

# 스마트 홈에서의 재실자 활동 인식

## - 센서 기반 재실자 활동 인식 확장 연구의 필요성 -

### Occupant Activity Recognition in Smart Homes

#### - The Need for Expanded Research on Sensor-Based Human Activity Recognition -

○장 재 영\* 김 현 우\*\* 차 승 현\*\*\*  
Jang, Jae Young Kim, Hyun Woo Cha, Seung Hyun

#### Abstract

Based on recent advancements in IoT technology, this paper conducts a comprehensive analysis of smart home technology, exploring the technical potential to offer diverse services such as healthcare monitoring, daily living assistance, energy management, and home security for occupants. With a specific focus on Human Activity Recognition (HAR) technology, it evaluates the trends and effectiveness of both Vision-based and Sensor-based HAR methods. Through detailed descriptions and analysis of Vision-based HAR and Sensor-based HAR, the paper explores the challenges and opportunities for HAR technology in smart home environments. In conclusion, it suggests the need for further research into new datasets and recognition models, especially for Sensor-based HAR, to facilitate the development of enhanced smart home environments, thus potentially improving the quality of life for occupants.

키워드 : 스마트 홈, 인간 활동 인식, 딥러닝, 활동 데이터셋, 인식 모델

Keywords : Smart Home, Human Activity Recognition, Deep Learning, Activity Dataset, Recognition Model

#### 1. 서론

최근 Internet of Things (IoT) 기술이 발전함에 따라, 스마트 홈에서 다양한 스마트 디바이스와 센서의 연결이 가능해졌다. 이를 통해, 재실자 활동 기반으로 스마트 디바이스를 제어하여 재실자들에게 헬스케어 모니터링, 일상생활 보조, 에너지 관리, 집 보안 등의 서비스를 제공한다[1]. 예를 들어, 헬스케어 모니터링 서비스는 고령자나 환자의 실시간 움직임 인식을 통한 활동량 분석 및 위험 감지[2]를 통해 환자나 고령자에게 안전하고 독립적인 생활을 지원한다. 일상생활 보조 서비스는 재실자들의 출입문 통과 분석을 통해 온도 및 조명 조절을 자동화하여[3] 재실자들에게 편안한 서비스를 제공한다. 에너지 관리 서비스는 Home Energy Management System (HEMS)[4] 기술을 활용하여 스마트 홈 내에서 재실자들의 에너지 사용을 효율적으로 관리한다. 집 보안 서비스는 개방형 인터넷으로 연결된 스마트 디바이스가 여러 해커의 공격 및 바이러스 감염으로 인해 발생할 수 있는 개인 정보 유출 문제를 블록체인 기술을[5] 활용하여 방지한다.

이러한 서비스 제공을 위한 스마트 홈의 핵심 기술은 재실자 활동 인식이며, 이는 Human Activity Recognition (HAR) 연구에 기반을 둔다. HAR은 디지털 디바이스로부터 측정된 시그널을 분석해 인간의 활동을 인식하는 연구이다[6]. HAR의 연구는 두 가지 카테고리로 나뉜다: Vision-based HAR, Sensor-based HAR.

Vision-based HAR은 카메라 영상을 사용하여 인간의 활동을 분류하고 인식하는 기술로, 인간의 활동을 Gesture, Action, Behavior, Interaction, Group Activity, 그리고 Event의 6가지 클래스로 분류하여 인식한다. 해당 기술은 스마트 홈의 주요 기술로써 활용되어 실시간으로 재실자들의 다양한 활동을 인식한다. 하지만, 높은 연산 처리 비용, 개인차로 인한 결과 편향성, 그리고 주거 공간 내 카메라 설치에 따른 프라이버시 문제로 실제 주거 공간 내 Vision-based HAR 기술 활용에는 한계가 있다.

Sensor-based HAR은 센서 데이터로부터 인간의 활동을 분석하고 인식하는 기술로 적은 연산 처리량과 높은 정확도를 수반한다. 특히, 해당 기술은 Vision-based HAR의 프라이버시 문제를 해결하여 스마트 홈에 주로 활용되어 재실자의 행동을 인식한다. 하지만, 현재 인식 가능한 재실자 활동의 종류가 제한적이고 인식 모델의 일반화 문제로 인해, Sensor-based HAR 기술이 적용된 스마트 홈은 재실자들에게 제한적인 서비스를 제공한다.

\* 한국과학기술원 메타버스대학원 석사과정

\*\* 한국과학기술원 문화기술대학원 석사과정

\*\*\* 한국과학기술원 문화기술대학원 부교수, 건축학박사

(Corresponding author : Department Culture of Technology, Advanced Institute of Science and Technology, shcha@kaist.ac.kr)  
이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2023-00212993).

## 2. 선행연구 고찰

### 2.1 Smart Homes

스마트 홈은 IoT 기술을 활용하여 주거 공간 내 다양한 장치를 자동화하고 상호 연결하여 편의성, 안전성, 에너지 효율성을 향상시킨 새로운 주거 시스템이다. 특히, 고령화 사회 진입, 개인 맞춤형 서비스 수요 증가, 에너지 효율성에 대한 관심 증가[7]는 스마트 홈 기술 발전을 더욱 가속화시키고 있다.

스마트홈 기술의 핵심 요소는 연결성, 자동화, 그리고 보안으로 구성되며, 각 핵심 요소별 기술 개발 진행 상황은 다음과 같다:

1) 연결성: 스마트 홈은 다양한 장치들이 서로 연결되어 정보를 주고받는 시스템으로, 강화된 유무선 통신 (Wi-Fi, Bluetooth, Zigbee, 등) 및 인터넷 연결 기술이 개발되고 있다.

2) 자동화: 스마트 홈은 재실자의 활동 및 주거 환경 조건을 인식하여 장치를 자동 제어하기 위해, 재실자의 정확한 활동 인식 및 패턴 도출 기술이 개발되고 있다.

3) 보안: 스마트 홈 기기와 시스템은 재실자의 개인 정보와 활동 데이터 등의 중요한 정보를 수집하고 처리한다. 이에 따라, 해당 데이터 보안을 위해 다양한 기술이 개발되고 있다.

최근 딥러닝 기술에 힘입어 고도화된 인간 활동 인식 기술이 개발됨에 따라, 스마트 홈에 적용되어 재실자들에게 편안한 서비스를 제공하고 있다. 하지만, 인간 활동 인식 종류가 제한적임에 따라, 재실자들에게 활동 지원을 위한 다양한 서비스 제공에는 여전히 한계가 있다.

### 2.2 HAR (Human Activity Recognition)

HAR은 인간의 활동을 식별하고 이해하는 것으로 다양한 분야에서 활용되고 있다. 이는 특히 노인케어나 헬스케어에 주로 활용되는 보조적인 기술로써, 스마트 홈과의 접목을 통해 재실자 활동 인식을 기반으로 다양한 서비스를 제공 가능하게 한다.

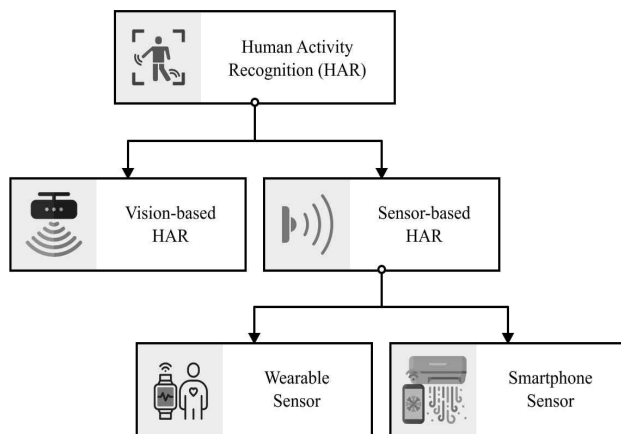


그림1. Human Activity Recognition 유형

### 2.2.1 Vision-based HAR

Vision-based HAR은 컴퓨터 비전 기술을 활용한 이미지 및 비디오 데이터 분석을 통해 인간의 활동을 인식하는 연구이다. 해당 연구는 인간의 활동을 Gesture, Action, Behavior, Interaction, Group Activity, 그리고 Event와 같이 6가지 클래스로 분류한다[8]. Gesture는 비언어적인 의사소통의 한 형태로, 손바닥을 마주치며 하이파이브를 하는 행위를 포함한다. Action은 작업을 완료하기 위한 몸의 움직임으로써, 걷기, 뛰기, 앉기 등이 이에 해당된다. 또한, Behavior는 화난 표정으로 목소리를 높이는 것과 같이, 사람의 전반적인 행동과 반응을 의미한다. Interaction은 둘 이상의 물체나 대상이 서로 영향을 주고받는 활동으로, 두 사람이 대화하거나 요리하는 것을 예로 들 수 있다. Group Activity는 다수의 사람에 의해 행해지는 활동으로, 농구, 축구 등이 이에 속한다. Event는 사람이 걷는 것과 같은 구체적인 활동이 아닌, 특정한 상황이나 사건(맥락)을 의미한다.

Vision-based HAR 연구에 사용되는 데이터셋은 RGB 및 RGB-D 카메라를 통해 수집된 Single View나 Multi View 데이터로 구성된다. HMDB51[9] 데이터셋은 RGB Single View로, 51가지의 Action, Behavior, Interaction, 그리고 Group Activity를 포함한다. UTKinect[10] 데이터셋은 RGB-D Single View로, 10가지의 서로 다른 Action을 포함한다. IXMAS[11] 데이터셋은 RGB Multi View로, 13가지의 서로 다른 Action을 포함한다. NTU RGB+D[12] 데이터셋은 RGB-D Multi View로, 60가지의 서로 다른 Action, Behavior, 그리고 interaction 데이터를 포함한다. 연구자들은 연구 목적에 맞게 데이터셋을 선택하고, 딥러닝 모델 특성에 맞춰 데이터를 전처리하여 HAR 연구를 진행한다.

딥러닝 모델은 여러 데이터 (e.g., 이미지, 텍스트, 사운드, 등)의 복잡한 패턴을 인식하여 인사이트를 도출하고 결과를 예측하는 도구로써, Vision-based HAR에도 흔히 활용되고 있다. Perrett, et al.[13]은 HMDB51 데이터셋을 활용하여 이미지 데이터의 시간적 관계를 모델링하기 위해 Transformer 모델을 변형한 Temporal-Relational CrossTransformers (TRX)를 활용하여 인간의 활동을 인식했다. Ghojogh, et al.[14]은 UTKinect 데이터셋을 Fisherposes 방법을 활용하여 각 프레임에서 포즈를 인식하고 구별하는 특징 벡터로 전처리한 후, Hidden Markov Model (HMM)을 사용하여 포즈 기반으로 인간 활동을 분류했다. Sabater, et al.[15]은 NTU RGB+D 데이터셋을 시간적 동작 세그먼트의 고정된 길이 벡터로 전처리하고, CNN (Convolutional Neural Network)의 변형인 Temporal Convolutional Network (TCN)를 활용하여 인간의 활동을 인식했다. Putra, et al.[16]은 feature extractor와 discriminator를 단일 모델로 통합하여 Deep Neural Network (DNN)을 구축하고, IXMAS 데이터셋을 사용하여 모델을 학습시켜 인간의 활동을 인식했다.

딥러닝 모델을 활용한 Vision-based HAR은 환자와 고령자들의 활동을 실시간 모니터링하는데 활용되어 헬스케어 및 위험 감지에 많은 기여를 하고 있다. 그럼에도 불구하고, 재실자 개인차 (e.g., 나이, 체형, 능력)에 따른 편향된 결과

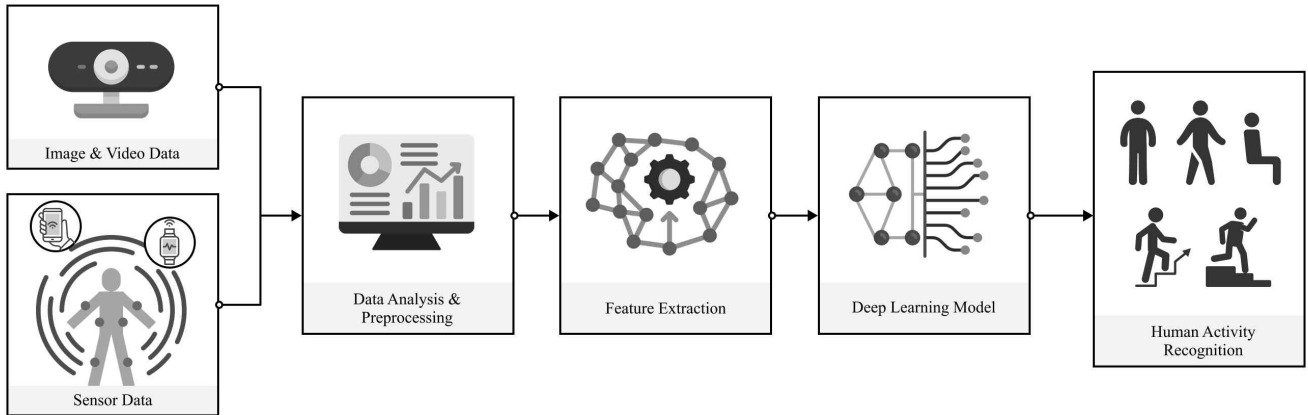


그림2. 일반적인 Human Activity Recognition 흐름도

의 한계와 맥락 인식 부족에 따른 정확한 활동 인식 및 예측의 한계가 존재한다. 또한, 주거 공간 내 재실자 활동 관찰을 위한 카메라 설치에 프라이버시 침해 문제와 직결되며, 이는 일상 주거 공간 내 Vision-based HAR 적용에 대한 어려움으로 이어진다.

### 2.2.2 Sensor-based HAR

Sensor-based HAR은 센서 디바이스를 통해 시그널을 확보하고 분석하여 인간의 행동을 인식하는 연구이다. 해당 연구는 기존 Vision-based HAR의 문제점인 고정된 카메라 위치에 따른 환경적 제약, 대상 이외의 비대상자를 포함하는 불필요한 정보 수집, 카메라를 통한 프라이버시 침해, 그리고 이미지 데이터 전처리의 복잡성과 높은 비용 문제를 센서의 활용을 통해 극복했다.

Sensor-based HAR 연구에 사용되는 데이터셋은 스마트 폰에 내장된 센서나 wearable 디바이스를 통해 수집 가능하며, 가속도계, 자이로스코프, 지자기계 데이터가 조합되어 구축된다. 예를 들어, HAR using smartphone[17] 데이터셋은 스마트 폰에 내장된 센서를 통해 수집되며 인간의 6가지 활동 (e.g., Walking, Sitting, 등)에 대한 가속도계 및 자이로스코프 데이터로 구성된다. PAMAP2[18] 데이터셋은 wearable 디바이스를 통해 수집되며 인간의 18가지의 활동 (e.g., Watching TV, Computer Work, 등)에 대한 가속도계 및 자이로스코프 데이터로 구성된다. 해당 데이터셋은 다양한 windowing 기법 (e.g., 행동 기반, 시간 기반, 센서 기반 등)이 활용되어 전처리 된다. WISDM[19] 데이터셋은 스마트 폰에 내장된 센서를 통해 수집되며 인간의 6가지 활동 (e.g., Jogging, Standing, 등)에 대한 가속도계 데이터로 구성된다. UniMiB SHAR[20] 데이터셋은 스마트 폰에 내장된 센서를 통해 수집되며 인간의 17가지의 활동 (e.g., Falling, Hitting Obstacle, 등)에 대한 가속도계 데이터로 구성된다.

전처리된 데이터셋은 시간적 의존성을 가지는 타임 시리즈 데이터로, 딥러닝 모델 중 시퀀스 데이터를 처리하기에 용이한 RNN (Recurrent Neural Network), LSTM (Long-Short Term Memory), GRU (Gated Recurrent Unit) 등이 학습에 활용되어

인간 활동 인식 연구에 사용된다. Tan et al.[21]은 데이터의 길이를 고정하는 Sliding Window 기법과 Overlapping 기법을 활용하여 HAR using smartphone 데이터셋을 전처리하고, GRU, CNN 등을 활용하여 인간의 활동을 6가지로 분류했다. 또한, Wan et al.[22]은 Sliding Window 기법을 활용하여 PAMAP2 데이터셋을 전처리한 후, training data와 test data를 7:3 비율로 나누어 LSTM, Bi-LSTM, CNN 등을 학습하고 검증했다. Ghatge, et al.[23]은 노이즈 제거, windowing과 segmentation 기법을 활용하여 WISDM 데이터셋을 전처리하고, 하이브리드 모델 (e.g., CNN + LSTM, CNN + Random Forest, 등)을 학습시켜 인간의 활동을 분류했다. Gattulli, et al.[24]은 UniMiB SHAR 데이터셋을 활용하여 CNN, Bi-LSTM 모델을 학습하고, 모델을 비교 평가하여 최적의 인간 활동 인식 모델을 구축했다.

Sensor-based HAR은 센서 디바이스를 통해 수집된 시그널 데이터를 활용함으로써 기존 Vision-based HAR의 문제들을 극복했다. 하지만, 통제된 실험실 환경에서 수집된 활동 데이터셋은 모델의 일반화를 제한했고, 이는 실제 주거 환경에 적용하기에 어려움이 있음을 시사한다. 또한, Sensor-based HAR 연구는 단순한 활동 인식에 초점이 맞춰져 있어, 일상 생활에서 수행 가능한 복잡한 활동 (e.g., 요리하기, 반려동물 먹이주기, 운동하기, 등) 인식에는 한계가 있다.

### 3. 결론

본 논문에서는 스마트 홈과 Human Activity Recognition 연구에 대한 기술 개발 동향 및 한계점을 심층적으로 분석했다. 이를 통해, 스마트 홈 기술의 발전이 재실자들에게 헬스케어 모니터링, 일상생활 보조, 에너지 관리, 집 보안 등의 다양한 서비스를 제공함에 있어 중요한 역할을 하고 있음을 확인했다. 특히, 스마트 홈 기술 중 재실자 활동 인식 기술에 핵심인 HAR 기술은 스마트 홈에 적용되어 재실자들에게 안전하고 편안한 서비스 제공이 가능함을 보였다. 하지만, Vision-based HAR은 주거 공간 내 카메라 설치에 따른 개인의 프라이버시 침해 문제에 직면했으며, Sensor-based HAR은 기존 Vision-based HAR 문제를 해결했으나, 인식 모델의 일반화 문제와 인식 가능한 재실자 활동

의 종류가 한정적임에 따라, 실제 스마트 홈에 활용되었을 때 제공 가능한 서비스가 제한적일 수 있음을 확인했다. 따라서, 재실자들에게 활동 보조를 위한 다양한 서비스를 제공하기 위해 Sensor-based HAR의 문제점을 고려하여 새로운 데이터셋 및 인식 모델 구축에 대한 연구가 필요하다. 이를 통해, 더 나은 스마트 홈 환경이 구축 가능하며, 이는 재실자들의 삶의 질을 향상시킬 것으로 기대된다.

## 참고문헌

- Liciotti, D., Bernardini, M., Romeo, L., & Frontoni, E. (2020). A sequential deep learning application for recognising human activities in smart homes. *Neurocomputing*, 396, 501-513.
- Kulurkar, P., kumar Dixit, C., Bharathi, V. C., Monikavishnuvarthini, A., Dhakne, A., & Preethi, P. (2023). AI based elderly fall prediction system using wearable sensors: A smart home-care technology with IOT. *Measurement: Sensors*, 25, 100614.
- Fontes, F., Antão, R., Mota, A., & Pedreiras, P. (2021). Improving the ambient temperature control performance in smart homes and buildings. *Sensors*, 21(2), 423.
- Mahapatra, B., & Nayyar, A. (2022). Home energy management system (HEMS): Concept, architecture, infrastructure, challenges and energy management schemes. *Energy Systems*, 13(3), 643-669.
- Ratkovic, N. (2022). Improving home security using blockchain. *International Journal of Computations, Information and Manufacturing (IJCIM)*, 2(1), 1.
- Dentamaro, V., Gattulli, V., Impedovo, D., & Manca, F. (2024). Human Activity Recognition with Smartphone-Integrated Sensors: A Survey. *Expert Systems with Applications*, 123143.
- Liciotti, D., Bernardini, M., Romeo, L., & Frontoni, E. (2020). A sequential deep learning application for recognising human activities in smart homes. *Neurocomputing*, 396, 501-513.
- Saleem, G., Bajwa, U. I., & Raza, R. H. (2023). Toward human activity recognition: a survey. *Neural Computing and Applications*, 35(5), 4145-4182.
- Kuehne, H., Jhuang, H., Garrote, E., Poggio, T., & Serre, T. (2011, November). HMDB: a large video database for human motion recognition. In *2011 International conference on computer vision* (pp. 2556-2563). IEEE.
- Xia, L., Chen, C. C., & Aggarwal, J. K. (2012, June). View invariant human action recognition using histograms of 3d joints. In *2012 IEEE computer society conference on computer vision and pattern recognition workshops* (pp. 20-27). IEEE.
- Weinland, D., Ronfard, R., & Boyer, E. (2006). Free viewpoint action recognition using motion history volumes. *Computer vision and image understanding*, 104(2-3), 249-257.
- Shahroudy, A., Liu, J., Ng, T. T., & Wang, G. (2016). Ntu rgb+d: A large scale dataset for 3d human activity analysis. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1010-1019).
- Perrett, T., Masullo, A., Burghardt, T., Mirmehdi, M., & Damen, D. (2021). Temporal-relational crosstransformers for few-shot action recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 475-484).
- Ghojogh, B., Mohammadzade, H., & Mokari, M. (2017). Fisherposes for human action recognition using kinect sensor data. *IEEE Sensors Journal*, 18(4), 1612-1627.
- Sabater, A., Santos, L., Santos-Victor, J., Bernardino, A., Montesano, L., & Murillo, A. C. (2021). One-shot action recognition in challenging therapy scenarios. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2777-2785).
- Putra, P. U., Shima, K., & Shimatani, K. (2018, April). Markerless human activity recognition method based on deep neural network model using multiple cameras. In *2018 5th International Conference on Control, Decision and Information Technologies (CoDIT)* (pp. 13-18). IEEE.
- Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013, April). A public domain dataset for human activity recognition using smartphones. In *Esann* (Vol. 3, p. 3).
- Reiss, A., & Stricker, D. (2012, June). Introducing a new benchmarked dataset for activity monitoring. In *2012 16th international symposium on wearable computers* (pp. 108-109). IEEE.
- Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2), 74-82.
- Micucci, D., Mobilio, M., & Napolitano, P. (2017). Unimib shar: A dataset for human activity recognition using acceleration data from smartphones. *Applied Sciences*, 7(10), 1101.
- Tan, T. H., Wu, J. Y., Liu, S. H., & Gochoo, M. (2022). Human activity recognition using an ensemble learning algorithm with smartphone sensor data. *Electronics*, 11(3), 322.
- Wan, S., Qi, L., Xu, X., Tong, C., & Gu, Z. (2020). Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications*, 25(2), 743-755.
- Hybrid deep learning approaches for smartphone sensor-based human activity recognition. *Multimedia Tools and Applications*, 80(28), 35585-35604.
- Gattulli, V., Impedovo, D., Pirlo, G., & Sarcinella, L. (2023). Human activity recognition for the identification of bullying and cyberbullying using smartphone sensors. *Electronics*, 12(2), 261.