

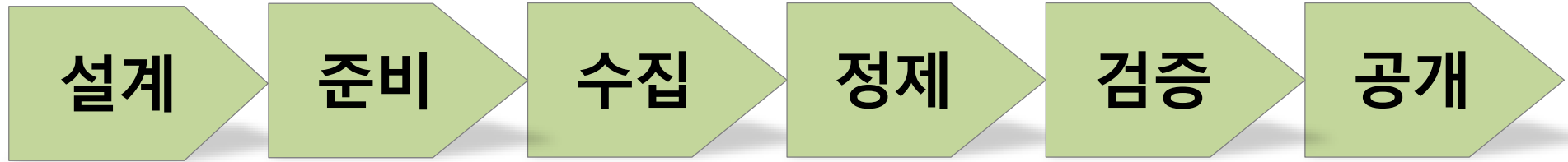
로봇 환경에서의 고령자 일상 행동 인식을 위한 대규모 3차원 영상 데이터 셋 구축과 공개

2019.10.16

김도형

한국전자통신연구원

목 차



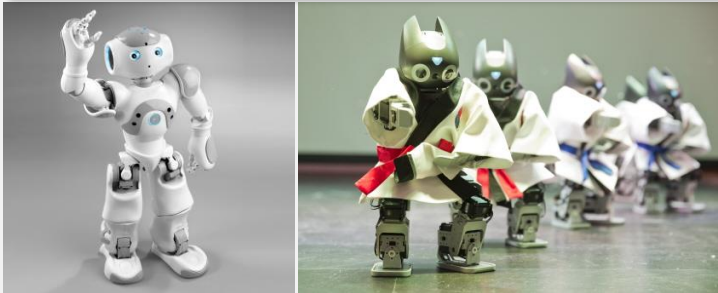
현재의 로봇 지능

동일하고 단발적인 반응을 반복



“주인님 안녕하세요, 좋은 아침이에요!”

규칙에 따라 사전 제작된 동작 수행



<얼굴검출> → [인사동작]+“안녕하세요”

흥미/유대감 저하 → 로봇 **효용성 감소**

미래의 로봇 지능

주인을 이해하고 관심을 표현



“주인님, 오랜만에 모자를 쓰셨네요. 페도라가 잘 어울려요.”

상대방의 정서를 인지하고 사람처럼 반응



외로워 보이면 위로하고 안아주는 시늉

유대관계 형성 → 로봇 **효용성 증가**

정서교감과 자연스러운 교류 기반
고령자 개인 맞춤형 서비스



정서 지원 건강 지원 생활 지원 인지 지원

휴먼케어로봇 서비스 통합 플랫폼과 고령자 특화 서비스



로봇-고령자간
유대감 향상과
관계의 장기화

핵심 가치

관심 / 정서 교감

신뢰 / 수용

“개인에 대해 심층 이해”
“마음을 담은 듯한 교류”

“건강단서의 정밀 감지”
“일상활동의 정확한 분석”

휴먼케어로봇용 4대 핵심 지능정보기술



개인 프로파일링 대인관계지능학습 의도이해와 예측 건강이상징후감지

고령자 도메인 특화 지능 훈련·성장

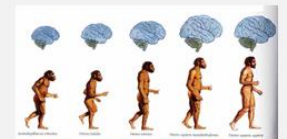


고령자 특성/행동/반응/건강 관련
멀티모달 센싱 데이터 스트림

핵심 지능정보기술
평가·정제 체계



소셜로봇기반 지능 검증

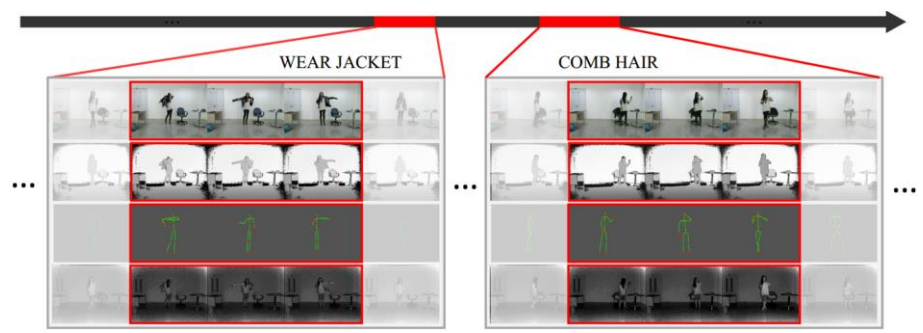
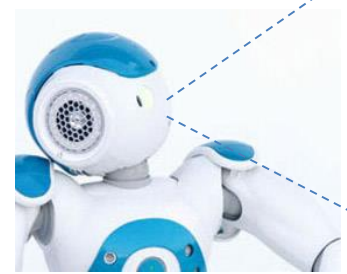


지능 평가 체계 개발



데이터 셋 구축 목적

로봇환경에서 획득한 다중 데이터를 분석하여 **고령자가 취하는 55가지 일상 행동의 아이디어를 인식하는 행동 인식 기술과 그 발생시점까지 검출하는 행동 검출 기술 개발**



이벤트번호	시작시간	종료시간	인식행동
...
...
120	11:32:10	11:32:50	물/음료 마시기
121	11:40:30	11:40:40	손씻기
122	12:05:11	12:05:30	전화받기
...
...

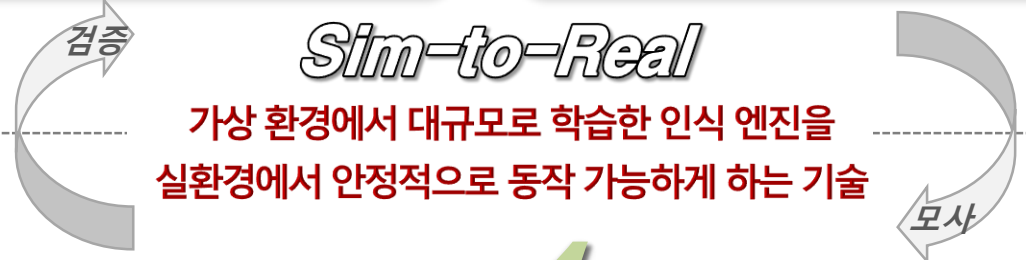
고령자 특화 대용량 데이터 확보 방안

Real world

Simulation

1 공개 데이터 확보 및 필요 데이터 확대 재생산

2 리빙랩 운영으로 장기적인 실제 고령자 데이터 확보



3 노인 외형/신체 모델링으로 대규모 가상 데이터 생성

4 고령자 주거 공간을 모사한 가상 주거 환경 구축

기존 데이터 셋 (2D)



ACTIVITYNET

Large Scale Activity Recognition Challenge



Washing dishes



Ironing clothes



Shotput



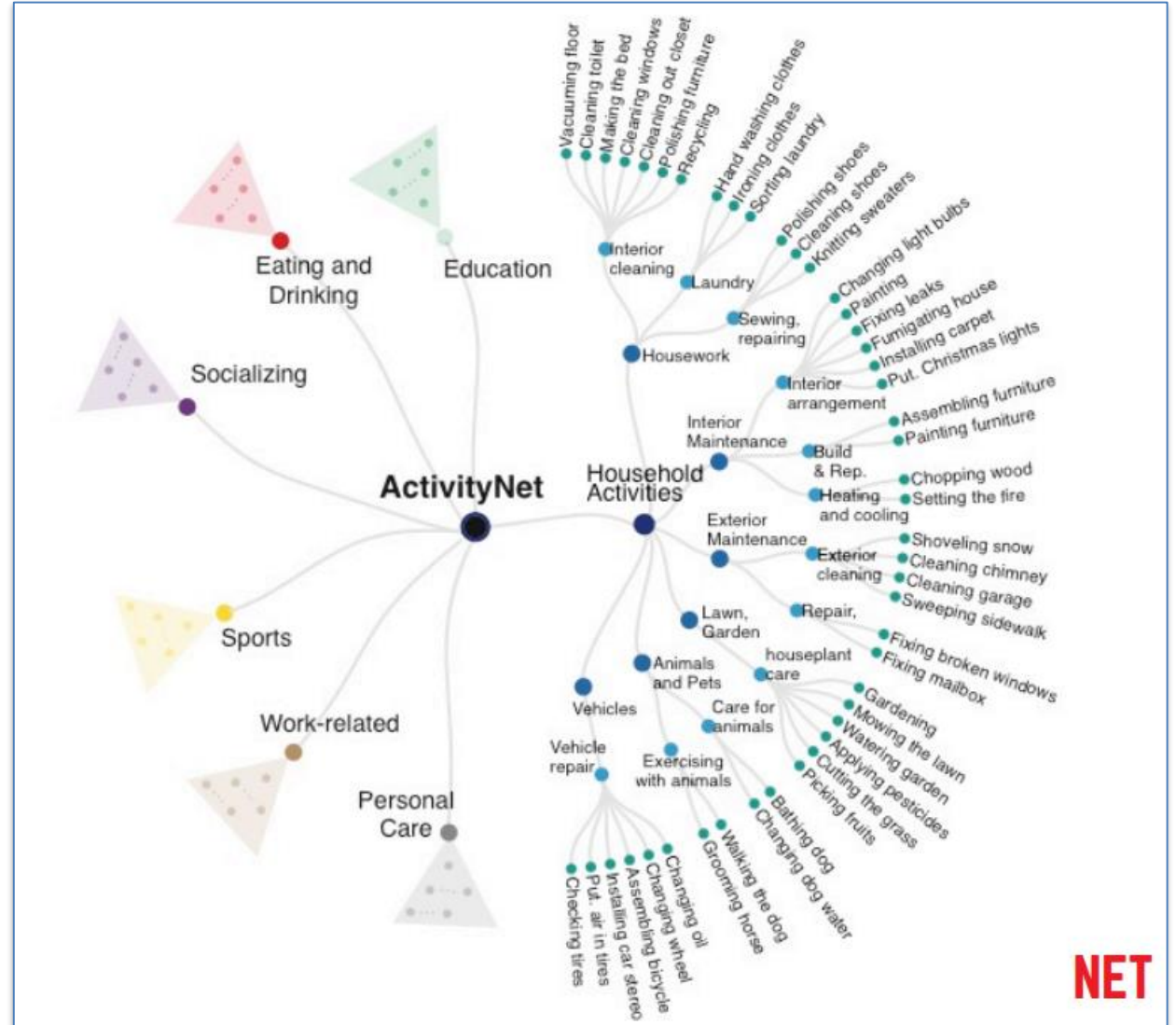
Water skiing



Rock climbing



Applying sunscreen



NET

기존 데이터 셋 (2D)

Table 1. Comparison of Charades with other video datasets.

Charades

	Actions per video	Classes	Labelled instances	Total videos	Origin	Type	Temporal localization
Charades v1.0	6.8	157	67K	10K	267 Homes	Daily Activities	Yes
ActivityNet [3]	1.4	203	39K	28K	YouTube	Human Activities	Yes
UCF101 [8]	1	101	13K	13K	YouTube	Sports	No
HMDB51 [7]	1	51	7K	7K	YouTube/Movies	Movies	No
THUMOS'15 [5]	1-2	101	21K+	24K	YouTube	Sports	Yes
Sports 1M [6]	1	487	1.1M	1.1M	YouTube	Sports	No
MPII-Cooking [14]	46	78	13K	273	30 In-house actors	Cooking	Yes
ADL [25]	22	32	436	20	20 Volunteers	Ego-centric	Yes
MPII-MD [11]	Captions	Captions	68K	94	Movies	Movies	No

- Holding a book
- Sitting on sofa/couch
- Sitting in a chair
- Someone is holding a paper/notebook
- Sitting in a bed
- Watching/Reading/Looking at a book
- Someone is standing up from somewhere
- Putting their paper/notebook somewhere
- Putting a book somewhere
- Walking through a doorway
- Making a sandwich
- Holding a dish
- Holding some food
- Taking food from somewhere
- Putting some food somewhere
- Holding a sandwich



Action Recognition Results

Rank	Team	Accuracy (mAP)
1	TeamKinetics	0.3441
2	DR/OBU	0.2974

Temporal Segmentation Results

Rank	Team	Accuracy (mAP)
1	TeamKinetics	0.2072
2	UMICH-VL	0.1803

<https://www.youtube.com/watch?v=x9AhZLDkbyc>

<http://vuchallenge.org/charades.html>

기존 데이터 셋 (3D)

NTU RGB+D

Datasets	Samples	Classes	Subjects	Views	Sensor	Modalities	Year
MSR-Action3D [19]	567	20	10	1	N/A	D+3DJoints	2010
CAD-60 [34]	60	12	4	-	Kinect v1	RGB+D+3DJoints	2011
RGBD-HuDaAct [23]	1189	13	30	1	Kinect v1	RGB+D	2011
MSRDailyActivity3D [38]	320	16	10	1	Kinect v1	RGB+D+3DJoints	2012
Act4 ² [6]	6844	14	24	4	Kinect v1	RGB+D	2012
CAD-120 [18]	120	10+10	4	-	Kinect v1	RGB+D+3DJoints	2013
3D Action Pairs [25]	360	12	10	1	Kinect v1	RGB+D+3DJoints	2013
Multiview 3D Event [43]	3815	8	8	3	Kinect v1	RGB+D+3DJoints	2013
Online RGB+D Action [46]	336	7	24	1	Kinect v1	RGB+D+3DJoints	2014
Northwestern-UCLA [40]	1475	10	10	3	Kinect v1	RGB+D+3DJoints	2014
UWA3D Multiview [28]	~900	30	10	1	Kinect v1	RGB+D+3DJoints	2014
Office Activity [41]	1180	20	10	3	Kinect v1	RGB+D	2014
UTD-MHAD [4]	861	27	8	1	Kinect v1+WIS	RGB+D+3DJoints+ID	2015
UWA3D Multiview II [26]	1075	30	10	5	Kinect v1	RGB+D+3DJoints	2015
NTU RGB+D	56880	60	40	80	Kinect v2	RGB+D+IR+3DJoints	2016

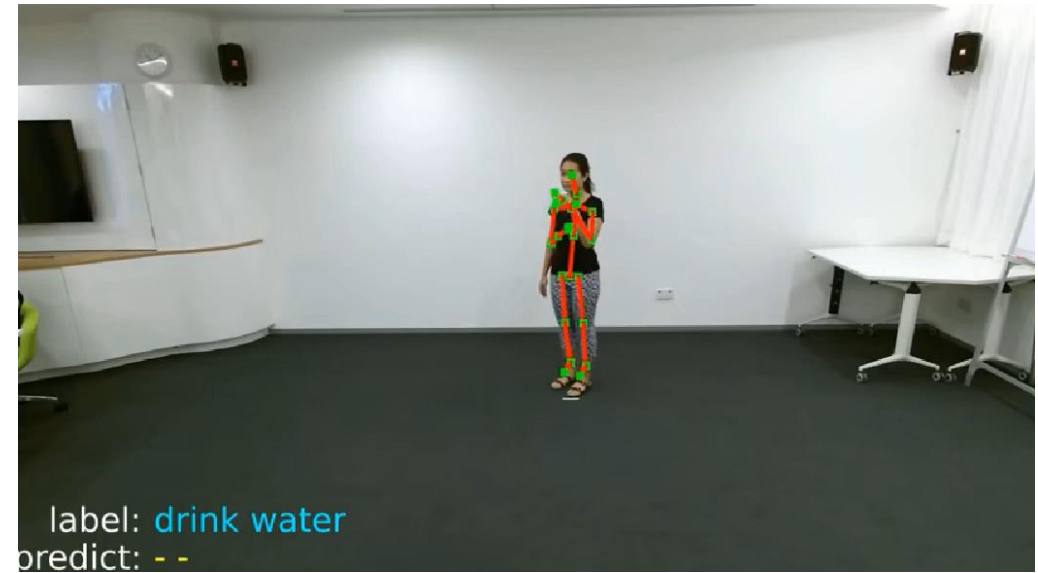


Table 1. Comparison between NTU RGB+D dataset and some of the other publicly available datasets for 3D action recognition. Our dataset provides many more samples, action classes, human subjects, and camera views in comparison with other available datasets for RGB+D action recognition.

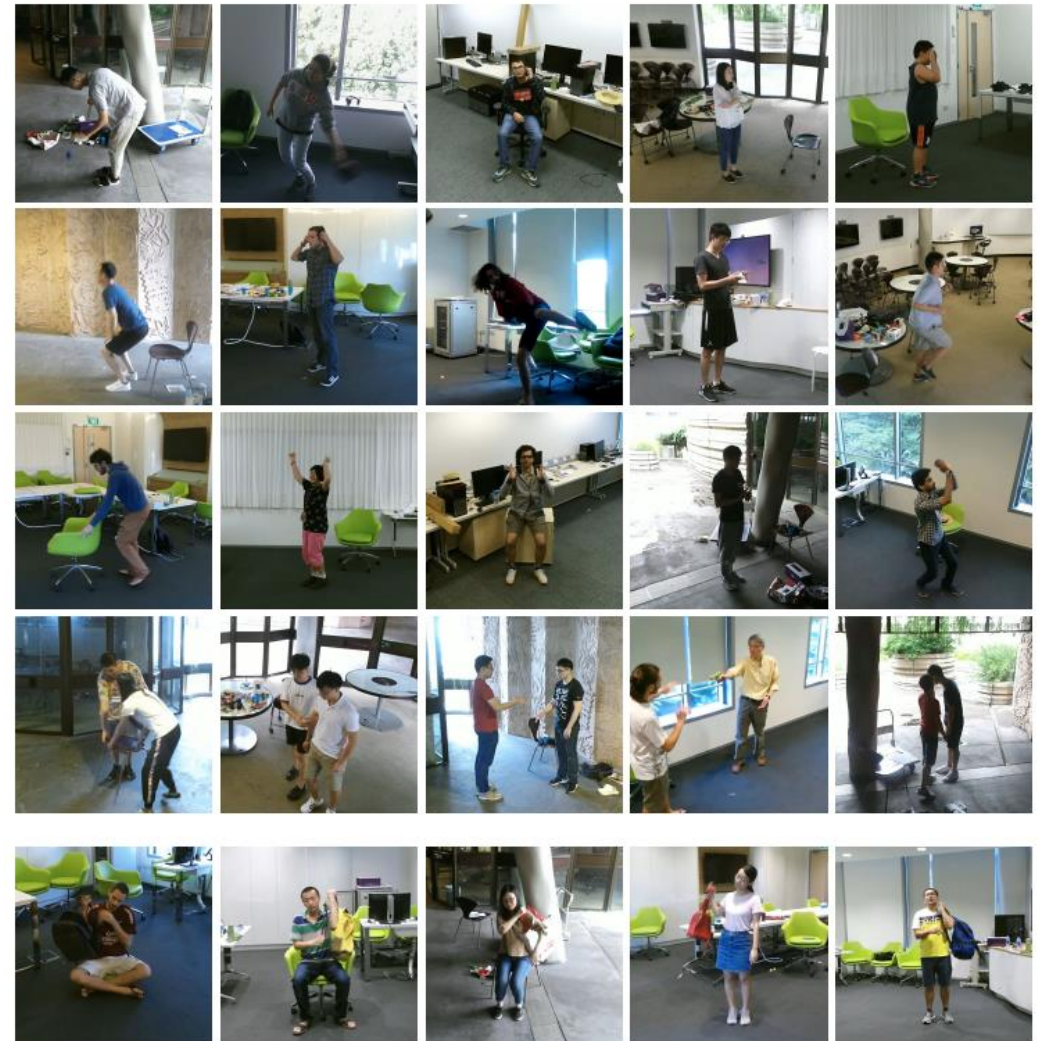
<https://www.youtube.com/watch?v=G0PXKCEgIoA&t=25s>

Amir Shahroudy, Jun Liu, Tian-Tsong Ng, Gang Wang, "NTU RGB+D: A Large Scale Dataset for 3D Human Activity Analysis", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

기존 데이터 셋 (3D)

NTU RGB+D 120

Datasets	#Videos	#Classes	#Subjects	#Views	Sensors	Data Modalities	Year	
MSR-Action3D	[11]	567	20	10	1	N/A	D+3D Joints	2010
CAD-60	[12]	60	12	4	-	Kinect v1	RGB+D+3D Joints	2011
RGBD-HuDaAct	[13]	1,189	13	30	1	Kinect v1	RGB+D	2011
MSRDailyActivity3D	[14]	320	16	10	1	Kinect v1	RGB+D+3D Joints	2012
UT-Kinect	[15]	200	10	10	4	Kinect v1	RGB+D+3D Joints	2012
Act4 ²	[16]	6,844	14	24	4	Kinect v1	RGB+D	2012
CAD-120	[17]	120	10+10	4	-	Kinect v1	RGB+D+3D Joints	2013
3D Action Pairs	[18]	360	12	10	1	Kinect v1	RGB+D+3D Joints	2013
Multiview 3D Event	[19]	3,815	8	8	3	Kinect v1	RGB+D+3D Joints	2013
Northwestern-UCLA	[20]	1,475	10	10	3	Kinect v1	RGB+D+3D Joints	2014
UWA3D Multiview	[21]	~900	30	10	1	Kinect v1	RGB+D+3D Joints	2014
Office Activity	[22]	1,180	20	10	3	Kinect v1	RGB+D	2014
UTD-MHAD	[23]	861	27	8	1	Kinect v1+WIS	RGB+D+3D Joints+ID	2015
UWA3D Multiview II	[24]	1,075	30	10	5	Kinect v1	RGB+D+3D Joints	2015
M ² I	[25]	~1,800	22	22	2	Kinect v1	RGB+D+3D Joints	2015
SYSU 3DHOI	[26]	480	12	40	1	Kinect v1	RGB+D+3D Joints	2017
NTU RGB+D 120		114,480	120	106	155	Kinect v2	RGB+D+3D Joints+IR	



Jun Liu, Amir Shahroudy, Mauricio Perez, Gang Wang, Ling-Yu Duan, Alex C. Kot, "NTU RGB+D 120: A Large-Scale Benchmark for 3D Human Activity Understanding", IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2019.

검출 대상 행동 55종

행동 및 소지품 인식대상 선정을 위한 고령자 관찰

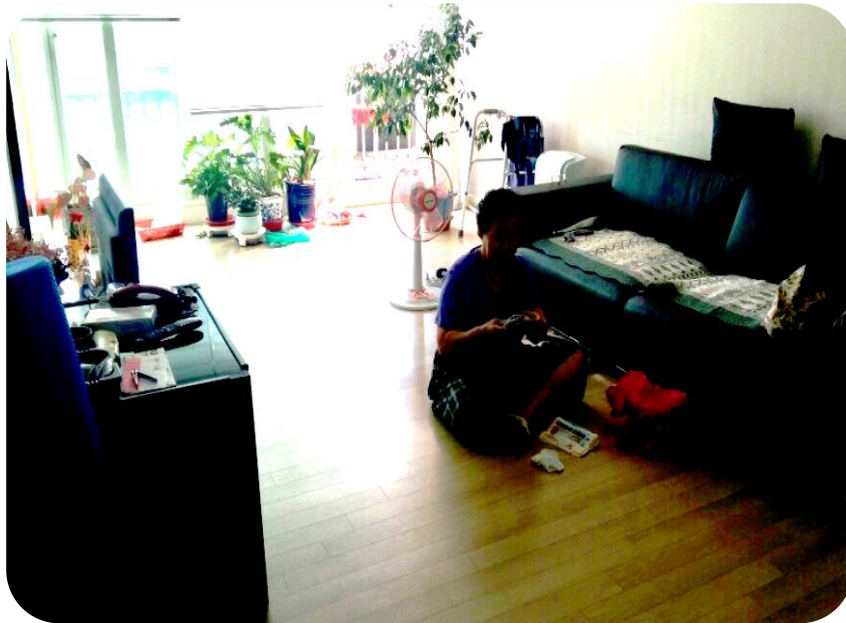
개요	목적	실제 고령자가 자주 하는 행동과 자주 사용하는 물건 관찰
	대상	고령자 53명 (70세 이상)
	조사기간	2017년 6월 16일 ~ 7월 5일
	조사방법	독거 고령자의 실 주거환경을 방문하여 하루 행동을 관찰 기록
관찰 결과	일상 행동 종류	총 245개의 일상 행동
	다빈도 활동	<ol style="list-style-type: none"> 1. TV 시청 2. 식사 관련 활동 (식사, 식사준비, 설거지, 식사 후 정리) 3. 용변보기 (화장실 사용) 4. 전화통화 5. 약복용 6. 세수 및 양치질 7. 옷 입고 벗기 (양말, 안경, 모자..)
	자주 사용하는 물건	핸드폰, 리모콘, 안경, 선풍기, 집구, 약통, 컵

검출 대상 행동 55종

구분	ID	인식대상 행동 (55종)	구분	ID	인식대상 행동 (55종)	구분	ID	인식대상 행동 (55종)
음식	1	수저/포크로 음식 집어먹기	가사	22	그릇 설거지 하기	건강	39	박수 치기
	2	물/음료를 컵에 따르기		23	진공청소기 돌리기		40	두손으로 얼굴비비기
	3	약을 뜯어서 삼키기		24	걸레로 옆드려서 바닥 닦기		41	맨손체조하기
	4	물/음료 마시기		25	식탁 걸레질 하기		42	목돌리기 운동하기
	5	냉장고 음식 넣고 꺼내기		26	창문/가구 등 닦기		43	어깨 셀프 안마하기
	6	콩나물/시금치/콩 다듬기		27	이불펴고/개기		44	고개숙여 인사하기
	7	과일깎기		28	손빨래하기		45	담소 나누기
	8	가스레인지 음식 데우기		29	빨래 널기		46	약수 하기
	9	칼로 도마위 음식 자르기		30	물건찾을려고 두리번거리기		47	포옹 하기
용모	10	이빨 닦기	31	리모콘으로 TV 컨트롤하기	HH 상호작용	48	서로 싸우기	
	11	손씻기	32	책 읽기		49	손을 좌우로 하기 (waving)	
	12	세수하기	33	신문 보기		HR 상호작용	50	이리오라고 손짓 하기 (calling)
	13	수건으로 얼굴/머리 닦기	34	글쓰기			51	손가락으로 가리키기 (pointing)
	14	화장품 바르기	35	전화걸기/받기	기타		52	문을 열고 들어가기
	15	립스틱 바르기	36	스마트폰 조작하기		53	쓰러지기	
	16	머리 빗기	37	컴퓨터 자판 치기		54	누워있다 일어나기(기상)	
	17	머리 드라이기로 말리기	38	담배피기		55	서있다가(앉아있다가) 눕기	
	18	상의 입기						
	19	상의 벗기						
	20	신발 신고 벗기						
	21	안경 쓰고 벗기						

데이터셋 촬영 환경

리빙랩



아파트 테스트베드



리빙랩 운영 계획

- 정의: 고령자가 실제 생활하는 주거 환경에 구축된 테스트베드
- 목적: 고령자의 실 데이터 획득 및 개발된 기술과 서비스의 적합성을 검증
- 운영계획
 - 운영기간: 2017.10~2020.12 (총 3년 3개월)
 - 가구수: 총 54 가구
 - 수집데이터: RGBD영상 데이터, 환경센서 데이터, 웨어러블센서 데이터

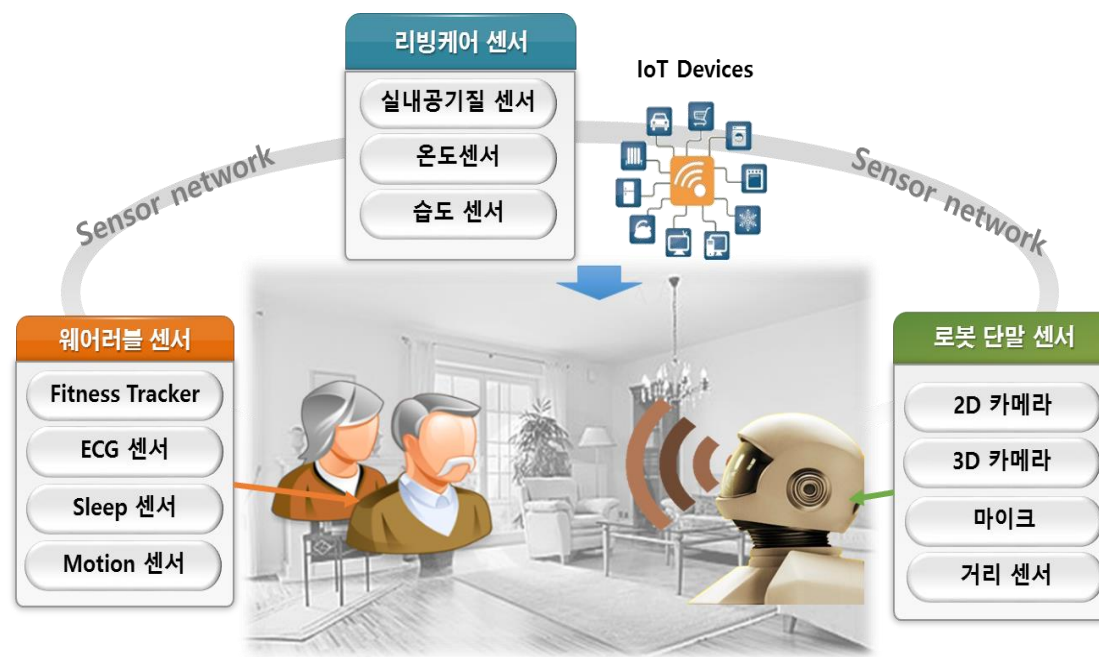
리빙랩(Living Lab) 역할

현장에서 요구되는 기술의 연구개발 반영

고령자를 대상으로 실제 데이터 획득

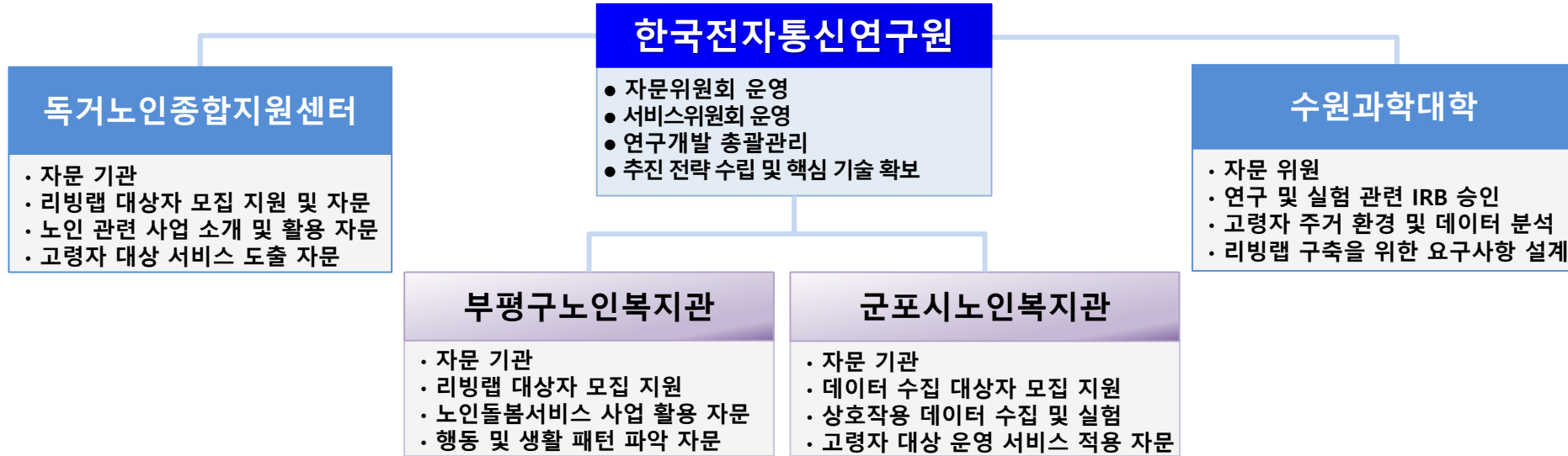
개발된 기술과 서비스의 적합성 검증

시장진입 간극을 해소하여 상용화 기간 단축





고령자 전문 자문 위원회 운영



노인 전문가 초청 세미나	<ul style="list-style-type: none"> • 노인의 실태와 돌봄의 필요성 세미나 • 노인의 사회적 관계망 특성 및 의사 소통 데이터 분석 • 독거노인 사랑잇기 사업 데이터 활용 방안 협의
리빙랩 대상자 모집 및 데이터 수집	<ul style="list-style-type: none"> • 리빙랩 대상자 모집 협의 및 자문 • 고령자 대상 상호작용 데이터 수집 관련 협의
고령자 대상 서비스 도출 및 적용	<ul style="list-style-type: none"> • 고령자 대상 돌봄 로봇 적용 협의 및 자문 • 고령자 대상 제공 서비스 검토 및 자문
전문가 위원회 개최	<ul style="list-style-type: none"> • 리빙랩 데이터 수집 관련 전문가 위원회 개최 • 2차년도 리빙랩 운영 관련 전문가 위원회 개최

리빙랩 모집 및 테스트 운영 준비 (2017.6~9)

- 고령자 데이터 획득 프로토콜 설계에 따른 IRB 획득 (2017.06)
- 리빙랩 참가자 모집 및 선정 (2017.08~09)
 - 독거노인종합지원센터, 부평노인복지관 협력
 - 거주 형태, 생활 패턴 및 사회적 관계 특성을 고려한 대상자 모집: 8가구 지원
 - 대상자 선정: 지원자의 인지 상태, 참여 의지 및 주거 형태를 고려한 최종 3 가구 선정
- 거주 환경
 - 20평~23평대 아파트와 빌라 (방 2, 주방, 거실)

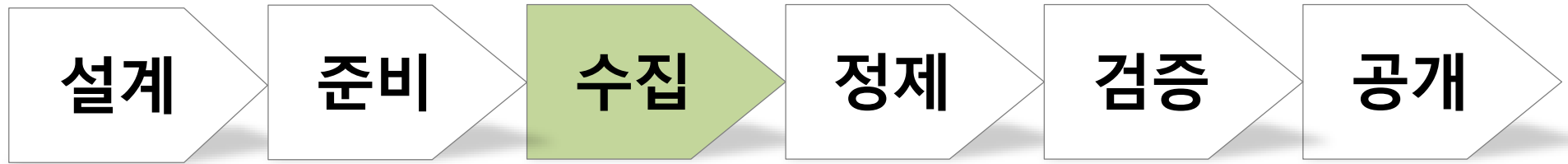
2017	3가구
2018	6가구
2019	24가구
2020	24가구



아파트 테스트베드 구축 및 고령자 섭외 (2018.6)

- 위치: 대전시 유성구 신성동 대림두레아파트 (31평형)
- 구축 목적
 - 실환경과 유사한 대용량 데이터 셋 수집
 - 실제 환경에 적용 가능한 상용화 기술 개발 및 평가
- 대상자 섭외: 유성구노인복지관 70세 이상 고령자 50명 (2018년), 일반성인 50명 (2019년)





리빙랩 테스트 운영 시작 (2017.10~12)

- 고정된 위치에서 총 3대의 Kinect 2.0 카메라 동시 촬영

시행착오

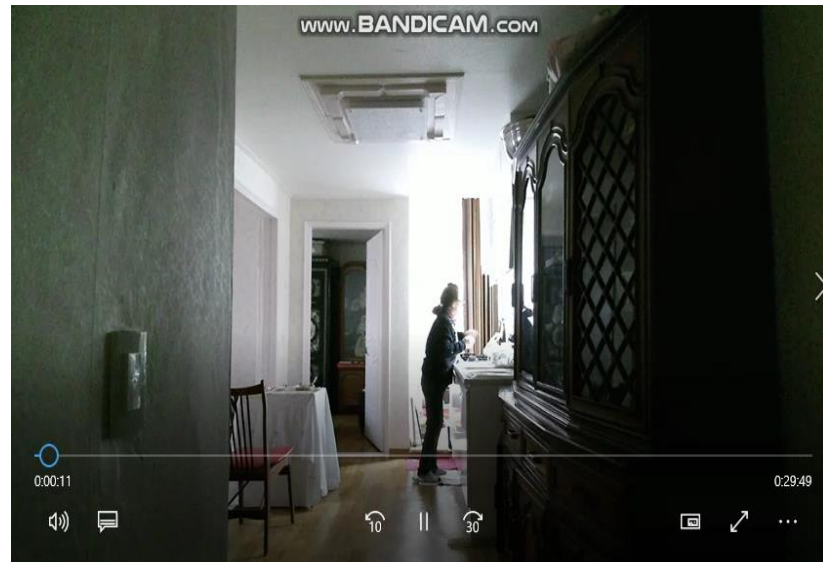


리빙랩 테스트 운영 시작 (2017.10~12)

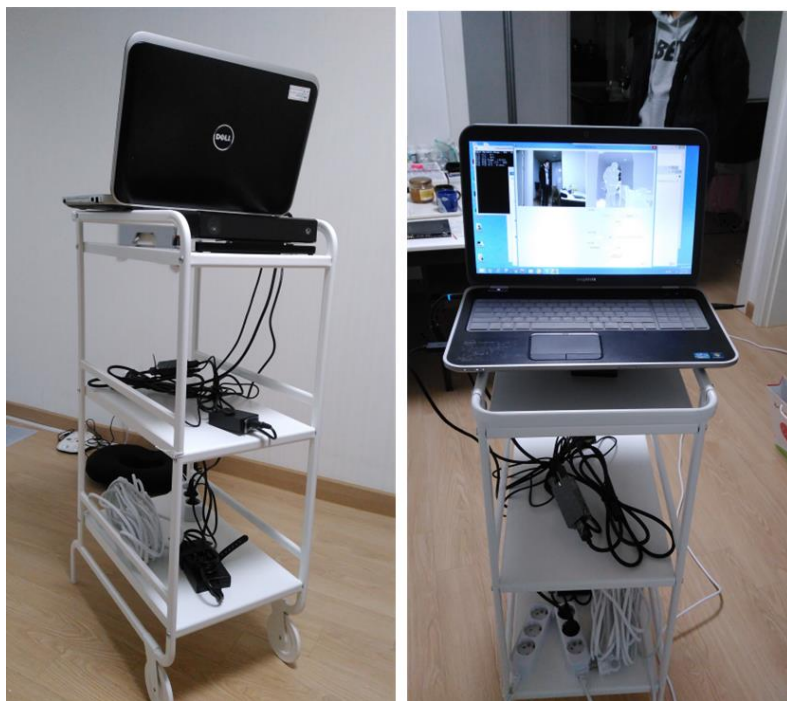


시행착오

- 로봇 서비스 환경 ?
- 데이터 확보 효율성 ?
- 어르신들의 협조 !!!



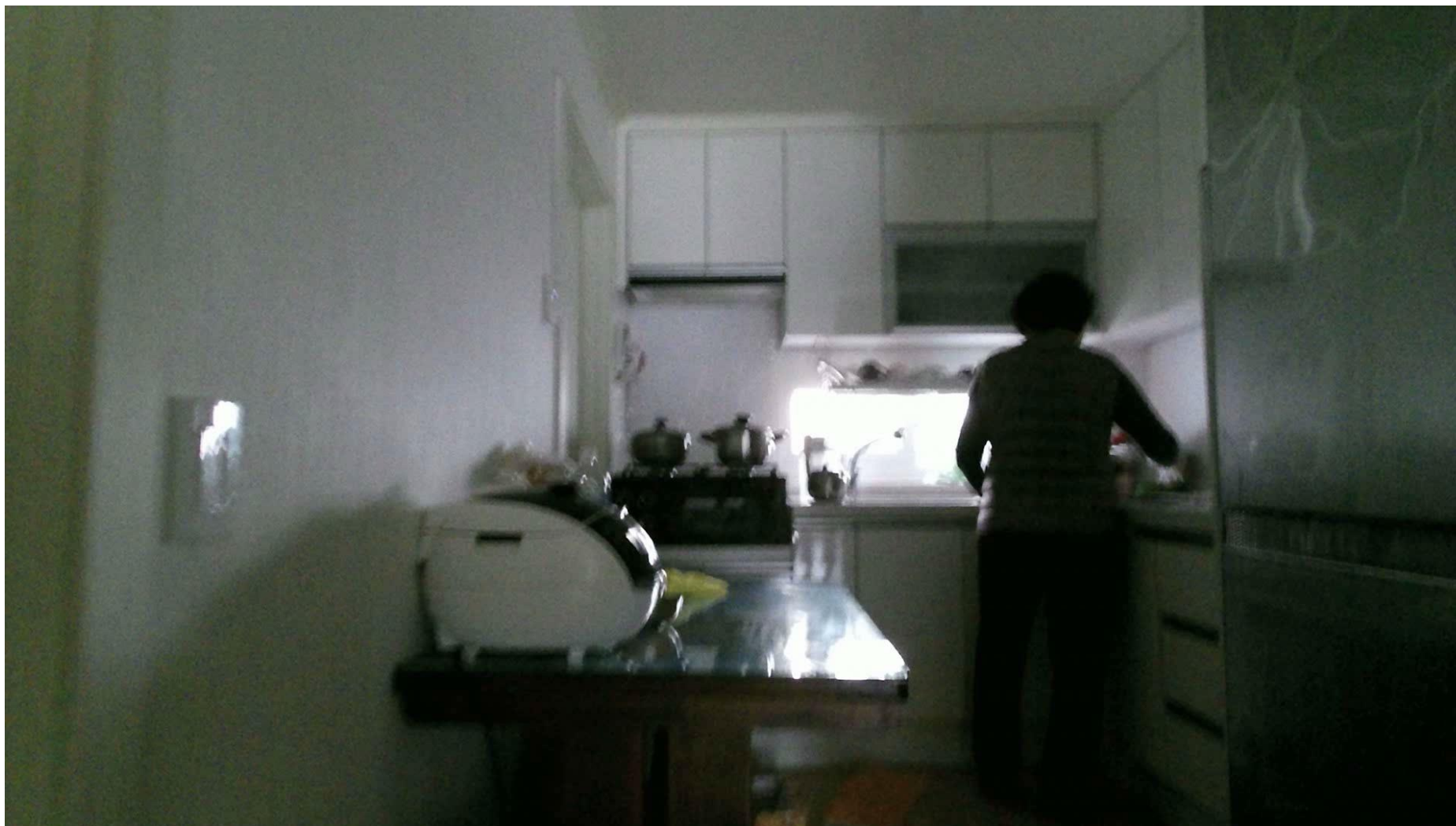
리빙랩 데이터셋 수집 시작 (2018.2~)



< 로봇 대응 촬영 장치 >

항목	설명
장치	3D 카메라가 장착된 로봇 대응품 활용
카메라 높이	2종류 (70cm, 120cm)
방법	촬영자가 어르신 근처의 적정 위치에 로봇을 끌고 다니면서 촬영
촬영거리	다양한 거리 (약 1m~3m 사이의 바디가 촬영될 수 있는 범위)
촬영각도	다양한 각도 (바디가 촬영될 수 있는 범위에서 로봇을 회전)
시나리오1	어르신에게 특별한 요구사항 없이 일상생활을 그대로 촬영 가상의 상황을 설정하고 그에 맞는 행동을 취해달라고 요구
시나리오2	어르신에게 55개의 행동을 취해달라고 요구하고 촬영

로봇 시각 기반 고령자 일상 행동 촬영



수집된 리빙랩 데이터셋 (2018.2~현재)

- **촬영환경:** 고령자 실거주 공간 (리빙랩)
- **구축 방법**
 - 독거 노인이 실제 거주하는 주거환경에서 고령자의 55종 행동을 촬영
 - 현재 18개 가구에 대하여 데이터 수집중
- **특징**
 - **고령자의 실거주환경에서 로봇시점으로 촬영한 대용량 RGBDS 행동인식 데이터 셋 (세계최초)**
- **데이터 형식:** RGB + Depth + Skeleton
- **데이터 규모:** 총 7,120개 데이터 셋 (280 시간 분량)

항 목	내 용
3D 카메라	Kinect v2
데이터 종류 및 해상도	1. RGB video (1920x1080) 2. Depth video (512x424) + BodyIndex 3. 3D Skeletal Joint 좌표 25종 + TrackingState 4. 2D Skeletal Joint 좌표 25종



리빙랩 데이터셋 샘플



리빙랩 데이터셋 샘플



아파트 테스트베드 데이터셋 촬영 장비



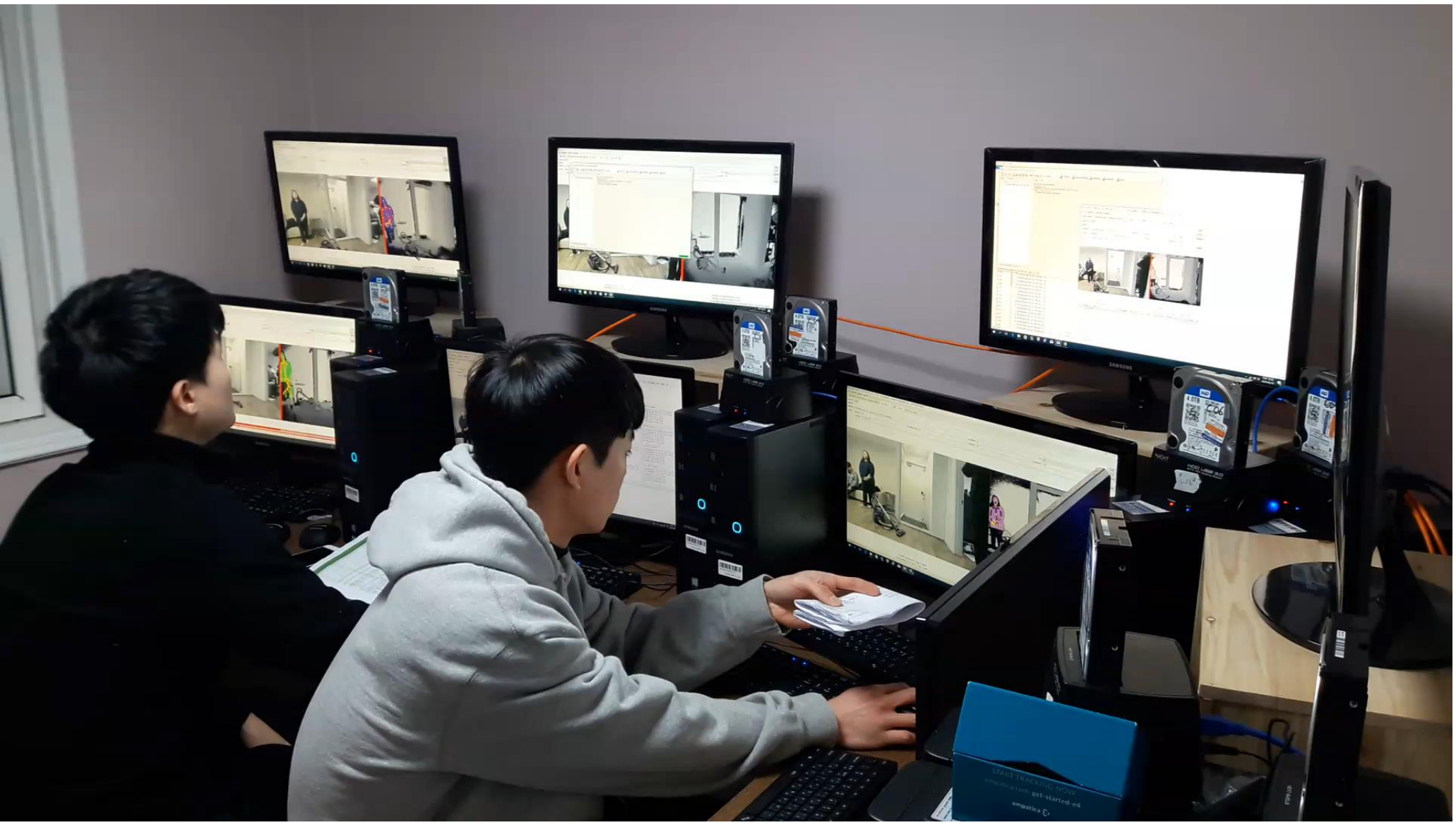
아파트 테스트베드 데이터셋 촬영 모습



아파트 테스트베드 데이터셋 촬영 시나리오

순서	장소	캠위치	Set개수	Group개수	그룹ID	행동ID	행동명	자세	구체적인 행동 수행방법 (자연스럽게)
1	싱크대		1	4	1 G01	A08	가스레인지	서서	(냄비+햇반)를 들고 가스레인지에 올리기(햇반 익힐 것)
2	싱크대		1	4	1 G02	A08	가스레인지	서서	후라이팬 등 다른 기구 사용
3	싱크대		1	4	1 G01	A09	도마위음식자르기	서서	채소를 도마위에 놓고 채소를 칼로 여러번 자르기
4	싱크대		1	4	1 G01	A11	손씻기	서서	물을 들고 손을 씻기
5	싱크대		1	4	1 G01	A22	그릇설거지	서서	수세미를 사용해서 그릇 씻기
6	싱크대		1	4	1 G02	A22	그릇설거지	서서	컵을 씻기
7	싱크대		1	4	1 G01	A28	손빨래	서서	물을 들고 행주를 빨기
8	싱크대		1	4	1 G01	A30	물건찾을려고두리번거리기	서서	주방문을 여기저기 열어보면서 물건 찾기
9	싱크대		1	1	0.25 G01	A50	이리 오라고 손짓하기	서서	싱크대 앞에 서서, 카메라를 정면으로 바라보고 (Set1)
10	싱크대		1	1	0.25 G02	A50	이리 오라고 손짓하기	서서	싱크대 앞에 서서, 다른 각도 카메라에 (Set2)
11	싱크대		1	4	1 G01	A02	물 따르기	서서	싱크대 앞에 서서
12	싱크대		1	4	1 G01	A03	약먹기	서서	싱크대 앞에 서서
13	싱크대		1	4	1 G01	A04	물마시기	서서	싱크대 앞에 서서
14	싱크대		1	4	1 G01	A42	목돌리기	서서	싱크대 앞에 서서
15	싱크대		1	4	1 G01	A43	어깨 셀프 안마하기	서서	싱크대 앞에 서서
16	싱크대		1	4	1 G01	A35	전화걸기받기	서서	휴대폰 사용. 뒷모습만 보이지는 않게 어느정도 돌아서서
17	싱크대		1	4	1 G01	A23	진공청소기 돌리기	서서	싱크대 앞 바닥에서(방향을 매 촬영마다 다양하게)
18	싱크대		1	4	1 G01	A38	담배피기	서서	싱크대 앞에 서서 담배피우기
19	냉장고		2	4	1 G01	A05	냉장고 음식 넣고 꺼내기	서서	냉장실에서 꺼내기 (아무거나 꺼내고 싶은거)
20	냉장고		2	4	1 G02	A05	냉장고 음식 넣고 꺼내기	서서	냉동실에서 꺼내기 (냉동식품)
21	냉장고		2	4	1 G01	A26	창문가구닫기	서서	냉장고를 아래 위로 닫기
22	냉장고		2	4	1 G01	A53	쓰러지기	누워서	근처에 쓰러져 있기
23	식탁		3	4	1 G01	A07	과일깎기	앉아서	과도로 사과/참외/배 등을 깎기
24	식탁		3	4	1 G02	A09	도마위음식자르기	앉아서	깎은 사과를 도마에 놓고 식칼로 자르기
25	식탁		3	4	1 G01	A01	수저/포크 음식 집어먹기	앉아서	숟가락으로 떠 먹기(햇반)
26	식탁		3	4	1 G02	A01	수저/포크 음식 집어먹기	앉아서	젓가락으로 집어 먹기
27	식탁		3	4	1 G02	A02	물 따르기	앉아서	물병을 들어서 컵에 물을 따르고 물병을 놓기
28	식탁		3	4	1 G02	A03	약먹기	앉아서	뜯어서 먹는 약 (젤리 등)
29	식탁		3	4	1 G03	A03	약먹기	앉아서	통에서 꺼내 먹는 약 (비타민 등)
30	식탁		3	4	1 G02	A04	물마시기	앉아서	컵을 들어서 물을 마시고 컵을 내리기
31	식탁		3	4	1 G01	A21	안경쓰고벗기	앉아서	카메라 방향으로 보면서 자연스럽게 안경 쓰기
32	식탁		3	4	1 G01	A32	책읽기	앉아서	책을 식탁에 두거나 손에 들고 읽기(책장 넘기기 필요)
33	식탁		3	4	1 G01	A33	신문보기	앉아서	신문을 식탁에 두거나 손에 들고 읽기(신문 넘기기 필요)
34	식탁		3	4	1 G02	A21	안경쓰고벗기	앉아서	카메라 방향으로 보면서 안경 벗기

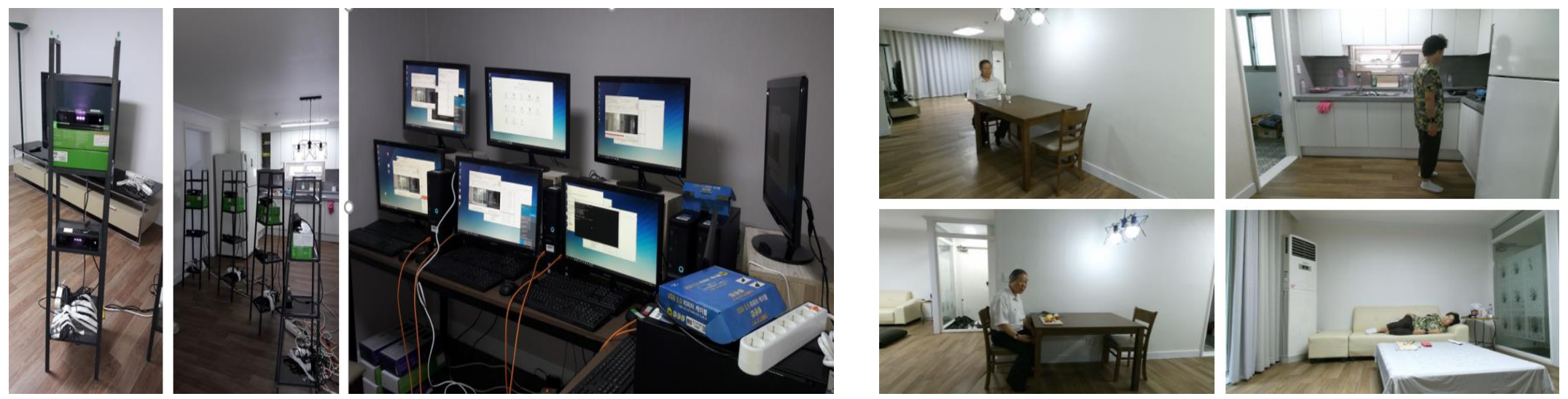
아파트 테스트베드 데이터셋 촬영 관제실



수집된 아파트 테스트베드 데이터셋 (2018.7~9 / 2019.1~3)

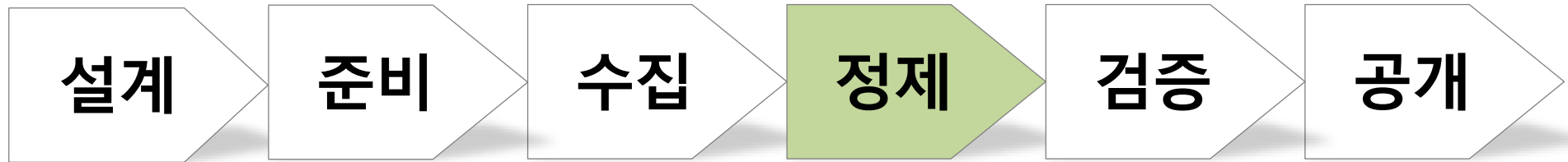
- 촬영환경: APT 테스트베드
- 구축 방법
 - 50명의 고령자와 50명의 성인을 섭외하여 55종의 일상 행동을 촬영
 - APT 테스트베드 환경에서 6개월간 촬영
- 특징
 - 다양한 로봇 환경에서 촬영한 대용량 RGBDS 행동인식 데이터 셋 (2019.11 공개 예정)
- 데이터 형식: RGB video + Depth video + Skeleton
- 데이터 규모: 총 112,658 데이터 셋

항 목	내 용
3D 카메라	Kinect v2
데이터 종류 및 해상도	1. RGB video (1920x1080) 2. Depth video (512x424) + BodyIndex 3. 3D Skeletal Joint 좌표 25종 + TrackingState 4. 2D Skeletal Joint 좌표 25종



아파트 테스트베드 데이터셋 샘플





주요 정제 작업

정제

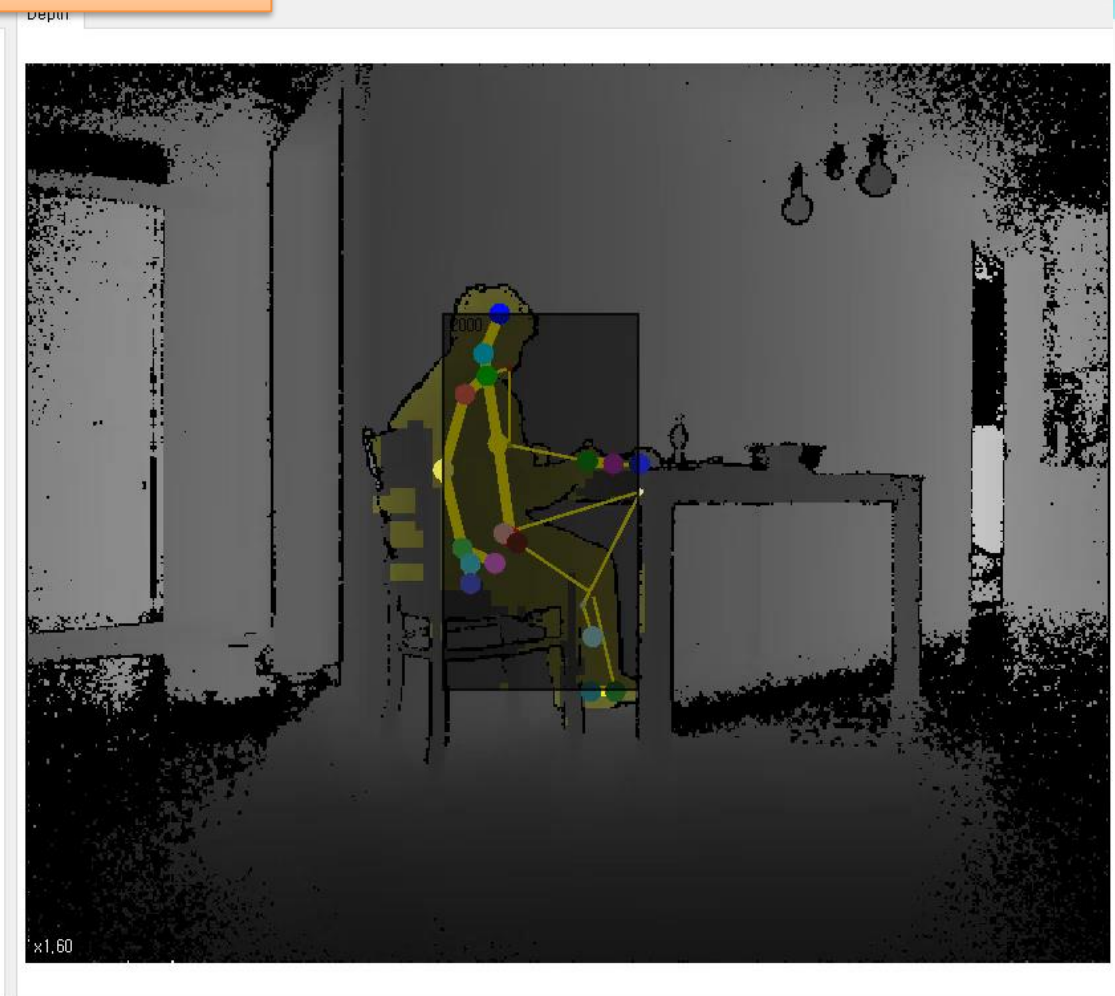
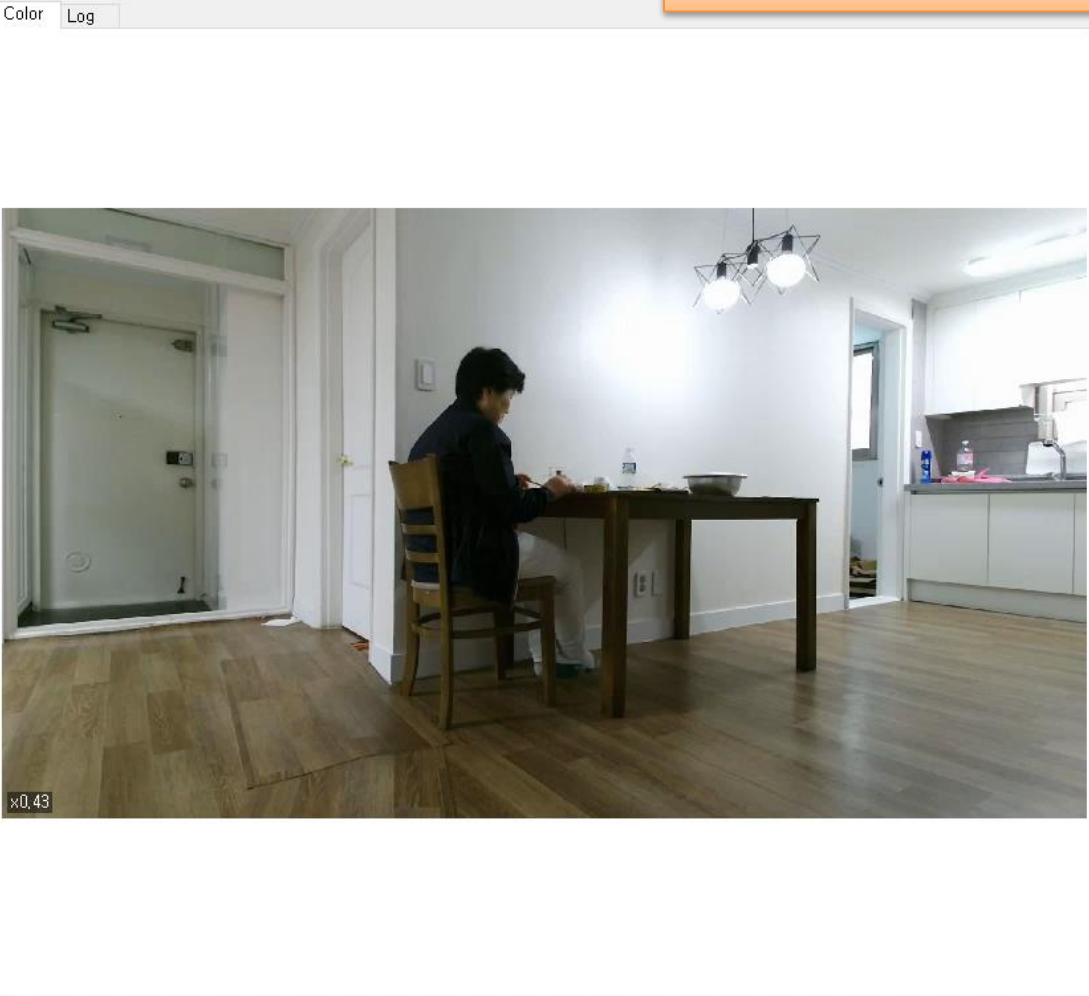


- 행동이 발생하는 정확한 구간 태깅
- 관절 추적 오류 리스트 작성
- Depthmap sequence를 무손실 PNG 파일로 압축 저장

데이터셋 정제툴

Files

- 19AM_S02_G01_P01_A18_C05-104016.avi
- 4AM_S02_G01_P48_A01_C01-111136.avi



0

[0] 00:00:00.000 - [279] 00:00:13.950

Sec 10 20 30 40 50 01:00 01:10

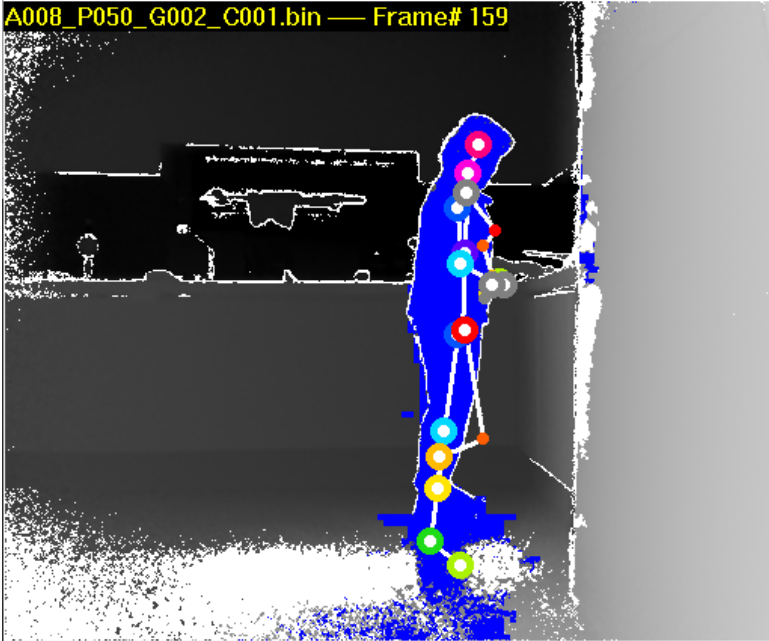
2.0

음식/수저, 포크로 음식 집어먹기

음식/수저, 포크로 음식 집어먹기

음식/수저, 포크로 음식 집어먹기

A008_P050_G002_C001.bin — Frame# 159



1. Rename All Files

A008_P050_G002_C001.avi
A008_P050_G002_C001.bin
A008_P050_G002_C001.body
A008_P050_G002_C001.joints
A044_P001_G002_C001.avi
A044_P001_G002_C001.bin
A044_P001_G002_C001.body
A044_P001_G002_C001.joints

2. Make Joint CSV files

A008_P050_G002_C001.joints
A044_P001_G002_C001.joints

3. Convert Depth/BodyIndex to PNG

A008_P050_G002_C001.bin
A044_P001_G002_C001.bin

4. Confirm / Select Depthmap Png Folder

depthmap027.png
depthmap028.png
depthmap029.png
depthmap030.png
depthmap031.png
depthmap032.png
depthmap033.png
depthmap034.png
depthmap035.png
depthmap036.png
depthmap037.png
depthmap038.png
depthmap039.png
depthmap040.png
depthmap041.png
depthmap042.png
depthmap043.png
depthmap044.png
depthmap045.png
depthmap046.png
depthmap047.png
depthmap048.png
depthmap049.png
depthmap050.png
depthmap051.png

데이터셋 정제툴



데이터 셋 품질 검증 절차



APT Testbed dataset 의 경우 약 9개월 소요 (1M/M 기준)

데이터셋 평가 프로토콜 – 아파트 테스트베드 데이터셋

- Cross-subject evaluation only
 - Training 70%, Test 30%
 - Test datasets: P03, P06 ... P99 (3의 배수로)
 - Training datasets: 나머지 datasets
 - 남녀 비율 고려
- 고령자 행동과 일반 성인 행동의 차이점 분석

	Train	Test
Protocol 1	고령자	고령자
Protocol 2	고령자	성인
Protocol 3	성인	성인
Protocol 4	성인	고령자

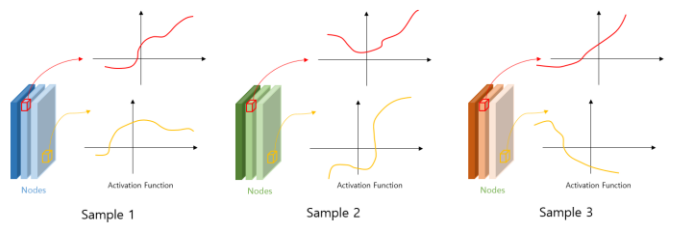
ETRI 행동 인식 엔진을 이용한 데이터 셋 검증

주요연구내용

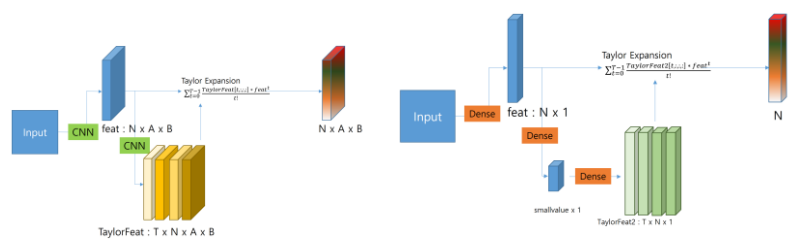
- **행동 인식용 Adaptive Activation Network 모델 개발**
 - Skeleton 정보를 시공간적 difference frame들을 만들어 사용
 - 데이터에 맞게 activation function을 생성하는 모듈 개발
 - Activation function에 Taylor Expansion을 적용해 높은 자유도의 네트워크 구성
 - GlobalMaxPooling을 사용해 비디오 길이와 상관없이 동작

기술의 특징 및 성능

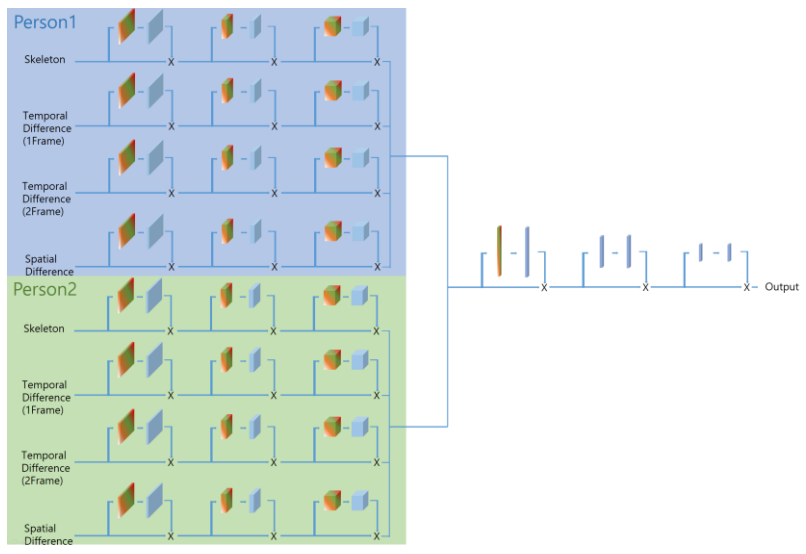
- **특징 및 우수성**
 - Activation Network를 사용해 데이터에 맞게 네트워크 재구성
 - Video 길이에 상관없이 인식이 가능한 네트워크
 - 약 480만 개 Weight의 가벼운 네트워크
- **성능 지표**
 - **행동 인식 성공률: 86.4%** (NTU daily action recognition DB 56,880 set 기준)
 - **행동 인식 성공률: 78.1%** (직접 촬영한 APT Testbed 고령자 DB 기준)



< 고유 특성 반영 Trainable Activation 방법 개념 >



< Trainable Activation 방법 >



< Trainable Activation을 사용한 행동인식기 구조도 >


< State-of-the-art 와의 성능 비교 >

기술	성능
PA-LSTM	70.3%
ST-LSTM+TG	69.2%
VA-LSTM	79.4%
ST-GCN	81.5%
TS-CNN	83.2%
HCN	86.5%
SLnL-rFA	89.1%
Ours w/o Activation Network(baseline)	82.6%
Ours	86.4%

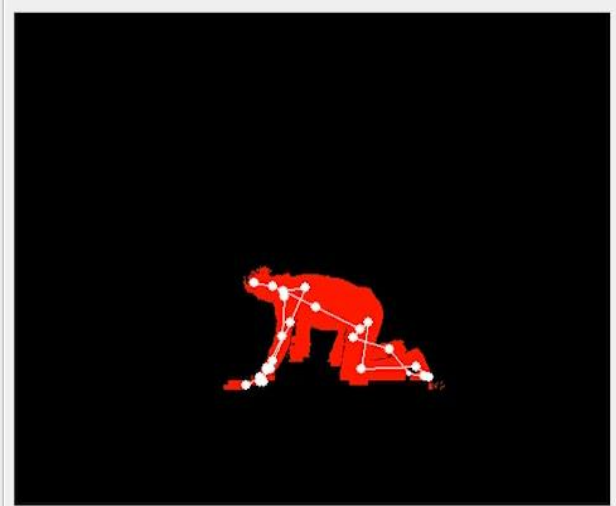
ETRI 행동 인식 엔진을 이용한 데이터 셋 검증

행동인식 - 걸레질

www.BANDICAM.COM



DepthView



Spotting

Spotted data is recording...frame #128

```

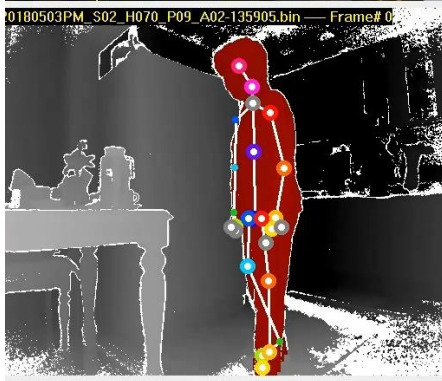
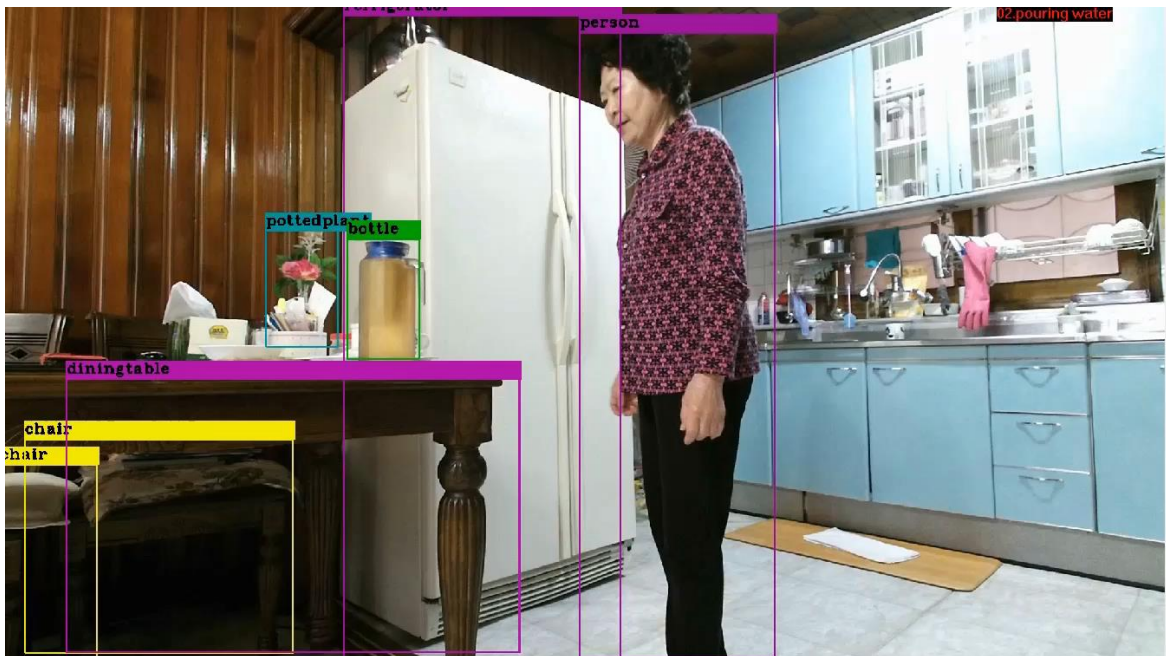
5.5471697e-04 3.0488596e-05 1.2079267e-01 7.3407357e-11 3.9099626e-07
3.5889833e-10 1.8869692e-11 5.4066314e-11 6.5444777e-07 1.9240609e-08
1.8006629e-06 4.6847163e-07 2.5590323e-05 9.1639459e-02 6.6834502e-08
3.5653789e-05 4.7275987e-02 2.9012686e-01 9.4851265e-03 1.4275947e-04
1.0854748e-05 1.5456728e-03 3.0260939e-05 6.4548232e-07 8.1795520e-07]]
[19]
0.04727599 0.09163946 0.12079267 0.29012686 0.40388945]
'포옹하기' '인사하기' '신문보기' '싸우기' '상의 벗기']
press enter
계속하려면 아무 키나 누르십시오 . . .
        
```

대표적인 행동 인식 기술을 이용한 데이터 셋 검증 (진행중)

- Benchmarking
 - Glimpse Clouds, CVPR, 2018
 - (Hierarchical) Co-occurrence Feature Learning from Skeleton Data, IJCAI, 2018
 - Recognizing Human Actions as the Evolution of Pose Estimation Maps, CVPR, 2018
 - Cooperative Training of Deep Aggregation Networks, AAAI, 2018
 - Deep Bilinear Learning for RGB-D Action Recognition, ECCV, 2018

Methods	Score	Dataset	Input	Note
Glimpse Clouds	89.9	NTU RGB+D	RGB	https://github.com/fabienbaradel/glimpse_clouds
HCN	91.1	NTU RGB+D	Skeleton	https://github.com/fandulu/Keras-for-Co-occurrence-Feature-Learning-from-Skeleton-Data-for-Action-Recognition
S2+H3	84.21 (H1+H3) 95.26 (S2+H3)	NTU RGB+D	RGB+S	https://github.com/nkliuyifang/Skeleton-based-Human-Action-Recognition
DDIf+VDIf+ DDIb+VDIf	86 (cross-subj) 89 (cross-view)	NTU RGB+D	RGB+D	N/A
Deep Bilinear	85 (cross-subj) 90 (cross-view)	NTU RGB+D	RGB+D+S	N/A https://github.com/laughtervv/DepthAwareCNN

데이터셋 간접 검증 - 기타 인식 엔진 (멀티모달행동인식)



Multimodal Human Action Recognition

- 01. eating food 0
- 02. pouring water 0
- 03. taking medicine 0
- 04. drinking 0
- 05. using refrigerator 0
- 06. trimming vegetable 0
- 07. peeling fruit 0
- 08. using gas stove 0
- 09. cutting vegetable 0
- 10. brushing teeth 0
- 11. washing hands 0
- 12. washing face 0
- 13. wiping face with towel 0
- 14. putting on cosmetics 0
- 15. putting on lipstick 0
- 16. brushing hair 0
- 17. using hair dryer 0
- 18. putting on clothes 0
- 19. taking off clothes 0
- 20. putting on/taking off shoes 0
- 21. putting on/taking off glasses 0
- 22. washing dishes 0
- 23. vacuuming 0
- 24. wiping floor 0
- 25. wiping dining table 0
- 26. wiping furniture 0
- 27. Making the bed 0
- 28. washing clothes by hand 0
- 29. hanging laundry 0
- 30. looking around for something 0
- 31. using remote control 0
- 32. reading book 0
- 33. reading newspaper 0
- 34. writing 0
- 35. talking on the phone 0
- 36. playing with smartphone 0
- 37. using computer 0
- 38. smoking 0
- 39. clapping 0
- 40. rubbing face 0
- 41. doing freehand exercise 0
- 42. doing neck roll exercise 0
- 43. massaging shoulder oneself 0
- 44. nod a greeting 0
- 45. talking to each other 0
- 46. handshaking 0
- 47. hugging each other 0
- 48. fighting each other 0
- 49. waving hand 0
- 50. sknocking hand up and down(call a robot) 0
- 51. pointing with finger 0
- 52. opening door and walking in 0
- 53. lie on the floor(falling down) 0
- 54. sitting up/standing up 0
- 55. lying on the bed 0

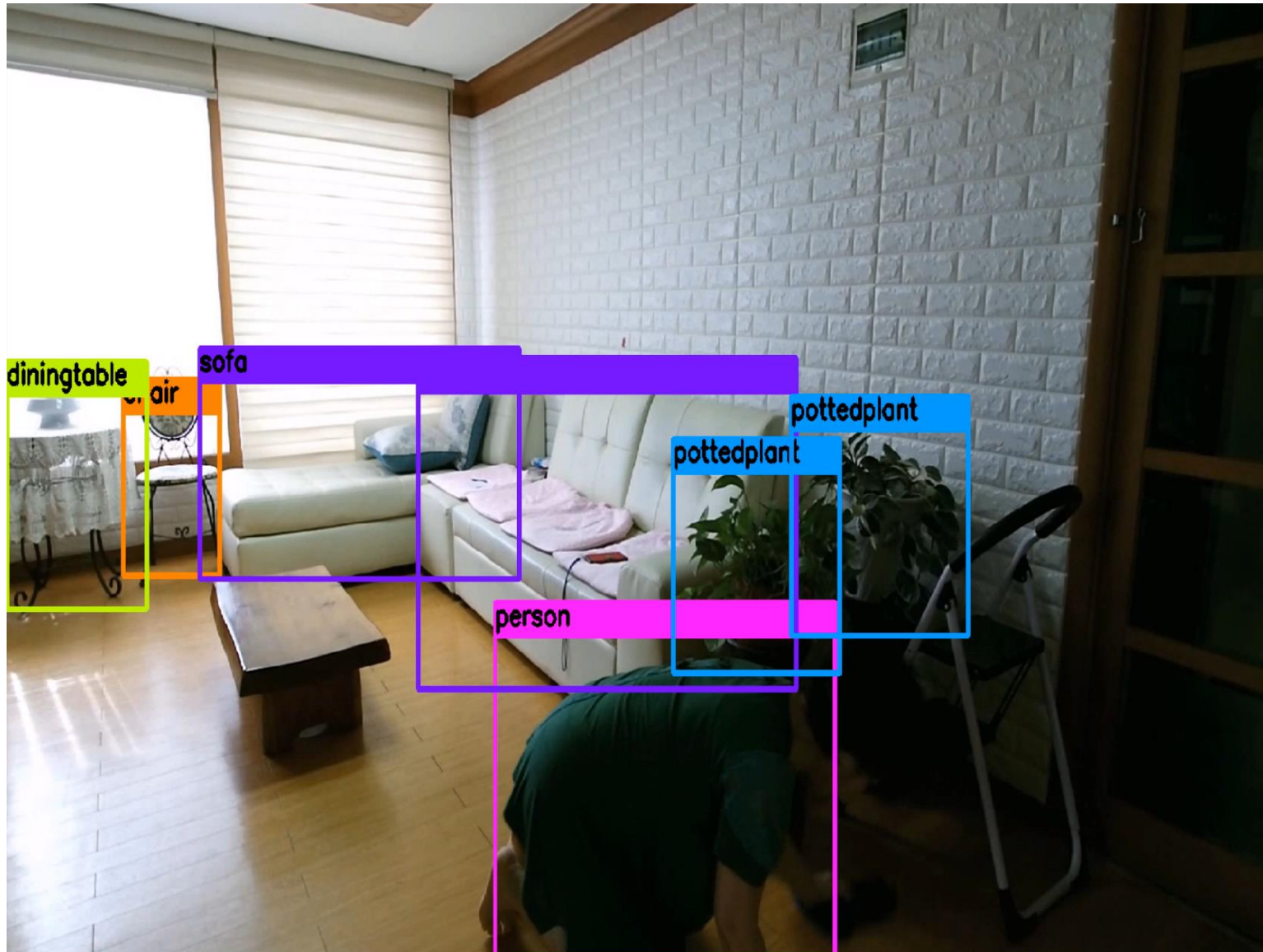
Vision-based Human Action Recognition

- 01. 0
- 02. 0
- 03. 0
- 04. 0
- 05. 0
- 06. 0
- 07. 0
- 08. 0
- 09. 0
- 10. 0
- 11. 0
- 12. 0
- 13. 0
- 14. 0
- 15. 0
- 16. 0
- 17. 0
- 18. 0
- 19. 0
- 20. 0
- 21. 0
- 22. 0
- 23. 0
- 24. 0
- 25. 0
- 26. 0
- 27. 0
- 28. 0
- 29. 0
- 30. 0
- 31. 0
- 32. 0
- 33. 0
- 34. 0
- 35. 0
- 36. 0
- 37. 0
- 38. 0
- 39. 0
- 40. 0
- 41. 0
- 42. 0
- 43. 0
- 44. 0
- 45. 0
- 46. 0
- 47. 0
- 48. 0
- 49. 0
- 50. 0
- 51. 0
- 52. 0
- 53. 0
- 54. 0
- 55. 0


Sensor-based Human Action Recognition

- 01. 0
- 02. 0
- 03. 0
- 04. 0
- 05. 0
- 06. 0
- 07. 0
- 08. 0
- 09. 0
- 10. 0
- 11. 0
- 12. 0
- 13. 0
- 14. 0
- 15. 0
- 16. 0
- 17. 0
- 18. 0
- 19. 0
- 20. 0
- 21. 0
- 22. 0
- 23. 0
- 24. 0
- 25. 0
- 26. 0
- 27. 0
- 28. 0
- 29. 0
- 30. 0
- 31. 0
- 32. 0
- 33. 0
- 34. 0
- 35. 0
- 36. 0
- 37. 0
- 38. 0
- 39. 0
- 40. 0
- 41. 0
- 42. 0
- 43. 0
- 44. 0
- 45. 0
- 46. 0
- 47. 0
- 48. 0
- 49. 0
- 50. 0
- 51. 0
- 52. 0
- 53. 0
- 54. 0
- 55. 0

데이터셋 간접 검증 - 기타 인식 엔진 (물체검출)



데이터셋 간접 검증 - 기타 인식 엔진 (휴먼 검출/추적)



The image displays two side-by-side video frames of a kitchen scene. In both frames, a person wearing a dark long-sleeved shirt and dark pants is bent over, appearing to be cleaning or interacting with the floor. The kitchen features a white countertop, a sink, a stove, and a refrigerator. A wooden dining table with chairs is visible in the foreground. The background includes a window and some kitchen items on the counter.

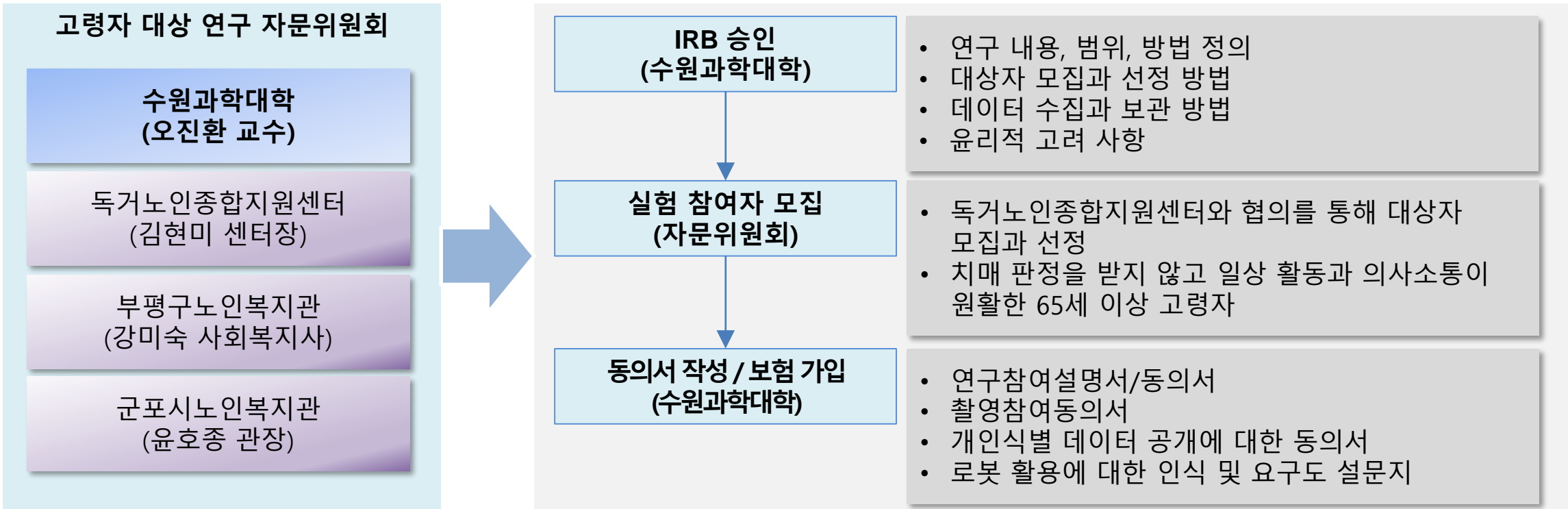
Below the video frames, there is a small, square inset image on the left side of a black rectangular area. This inset image shows a smaller version of the person from the video frames, positioned in the same kitchen setting. The rest of the black area is empty, representing a tracking or detection output.



데이터 공개를 위한 준비 활동과 성과 (1/2)

개인 데이터 수집 공개를 위한 수행 체계 구축

- 로봇을 이용한 사용자 연구 전문가와 고령자 유관 기관으로 구성된 자문위원회 운영 (2017 ~)
- 개인정보 수집과 고령자 대상 연구 수행을 위한 IRB 승인 (2017년 이후 매년 심의 승인)
- 수집 데이터 공개를 위한 동의서 양식 정의 활용: 개인정보 공개/비공개 선택 가능



데이터 공개를 위한 준비 활동과 성과 (2/2)

데이터 공개를 위한 관리 체계 구축

- 데이터 수집과 보관은 주관기관인 ETRI가 총괄 관리
- 참여기관은 데이터 보관과 활용을 위한 “개인정보보호서약서”에 서명하고 규약 준수
- 일반 공개할 데이터와 개인정보보호 대상 데이터를 구분하여 공개 관리

개인정보보호서약서 주요 내용

- 본 과제 발생 데이터를 제3자에 제공 시 ETRI의 사전 승인 필요
- 참여기관은 개인정보 보유현황을 보관 관리
- 참여기관은 개인정보 보유현황, 제3자 제공 시 ETRI의 사전 승인서와 제공 내역을 보관 관리
- ETRI는 참여기관별 연구데이터 보유 현황 파악 관리
- 본 과제 발생 데이터는 과제 종료(2021년) 이후 5년인 2026년에 폐기를 원칙으로 함
- 개인정보 유출로 인한 모든 법적 책임은 유출 기관에 있음

데이터 공개 방식 구분

공개 방식	방법	해당 데이터
제한 공개	EULA ¹⁾ 접수 후 다운로드 링크 제공 (피제공자 정보 기록 관리)	<ul style="list-style-type: none"> • 고령자 실주거 환경 장기 수집 데이터 • 고령자 55종 행동 데이터 • 고령자 교류 행위 데이터 등
일반 공개	개방 다운로드 링크 제공	<ul style="list-style-type: none"> • TED 발화 제스처 데이터 • 가상환경 데이터 (외형특징, 행동, 고령자 특화 사물) 등

1) End User License Agreement

행동인식 데이터셋의 공개 의의 및 계획

- 로봇지능 개발을 위한 **고령자 특화 대규모 실환경 및 가상합성** 데이터베이스 토털 패키지 구축
 - APT 테스트베드 환경
 - 고령자 실거주 리빙랩 환경
 - 가상 합성 데이터 셋 및 가상 주거 환경
- **세계 최초의 로봇환경에서의 고령자 특화 대용량 데이터베이스** 공개
- 대규모 기계 학습용 DB를 학교, 연구소, 산업계에 공개함으로써 관련 연구 커뮤니티의 성장에 기여

데이터 셋	특 징	공개 내용	규모	공개 시점
행동인식 (@APT테스트베드)	<ul style="list-style-type: none"> • Action Recognition • 로봇시점, 고령자 대상 (100명) • 멀티모달 데이터 	<ul style="list-style-type: none"> • 2D RGB video • 3D depth image sequence • Skeleton data sequence 	112,658 set (약 6TB)	2019.11
행동인식/검출 (@리빙랩)	<ul style="list-style-type: none"> • Action Recognition & Detection • 로봇 시점, 고령자 대상 • 실제 고령자가 사는 주거 환경 	<ul style="list-style-type: none"> • 2D RGB video • 3D depth image sequence • Skeleton data sequence 	300 시간 분량	2020 2021

2019년 공개 대상 데이터셋

순번	데이터 셋	구축기관
1	로봇 발화 제스처 생성을 위한 TED 제스처 데이터 셋 - 1,766 TED 영상, 106.1 시간	한국전자통신연구원
2	고령자 관련 일상 물건 인스턴스 데이터 셋 - 대상물체 15종, 총 830개 RGBD 동영상	한국전자통신연구원
3	AIR-Act2Act: 노인과의 소셜 상호작용 데이터 셋 - 고령자 100 그룹, 총 15,000 RGBDS 셋	한국전자통신연구원
4	교류행위 시점 인식을 위한 데이터 셋 - 고령자 100명, 총 33시간 분량 영상/음성 데이터	한국전자통신연구원
5	웨어러블 센서기반 일상생활 행동 데이터 셋 - 리빙랩 8 장소, 총 168,890 개의 Motion/wearable/IoT 센서데이터	전자부품연구원
6	생활패턴 추론을 위한 지각 정보 기반 데이터 셋 - 리빙랩 3 장소, 총 660 시간 분량의 percept data 및 의도 태깅 데이터	송실대학교
7	휴먼케어 로봇 기술 개발을 위한 고령자 음성 데이터 셋 - 전국 10개 지역, 고령자 대화체 음성/텍스트 데이터 총 400시간	마인즈랩

감사합니다

AiR
Artificial Intelligence for Robots

