REAL-TIME FALL DETECTION USING MMWAVE RADAR

Wenxuan Li, Dongheng Zhang,Yadong Li, Zhi Wu, Jinbo Chen, Dong Zhang, Yang Hu, Qibin Sun, Yan Chen

School of Cyber Science and Technology, University of Science and Technology of China, Hefei, China

ABSTRACT

Fall is a severe health threat for elders' health care. While existing systems could achieve promising performance under specific scenarios, the required computing resources are usually not affordable, which is not applicable for real-time detection. In this paper, we propose mmFall, a real time fall detection system using millimeter wave signal which can achieve impressive accuracy with low computation complexity. Specifically, we first extract the signal variation corresponding to human activity with spatial-temporal processing. To enhance the system performance and robustness, we perform data augmentation by shifting, flipping, extracting and interpolating the signal. Finally, we design a light-weight convolutional neural network to achieve real-time fall detection. Extensive experimental results demonstrate that the proposed system could achieve state-of-the-art performance with limited computation complexity.

Index Terms— Fall Detection, Radar, Deep Learning

1. INTRODUCTION

With the aging of the population over the world, the health care of elderly has been a great challenge for the human society. According to statistics, each year about 684,000 individuals die from falls globally [1], and thus accidental falls become a major threat to the safety of the elderly. However, it is too hard to prevent falls under the help of nurses since this would introduce unaffordable cost. Hence, we need to continuously monitor the status of elderly and alarm in time once the fall is detected.

To tackle this challenge, lots of methods has been investigated in recent years. Based on the sensor utilized for sensing, existing methods can be generally categorized into two types: wearable devices based and camera based ones. Wearable devices including accelerometers [2] [3] , gyroscopes [4] , RFID [5] have all been investigated for fall detection. However, the survey shows that wearable devices are not suitable for elderly [6], which tend to be left or forgot by elder users. Besides, the physical contact between the sensor and user also decreases the acceptance level of wearable devices.

With the rapid development of deep neural network, vision-based fall detection systems have achieved contactless fall detection in specific environments [7] [8] [9] [10]. However, the severe privacy concerns introduced by the deployment of cameras have limited the application of such systems.

To break the limitations of wearable devices and cameras, Radio Frequency (RF) based systems has attracted attentions in recent years. WiFi based systems utilize channel state information (CSI) to achieve fall detection [11]. However, due to the limited bandwidth and number of antennas, WiFi based systems are prone to produce false alarms, which limits their applications. Similarly, doppler radars measure human movement velocity to achieve fall detection, which tends to be affected by the non-fall activities of human. To address this problem, Ma et al. [12] adopt Ultra-Wideband (UWB) Monostatic Radar and Long short-term memory (LSTM) to achieve fall detection. However, since the impact of environment has not been removed, the adaption ability to new people and environment is limited. To address the aforementioned problems, Tian et al. [13] leverage two angle-range heatmaps perpendicular to each other to distinguish human daily activity and falls. To eliminate the impact of the environment, a large scale dataset has been collected under lots of scenarios. However, achieving such a system is costly in both hardware and dataset, which is impractical in some scenarios.

In this paper, we propose mmFall, a real-time fall detection system using millimeter wave signal which can achieve impressive accuracy with low computation complexity. Specifically, to separate the signal from different spatial locations, we adopt a millimeter Frequency Modulated Continuous Wave (FMCW) to capture human reflections. With the captured data, we first perform 3D-FFT to extract the signal corresponding to human activity and suppress the noise and interference introduced by the environment. To further enhance the system performance, we perform data augmentation by shifting, flipping, extracting and interpolating the data to increase its diversity. Finally, we design a light-weight convolutional neural network to achieve real-time fall detection. To evaluate the proposed system, we collect dataset from 10 subjects with different height and weight. Extensive experimental results demonstrate that the proposed system could achieve state-of-the-art performance with limited computation complexity.

The rest of this paper is organized as follows. In section 2, we introduce the proposed mmFall system in detail. In section 3, we discuss the experimental results. Finally, we draw conclusion in section 4.

2. MMFALL SYSTEM

The overview of the proposed mmFall system is shown in Fig. 1. Our system acquires relected signal from the mmWave radar, transforms the signal into range-angle domain and finally performs fall detection with a neural network.

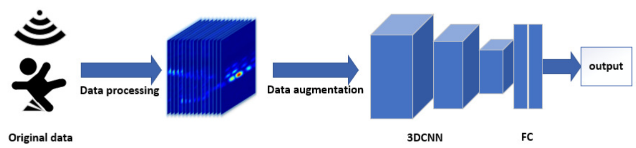

Fig. 1. System Overview

2.1. Data Collection

To separate the signals from different spatial locations, we adopt Frequency-modulated continuous-wave(FMCW) radar to capture human reflections. Because falls are always accompanied by signal variation in height and velocity. We mount the radar on the ceiling, so that the radar could better capture the signal variation caused by falls.

2.2. Data Processing

FMCW radar transmits chirps by TX antenna and receives reflected signal by RX antenna. TX signal and RX signal will be mixed to produce intermediate frequency (IF) signal. To extract spatio-temporal features, we performed 3D Fast Fourier Transform (FFT) on IF signal [14, 15, 16]. Specifically, we first perform Range-FFT to separate the signals from different ranges. The relationship between the signal frequency and signal range is given by

$$d = \frac{f_{IF}c}{2k}, \tag{1}$$

where k is the slope of chirp and c is speed of signal. Then, we perform Doppler-FFT to extract the Doppler information caused by human movement. The relationship between the movement speed and doppler frequency is given by

$$v = \frac{\lambda\omega}{4\pi T_c}, \tag{2}$$

where T_c is the interval between two chirps and ω is the phase difference between two adjacent chirps. Since falls tend to be accompanied with high moving speed, we adopt high-pass filtering to suppress the velocity less than 0.22m/s.

Finally, we perform Angle-FFT to extract the angle information. The phase shift of the signal among different antennas is given by

$$\theta = \arcsin\frac{\lambda\omega}{2\pi l}, \tag{3}$$

where l is the distance between adjacent receiving antennas. After the 3D-FFT, the input signal has been transformed into a 3D tensor, which is too complex for real-time processing. To alleviate this problem, we sum the 3D tensor among the velocity dimension. Because of the partial loss of velocity information, we stack 40 frames as a input to CNN. The output of the data processing is shown in Fig. 2.

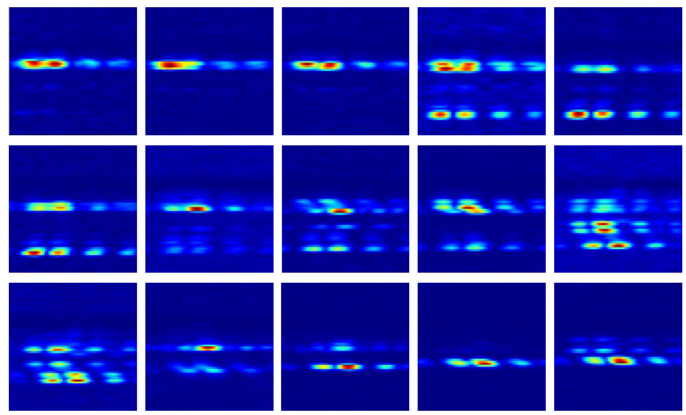

Fig. 2. Input image sequence of a fall, where the horizontal is angle, vertical is distance, and the value of pixel is the weight of velocity and power destiny.

2.3. Neural Network

The range-angle heatmap as shown in Fig. 2 is difficult to interpret intuitively. Hence, we adopt convolutional neural network (CNN) to achieve fall detection with the range-angle heatmap as input. To utilize the information among time axis, we design a 3D CNN rather than 2D ones to achieve information aggregation. The detailed architecture is demonstrated in Table 1.

Layer	Layer type	kernel size	stride	Output
0	Input	-	-	[1,40,64,64]
1	Conv.	7	(1,2,2)	[8,40,32,32]
2	Conv.	3	2	[16,20,16,16]
3	Conv.	3	2	[32,10,8,8]
4	MP	-	2	[32,5,4,4]
5	FC	-	-	64
6	FC	-	-	2

Table 1. Our model structure.

In elders' daily life, the amount of non-fall activities is much larger than the falls. As a result, the collected dataset is severely imbalance, of which the amount of non-fall data is eight times of fall data. To tackle this problem, we adopt Focal loss to train the neural network [17], which could make the neural network focus learning on hard negative examples. The loss function is defined as follows

$$L_{fl} = \begin{cases} -\alpha(1-y')^{\gamma} log y' & y=1 \\ -(1-\alpha)y'^{\gamma} log(1-y') & y=0, \end{cases} \quad (4)$$

where y' is prediction of y, γ is a hyper-parameter which can reduce the weight of easy positive and easy negative, and α can balance the disproportionate of positive and negative. In the proposed system, we set $\gamma = 2$ and $\alpha = 0.88$.

2.4. Data Augmentation

To further enhance the performance and robustness of the proposed system, we perform data augmentation to improve the diversity of the dataset.[18] We use two kinds of transformation to simulate the signal under different cases: *(1) Shifting and Flipping.* We perform shifting and flipping on the heatmap to simulate the fall at different locations. *(2) Interpolating and Decimating.* To simulate the speed variation of falls, we change the length of the input by interpolating and decimating the processing signal. The dataset size after data augmentation is 7 times larger than the original dataset.

2.5. Real-time Detection

We adopt a sliding window to select the signal and perform fall detection. If a fall occurs at the edge of the window, the misjudgment of the signal would lead to the performance degradation. Hence, the overlap of adjacent windows is set to 80% to prevent such case. To reduce the disturbing introduced by false positive detection, a fall is confirmed if it is detected in four consecutive time windows, or it would be ignored as a false positive.

3. EXPERIMENTS

In this section, we present the detailed design of and the results of the experiments.

3.1. Dataset

Our dataset is collected from 10 subjects with different height and weight. During data collection, the radar is fixed on the ceiling as shown in Fig. 3. The data is collected from 3 different environments with different height. The distance between radar and floor is 3.1m, 2.2m and 2.9m, respectively.

We collect 4 patterns of falls, falling forward, falling backwards, falling sitting in empty chair and falling sideward [19]. Each type of fall is performed 5 times in 4 different direction. Our dataset also includes 11 kinds of activities of daily living which are listed as follows: lie down, get up, squat, stand, stand up, sit, walk, run, jump, pick up and stoop. A total of 1600 fall sample and 12000 non-fall samples are collected. Specifically, there are 6 subjects are sampled in environment 1, 4 subjects in environment 2, and all of 10 subjects in environment 3, which compose 20 parts of dataset.

(a) Environment 1 (b) Environment 2

Fig. 3. Data collection environment.

3.2. Implementation Details

We use TI AWR1843 [20] Evaluation Board to collect FMCW radio data. The slope of the FMCW signal is 100MHz/us and a chirp last for 40 us. Each chirp is sampled 128 times. There are 64 chirps transmitted by per TX, and the interval between adjacent chirp is 160 us. The duration of each frame is 50ms. With these parameters, the maximum detection range is about 7m, and range resolution is about 5.7 cm. Max velocity is 1.8m/s and velocity resolution is 0.056m/s.

3.3. Evaluation Metrics

Due to the imbalance between positive samples and negative samples, detection accuracy could not characterize the system performance effectively. Since the output of the proposed system could be generally categorized in to true positives(TP), false positives (FP), true negatives (TN), and false negatives (FN), we adopt the metrics similar to [13] to evaluate the proposed system:

- Precision: $P = \frac{TP}{TP+FP}$
- Recall: $R = \frac{TP}{TP+FN}$
- F1 score: $F1 = \frac{2*p*r}{p+r}$.

3.4. Performance Evaluation

We evaluate the proposed system with different subjects and environments. Besides our model, we also implement SVM and a 2DCNN+LSTM model for comparison. We first roughly divide the dataset into non-test and testset. Then,

	Accuracy	Precision	Recall	F1 score	Parameters	FLOPs	Run time
SVM	0.869	0.371	0.165	0.228	-	-	-
2DCNN+LSTM	0.968	0.846	0.896	0.871	339K	37M	19.5ms
Resnet	0.987	0.915	0.986	0.949	14M	16G	106.5ms
Ours	0.988	0.936	0.969	0.952	184K	140M	3.9ms

Table 2. Performance and complexity comparison with other models.

80% data from non-test set are utilized for training and 20% for validation and parameter tuning.

3.5. Performance of different cases

To evaluate the performance of the proposed system under different cases, we adopt four kinds of methods to divide the training and testing dataset. *(1) Case 1.* Both the people and environment for data collection have not been seen in the training set. *(2) Case 2.* The environment keeps the same while the people has not been seen in the training set. *(3) Case 3.* The people keeps the same while the environment has changed. *(4) Case 4.* Both environment and people keep the same in training and testing set. The results are summarized in Table 3.

	Accuracy	Precision	Recall	F1 score
Case 1	0.970	0.954	0.783	0.870
Case 2	0.995	0.995	0.962	0.978
Case 3	0.985	0.992	0.884	0.936
Case 4	0.999	0.996	0.999	0.998

Table 3. Performance in different cases.

3.5.1. Evaluation of data augmentation

In this section, we evaluate the impact of data augmentation. The results are summarized in Table 4 and the ROC curve is shown in Fig 4. As shown in the table and figure, the data augmentation provides remarkable performance enhancement.

	Accuracy	Precision	Recall	F1 score
Without augment	0.970	0.954	0.783	0.870
With augment	0.988	0.936	0.969	0.952

Table 4. Performance of data augmentation.

3.5.2. Evaluation of model

We compared our 3DCNN model with a 2DCNN+LSTM model, SVM and cascade 3DResnet in [13]. All experiments are done in different people and environment after data augmentation. Model performance are shown by ROC curve in Fig.5. Since GPU may not be availble in practice, runtime is calculated on the CPU Intel i7-11700K. As shown in Table 2 and Fig. 5, the performance of our system is better than the others. What's more, the processing speed and model complexity are significantly better than that of 3DResnet in [13], which makes the proposed system be capable of achieving real-time fall detection.

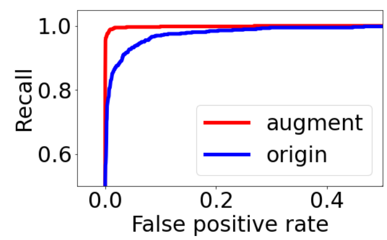

Fig. 4. Comparison of data augmentation by ROC curve.

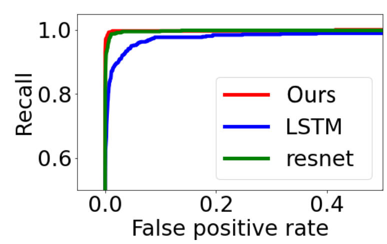

Fig. 5. Comparison of different baseline by ROC curve.

4. CONCLUSIONS

In this paper, we proposed a real-time fall detection system, mmFall, to utilize the millimeter FMCW radar to detect fall for elder care. The proposed mmFall first utilized the spatial-temporal signal processing to extract the signals related to the human activity, and then extracted the complex features of different fall pattern with a light-weight 3DCNN. Data augmentation was adopted to enhance the generalization of mmFall. By evaluating under different environments and people, mmFall could detect falls accurately in real time.

5. REFERENCES

[1] World Health Organization, "Falls," `https://www.who.int/news-room/fact-sheets/detail/falls` April. 26, 2021.

[2] Jay Chen, Karric Kwong, Dennis Chang, Jerry Luk, and Ruzena Bajcsy, "Wearable sensors for reliable fall detection," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, 2006, pp. 3551–3554.

[3] AK Bourke, JV O'brien, and GM Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm," *Gait & posture*, vol. 26, no. 2, pp. 194–199, 2007.

[4] Alan K Bourke and Gerald M Lyons, "A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor," *Medical engineering & physics*, vol. 30, no. 1, pp. 84–90, 2008.

[5] Yung-Chin Chen and Yi-Wen Lin, "Indoor rfid gait monitoring system for fall detection," in *2010 2nd International Symposium on Aware Computing*. IEEE, 2010, pp. 207–212.

[6] Robert Steele, Amanda Lo, Chris Secombe, and Yuk Kuen Wong, "Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare," *International journal of medical informatics*, vol. 78, no. 12, pp. 788–801, 2009.

[7] Giovanni Diraco, Alessandro Leone, and Pietro Siciliano, "An active vision system for fall detection and posture recognition in elderly healthcare," in *2010 Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)*. IEEE, 2010, pp. 1536–1541.

[8] Zhong Zhang, Christopher Conly, and Vassilis Athitsos, "A survey on vision-based fall detection," in *Proceedings of the 8th ACM international conference on PErvasive technologies related to assistive environments*, 2015, pp. 1–7.

[9] Adrián Núñez-Marcos, Gorka Azkune, and Ignacio Arganda-Carreras, "Vision-based fall detection with convolutional neural networks," *Wireless communications and mobile computing*, vol. 2017, 2017.

[10] Heilym Ramirez, Sergio A Velastin, Ignacio Meza, Ernesto Fabregas, Dimitrios Makris, and Gonzalo Farias, "Fall detection and activity recognition using human skeleton features," *IEEE Access*, vol. 9, pp. 33532–33542, 2021.

[11] Yuxi Wang, Kaishun Wu, and Lionel M Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2016.

[12] Liang Ma, Meng Liu, Na Wang, Lu Wang, Yang Yang, and Hongjun Wang, "Room-level fall detection based on ultra-wideband (uwb) monostatic radar and convolutional long short-term memory (lstm)," *Sensors*, vol. 20, no. 4, pp. 1105, 2020.

[13] Yonglong Tian, Guang-He Lee, Hao He, Chen-Yu Hsu, and Dina Katabi, "Rf-based fall monitoring using convolutional neural networks," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–24, 2018.

[14] Dongheng Zhang, Yang Hu, and Yan Chen, "Mtrack: Tracking multiperson moving trajectories and vital signs with radio signals," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3904–3914, 2020.

[15] Yan Chen, Hongyu Deng, Dongheng Zhang, and Yang Hu, "Speednet: Indoor speed estimation with radio signals," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2762–2774, 2020.

[16] Dongheng Zhang, Ying He, Xinyu Gong, Yang Hu, Yan Chen, and Bing Zeng, "Multitarget aoa estimation using wideband lfmcw signal and two receiver antennas," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7101–7112, 2018.

[17] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[18] Yadong Li, Dongheng Zhang, Jinbo Chen, Jinwei Wan, Dong Zhang, Yang Hu, Qibin Sun, and Yan Chen, "Towards domain-independent and real-time gesture recognition using mmwave signal," *arXiv preprint arXiv:2111.06195*, 2021.

[19] Lourdes Martínez-Villaseñor, Hiram Ponce, Jorge Brieva, Ernesto Moya-Albor, José Núñez-Martínez, and Carlos Peñafort-Asturiano, "Up-fall detection dataset: A multimodal approach," *Sensors*, vol. 19, no. 9, pp. 1988, 2019.

[20] Texas Instruments, "Awr1843 single-chip 76-ghz to 81-ghz automotive radar sensor integrating dsp, mcu and radar accelerator," `http://www.ti.com/product/AWR1843` Nov. 18, 2019.