# Privacy-Preserving Fall Detection with Deep Learning on mmWave Radar Signal

Yangfan Sun [#1], Renlong Hang [#2], Zhu Li [#2], Mouqing Jin [*1], Kelvin Xu [*3]

[#] *Department of Computer Science and Electrical Engineering, University of Missouri-Kansas City, USA*

[*] *Ric Semiconductor, LLC, Texas, USA*

[1] {ysb5b, mqjwb8}@mail.umkc.edu

[2] {hangr, lizhu}@umkc.edu

[3] kelvin.xu@ric-semiconductor.com

***Abstract*—Fall is one of the main reasons for body injuries among seniors. Traditional fall detection methods are mainly achieved by wearable and non-wearable techniques, which may cause skin discomfort or invasion of privacy to users. In this paper, we propose an automatic fall detection method with the assist of the mmWave radar signal to solve the aforementioned issues. The radar devices are capable to record the reflection from objects in both the spatial and temporal domain, which can be used to depict the activities of users with the support of a recurrent neural network (RNN) with long-short-term memory (LSTM) units. First, we employ the radar low-dimension embedding (RLDE) algorithm to preprocess the Range-angle reflection heatmap sequence converted from the raw radar signal for reducing the redundancy in the spatial domain. Then, the processed sequence is split into frames for inputting LSTM units one by one. Eventually, the output from the last LSTM unit is fed in a Softmax layer for classifying different activities. To validate the effectiveness of our proposed method, we construct a radar dataset with the assist of market radar module devices, to implement several experiments. The experimental results demonstrate that, compared to LSTM only and the widely used 3-D convolutional neural network (3-D CNN), combining RLDE and LSTM can achieve the best detection results with much less computational time consumption. In addition, we extend the proposed method to classify multiple human activities simultaneously and the satisfied performances are observed.**

***Index Terms*—Fall detection, mmWave radar signal, radar low-dimension embedding (RFLE) algorithm, long-short-term memory (LSTM), human activities detection**

## I. Introduction

Fall activity is one of the leading causes of accidental death and injury for seniors and nearly $34 billion direct medical costs annually are caused by this [1]. Researchers pay more attention to this issue recently while the aging of population becomes a serious social problem. Traditional fall detection methods can be categorized into wearable and non-wearable techniques. For wearable methods, they are only functional while carrying portable devices, which are easy to cause skin discomfort or inconvenience [2]. Non-wearable methods are mainly motivated by modern computer version technologies, which can automatically alarm falls through surveillance equipments [3]. Even though they can detect fall activities accurately, high power consumption cannot be ignored. More importantly, the invasion of privacy [4] and the limited Field of View (FoV) are insufferable.

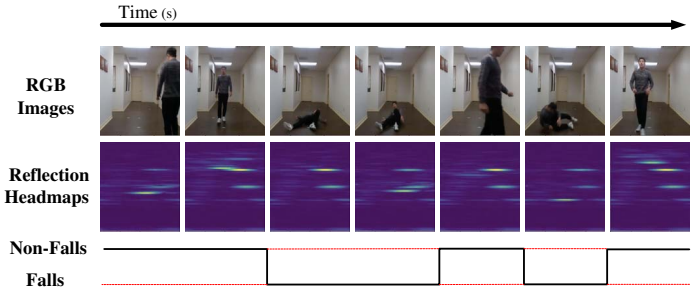


Fig. 1. The camera-based RGB images and corresponding range-angle reflection heatmaps.

In order to overcome aforementioned disadvantages, the radar-based fall detection methods have been proposed [5] [6]. The general idea of these methods is to depict human activities by recording the changes in the received specular signal reflected from the human body [7]. Less visual information compared to traditional non-wearable methods is generated by radar devices while implementing these methods, leading a result of reducing computational complexity and memories consumption of chips. In addition, users' privacy is perfectly preserved due to the impossibility on face identification, shown in Fig. 1. Earlier researches category radar-based fall detection methods into two classes: Doppler signal based methods and WiFi channel based methods. The former methods execute through the association between Doppler frequency and motion velocity [8], while the latter methods compute the changes of different signal in the WiFi channel [9]. Neither of them takes the locality component into account but solely the variations in the speed of motions.

Motivated by the method proposed in [10], we propose a fall detection method based on the mmWave radar signal, which considers locality and velocity components simultaneously. Different from the existing tasks, our method is to characterize mmWave radar reflections based on distance from the human body along with the elevation and azimuthal angles of arrays, which can be aggregated as spatiotemporal patterns, denoted as the range-angle reflection heatmaps. Fig. 1 shows the associations between reflection heatmaps and their correspond-

**978-1-7281-3723-0/19/$31.00 ©2019 IEEE**

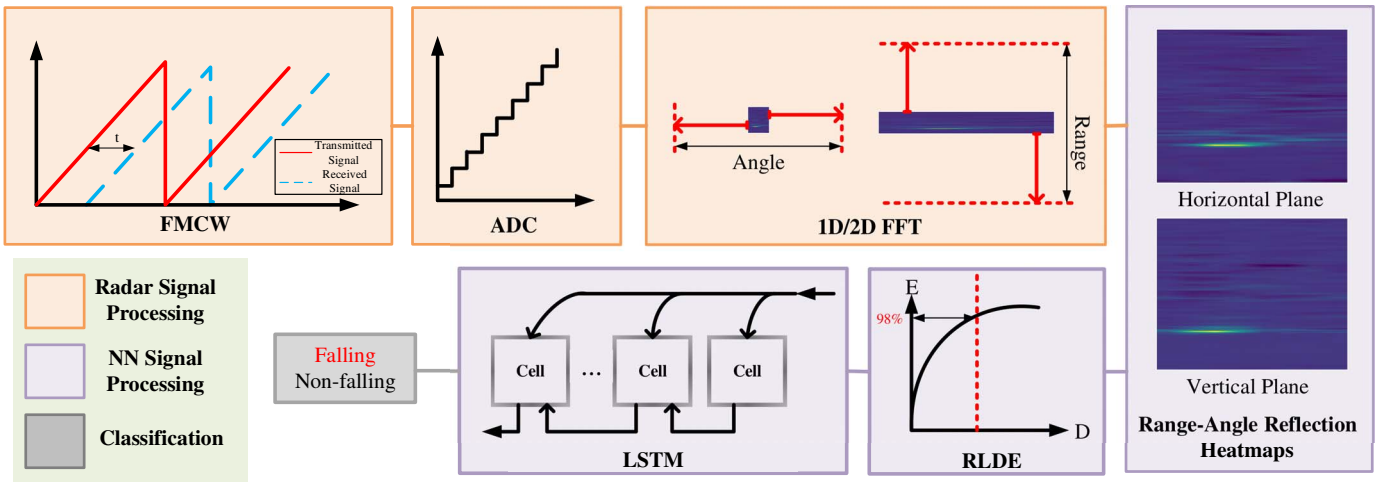


Fig. 2. The framework of proposed radar-based fall detection method

ing human activities, which can be observed through related camera-based images. Therefore, reflection heatmaps can be utilized to estimate fall activities in the presence of other sources of activities by neural network algorithms. Unlike [10], we jointly apply the radar low-dimension embedding (RLDE) algorithm and recurrent neural network (RNN) with long-short-term memory (LSTM) units as the substitution of 3-D convolutional neural network (3-D CNN) for reduction at the computational complexity and memory consumption.

## II. Proposed Method

Essentially, human activities, e.g., fall or non-fall, can be regarded as the changes of motion in term of range, angle and speed. These fundamental motion attributes can be derived through recording the variations in the signal of Receiver (RX) and Transmitter (TX) Antennas equipped at the IWR1642 device [11], which include time interval and intensity of signal between RX and TX. With the support of neural network algorithms, we can easily establish the correlation between human activities and corresponding attributes, further to design a radar-based human activity detection method, in this case, a fall detection method. Fig. 2 illustrates the framework of the proposed fall detection method, which consists of two parts: radar signal processing and neural network signal processing.

### A. Radar Signal Process

In this subsection, we aim to implement the radar signal conversion for the next processing, which consists of two following procedures:

- Analog-Digital Convert (ADC): Modulate radar signal from continuous format to discrete format.
- Fast Fourier Transform (FFT): Convert the discrete radar signal from the representation from the frequency to spatial domain.

In general, the transmitted and received mmWave radar signal is the analog frequency format signal. For our purpose, radar signal processing need be utilized to transform the raw signal into the discrete spatiotemporal signal, shown in Fig. 2. In this process, we implement FFT twice in order to achieve the expansions in both spatial domains (range and angle domain). Eventually, the processed data can be visualized as a set of reflection heatmaps, which is extracted from two separate perpendicular planes (vertical and horizontal plane) simultaneously.

### B. Neural Network Signal Processing

It is noted that human activities (fall or non-fall) are considered as continuous dynamic patterns, which are provided by spatiotemporal dependency. The concatenation of multiple successive reflection heatmaps can recognize both spatial and temporal dependency, while individual reflection heatmap merely is a representation of spatial expansion. Therefore, with the attention to the continuity of activities, we merge $n$ frames of reflection heatmaps to match the average duration of activities.

Fig. 1 visualizes several successive reflection heatmaps at a certain time interval and their corresponding camera-based images and labeled activities. We can observe that the changes in motion mainly concentrate on several specific regions, which indicates the existence of spatial redundancy. It is necessary to apply the RLDE algorithm to eliminate the redundancy. Assume the $t$-th frame of reflection heatmaps at horizontal and vertical planes as $H_t \in \mathbb{R}^{N\times M}$ and $V_t \in \mathbb{R}^{N\times M}$, respectively, where $N$ and $M$ are denoted as the maximal range and angle recorded. The RLDE Algorithm can be utilized to project reflection heatmaps to a low-dimension subspace $P \in \mathbb{R}^{\widetilde{N}\times N}$ as Equation (1),

$$\widetilde{H_t} = P * H_t,\ \widetilde{V_t} = P * V_t \tag{1}$$

where $\widetilde{N}$ represents the dimension of projected mapping. $\widetilde{H_t}$ and $\widetilde{V_t}$ are denoted as the $t$-th frame of projected reflection heatmaps in $\mathbb{R}^{\widetilde{N}\times M}$. We select the principal components analysis (PCA) algorithm to derive the low-dimension subspace $P$ linearly from our training set due to its low computational complexity.

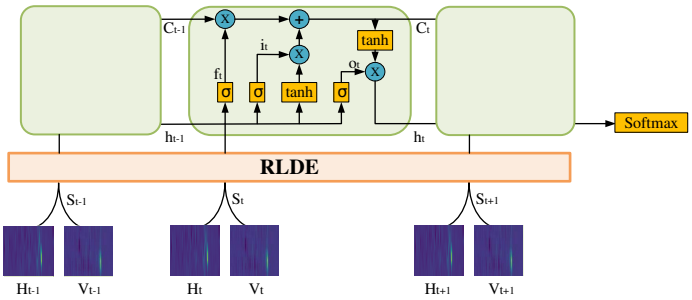


Fig. 3. The Proposed Network Architecture for Fall Detection Method

After data preprocessing, the neural network signal processing is elaborated. The sequential structure of RNN with LSTM units determines the advantages on applications of time dependency [12]. So, we input preprocessed heatmaps $\widetilde{H}_t$ and $\widetilde{V}_t$ into the a LSTM network, illustrated in Fig. 3. To make the network easier to train, we normalize the pixel values to the range of $[0, 1]$ and concatenate preprocessed heatmaps of two planes as $S_t \in \mathbb{R}^{\widetilde{N}\times 2M}$. Each concatenated heatmap $S_t$ is fed in each LSTM cell in sequence, undergoing three build-in operative gates (input, forget and output). First, the input gates are used to input $S_t$ in current LSTM cell,

$$i_t = \sigma(W_i \cdot [h_{t-1}, S_t] + b_i) \tag{2}$$

where $W_i$ and $b_i$ are weights and biases of input gates, respectively. $\sigma(.)$ is a sigmoid activation function. $h_{t-1}$ is the output of last LSTM cell. The boundary condition is $1 < t \leq n$. Second, the forget gates are to determine whether $S_t$ is discarded for the current cell,

$$f_t = \sigma(W_f \cdot [h_{t-1}, S_t] + b_f) \tag{3}$$

where $W_f$ and $b_f$ are weights and biases of forget gates, respectively. Third, the output gates are to send chosen pieces of the message from $S_t$ to the next cell,

$$o_t = \sigma(W_o \cdot [h_{t-1}, S_t] + b_o) \tag{4}$$

where $W_o$ and $b_o$ are weighs and biases of output gates, respectively. These gates jointly influence the output of each cell through $t$-th cell state vector $C_t$,

$$C_t = f_t * C_{t-1} + i_t * tanh(W_c \cdot [h_{t-1}, S_t] + b_c) \tag{5}$$

where $W_c$, $b_c$ are the weights and biases of cell state vector.

$$h_t = o_t * tanh(C_t) \tag{6}$$

$\tanh(.)$ is the Tanh activation function. While $t = n$, we output $h_t$ as the result of LSTM layer for classifying through a Softmax function.

The proposed network aims at learning a label $Y_0$ from the preprocessed heatmaps $H$ and $H$ through the estimated network parameter $\Theta = \{W_i, b_i, W_f, b_f, W_o, b_o, W_c, b_c\}$ with predefined loss function,

$$Y = f([H, V], \Theta) \tag{7}$$

which is achieved by minimizing the loss between the predicted labels $Y$ and their corresponding ground truth $Y_0$. The cross-entropy function is adopt as the objective function,

$$L(\Theta) = \sum_{i=1}^{m} [Y_{0i} \cdot log(Y_i) + (1 - Y_{0i}) \cdot log(1 - Y_i)] \tag{8}$$

where, $m$ is denoted as the number of labeled category. Adam optimizer is utilized to minimize the loss during the training process by updating the set of network parameter set $\Theta$ through the initialization of random Gaussian distribution with zero mean and standard deviation of 0.1.

## III. Dataset

We utilize a pair of TI's IWR1642 evaluation module devices [11] to produce FMCW mmWave radar signal. The values of related parameters are listed in Table I. Nearly 30% reductions on Rx and Tx antennas, compared to previous task [10], provide a wider-range of practical applications. 4,126 samples (2.56 seconds for each sample) consist of 128 frames of reflection heatmaps before concatenation for each. Except for 711 fall activities, we also collect various non-fall activities as the purpose of increasing the sample diversity, e.g., walking, pickup, standup, boxing, sitting and jogging.

TABLE I
Foundational Parameters of Radar Signal Device

| Parameters | Values | Parameters | Values |
|---|---|---|---|
| Maximum Detectable Range | 10 m | Wave Technology | FMCW |
| Range Resolution | 4 cm | Frequency Range | 77-81 GHz |
| Number of RX Antennas | 8 | Number of TX Antennas | 4 |
| Field of View | 120° | Angular Resolution | 15° |
| Maximum Velocity | 6.5 m/s | Maximum Bandwidth | 3750 MHz |
| Velocity Resolution | 0.2 m/s | ADC samples | 256 |
| Wavelength | 3.9 mm | Frame Rate | 25 f/s |

## IV. Experiment

In this section, we design a series of experiments to evaluate the performance of our proposed fall detection method compared to the state-of-the-art. Two labeled categories (falls and non-falls) dataset is established and split into 80%, 10% and 10% as the training, validation and testing set. The proposed method executes with $1e^{-4}$ of the learning rate, 100 epochs and 12 for each batch on Tensorflow.

TABLE II
The Comparison between Baseline and Proposed Method with and without RLDE Implementation

| Methods | Precision | Recall | $F1_{score}$ | Training Time(s) |
|---|---|---|---|---|
| 3-D CNN | 95.3% | 96.6% | 96.0% | 181.21 |
| LSTM | 100.0% | 93.6% | 96.7% | 94.29 |
| $LSTM^{64}$ | 100.0% | 97.9% | 98.9% | 56.83 |
| $LSTM^{32}$ | 100.0% | 95.8% | 97.8% | 37.22 |
| $LSTM^{16}$ | 100.0% | 97.7% | 98.9% | 22.21 |
| $LSTM^{8}$ | 97.9% | 100.0% | 98.9% | 20.33 |
| $LSTM^{4}$ | 100.0% | 97.7% | 98.9% | 17.08 |
| $LSTM^{2}$ | 97.5% | 88.6% | 92.9% | 15.12 |

To validate the performance in the case of the unbalanced dataset, we introduce the evaluated metrics of $Precision$, $Recall$ and $F1_{score}$ as the accurate measurement. In addition, the average training times on each epoch are recorded to show the benefits of computational simplification in Table II.

First, the performance of LSTM$^4$ and 3-D CNN can be observed. LSTM$^4$ represents a method of combining RLDE and LSTM, preserving four dimensions projected mapping in spatial domain. Nearly 3% increasing on $F1_{score}$ demonstrates the effectiveness of LSTM$^4$ compared to 3-D CNN. Meanwhile, 4.7% and 1.1% improve on $Precision$ and $Recall$, respectively. More important, LSTM$^4$ only need less than 10% training time compared to the 3-D CNN that proves the low consumption on LSTM$^4$, which is critical for real-time applications.

Second, we evaluate the benefits of LSTM network individually. The comparison of the LSTM without RLDE implementation and the 3-D CNN shows an increase of less than 1% on $F1_{score}$. Still, the efficiency of training processing develops approximate 50% that the same performance can be achieved with merely half time consumption.

Third, the performance also demonstrates the effectiveness of RLDE in Table II. The peak performance is occurred at LSTM$^4$ with 2.2% increases on $F1_{score}$ and 20% training time demand compared to the LSTM only. It is noticeable that an obvious drop on evaluated results and approaching saturation on the reduction of training time at LSTM$^2$, which explains the threshold of preserved dimensions.

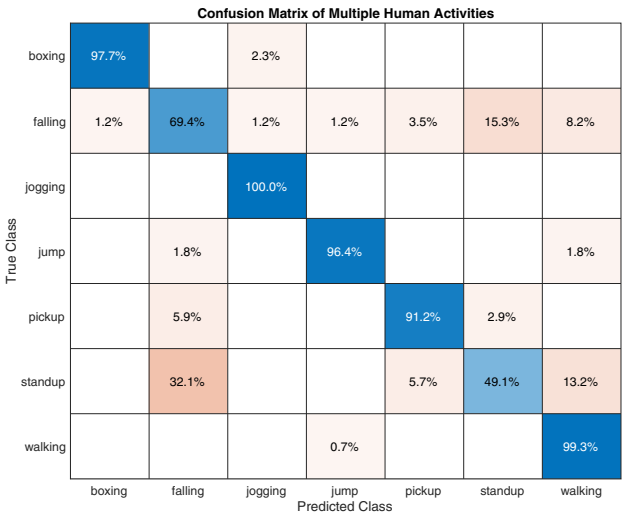


Fig. 4. Confusion Matrix of Multiple Activities Detection

Besides, we achieve an application of multiple activity detection with the proposed method. The samples are re-labeled into seven categories of activities, shown in Fig. 4. The result shows overall 80% accuracy on activities detection through the confusion matrix. Understandably, a degraded precision occurs at this detailed classification compared to the two categories classification. We can investigate that the proposed method has shortcomings in identifying human activities with similar motion attributes. Therefore, a more effective method will be considered in future work for more efficiently preserving temporal information from the mmWave radar signal.

## V. Conclusion

This paper proposes a privacy-preserving radar-based fall detection method, which utilizes an RNN with LSTM units architecture to learn human activities through spatiotemporal patterns extracted from the reflection heatmaps caught by radar devices. The experimental results show that our proposed RLDE+LSTM method outperforms the state-of-the-art in 3% increases on $F1_{score}$ with barely 10% training time needed. Our method highly reduces the number of network neurons and memory consumption, which makes real-time detection achievable. Besides, we extend the application of the proposed method on multiple activities detection that proves the latent capacity of more complexity detection.

## Acknowledgment

This project is supported in part by a National Science Foundation (NSF) I/UCRC award 1747751.

## References

[1] J. A. Stevens, P. S. Corso, E. A. Finkelstein, and T. R. Miller, "The costs of fatal and non-fatal falls among older adults," *Injury prevention*, vol. 12, no. 5, pp. 290–295, 2006.

[2] J. Dai, X. Bai, Z. Yang, Z. Shen, and D. Xuan, "A pervasive fall detection system using mobile phones," in *IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 292–297.

[3] T. Lee and A. Mihailidis, "An intelligent emergency response system: preliminary development and testing of automated fall detection," *Journal of telemedicine and telecare*, vol. 11, no. 4, pp. 194–198, 2005.

[4] D. Lyon, *Surveillance as social sorting: Privacy, risk, and digital discrimination*. Psychology Press, 2003.

[5] H. Wang and L. Cai, "On adaptive spatial-temporal processing for airborne surveillance radar systems," *IEEE Transactions on aerospace and electronic systems*, vol. 30, no. 3, pp. 660–670, 1994.

[6] B. Y. Su, K. Ho, M. J. Rantz, and M. Skubic, "Doppler radar fall activity detection using the wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 3, pp. 865–875, 2014.

[7] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, p. 219, 2015.

[8] L. Liu, M. Popescu, M. Skubic, M. Rantz, T. Yardibi, and P. Cuddihy, "Automatic fall detection based on doppler radar motion signature," in *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. IEEE, 2011, pp. 222–225.

[9] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "Falldefi: Ubiquitous fall detection using commodity wi-fi devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, p. 155, 2018.

[10] Y. Tian, G.-H. Lee, H. He, C.-Y. Hsu, and D. Katabi, "Rf-based fall monitoring using convolutional neural networks," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, p. 137, 2018.

[11] K. Ramasubramanian and B. Ginsburg, "Awr1243 sensor: Highly integrated 76–81-ghz radar front-end for emerging adas applications," 2017.

[12] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," 1999.