠 형상이나 금속 물체에서는 그리퍼를 기울여 접근하는 것이 충돌 회피를 넘어 파지 안정성을 높이는 수단이 될 수 있다.

이러한 기울임 증강의 타당성은 렌치 공간 분석으로 설명할 수 있다. 먼저 antipodal 조건은 양 손가락이 누르는 방향이 두 접촉점의 표면 법선과 거의 일직선을 이루도록 하는 제약으로, 평행 그리퍼가 마찰만으로 안정적인 force closure를 형성하기 위한 기하학적 요건에 해당한다. antipodal 파지에서 파지의 안정성은 두 접촉점 C1, C2의 위치와 접촉면 법선이 형성하는 마찰 원뿔(friction cone), 그리고 이들이 만들어내는 렌치 공간에 의해 결정된다.

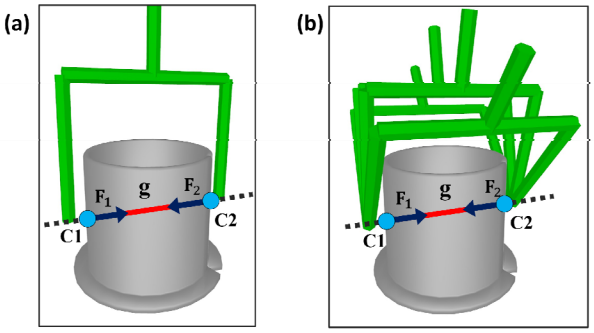


Fig. 3 Grasp axis-invariant tilt augmentation

Fig. 3에서 (a)는 수직 파지를, (b)는 동일한 접촉점 C1, C2를 유지한 채 파지축을 중심으로 그리퍼를 회전(tilt)시킨 파지들을 나타낸다. 그리퍼를 파지축 주위로 회전시키더라도 접촉점의 위치와 법선, 그리고 접촉력 F1, F2의 작용선이 변하지 않으므로, 이상적인 점 접촉을 가정했을 때 회전된 파지들은 원래 파지와 동일한 렌치 공간을 가지며 이에 따라 파지 품질(quality)이 보존된다.

해당 이론에 근거하여, 기존 파지 후보의 라벨을 재사용하면서 다양한 기울임 자세를 적은 비용으로 데이터셋에 증강할 수 있다. 새로운 기울임 자세에 대해 별도의 라벨링이나 파지 성공률 계산이 필요하지 않다. 본 연구에서는 각 파지 자세를 -90° ~ 90° 범위로 회전시켜 기울임 각을 확장하고 검증 실험에서는 0°, 15°, 30°의 세 각도 조건을 대상으로 성능을 평가하였다. 추론 단계에서는 각 기울임 후보에 대한 파지 품질을 평가하고 그중 가장 높은 자세를 선택함으로써, 경량 네트워크만으로 5자유도 파지를 수행한다.

2.4 GQCNN 기반 파지 품질 평가

파지 품질 평가에는 깊이 패치로부터 파지 성공 확률을 예측하는 grasp quality CNN (GQCNN)을 활용하였다[4]. Fig. 4와 같이 먼저 입력 깊이 영상에서 antipodal 조건을 만족하는 다수의 파지 후보를 샘플링한다. 이렇게 얻은 각 파지 후보를 중심으로 깊이

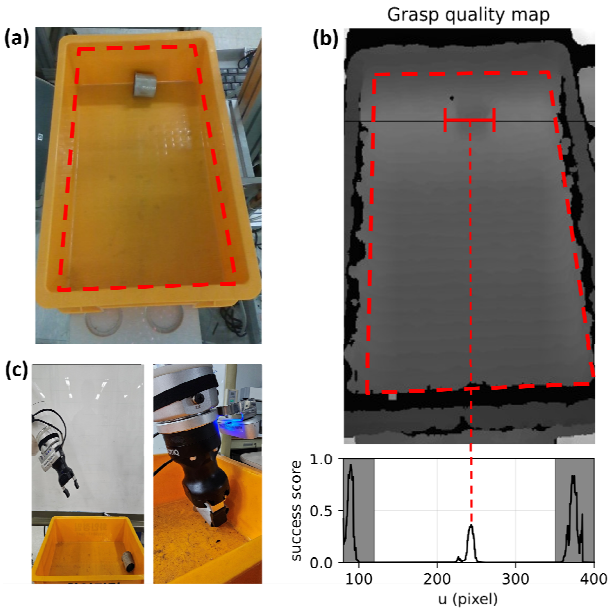


Fig. 5 Grasp-quality result of the trained network in flange

영상을 잘라내어 깊이 패치로 변환한다. 이렇게 정렬된 깊이 패치와 함께 파지 깊이(grasp depth)를 네트워크 입력으로 사용함으로써, 깊이 영상 상의 파지 위치와 회전에 무관하게 일관된 품질 평가가 가능하도록 하였다.

이러한 후보별 평가 결과를 영상 공간에 시각화하면 Fig. 5와 같이 나타낼 수 있다. Fig. 5(b)는 기울임 각도를 0°로 고정한 조건에서 입력 영상(Fig. 5(a)) 내 위치에 따라 파지 성공 확률이 어떻게 분포하는지를 나타내며, 점수가 가장 높은 위치가 최적 파지점으로 선택된다(Fig. 5(c)). 영상 내에서 파지 형상이 안정적인 영역에 높은 점수가 집중됨을 확인할 수 있다.

파지의 성공 여부는 주로 그리퍼의 양 손가락이 접촉하는 국소 영역의 형상에 의해 결정되므로, 평가 범위를 파지점 주변으로 한정하면 배경이나 원거리 영역의 노이즈에 영향을 받지 않고 접촉

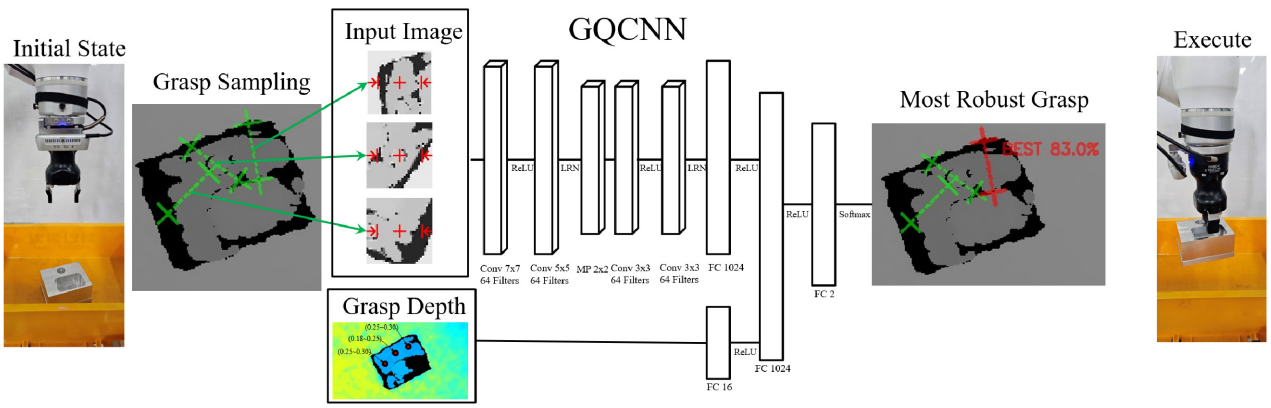


Fig. 4 GQCNN-based grasp quality evaluation

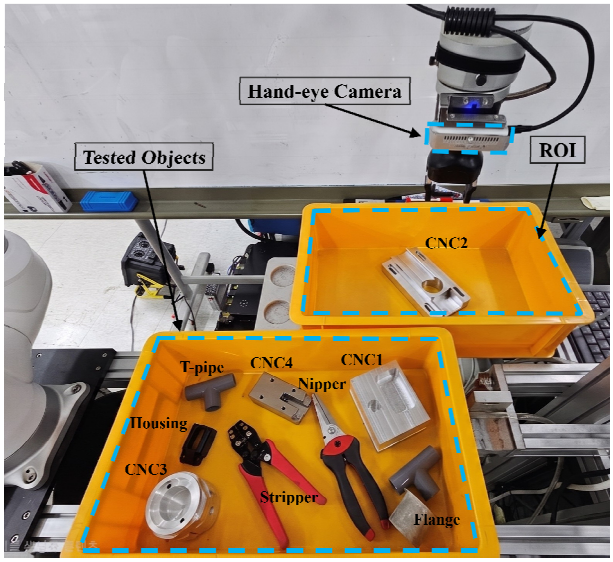


Fig. 6 Experimental setup and the tested objects

부의 기하 정보에 집중하여 품질을 평가할 수 있다. 또한 학습에 2.2절의 노이즈가 포함된 패치를 사용하므로, 네트워크는 홀과 왜곡이 나타나는 양상 자체를 파지 가능 영역을 가늠하는 형상 단서로 활용하여 안정적인 파지점을 식별하도록 학습한다.

제안 구조의 또 다른 장점은 계산 효율성에 있다. GQCNN은 비교적 얕은 합성곱 구조를 가지므로 다수의 파지 후보와 기울임 후보를 빠르게 일괄 평가할 수 있다. 이러한 경량성은 6자유도 파지 생성 기법이 일반적으로 요구하는 높은 계산 비용 없이도, 앞서 도입한 기울임 증강과 결합하여 5자유도 파지를 실용적인 속도로 수행할 수 있게 한다.

3. 실험 및 결과

3.1 실험 환경 및 방법

제안 방법의 성능을 실제 환경에서 검증하기 위해 Fig. 6과 같은 로봇 파지 시스템을 구성하였다. 로봇 말단부에 RealSense D435i 깊이 카메라를 장착하였고, 시험 물체로는 금속 6종(CNC1–CNC4, Flange, Stamp)과 비금속 4종(T-pipe, Nipper, Stripper, Housing)으로 총 10종을 선정하였다. 해당 물체는 실제 가공 및 조립 현장에서 빈번히 취급되는 부품들로 구성하였으며, 금속과 비금속은 깊이 센서에서 나타나는 노이즈 양상이 서로 다르므로 두 종류를 함께 포함하여 다양한 표면 특성과 형상에 대한 강건성을 평가하고자 하였다. 또한 모두 학습 데이터셋에 포함되지 않은 물체로 구성하여 처음 보는 물체에 대한 일반화 성능을 평가하였다.

실험은 별도의 조명 장치 없이 일반적인 실내 조명 환경에서 수행하였다. 시험 물체는 작업면에 고정된 상자 안에 두고 해당 상자 내부만을 region of interest (ROI)로 설정하였다. 각 시행에서는 물체를 상자 안에 다시 내려놓아 안정적인 정지 자세로 설정하고, 그 자세에서 상자 내에서의 위치와 회전을 임의로 바꾸어 가며 반복하여 특정 자세와 위치에 치우치지 않도록 하였다. 본 평가는 평행 그리퍼로 파지가 가능한 자세를 대상으로 수행하였다. 학습된 네트워크는 입력 깊이 영상의 픽셀 좌표마다 파지 성공 확률을 산출하며, 가장 높은 점수를 갖는 위치가 최적 파지점으로 선택된다. 파지 추론은 RTX 3070 Laptop GPU에서 수행하였다.

제안한 물리 기반 노이즈 모델의 효과를 분리하여 평가하기 위해, 동일한 GQCNN 구조를 사용하되 학습 깊이 데이터의 노이즈 조건만 달리한 세 모델을 각각 학습하였다. 첫 번째 모델은 노이즈가 없는 기하학적 깊이 영상(clean depth), 두 번째 모델은 clean depth의 유효 픽셀에 독립적인 가산성 Gaussian 교란 $\varepsilon \sim N(0, \sigma^2)$을 적용한 영상, 세 번째 모델은 SAPIEN active stereo와 Landau 노이즈 모델을 결합한 제안 영상으로 학습하였다. 세 조건에서 객체 메쉬, antipodal 파지 후보, 파지 라벨, 학습 표본 수, train/validation 분할, 네트워크 구조 및 학습 일정은 동일하게 유지하였다. Gaussian noise의 표준편차는 D435의 거리 의존적 깊이 노이즈 모델 $\sigma(z) = 0.004z^2$을 적용하여 설정하였다[17]. 여기서 z와 σ의 단위는 m이며, 깊이 영상 획득 시 카메라와 물체 사이의 거리가 0.5 m로 고정되어 있으므로 $\sigma = 0.001$ m(1.0 mm)를 사용하였다. 노이즈 학습 조건에 따른 차이를 기울임 효과와 분리하기 위해 비교실험은 0^o 접근으로 고정하고, 10종 물체에 대해 물체별 23회씩 동일한 무작위 재배치 절차로 수행하였다. 각 조건의 성공률은 성공 횟수와 전체 시행 횟수(n/N) 및 백분율로 제시하였다.

파지 실행 시 기울임 각도는 0^o, 15^o, 30^o의 세 조건으로 설정하였다. 0^o와 15^o 조건은 공간적 제약이 없는 일반적인 상황으로 시험 물체 10종 전체에 대해 물체별 23회씩, 30^o 조건은 후술하는 일부 물체에 한해 물체별 20회씩 반복 시행하여 성공률을 측정하였다. 파지 성공은 그리퍼가 접근 및 파지 과정에서 물체나 주변과 충돌하지 않고, 파지한 물체를 옆에 놓인 상자로 이송하여 내려놓을 때까지 미끄러짐이나 이탈 없이 파지 상태를 유지한 경우로 정의하였다.

30^o 조건의 대상은 크기가 작은 Flange, T-pipe, Housing으로, 이들을 상자 벽면에 밀착시킨 상태에서 수행하였다. 이때 벽면과의 거리와 방향은 시행마다 달리하여 다양한 정도의 제약 상황을 포괄하였다. 벽면에 밀착된 물체는 그리퍼가 수직으로 접근하면 손가락이 벽과 충돌하여 유효한 파지 자세가 존재하지 않고, 비스듬히 기울인 접근만이 충돌 없는 해를 가진다. 이처럼 수직 파지로는 해가 존재하지 않는 배치를 의도적으로 포함시켜, 기울임 기반 5자유도 확장이 파지 가능 영역을 실제로 넓히는지를 검증하고자 하였다.

3.2 파지 성능 평가 및 분석

제안 모델을 이용한 총 520회의 파지 시행에서 470회가 성공하여 전체 파지 성공률은 90.4%였고, 단일 시도당 GQCNN 평균 추론 시간은 0.3초였다. 재질별로는 금속 264/296회(89.2%), 비금속 206/224회(92.0%)로 2.8%p의 차이를 보였다. 이는 제안 모델이 광학 노이즈가 큰 금속 물체에서도 비금속 물체와 유사한 수준의 성능을 유지했음을 보여준다. 다만 물리 기반 노이즈 모델 자체의 기여는 전체 성공률만으로 판단하지 않고, 세 학습 조건의 통제된 비교 결과를 근거로 해석하였다.

Table 1은 학습 깊이 조건만 변경한 세 모델의 실제 파지 성공률을 비교한다. 제안 모델은 0° 조건에서 금속 물체 121/138회(87.7%), 비금속 물체 85/92회(92.4%), 전체 206/230회(89.6%)의 파지에 성공하였다. Clean depth 및 Gaussian noise 학습 모델의 결과를 동일한 형식으로 제시함으로써, 실제 환경에서 관찰되는 성능 차이가 네트워크 구조나 파지 라벨이 아니라 학습 깊이의 노이즈 모델에서 기인하는지를 평가한다.

Clean depth 모델은 금속 물체에서 97/138회(70.3%), 비금속 물체에서 78/92회(84.8%), 전체 175/230회(76.1%)의 파지에 성공하였다. Gaussian noise 모델은 각각 106/138회(76.8%), 82/92회(89.1%), 전체 188/230회(81.7%)였으며, 제안 모델은 각각 121/138회(87.7%), 85/92회(92.4%), 전체 206/230회(89.6%)였다. 제안 모델의 전체 성공률은 clean depth 모델보다 13.5%p, Gaussian noise 모델보다 7.8%p 높았다. 개선 폭은 금속 물체에서 각각 17.4%p 및 10.9%p로, 비금속 물체의 7.6%p 및 3.3%p보다 크게 나타났다. 이는 단순한 독립 Gaussian 교란보다 active stereo 정합 실패와 금속 반사에 따른 구조적 홀·경계 왜곡을 포함한 학습 데이터가 실제 D435i 입력에 대한 강건성 향상에 더 효과적이었음을 보여준다.

공간적 제약이 없는 동일 시험 조건에서 0°는 206/230회(89.6%), 15°는 209/230회(90.9%) 성공하여 1.3%p의 차이를 보였다. 전체 성공률의 차이가 작으므로 15°가 0°보다 우수하다고 단정하기보다는, 기울임 변화에도 성능이 안정적으로 유지된 것으로 해석하였다. 물체별로는 Flange가 19회에서 22회로 3회, Stamp가 20회에서 22회로 2회 증가하였고, CNC3와 Nipper는 각각 1회 증가하였다. Flange의 환형 형상과 Stamp의 낮고 얇은 형상에서는 15° 접근이 그리퍼 손가락의 삽입 여유와 가장자리 접촉 조건을 개선했을 가능성이 있으나, 해당 해석은 물체별 실패 유형이 완전히 분리되지 않은 상태에서 관찰된 경향이다. 반면 CNC2, CNC4, T-pipe 및 Housing은 각각 1회 감소했으며, CNC1과 Stripper는 두 각도에서 동일한 성공 횟수를 보였다.

Table 2는 물체별·기울임 각도별 파지 성공률을 정리한 것이다. 0° 및 15°의 대부분 조건에서 85% 내외 이상의 성공률을 유지하였다. 전체 50회의 실패는 충돌 24회(48%), 빈 파지 21회(42%), 미끄러짐 5회(10%)로 분류되었다. 충돌과 빈 파지가 실패의 90%를 차지한 반면 미끄러짐은 제한적이었으므로, 주요 한계는 파지 후 접촉 안정성보다는 좁은 접근 간격, 반사 노이즈로 인한 깊이 홀, 그리고 그에 따른 파지 위치 오차에 있었다. 다만 현재 실패 분류는 전체 시행을 합산한 값이므로 개별 물체의 0°~15° 증감 원인을 특정 실패 유형과 직접 연결하지 않았으며, 각도별 차이가 1회 수준인 물체에서는 우연 변동 가능성을 함께 고려하였다.

30° 조건은 0° 및 15°와 달리 Flange, T-pipe 및 Housing을 상자 벽면에 밀착시킨 별도의 공간 제약 실험이므로, 각도 증가에 따른 성공률 우열을 직접 비교하기 위한 조건이 아니다. 총 60회 중 55회가 성공하여 91.7%의 성공률을 보였으며, 물체별로는 Flange 18/20회(90%), T-pipe 19/20회(95%), Housing 18/20회(90%)였다. 이 배치에서는 수직 접근 시 그리퍼 손가락이 벽과 충돌하여 충돌 없는 파지 해가 존재하지 않았으나, 30° 기울임을 통해 접근 가능한 자세를 생성할 수 있었다. 따라서 이 결과는 30°가 일반 조건에서 더 높은 성공률을 보인다는 의미가 아니라, 기울임 기반 5자유도 확장이 수직 파지의 도달 가능 영역을 실제로 넓혔음을 보여주는 feasibility 검증으로 해석하였다.

Table 1 Grasp success rate by depth condition (0°)

Training depth condition	Metallic n/N (%)	Non-metallic n/N (%)	Overall n/N (%)
Clean depth	97/138 (70.3)	78/92 (84.8)	175/230 (76.1)
Gaussian noise	106/138 (76.8)	82/92 (89.1)	188/230 (81.7)
Proposed physics-based	121/138 (87.7)	85/92 (92.4)	206/230 (89.6)

Table 2 Grasp success rate (%) by object and tilt angle for metallic and non-metallic objects

Tilt angle(°)	Metallic success rate (%)							Non-metallic success rate (%)				
	CNC1	CNC2	CNC3	CNC4	Flange	Stamp	Avg	T-pipe	Nipper	Stripper	Housing	Avg
0	87.0	95.7	87.0	87.0	82.6	87.0	87.7	95.7	91.3	95.7	87.0	92.4
15	87.0	91.3	91.3	82.6	95.7	95.7	90.6	91.3	95.7	95.7	82.6	91.3
30*	–	–	–	–	90.0	–	–	95.0	–	–	90.0	–

*30° trials were conducted only for small objects placed against a wall.

4. 결 론

본 연구에서는 금속 표면의 반사로 인한 깊이 노이즈와 수직 파지의 공간적 제약을 동시에 다루기 위해, active stereo 원리와 금속 반사에 따른 정합 실패를 반영한 합성 깊이 데이터를 구축하고 이를 GQCNN 학습에 활용하였다. 또한 렌치 공간 분석에 근거한 기울임 증강을 통해 추가적인 라벨링이나 파지 품질 재계산 없이 파지 자유도를 5자유도로 확장하였다. 물리 기반 노이즈 모델의 독립적 효과는 clean depth 및 Gaussian noise 학습 모델과 네트워크·라벨·시험 조건을 통제한 실제 파지 비교로 평가하였다.

0° 통제 비교에서 clean depth 모델, Gaussian noise 모델 및 제안 모델의 전체 성공률은 각각 175/230회(76.1%), 188/230회(81.7%) 및 206/230회(89.6%)였다. 제안 모델은 두 비교군보다 각각 13.5%p 및 7.8%p 높았다. 제안 모델의 전체 520회 실험 성공률은 90.4%였고, 금속과 비금속 성공률은 각각 89.2%와 92.0%였다. 0°와 15°의 전체 성공률 차이는 1.3%p로 작아 기울임 변화에도 성능이 유지되는 것으로 해석하였다. 한편 벽면 인접 물체의 30° 실험에서는 55/60회(91.7%) 성공하여, 수직 접근으로 해가 존재하지 않는 배치에서 5자유도 확장의 실효성을 확인하였다.

제안한 방법은 경량 네트워크 구조를 기반으로 평균 0.3초의 추론 속도와 높은 실제 파지 성공률을 동시에 확보하므로, 빈피킹과 같이 반사성 부품과 공간 제약이 함께 존재하는 제조 환경에 적용될 수 있다. 다만 현재 모델은 이산화된 기울임 각도를 명시적으로 선택하며, 물체별·각도별 실패 원인의 세부 분리가 제한적이다. 향후에는 다중 작업 학습(multi-task learning)을 통해 최적 접근 각도와 충돌 확률을 단일 네트워크에서 추론하고, 복수의 학습 seed와 추가 물체를 이용하여 노이즈 모델 비교의 재현성을 확장할 계획이다.

후 기

이 연구는 산업통상부 소재부품기술개발사업(No.20026387)의 지원을 받아 수행되었음.

References

[1] Lenz, I., Lee, H., Saxena, A., 2015, Deep Learning for Detecting Robotic Grasps, Int. J. Robot. Res., 34:4-5 705-724, https://doi.org/10.1177/0278364914549607.

[2] Bohg, J., Morales, A., Asfour, T., Kragic, D., 2014, Data-Driven Grasp Synthesis—A Survey, IEEE Trans. Robot., 30:2 289-309, https://doi.org/10.1109/TRO.2013.2289018.

[3] Mahler, J., Pokorny, F. T., Hou, B., Roderick, M., Laskey, M., Aubry, M., Kohlhoff, K., Kröger, T., Kuffner, J., Goldberg, K., 2016, Dex-Net 1.0: A Cloud-Based Network of 3D Objects for Robust Grasp Planning Using a Multi-Armed Bandit Model with Correlated Rewards, Proc. IEEE Int. Conf. Robotics and Automation (ICRA), 1957-1964, https://doi.org/10.1109/ICRA.2016.7487342.

[4] Mahler, J., Liang, J., Niyaz, S., Laskey, M., Doan, R., Liu, X., Aparicio Ojea, J., Goldberg, K., 2017, Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics, Proc. Robotics: Science and Systems (RSS) XIII, 58, https://doi.org/10.15607/RSS.2017.XIII.058.

[5] Mahler, J., Goldberg, K., 2017, Learning Deep Policies for Robot Bin Picking by Simulating Robust Grasping Sequences, Proc. Conf. Robot Learning (CoRL), 515-524, <https://proceedings.mlr.press/v78/mahler17a.html>.

[6] Mahler, J., Matl, M., Liu, X., Li, A., Gealy, D., Goldberg, K., 2018, Dex-Net 3.0: Computing Robust Vacuum Suction Grasp Targets in Point Clouds Using a New Analytic Model and Deep Learning, Proc. IEEE Int. Conf. Robotics and Automation (ICRA), 5620-5627, 5620-5627, https://doi.org/10.1109/ICRA.2018.8460887.

[7] Mahler, J., Matl, M., Satish, V., Danielczuk, M., DeRose, B., McKinley, S., Goldberg, K., 2019, Learning Ambidextrous Robot Grasping Policies, Sci. Robot., 4:26 eaau4984, https://doi.org/10.1126/scirobotics.aau4984.

[8] Morrison, D., Corke, P., Leitner, J., 2018, Closing the Loop for Robotic Grasping: A Real-time, Generative Grasp Synthesis Approach, Proc. Robotics: Science and Systems (RSS) XIV, 21, https://doi.org/10.15607/RSS.2018.XIV.021.

[9] Liang, H., Ma, X., Li, S., Görner, M., Tang, S., Fang, B., Sun, F., Zhang, J., 2019, PointNetGPD: Detecting Grasp Configurations from Point Sets, Proc. IEEE Int. Conf. Robotics and Automation (ICRA), 3629-3635, https://doi.org/10.1109/ICRA.2019.8794435.

[10] Sundermeyer, M., Mousavian, A., Triebel, R., Fox, D., 2021, Contact-GraspNet: Efficient 6-DoF Grasp Generation in Cluttered Scenes, Proc. IEEE Int. Conf. Robotics and Automation (ICRA), 13438-13444, https://doi.org/10.1109/ICRA48506.2021.9561877.

[11] Fang, H.-S., Wang, C., Fang, H., Gou, M., Liu, J., Yan, H., Liu,

W., Xie, Y., Lu, C., 2023, AnyGrasp: Robust and Efficient Grasp Perception in Spatial and Temporal Domains, IEEE Trans. Robot., 39:5 3929-3945, https://doi.org/10.1109/TRO.2023.3281153.

[12] Zhang, J., Liu, H., Li, D., Yu, X., Geng, H., Ding, Y., Chen, J., Wang, H., 2024, DexGraspNet 2.0: Learning Generative Dexterous Grasping in Large-scale Synthetic Cluttered Scenes, Proc. Conf. Robot Learning (CoRL)., https://doi.org/10.48550/arXiv.2410.23004.

[13] Dai, Q., Zhang, J., Li, Q., Wu, T., Dong, H., Liu, Z., Tan, P., Wang, H., 2022, Domain Randomization-Enhanced Depth Simulation and Restoration for Perceiving and Grasping Specular and Transparent Objects, Proc. European Conf. Computer Vision (ECCV), LNCS, 13699 374-391, https://doi.org/10.1007/978-3-031-19842-7_22.

[14] Xiang, F., Qin, Y., Mo, K., Xia, Y., Zhu, H., Liu, F., Liu, M., Jiang, H., Yuan, Y., Wang, H., Yi, L., Chang, A. X., Guibas, L. J., Su, H., 2020, SAPIEN: A Simulated Part-based Interactive Environment, Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 11097-11107, https://doi.org/10.1109/CVPR42600.2020.01111.

[15] Landau, M. J., Choo, B. Y., Beling, P. A., 2016, Simulating Kinect Infrared and Depth Images, IEEE Trans. Cybern., 46:12 3018-3031, https://doi.org/10.1109/TCYB.2015.2494877.

[16] Srinivas, S. K., Shukla, Y., Arnold, A., Chitta, S., 2025, GraspFactory: A Large Object-Centric Grasping Dataset, arXiv:2509.20550, https://doi.org/10.48550/arXiv.2509.20550.

[17] Ahn, M. S., Chae, H., Noh, D., Nam, H., Hong, D., 2019, Analysis and Noise Modeling of the Intel RealSense D435 for Mobile Robots, Proc. 16th Int. Conf. Ubiquitous Robots (UR), 707-711, https://doi.org/10.1109/URAI.2019.8768489.



Minsoo Oh
Student Researcher, Autonomous Manufacturing & Process Research Division, Korea Institute of Industrial Technology (KITECH). His research interests include robotic grasping and manipulation.
E-mail: minsoo5849@gmail.com



Jaehak Lee
Principal Researcher, Autonomous Manufacturing & Process Research Division, Korea Institute of Industrial Technology (KITECH). His research interests include machine tool and automation.
E-mail: ljh1125@kitech.re.kr