

인간형 로봇 손 제어를 위한 손동작 모방 및 합성 알고리즘

Gesture Imitation and Composition Algorithm for Anthropomorphic Robot Hand Control

선 경 민¹, 신 상 훈¹, 양 석¹, 이 수 영^{*1}

(Kyung-Min Sun¹, Sang-Hun Sin¹, Seok Yang¹, and Soo-Yeong Yi^{1,*})

¹Department of Electrical and Information Engineering, Seoul National University of Science and Technology

Abstract: Controlling anthropomorphic robot hands with multiple joints is challenging due to their high degrees of freedom, which makes it difficult to provide effective demonstrations for dexterous manipulation. Moreover, existing approaches often require large training datasets or suffer from computational delays due to intensive calculations. This paper presents a study on imitation learning for the control of a prosthetic robot hand with 21 degrees of freedom, including wrist joints. By adopting a nearest-neighbor-based algorithm, the proposed control system achieves efficient learning with a small amount of data and provides an intuitive demonstration interface. Experimental results for three different tasks using the robot hand are also presented.

Keywords: anthropomorphic robot hand, k-nearest neighbors algorithm, gesture imitation

1. 서론

사람의 손은 높은 자유도를 가진 매우 정교한 기관으로, 다양한 작업을 수행할 수 있게 하며 인간 지능 발달의 핵심적인 역할을 수행해 왔다. 그러나 오늘날 산업 현장이나 일상 환경에서 활용되는 대부분의 로봇은 단순한 그리퍼(Gripper)를 장착하고 있어, 수행할 수 있는 작업이 매우 제한적이다. 로봇이 고도화된 인간의 작업을 대체하고 사람처럼 유연하게 상호작용을 하기 위해서 인간의 손을 모방한 다관절 로봇손의 필요성이 대두되고 있으며, 관련 연구 또한 주목받고 있다. 하지만 다관절 로봇손은 자유도가 매우 높아 정교한 조작을 할 수 있도록 교시하는 데에 어려움이 따른다. 이에 따라 로봇의 모든 관절 움직임을 일일이 정의하는 전통적인 프로그래밍 방식은 막대한 시간과 노력을 소모하게 되며, 고정된 환경을 넘어 다양한 환경에서 작동하기가 어렵다는 문제가 있다. 강화 학습(Reinforcement learning)이 이러한 문제의 대안으로 떠올랐으나, 로봇손의 높은 자유도는 탐색 공간을 방대하게 하기 때문에 표본 효율성이 낮고, 학습에 과도한 시간과 노력이 소요되는 한계점이 있다. 이러한 한계를 극복하는 방안으로, 연구자들은 모방학습(Imitation learning)에 주목하기 시작했다. 모방학습은 전문가의 시범 데이터를 관찰하고 그 행동을 흉내 내며 지식을 습득하는 방식으로, 사람이 직접 시범을 보이며 데이터를 제공함으로써 복잡한 탐색 과정을 대폭 단축하고, 수많은 시행착오를 거쳐야 하는 순수 강화 학습의 비효율성을 개선할 수 있다는 장점이 있다[1].

다관절 로봇손을 위한 모방학습 연구는 대표적으로 행동 복제(Behavior cloning)와 강화 학습을 결합한 방식으로 구분된다. 먼저, 행동 복제는 전문가의 시범 데이터에서 관찰된 상태와 행동을 짝지어 지도 학습(Supervised learning)하는 방식으로, 관련 연구가 다수 진행되었다[2,3]. 이 방식은 주어진 상태에 대해 전문가의 행동을 재연할 확률을 최대화하도록 정책을 학습한다. 그러나 행동 복제는 학습 데이터가 부족할 경우, 시범 데이터에서 관찰되지 않은 새로운 상황을 마주했을 때 효과적으로 대응하지 못하여 작업 성공률이 저하되는 한계를 지닌다[4]. 한편, 모방학습과 강화 학습을 결합하는 연구 또한 활발히 진행 중이다. 이 방식은 모방학습을 초기 가이드로 활용하여 전문가 시범을 통해 정책의 방향성을 설정하고, 이후 강화 학습을 통해 정책을 조정함으로써 불필요한 탐색 시간을 줄이고 학습 효율을 높인다. 대표적으로 DAPG (Demonstration Augmented Policy Gradient)와 같은 알고리즘을 기반으로 한 연구가 다수 수행되었다[5,6,7,8]. 하지만 이러한 접근법 역시 여전히 상당한 양의 시뮬레이션 기반 학습을 필요로 하며, 시뮬레이션에서 학습한 정책을 현실 세계에 적용하기 어렵다는 문제를 지닌다. 이는 접촉력, 마찰, 물체 변형과 같이 시뮬레이션에서 구현되는 물리적 특성과 현실 물리 환경 간의 차이에서 비롯되며, 이로 인해 시뮬레이션에서 최적화된 정책이 현실에서는 실패하는 경우가 빈번하다[9]. 이러한 시뮬레이션과 실제 상황 사이의 간극을 줄이기 위한 연구가 다각도로 진행되고 있으나[17,18], 정밀한 다

* Corresponding Author

Manuscript received December 4, 2025; revised January 13, 2026; accepted January 27, 2026

선경민: 서울과학기술대학교 전기정보공학과 학부생(sunm1229@seoultech.ac.kr, ORCID[®] 0009-0001-9754-5968)

신상훈: 서울과학기술대학교 전기정보공학과 학부생(mlight1423@naver.com, ORCID[®] 0009-0001-9172-4797)

양석: 서울과학기술대학교 전기정보공학과 학부생(rhscnd24@seoultech.ac.kr, ORCID[®] 0009-0001-2237-6178)

이수영: 서울과학기술대학교 전기정보공학과 교수(suylee@seoultech.ac.kr, ORCID[®] 0000-0001-8110-1468)

※ 상기 논문은 2025년도 ICROS-KROS 호남제주 그랜드 합동 학술대회에서 초안이 발표되었음[18].

※ 이 연구는 서울과학기술대학교 교내연구비의 지원으로 수행되었습니다.

관절 로봇 손 제어에 있어 그 격차를 온전히 해소하는 데에는 여전히 한계가 있다. 또한, 근본적으로 다관절 로봇손의 높은 자유도는 학습 효율을 떨어뜨리고, 최적의 동작을 계획하는 것을 어렵게 만든다.

강화 학습은 막대한 양의 데이터가 필요하므로, 실제 로봇을 이용해 직접 훈련하는 것은 비효율적일 뿐만 아니라, 학습 과정에서 예측 불가능한 행동으로 인해 로봇손이 파손될 위험이 있다. 현실 세계에서 더 적은 데이터로 효율적인 학습을 달성하기 위해, [4]에서는 '최근접 이웃(k-NN, k-Nearest Neighbors)' 방식을 활용하는 비모수(Non-parametric) 접근법을 시도하였다. 이 방식은 별도의 학습 과정 없이, 전문가 시연 데이터베이스 내에서 현재 관측된 상황과 가장 가까운 k개의 이웃 데이터를 찾아 그 행동들의 평균을 취해 로봇의 다음 행동을 결정한다. '최근접 이웃' 방식은 학습 데이터가 적은 상황에서도 강점을 보이며, 행동 복제와 달리 훈련 데이터에서 보지 못한 상황을 마주하더라도 가장 유사한 상황을 참조하여 작업 수행이 가능하다는 장점이 있다[10,11]. 실제로 [4]에서는 이 접근법을 통해 물체를 뒤집고(Flipping), 돌리고(Spinning), 회전시키는(Rotating) 등의 작업을 성공적으로 수행할 수 있음을 입증하였다. 하지만 [4]의 연구에서 사용한 Allegro Hand는 손가락이 4개이고 부피가 커, 5개 손가락을 가진 사람의 손동작을 로봇손의 움직임으로 변환하는 과정이 직관적이지 않다는 한계가 있다. 이는 조작자가 자신의 손 움직임과 로봇손의 실제 움직임 간의 괴리로 인해, 로봇손의 동작을 예측하면서 조종해야 하는 부담을 초래한다. 또한, [4]에서는 손가락 끝의 좌표를 기준으로 역기구학(Inverse kinematics)을 계산하여 로봇손의 관절 각도를 제어하는데, 이러한 방식은 계산 지연을 유발할 수 있다. 특히 정교한 손 조작은 종종 로봇의 특이점(singularity) 근처에서 이루어지는데, 자유도가 높아질수록 역기구학은 이러한 특이점 상황에서 불안정해지거나 실패할 가능성이 있다[12,13].

본 논문에서는 이러한 문제점들을 해결하기 위해, [14]의 케이블 구동형 로봇손을 기반으로 '최근접 이웃' 방법론을 적용한 모방학습 시스템을 구축한다. 해당 로봇손은 사람의 손을 모방하여 각 손가락 마디의 길이와 관절 가동 범위가 설계되었기 때문에, 사람 손의 움직임을 로봇손의 움직임으로 직관적으로 변환할 수 있다. 이는 조작자가 시연 과정에서 느끼는 인지적 부담을 줄여준다. 또한, 역기구학과 같은 복잡한 계산 과정 없이, 1:1 관절 각도 매핑을 통해 관절 공간(Joint space)에서 로봇을 제어하는 시연 시스템을 구현하였다. 본 연구는 이러한 직관적인 시연 시스템을 통해 비전문가라도 쉽고 빠르게 시연 데이터베이스를 구축하고 '최근접 이웃' 기반의 모방학습을 적용할 수 있는 환경을 구축하는 것을 목표로 한다. 나아가 [14]의 로봇손에 손목 관절을 추가하여, 더욱 다양한 동작을 수행할 수 있도록 구조를 확장하였다.

본 논문의 구성은 다음과 같다. 2장에서는 본 연구에 사용된 케이블 구동형 로봇손과 추가된 손목 관절을 소개한다. 3장에서는 MediaPipe Hands[15]를 기반으로 구축한 직관적이고 효율적인 시연 데이터 수집 시스템에 대해 설명한다. 4장에서는 본 연구에 적용된 '최근접 이웃' 기반의 알고리즘과 그 적용 방법에 대해 기술한다. 5장에서는 검증을 위해 정의된 태

스크와 실제 로봇을 이용한 실험 결과를 통합하여 제시하며, 마지막 6장에서는 결론으로 마무리한다.

II. 로봇 손 하드웨어 구성

2장에서는 제안하는 모방학습 시스템을 구축하기 위해 사용된 하드웨어 플랫폼에 관해 기술한다. 본 연구에서는 인간의 손과 유사한 기구학적 구조를 가진 [14]의 케이블 구동형 로봇손을 채택하였으며, 작업 반경과 동작의 다양성을 확장하기 위해 추가적인 손목 관절을 설계하여 통합하였다.

1. 케이블 구동형 로봇 손

본 연구에서 대상으로 하는 로봇손 플랫폼[14]은 각 관절에 연결된 양방향의 케이블 길이를 서로 당기거나 풀어주는 길항 구동(Antagonistic actuation) 방식을 통해 손가락의 움직임을 제어한다(그림 1). 하드웨어의 주요 구성 요소는 3D 프린팅 된 부품과 동력 전달을 위한 와이어(케이블)로 이루어져 있어, 제작 비용이 저렴하고 유지 보수가 용이하며 누구나 쉽게 복제하여 구현할 수 있다는 장점이 있다.

본 논문의 대상 로봇손 플랫폼은 무엇보다 인간 손의 해부학적 구조를 정밀하게 모방하여 설계되었다는 점에서 본 논문의 모방 학습 시스템 연구에 최적화되어 있다. 각 손가락 마디의 길이와 관절의 가동 범위는 실제 사람 손의 평균적인 치수와 운동 범위를 기반으로 설계되었다. 구체적으로 사람 손의 운동학적 구조와 동일하게, 중수 수지 관절(MCP)은 굽히기(Flexion) 및 펴기(Extension)와 벌리기(Abduction) 및 모으기(Adduction)가 가능한 2자유도(2-DOF)를 가지며, 근위 지간 관절(PIP)과 원위 지간 관절(DIP)에 해당하는 관절은 굽히기 및 펴기 동작을 수행하는 1자유도(1-DOF)를 갖는다.

이러한 기구학적 유사성은 다음과 같은 이점을 제공한다. 첫째, 사람 손의 움직임을 로봇손의 움직임으로 변환하는 과정이 매우 직관적이다. 본 로봇손은 별도의 변환 과정 없이 관절 공간에서의 1:1 매핑할 수 있으며, 이는 데이터 처리 과정을 단순화할 뿐만 아니라, 변환 과정에서 발생할 수 있는 동작의 왜곡을 최소화하여 정보를 전달할 수 있다. 둘째, 시연자가 느끼는 인지적 부담을 최소화한다. 시연자는 자신의 손동작이 로봇에게 어떻게 적용될지 즉각적으로 예측할 수 있으므로, 로봇의 기구학적 제약을 고려하며 조작해야 하는 어려움 없이 자연스러운 시연 데이터를 생성할 수 있다.



그림 1. 케이블 구동형 인간형 로봇손[14].

Fig. 1. Cable-driven anthropomorphic robot hand [14].

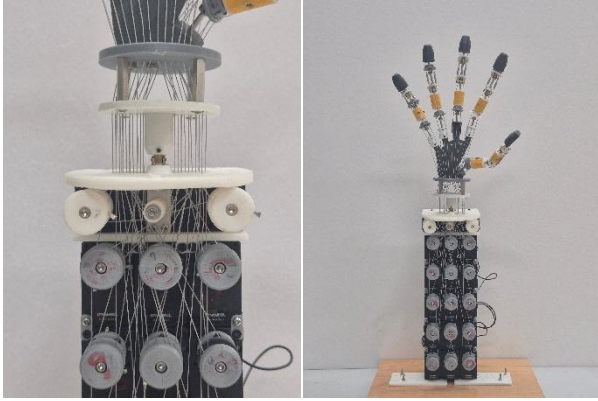


그림 2. 손목 관절과 결합된 로봇손의 설계 및 시제품.

Fig. 2. Design and prototype of the robotic hand integrated with a wrist joint.

2. 손목 관절

[14]에서 제안된 로봇손은 총 20자유도(20-DOF)를 가지며, 이를 제어하기 위해 15개의 구동모터를 사용한다. 다만 손목이 구현되어 있지 않아, 쥐기와 놓기 등의 파지 동작에는 문제가 없지만 파지 후 들어 올리는 동작을 포함해 다양한 동작 수행에 어려움이 있어, 본 연구에서 손목 관절을 추가하는 방향으로 수정, 제작하였다. 이를 위해 2개의 손목 관절 구동모터를 추가로 결합하였다.

사람의 손목 즉, 수근 골 관절(Intercarpal joints)은 2자유도(2-DOF)를 가지며 굽히기 및 펴기 동작과 엄지 방향으로 기울어지는 노쪽 치우침(Radial deviation) 및 소지 방향으로 기울어지는 자쪽 치우침(Ulnar deviation) 동작을 수행할 수 있다. 굽히기는 최대 90° , 펴기는 최대 70° 이고, 노쪽 치우침과 자쪽 치우침 동작은 각각 25° , 35° 까지 가동이 가능하다. 이러한 손목 관절의 동작은 손가락의 중수 수지 관절(MCP)과 유사하다. 차이점은 손목의 경우, 펴기 동작이 양각 70° 까지 가능하다는 것으로 펴기 동작이 최대 0° 까지인 손가락과 비교된다.

추가된 손목 관절의 형태를 그림 2에 보인다. 로봇손의 손가락 관절과 동일하게 보빈에 연결된 케이블의 길항구동으로 제어된다. 로봇손의 중수 수지 관절(MCP)과 유사한 형태를 가지며, 부양각이 각각 최대 60° 로 굽히기와 펴기에 해당하는 피치(Pitch) 동작이 가능하다. 피치 동작시에 손가락의 움직임에 가해지는 변형을 최소화하기 위해, 손가락을 제어하는 30개의 와이어가 손목의 회전축을 통과하도록 설계하였다. 그러나 치우침에 해당하는 요(Yaw) 동작시에는 전체적인 손가락의 위치에 변형이 가해진다. 이러한 전체적인 변형을 역기구학으로 푸는데 한계가 있기에, 손목 관절의 주요한 동작인 피치 동작만 가능하도록 제한하였다.

본 연구에서 사용된 부품들은 [14]와 동일하게 테이블탑 3D 프린터로 설계 및 제작하였다. 해당 3D 프린터는 용융적층 모델링(FDM) 방식이며, 수직 해상도가 0.1mm 인 직경 0.4mm 의 노즐을 사용하였다.

III. 시연 시스템 구축

3장에서는 케이블 구동형 로봇손을 제어하기 위한 직관적

인 시연 데이터 수집 시스템에 관해 기술한다. 본 연구의 핵심은 복잡한 재표적화(Retargeting) 과정 없이, 사람 손의 관절 각도를 로봇 손의 관절 각도로 1:1 매핑하여 제어하는 것이다. 이를 위해 단일 RGB 카메라만으로 손의 관절 정보를 정밀하게 추출할 수 있는 구글의 MediaPipe Hands [15]를 활용하였다.

1. MediaPipe Hands

MediaPipe Hands는 입력된 이미지로부터 손가락들의 21개 랜드마크(3D 좌표)를 실시간으로 추정한다. 21개의 랜드마크는 그림 3과 같이 손목을 시작점(0번)으로 하여, 각 손가락의 중수 수지 관절(MCP), 근위 지간 관절(PIP), 원위 지간 관절(DIP), 그리고 손가락 끝(TIP)에 해당하는 지점들로 구성된다. 이 모델은 각 관절의 x, y, z 좌표를 제공하므로, 본 연구에서는 이 좌표 간의 벡터 연산을 통해 로봇 제어에 필요한 관절 각도를 산출하였다.

2. 관절 각도 계산 알고리즘

제안하는 시연 시스템을 구현하기 위해서는, 손가락 관절들의 랜드마크 정보로부터 로봇 제어에 필요한 관절 각도를 도출하는 과정이 선행되어야 한다. 각 관절의 자유도 특성에 따라 그 계산 방식을 달리 적용해야 한다.

2.1 자유도 관절 (PIP, DIP)

근위 지간 관절(PIP)과 원위 지간 관절(DIP)은 굽히기 및 펴기 동작만을 수행하는 1자유도 관절이다. 이 경우, 세 개의 랜드마크가 이루는 사잇각을 계산하는 단순한 방식으로 각도를 구할 수 있다. 검지의 원위 지간 관절(DIP)을 예로 들면, 6번(PIP), 7번(DIP), 8번(TIP) 랜드마크를 이용한다. 7번 좌표를 기준으로 6번과 8번을 향하는 두 벡터(v_{76} , v_{78})를 생성하고, 두 벡터의 내적을 통해 사이 각도를 산출한다.

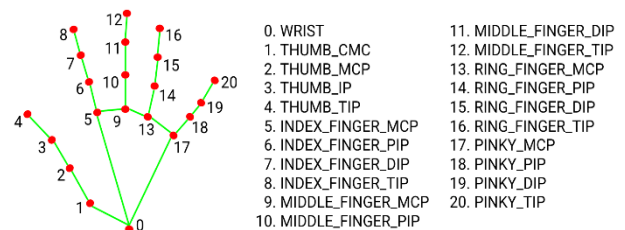


그림 3. MediaPipe Hands로 얻을 수 있는 손 관절 랜드마크.

Fig. 3. Hand landmarks provided by MediaPipe Hands.

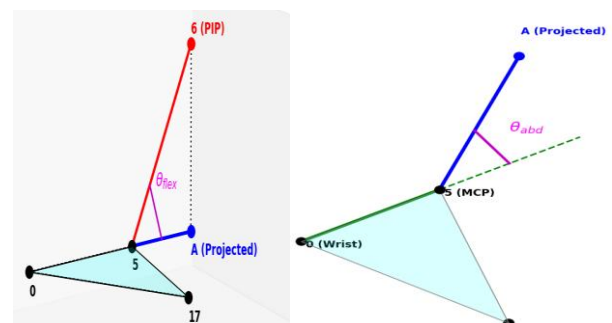


그림 4. 손바닥 평면 위의 사영을 사용한 각도 계산.

Fig. 4. Calculation of angle using vector projection on the palm plane.

2.2 자유도 관절(MCP)

중수 수지 관절(MCP)은 굽히기 및 펴기 동작뿐만 아니라 벌리기 및 모으기 동작까지 수행하는 2자유도 관절이다. 따라서 단순한 세 점 사이의 각도 계산만으로는 두 가지 회전 성분을 분리해 낼 수 없어 추가 계산 과정이 필요하다. 검지 손가락의 중수 수지 관절(MCP)을 기준으로 한 각도 추출 과정은 다음과 같다.

- 1) 손바닥 평면 정의: 먼저 손목(0), 검지 MCP (5), 소지 MCP (17)의 세 점을 잇는 삼각형을 기준으로 손바닥이 이루는 기준 평면을 정의한다.
- 2) 벡터 정사영: 검지의 방향을 나타내는 5-6번 벡터(v_{56})를 앞서 정의한 손바닥 평면 위로 정사영(Projection)시킨다.
- 3) 굽힘 각도 산출: 손바닥 평면 위에 투영된 점을 점 A라고 하자. 측면에서 손을 바라보았을 때, 실제 손가락 벡터(v_{56})와 손바닥 평면에 투영된 벡터(v_{5A}) 사이의 각도는 손가락이 손바닥 평면에서 얼마나 들어 올려졌는지를 나타낸다. 즉, 이 각도가 MCP 관절의 굽힘 각도가 된다.
- 4) 벌림 및 모음 각도 산출: 다시 손바닥을 정면에서 바라보았을 때, 손바닥 평면 위에서 투영된 벡터(v_{5A})와 손목에서 관절로 이어지는 기준 벡터(v_{05}) 사이의 각도를 측정한다. 이 각도는 손가락이 기준 축으로부터 좌우로 얼마나 회전했는지를 나타내므로, 벌림/모음 각도로 정의할 수 있다.

2.3 손목 관절

손목 관절의 굽힘 각도를 알기 위해서는 팔꿈치까지 확장된 랜드마크가 필요하지만, 새로운 인식 모델을 만들거나 전신 랜드마크 인식 모델을 추가하는 것은 본 연구 환경에서는 제한적이다. 따라서 손목을 굽히지 않은 상태인 굽힘 각도가 0도일 때, 손바닥이 카메라와 정면으로 마주보는 것을 조건으로 하여 다음의 방법을 적용한다.

9번 랜드마크를 x-y 평면에 평행하면서 0번 랜드마크가 포함된 평면에 사영한 점을 B라고 하자. MediaPipe Hands에서 제공하는 랜드마크의 z좌표 값은 손목(0번)의 랜드마크를 기준($z = 0$)으로 하므로 위 과정은 9번 랜드마크를 x-y 평면에 사영하는 것과 같다. 0번 랜드마크에서 점 B로 향하는 벡터(v_{0B}), 0번 랜드마크에서 9번 랜드마크로 향하는 벡터(v_{09})가 이루는 각을 손목의 굽힘 각도로 한다.

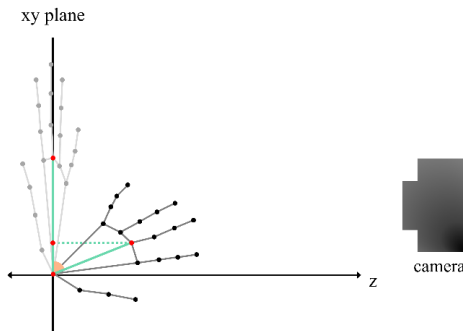


그림 5. 손목 관절의 굽힘 각도 계산.
Fig. 5. Calculation of wrist flexion angle.

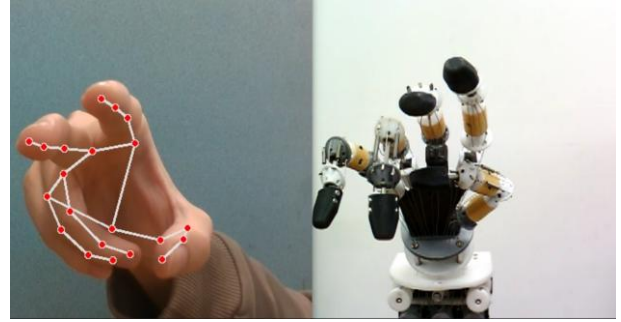


그림 6. 실시간 손동작 모방 시스템.
Fig. 6. Real-time hand gesture imitation system.

2.4 정리

이와 같은 기하학적 계산 과정을 통해, 3차원 공간상의 손가락 관절들의 랜드마크로부터 복잡한 역기구학 풀이 없이도 로봇 제어에 필요한 21자유도에 대한 각도 정보를 직관적이고 정확하게 산출할 수 있다. 사람 손동작 동영상 프레임마다 앞서 기술한 연산 과정을 실시간으로 수행하여, 산출된 손가락 관절 각도 정보를 로봇손에 전달하는 과정을 반복한다. 이러한 반복 과정을 통해 로봇손은 시연자의 손동작을 실시간으로 모사할 수 있으며, 최종적으로 본 시스템은 추후 모방 학습을 위한 데이터 세트를 생성하는 시연 시스템으로 활용된다.

IV. 선형 보간 기반 모방학습 알고리즘

4장에서는 3장에서 구축한 시연 시스템을 통해 수집한 사람 손동작 데이터를 바탕으로, 실제 로봇손을 제어하기 위한 모방 학습 알고리즘에 관해 기술한다. 본 연구에서는 [4]의 연구에 착안하여, 적은 양의 데이터로도 효율적인 학습이 가능한 비모수(non-parametric) 접근법인 '최근접 이웃' 방식을 채택하였다. 특히 데이터가 부재한 구간에서도 부드럽고 정확한 동작을 생성하기 위해 선형 보간 기법을 도입하였다.

1. 최근접 이웃 기반의 모방학습 시스템

'최근접 이웃(k-NN, k-Nearest Neighbors)' 알고리즘 기반의 모방학습은 별도의 학습 과정 없이 데이터베이스에 저장된 사람 손 시연 데이터 중 현재 로봇이 마주한 상황과 가장 유사한 k개의 이웃 데이터를 실시간으로 검색하여 다음 행동을 결정하는 방식이다. 이는 데이터 효율성이 높을 뿐만 아니라, 훈련 데이터에서 보지 못한 새로운 상황을 마주하더라도 가장 유사한 과거의 경험을 참조하여 즉각적인 대응이 가능하다는 장점이 있다[10,11].

2. 선형 보간을 이용한 동작 생성

[4]에서 활용한 k-NN 방식은 단순히 검색된 k개 이웃 행동의 평균을 내는 방식을 주로 사용한다. 그러나 이 방식은 데이터 포인트 사이의 간격이 넓은 경우, 동작이 불연속적이거나 정밀도가 떨어지는 한계가 있다. 본 연구에서는 이러한 문제를 개선하고, 데이터가 없는 구간에서도 효과적인 손동작을 생성하기 위해 선형 보간(linear interpolation) 기법을 적용하였다.

구체적으로, 새로운 상황 A_{new} 가 관측되면 데이터베이스 내에서 새로 관측된 상황과 가장 인접하면서도 작은 값을 가

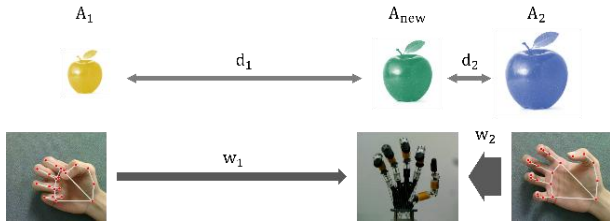


그림 7. 선형 보간법을 사용한 손동작 생성 알고리즘.

Fig. 7. Hand motion generation algorithm using linear interpolation.

진 데이터 A_1 과, 큰 값을 가진 데이터 A_2 를 선택한다. 이때 새로운 상황과 A_1 , A_2 와의 데이터 값 크기 차이를 각각 d_1 , d_2 라 정의한다.

로봇손의 다음 동작은 두 이웃 데이터의 행동에 가중치 w_1 , w_2 를 적용하여 생성된다. 이때 가중치는 각 데이터까지의 거리에 반비례하여 설정되며, 구체적으로는 d_1 에 대한 d_2 의 비율을 고려하여 다음과 같이 결정된다.

$$w_1 = \frac{d_2}{d_1 + d_2}, \quad w_2 = \frac{d_1}{d_1 + d_2} \quad (1)$$

현재 상황이 A_1 에 더 가까울수록 A_1 의 동작에 더 큰 가중치를 부여하여 자연스러운 동작 생성을 유도한다. 이와 같은 선형 보간 기반의 접근법은 단순히 이웃 데이터의 평균을 취하는 방식보다 데이터 부재 구간을 수학적으로 보정해 주므로, 시연 데이터가 충분하지 않은 상황에서도 더 부드럽고 정확한 손동작 생성을 가능하게 한다.

V. 실험 및 결과

5장에서는 4장에서 제안한 선형 보간 기반의 최근접 이웃 방식의 성능을 검증하기 위해, 정의된 세 가지 태스크에 관해 기술하고 그에 대한 실험 결과를 제시한다.

1. 태스크 구성

검증을 위해 정의된 태스크는 테이프 파지, 막대 균형 맞추기, 그리고 공 집어 올리기의 세 가지 작업이다.

1.1 다양한 크기의 테이프 파지

첫 번째 태스크는 다양한 크기의 원형 테이프를 안정적으로 파지하는 작업이다. 이 태스크에서 행동을 생성하는 기준은 테이프의 너비(width)와 높이(height)로 정의하였다. 테이프의 너비와 높이는 객체 검출 모델인 'YOLO11' [16]을 활용하여 추출된 테이프의 바운딩 박스 정보를 기반으로 획득하였다. 바운딩 박스의 픽셀 정보는 Depth 카메라의 깊이 정보를 활용하여 미터 단위 좌표계로 변환된다. 이러한 좌표 변환 과정을 통해, 카메라와 물체 사이의 거리가 변하더라도 일관된 물리적 크기를 검출하여 절대적인 기준으로 활용할 수 있도록 하였다.

먼저 그림 8과 같이 시연 시스템을 이용하여 다양한 크기의 테이프를 잡는 사람의 손동작 데이터를 기록하여 데이터 베이스를 구축하였다. 이후 실제 수행 단계에서는 학습 데이터에 포함되지 않은 새로운 크기의 테이프가 제시되었을 때, 그림 9와 같이 4장에서 제안한 최근접 이웃-선형보간 알고리즘을 통해 관측된 크기 정보에 맞는 파지 동작을 스스로 생성하여 수행하도록 하였다. 부록의 표 A-1은 인식된 테이프

의 크기에 따라 생성된 로봇 손의 관절 각도 데이터의 예시를 보여준다.

1.2 막대 균형 맞추기

두 번째 태스크는 기울어진 막대를 손가락을 이용해 균형을 맞추는 작업이다. 행동을 생성하는 기준은 막대의 기울어진 각도로 정의하였다. 이때, 막대의 형상적 특징으로 인해 바운딩 박스를 추출하기 어려워, 막대의 양 끝단에 ArUco 마커를 부착하는 방식을 채택하였다. 양 끝 마커의 중심 좌표를 검출하고, 두 중심점을 잇는 선분이 이루는 각도를 계산하여 기준으로 활용하였다.

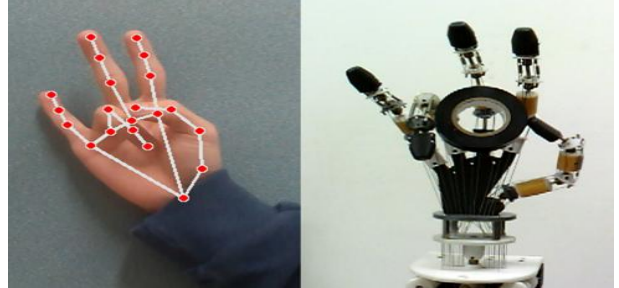


그림 8. 시연을 통한 행동 데이터 수집: 원형 테이프 파지.

Fig. 8. Data collection via demonstration system: circular tape grasping.



그림 9. k-최근접 이웃을 이용한 테이프 파지 작업 실험.

Fig. 9. Task execution using k-NN for tape grasping.



그림 10. 시연을 통한 행동 데이터 수집: 막대 균형 맞추기.

Fig. 10. Data collection via demonstration system: balancing a stick.



그림 11. k-최근접 이웃을 이용한 막대 균형 맞추기 작업 실험.

Fig. 11. Task execution using k-NN for stick balancing.

다양한 각도로 기울어진 막대에 대응하여 손가락을 움직여 균형을 맞추는 시연 데이터를 수집하였다(그림 10). 이후 임의의 각도로 기울어진 막대가 제시되었을 때, 선행 보간을 통해 가장 인접한 각도 데이터들의 행동을 조합하여 막대를 수평으로 맞추는 동작을 반복적으로 수행하도록 하였다. 부록의 표 A-2는 인식된 막대 각도에 따라 생성된 로봇 손의 관절 각도 데이터의 예시를 보여준다.

1.3. 높이별 공 집어 올리기

세 번째 태스크는 공의 높이에 따라 공을 집어 올리는 태스크로, 추가된 손목 관절의 자유도를 적극적으로 활용할 수 있는 작업이다. 행동을 생성하는 기준은 인식된 공의 높이 정보다. 높이는 YOLO에서 인식된 공의 바운딩 박스의 하단 y 좌표를 기준으로 한다. 일정한 거리에 있는 ArUco 마커를 카메라 화면 상에 인식시킨 뒤 이를 이용하여 카메라 화면 최하단과 바운딩 박스 하단 사이의 실제 거리를 구한다.

다양한 높이 위치의 공에 대응하여 손목과 손가락의 움직임을 수행하는 시연 데이터를 구축하였다. 이를 바탕으로 로봇은 임의의 높이로 인식된 공에 대해, 데이터베이스 내의 인접한 데이터들을 조합하여 손목과 손가락을 동시에 제어하여 성공적으로 공을 잡고 들어 올리는 동작을 수행하도록 하였다. 부록의 표 A-3은 인식된 공의 높이에 따라 손목 관절을 포함하여 생성된 로봇 손의 관절 각도 데이터의 예시를 보여준다.

2. 실험 결과 및 분석

각 태스크에 대해 총 10회 수행하였으며, 작업 성공 여부에 따른 성공률을 측정하였다. 실험 결과는 표 1과 같다. 테이프 파지 태스크는 70%의 성공률을 보였다. 실패 원인으로는 테이프의 폭이 좁아 로봇손의 가동 범위를 벗어나거나, 표면이 미끄러운 테이프를 놓치는 현상이 있었다. 막대 균형

표 1. 각 검증 작업별 성공률(10회 시도).

Table 1. Success rates for each task (10 trials).

	테이프 잡기	막대 균형	공 집어 올리기
성공률	70 %	100 %	70 %

유지 태스크는 100%의 성공률을 기록하였다. 10회의 모든 시도에서 로봇손은 막대의 기울어진 각도에 대응하여 안정적으로 균형을 맞추었다. 공 집어 올리기 태스크는 70%의 성공률을 보였다. 실패 원인으로는, 공을 들어 올리는 과정에서 로봇손의 악력이 부족하여 공을 놓치는 경우가 있었다.

종합적으로, 세 가지 태스크 모두에서 시연된 동작을 기반으로 새로운 상황에 대해 적절한 행동을 생성해 내었다. 이는 본 연구에서 제안한 모방 학습 시스템이 데이터가 부족한 상황에서도 선행 보간을 통해 성공적으로 동작을 수행할 수 있음을 실험적으로 검증하였다.

VI. 결론

본 연구는 케이블 구동형 로봇손과 MediaPipe Hands를 활용한 관절 각도 추출 방식을 연동함으로써, 복잡한 기구학·역기구학 과정 없이 관절 공간에서 1:1 매핑이 가능한 직관적인 시연 환경을 구축하였다. 이는 로봇 비전문가라도 다관절 인간형 로봇손을 쉽고 빠르게 제어할 수 있게 하여 데이터 수집의 난도를 크게 낮추었다는 데 의의가 있다. 또한 단순히 수집된 데이터를 재현하는 것을 넘어, 선행 보간을 결합한 k-NN 알고리즘을 적용함으로써 데이터 부재 구간에서도 주어진 상황에 맞게 유연하고 연속적인 동작을 생성할 수 있는 효율적인 모방학습 제어 시스템을 제시하였다. 제안된 시스템의 유효성은 세 가지 작업 수행 실험을 통해 검증되었다. 이러한 시스템은 로봇손 뿐만 아니라 로봇 팔이나 보행 로봇처럼 정교한 관절 제어가 요구되는 다양한 분야에서 활용 가능성이 있을 것으로 기대된다.

REFERENCES

- [1] S. An, Z. Meng, C. Tang, Y. Zhou, T. Liu, F. Ding, S. Zhang, Y. Mu, R. Song, W. Zhang, Z. Hou, and H. Zhang, "Dexterous manipulation through imitation learning: a survey," arXiv preprint arXiv: 2504.03515, Apr. 2025.
doi: <https://doi.org/10.48550/arXiv.2504.03515>
- [2] Y. Li, C. Guo, J. Ren, B. Chen, C. Cheng, H. Zhang, and H. Lu, "SoftGrasp: Adaptive grasping for dexterous hand based on multimodal imitation learning," *Biomimetic Intelligence and Robotics*, vol. 5, no. 2, Jun. 2025.
doi: <https://doi.org/10.1016/j.birob.2025.100217>
- [3] D. Liconti, Y. Toshimitsu, and R. Katzschmann, "Leveraging pretrained latent representations for few-shot imitation learning on a dexterous robotic hand," arXiv preprint arXiv: 2404.16483, Apr. 2024.
doi: <https://doi.org/10.48550/arXiv.2404.16483>
- [4] S. Arunachalam, S. Silwal, B. Evans, and L. Pinto, "Dexterous imitation made easy: a learning-based framework for efficient dexterous manipulation," *Proc. of 2023 IEEE International Conference on Robotics and Automation (ICRA)*, London, United Kingdom, pp. 5954-5961, May 2023.

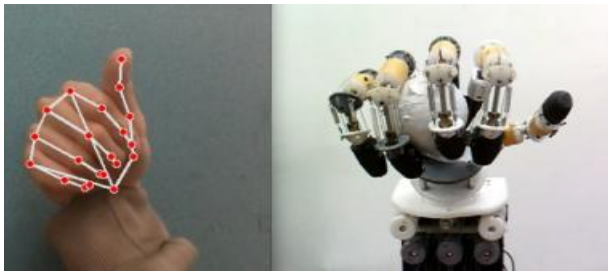


그림 12. 시연을 통한 행동 데이터 수집: 공 집어 올리기.
Fig. 12. Data collection via demonstration system: ball lifting.



그림 13. k-최근접 이웃을 이용한 공 집어 올리기 작업 실행.
Fig. 13. Task execution using k-NN for ball lifting.

- doi: <https://doi.org/10.1109/ICRA48891.2023.10160275>
- [5] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning complex dexterous manipulation with deep reinforcement learning and demonstrations," *Proc. of Robotics: Science and Systems (RSS)*, vol. 14, Pennsylvania, USA, Jun. 2018.
doi: <https://doi.org/10.15607/RSS.2018.XIV.049>
- [6] Y. Qin, H. Su, and X. Wang, "From one hand to multiple hands: imitation learning for dexterous manipulation from single-camera teleoperation," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10873-10881, Aug. 2022.
doi: <https://doi.org/10.1109/LRA.2022.3196104>
- [7] A. Prados, G. Espinoza, N. Fernandez, and R. Barber, "Imitation learning for low-cost dexterous hand using RGBD camera," *Proc. of 2024 7th Iberian Robotics Conference (ROBOT)*, Madrid, Spain, pp. 1-6, Nov. 2024.
doi: <https://doi.org/10.1109/ROBOT61475.2024.10796919>
- [8] Y. Qin, Y. Wu, S. Liu, H. Jiang, R. Yang, Y. Fu, and X. Wang, "DexMV: Imitation learning for dexterous manipulation from human videos," arXiv preprint arXiv: 2108.05877, Aug. 2021.
doi: <https://doi.org/10.48550/arXiv.2108.05877>
- [9] A. Weinberg, A. Shirizly, O. Azulay, and A. Sintov, "Survey of learning-based approaches for robotic in-hand manipulation," *Frontiers in Robotics and AI*, vol. 11, Nov. 2024.
doi: <https://doi.org/10.3389/frobt.2024.1455431>
- [10] S. Schaal and C.G. Atkeson, "Robot juggling: implementation of memory-based learning," *IEEE Control Systems Magazine*, vol. 14, no. 1, pp. 57-71, Feb. 1994.
doi: <https://doi.org/10.1109/37.257895>
- [11] J. Pari, N. Shafiullah, S. Arunachalam, and L. Pinto, "The Surprising Effectiveness of Representation Learning for Visual Imitation," arXiv preprint arXiv: 2112.01511, Dec. 2021.
doi: <https://doi.org/10.48550/arXiv.2112.01511>
- [12] T. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," *arxiv preprint arXiv: 2304.13705*, Apr. 2023.
doi: <https://doi.org/10.48550/arXiv.2304.13705>
- [13] P. Wu, Y. Shentu, Z. Yi, X. Lin, and P. Abbeel, "GELLO: A general, low-cost, and intuitive teleoperation framework for robot manipulators," *Proc. of 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Abu Dhabi, United Arab Emirates, pp. 12156-12163, Oct. 2024.
doi: <https://doi.org/10.1109/IROS58592.2024.10801581>
- [14] S. Min and S. Yi, "Development of cable-driven anthropomorphic robot hand," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1176-1183, Feb. 2021.
doi: <https://doi.org/10.1109/LRA.2021.3056375>
- [15] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C. Chang, and M. Grundmann, "MediaPipe Hands: On-device real-time hand tracking," arXiv preprint arXiv: 2006.10214, Jun. 2020.
doi: <https://doi.org/10.48550/arXiv.2006.10214>
- [16] R. Khanam and M. Hussain, "YOLOv11: An overview of the key architectural enhancements," arXiv preprint arXiv: 2410.17725, Oct. 2024.
doi: <https://doi.org/10.48550/arXiv.2410.17725>
- [17] D. Kang, J. Park, and P. Kim, "A study on bridging the gap with reinforcement learning: TD3 optimization for sim-to-real transfer from gazebo to isac sim," *Journal of Institute of Control,*

Robotics and Systems (in Korean), vol. 31, no. 11, pp. 1290-1298, Nov. 2025.

doi: <https://doi.org/10.5302/I.ICROS.2025.25.0127>

- [18] H. Shin, S. Yang, K. Sun, and S. Yi, "Prosthetic robot hand control via gesture imitation and composition," *Proc. of 2025 ICROS-KROS Honam Jeju Regional Conference (in Korean)*, 2025.



선 경 민

2026년 2월 서울과학기술대학교 전기정보공학과(공학사).



신 상 훈

2026년 2월 서울과학기술대학교 전기정보공학과(공학사).



양 석

2019년 3월~현재 서울과학기술대학교 전기정보공학과. 관심분야는 안테나.



이 수 영

1988년 2월 연세대학교 전자공학과(공학사). 1990년 2월 KAIST 전기 및 전자공학과(공학석사). 1994년 8월 KAIST 전기 및 전자공학과(공학박사). 1995년 3월~1999년 8월 KIST 시스템연구부 선임연구원. 1997년 2월~1998년 2월 Univ. of Southern California 박사 후 과정. 1999년 9월~2007년 2월 전북대학교 전자정보공학부 부교수. 2005년 6월~2006년 8월 Univ. of Illinois at Urbana-Champaign 방문교수. 2007년 3월~현재 서울과학기술대학교 전기정보공학과 교수. 관심분야는 보행로봇, 로봇비전, 이동로봇.

Appendix

표 A-1. 인식된 테이프 크기에 따라 생성된 관절각 데이터.

Table A-1. Generated joint angles based on detected tape size.

	시연(학습) 데이터					검증작업
	36x35 [mm]	59x58 [mm]	60x59 [mm]	76x78 [mm]	100x95 [mm]	69x70 [mm]
관절 각 1	20.7°	17.6°	53.1°	39.5°	-2.6°	45.3°
관절 각 2	-8.3°	-17.0°	-10.5°	-3.1°	-100.4°	-6.3°
관절 각 3	23.8°	9.6°	55.2°	58.3°	97.1°	57.0°
관절 각 4	29.2°	29.7°	57.4°	38.7°	94.2°	46.7°
관절 각 5	16.8°	-1.7°	58.4°	40.5°	-2.0°	48.2°
관절 각 6	1.0°	-7.3°	-24.4°	-25.0°	-15.6°	-24.7°
관절 각 7	-6.7°	-8.0°	77.3°	75.2°	56.8°	76.1°
관절 각 8	6.2°	5.4°	25.5°	53.7°	124.8°	41.6°
관절 각 9	20.0°	8.3°	60.7°	45.2°	22.5°	51.9°
관절 각 10	-8.6°	-6.7°	-23.2°	-32.3°	-19.6°	-28.4°
관절 각 11	4.2°	1.9°	79.0°	93.3°	95.5°	87.2°
관절 각 12	8.3°	12.5°	91.7°	59.0°	31.5°	73.1°
관절 각 13	45.1°	46.3°	75.6°	44.8°	-9.0°	58.0°
관절 각 14	-3.9°	5.3°	-32.7°	-15.5°	102.1°	-22.9°
관절 각 15	48.7°	54.1°	64.8°	83.5°	147.3°	75.5°
관절 각 16	12.0°	31.7°	84.5°	57.3°	123.6°	69.0°
관절 각 17	15.2°	11.1°	60.7°	33.4°	7.7°	45.1°
관절 각 18	-9.2°	3.8°	19.6°	1.9°	99.1°	9.5°
관절 각 19	-6.8°	-1.0°	66.1°	49.3°	99.7°	56.5°
관절 각 20	0.2°	-5.6°	71.0°	60.6°	92.7°	65.1°

표 A-2. 인식된 막대 각도에 따라 생성된 관절 각도 데이터.

Table A-2. Generated joint angles based on detected stick angles.

	시연(학습) 데이터								검증작업
	-90 [°]	-55 [°]	-54 [°]	-1 [°]	1 [°]	49 [°]	50 [°]	90 [°]	78 [°]
관절 각 1	42.6°	-4.1°	-6.6°	0.2°	4.5°	7.2°	43.1°	60.0°	54.9°
관절 각 2	-30.9°	-18.2°	-18.0°	-18.3°	-18.6°	-14.8°	8.3°	15.0°	13.0°
관절 각 3	19.9°	4.6°	12.3°	10.0°	13.2°	8.9°	58.3°	71.1°	67.3°
관절 각 4	30.7°	10.7°	6.0°	17.8°	7.0°	9.2°	15.2°	31.4°	26.5°
관절 각 5	6.7°	35.3°	34.9°	43.7°	-4.7°	2.8°	1.3°	8.6°	6.4°
관절 각 6	-16.5°	2.5°	-7.0°	-18.3°	-9.0°	-8.4°	-16.1°	-18.4°	-17.7°
관절 각 7	-9.4°	-12.6°	0.1°	22.4°	10.0°	-11.7°	-15.3°	-11.3°	-12.5°
관절 각 8	-16.0°	-1.5°	0.1°	-20.0°	8.6°	7.9°	-3.9°	-15.6°	-12.1°
관절 각 9	3.6°	5.8°	27.3°	4.7°	10.4°	11.2°	-0.3°	6.0°	4.1°
관절 각 10	-9.4°	-0.3°	-11.4°	0.7°	-7.2°	-5.1°	-5.0°	-10.4°	-8.8°
관절 각 11	-12.1°	3.0°	3.4°	5.5°	8.9°	8.1°	-11.0	-7.3°	-8.4°
관절 각 12	5.7°	1.4°	10.0°	2.0°	0°	6.8°	14.2°	6.4°	8.7°
관절 각 13	2.2°	4.2°	20.8°	-4.1°	38.9°	38.8°	-8.1°	6.2°	1.9°
관절 각 14	-1.1°	1.8°	-1.9°	8.6°	-24.2°	-15.6°	7.5°	-3.1°	0.1°
관절 각 15	2.3°	6.7°	11.3°	9.5°	53.1°	28.1°	0.0°	-1.8°	-1.3°
관절 각 16	0.5°	-0.1°	0.8°	7.3°	-9.3°	0.9°	-4.0°	-5.2	-4.8°
관절 각 17	2.3°	-2.3°	7.9°	5.3°	-1.3°	4.1°	6.7°	4.9°	5.4°
관절 각 18	5.7°	-6.9°	-3.0°	1.3°	4.2°	-5.6°	4.5°	10.9°	9.0°
관절 각 19	-11.0°	-5.7°	-10.0°	-3.0°	-4.7°	-10.3°	3.6°	5.4°	4.9°
관절 각 20	-5.5°	6.8°	-0.6°	4.7°	2.1°	7.8°	8.2°	3.1°	4.6°

표 A-3. 인식된 공 높이에 따른 생성된 관절 각도 데이터.

Table A-3. Generated joint angles based on detected ball height.

	시연(학습) 데이터		검증작업
	120.3 [mm]	265.7 [mm]	152.6 [mm]
관절 각 1	55.2°	60.0°	56.3°
관절 각 2	10.3°	8.6°	9.9°
관절 각 3	66.6°	63.3°	65.9°
관절 각 4	21.5°	42.5°	26.1°
관절 각 5	65.2°	49.1°	61.7°
관절 각 6	-0.1°	-4.3°	-1.0°
관절 각 7	84.7°	71.2°	81.7°
관절 각 8	10.9°	31.1°	15.3°
관절 각 9	46.5°	57.6°	48.9°
관절 각 10	41.7°	52.9°	44.2°
관절 각 11	62.5°	69.1°	64.0°
관절 각 12	89.6°	101.4°	92.2°
관절 각 13	50.4°	51.6°	50.7°
관절 각 14	61.0°	63.2°	61.5°
관절 각 15	76.4°	77.3°	76.6°
관절 각 16	59.2°	69.4°	61.4°
관절 각 17	66.6°	67.4°	66.8°
관절 각 18	11.6°	13.1°	11.9°
관절 각 19	73.6°	65.1°	71.7°
관절 각 20	72.9°	73.1°	72.9°
손목 각도	30.0°	1.5°	23.7°