

tinyRadar: mmWave Radar based Human Activity Classification for Edge Computing

Satyapreet Singh Yadav^{1*}, Radha Agarwal^{1*}, Kola Bharath¹, Sandeep Rao², Chetan Singh Thakur¹
{satyapreets, radhaagarwal, kolabharath, csthakur}@iisc.ac.in, s-rao@ti.com

¹Department of Electronic Systems Engineering, Indian Institute of Science, Bangalore, India, 560012

²Texas Instruments, Bangalore, India, 560093

Abstract—The rising need for elderly care, child care, and intrusion detection challenges the sustainability of traditional systems that depend on in-person monitoring and surveillance. The current state-of-the-art technology heavily relies on Infra-Red (IR) and camera-based systems, which often require cloud computing. It can lead to higher latency, data theft, and privacy issues of being continuously monitored. This paper proposes a novel tinyML-based single-chip radar solution for on-edge sensing and detection of human activity. Edge computing within a small form factor solves the issue of data theft and privacy concerns as radar provides point cloud information. Also, it can operate in adverse environmental conditions like fog, dust, and low light. This work used the Texas Instruments IWR6843 millimeter wave (mmWave) radar board to implement signal processing and Convolutional Neural Network (CNN) for human activity classification. A dataset for four different human activities generalized over six subjects was collected to train the 8-bit quantized CNN model. The real-time inference engine implemented on Cortex®-R4F using CMSIS-NN framework has a model size of 1.44KB, gives the classification result after every 120ms, and has an overall subject-independent accuracy of 96.43%.

Index Terms—tinyML, Edge computing, TI mmWave radar IWR6843, velocity-time map, CMSIS-NN, CNN

I. INTRODUCTION

A real-time surveillance system is needed for continuous monitoring of human activities. In the modern era, remote surveillance is desired to ensure the safety of children, the elderly, and physically-challenged living alone at home and detect unintended human presence in restricted areas.

A remote human activity monitoring framework requires data security, privacy, low latency, robustness to adversarial attacks [1], easy portability, reliability in all weather conditions, and the ability to function in darkness [2]. A cost-effective non-contact solution capable of fulfilling the outlined monitoring requirements is met by the recent advances in mmWave radar [3]–[7], and Machine Learning (ML) techniques deployable on-edge. Alarming activities such as falls in the case of the elderly, no movement in the case of bed-ridden patients [5] can be easily detected using mmWave radar technology. Availability of Digital Signal Processor (DSP), Hardware Accelerator (HWA), and Cortex®-R4F Micro-controller (MCU) on a single-chip [8] makes on-edge processing possible on mmWave radar, leading to lower latency and enhancing the portability, making it deployable in small form factor.

The majority of the mmWave radar-based systems acquire raw data from the sensor and perform classification tasks off-edge [4], [9], [10]. There are few works related to on-edge processing on mmWave radar board, such as classifying human gestures [3], wherein hand-crafted features were used to train a simple classification network. But, classification using hand-crafted features is not scalable.

This work proposes a novel real-time, single-chip mmWave radar-based on-edge signal processing and classification system. The mmWave radar sensor provides information of the target environment. This information is processed to extract the signature (velocity-time) maps of the target’s activity. IWR6843 mmWave radar, which operates in 60-64 GHz frequency range, can be configured to provide excellent range and velocity resolution, which is critical to differentiate various human activities. We implemented the inference engine on the radar board for the classification of four human activities using a lightweight CNN model deployed using CMSIS-NN framework [11]. We have made our human activity data set publicly available [12].

II. SYSTEM DESCRIPTION

An mmWave radar sensor transmits frequency-modulated continuous waves (FMCW), called chirps, to the target environment using transmitting (TX) antennas, and the reflected signals from objects are captured by receiving (RX) antennas. The inner details of the proposed system are shown in Fig. 1. The onboard Cortex®-R4F MCU is programmed to configure the chirp parameters for transmission. The received and transmitted chirp signals are mixed to obtain Intermediate Frequency (IF) signals. Analog IF signals are converted into digital signals using an on-chip Analog to Digital Converter (ADC). The frequency of digitized IF signals holds the information about the range of the objects, and the phase difference of two consecutive IF signals contain the information about the objects’ velocity. The IF signals of two different RX antennas are used to measure the angle of the objects.

The digitized data is passed through HWA to perform 1D-Fast Fourier Transform (range-FFT) to generate range-time maps. The range-time maps are passed to DSP C67x for performing 2D-FFT (velocity-FFT) to generate range-velocity maps and then velocity-time maps. The velocity-time map is passed from DSP C67x to the inference network, which is deployed onto the Cortex®-R4F MCU for classification. The

*Equal contribution

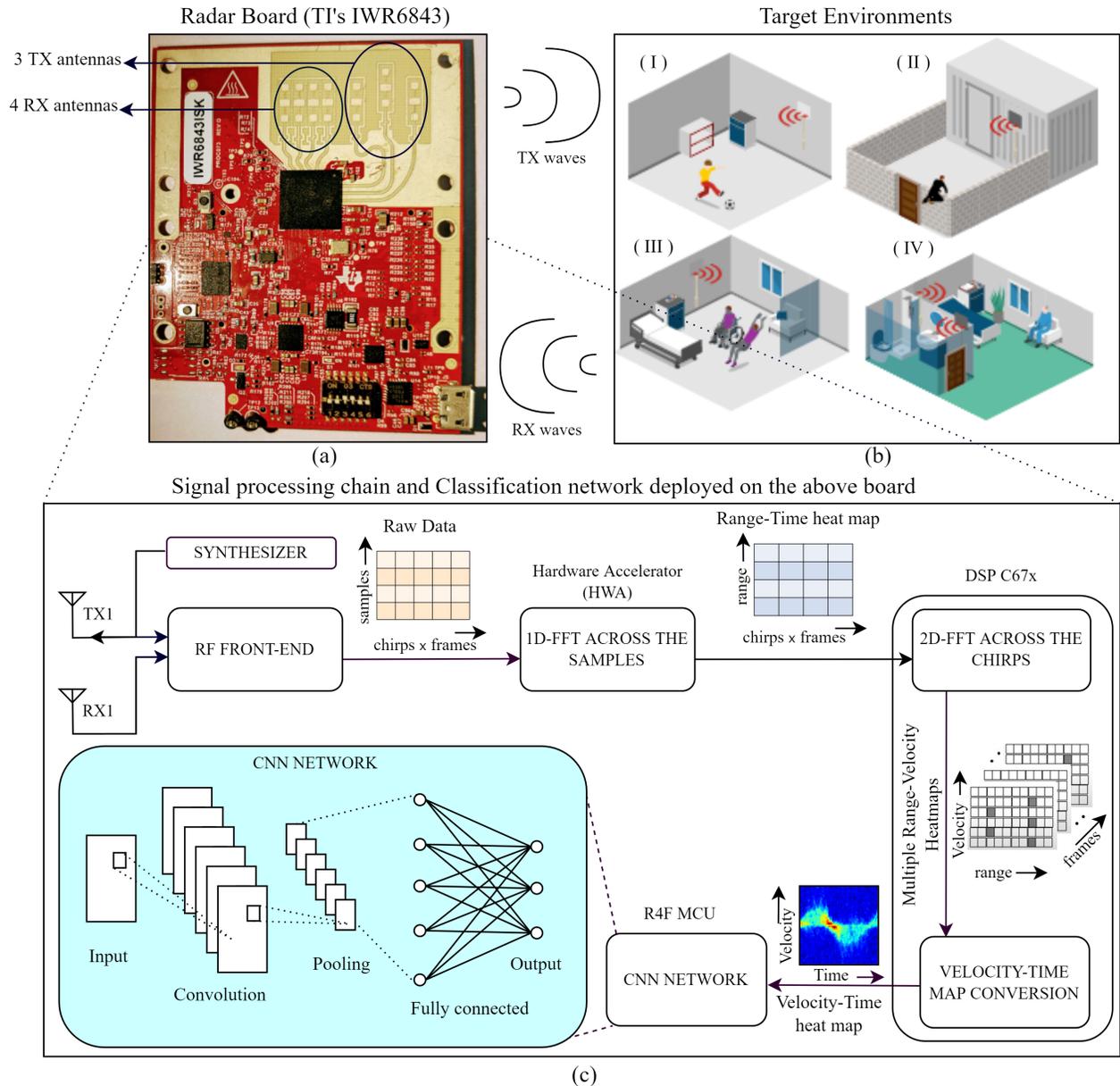


Fig. 1. Proposed block diagram showing the flow from sensing of the environment to implementation of signal processing chain and classification algorithm on radar board. (a) IWR6843 radar sensor (we used one TX and one RX antennas) (b) Potential target environments for surveillance, (I) Child monitoring (II) Intrusion detection (III) Patient monitoring and Fall detection (IV) Psychiatric ward monitoring. (c) Block diagram describing the implementation of the signal processing chain (on HWA and DSP C67x) for generation of velocity-time maps, and the CNN network for classification (on Cortex@-R4F MCU).

velocity-time maps of each human activity were recorded for the offline training of the CNN model.

III. DATASET DESCRIPTION

The chirp configurations used for the data collection are shown in TABLE I. The higher the number of chirp loops or the longer the chirp duration (Idle time + Ramp time), the better the velocity resolution. However, increasing the chirp duration decreases the maximum measurable velocity. Steeper frequency slope or longer ramp time provides better range resolution. However, increasing the frequency slope decreases the maximum measurable range. The maximum measurable

range can be increased by increasing the sampling rate to the maximum allowable limit but at the cost of a decrease in SNR. Frame periodicity was chosen such that the duty cycle is below 50%, which is a hardware limitation. Increasing the number of RX antennas increases the SNR and the angular resolution but also increases the data storage. The RX antennas can be chosen between one to four depending on the application. Applications like gesture recognition require finer velocity resolution and higher SNR, which can be achieved by increasing the chirps in a frame or the chirp duration [13].

We chose 64 chirps per frame for this application with

a chirp duration of $287\mu\text{s}$ to achieve a velocity resolution of 0.132m/s . This enables us to capture the relevant human activity signatures in a velocity-time map of size 64×64 of duration 2.56 seconds. We designed our chirp parameters, namely frequency slope, sample rate, and ADC samples, such that we meet high SNR to measure human activity within the 7m range. We chose one TX and one RX antenna to reduce 3D radar cube data size, which reduces the memory footprint on the radar board.

TABLE I
CHIRP CONFIGURATIONS

Parameter	Value
Frequency slope	$45.624 \text{ MHz}/\mu\text{s}$
Idle time	$199.8 \mu\text{s}$
ADC samples	241
Sample rate	3000 ksps
Ramp end time	$87 \mu\text{s}$
Chirps per frame	64
Frame periodicity	40 ms
Number of Frames	64
TX antennas used	1
RX antennas used	1
Maximum range	7.89 m
Range resolution	4.09 cm
Maximum velocity	4.23 m/s
Velocity resolution	0.132 m/s

IV. CNN BASED ACTIVITY RECOGNITION NETWORK

A. Radar Signal Processing

In FMCW radar, the received signals, after digitization, forms a 3D radar cube with dimensions as the number of samples, number of chirps, and number of receive antennas. First, 1D-FFT across chirp samples is executed on the 3D radar cube to obtain the range-time map. The range-time maps are further processed to identify moving targets against static clutter. Then, 2D-FFT is performed across the chirps to obtain range-velocity maps. Each range-velocity map is incoherently added in the range dimension to obtain one velocity column. This process is repeated over a sequence of frames to get a velocity-time map. The normalized velocity-time maps for different human activities are shown in Fig. 2. These maps contain the time-varying frequency modulations imposed on the received signals because of human body motion.

B. CNN Architecture

The key features of a CNN are translational and rotational invariance. We found that velocity-time maps contain unique signatures for different human activities, enabling us to use a small CNN network, making it a perfect candidate for hardware implementation.

The network architecture shown in Fig. 3 describes the CNN model trained on normalized velocity-time maps of human activities. The CNN model was trained on 932 velocity-time maps generated for five different human subjects and four different activities on Caffe [14]. Adam optimizer, multinomial logistic loss function, L2 regularization, and fine-tuned hyper-parameters were used to achieve the best possible accuracy.

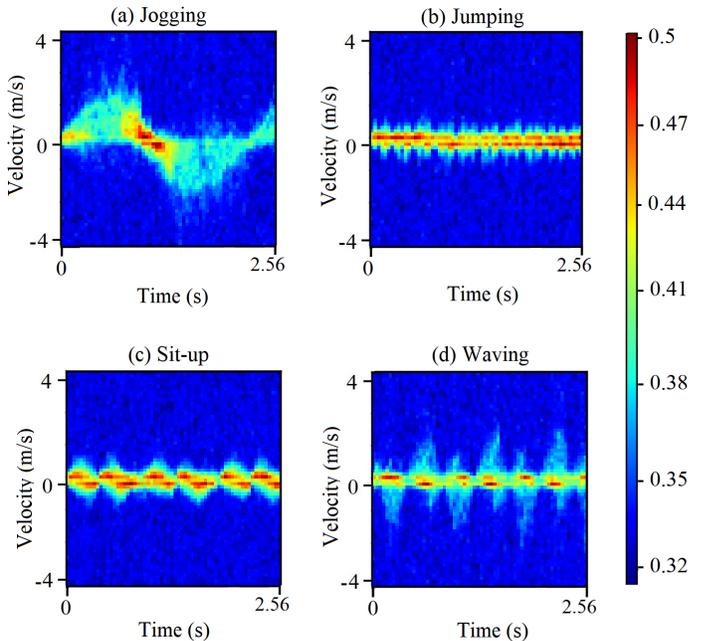


Fig. 2. Normalized velocity-time heat maps of 4 human activities: (a) Jogging, (b) Jumping, (c) Sit-up, and (d) Waving hands.

Post-training quantization was done to get the weights and activations in an 8-bit integer format in order to port the network to hardware. The CNN model was tested on different human subjects to validate its subject independent nature.

For inference, the quantized weights and biases were used to deploy the model on the radar board’s Cortex®-R4F MCU. The CNN architecture was built using CMSIS-NN’s custom Application Programming Interface (API).

V. RESULTS

We implemented the radar signal processing and classification pipeline on IWR6843. The signal processing chain utilized a total memory footprint of 112KB on L3RAM, which has a total storage capacity of 768KB. L3RAM is a shared memory unit between DSP C67x and Cortex®-R4F MCU. The classification network utilized a total memory footprint of 11.07KB on Cortex®-R4F MCU. A sliding window of 3 frames was used to get the classification result after every 120ms ($40\text{ms} \times 3$ frames). TABLE II contains details of the memory footprint and processing latency of each stage.

The Cortex®-R4F MCU takes 80mW of power. CNN computation results in 60.8% of CPU load, resulting in 48.64mW ($0.608 \times 80\text{mW}$) of power consumption.

We obtained a classification accuracy of 96.43% on the test dataset comprising of 224 velocity-time maps with the 8-bit quantized CNN network. The confusion matrix highlights the accuracy per class for human activity recognition on hardware is shown in TABLE III. Here, ‘No activity’ indicates no human movement in front of the radar. Jogging and Waving were 100% classified, whereas the other activities, i.e., Sit-up and

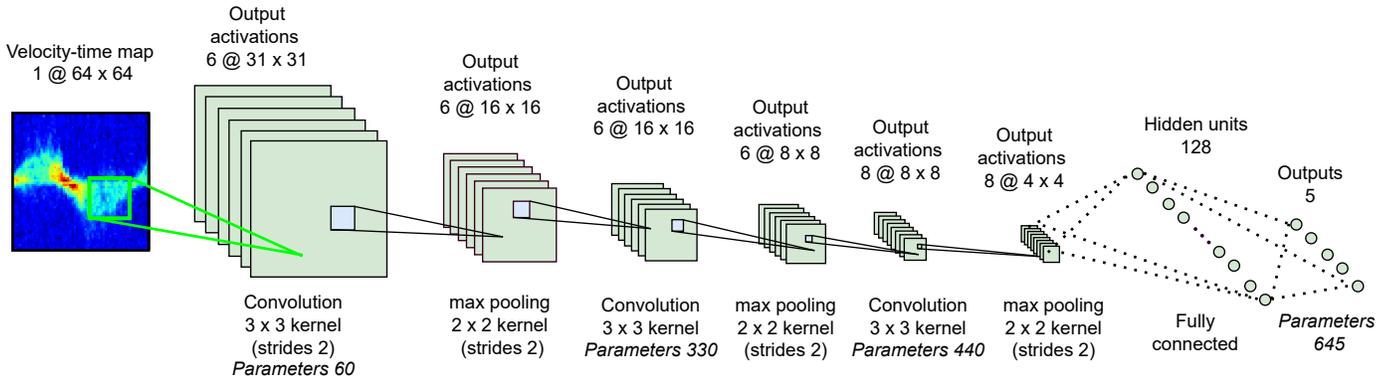


Fig. 3. Proposed CNN architecture

TABLE II
MEMORY CONSUMPTION AND PROCESSING LATENCY

Processing stage	Core	Memory utilization (KB)	Processing latency (ms)
ID-FFT	HWA	64	~18.37
2D-FFT	C67x	32	~1.08
Velocity-time map	C67x	16	~1.78
Classification network	Parameters	1.44	~75
	Output activations	9.63	

Jumping, were classified with an accuracy of 89.13% and 97.83%, respectively.

TABLE III
CONFUSION MATRIX IN PERCENTAGE

Predicted \ Actual	Jogging	Jumping	Sit-up	Waving	No activity
Jogging	100	0	0	0	0
Jumping	0	97.83	0	0	2.17
Sit-up	0	0	89.13	10.87	0
Waving	0	0	0	100	0
No activity	0	5	0	0	95

VI. CONCLUSION

We have proposed a tinyML mmWave radar-based system for the human activity classification task. This is done by integrating the signal processing chain on DSP C67x with the classification network on Cortex®-R4F MCU for real-time inference. We obtained an accuracy of 96.43% with the 8-bit quantized CNN model on 224 velocity-time maps comprising of five different classes. It acquires a model size of 1.44KB, provides classification output every 120ms, and consumes 48.64mW of power. The proposed architecture can be used for various applications by retraining the CNN model. We will also explore the sparse CNN architecture to exploit inherent sparsity present in the radar data [15].

REFERENCES

- [1] Yanmao Man, Ming Li, and Ryan Gerdes. GhostImage: Remote perception attacks against camera-based image classification systems. In *23rd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2020)*, pages 317–332, San Sebastian, October 2020. USENIX Association.
- [2] Bibrat Ranjan Pradhan, Yeshwanth Bethi, Sathyaprakash Narayanan, Anirban Chakraborty, and Chetan Singh Thakur. N-HAR: A neuro-morphic event-based human activity recognition system using memory surfaces. *Proceedings - IEEE International Symposium on Circuits and Systems*, 2019-May, 2019.
- [3] Piyali Goswami, Sandeep Rao, Sachin Bharadwaj, and Amanda Nguyen. Real-time multi-gesture recognition using 77 ghz fmcw mimo single chip radar. In *2019 IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–4, 2019.
- [4] Manuel Eggmann, Jonas Erb, Philipp Mayer, Michele Magno, and Luca Benini. Low power embedded gesture recognition using novel short-range radar sensors. In *2019 IEEE SENSORS*, pages 1–4, 2019.
- [5] Feng Jin, Renyuan Zhang, Arindam Sengupta, Siyang Cao, Salim Hariri, Nimit K. Agarwal, and Sumit K. Agarwal. Multiple patients behavior detection in real-time using mmwave radar and deep cnns. In *2019 IEEE Radar Conference (RadarConf)*, pages 1–6, 2019.
- [6] Huiling Cai, Fei Li, Dianzhu Gao, Yingze Yang, Shuo Li, Kai Gao, Aina Qin, Chao Hu, and Zhiwu Huang. Foreign objects intrusion detection using millimeter wave radar on railway crossings. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2776–2781, 2020.
- [7] Renyuan Zhang and Siyang Cao. Real-time human motion behavior detection via cnn using mmwave radar. *IEEE Sensors Letters*, 3(2):1–4, 2019.
- [8] Chethan Kumar Y B. 60GHz mmWave Sensor EVMs. pages 1–81, 2020.
- [9] Zhi Zhou, Zongjie Cao, and Yiming Pi. Dynamic gesture recognition with a terahertz radar based on range profile sequences and doppler signatures. *Sensors*, 18(1), 2018.
- [10] Xinyu Li, Yuan He, and Xiaojun Jing. A survey of deep learning-based human activity recognition in radar. *Remote Sensing*, 11(9), 2019.
- [11] Liangzhen Lai, Naveen Suda, and Vikas Chandra. CMSIS-NN: efficient neural network kernels for arm cortex-m cpus. *CoRR*, abs/1801.06601, 2018.
- [12] IISc Bangalore NeuRonICS Lab. Human activity recognition dataset. <http://gofile.me/4T44b/DM8DjmQJx>, 2021.
- [13] Sandeep Rao and Texas Instruments. Introduction to mmWave radar sensing: FMCW radars — TI.com Training Series.
- [14] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.
- [15] Benjamin Graham and Laurens van der Maaten. Submanifold sparse convolutional networks. <http://arxiv.org/abs/1706.01307>, 2017.