GR-Fall: A Fall Detection System with Gait Recognition for Indoor Environments Using SISO mmWave Radar

CHENGZHEN MENG, University of Science and Technology of China, China
CHENMING HE, University of Science and Technology of China, China
DEQUAN WANG, University of Science and Technology of China, China
YUXUAN XIAO, University of Science and Technology of China, China
LINGYU WANG, University of Science and Technology of China, China
XIAORAN FAN, Independent Researcher, USA
LU ZHANG, Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China
YANYONG ZHANG*, University of Science and Technology of China, China and Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China

Fall detection is essential for safeguarding the health of elderly persons, enabling timely alerts to family members or the community. Millimeter-wave (mmWave) radar offers an effective solution, as it is privacy-preserving, non-invasive, and highly sensitive to motion. However, most existing approaches rely on multi-input, multi-output mmWave radar to generate 4D point clouds or range-angle heatmaps, significantly raising device costs. In this paper, we propose GR-Fall, a fall detection system with integrated gait recognition designed for indoor environments using single-input, single-output mmWave radar. To achieve high performance in various environments, we develop a data augmentation algorithm for target heatmaps and a cross-attention-based heatmap fusion framework for efficient fall detection. Furthermore, we introduce an innovative fall alarm mechanism based on joint fall-gait detection. This mechanism activates alerts when a person is detected having difficulty moving after a fall, thus minimizing unnecessary alarms and reducing strain on community resources. To evaluate GR-Fall, we recruit 33 volunteers and collect 5,799 instances across four different environments. Experimental results show that GR-Fall achieves 98.1% precision and 98.7% recall in new environments and with new participants, outperforming other state-of-the-art heatmap-based methods.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**.

Additional Key Words and Phrases: Wireless Sensing, mmWave Radar, Fall Detection

---

*Corresponding author.

Authors' Contact Information: Chengzhen Meng, University of Science and Technology of China, Hefei, Anhui, China, czmeng@mail.ustc.edu.cn; Chenming He, University of Science and Technology of China, Hefei, Anhui, China, hechenming@mail.ustc.edu.cn; Dequan Wang, University of Science and Technology of China, Hefei, Anhui, China, wdq15588@mail.ustc.edu.cn; Yuxuan Xiao, University of Science and Technology of China, Hefei, Anhui, China, xiaoyx@mail.ustc.edu.cn; Lingyu Wang, University of Science and Technology of China, Hefei, Anhui, China, wly18654382511@gmail.com; Xiaoran Fan, Independent Researcher, Sunnyvale, California, USA, gunanjiluzhe@gmail.com; Lu Zhang, Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, Anhui, China, luzha@ustc.edu.cn; Yanyong Zhang, University of Science and Technology of China, Hefei, Anhui, China and Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, Anhui, China, yanyongz@ustc.edu.cn.

ACM 2474-9567/2025/9-ART118
https://doi.org/10.1145/3749471

ACM Reference Format:
Chengzhen Meng, Chenming He, Dequan Wang, Yuxuan Xiao, Lingyu Wang, Xiaoran Fan, Lu Zhang, and Yanyong Zhang. 2025. GR-Fall: A Fall Detection System with Gait Recognition for Indoor Environments Using SISO mmWave Radar. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 3, Article 118 (September 2025), 26 pages. https://doi.org/10.1145/3749471

1 INTRODUCTION

Falls are the predominant cause of injury and death in the elderly population (aged 65 and older) in the United States according to the American Medical Association (AMA) [44]. Furthermore, the Centers for Disease Control and Prevention (CDC) report that approximately one-third of elders suffer at least one fall each year [37]. Therefore, fall detection plays a crucial role in safeguarding the health of the elderly. By promptly identifying falls, alerts can be sent to family members or community caregivers, enabling timely medical intervention and reducing the risk of severe complications.

Many sensors can be used for fall detection, which are typically classified into wearable and non-wearable device-based systems [48, 69]. Firstly, wearable devices, such as accelerometers [25, 33] and RFID systems [7, 43], may cause discomfort to users when used for extended periods. Additionally, these systems may fail if elderly individuals forget to wear or charge the devices. These limitations of wearable technologies can be addressed by using non-wearable alternatives. For example, camera [2, 14] is often used for fall detection by extracting features from multiple consecutive frame images. However, it is sensitive to lighting conditions and raises privacy concerns. Recently, the use of radio frequency (RF) signals, such as Wi-Fi and mmWave, for fall detection [9, 24, 27, 31, 39, 42, 48, 59, 61, 69] have garnered significant attention from researchers due to its non-invasive nature and privacy-preserving features. Fall detection using Wi-Fi [11, 39, 59, 61] analyzes changes in channel state information (CSI) to detect falls. However, its performance is susceptible to environmental variations, which limits its reliability. Moreover, mmWave radar, operating at high frequencies (GHz), stands out among RF-based devices due to its exceptional range resolution and precise motion detection capabilities [15, 30, 35, 36, 46, 63, 68, 73, 76]. Its ability to measure velocity enables it to effectively distinguish dynamic targets from the static environment, making it less susceptible to environmental changes. These advantages, combined with its non-invasive nature and privacy-preserving features, make mmWave radar a highly effective and reliable solution for fall detection.

Current mmWave radar fall detection techniques primarily rely on two types of data: point clouds and heatmaps. Point cloud-based methods [24, 26, 42, 69, 72] typically use height thresholds or traditional classifiers, such as Support Vector Machines (SVM), to identify falls. On the other hand, heatmap-based approaches [9, 10, 27, 47, 48] employ either traditional classifiers (e.g., KNN) or deep neural networks (e.g., CNN) to extract features from mmWave radar range-angle or range-Doppler heatmaps for fall detection. However, most of these methods require *Multi-Input Multi-Output (MIMO)* mmWave radar to generate 4D point clouds or range-angle heatmaps. This reliance on MIMO technology significantly increases device costs and power consumption, making these solutions less feasible for large-scale deployment.

Our goal is to design a low-cost fall detection system for indoor environments, leveraging *Single Input Single Output (SISO)* mmWave radar, which costs only a quarter of the price of MIMO mmWave radar [18, 20]. However, previous methods have not sufficiently accounted for environmental variations, fall location, or direction variations. Additionally, the limited information provided by SISO mmWave radar can lead to reduced performance. Moreover, accurately assessing the severity of falls remains unsolved. Specifically, we highlight the above issues and system design goals as follows:

- **Robust**: To reliably ensure personal safety, a fall detection system must be robust, delivering consistent performance despite variations in environment, fall location, or direction. These factors significantly affect the heatmap data generated by mmWave radar. However, current heatmap-based approaches [27, 48, 64] typically use heatmaps as direct input to the fall detection network without accounting for such variability.

Although these methods often involve collecting fall data across various conditions to improve system robustness, this process is labor-intensive and challenging to complete comprehensively.

- **Accurate**: Accuracy is essential for any fall detection system. However, SISO mmWave radar captures limited data, which can reduce performance when existing methods are applied. Specifically, point cloud-based approaches [24, 69] depend on 4D point clouds that incorporate a height dimension, while current heatmap-based techniques [27, 48, 64] typically employ range-angle heatmaps. Since SISO mmWave radar cannot generate either 4D point clouds or range-angle heatmaps, these methods are not directly applicable to SISO mmWave radar.
- **Practical**: Triggering an alarm for every fall is impractical, as not all falls are severe. For instance, in minor falls where the person is unharmed, triggering an alarm is unnecessary and consumes community resources. Some studies [48, 69] have proposed assessing fall severity by checking if the person stands up afterward. However, standing up does not guarantee the person's safety, as they may still struggle to move and require medical intervention.

In this paper, we present GR-Fall, a fall detection system with gait recognition designed for indoor environments, utilizing SISO mmWave radar. Given the absence of height information in the point cloud data from SISO mmWave radar, our approach leverages heatmap data for effective fall detection. GR-Fall offers a low-cost, robust, high-precision, and practical solution. The system consists of the following core components: (1) The *target extraction module* generates environment-independent target heatmaps by integrating clutter and noise filtering, clustering, and cropping of the Range-Doppler (RD) heatmap. This process effectively isolates the target from the original RD heatmap, thereby enhancing the robustness of GR-Fall. This process extracts the target's RD heatmap sequence (RDSeq) and micro-Doppler (MD) heatmap, which are subsequently used by the following modules. The RDSeq captures variations in the target's range and velocity over time through consecutive multi-frame heatmaps, while the MD heatmap directly reflects changes in velocity over time. (2) In the *fall detection module*, we begin with data augmentation that takes into account variations in the location and direction of fall data. This enriched dataset enhances the robustness of the fall detection network against different fall locations and directions. We then apply an attention mechanism to extract and fuse features from the target's RDSeq and MD heatmaps, fully leveraging information from both heatmap types to improve fall detection accuracy. (3) To further reduce unnecessary strain on community resources, we propose *an alarm mechanism based on joint fall-gait detection*. GR-Fall incorporates gait recognition to assess whether the target resumes normal walking patterns following a fall. Specifically, it calculates temporal and velocity variations from the MD heatmap as motion features and uses a Support Vector Machine (SVM) to determine if the person is walking normally. GR-Fall 's state machine then integrates outputs from both fall detection and gait recognition, triggering alarms only if normal gait is not detected following a fall. This design minimizes false alarms and unnecessary responses, ensuring the system is both reliable and practical.

We implement GR-Fall and deploy it in four distinct environments—a conference room, a break room, an office, and a living room—to thoroughly evaluate its effectiveness. The study involves 33 volunteers, consisting of 23 males and 10 females. During the experiments, we collect 901 instances of severe falls, 303 instances of non-severe falls, 952 instances of fall-like behaviors, and 3,643 instances of daily activities to assess GR-Fall's performance. The experimental results demonstrate that GR-Fall achieves 98.1% precision and 98.7% recall in new environments and with new participants, outperforming other state-of-the-art (SOTA) heatmap-based methods [27, 48].

In summary, the main contributions of our GR-Fall can be summarized as follows:

- We propose GR-Fall, a fall detection system with integrated gait recognition for indoor environments using SISO mmWave radar. Our system introduces a novel fall alarm mechanism that triggers an alert only if normal gait is not detected following a fall, thereby minimizing strain on community resources.

- We develop a set of techniques to improve the robustness, accuracy, and practicality of GR-Fall, including data augmentation for target heatmaps, cross-attention-based heatmap fusion for fall detection, and velocity-time variation-based gait recognition.
- We implement GR-Fall using commercial off-the-shelf (COTS) mmWave radar and evaluate its performance on a self-collected dataset spanning four distinct scenarios captured from 33 volunteers. Extensive experimental results demonstrate that GR-Fall achieves 98.1% precision and 98.7% recall in new environments and with new participants, outperforming other state-of-the-art (SOTA) heatmap-based methods.

2 PRELIMINARY

In this section, we begin by introducing the principles of mmWave radar sensing, with a focus on range and velocity estimation. Following that, we demonstrate the feasibility of using mmWave radar for fall detection.

2.1 Principles of mmWave Radar Sensing

The mmWave radar can effectively estimate the range and velocity of the target. It typically includes both transmitting and receiving antennas. The transmitting antenna (TX) sends out multiple chirps, which are reflected by objects and then captured by the receiving antenna (RX). These transmitted and received signals are mixed to produce the intermediate frequency (IF) signal. This IF signal is subsequently processed to estimate the target's range and velocity. The details of the range and velocity estimation processes are provided below.

Range Estimation. Due to the flight time, a frequency shift occurs between the transmitted and received signals. Consequently, the frequency of the IF signal (f_{IF}) is proportional to the distance (d) between the radar and the object [21, 28]. The formula for calculating the distance is given by:

$$f_{IF} = S\tau = \frac{S2d}{c} \Rightarrow d = \frac{f_{IF}c}{2S}, \tag{1}$$

where S is the slope of the chirp frequency change, τ is the round trip time, and c is the speed of light (3×10^8m/s).

Velocity Estimation. For a moving object, the chirp's phase varies with the range between the radar and the object. This results in a phase difference (ω) between adjacent chirps, which is proportional to the object's radial velocity (v), also referred to as Doppler velocity [21, 28]. The formula for calculating the velocity is given by:

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

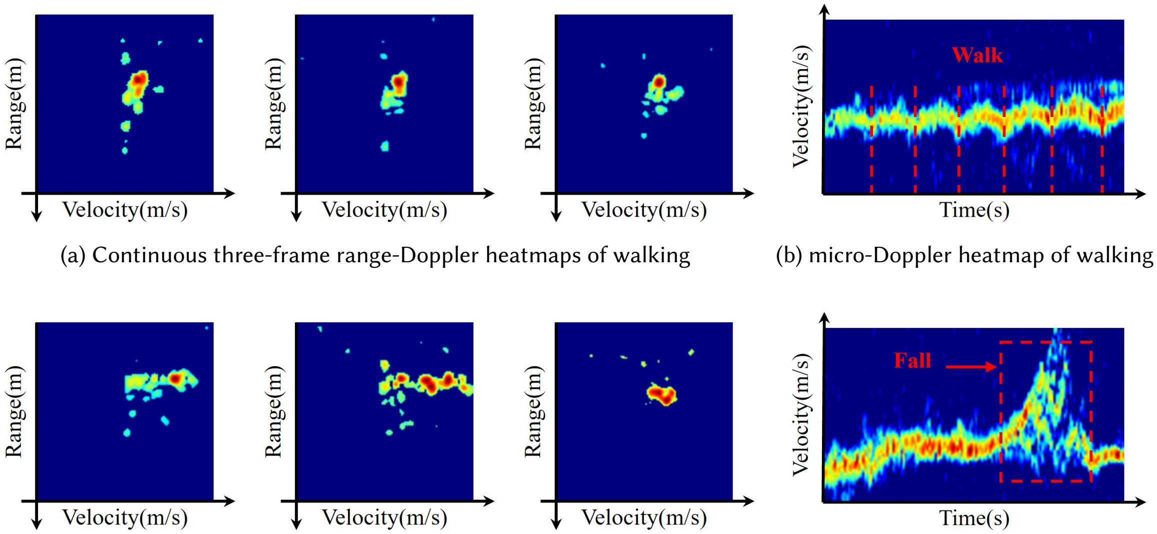
where λ is the wavelength, and T_c is the time interval for transmitting multiple chirps.

2.2 Feasibility of Detecting Falls Using mmWave Radar

Motion features, such as velocity, can be extracted from the range-Doppler (RD) heatmap [8] and micro-Doppler (MD) heatmap [12]. Specifically, building on the range estimation and velocity estimation discussed in Sec. 2.1, the RD heatmap is generated by performing two Fast Fourier Transform (FFT) on the sampled IF signal (I). The formula for calculating the RD heatmap (H) is as follows:

$$H = \left| \underset{j=1}{\overset{N_C}{F}} (\underset{i=1}{\overset{N_S}{F}} (I(i,j))) \right|, \ I \in C^{N_S \times N_C}, \ H \in R^{N_R \times N_V}. \tag{3}$$

Here, $F(\cdot)$ represents the Fast Fourier Transform (FFT), N_S is the analog-to-digital (ADC) sampling rate, N_C is the number of chirps, N_R is the dimension of range, and N_V is the dimension of velocity. As illustrated in Fig. 1(a) and Fig. 1(c), this RD heatmap is presented as a two-dimensional matrix. In this matrix, higher values indicate a higher probability of detecting a target at specific range and velocity coordinates. By leveraging multi-frame RD heatmaps, we can effectively distinguish between falls and other human activities, such as walking.



(a) Continuous three-frame range-Doppler heatmaps of walking (b) micro-Doppler heatmap of walking

(c) Continuous three-frame range-Doppler heatmaps of falling (d) micro-Doppler heatmap of falling

Fig. 1. Comparison of heatmaps of walking and falling. (a) and (c) display continuous three-frame range-Doppler heatmaps for walking and falling, respectively. (b) and (d) show micro-Doppler heatmaps for walking and falling, respectively.

Furthermore, the MD heatmap can be generated using FFT and Short-Time Fourier Transform (STFT) on the multi-frame IF signal. This heatmap, represented as a two-dimensional matrix with time and velocity dimensions, effectively captures variations in the target's velocity over time. As illustrated in Fig. 1(b) and Fig. 1(d), the velocity is relatively stable during walking and exhibits a distinct peak during a fall. The MD heatmap can also help distinguish between falls and other human activities.

3 SYSTEM DESIGN

3.1 Design Goals

To achieve a robust, accurate, and practical fall detection system for indoor environments using SISO mmWave radar, we have established the following goals for GR-Fall:

- GR-Fall should demonstrate robustness, maintaining consistent performance across different environments, fall locations, and directions.
- GR-Fall should achieve high precision in fall detection by effectively leveraging multiple heatmaps from the SISO mmWave radar.
- GR-Fall should provide practicality by activating alarms only for severe falls that significantly impair a person's mobility, thereby reducing unnecessary strain on community resources.

3.2 System Overview

To achieve the design goals mentioned above, we propose GR-Fall, a fall detection system with integrated gait recognition for indoor environments. Leveraging a SISO mmWave radar, GR-Fall provides a cost-effective alternative to MIMO mmWave radar systems for fall detection. Specifically, GR-Fall can detect falls across various environments, locations, and directions by employing heatmap-based target extraction and data augmentation

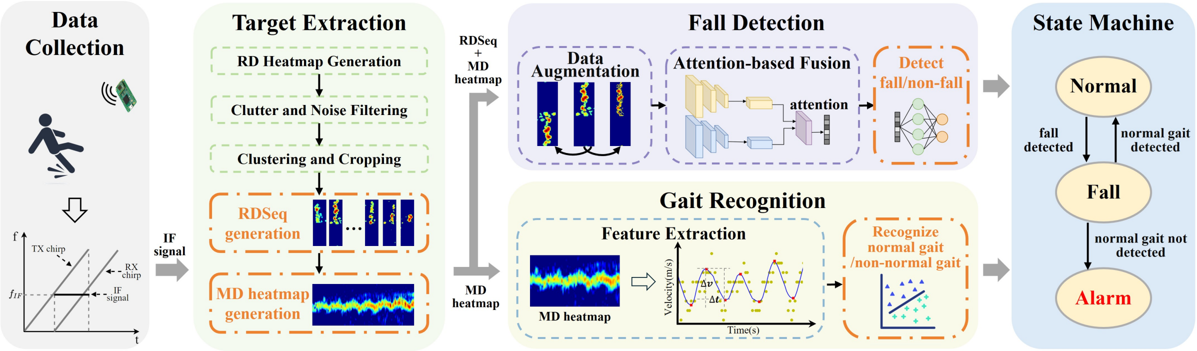


Fig. 2. GR-Fall includes several key components: target extraction, fall detection, and gait recognition. GR-Fall incorporates an alarm mechanism based on joint fall-gait detection, triggering an alarm only when the person is unable to walk normally after a fall. This approach effectively reduces the unnecessary use of community resources.

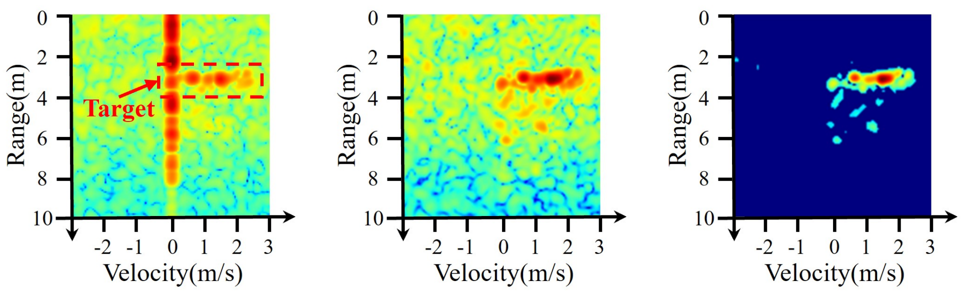
techniques (discussed in Sec. 3.3 and Sec. 3.4). In addition, GR-Fall harnesses multi-type heatmap data from the SISO mmWave radar, incorporating an attention mechanism to achieve high precision in fall detection (discussed in Sec. 3.4). Finally, GR-Fall includes an alarm mechanism based on joint fall-gait detection, integrating gait recognition to determine if normal walking patterns resume after a fall. An alarm is triggered only when the person cannot walk normally after falling, thereby minimizing unnecessary strain on community resources (discussed in Sec. 3.5).

Fig. 2 illustrates the overall architecture of GR-Fall, comprising five components: data collection, target extraction, fall detection, gait recognition, and a state machine. First, the data collection module captures the IF signal from the SISO mmWave radar. The target extraction module then processes this signal to generate the RD heatmap and further extract the target's RDSeq and MD heatmap through clutter and noise filtering, clustering, and cropping. In the fall detection module, data augmentation is first applied to expand fall data across various locations and directions by adjusting target energy intensity and flipping heatmap dimensions. An attention mechanism is then utilized to extract and fuse features from the RDSeq and MD heatmaps, significantly improving fall detection accuracy. Subsequently, the gait recognition module extracts motion features by analyzing gait velocity-time variations from the MD heatmap and employs a Support Vector Machine (SVM) to determine if normal walking behavior is present. Finally, the state machine module manages state transitions based on the results of fall detection and gait recognition, triggering an alarm only when a fall is detected and normal gait is absent. This approach effectively minimizes unnecessary strain on community resources.

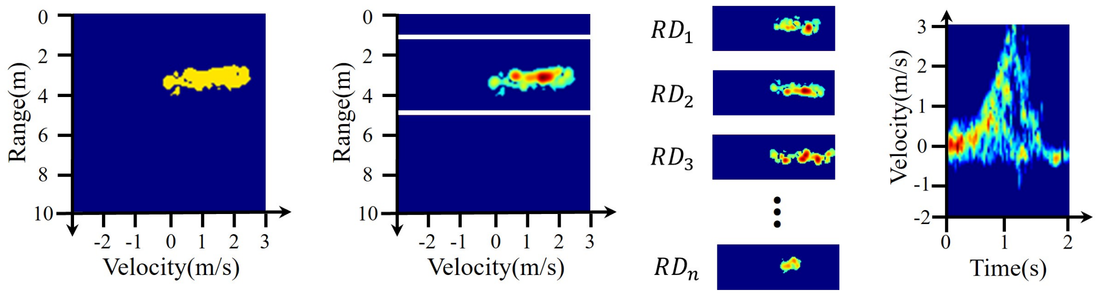
3.3 Heatmap-based Target Extraction

In this section, we extract the target's RD heatmap sequence (RDSeq) and MD heatmap for both the fall detection (Sec. 3.4) and gait recognition (Sec. 3.5). This is achieved through clutter and noise filtering, clustering, and cropping to the RD heatmap generated as described in Sec. 2.2. These processes ensure that the fall detection and gait recognition modules remain robust, consistently performing well regardless of environmental variations. Further details on this module are provided below.

RD Heatmap Generation. As described in Sec. 2.2, we first apply two FFTs to the mmWave radar IF signal to generate the RD heatmap. This RD heatmap is a two-dimensional matrix where higher values represent a greater likelihood of detecting a target at specific range and velocity coordinates. As illustrated in Fig. 3(a), this heatmap



(a) RD Heamap (b) Clutter Filtering (c) Noise Filtering



(d) Clustering (e) Cropping (f) RDSeq (g) MD Heatmap

Fig. 3. Results from each stage of the heatmap-based target extraction module. This module performs clutter and noise filtering, clustering, and cropping on the RD heatmap to generate the target's RDSeq and MD heatmap.

includes dynamic target, static clutter, and background noise. We denote the RD heatmap (H) as:

$$H = H_d + H_s + H_b. \tag{4}$$

Here, H_d represents the dynamic target, H_s denotes static clutter caused by radar reflections off stationary objects like walls and the ground, and H_b refers to background noise, which arises from weak signals generated in areas without objects, resulting from the radar's inherent electronic noise. In the following steps, we eliminate the static clutter and background noise to extract the target from the heatmap.

Clutter and Noise Filtering. We sequentially filter out the static clutter and the background noise from the RD heatmap (H). Firstly, we employ the Moving Target Indication (MTI) algorithm [3] to filter out static clutter (H_s). Since stationary targets do not exhibit phase changes between consecutive frames, subtracting adjacent frames effectively eliminates these stationary targets, effectively retaining the moving target. The formula is given by:

$$H_{d+b}(t) = H(t) - H(t-1). \tag{5}$$

Subsequently, we filter out background noise. During human activity, body velocity follows a continuous distribution, which is visualized on an RD heatmap by the high intensity of several adjacent Doppler bins. Since CFAR relies on local reference windows to estimate noise, if numerous bins in the Doppler domain exhibit high intensities, CFAR will select the bins with the highest intensities as valid points, potentially overlooking other bins containing velocity information. To address this, we remove weak background noise (H_b) by applying a predefined intensity threshold ($I_{threshold}$). After normalizing the heatmap values to a range between 0 and 1, we set any position in the heatmap that falls below this threshold to zero. In our implementation, $I_{threshold}$ is empirically set to 0.6. The formula is as follows:

$$H_d(i,j) = \begin{cases} H_{d+b}(i,j), & H_{d+b}(i,j) \geq I_{threshold} \\ 0, & H_{d+b}(i,j) < I_{threshold} \end{cases}, \tag{6}$$

As illustrated in Fig. 3(b) and Fig. 3(c), we effectively eliminate the static clutter with higher intensity and the background noise with lower intensity from the RD heatmap.

Clustering and Cropping. We further process the filtered heatmap (H_d) by clustering and cropping to extract the target. Firstly, we employ Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [13] to group data based on density in the range-Doppler space. This identifies dense regions as clusters (targets) and sparse regions as noise. Subsequently, we identify the position with the highest intensity within each cluster as the center and crop the area of the heatmap for the target. The formula is given by:

$$\begin{cases} i_{start} = \underset{i}{argmax}\ D(H_d)_{i,j} - N_L/2 \\ i_{end}\ = \underset{i}{argmax}\ D(H_d)_{i,j} + N_L/2 \end{cases}, \ H_d \in R^{N_R \times N_V}, \tag{7}$$

$$RD(i,j) = H_d(i_{start} + i, j), \text{ for } i = 0 \text{ to } i_{end} - i_{start}, \tag{8}$$

where $D(\cdot)$ denotes the DBSCAN clustering process, $i_{start} \sim i_{end}$ indicates the cropping range, N_R represents the range dimension, N_V represents the velocity dimension, and N_L specifies the length after cropping. In our implementation, we set $N_L = 40$. As illustrated in Fig. 3(d) and Fig. 3(e), we cluster to identify the target and then crop it from the heatmap.

RDSeq and MD Heatmap Generation. Finally, we generate the RDSeq and MD heatmap of the target. We set a time window and crop all frames within this window based on the cropping parameters from the data of the first frame. As illustrated in Fig. 3(f), the resulting heatmaps, known as the target's RDSeq, effectively capture variations in the target's range and velocity compared to the first frame of the time window. The sequence is defined as follows:

$$\text{RDSeq} = \{RD_i \mid 1 \leq i \leq N_T\},\ RD_i \in R^{N_L \times N_V}, \tag{9}$$

where N_T represents the number of frames in the time window. In our implementation, we set $N_T = 40$.

In addition, we sum each frame in the RDSeq along the range dimension and stitch them together in time order to generate the MD heatmap. As illustrated in Fig. 3(g), the MD heatmap (H_{MD}) highlights changes in velocity over time. The corresponding formula is:

$$H_{MD}(i,j) = \sum_{k=1}^{N_L} RD_i(k,j),\ RD_i \in \text{RDSeq},\ H_{MD} \in R^{N_T \times N_V}. \tag{10}$$

In summary, RDSeq is especially valuable as it captures variations in target range and velocity over time through consecutive multi-frame heatmaps, while the MD heatmap directly reflects changes in velocity over time. Both heatmaps respond to the target's movement and can be effectively leveraged for subsequent fall detection and gait recognition tasks.

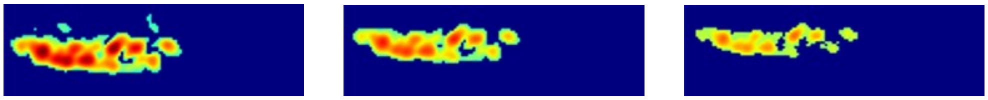
3.4 Cross-attention-based Heatmap Fusion for Fall Detection

After performing heatmap-based target extraction as described in Sec. 3.3, we obtain the target's RDSeq and MD heatmap. In this section, we first apply data augmentation to create training sets that encompass various fall locations and directions, thereby enhancing the robustness of our fall detection network. Next, we employ an attention mechanism to extract and fuse features from the RDSeq and MD heatmaps, further improving the precision of fall detection. Further details are provided below.

(a) Original heatmap

(b) Medium-range energy attenuation

(c) Long-range energy attenuation



(d) Original heatmap with dimension flip

(e) Medium-range energy attenuation with dimension flip

(f) Long-range energy attenuation with dimension flip

Fig. 4. Results of data augmentation: (a) shows the original target's heatmap; (b) and (c) illustrate heatmaps augmented by range; (d), (e), and (f) depict heatmaps augmented by direction.

Data Augmentation. The robustness of neural networks heavily relies on the quantity and quality of the training dataset [30]. However, collecting fall data from all possible locations is labor-intensive and challenging to achieve fully. To overcome this, we propose a data augmentation framework to generate a large volume of effective training data encompassing diverse fall location variations based on the limited available fall data. As a result, the amount of data after augmentation is five times greater than before, significantly enhancing the diversity and effectiveness of the training data.

Firstly, according to the mmWave radar signal power decay equation [41], the power received by the radar's receiving antenna is inversely proportional to the fourth power of the target's range from the radar. Therefore, for the current fall location data, we can get data for different fall locations by adjusting the corresponding power. The formula is given by:

$$H_A(i,j) = H(i,j) - 10 \times log(\frac{(d_i + \triangle d)^4}{d_i^4}),\ H \in R^{N_R \times N_V}, \tag{11}$$

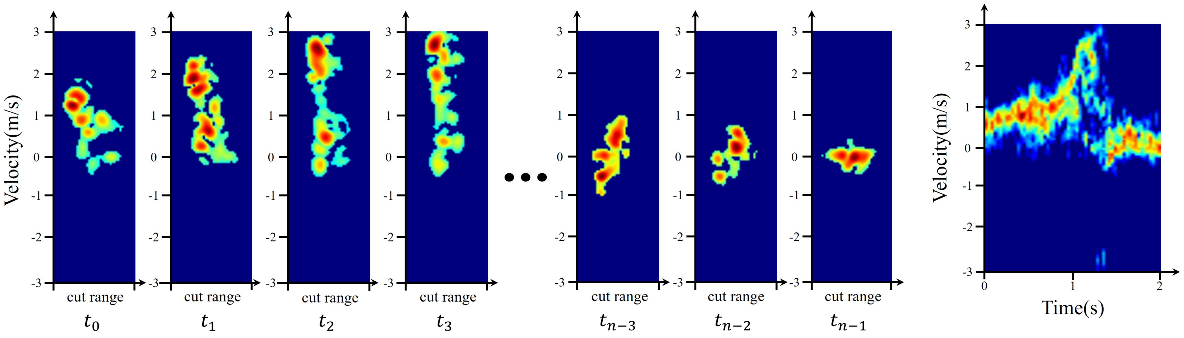
where H represents the RD heatmap, d_i represents the distance between the target located in row i of the heatmap and the radar, and $\triangle d$ represents a change in distance. The augmented data are illustrated in Fig. 4(b) and Fig. 4(c).

In addition, by flipping the range and velocity dimensions of the heatmap, we can simulate fall data for subjects facing the radar from the opposite direction, thereby enriching our training dataset. The corresponding formula is as follows:

$$H_A(i,j) = H(N_R - i, N_V - j),\ H \in R^{N_R \times N_V}, \tag{12}$$

where $(N_R - i)$ represents the flip of the range dimension and $(N_V - j)$ represents the flip of the velocity dimension. The augmented data are illustrated in Fig. 4(d), Fig. 4(e), and Fig. 4(f).

Attention-based Encoder and Fusion. Both the RDSeq and MD heatmap generated in Sec. 3.3 capture critical motion features of the target, which are essential for accurate fall detection. As shown in Fig. 5(a), RDSeq is particularly informative, capturing variations in target range and velocity over time through consecutive multi-frame heatmaps, consisting of 40 frames in our setup. In contrast, Fig. 5(b) illustrates that the MD heatmap offers a more intuitive representation by condensing the entire process into a single frame, directly reflecting changes in velocity over time. To leverage both strengths, we design an attention-based fall detection network that extracts and fuses features from RDSeq and MD heatmap for precise fall detection. The overview of our attention-based fall detection network is presented in Fig. 6.



(a) RD Heatmap Sequence (RDSeq) (b) MD Heatmap

Fig. 5. RDSeq and MD heatmap of the target during a fall. RDSeq effectively captures changes in target range and velocity over time through multi-frame sequences (40 frames in our setup), while the MD heatmap condenses the entire speed change process into a single frame, providing an intuitive representation of velocity shifts over time.

We begin by encoding the features of the RDSeq and MD heatmap separately. Since heatmaps are two-dimensional arrays, 2D CNNs are well-suited for extracting relevant information. For the RDSeq, we use a pre-trained ResNet-18 [17], excluding the final global pooling and fully connected layers, to extract features within the range-velocity space of the heatmaps. To capture the temporal information from the multi-frame data in RDSeq, we further utilize a transformer encoder with an attention mechanism [55] for feature extraction, denoted as F_{RDSeq}. Similarly, to capture the range and temporal information of the MD heatmap (H_{MD}), we also use a pre-trained ResNet-18 [17], supplemented by a multilayer perceptron (MLP), to extract features denoted as F_{MD}. The formula is provided below:

$$F_{RDSeq} = TE(Concat(HE(RD_1), ..., HE(RD_n))),\ RD_i \in \text{RDSeq}, \tag{13}$$

$$F_{MD} = MLP(HE(H_{MD})),\ H_{MD} \in R^{N_T \times N_V}, \tag{14}$$

where $HE(\cdot)$ denotes the heatmap encoder, $TE(\cdot)$ denotes the transofemer encoder, $Concat(\cdot)$ represents the feature vector concatenation, and $MLP(\cdot)$ represents the multilayer perception.

Additionally, we need to fuse the features extracted from the RDSeq (F_{RDSeq}) and the MD heatmap (F_{MD}). Direct vector concatenation alone is insufficient for effective information transfer between the two data sources. Inspired by [4], we employ a cross-attention mechanism for feature fusion, allowing comprehensive integration of information from both feature sets to enhance fall detection accuracy. In the cross-attention process, the query (Q) represents the critical information needed for the current task. During computation, Q is matched with the key (K) to identify the most relevant information from another data source, the value (V), which contains the specific information corresponding to K. When Q finds a match with the K, the content of V is used to respond to the query, effectively integrating relevant information. A crucial aspect of this approach is the selection of the Q, K, and V vectors. In our task, the MD heatmap more directly reflects movement changes while the RDSeq captures richer movement details. Therefore, we designate F_{MD} as Q and use F_{RDSeq} for both K and V, enabling F_{RDSeq} to query relevant information from F_{MD}, thereby extracting more efficient information from F_{RDSeq} with the support of F_{MD}. The following formula represents the process:

$$Q = F_{MD},\ K = V = F_{RDSeq}, \tag{15}$$

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V, \tag{16}$$

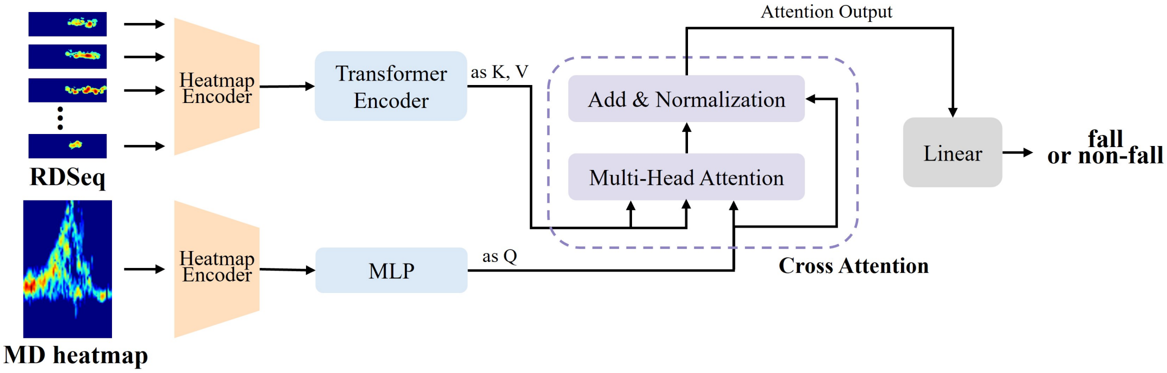


Fig. 6. The overview of our attention-based fall detection network. It uses an attention mechanism to extract and fuse features from the RDSeq and MD heatmap to detect falls.

where $\frac{1}{\sqrt{d_k}}$ represents the scaling factor.

Finally, after fusing the features through a cross-attention mechanism, we apply a linear layer followed by a softmax function to classify fall and non-fall events.

Loss Function. Due to the collection of fewer fall events compared to non-fall events, resulting in a sample imbalance in the data, we employed Focal Loss [32] as the loss function. This function can effectively reduce the impact of easy-to-classify samples on the overall loss and increase the emphasis on those that are difficult to classify. The formula is as follows:

$$FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} log(p_t), \tag{17}$$

where α_t are the weighting factors used to balance the positive and negative samples, and γ is employed to adjust the loss function's focus on samples of varying difficulty.

3.5 Velocity-time Variation-based Gait Recognition

In this section, we recognize the target's gait using motion features extracted from the MD heatmap generated in Sec. 3.3. As discussed in [69], focusing on life-threatening falls is essential to conserve precious community resources. If person can walk normally after a fall, they are either uninjured or capable of seeking help on their own, making an alarm unnecessary. To this end, we design a gait recognition module in GR-Fall to assess whether the target resumes normal walking patterns after a fall. Then, our GR-Fall triggers alarms only in cases of significant falls that impair the person's ability to walk, thus reducing the unnecessary waste of community resources. Detailed information is provided below.

Firstly, since the target's MD heatmap directly reflects changes in velocity over time, we extract the velocity of the target torso using the percentile method described in [54]. Specifically, this method calculates the weighted median of target velocities from the MD heatmap to accurately determine the target's torso velocity. The formula is provided below:

$$P(v,t) = \frac{\sum_{v=v_{\min}}^{v} H_{MD}(v,t)}{\sum_{v=v_{\min}}^{v_{\max}} H_{MD}(v,t)}, \tag{18}$$

where $P(v,t)$ represents the percentile in velocity at v and time at t, and $H_{MD}(v,t)$ denotes the intensity in velocity at v and time at t of the MD heatmap. The torso velocity is determined when $P(v,t) = 50\%$. The target's

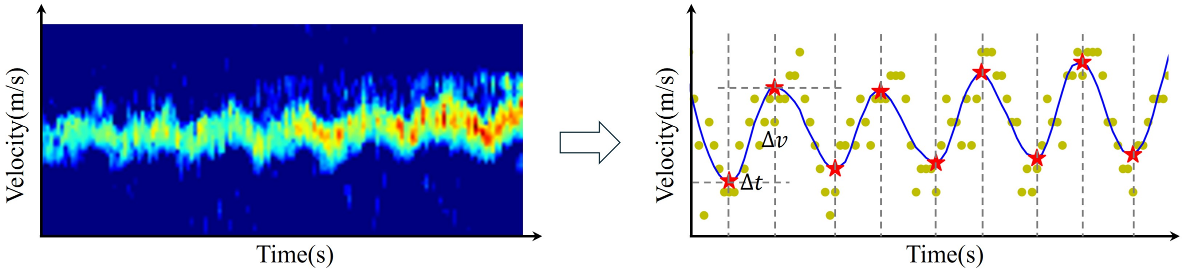


Fig. 7. Motion features extracted from the MD heatmap in the gait recognition module. The MD heatmap is displayed on the left, with yellow dots indicating torso velocities, blue dots representing smoothed data, and red pentagons marking peak velocities on the right.

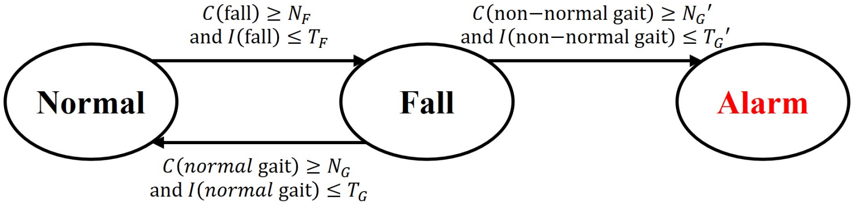


Fig. 8. The state machine of our GR-Fall consists of three states: normal, fall, and alarm. Here, $C(\cdot)$ represents the number of detected frames, while $I(\cdot)$ denotes the time interval between two consecutive detected frames. An alarm is triggered by GR-Fall only if normal gait is not detected after a fall.

MD heatmap (H_{MD}) is displayed on the left side of Fig. 7, and the extracted torso velocities are marked by yellow dots on the right side of the figure.

Furthermore, we employ the Savitzky-Golay filter [45] to smooth the data and utilize the Scipy [56] to identify the peak velocities of the target's movement. These peak velocities can be used to construct a motion signature for gait recognition. The formula for the maximum peak velocity is as follows:

$$v[p] > v[p-1],\ v[p] > v[p+1],\ \text{and } v[p] - \min(v[l], v[r]) \geq \text{prom}_{\text{min}}, \tag{19}$$

Here, $v[p]$ represents the peak velocity, while l and r denote the lowest valley positions on the left and right sides, respectively. The parameter prom_{min} is the minimum prominence threshold, which we empirically set to 1. The minimum peak is calculated similarly. As illustrated on the right side of Fig. 7, the blue dots represent the smoothed velocity data, while the red pentagon marks the detected peak velocities.

We then calculate the temporal and velocity variations ($\triangle t$ and $\triangle v$) between consecutive peaks to construct the target's motion features for gait recognition. This can capture variations in the target torso velocity during movement. The formula is defined as follows:

$$F = \{v_1 - v_0,\ t_1 - t_0,\ v_2 - v_1,\ t_2 - t_1,\ \ldots,\ v_n - v_{n-1},\ t_n - t_{n-1}\}, \tag{20}$$

where $(v_0, \ldots, v_n)$ denote the peak velocities, $(t_0, \ldots, t_n)$ are the corresponding times to these peak velocities, and n denotes the number of peaks. Utilizing these motion features, we employ an SVM classifier [5] to determine whether the target's gait is normal.

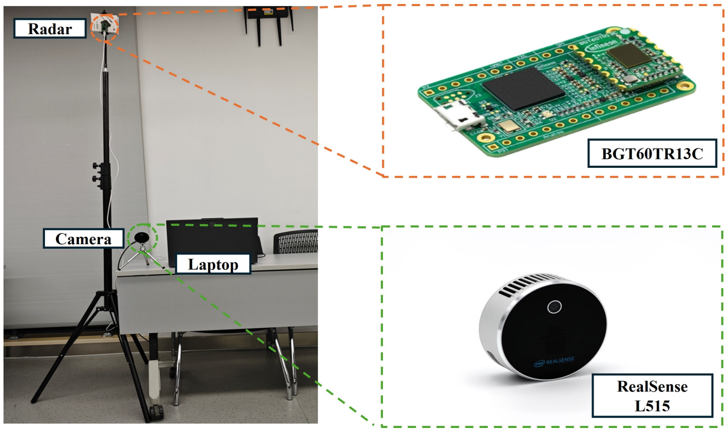


Fig. 9. The data collection platform of our GR-Fall. In this setup, the radar data is used as input for GR-Fall, while the camera is utilized solely for labeling the radar data to generate the ground truth.

Finally, we design a state machine based on the results from our fall detection module (Sec. 3.4) and gait recognition module (Sec. 3.5). Initially set to the *normal state*, the machine transitions to the *fall state* if the fall is detected over N_F consecutive frames. While in the *fall state*, if the normal gait is recognized in N_G consecutive frames, the machine returns to the *normal state*; otherwise, it shifts to the *alarm state*. This approach ensures that GR-Fall triggers alarms only when normal gait is not detected following a fall, thereby minimizing unnecessary use of community resources. The state transition process is illustrated in Fig. 8, with the formula given by:

$$\text{Normal} \Rightarrow \text{Fall:} \quad C(\text{fall}) \geq N_F \text{ and } I(\text{fall}) \leq T_F, \tag{21}$$

$$\text{Fall} \Rightarrow \text{Normal:} \quad C(\text{normal gait}) \geq N_G \text{ and } I(\text{normal gait}) \leq T_G, \tag{22}$$

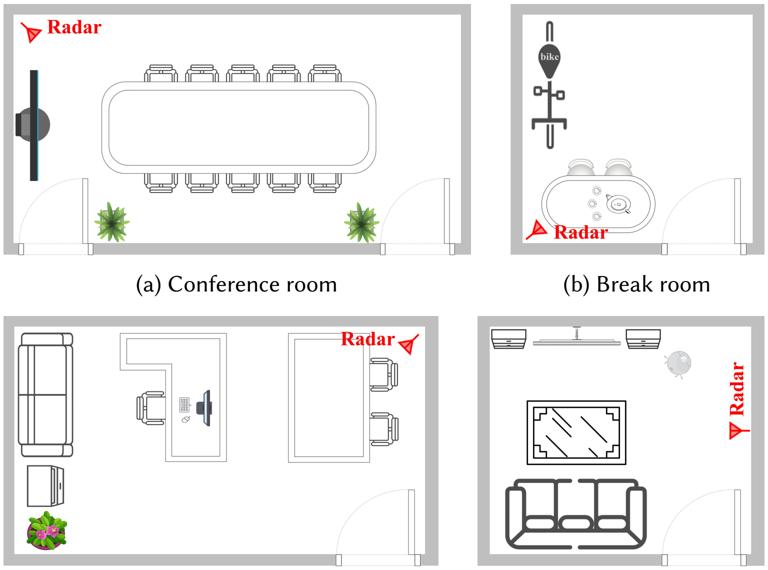
$$\text{Fall} \Rightarrow \text{Alarm:} \quad C(\text{non-normal gait}) \geq N'_G \text{ and } I(\text{non-normal gait}) \leq T'_G, \tag{23}$$

where $C(\cdot)$ reprsents the number of detected frames and $I(\cdot)$ denotes the time interval between two consecutive detected frames.

4 IMPLEMENTATION

Our GR-Fall is built on the Infineon's BGT60TR13C [19], a single-chip mmWave radar. In particular, we utilize this radar's one transmitting antenna and one receiving antenna (TX1 and RX1). The TX1 emits radio signals into the environment, and the RX1 captures the signals reflected by targets. These signals are then mixed to produce the IF signal, which is sampled and sent to the laptop for further analysis. Furthermore, the radar start and end frequencies are set to 60.5 GHz and 61.5 GHz, respectively. The number of samples per chirp is set to 128, and the number of chirps per frame is set to 64. The radar operates at a frame rate of 30 fps. With this configuration, the mmWave radar system achieves a range resolution of 0.15 m and a velocity resolution of 0.09 m/s.

In addition, our GR-Fall's fall detection network is implemented by using the PyTorch framework [40]. Our experimental platform consists of an Intel Core i5-11500 CPU [22] and an NVIDIA GeForce RTX 3060 GPU [38]. On this platform, the training time is 40 minutes per epoch. During training, we set the batch size to 8, use the Adam optimizer, and initialize the learning rate at 1×10^{-5}, which decreases following a cosine annealing schedule. The training process is conducted over a maximum of 50 epochs. For the loss function, as defined in Eq. (17), the parameters are set to $\alpha_t = 0.75$ and $\gamma = 2$.



(a) Conference room (b) Break room

(c) Office (d) Living room

Fig. 10. The layout of four experimental scenarios for evaluation: a conference room ($7m \times 8m$), a break room ($3.5m \times 3.5m$), an office ($4m \times 5m$), and a living room ($3m \times 4.5m$). The living room contains a dynamic sweeping robot. Red squares indicate the placement locations of the mmWave radar.

5 EVALUATION

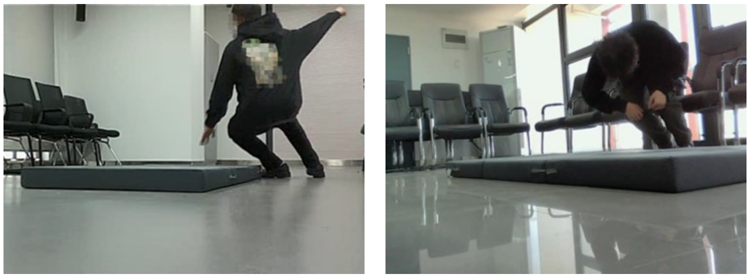
5.1 Experiment Setup

Data Collection. Our data collection platform, shown in Fig. 9, collects data from both the mmWave radar (Infineon BGT60TR13C [19]) and the camera (Intel RealSense L515 [23]). We ensure time synchronization between these two sensors, which is crucial for subsequent training and testing. In this configuration, the mmWave radar data serves as the input for GR-Fall, while the camera images are used exclusively for labeling the radar data to establish ground truth. This ground truth is essential for GR-Fall's network training and performance evaluation.

Furthermore, to comprehensively evaluate the performance of our GR-Fall, we deploy it across four distinct scenarios: a conference room ($7m \times 8m$), a break room ($3.5m \times 3.5m$), an office ($4m \times 5m$), and a living room ($3m \times 4.5m$). The living room contains a dynamic sweeping robot. The layout of each scenario is depicted in Fig. 10. Notably, our study has received approval from the Institutional Review Board (IRB). We recruit 33 volunteers, 23 males and 10 females, who participate in severe falls, non-severe falls, fall-like behaviors, and daily activities within each scenario. To ensure the realism of the falling actions performed by the volunteers, we collect a set of representative fall-related videos [49–52, 65–67] from YouTube and TikTok for them to study and replicate. Furthermore, the supplementary material includes a video recording that documents the data collection process. The specific actions performed are as follows:

- *Severe falls:* Falling patterns include slipping, tripping, losing consciousness, losing balance, and falling from a chair. Specifically, volunteers remain on the ground for an extended period after falling (over 15 seconds) or pretend to have injured a leg upon getting up, walking with a limp;
- *Non-severe falls:* The volunteers stand up after falling and walk normally, with the falling patterns being the same as described above (slipping, tripping, losing consciousness, etc.);
- *Fall-like behaviors:* These include sitting down, picking up objects, and crouching down to tie shoes;

(a) Slip (b) Trip (c) Lose consciousness



(d) Lose balance (e) Fall from chair

Fig. 11. Sampled fall patterns in our dataset.

- *Daily activities:* This category encompasses walking, eating, drinking, sweeping the floor, opening the door, sitting and playing with a mobile phone, and stretching arms and legs, among others.

Train/Test Split. Our dataset comprises 901 instances of severe falls, 303 instances of non-severe falls, 952 instances of fall-like behaviors, and 3,643 instances of daily activities. In our subsequent experiment, we select 336 sets of fall data, 330 sets of fall-like behavior data, and 1,528 sets of daily activity data from the conference room scenario for training (Fig. 10(a)). Meanwhile, data from the other three scenarios (Fig. 10(b), Fig. 10(c), and Fig. 10(d)) are reserved for testing purposes. Additionally, examples of falls from our dataset are shown in Fig. 11.

5.2 Evaluation Metrics

Since fall actions are low-probability events, the fall data in our collected dataset is less frequent than the non-fall data, which closely mirrors real-life scenarios. Although non-fall data is more abundant, our evaluation metrics prioritize fall data to ensure that excessive non-fall data does not interfere with our assessment of fall detection performance. To clarify the metrics, we abbreviate True Positives, False Negatives, True Negatives, and False Positives as TP, FN, TN, and FP, respectively. Given that falls occur much less frequently than non-fall behaviors in real-life scenarios, accuracy ($\frac{TP+TN}{TP+TN+FP+FN}$) is no longer a reliable metric for fall detection. True Negatives (TN) dominate the calculation as long as the non-fall data are correctly classified, meaning that even if all data are classified as non-falls, accuracy will remain high. Therefore, we exclude TN from the evaluation metric formula and focus on fall actions, specifically the ratio of correctly detected falls to all detected falls (precision) and the ratio of correctly detected falls to the total number of falls (recall).

- *Precision:* The ratio of correctly detected falls to all detected falls, defined as $p = \frac{TP}{TP+FP}$.
- *Recall:* The ratio of correctly detected falls to the total number of falls, defined as $r = \frac{TP}{TP+FN}$.
- *F1 Score:* The harmonic mean of precision and recall, calculated as $F1 = \frac{2\times p\times r}{p+r}$.

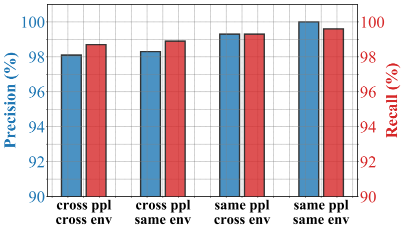


Fig. 12. Impact of the environments (env) and people (ppl).

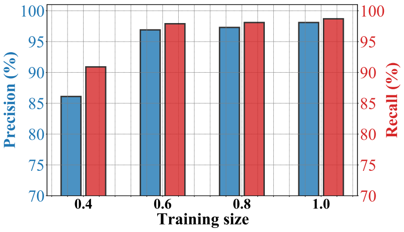


Fig. 13. Impact of the training size.

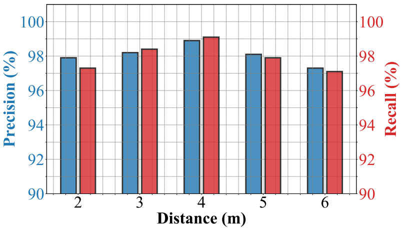


Fig. 14. Impact of the distance.

5.3 Performance of GR-Fall

Overall Performance. We evaluate the overall performance of GR-Fall and compare it to state-of-the-art (SOTA) heatmap-based methods [27, 48] on our dataset. Specifically, mmFall [27] utilizes a 3D CNN to extract features from multi-frame range-angle heatmaps, which we apply to the RDSeq due to the lack of angle information in SISO radar. Similarly, Aryokee [48] employs a 3D CNN to extract features from two types of range-angle heatmaps, followed by feature concatenation, which we apply to the RDSeq and MD heatmap.

The results presented in Tab. 1 are based on the aforementioned evaluation metrics. Our GR-Fall achieves an F1 Score of 98.4%, a precision of 98.1%, and a recall of 98.7%, demonstrating the best performance among all baselines. In comparison, the performance of other methods is as follows: mmFall (87.4% F1 Score, 86.4% precision, 88.3% recall) and Aryokee (92.0% F1 Score, 91.3% precision, 92.8% recall). Their poor performance can be attributed to the lack of target extraction, as these methods utilize the raw heatmap directly as input to the fall detection network, making them susceptible to environmental influences. Furthermore, Aryokee simply concatenates the extracted features, failing to fully leverage the information from multiple types of heatmaps. Whereas, our GR-Fall employs clutter and noise filtering, clustering, and cropping to effectively extract targets, ensuring robustness and consistent performance despite environmental variations (as discussed in Sec. 3.3). Additionally, we utilize an attention mechanism to extract and fuse features from both the RDSeq and MD heatmap, enhancing the precision of fall detection (as discussed in Sec. 3.4).

Impact of Different Environments and People. We also analyze the performance of GR-Fall when tested in the same vs. cross environments and people. The training dataset remains unchanged, consisting solely of data from the conference room (Fig. 10(a)). The test data, however, includes variations such as new people in new environments (Fig. 10(b) and Fig. 10(c)), new people in the same environment, same people in new environments, and same people in the same environment. As shown in Fig. 12, GR-Fall achieves over 98% precision and recall across all test cases. Specifically, under the cross-people and cross-environment setting, GR-Fall achieves 98.1% precision and 98.7% recall. If we further change to the same environment, precision and recall improve slightly by

Table 1. Overall performance of GR-Fall.

Method	F1 Score	Precision	Recall
mmFall [27]	87.4%	86.4%	88.3%
Aryokee [48]	92.0%	91.3%	92.8%
GR-Fall	**98.4%**	**98.1%**	**98.7%**

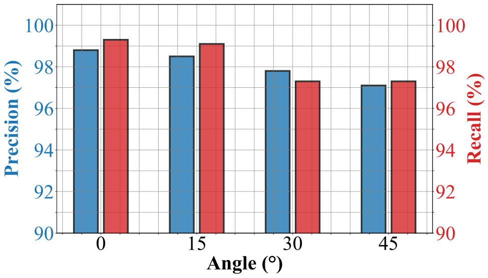


Fig. 15. Impact of the angle.

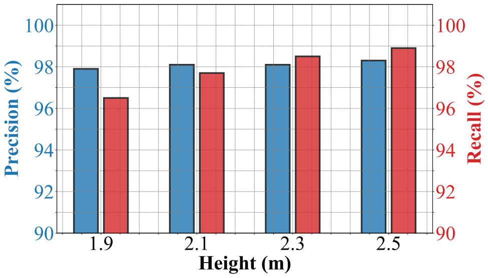


Fig. 16. Impact of the height.

1.2% and 0.6%, respectively. The results confirm that since GR-Fall works well across environments and people, it also works well on the same environments and people, and its performance is even slightly higher. This consistent performance is attributed to our heatmap-based target extraction module (Sec. 3.3), which ensures GR-Fall's resilience to environmental changes.

Impact of Different Training Size. To examine the impact of training set size on the GR-Fall, we vary the training data size from 40% to 100% of the total samples, increasing in 20% increments. Fig. 13 presents the precision and recall values for fall detection at each training set proportion. The results indicate that when the training set size increases from 40% to 60%, precision and recall quickly exceed 95%. Beyond a training set proportion of 60%, the improvement stabilizes. This suggests that by the time the training set reaches 60%, GR-Fall has already learned the majority of fall-related features.

Impact of Different Distances. We evaluate the performance of GR-Fall at varying distances (2m, 3m, 4m, 5m, and 6m), as shown in Fig. 14. The distance affects the intensity of the reflected signals, with signal strength decreasing as distance increases. However, thanks to GR-Fall's data augmentation module during training (Sec. 3.4), GR-Fall consistently achieves over 97% precision and recall even at greater distances. This demonstrates the robustness of GR-Fall for varying distances between the radar and the targets. The best performance is observed at 4 meters, as the radar's receiving antenna exhibits varying gain levels in different directions, with the highest gain directly in front of the radar antenna.

Impact of Different Angles. We evaluate the performance of GR-Fall at varying angles (0°, 15°, 30° and 45°), as shown in Fig. 15. Influenced by the radar beam's directionality and the presence of sidelobes, the energy of radar signals varies at different angles, attenuating significantly as the angles increase. The experimental results demonstrate that GR-Fall consistently achieves over 97% precision and recall, even at larger angles. The best performance occurs at 0°, where the radar directly faces the target.

Impact of Different Heights. We evaluate the performance of GR-Fall at different heights (1.9m, 2.1m, 2.3m, and 2.5m), as illustrated in Fig. 16. The radar's height primarily affects its sensing area, which is smaller at lower heights. The experimental results show that GR-Fall maintains a precision of over 97% across all heights. Moreover, as the height increases to 2.1m, the recall of GR-Fall exceeds 97%, attributed to the expanded sensing area. These results demonstrate GR-Fall's robustness to variations in radar height.

Sensing Coverage. The sensing coverage of the mmWave radar depends on its sensing distance, field of view (FOV), and placement configuration. Our experimental results (Fig. 14, Fig. 15, and Fig. 16) show that our GR-Fall achieves a sensing range of 6m with a FOV of ±45°. When mounted at a height of 2.5m and tilted diagonally downward, it covers a square area on the ground with a side length of 4.5m, enabling effective coverage in typical indoor environments such as the bedroom, study, and living room.

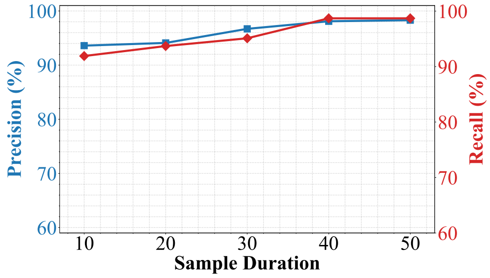


Fig. 17. Impact of the sample duration.

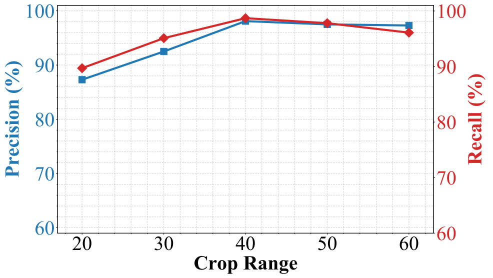


Fig. 18. Impact of the crop range.

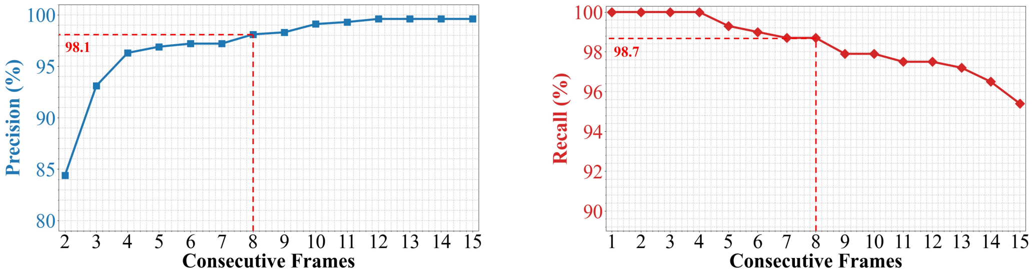


Fig. 19. Impact of the consecutive frames on detecting falls. The precision and recall of GR-Fall exceed 98% when the number of consecutive frames is set to 8.

Impact of Different Sample Duration. To investigate the impact of sample duration (N_T) on the performance of GR-Fall, we conduct experiments using sample sizes of 10, 20, 30, 40, and 50 frames. A smaller N_T may result in insufficient data, while a larger N_T can lead to delayed responses. Thus, selecting an optimal sample duration is crucial. The results, illustrated in Fig. 17, indicate that performance increases with the addition of frames, which also extends the response time. Especially, the improvement in both precision and recall when transitioning from 40 to 50 frames is minimal, at less than 1%. Therefore, we set the sample duration to 40 frames in GR-Fall to balance the performance and response time.

Impact of Different Crop Ranges. To investigate the impact of crop range (N_L) on the performance of GR-Fall, we conduct experiments with crop ranges of 20, 30, 40, 50, and 60. As shown in Fig. 18, the system performs best when N_L is set to 40. This is because, as N_L increases from 20, it captures progressively more information about the target, leading to improved performance. However, when N_L exceeds 40, the percentage of the target in the heatmap decreases, causing a decline in performance.

Impact of Different Consecutive Frames on Detecting Falls. As GR-Fall's state machine transitions to the *fall state* upon detecting a fall in N_F consecutive frames, selecting an appropriate value for N_F is critical. A small N_F can lead to a high false alarm rate, while a large N_F may result in missed alarms and delayed responses. As illustrated in Fig. 19, we calculate the precision and recall of GR-Fall for various values of N_F. When N_F increases from 8 to 9, precision improves only slightly, while recall decreases more significantly. Therefore, we choose $N_F = 8$, achieving a precision of 98.1% and maintaining a high recall of 98.7%.

Impact of Different Consecutive Frames on Recognizing Gaits. As GR-Fall's state machine transitions to the *alarm state* after detecting non-gait in N_G consecutive frames, choosing an appropriate value for N_G is essential. A smaller N_G could lead to more false gait recognitions, while a larger N_G might result in missed recognitions and

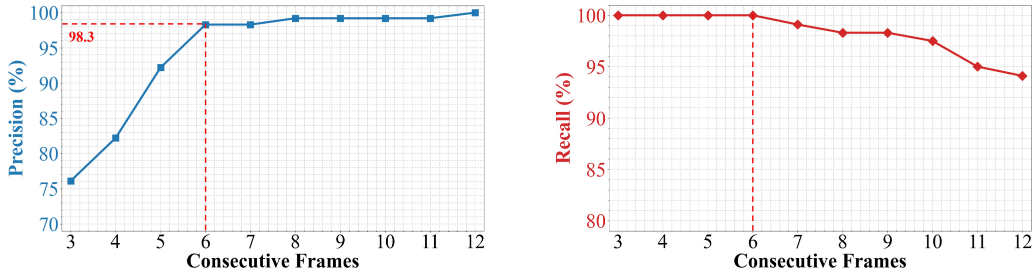
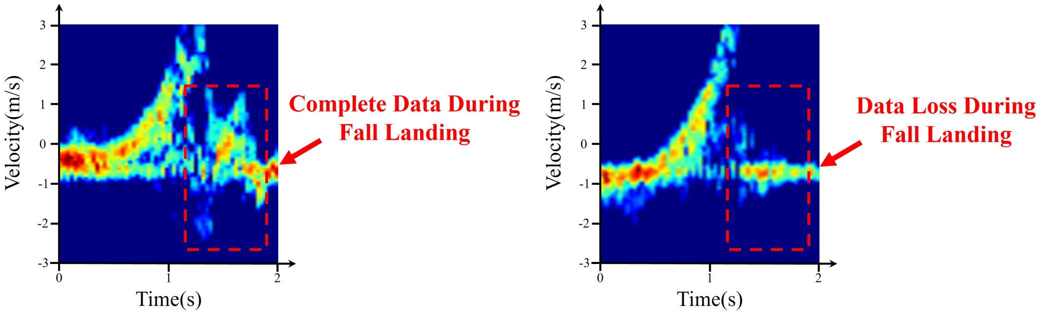


Fig. 20. Impact of the consecutive frames on recognizing gaits. The precision exceeds 98% and recall reaches 100% when the number of consecutive frames is set to 6.

Fig. 21. Occlusion of different objects.



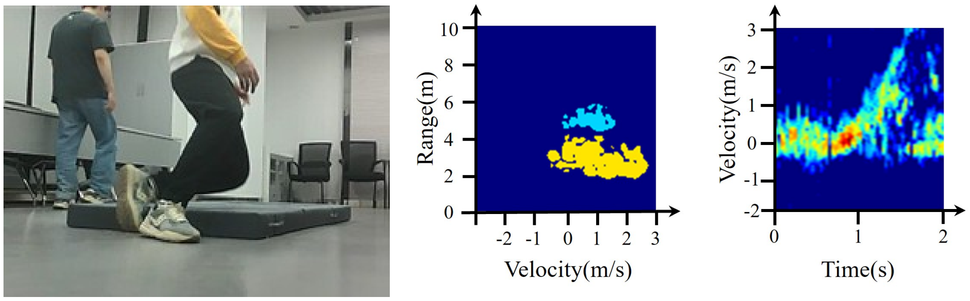
(a) Fall without occlusion.

(b) Fall with occlusion.

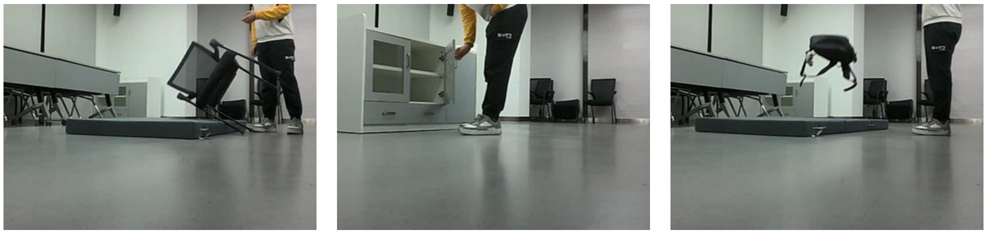
Fig. 22. Impact of object occlusion on heatmap data. When an object partially obstructs the radar's view, the radar is unable to capture motion data as the person approaches the ground during a fall, as most of the body is obscured during this period.

delayed responses. We evaluate gait recognition performance across various N_G values, as shown in Figure 20. With N_G set to 6, the system achieves a 100% recall in recognizing gaits. Therefore, we conclude that $N_G = 6$ is an optimal threshold, as it maintains 100% recall while keeping the false alarm rate below 2%.

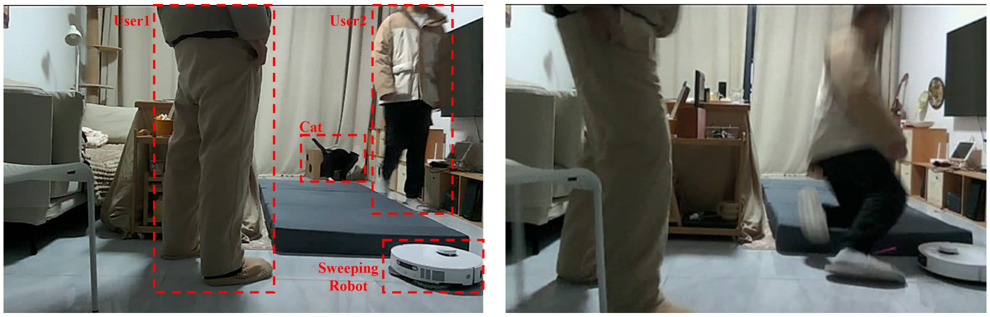
Impact of Object Occlusion. As shown in Fig. 21, we further assess GR-Fall's performance in environments with object occlusion, involving obstacles like tables, chairs, and similar items. Fig. 22 presents the target MD heatmap with and without occlusion at the same location. When an object partially obstructs the radar's view,



(a) Fall in the presence of another person. (b) RD Heatmap. (c) MD Heatmap.



(d) Chair tipping over. (e) Cabinet door opening. (f) Object dropping.



(g) Dynamic environment. (h) Fall in the dynamic environment.

Fig. 23. Impact of dynamic environments. When another moving person is present, as illustrated in (a), GR-Fall's target extraction module independently extracts data for each target, effectively distinguishing between different objects, as shown in (b). The data corresponding to the fall target remains unaffected, as depicted in (c). Additionally, (d), (e), and (f) illustrate dynamic changes within the environment, such as a chair tipping over, a cabinet door opening, and an object dropping. Moreover, as shown in (g) and (h), the dynamic environment includes two users, a sweeping robot, and a pet.

the radar is unable to capture motion data as the person approaches the ground during a fall, as most of the body is obscured during this period. Nonetheless, our system continues to detect falls accurately based on heatmap data, even when information is incomplete. Experimental results indicate that GR-Fall achieves an F1 score of 97.7%, a precision of 97.2%, and a recall of 98.1%, highlighting its robustness in occluded environments.

Impact of Dynamic Environments. Elderly falls may occur in dynamic environments. To evaluate GR-Fall's performance under such conditions, we evaluate it in scenarios involving environmental changes, such as a chair tipping over, a cabinet door opening, and an object dropping, as illustrated in Fig. 23(d), 23(e), and 23(f). In addition, we also evaluate its robustness in the presence of other moving objects within the room, such as a sweeping robot, a pet, or people, as illustrated in Fig. 23(g).

As shown in Fig. 23(a), when another moving person is present, GR-Fall's target extraction module (Sec. 3.3) independently extracts data for each target, ensuring accurate detection for all targets. Fig. 23(b) illustrates that the yellow region corresponds to the fall target, while the blue region corresponds to another target. Furthermore, Fig. 23(c) demonstrates that the fall target's MD heatmap remains unaffected. Experimental results show that GR-Fall achieves an F1 score of 97.2%, a precision of 97.8%, and a recall of 96.7%. Observations indicate a slight performance decrease due to the moving human body obstructing the line of sight between the radar and the faller. Nevertheless, the system consistently maintains good performance in dynamic environments.

Realistic Fall Simulations by Actors. To further enhance the realism of the collected data, we invite five experienced actors from the university's acting club to study and replicate real-world falls. Each actor specializes in simulating a specific type of fall—slipping, tripping, losing consciousness, losing balance, or falling from a chair—based on fall videos sourced from YouTube and TikTok [49–52, 65–67]. Using the pre-trained model, the newly collected actor data is used exclusively for testing. Experimental results demonstrate that GR-Fall consistently maintains good performance, with an F1 score of 97.7%, a precision of 98.1%, and a recall of 97.2%.

Ablation Study. GR-Fall consists of three key components: target extraction, fall detection, and gait recognition. After clutter filtering, clustering, and cropping, the target extraction module processes the IF signal to generate the target's RDSeq and MD heatmap. The fall detection module begins with data augmentation to expand the fall data and then uses a cross-attention mechanism to fuse features from RDSeq and MD heatmap for accurate fall detection. The gait recognition module extracts gait velocity-time variation as motion features from the MD heatmap and employs an SVM to determine if the user is walking normally. To assess the effectiveness of each component, we validate our GR-Fall by ablating specific components:

- **W/O Target Extraction:** It refers to skipping the target extraction process, using the original RD heatmap directly as input for subsequent modules.
- **W/O Data Augmentation:** It refers to training the model solely with the collected data, without applying any data augmentation techniques.
- **W/O MD Heatmap:** It refers to the removal of only the network branch responsible for extracting features from the MD heatmap.

Table 2. Ablation study of GR-Fall.

Method	F1 Score	Precision	Recall
W/O Target Extraction	93.3%	93.9%	92.7%
W/O Data Augmentation	94.9%	96.0%	93.8%
W/O MD Heatmap	94.8%	94.9%	94.8%
W/O RDSeq	94.5%	93.4%	95.8%
W/O Cross-Attention	95.2%	96.3%	94.3%
W/O Gait Recognition	82.5%	70.6%	99.4%
GR-Fall	**98.4%**	**98.1%**	**98.7%**

- **W/O RDSeq:** It refers to the removal of only the network branch responsible for extracting features from the RDSeq data.
- **W/O Cross-Attention:** It refers to replacing the network's attention mechanism for feature fusion with a simple concatenation of feature vectors.
- **W/O Gait Recognition:** It refers to the approach of immediately raising alarms upon detecting a fall, without assessing the severity of the fall.

The experimental results are presented in Tab. 2. Specifically, when target extraction is omitted, environment-independent target heatmaps cannot be obtained, resulting in more missed detections under changing conditions and a significant decrease in recall. Similarly, without data augmentation, the model lacks diverse fall data for training, reducing both robustness and overall performance. Furthermore, using a single heatmap or merely concatenating two heatmap features yields lower performance compared to attention-based feature fusion. Without gait recognition, distinguishing between severe and non-severe falls becomes impossible, causing non-severe falls to trigger false alarms and reducing precision. Consequently, our system integrates all these components to achieve optimal performance.

6 RELATED WORK

In this section, we review related work on fall detection, with a focus on both sensor-based and mmWave-based approaches. Detailed information is provided below.

6.1 Sensor-based Fall Detection

Fall detection systems can be categorized into wearable and non-wearable device-based [69]. Wearable systems include those using accelerometers [25, 33], RFID [7, 43], and similar devices [1, 53]. However, prolonged wear may cause discomfort to the user, and these systems may fail if the user forgets to wear or charge them. The limitations of wearable technologies can be addressed by using non-wearable alternatives. For example, camera [2, 14] is often used for fall detection by extracting features from multiple consecutive frame images. However, it is sensitive to lighting conditions and raises privacy concerns. In addition, WiFi-based fall detection methods [11, 39, 59, 61] offer non-invasive and privacy-preserving solutions. They detect falls by analyzing variations in channel state information (CSI). However, their effectiveness is significantly affected by environmental dynamics.

6.2 mmWave-based Fall Detection

The mmWave radar operates using high-frequency RF signals (GHz), making it non-invasive, privacy-preserving, and capable of providing high-range resolution and accurate motion detection [16, 29, 30, 46, 62, 63, 68, 71, 73]. Recent advancements in mmWave radar technology have significantly enhanced human sensing capabilities [6, 34, 57, 58, 60, 70, 74, 75]. Current mmWave radar fall detection techniques primarily utilize two types of data: point clouds and heatmaps.

The first approach to fall detection utilizes mmWave radar point clouds. Jin et al. [24] detect falls using a hybrid variational RNN autoencoder based on human body point clouds. Kittiyanpunya et al. [26] employ an LSTM model to detect falls using point cloud data and Doppler velocity. Zhang et al. [69] propose a two-stage fall detection system that identifies life-threatening falls by continuously monitoring the body's state after a fall. The second approach to fall detection utilizes mmWave radar heatmaps. Tian et al. [48] use a CNN and a state machine to detect falls based on range-angle heatmaps. Ding [10] employs KNN to extract features from range-Doppler heatmaps for fall detection. Sun et al.[47] leverage an LSTM to detect falls based on range-angle heatmaps. Li et al. [27] use a 3D CNN to detect falls based on range-angle heatmaps. However, most of these methods require MIMO mmWave radar, which leads to increased device costs and higher data transmission volume, making it less favorable for large-scale deployment of the equipment.

In contrast, we propose GR-Fall, an mmWave-based fall detection system with integrated gait recognition designed for indoor environments By utilizing a low-cost SISO mmWave radar, GR-Fall offers a more economical alternative to MIMO radar systems [18, 20]. To ensure robustness, accuracy, and practicality, our system incorporates several key components: target extraction, fall detection, and gait recognition. Importantly, GR-Fall triggers an alarm only when both a fall is detected and normal gait is absent, effectively reducing unnecessary strain on community resources.

7 DISCUSSION

Hardware Support. Our GR-Fall is built on Infineon's BGT60TR13C [19]. Specifically, we utilize the radar's single transmitting antenna (TX1) and single receiving antenna (RX1). The range and velocity resolutions of the mmWave radar are determined by its bandwidth and frame duration, respectively [41, 69]. Consequently, this radar configuration can also be applied to other SISO radars, such as Infineon's BGT60UTR11AIP [18], to achieve comparable range and velocity resolutions as well as overall performance.

Computational Overhead. Since the GPU may be less suitable for deployment due to its high cost, we utilize it solely during the training phase before deployment to accelerate model development. During actual deployment, inference is carried out on the CPU to reduce costs. Specifically, we evaluate the runtime of GR-Fall on the Intel Core i5-11500 CPU [22], where the processing time is 80 ms—equivalent to 12 fps—enabling the proposed system to achieve real-time fall detection.

Potential Application. GR-Fall's technologies have promising applications in smart homes, health monitoring, and related fields. For example, the attention-based heatmap feature extraction and fusion technique can be extended to other perception tasks, such as gesture recognition and pose estimation, by integrating data from multiple heatmaps. Additionally, the velocity-time variation-based gait recognition can be leveraged for applications like user authentication.

Limitations and Future Work. GR-Fall covers a square area with a side length of 4.5m, ensuring effective coverage in typical indoor environments such as bedrooms, studies, and living rooms. However, the sensing range of a single radar may be limited in larger spaces, leading to the potential failure to detect targets outside this range. Deploying multiple radars offers a practical solution to this limitation. In future work, we plan to explore methods for integrating multiple radars to achieve seamless and comprehensive coverage.

8 CONCLUSION

In this paper, we propose GR-Fall, a fall detection system with integrated gait recognition, designed for indoor environments. Leveraging a SISO mmWave radar, GR-Fall provides a cost-effective alternative to MIMO radar systems. GR-Fall introduces an innovative fall alarm mechanism that triggers alerts only if normal gait is not detected after a fall, effectively reducing strain on community resources. The system integrates several key components: heatmap-based target extraction, cross-attention-based heatmap fusion for fall detection, and velocity-time variation-based gait recognition, delivering robustness, high accuracy, and practicality. To evaluate GR-Fall, we recruit 33 volunteers and collect 5,799 instances across four different environments. Experimental results demonstrate that GR-Fall outperforms state-of-the-art methods, particularly in new environments and with new participants.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 62332016), the Key Research Program of Frontier Sciences, CAS (No. ZDBS-LY-JSC001), and the National Science and Technology Major Project (No. 2022ZD0114905).

References

[1] Stefano Abbate, Marco Avvenuti, Francesco Bonatesta, Guglielmo Cola, Paolo Corsini, and Alessio Vecchio. 2012. A smartphone-based fall detection system. *Pervasive and Mobile Computing* 8, 6 (2012), 883–899.

[2] Ekram Alam, Abu Sufian, Paramartha Dutta, and Marco Leo. 2022. Vision-based human fall detection systems using deep learning: A review. *Computers in biology and medicine* 146 (2022), 105626.

[3] Augusto Aubry, Antonio De Maio, Vincenzo Carotenuto, and Alfonso Farina. 2016. Radar phase noise modeling and effects-part I: MTI filters. *IEEE Trans. Aerospace Electron. Systems* 52, 2 (2016), 698–711.

[4] Xuyang Bai, Zeyu Hu, Xinge Zhu, Qingqiu Huang, Yilun Chen, and Hongbo Fu. 2022. Transfusion: Robust lidar-camera fusion for 3d object detection with transformers. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 1090–1099.

[5] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. 1992. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*. 144–152.

[6] Weiyan Chen, Hongliu Yang, Xiaoyang Bi, Rong Zheng, Fusang Zhang, Peng Bao, Zhaoxin Chang, Xujun Ma, and Daqing Zhang. 2023. Environment-aware Multi-person Tracking in Indoor Environments with MmWave Radars. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (2023), 1–29.

[7] Yung-Chin Chen and Yi-Wen Lin. 2010. Indoor RFID gait monitoring system for fall detection. In *International Symposium on Aware Computing*. IEEE, 207–212.

[8] Marcio L Lima de Oliveira and Marco JG Bekooij. 2020. Deep convolutional autoencoder applied for noise reduction in range-Doppler maps of FMCW radars. In *IEEE International Radar Conference*. IEEE, 630–635.

[9] Chuanwei Ding, Hong Hong, Yu Zou, Hui Chu, Xiaohua Zhu, Francesco Fioranelli, Julien Le Kernec, and Changzhi Li. 2019. Continuous human motion recognition with a dynamic range-Doppler trajectory method based on FMCW radar. *IEEE Transactions on Geoscience and Remote Sensing* 57, 9 (2019), 6821–6831.

[10] Chuanwei Ding, Yu Zou, Li Sun, Hong Hong, Xiaohua Zhu, and Changzhi Li. 2019. Fall detection with multi-domain features by a portable FMCW radar. In *IEEE MTT-S International Wireless Symposium (IWS)*. IEEE, 1–3.

[11] Jianyang Ding and Yong Wang. 2020. A WiFi-based smart home fall detection system using recurrent neural network. *IEEE Transactions on Consumer Electronics* 66, 4 (2020), 308–317.

[12] Baris Erol and Moeness G Amin. 2018. Radar data cube analysis for fall detection. In *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2446–2450.

[13] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*. 226–231.

[14] Homa Foroughi, Baharak Shakeri Aski, and Hamidreza Pourreza. 2008. Intelligent video surveillance for monitoring fall detection of elderly in home environments. In *International Conference on Computer and Information Technology*. IEEE, 219–224.

[15] Chenming He, Chengzhen Meng, Chunwang He, Xiaoran Fan, Beibei Wang, Yubo Yan, and Yanyong Zhang. 2024. See Through Vehicles: Fully Occluded Vehicle Detection with Millimeter Wave Radar. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 740–754.

[16] Chenming He, Rui Xia, Chengzhen Meng, Xiaoran Fan, Dequan Wang, Haojie Ren, Jianmin Ji, and Yanyong Zhang. 2025. Ghost Points Matter: Far-Range Vehicle Detection with a Single mmWave Radar in Tunnel. In *Proceedings of the 31th Annual International Conference on Mobile Computing and Networking*. 1–16.

[17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[18] Infineon. 2021. BGT60LTR11AIP. [Online]. https://www.infineon.com/cms/en/product/sensor/radar-sensors/radar-sensors-for-iot/60ghz-radar/bgt60ltr11aip/.

[19] Infineon. 2021. BGT60TR13C. [Online]. https://www.infineon.com/cms/en/product/evaluation-boards/demo-bgt60tr13c/.

[20] Infineon. 2023. BGT60ATR24C. [Online]. https://www.infineon.com/cms/en/product/sensor/radar-sensors/radar-sensors-for-automotive/60ghz-radar/bgt60atr24c/.

[21] Texas Instruments. 2020. The fundamentals of millimeter wave radar sensors.

[22] Intel. 2021. i5-11500. [Online]. https://www.intel.com/content/www/us/en/products/sku/212277/intel-core-i511500-processor-12m-cache-up-to-4-60-ghz/specifications.html.

[23] Intel. 2021. L515. [Online]. https://www.intelrealsense.com/lidar-camera-l515/.

[24] Feng Jin, Arindam Sengupta, and Siyang Cao. 2020. mmfall: Fall detection using 4-d mmwave radar and a hybrid variational rnn autoencoder. *IEEE Transactions on Automation Science and Engineering* 19, 2 (2020), 1245–1257.

[25] Maarit Kangas, Antti Konttila, Per Lindgren, Ilkka Winblad, and Timo Jämsä. 2008. Comparison of low-complexity fall detection algorithms for body attached accelerometers. *Gait & posture* 28, 2 (2008), 285–291.

[26] Chainarong Kittiyanpunya, Pongsathorn Chomdee, Akkarat Boonpoonga, and Danai Torrungrueng. 2023. Millimeter-Wave Radar-Based Elderly Fall Detection Fed by One-Dimensional Point Cloud and Doppler. *IEEE Access* 11 (2023), 76269–76283.

[27] Wenxuan Li, Dongheng Zhang, Yadong Li, Zhi Wu, Jinbo Chen, Dong Zhang, Yang Hu, Qibin Sun, and Yan Chen. 2022. Real-time fall detection using mmWave radar. In *IEEE International Conference on Acoustics, Speech and Signal Processing*. 16–20.

[28] Xinrong Li, Xiaodong Wang, Qing Yang, and Song Fu. 2021. Signal Processing for TDM MIMO FMCW Millimeter-Wave Radar Sensors. *IEEE Access* 9 (2021), 167959–167971.

[29] Yao Li, Yingjie Wang, Chengzhen Meng, Yifan Duan, Jianmin Ji, Yu Zhang, and Yanyong Zhang. 2024. FARFusion: A practical roadside radar-camera fusion system for far-range perception. *IEEE Robotics and Automation Letters* 9, 6 (2024), 5433–5440.

[30] Yadong Li, Dongheng Zhang, Jinbo Chen, Jinwei Wan, Dong Zhang, Yang Hu, Qibin Sun, and Yan Chen. 2022. Towards domain-independent and real-time gesture recognition using mmwave signal. *IEEE Transactions on Mobile Computing* 22, 12 (2022), 7355–7369.

[31] Jie Lian, Xu Yuan, Ming Li, and Nian-Feng Tzeng. 2021. Fall detection via inaudible acoustic sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–21.

[32] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*. 2980–2988.

[33] Ulrich Lindemann, A Hock, M Stuber, W Keck, and Clemens Becker. 2005. Evaluation of a fall detector based on accelerometers: A pilot study. *Medical and Biological engineering and computing* 43 (2005), 548–551.

[34] Alan Liu, Yu-Tai Lin, and Karthikeyan Sundaresan. 2024. View-agnostic Human Exercise Cataloging with Single MmWave Radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–23.

[35] Haipeng Liu, Kening Cui, Kaiyuan Hu, Yuheng Wang, Anfu Zhou, Liang Liu, and Huadong Ma. 2022. mTransSee: Enabling environment-independent mmWave sensing based gesture recognition via transfer learning. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–28.

[36] Chengzhen Meng, Yifan Duan, Chenming He, Dequan Wang, Xiaoran Fan, and Yanyong Zhang. 2024. mmplace: Robust place recognition with intermediate frequency signal of low-cost single-chip millimeter wave radar. *IEEE Robotics and Automation Letters* 9, 6 (2024), 4878–4885.

[37] Briana Moreland. 2020. Trends in nonfatal falls and fall-related injuries among adults aged >= 65 years—United States, 2012–2018. *MMWR. Morbidity and mortality weekly report* 69 (2020), 875–881.

[38] NVIDIA. 2021. RTX3060. [Online]. https://www.nvidia.com/en-me/geforce/graphics-cards/30-series/rtx-3060-3060ti/.

[39] Sameera Palipana, David Rojas, Piyush Agrawal, and Dirk Pesch. 2018. FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–25.

[40] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* 32 (2019), 8024–8035.

[41] Sandeep Rao. 2017. Introduction to mmWave sensing: FMCW radars. *Texas Instruments (TI) mmWave Training Series* (2017), 1–11.

[42] Ariyamehr Rezaei, Alessandro Mascheroni, Michael C Stevens, Reza Argha, Michela Papandrea, and Alessandro Puiatti. 2023. Unobtrusive human fall detection system using mmwave radar and data driven methods. *IEEE Sensors Journal* 23, 7 (2023), 7968–7976.

[43] Wenjie Ruan, Lina Yao, Quan Z Sheng, Nickolas JG Falkner, Xue Li, and Tao Gu. 2015. Tagfall: Towards unobstructive fine-grained fall detection based on uhf passive rfid tags. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. 140–149.

[44] Alexis R Santos-Lozada. 2023. Trends in deaths from falls among adults aged 65 years or older in the US, 1999-2020. *JAMA* 329, 18 (2023), 1605–1607.

[45] Abraham Savitzky and Marcel JE Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry* 36, 8 (1964), 1627–1639.

[46] Akash Deep Singh, Sandeep Singh Sandha, Luis Garcia, and Mani Srivastava. 2019. Radhar: Human activity recognition from point clouds generated through a millimeter-wave radar. In *Proceedings of the 3rd ACM Workshop on Millimeter-wave Networks and Sensing Systems*. 51–56.

[47] Yangfan Sun, Renlong Hang, Zhu Li, Mouqing Jin, and Kelvin Xu. 2019. Privacy-preserving fall detection with deep learning on mmWave radar signal. In *IEEE Visual Communications and Image Processing (VCIP)*. IEEE, 1–4.

[48] Yonglong Tian, Guang-He Lee, Hao He, Chen-Yu Hsu, and Dina Katabi. 2018. RF-based fall monitoring using convolutional neural networks. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–24.

[49] TikTok. 2022. Fall. [Online]. https://www.tiktok.com/@yaitsmelauren/video/7091296869707599150.

[50] TikTok. 2023. Fall. [Online]. https://www.tiktok.com/@tinytube3_/video/7225636375079570731.

[51] TikTok. 2024. Fall. [Online]. https://www.tiktok.com/@whalevision.official/video/7218897662152740142.

[52] TikTok. 2024. Fall. [Online]. https://www.tiktok.com/@plutocat777/video/7458693786542066976.

[53] Panagiotis Tsinganos and Athanassios Skodras. 2017. A smartphone-based fall detection system for the elderly. In *Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis*. IEEE, 53–58.

[54] Philip van Dorp and FCA Groen. 2008. Feature-based human motion parameter estimation with radar. *IET Radar, Sonar & Navigation* 2, 2 (2008), 135–145.

[55] A Vaswani. 2017. Attention is all you need. *Advances in Neural Information Processing Systems* (2017), 5998–6008.

[56] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272. https://doi.org/10.1038/s41592-019-0686-2

[57] Chao Wang, Feng Lin, Zhongjie Ba, Fan Zhang, Wenyao Xu, and Kui Ren. 2022. Wavesdropper: Through-wall word detection of human speech via commercial mmWave devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–26.

[58] Dequan Wang, Xinran Zhang, Kai Wang, Lingyu Wang, Xiaoran Fan, and Yanyong Zhang. 2024. RDGait: A mmWave Based Gait User Recognition System for Complex Indoor Environments Using Single-chip Radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–31.

[59] Hao Wang, Daqing Zhang, Yasha Wang, Junyi Ma, Yuxiang Wang, and Shengjie Li. 2016. RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices. *IEEE Transactions on Mobile Computing* 16, 2 (2016), 511–526.

[60] Shuai Wang, Dongjiang Cao, Ruofeng Liu, Wenchao Jiang, Tianshun Yao, and Chris Xiaoxuan Lu. 2023. Human parsing with joint learning for dynamic mmwave radar point cloud. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–22.

[61] Yuxi Wang, Kaishun Wu, and Lionel M Ni. 2016. Wifall: Device-free fall detection by wireless networks. *IEEE Transactions on Mobile Computing* 16, 2 (2016), 581–594.

[62] Guangyu Wu, Fuhui Zhou, Chengzhen Meng, and Xiang-Yang Li. 2023. Precise UAV MMW-vision positioning: A modal-oriented self-tuning fusion framework. *IEEE Journal on Selected Areas in Communications* 42, 1 (2023), 6–20.

[63] Chunyang Xie, Dongheng Zhang, Zhi Wu, Cong Yu, Yang Hu, and Yan Chen. 2023. RPM 2.0: RF-Based Pose Machines for Multi-Person 3D Pose Estimation. *IEEE Transactions on Circuits and Systems for Video Technology* 34, 1 (2023), 490–503.

[64] Yicheng Yao, Changyu Liu, Hao Zhang, Baiju Yan, Pu Jian, Peng Wang, Lidong Du, Xianxiang Chen, Baoshi Han, and Zhen Fang. 2022. Fall detection system using millimeter-wave radar based on neural network and information fusion. *IEEE Internet of Things Journal* 9, 21 (2022), 21038–21050.

[65] YouTube. 2023. Fall. [Online]. https://www.youtube.com/watch?v=p3tTZcLhJic.

[66] YouTube. 2024. Fall. [Online]. https://www.youtube.com/shorts/XUKRNHkPrEQ.

[67] YouTube. 2024. Fall. [Online]. https://www.youtube.com/shorts/lJKXwMqIQXM.

[68] Bin-Bin Zhang, Dongheng Zhang, Yadong Li, Yang Hu, and Yan Chen. 2023. Unsupervised domain adaptation for rf-based gesture recognition. *IEEE Internet of Things Journal* 10, 23 (2023), 21026–21038.

[69] Duo Zhang, Xusheng Zhang, Shengjie Li, Yaxiong Xie, Yang Li, Xuanzhi Wang, and Daqing Zhang. 2023. Lt-fall: The design and implementation of a life-threatening fall detection and alarming system. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–24.

[70] Duo Zhang, Xusheng Zhang, Yaxiong Xie, Fusang Zhang, Xuanzhi Wang, Yang Li, and Daqing Zhang. 2024. Local: An automatic location attribute calibration approach for large-scale deployment of mmwave-based sensing systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–27.

[71] Duo Zhang, Xusheng Zhang, Yaxiong Xie, Fusang Zhang, Hongliu Yang, and Daqing Zhang. 2024. From single-point to multi-point reflection modeling: Robust vital signs monitoring via mmwave sensing. *IEEE Transactions on Mobile Computing* (2024).

[72] Duo Zhang, Xusheng Zhang, Zhehui Yin, Yaxiong Xie, Hewen Wei, Zhaoxin Chang, Wenwei Li, and Daqing Zhang. 2025. mmRotation: Unlocking Versatility of a Single mmWave Radar via Azimuth Panning and Elevation Tilting. *IEEE Transactions on Mobile Computing* 24, 07 (2025), 6045–6061.

[73] Jia Zhang, Rui Xi, Yuan He, Yimiao Sun, Xiuzhen Guo, Weiguo Wang, Xin Na, Yunhao Liu, Zhenguo Shi, and Tao Gu. 2023. A survey of mmWave-based human sensing: Technology, platforms and applications. *IEEE Communications Surveys & Tutorials* 25, 4 (2023), 2052–2087.

[74] Xusheng Zhang, Duo Zhang, Yaxiong Xie, Dan Wu, Yang Li, and Daqing Zhang. 2024. Waffle: A Waterproof mmWave-based Human Sensing System inside Bathrooms with Running Water. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–29.

[75] Langcheng Zhao, Rui Lyu, Hang Lei, Qi Lin, Anfu Zhou, Huadong Ma, Jingjia Wang, Xiangbin Meng, Chunli Shao, Yida Tang, et al. 2024. AirECG: Contactless Electrocardiogram for Cardiac Disease Monitoring via mmWave Sensing and Cross-domain Diffusion Model. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–27.

[76] Langcheng Zhao, Rui Lyu, Qi Lin, Anfu Zhou, Huanhuan Zhang, Huadong Ma, Jingjia Wang, Chunli Shao, and Yida Tang. 2024. mmArrhythmia: Contactless Arrhythmia Detection via mmWave Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–25.