

다중 입출력 FMCW 레이다를 활용한 합성곱 신경망 기반 사람 동작 인식 시스템

CNN Based Human Activity Recognition System Using MIMO FMCW Radar

김준성 · 심재용 · 장수림 · 임승찬 · 정운호*
한국항공대학교 항공전자정보공학부

Joon-sung Kim · Jae-yong Sim · Su-lim Jang · Seung-chan Lim · Yunho Jung*

School of Electronics and Information Engineering, Korea Aerospace University, Gyeonggi-do, 10540, Korea

[요약]

본 논문에서는 다중 입출력 주파수 변조 연속파 (MIMO FMCW; multiple input multiple output frequency modulation continuous wave) 레이다 기반 HAR (human activity recognition) 시스템의 설계 및 구현 결과를 제시하였다. 다중 입력 다중 출력 레이다 센서를 통한 포인트 클라우드 데이터를 활용하여 HAR 시스템을 구현하면 사생활 보호와 함께, 안전성 및 정확성 측면에서 장점이 있다. 본 논문에서는, MIMO FMCW 레이다 센서로부터의 포인트클라우드 데이터 기반 HAR을 위해 PointPillars와 DS-CNN (depthwise separable convolutional neural network)을 기반으로 최적 경량 네트워크를 개발하였다. 경량화된 네트워크를 통해 고해상도 포인트 클라우드 데이터를 처리하여 높은 인식 정확도와 함께 효율성을 달성하였다. 결과적으로, 98.27%의 정확도와 11.27M Macs (multiply-accumulates) 연산 복잡도로 구현 가능성을 확인하였다. 또한, 개발한 모델을 라즈베리파이(Raspberry-Pi) 시스템에 구현하여 최대 8 fps의 속도로 포인트 클라우드 데이터 처리가 가능함을 확인하였다.

[Abstract]

In this paper, a human activity regeneration (HAR) system based on multiple input multiple output frequency modulation continuous wave (MIMO FMCW) radar was designed and implemented. Using point cloud data from MIMO radar sensors has advantages in terms of privacy, safety, and accuracy. For the implementation of the HAR system, a customized neural network based on PointPillars and depthwise separate convolutional neural network (DS-CNN) was developed. By processing high-resolution point cloud data through a lightweight network, high accuracy and efficiency were achieved. As a result, the accuracy of 98.27% and the computational complexity of 11.27M multiply-accumulates (Macs) were achieved. In addition, the developed neural network model was implemented on Raspberry-Pi embedded system and it was confirmed that point cloud data can be processed at a speed of up to 8 fps.

Key word : MIMO FMCW Radar, Human Activity Recognition, Point Cloud, PointPillars, CNN.

<http://dx.doi.org/10.12673/jant.2024.28.4.428>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 1 August 2024; Revised 19 August 2024
Accepted (Publication) 25 August 2024 (30 August 2024)

*Corresponding Author; Yunho Jung

Tel: +82-2-300-0133

E-mail: yjung@kau.ac.kr

I. 서론

HAR (human activity recognition) 기술은 현대 사회에서 사용자의 안전과 건강 관리를 돕는 중요한 역할을 한다. 이러한 추세에 따라 다양한 센서로부터 수집된 데이터에서 인공지능 (AI; artificial intelligence)을 사용하여 활동을 식별하는 HAR 시스템 연구가 활발히 진행되고 있다 [1]-[4]. 특히 고령화 사회에 발 맞추어, 노년층의 일상생활 속 건강 관리에 도움을 줄 수 있고, 응급 상황에서의 신속한 대응을 도울 수 있는 HAR 시스템의 중요성이 부각되고 있다 [5],[6].

하지만, 기존 HAR 시스템은 주로 일상생활에서 자주 발생하지 않고 분류가 쉬운 특정 동작에 집중되거나 [1],[2], 일상생활에서 자주 사용하는 동작임에도 다양한 동작을 분류하지는 못하였다 [3]-[11]. 이에 본 논문은 일상생활에서 자주 발생하는 다양한 동작을 더 정확하고 효과적으로 인식하고 분류할 수 있는 HAR 시스템을 개발하는 것을 목표로 하여 기존 연구의 한계를 극복할 수 있는 새로운 시스템을 제안한다.

여러 동작을 분류하는 HAR 시스템을 구현하기 위하여 다양한 센서가 사용되고 있는데, 그중 대표적으로 카메라 센서가 주로 사용되고 있다 [7]-[9]. 카메라 센서는 고해상도 이미지를 제공하여 동작을 정확하게 인식할 수 있는 장점이 있지만, 개인의 일상생활을 지속적으로 모니터링하는 과정에서 사생활 침해 우려가 있다. 하지만 레이더 센서는 개인의 신체 이미지를 직접적으로 촬영하지 않는다는 측면에서 카메라 센서와 비교해 사생활 보호 측면에서 유리하다고 할 수 있다. 또한 라이다 센서와 비교해 환경에 대한 민감도가 낮아 다양한 환경에서 안정적으로 작동할 수 있다. 추가적으로, 레이더 센서는 고속 데이터 처리가 가능하여 실시간 응용 분야에 적합하고, 낮은 전력 소비와 저렴한 비용으로 효율적인 시스템 구성이 가능하다는 장점이 있다. 하지만 레이더 센서는 실내환경과같이 복잡한 환경에서 반사파가 다중 경로로 전파되면서 신호 간섭이 발생할 수 있고, 해상도가 제한적이기 때문에 작은 물체나 복잡한 형태의 물체를 정확하게 식별하는 데 어려움이 있다. 이러한 단점을 극복하기 위해, 본 논문에서는 다중 송수신 안테나를 통해 실내환경에서 타깃의 거리, 속도, 방향 등을 고해상도로 감지할 수 있고, 다중 경로 문제를 완화시킬 수 있는 MIMO (multiple input multiple output) 레이더를 채택하여 시스템에 적용하였다. 특히 MIMO 레이더는 고품질의 포인트 클라우드 데이터를 생성하여 알고리즘의 성능을 크게 향상시킬 수 있어 본 논문의 시스템에 선정하였다.

PointPillars는 3D 포인트 클라우드 데이터를 2D 이미지로 변환하여 공간 구조와 특징을 잘 보존하면서도 연산량을 줄여 실시간 처리가 필요한 응용 분야에 매우 적합한 기술이다. 이는 2D 합성곱 신경망(CNN; convolutional neural network)을 사용한 데이터 처리가 훨씬 효율적으로 이루어지게 한다. 따라서 본 논문에서는 PointPillars 내부의 백본(backbone) 구조를 정확도와 네트워크의 복잡성의 tradeoff 을 고려해 최적화하여 최적의

알고리즘 모델인 맞춤형 네트워크를 개발하여 선정하고, 이를 통해 정확도가 높은 동작 인식 시스템을 구현하였다.

선행 HAR 시스템 연구들은 주로 센서와 데스크톱 컴퓨터 간의 연결을 통해 데이터를 처리하여, 비용과 크기 면에서 한계가 있다 [10],[11]. 이에 본 논문에서는 센서를 라즈베리파이 (Raspberry-Pi)와 직접 연결하여 모바일 및 임베디드 디바이스 환경에서도 실시간 데이터 처리가 가능하도록 구현하였다.

본 논문의 II장에서는 시스템을 구성하는 MIMO 레이더 및 PointPillars와 DS-CNN (depthwise separable convolutional neural network)에 대해 설명한다. III장에서는 제안된 시스템의 전반적인 개요와 시스템에 사용된 데이터셋의 정의, 데이터의 전처리 과정 및 구성 등을 제시한다. IV장에서는 제안된 시스템의 성능평가 결과를 설명한다. V장에서는 시스템의 구현에 대해 설명하고, 끝으로 VI장에서 본 논문의 결론을 맺는다.

II. Backgrounds

2-1 MIMO FMCW 레이더

MIMO FMCW 레이더는 다중 송신 안테나와 다중 수신 안테나를 사용하여 높은 해상도와 정확도를 제공하는 레이더 시스템이다. 이 시스템은 다중 안테나 배열을 통해 공간적으로 더 많은 정보를 수집하여 물체의 위치와 속도를 정밀하게 측정할 수 있다. 각 송신 안테나에서 다른 신호를 송출하고, 수신 안테나에서 이를 받아 분석함으로써 신호 간섭을 줄이고 명확한 목표물 검출이 가능하다.

2-2 PointPillars

PointPillars는 포인트 클라우드 데이터를 효과적으로 처리하기 위해 개발된 딥러닝 모델로, 3D 공간의 점들을 'pillar'라는 3D 격자 구조로 변환하고, 각 pillar 내의 포인트들을 집계하여 특징 벡터를 생성한다 [12]. 이 특징 벡터는 Pillar Feature Net (PFN)에서 포인트의 희소성과 불규칙성을 해결하여 균일한 형식으로 변환한 후, 백본 네트워크에 입력되어 효율적인 특징 추출과 객체 탐지를 수행한다. PFN은 포인트 집계를 담당하며, 백본 네트워크는 주로 2D 합성곱 신경망을 활용하여 데이터의 공간적 구조와 특징을 유지하면서 연산량을 줄인다. 이러한 접근 방식은 계산 효율성을 크게 향상시키고, 실시간 처리 및 다양한 응용 분야에서 높은 성능을 제공한다.

2-3 DS-CNN

DS-CNN은 경량화되고 효율적인 합성곱 신경망 모델로 모바일 및 임베디드 시스템에서 실시간 처리 작업과 같이 제한된 계산 자원과 전력 소비를 고려하여 만들어졌다. 이는 깊이별 합성곱 (depthwise convolution)과 포인트별 합성곱 (pointwise

convolution)의 조합으로 구성된다.

깊이별 합성곱을 통해 각 채널별로 연산 되면서 하나의 입력 채널마다 하나의 필터가 적용된다. 깊이별 합성곱은 식 (1)의 계산 비용을 가진다.

$$C_{depthwise} = D_K \cdot D_K \cdot M \cdot D_F \cdot D_F \quad (1)$$

깊이별 합성곱의 경우 일반적인 합성곱에 비해 효율적이지만 새로운 특징을 생성하지 못하므로 1x1 합성곱을 통해 출력의 선형 결합을 연산하는 포인트별 합성곱을 수행해야 한다. 포인트별 합성곱은 식 (2)의 계산 비용을 가진다.

$$C_{pointwise} = M \cdot N \cdot D_F \cdot D_F \quad (2)$$

깊이별 합성곱과 포인트별 합성곱을 진행할 때의 연산량을 일반적인 합성곱의 연산량인 식 (3)과 비교하면 식 (4)로 알 수 있으며, 3x3의 합성곱 연산에서 연산량이 대략 8~9배 감소하게 된다.

$$C_{standard} = D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F \quad (3)$$

$$\frac{C_{depthwise} + C_{pointwise}}{C_{standard}} = \frac{1}{N} + \frac{1}{D_K^2} \quad (4)$$

D_K : kernel size, D_F ; output size

M : input channel, N : output channel

III. 제안된 시스템의 개요

3-1 시스템 구조

제안된 시스템의 개략적인 구조는 그림 1과 같다. 사용자가 MIMO 레이더 센서 앞에서 동작을 취하면, 센서는 해당 데이터를 수집한다. 수집된 레이더 데이터는 와이파이(Wi-Fi)를 통해 라즈베리파이로 전송된다. 용량이 한정된 라즈베리파이의 특성과, 필요한 데이터만 추출하여 사용하기 위해 Python 코드를 사용하여 데이터 전처리 과정을 거친다. 이후, 선정된 알고리즘 모델을 사용하여 사용자의 자세를 분류한 후, 그 결과를 사진

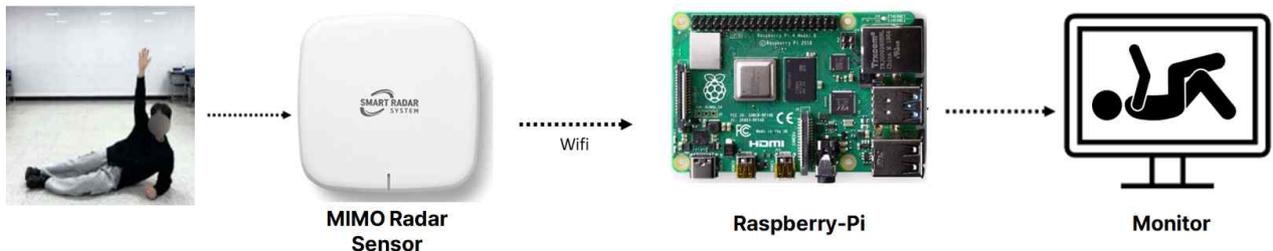


그림 1. 제안된 사람 동작 인식 시스템의 구조도

Fig. 1. Structure of the proposed human activity recognition system.

형태로 모니터에 출력한다.

이러한 시스템 구조는 실시간 데이터 수집과 처리를 가능하게 하여 시스템의 효율성을 높인다. 또한, 라즈베리파이의 소형화된 크기와 저전력 소모로 시스템을 간편하게 설치하고 운용할 수 있게 한다. 와이파이 통신을 통해 데이터를 전송함으로써 유선 연결의 번거로움을 줄이고, 더욱 유연한 설치 환경을 제공한다.

3-2 데이터셋 정의 및 구성

본 논문의 시스템에서는 Smart Radar System사의 RETINA-4SN 모델을 사용하였다. 해당 모델은 송신 안테나(TX)와 수신 안테나(RX)가 각각 3개와 4개로 구성된 MIMO 레이더 센서로, 4D 이미지를 제공하며, 77~81GHz의 주파수 대역을 통해 와이파이 통신을 지원한다. 해당 레이더 센서는 초당 20프레임의 속도로 데이터를 처리하며, 프레임당 평균 600~1,200개의 포인트가 추출된다. 각 포인트 데이터는 x, y, z 좌표와 도플러 값으로 구성된다.

데이터 측정 환경은 그림 2와 같이 실내에서의 일반적인 이동 범위와 높이를 고려하여 다양한 실생활 상황을 재현하기 위해 설정하였다. 이에 따라 센서 위치 (x, y, z) = (0, 0, 0)를 기준으로 하여 x는 가로축으로 -2 m ~ 2 m (총 4 m)까지, y는 세로축으로 0~5 m (총 5 m), z는 높이 축으로 -0.8 m~1.5 m (총 2.3 m)까지 측정 환경을 설정하였다.

남성1 (177 cm), 남성2 (179 cm), 여성1 (168 cm)을 대상으로 데이터를 추출하였으며, 실험 동작은 그림 3에 제시된 바와 같다.; 0. 서기 (Standing), 1. 의자에 앉기 (Sitting on a chair), 2. 바닥에 앉기 (Sitting on the floor), 3. 눕기 (Lying down), 4. 물건 줍기 (Picking up object), 5. 걸레질하는 동작 (Crawling), 6. 쪼그려 앉기 (Squatting), 7. 누워서 손들기 / 응급 상황 인식 (Lying wave hands), 8. 신문보기 (Reading newspaper), 9. 약 먹기 (Taking medicine)로 총 10개의 동작을 선정하였다. 본 논문에서 선정한 동작들은 노년층들의 일상생활에서 빈번히 발생하는 동작들뿐만 아니라, 응급 상황을 인식할 수 있는 동작까지 포함하여 구성하였다.

다양한 포인트 클라우드 데이터의 분포를 학습시키기 위해 각 동작당 약 1,800프레임의 데이터를 수집하여, 총 약 18,000프레임의 데이터를 확보하였다. 해당 데이터셋은 딥러닝 알고

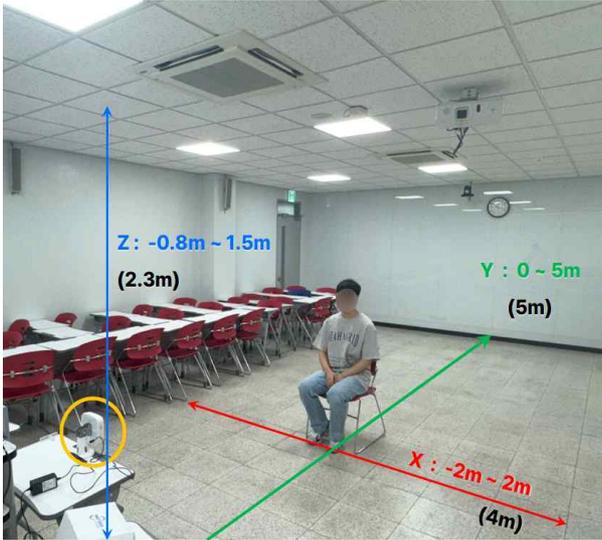


그림 2. 데이터셋 취득 실험 환경
Fig. 2. Experimental environment for dataset acquisition.

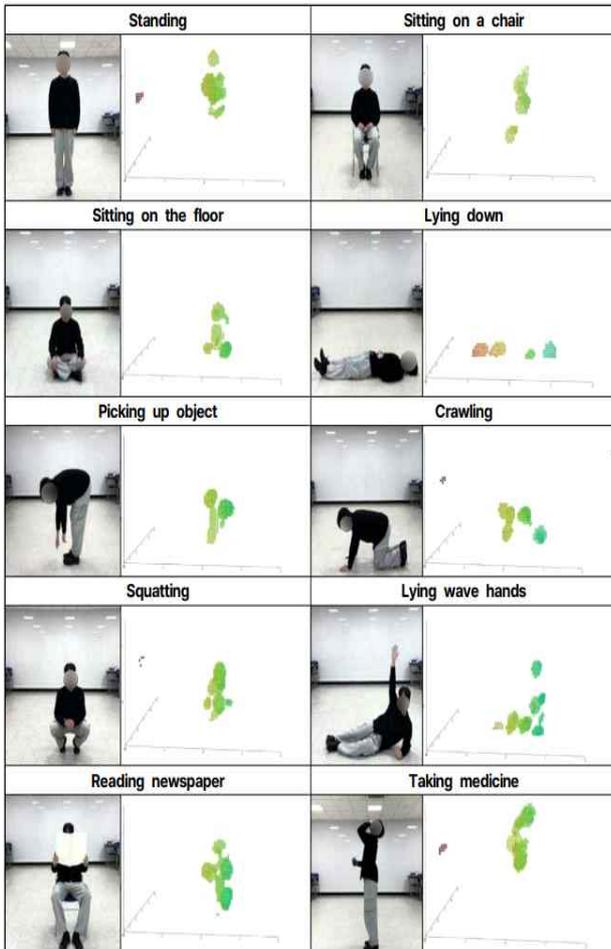


그림 3. 10가지 activity에 대한 point cloud data 예시
Fig. 3. Examples of point cloud data for 10 activities.

리듬 성능평가를 위해 90%는 학습용으로, 10%는 추론용으로 나누어 사용되었다.

3-3 데이터 전처리 과정

와이파이로 연결된 레이더 센서를 ROS(robot operating system)를 통해 센서와 노드 간에 연속적인 데이터 전송이 가능하게 한다. 이때 라즈베리파이의 제한된 메모리 용량 문제를 해결하고, 연속적으로 수집된 데이터 중 필요한 데이터만을 PointPillars에서 사용하기 위해 Python 코드를 활용하여 데이터 전처리 과정을 진행한다.

수집되는 포인트 클라우드 데이터는 매우 큰 용량을 차지하기 때문에, 메모리 오버플로우를 방지하고 실시간 처리를 가능하도록 본 논문에서는 수신된 데이터 중 최근 50프레임만 유지하고 이전의 데이터는 순차적으로 제거하는 방식을 채택하였다. 실시간 처리 시스템에서 모든 데이터를 분석하는 것은 불필요한 계산 부담을 야기할 수 있다. 따라서 최근 50프레임 중 동작 식별을 위해 필요한 데이터만을 추출하는 전략을 사용하였다. 구체적으로, 동작 감지 및 식별을 위해 최신 30프레임의 데이터를 선택하여 처리한다. 이 방식은 최근의 변화와 움직임을 반영할 수 있을 만큼 충분한 데이터를 제공하면서도, 불필요한 과거 데이터를 제외하여 계산 비용을 줄이는 효과를 얻을 수 있다. 이를 통해 실시간 데이터 처리 시스템의 효율성을 유지하면서도 신뢰성 있는 동작 인식이 가능하다.

또한 본 논문에서는 처리해야 할 데이터양을 줄이고, 효율적인 메모리 관리 및 처리성을 확보하기 위해 수집된 포인트 클라우드 데이터에서 각 포인트의 x, y, z 좌표와 도플러 값만 선별적으로 추출하였다. 각 포인트의 x, y, z 좌표는 객체의 공간적 위치를, 도플러 값은 객체의 속도를 추정하는 데 사용된다.

포인트 클라우드 데이터는 객체 인식과 같은 딥러닝 모델에서 정확한 처리를 위해 높은 정밀도를 요구한다. 이를 위해 본 논문에서는 수집된 x, y, z 좌표와 도플러 값을 32비트 부동 소수점 형식으로 변환하였다. 32비트 부동 소수점은 충분한 정밀도를 제공하면서도 64비트 형식보다 메모리 사용량이 적다. 이는 라즈베리파이와 같은 제한된 메모리 환경에서도 정확한 데이터를 처리할 수 있게 하며, 계산 자원의 효율적인 활용을 가능하게 한다. 마지막으로, 이 데이터를 이진 파일(binary file) 형태로 변환하여 PointPillars에서 사용한다. 결과적으로 이러한 데이터 전처리 과정은 데이터의 정확도와 범용성을 확보할 수 있는 장점이 있다.

IV. 제안된 시스템의 성능평가

4-1 연산량 및 파라미터 수, 정확도 비교

높은 정확도와 낮은 연산 복잡도, 적은 파라미터 수 등 경량화 목적에 달성하기 위한 최적의 네트워크를 찾기 위해

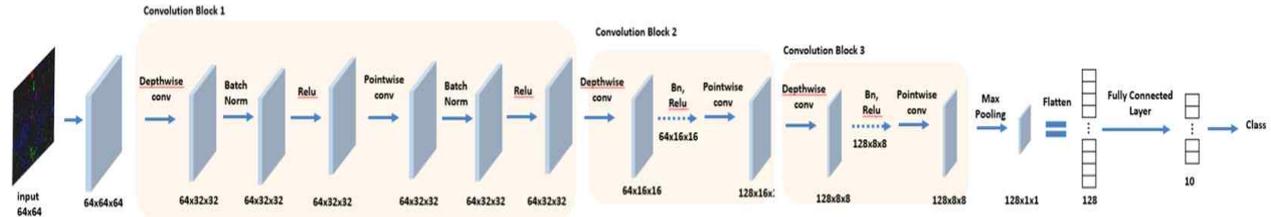


그림 4. 맞춤형 DS-CNN 모델 구조
 Fig. 4. Customized DS-CNN model architecture.

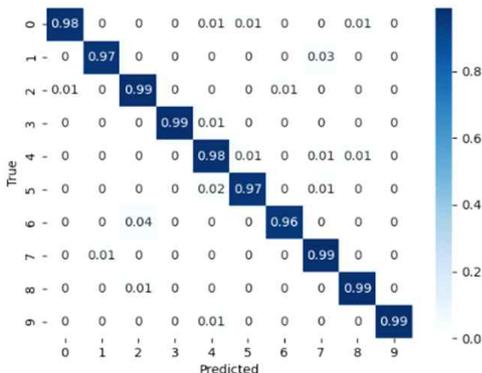


그림 5. 맞춤형 DS-CNN 모델 정확도
 Fig. 5. Customized DS-CNN model accuracy

표 1. LeNet-5와 DS-CNN 모델 정확도, 연산량, 파라미터 수 비교

Table 1. Comparison of LeNet-5 and DS-CNN models in terms of accuracy, computational complexity, and number of parameters.

Network model	Accuracy [%]	Complexity [MMacs]	Parameters [k]
LeNet-5 model	92.5	52.94	236.45
DS-CNN model	88.73	4.75	46.41

표 2. 다양한 DS-CNN의 정확도와 연산 복잡도, 파라미터 수

Table 2. Accuracy, computational complexity, and number of parameters of various DS-CNN models.

Input Size	c1	c2	c3	c4	c5	c6	fc	A [%]	C [MMacs]	P [k]
32x32	64	64	64	128	-	-	128	86.64	2.53	24.27
	64	64	128	128	-	-	256	88.73	4.75	46.41
	64	64	64	128	128	128	128	91.21	2.82	42.31
	64	64	128	128	256	256	256	92.43	5.84	115.3
64x64	64	64	64	128	-	-	128	93.32	10.13	24.27
	64	64	128	128	-	-	256	95.58	18.98	46.41
	64	64	64	128	128	128	128	98.28	11.27	42.31
	64	64	128	128	256	256	256	98.42	23.35	115.3

A: Accuracy, C: Complexity, P: Parameters

c1: Convolution 1 input channel, c2: Convolution 2 input channel, c3: Convolution 3 input channel, c4: Convolution 4 input channel, c5: Convolution 5 input channel, c6: Convolution 6 input channel, fc: Fully Connected Layer input channel

LeNet-5와 DS-CNN을 기반으로 다양하게 모델을 구성하고 학습을 진행하였다. 모델 학습은 PyTorch 딥러닝 프레임워크 기반 PointPillars 알고리즘을 사용하여 진행하였다. 손실 함수로는 NLL Loss (negative log likelihood loss)를 사용하였으며, 배치 크기는 32 또는 64, Epoch은 200으로 학습을 진행하였다.

기본 LeNet-5 모델과 DS-CNN 모델의 정확도, 연산 복잡도, 파라미터 수는 표 1과 같이 측정되었다. LeNet-5 모델은 높은 정확도를 보였으나, 연산 복잡도와 파라미터 수가 커서 제한된 환경에서의 동작에는 적합하지 않다고 판단하여 DS-CNN 모델 최적화를 진행하였다.

DS-CNN 모델의 특성을 유지하기 위해, 포인트별 합성곱 계층 (convolution layer 2, convolution layer 4, convolution layer 6)의 스트라이드 크기와 커널 크기는 각각 1로 고정하였다. 입력 pseudo 이미지 크기와 합성곱 계층의 개수에 따른 정확도, 연산 복잡도, 파라미터 수가 표 2와 같이 측정되었다.

따라서, 정확도와 연산 복잡도, 파라미터 수의 tradeoff을 고려하였을 때 최종적으로 그림 4와 같은 구조의 네트워크를 선정하였으며, 이를 맞춤형 DS-CNN으로 지칭한다. 이 모델의 입력 크기는 64x64이고 6개의 합성곱 계층과 1개의 완전 연결 계층 (fully connected layer)로 구성되어 있으며, 세 번째 합성곱 계층의 입력 채널이 64이다. 모든 합성곱 계층 이후에는 배치 정규화와 Relu 함수가 적용되며, 마지막 합성곱 계층 다음 Max pooling이 추가적으로 배치된다.

1,800개의 추론 데이터를 대상으로 맞춤형 DS-CNN 성능을 평가한 결과는 그림 5의 Confusion Matrix와 같다. 10가지 동작에 대해 각각 0.01%, 0.02%, 0.03%, 0.04% 정도의 오차를 보이며 최종 평균 98.28%의 정확도로 분류되는 것을 확인할 수 있다.

V. 제안된 시스템의 구현

5-1 소프트웨어 설계 및 구현

본 논문에서 제안하는 시스템은 그림 6과 같이 라즈베리파이 에 앞선 성능평가에서 최종 선정된 알고리즘 모델을 활용하여 구현하였다. MIMO 레이다와 라즈베리파이간 와이파이 연결, ROS 사용, 각종 Python 라이브러리 사용 등의 다양한 호환성 문제를 해결하기 위해 Ubuntu 20.04 운영체제를 사용하였다.

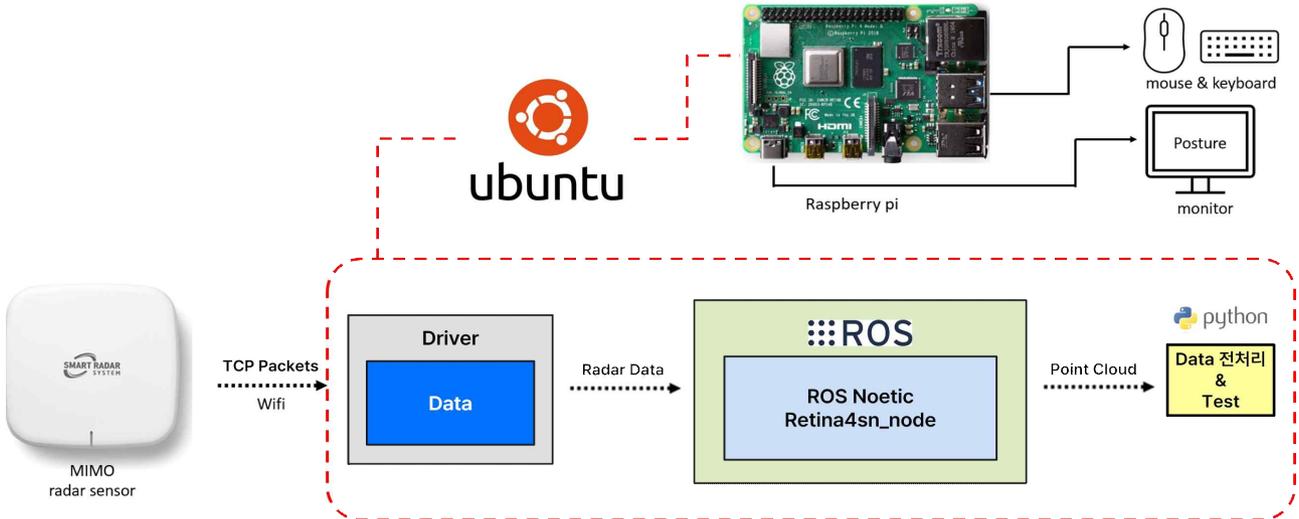


그림 6. 라즈베리파이(Raspberry-Pi)에 구현된 시스템 구조
 Fig. 6. System structure implemented on Raspberry-Pi.

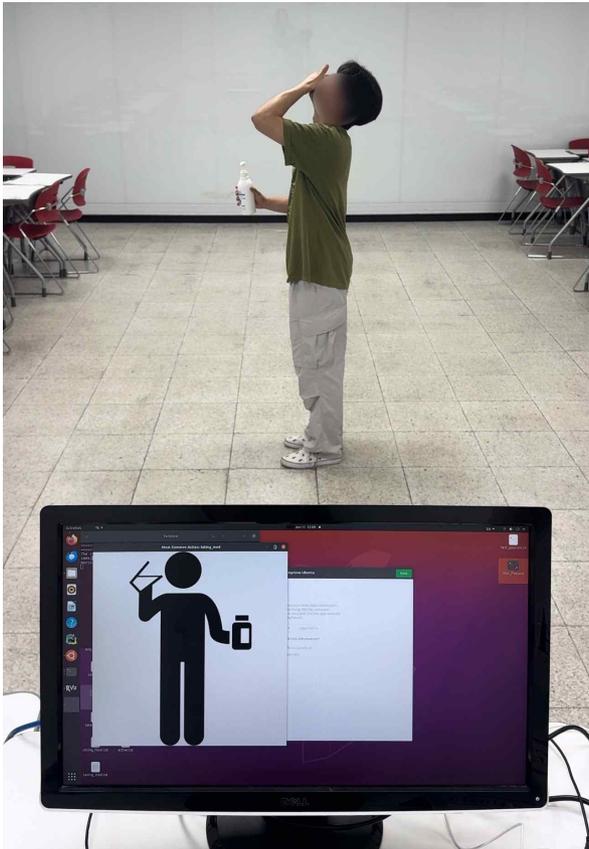


그림 7. 제안된 HAR 시스템 실험 장면
 Fig. 7. Demonstration scene of the proposed HAR system.

수신된 레이더 데이터를 포인트 클라우드 데이터로 변환하기 위해 ROS의 Noetic 버전을 사용하였다. 최종적으로 Python을 이용하여 포인트 클라우드 데이터를 이진 파일로 전처리하고, 결과를 모니터에 시각화하는 검증 시스템을 구현하였다. 결과

적으로 제안된 시스템은 데이터 취득부터 검증 과정까지 하나로 연결되어 동작한다. 이 시스템은 실제 환경에서 데이터 수집과 분석을 실시간으로 수행할 수 있는 강력한 도구로서 기능한다.

5-2 실험 결과 및 분석

본 논문에서 제안한 시스템을 사용하여 실제 동작 분류 작업을 수행하였다. 그림 7은 동작 분류 작업을 수행하는 모습을 보여준다. 센서는 사람의 동작에 따른 데이터를 실시간으로 수신하고 처리하여, 10개의 동작 중 하나를 분류한 후, 추론 결과를 모니터에 사진 형태로 출력한다.

실험 결과, 본 논문의 시스템은 최대 8 fps의 속도로 포인트 클라우드 데이터 처리가 가능함을 확인하였다. 이러한 결과는 실생활에서 본 논문의 시스템이 사용자가 느끼기에 실시간 데이터를 처리에 있어 성능이 다소 부족하다는 것을 시사한다. 따라서 이는 향후 시스템의 개선 사항으로 지적되며, 처리 속도 개선을 위해 하드웨어를 통한 가속이 필요함을 알 수 있다.

또한, 실생활에서는 본 논문에서 다룬 10가지의 동작 이외에도 매우 다양한 동작들이 발생할 수 있다. 따라서 실제 환경에서 본 논문의 시스템의 응용도를 높이기 위해서는 현재보다 훨씬 더 다양한 동작에 대한 데이터를 수집하고, 이를 바탕으로 알고리즘을 최적화하여 더 다양한 동작을 정확하게 인식하고 구분할 수 있는 시스템으로 발전이 필요하다.

향후 연구에서는 앞서 언급한 한계점을 개선하는 동시에, 시스템의 일반화 성능을 높이기 위해 다양한 실제 환경에서의 실험을 추가하여 제안된 시스템이 신뢰성 있게 동작할 수 있도록 범용성을 검증하고자 한다.

VI. 결 론

본 논문에서는 다중 입출력 FMCW 레이더와 합성곱 신경망을 기반으로 효율적인 사람 동작 인식 시스템을 제안하였다. 일상생활에서 자주 사용되는 동작으로 클래스를 선정하여 경쟁력 있는 시스템을 설계하였다. 정확도와 연산 복잡도 측면에서 최적화된 딥러닝 모델인 맞춤형 DS-CNN을 포함한 시스템을 설계 및 구현하였다. 실험 결과, 동작 분류 정확도는 98.28%를 달성하였으며, 연산 복잡도는 곱셈 후 누적(Macs; mutiply-accumulates) 연산에서 11.27M 수준으로 동작 분류 측면에서 경쟁력 있는 모델임을 확인하였다. 최종적으로 개발한 모델을 라즈베리파이에 구현하여 최대 8 fps의 속도로 포인트 클라우드 데이터 처리가 가능함을 확인하였다.

Acknowledgments

본 연구는 2024년도 산업통상자원부 및 한국산업기술평가관리원 시장주도형 K-센서 기술개발 사업 (00144290)의 일환으로 수행되었으며, CAD tool은 IDEC에 의해 지원되었음.

References

[1] X. Zhou, X. Meng, J. Zheng, G. Fang, and T. Guo “Human body recognition based on the sparse point cloud data from MIMO millimeter-wave radar for smart home,” *Multimedia Tools and Applications*, Vol. 83, No. 8, pp. 22055-22074, Aug, 2023. DOI: <https://doi.org/10.1007/s11042-023-15700-7>.

[2] Z. Wu, Z. Cao, X. Yu, J. Zhu, C. Song, and Z. Xu “A novel multiperson activity recognition algorithm based on point clouds measured by millimeter-wave MIMO radar,” *IEEE Sensors Journal*, Vol. 23, Issue. 17, pp. 19509-19523, June, 2023. DOI: <https://doi.org/10.1109/JSEN.2023.3283778>.

[3] S. M. Mohammadi, M. Alnowami, S. Khan, D.-J. Dijk, A. Hilton, and K. Wells, “Sleep posture classification using a convolutional neural network,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu: HI, pp. 3501-3504, July. 2018. DOI: <https://doi.org/10.1109/EMBC.2018.8513009>.

[4] Y. M. Kim, Y. D. Son, W. J. Kim, B. K. Jin, and M. H. Yun, “Classification of children’s sitting postures using machine learning algorithms,” *Applied Sciences*, Vol. 8, Issue. 8, July. 2018. DOI: <https://doi.org/10.3390/app8081280>.

[5] A. Hayat, F. Morgado-Dias, B. P. Bhuyan, and R. Tomar, “Human activity recognition for elderly people using machine and deep learning approaches,” *Information*, Vol. 13, Issue. 6, May. 2022. DOI: <https://doi.org/10.3390/info13060275>.

[6] G. Sebestyen, I. Stoica, A. Hangan, “Human activity recognition and monitoring for elderly people,” in *IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, Cluj-Napoca, pp. 341-347, Sep. 2016. DOI: <https://doi.org/10.1109/ICCP.2016.7737171>.

[7] L. Yang, G. Li, M. Ritchie, F. Fioranelli and H. Griffiths, “Gait classification based on micro-doppler features,” in *CIE International Conference on Radar*, Guangzhou: China, pp. 1-4, 2016. DOI: <https://doi.org/10.1109/RADAR.2016.8059301>.

[8] R. J. Javier, Y. Kim, “Application of linear predictive coding for human activity classification,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 11, No. 10, pp. 1831-1834, Oct. 2014. DOI: <https://doi.org/10.1109/LGRS.2014.2311819>.

[9] X. Li, Y. He and X. Jing, “A deep multi-task network for activity classification and person identification with micro-doppler signatures,” in *International Radar Conference*, Toulon: France, pp. 1-5, Sept. 2019. DOI: <https://doi.org/10.1109/RADAR41533.2019.171263>.

[10] Z. Yu, A. Taha, W. Taylor, A. Zahid, K. Rajab, H. Heidari, ..., Q. H. Abbasi, “A radar-based human activity recognition using a novel 3-D point cloud classifier,” *IEEE Sensor Journal*, Vol. 22, Issue. 19, pp. 18218 - 18227, Oct, 2022. DOI: <https://doi.org/10.1109/JSEN.2022.3198395>.

[11] X. Dang, P. Jin, Z. Hao, W. Ke, H. Deng, L. Wang, “Human movement recognition based on 3D point cloud spatiotemporal information from millimeter-wave radar,” *Sensors*, Vol. 23, Issue. 23, Nov, 2023. DOI: <https://doi.org/10.3390/s23239430>.

[12] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, “Pointpillars: fast encoders for object detection from point clouds,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach: CA, pp. 12689-12697, Jan. 2020. DOI: <https://doi.org/10.1109/CVPR.2019.01298>.

[13] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, Hartwig Adam, “MobileNets: efficient convolutional neural networks for mobile vision applications,” *arXiv*, April. 2017. DOI: <https://arxiv.org/abs/1704.04861>.



김 준 성 (Joon-sung Kim)

2019년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부
※관심분야 : SoC 설계, 임베디드 시스템



심 재 용 (Jae-yong Sim)

2019년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부
※관심분야 : SoC 설계, 임베디드 시스템



장 수 림 (Su-lim Jang)

2020년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부
※관심분야 : SoC 설계, 임베디드 시스템, 딥러닝



임 승 찬 (Seung-chan Lim)

2024년 2월 : 한국항공대학교 항공전자정보공학부 (공학사)
2024년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 석사
※관심분야 : SoC 설계, 임베디드 시스템, 딥러닝



정 윤 호 (Yun-ho Jung)

1998년 2월 : 연세대학교 전자공학과 (공학사), 2000년 2월 : 연세대학교 전기전자공학과 (공학석사)
2005년 2월 : 연세대학교 전기전자공학과 (공학박사), 2005년 ~ 2007년 : 삼성전자 책임연구원
2007년 ~ 2008년 : 연세대학교 연구교수
2008년 ~ 현재 : 한국항공대학교 항공전자정보공학부 교수
※관심분야 : 무선통신 시스템, 항공통신 시스템, 레이더시스템, 인공지능 시스템 SoC 설계