

## Article

# Fall Detection System Based on Point Cloud Enhancement Model for 24 GHz FMCW Radar

Tingxuan Liang <sup>1</sup>, Ruizhi Liu <sup>1</sup>, Lei Yang <sup>2</sup>, Yue Lin <sup>2</sup>, C.-J. Richard Shi <sup>3</sup> and Hongtao Xu <sup>1,\*</sup>

<sup>1</sup> State Key Laboratory of Integrated Chips and Systems, Fudan University, Shanghai 201203, China; 18112020032@fudan.edu.cn (T.L.)

<sup>2</sup> ICLegend Micro, Shanghai 201203, China

<sup>3</sup> Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195, USA; cjshi.fudan@gmail.com

\* Correspondence: hongtao@fudan.edu.cn

**Abstract:** Automatic fall detection plays a significant role in monitoring the health of senior citizens. In particular, millimeter-wave radar sensors are relevant for human pose recognition in an indoor environment due to their advantages of privacy protection, low hardware cost, and wide range of working conditions. However, low-quality point clouds from 4D radar diminish the reliability of fall detection. To improve the detection accuracy, conventional methods utilize more costly hardware. In this study, we propose a model that can provide high-quality three-dimensional point cloud images of the human body at a low cost. To improve the accuracy and effectiveness of fall detection, a system that extracts distribution features through small radar antenna arrays is developed. The proposed system achieved 99.1% and 98.9% accuracy on test datasets pertaining to new subjects and new environments, respectively.

**Keywords:** radar; fall detection; machine learning



**Citation:** Liang, T.; Liu, R.; Yang, L.; Lin, Y.; Shi, C.-J.R.; Xu, H. Fall Detection System Based on Point Cloud Enhancement Model for 24 GHz FMCW Radar. *Sensors* **2024**, *24*, 648. <https://doi.org/10.3390/s24020648>

Academic Editor: Bijan Najafi

Received: 11 December 2023

Revised: 5 January 2024

Accepted: 17 January 2024

Published: 19 January 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

According to a report from the World Health Organization, approximately 28–35% of older adults fall each year, leading to serious injury or death [1]. Therefore, intelligently detecting falls in indoor conditions can reduce the risk of the elderly injuring themselves.

Various technologies have been adopted to detect falls. Existing fall detection methods require wearable sensors [2]. Accelerometers have been widely used in wearable methods, and a velocity threshold can be set to detect fall events [3,4]; however, these may be forgotten because of their inconvenience. Vision-based methods eliminate the need to wear something, but they are costly, sensitive to the lighting conditions, and invade privacy [5,6]. Recently, radar sensors have become more popular in fall detection system due to the advantages compared with other sensing technologies: (a) convenience over wearable technologies [7]; (b) high sensitivity to motion compared to depth sensors in complex living environments and weak lighting conditions; (c) privacy compliance over vision sensors [8]; and (d) low hardware cost compared with other sensors [9]. Typical radars for human fall detection are continuous wave (CW) radars and frequency-modulated continuous wave (FMCW) radars. A CW radar signal was converted into the time–frequency domain and extracted artificial features for detecting a falling person [10,11]. Doppler-time signatures recorded in the CW signals were used for the training of a machine learning algorithm [12]. However, CW radar can only provide velocity information. Due to the lack of information richness, actions with similar postures, such as sitting and squatting, may lead to inaccuracies. A better choice is to use an FMCW radar, which can simultaneously provide the range, Doppler, and angle information of the targets and also high sensitivity to motion [13].

Traditionally, researchers have explored several methods based on FMCW radars, which range from 57–85 GHz [14]. The Doppler information could describe the velocity

attribute of a motion; thus, the range-Doppler map has been widely used in FMCW radar-based fall detection methods proposed in the literature [15–18]. The Doppler-time map, including time information, was directly used as a feature to detect fall events [18]. Many studies on FMCW radar-based fall detection rely on the time–frequency characteristics of the FMCW radar return signal, including the frequency magnitude, frequency ratio, and the duration of the motion [19]. However, similar to CW radar-based methods, these methods cannot provide spatial information, and similar motions may lead to inaccuracies. Micro-Doppler and spatial information have been used to achieve high accuracy, proving that deep learning methods are superior to traditional artificial feature extraction methods [20]. An FMCW radio has been used to obtain the 3D position information of the human body and heatmaps in both horizontal and vertical directions [17]. However, there is still a problem in combining 3D location information: achieving high angular resolution using radars with large antenna arrays.

To utilize 3D spatial information, recent innovations in human activity detection have explored point clouds from radar [21–23], in which each point contains a 3D position in space. However, in contrast to LiDAR and camera sensors, there are two main challenges in these studies: (1) the point clouds generated by mmWave radar are usually sparse and of low resolution, and (2) the point clouds include many ghost points caused by the multipath effect. As a result, the classification accuracy and reliability may be negatively affected. To address these challenges, several methods have been designed for use in conjunction with high-resolution radars. 12Tx-16Rx antenna arrays have been used to generate high-quality point clouds [22]. Hawkeye generated 2D depth images using radar intensity maps obtained from SAR scans [23]. However, although large antenna arrays and SAR technologies can improve the resolution, they are still very slow and may not be practical in many applications that require a short response time and low-cost hardware. In addition, sparsity-related methods and deep learning-based methods have been used for point clouds' quality enhancement [24]. Some sparsity-related methods, such as K-mean [25] and density-based spatial clustering of applications with noise (DBSCAN) algorithm [26], have been used to cluster firstly to remove outliers in the point clouds. However, these technologies could not adequately remove a sufficient number of outlier points. In recent studies, a few deep learning-based methods have been developed based on PointNet [27]. After learning a mapping from the noisy input, they can automatically generate a set of clean points. Inspired by the PointNet, PCN combined a permutation invariant and non-convolutional feature extractor to complete a point cloud from a partial input, and then used a refinement block to denoise the prediction to produce the final point cloud [28]. GPDNet was applicable to denoise point clouds based on graph-convolutional layers [29]. However, most of these methods used LiDAR or other sensor applications and extracted pointwise features. Hence, they may not be efficient for radar-based point clouds because of their very low resolution.

In this study, we propose an FMCW radar-based fall detection method that investigates 3D point clouds while operating at 24 GHz. These systems have not been studied well owing to hardware limitations. First, we obtained raw point clouds from the radar. We then designed a new model to transform the raw points into high-quality point clouds that are closer to the ground truth. Next, we estimated the distribution parameters in the point clouds for classification.

The main contributions of this paper are as follows:

- (1) We propose an efficient fall detection system that uses a small, low-cost radar. As shown in Figure 1, the novel framework is primarily composed of three parts: point cloud enhancement (PCE) for point cloud quality improvement, a feature extractor for human pose parameter extraction, and a classifier for classifying normal events and fall events;
- (2) A PCE model is introduced to transform low-quality point clouds into high-quality point clouds and to reconstruct the shape of the point clouds using the shape of human

- body. A novel 3D point-box hybrid regression loss function based on pointwise and 3D bounding box features is proposed as a substitute for the traditional loss function;
- (3) Our system works on sparse and noisy raw radar data without using expensive hardware or synthetic aperture radar (SAR) scans.

The remainder of this article is organized as follows. Section 2 provides an overview of the radar system and signal processing flow. In Section 3, the details of the proposed method are presented. The results are discussed in Section 4. Finally, Section 5 concludes this study.

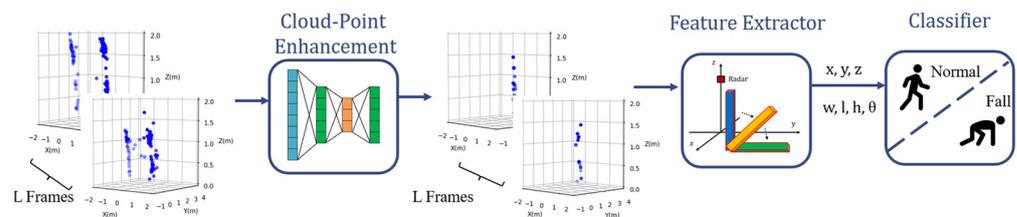


Figure 1. Proposed fall detection system based on PCE model.

## 2. Radar System Design

An FMCW radar is often used to estimate the range and velocity of a target using its radio frequency signals. Each transmitted signal  $s(t)$  is a chirp signal, the analytical expression [30] of which is

$$s(t) = \exp\left(j2\pi\left(f_0t + \frac{1}{2}Kt^2\right)\right) \quad (1)$$

where  $f_0$  denotes the carrier frequency, and  $K$  is the chirp constant.

Referring to the Boulic model [31], the echo from the  $i$ th human body part is a time-delayed version of the transmitted signal, and the received baseband signal can be expressed as

$$S_i(t) = \exp\left(j2\pi f_0\left(t - \frac{2R_i}{c}\right)\right) \quad (2)$$

where  $R_i$  is the distance between the  $i$ th ellipsoidal center and the radar, and  $c$  is the speed of light. The echo from the entire human body can be expressed as

$$S_{all}(t) = \sum_{i=1}^M \eta_i S_i(t) \quad (3)$$

where  $\eta_i$  is the attenuation coefficient, which is governed by the radar cross section for different body regions, and  $M$  is the number of scattering points of human body parts.

As shown in Figure 2, the raw data from each channel generated a 3D data cube. FMCW radar signal processing started with the sampled echoes that were transferred to the range-Doppler matrix [32]. First, the range of the fast Fourier transform (FFT) assisted in estimating the range of the targets, and the second FFT determined the Doppler frequency. The moving target indicator distinguished the targets from the background. For more reliable detection, a constant false alarm rate was used to detect the targets against the background noise. A direction-of-arrival (DOA) estimation algorithm was used to estimate the angle of the target.

For angle estimation, both 3D FFT and the multiple signal classification (MUSIC) algorithm [33] are popular methods. Compared with the MUSIC algorithm, 3D FFT has computing efficiency, but the resolution of 3D FFT is not sufficient for detecting closely spaced points. To obtain a better image of the human body and low computing cost, based on TDM-MIMO antenna arrays in Figure 3, after we obtained eight virtual receiver samples, a 3D FFT was used to obtain the azimuth angle  $\theta_{az}$ , and the MUSIC algorithm was used

to estimate the elevation angle  $\theta_{el}$ . The output spherical coordinates were converted into Cartesian coordinates using transfer matrix T:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_t & \sin\theta_t \\ 0 & -\sin\theta_t & \cos\theta_t \end{bmatrix} \begin{bmatrix} R\cos\theta_{el}\sin\theta_{az} \\ R\cos\theta_{el}\cos\theta_{az} \\ R\sin\theta_{el} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} \quad (4)$$

where  $R$  is the distance between the target and the radar. After the transformation, we obtained the position values  $[x, y, z]$  of each point in each frame in Cartesian coordinates.

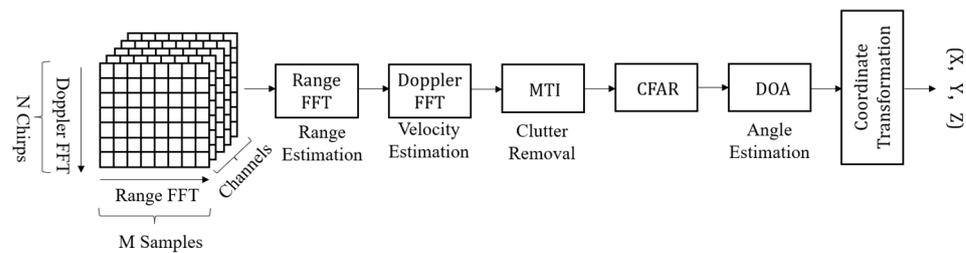


Figure 2. Signal processing flow.

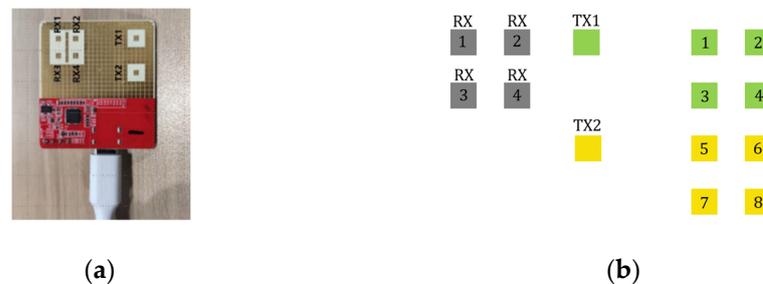


Figure 3. TDM-MIMO antenna arrays. (a) Radar diagram and real antenna array; (b) virtual antenna array.

### 3. Proposed Method

The point clouds detected from the range-Doppler map are sensitive to the environment. Figure 4 shows different point distributions generated in different indoor environments, in which obstacles cause severe multipath interference and affect the performance of fall detection systems in new environments.

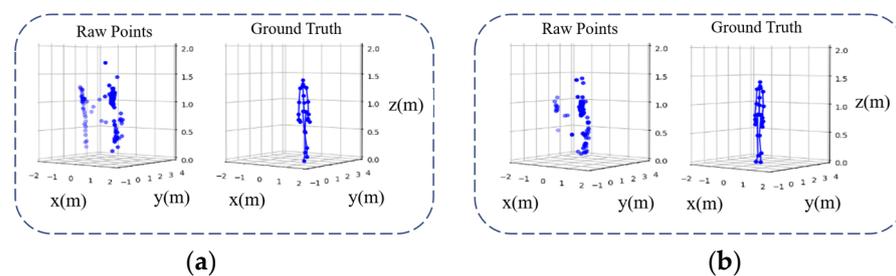
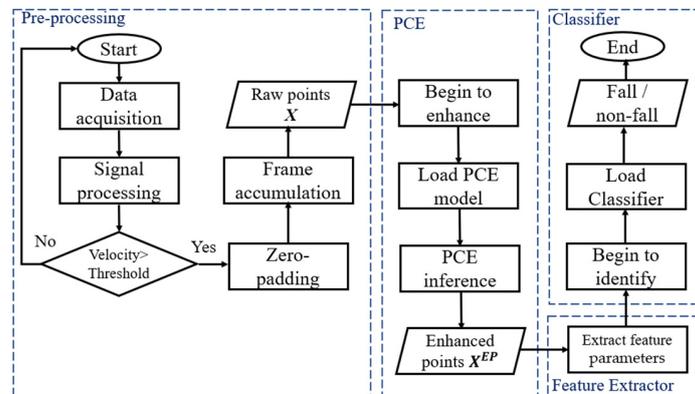


Figure 4. Raw point clouds in different environments. (a) Raw point clouds from radar collected in a relatively empty room; the number of points is 85. (b) Raw point clouds collected in a real bedroom; the number of points is 125.

Therefore, we propose a fall detection method based on a PCE model. A flowchart of the proposed method is presented in Figure 5. The flow of proposed method consists of four steps: (1) to satisfy the input of the PCE model, the number of point clouds of a motion pattern could be extended to fixed number; (2) after raw point clouds from the radar in pre-processing, PCE model removes noise points and generates high-quality point clouds; (3) these point clouds are then fed into the feature extractor for human pose parameter

extraction; (4) a lightweight classifier is used for classifying normal and fall events. In this section, the functionality of each component is described. More detailed descriptions are as follows.



**Figure 5.** The flowchart of proposed system.

### 3.1. Pre-Processing

The motion window accumulates  $L$  frames of the point cloud from the radar, where  $L$  is determined by the length of an action such as walking, squatting, and falling. For each action pattern, we extracted the 3D vectors of  $[x, y, z]$  for each point in every frame. In each frame, the number of radar point clouds was random owing to the nature of the FMCW radar measurement. To satisfy the input of the PCE model, zero padding was used for data oversampling. Thus, the number of point clouds can be extended to a fixed number. We obtained the motion pattern

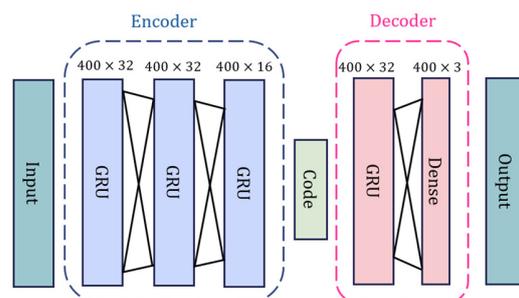
$$X = \left\{ \left\{ x_m^i \right\}_{m=1}^M \right\}_{l=1}^L \quad (5)$$

where  $l$  is the frame index of the motion,  $M$  is the number of points from radar in each frame, and  $p_m^l$  is the  $m^{\text{th}}$  point in the  $l^{\text{th}}$  frame, which is a vector of  $[x_m^l, y_m^l, z_m^l]$ .

Simultaneously, the Azure Kinect was used as a reference sensor because of its high precision in extracting human skeletal keypoints. Ignoring some inessential points, the desired 27 skeletal keypoints returned from Azure Kinect were also accompanied by a radar time stamp in system time and served to label the ground truth in the training process.

### 3.2. PCE Model

Although point clouds have already been obtained from range-Doppler maps, the resulting points are still low-resolution, very noisy, and affect the classification accuracy and reliability. In addition, it is unnecessary for a fall detection system to increase the computational cost of reconstructing each point. Thus, to improve the quality of these point clouds and increase classification reliability, we propose a PCE model that aims to minimize the difference between the shapes of the radar point clouds and the ground truth. An overview of the proposed PCE architecture is shown in Figure 6.



**Figure 6.** Proposed PCE model architecture.

### 3.2.1. Encoder

The encoder is responsible for extracting global features from incomplete point clouds. As shown in Figure 6, the encoder comprises three consecutive gate recurrent units (GRU) [34] followed by a dense layer. Using GRU mechanisms in a machine learning network is a relatively new time-series architecture compared with recurrent neural networks (RNNs) and long short-term memories (LSTMs). GRUs can retain long-term dependencies and are also faster for training compared with traditional LSTMs because there are fewer gates to update. As shown in Figure 6, the units of the three GRU layers are 32, 32, and 16. The input data  $X^{RP} = \{x_1^{RP}, x_2^{RP}, \dots, x_N^{RP}\}$  are first reconstructed by the hidden parts, and the output of the encoder is a series of vectors  $\{h_1, h_2, \dots, h_b\}$ . Thus, the encoder process is given by

$$h_b = \text{GRU}(W_{e,b}X^{RP} + b_{e,b}) \quad (6)$$

where  $W_{e,b}$  and  $b_{e,b}$  are the encoder weight matrix and bias vector for the  $b$ th node ( $b = 1, 2, \dots, B$ ).

### 3.2.2. Decoder

The decoder consists of three consecutive GRUs followed by a dense layer. The units of the three GRU layers are 32. The output from the GRU is

$$u_i = \text{GRU}(W_{d,b}w_i + b_{d,b}) \quad (7)$$

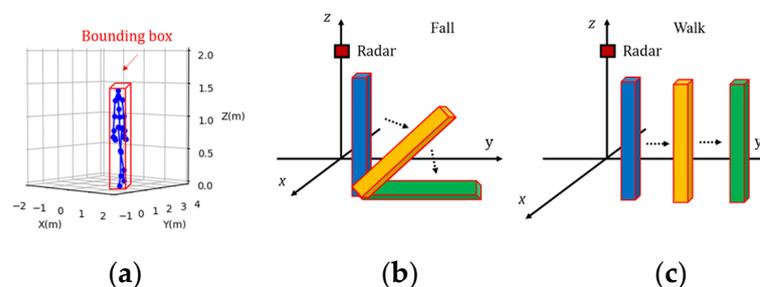
where  $W_{d,b}$  and  $b_{d,b}$  are the decoder weight matrix and bias vector for the  $b$ th node ( $b = 1, 2, \dots, B$ ), and  $w_i$  is the input layer of the decoder. The recovered point clouds  $\hat{X}$  through the dense layer is obtained via

$$\hat{X} = W_d u_i + b_d \quad (8)$$

where  $W_d$  and  $b_d$  are the weights and biases of the output layer, respectively. After the sampling layer, an enhanced point cloud  $X^{EP}$  is obtained.

### 3.2.3. Loss Function

A simple method for extracting pointwise features in 3D space is supervised by the mean squared error (MSE). However, directly supervising the pointwise features of point clouds may not utilize 3D receptive field information. The observations in Figure 7 indicate that the change in the 3D receptive field bounding box of a falling human body is different from that of a walking human body. Specifically, changes in the bounding box of a human body are related to its pose. Therefore, we not only consider pointwise features but also detect fall events by characterizing the uniqueness of bounding box changes for such poses.



**Figure 7.** (a) Pose based on the 3D bounding box of a standing human body; (b) change in the 3D bounding box of a falling human body; (c) change in the 3D bounding box of a walking human body.

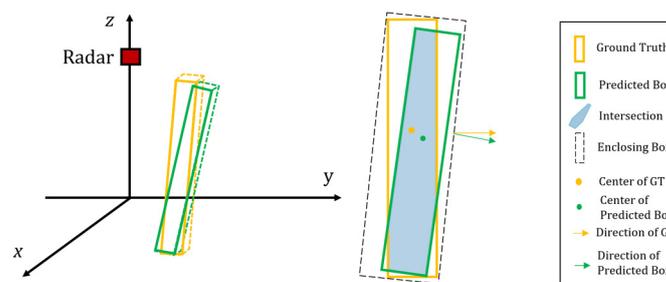
To use the bounding box of the radar point clouds for fall detection, the position and shape of the predicted box should be closely related to the corresponding ground truth. In this manner, we propose a 3D point-box hybrid regression loss to reduce the error

between the predicted radar bounding box and ground truth. Previous studies [35,36] found intersection over union (IoU) loss and generalized IoU functions for 2D boxes using only the width and height of the boxes without considering the direction of the boxes. In addition, it is difficult to provide a specific formula describing the intersection between two 3D bounding boxes because a variety of cases must be considered. Some previous studies projected a 3D bounding box onto two 2D bounding boxes, but this did not increase the accuracy because of lack of direction [37].

In this study, because most of the measured human motion states were symmetrical along the x-axis (as shown in Figure 8), a 3D bounding box was obtained based on the 2D rotated IoU by multiplying the length along the x-axis. As shown in Figure 9, the intersection of the predicted box and ground truth included a variety of polygons, which was the sum of the areas of all the triangles. Therefore, the 3D IoU can be expressed as

$$IoU_{3D} = \frac{Area_{OL} \cdot x_{OL}}{Area_r \cdot x_r + Area_{GT} \cdot x_{GT} - Area_{OL} \cdot x_{OL}} \quad (9)$$

where  $Area_{OL}$  is the area of overlap between the two boxes;  $Area_r$  and  $Area_{GT}$  are the areas of the predicted box from radar and ground truth, respectively;  $x_{OL}$  is the overlap on the x-axis; and  $x_r$  and  $x_{GT}$  are the lengths of the predicted box and ground truth along the x-axis, respectively.



**Figure 8.** Example of a polygon formed by the overlap of the predicted box and ground truth. The 3D IoU can be approximately achieved by multiplying the length along the x-axis with the directed 2D IoU.

Additionally, an accurate center  $[x_{center}, y_{center}, z_{center}]$  is a prerequisite for predicting an accurate bounding box. An accurate width along the y-axis and an accurate height along the z-axis also contribute to maximizing the fall detection accuracy. Based on the above observations, the box-based loss function is expressed as

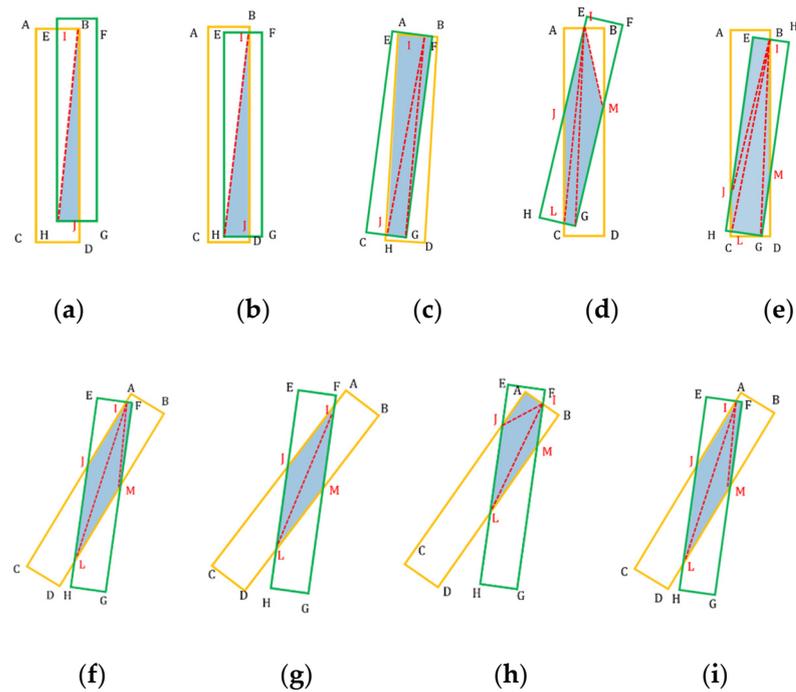
$$\begin{aligned} Loss_{box} &= Loss_{IoU_{3D}} + Loss_{center_{dis}} + Loss_{line} \\ &= 1 - IoU_{3D} + \frac{d^2(b^r, b^{GT})}{(l^e)^2 + (w^e)^2 + (h^e)^2} \\ &\quad + \frac{d^2(w^r, w^{GT})}{(w^e)^2} + \frac{d^2(h^r, h^{GT})}{(h^e)^2} \end{aligned} \quad (10)$$

where  $b^r$  and  $b^{GT}$  are the center positions of the predicted box and ground truth box, respectively;  $l^e$ ,  $w^e$ , and  $h^e$  are the length, width, and height of the enclosing box, respectively;  $l^r$ ,  $w^r$ , and  $h^r$  are the length, width, and height of the predicted box, respectively; and  $l^{GT}$ ,  $w^{GT}$ , and  $h^{GT}$  are the length, width, and height of the ground truth box, respectively.

In summary, the 3D point-box hybrid regression loss consisted of a 3D bounding box IoU loss and a pointwise loss, which can be described as

$$Loss_{HyLoss} = Loss_{point} + Loss_{box} \quad (11)$$

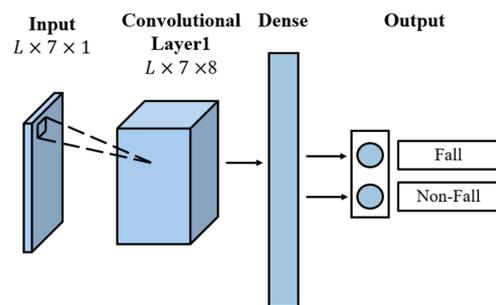
where  $Loss_{point}$  is the position loss of the points for optimizing the IoU of the human body and the ground reflection points.



**Figure 9.** (a–i) are examples of the directed 2D IoU with different numbers of intersection points. The insertion area is obtained by summing the triangular areas.

### 3.3. Feature Extraction

We used a lightweight CNN for classification. The architecture of the lightweight CNN is shown in Figure 10. The CNN included an input layer sequence, one convolution layer, one dense layer, and a softmax layer. The feature parameters from one frame included  $[x_{center}, y_{center}, z_{center}, w, l, h, \theta]$ , and  $L$  frames of  $7 \times 1$  feature parameters were first subjected to the convolution layer. The convolution layer captured the movement features, consisting of eight hidden neurons with a kernel size of eight. The convolution layer employed rectified linear units as activation functions for the hidden layers. The  $L \times 7 \times 8$  output from the convolution layer was fed into the dense layer. The softmax function was used in the final dense layer for classification.



**Figure 10.** Architecture of the lightweight CNN.

## 4. Implementation and Evaluation

### 4.1. Setup and Data Collection

We used the 24 GHz mmWave FMCW radar from ICLegend Micro for data collection. The radar sensor had two transmitting antennas and four receiving antenna channels. The radar parameter configurations are listed in Table 1.

To evaluate the proposed system, we set up the experiments and collected data in two different indoor environments. The experimental setup is shown in Figure 11. Room A, shown in Figure 11b, was a relatively empty office, and room B, shown in Figure 11c, was a

bedroom where obstacles caused severe multipath and motion interference. The elevation field of view (FOV) of the Kinect sensor is 135 degrees, and the elevation FOV of the radar is 90 degrees. To capture more of a scene, the radar and Kinect sensors were both mounted at heights of 1.3 m. We collected 17,600 frames of data from 13 subjects performing fall and non-fall actions. The participants were aged between 20 and 30 years, and their heights were between 165 and 177 cm.

**Table 1.** Unit parameter configuration of the radar.

Parameter	Value
Carrier frequency	24 GHz
Bandwidth	3.17 GHz
Duration of a chirp	460 $\mu$ s
Number of chirps per frame	254
Chirps per CPI	64
Duration of a frame	117 ms

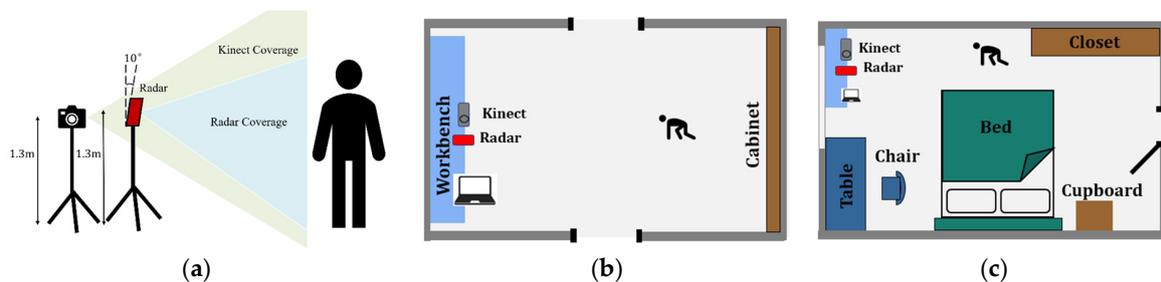
Furthermore, we divided these frames into three datasets:

**Dataset0.** This dataset included 10,860 frames from five participants who performed the experiments in room A. After data argumentation, including shifting, rotating, and adding noise points, the dataset included 13,560 frames.

**Dataset1.** This dataset included 3040 frames from four participants who performed the experiments in room A.

**Dataset2.** This dataset included 3700 frames from four participants who performed the experiments in room B.

For each dataset, the motion included falling backward and falling forward, sitting on a chair, jumping, walking, or squatting. The ratio of the number of fall samples to non-fall samples was 1:2.



**Figure 11.** (a) Collected movement in dataset; (b) evaluation scenario: room A (office); (c) evaluation scenario: room B (bedroom).

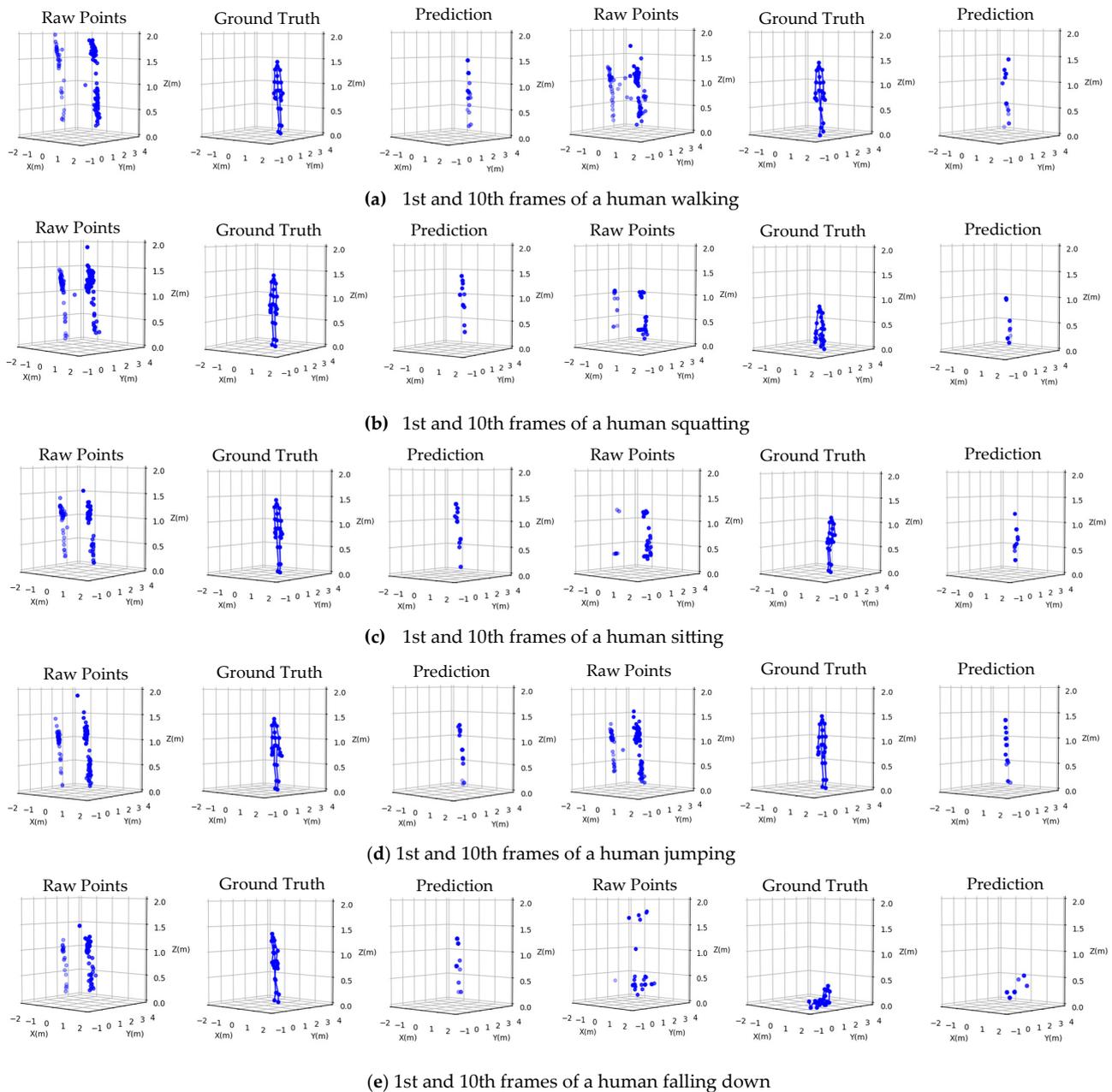
## 4.2. Evaluation

In this section, we present the experimental results of the proposed model in a fall detection system. All algorithms were realized based on Python code. The computation platform used was a laptop with a 2.60 GHz Intel(R) Core (TM) i7-10750H CPU.

### 4.2.1. Point Cloud Quality

Before training the proposed PCE model, the 3D radar data and ground truth skeletal points from the Azure Kinect were extracted, as described in Section IV. Then,  $L$  frames of data from the radar were combined sequentially to obtain a 3D tensor ( $L \times M \times 3$ ). We then evaluated the PCE model for  $L = \{1, 2, \dots, 10\}$  by generating 10 distinct datasets,  $Dataset0_i$ ,  $Dataset1_i$  and  $Dataset2_i$ , where  $i \in \{1, 2, \dots, 10\}$  is the frame index. The proposed PCE model  $M_i$  was trained on  $Dataset0_i$ , and the trained PCE model  $M_i$  was subjected to  $Dataset1_i$  and  $Dataset2_i$ , which did not participate in the training. Figure 12 shows predicted points. We selected two frames of each motion for comparison against the ground

truth. In addition, because of the nature of the FMCW radar measurement, an exact and complete match of  $[x, y, z]$  between the predicted and ground truth keypoints was simply unrealistic; therefore, finding a method for evaluating the predicted results was crucial.



**Figure 12.** Comparison between raw point clouds, the ground truth, and the PCE model's predicted points in two frames (1st and 10th) for five different actions (walking, squatting, sitting, jumping, and falling).

To evaluate the quality of the point clouds, we calculated the average 3D IoU for every L frame scenario across the  $f$  test samples in the dataset according to the formula  $3DIoU_i = \frac{1}{N} \sum_{j=1}^N 3DIoU_{ij}, \forall i \in \{1, 2, \dots, 10\}$ . The average 3D IoU results for all L frames from Dataset1 and Dataset2 are outlined in Tables 2 and 3, respectively. In the 10th frame, the average 3D IoU is lower than that in the other frames. The reason for this result was that sometimes the number of points was too small because the motion may have already been completed in advance. We also compared the raw point clouds from the radar and the DBSCAN method, which is the most popular method for removing outliers and improving

the quality of point clouds. The proposed method outperformed the other methods. In addition, the PCE with the proposed HyLoss function outperformed the traditional MSE loss function.

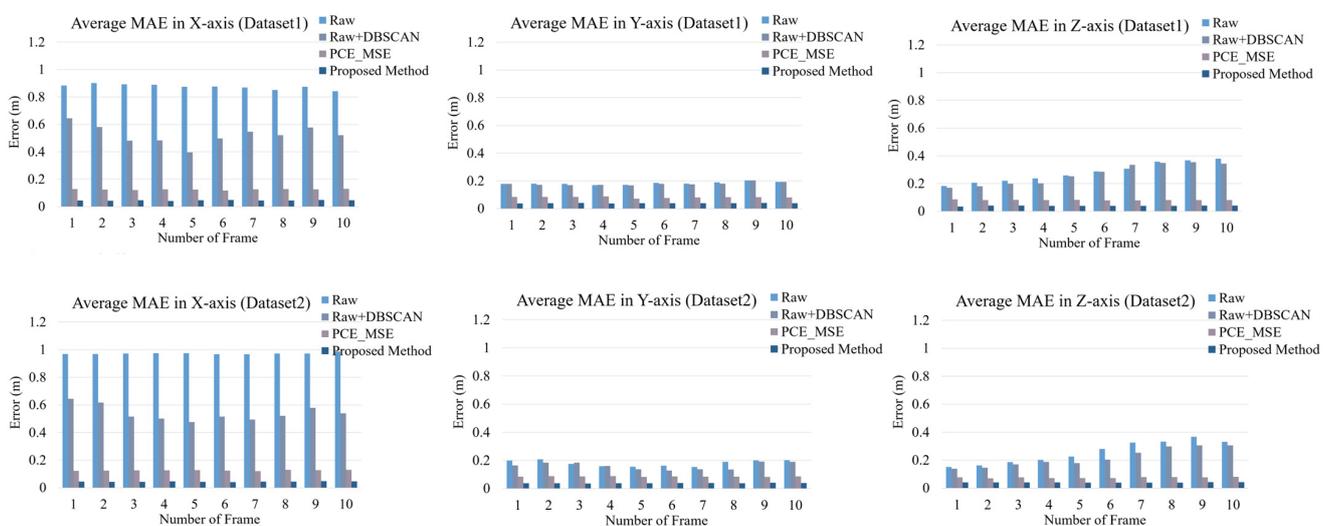
**Table 2.** Average 3D IoU of Dataset1 in each frame.

	1	2	3	4	5	6	7	8	9	10	Mean
Raw (baseline)	0.276	0.265	0.262	0.249	0.269	0.233	0.215	0.194	0.202	0.183	<b>0.234</b>
Raw + DBSCAN [26]	0.300	0.287	0.266	0.257	0.249	0.225	0.212	0.207	0.191	0.189	<b>0.259</b>
PCE_MSE	0.609	0.633	0.623	0.636	0.618	0.597	0.596	0.579	0.574	0.525	<b>0.599</b>
Proposed PCE_HyLoss	0.635	0.669	0.687	0.681	0.664	0.652	0.653	0.627	0.628	0.619	<b>0.651</b>

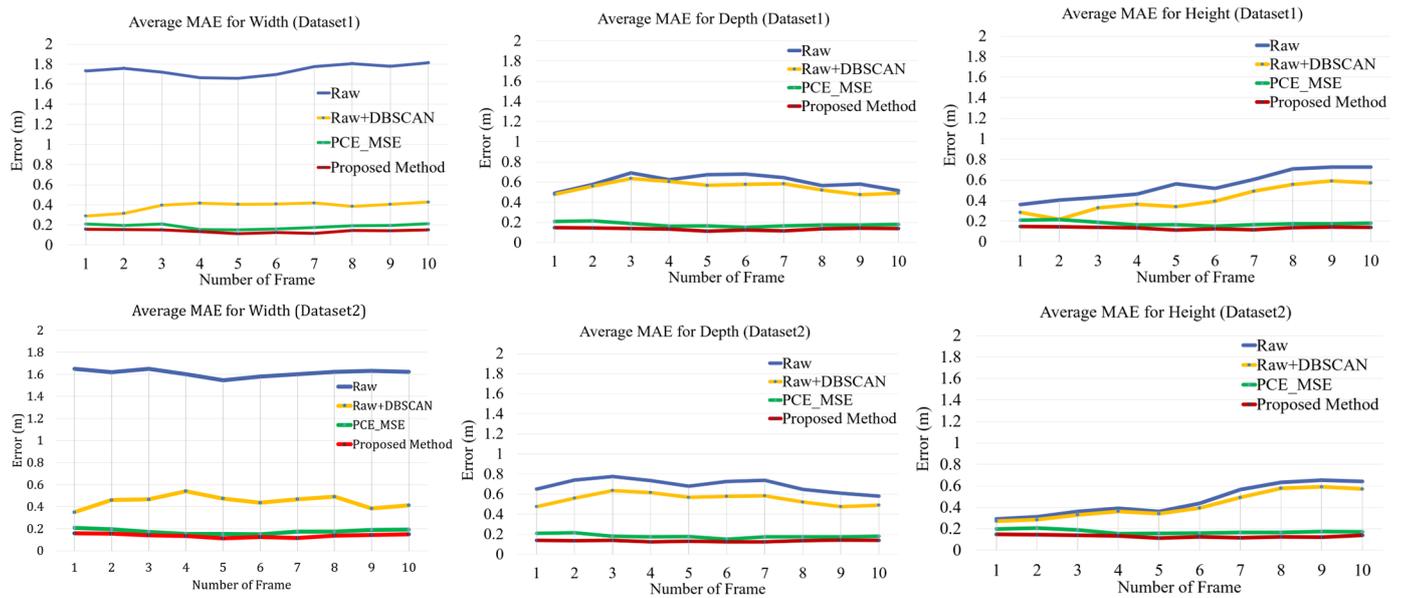
**Table 3.** Average 3D IoU of Dataset2 in each frame.

	1	2	3	4	5	6	7	8	9	10	Mean
Raw	0.300	0.287	0.266	0.257	0.249	0.225	0.212	0.207	0.191	0.189	<b>0.238</b>
Raw + DBSCAN [26]	0.276	0.265	0.262	0.249	0.269	0.233	0.235	0.224	0.222	0.193	<b>0.242</b>
PCE_MSE	0.591	0.613	0.625	0.614	0.612	0.628	0.591	0.571	0.574	0.526	<b>0.595</b>
Proposed PCE_HyLoss	0.649	0.673	0.677	0.697	0.652	0.665	0.624	0.629	0.614	0.618	<b>0.650</b>

Furthermore, we computed the centroids of the coordinates of the point clouds. For accuracy, the centroids of the coordinates of the predicted point clouds were compared with the centroids of the coordinates of the ground truth labels by calculating the mean absolute error (MAE) in the x-, y-, and z-coordinates. The centroid of the coordinates was  $\left(\frac{\max(x)+\min(x)}{2}, \frac{\max(y)+\min(y)}{2}, \frac{\max(z)+\min(z)}{2}\right)$ . The average MAEs for all L frames tested in Dataset1 and Dataset2 are shown in Figure 13. We also compared the raw point clouds from the radar, DBSCAN, and PCE models using the traditional MSE loss function instead of the proposed HyLoss. The average MAE of the predicted method was comparable to that of the other methods. In other words, the localization of the centroid from the PCE model was the closest to the ground truth. In addition, we computed the average MAE of the width, depth, and height of the bounding box along the x-, y-, and z-axis. As shown in Figure 14, the width, depth, and height of the predicted bounding box were significantly lower than those of the other methods for both Dataset1 and Dataset2. This result indicates that the PCE model with the proposed HyLoss offers a better image of the human body for a wide range of people and environments.



**Figure 13.** Localization error of the centroid of the points cloud compared to the baseline.



**Figure 14.** MAE of the width, depth, and height of the bounding box from the points cloud compared to the baseline.

#### 4.2.2. Classification for Fall Detection

In terms of fall detection, the predicted high-quality point clouds of Dataset0, Dataset1, and Dataset2 were fed into the feature extractor to obtain the human pose parameters. The lightweight CNN shown in Figure 9 was used to classify normal and fall events. Dataset0 was used to train the lightweight CNN classifier, Dataset1 was used for validation with new people, and Dataset2 was used for validation with new people and new environments. Neither Dataset1 nor Dataset2 were involved in the entire training process. To validate the choice of PCE architecture parameters, we compared the average accuracy, recall, and F1 of the PCE model with various parameters. The results are summarized in Table 4. Every encoder is denoted by E, and every decoder is denoted by D. For example, the proposed PCE model can be described as raw+32E+32E+16E+32D. Clearly, the proposed architecture is suitable. In addition, a comparison of the PCE with the traditional MSE loss function revealed that the accuracy, recall, and F1 score of the proposed HyLoss function outperformed those of the traditional function, which shows that the proposed HyLoss function improves the fall detection system.

**Table 4.** Classification result of the proposed PCE model using different architectures.

	Dataset0			Dataset1			Dataset2		
	Acc	Recall	F1	Acc	Recall	F1	Acc	Recall	F1
64E + 64E + 32E + 32D	0.993	0.994	0.995	0.976	0.977	0.977	0.972	0.973	0.972
16E + 16E + 16E + 32D	0.986	0.984	0.985	0.971	0.973	0.973	0.971	0.973	0.972
32E + 32E + 16E + 16E + 32D	1.000	0.994	0.994	0.991	0.993	0.993	0.988	0.988	0.988
32E + 32E + 16E + 32D + 32D	1.000	0.990	0.989	0.989	0.989	0.989	0.986	0.987	0.988
PCE_MSE (with traditional loss)	0.997	0.995	0.995	0.975	0.976	0.976	0.970	0.971	0.972
PCE with HyLoss (ours)	0.993	0.992	0.993	0.991	0.992	0.993	0.989	0.988	0.988

The purpose of PCE is to improve the reliability of the classification during fall detection tasks. Therefore, we compared the performance of the PCE classification reliability (instead of position errors or IoU) with existing point cloud models. First, the predicted points are shown in Figure 9. The results are listed in Table 5. The accuracy of the raw points and DBSCAN method for Dataset0 were higher than 0.98 because the test data included some people in the training set. The accuracy of the raw points (baseline) and DBSCAN

method for Dataset1 and Dataset2 were lower than 0.8 because the raw data from the radar detection contained many invalid points from new test data, and DBSCAN could not adequately remove a sufficient number of outlier points. However, the performance of the proposed method exhibited no obvious changes for any of the three datasets. This means that the point clouds predicted by the PCE could improve the performance of classification when new people and environments are involved. For computing the complexity analysis, the proposed method had 0.017 million parameters, which was a much smaller number than that of the other methods. The floating-point operations (FLOPs) associated with the proposed method was 0.303M, which was also lower than that of the other methods. In addition, the fall detection system achieved high accuracy and required less response time. For the response times of all the competing methods, we ran them on the same platform. Because all of them required the same computing time for feature extraction and classification, we only calculated the computing time of PCE on a single sample at a time. Although the accuracy was not favorable for new people and environments, the response time of the proposed method was shorter than that of the others. In summary, our work balances accuracy and speed.

**Table 5.** Classification performance of same datasets using current point cloud generation methods.

	Params (M)	FLOPs (M)	Response Time (ms)	Dataset0			Dataset1			Dataset2		
				Acc	Recall	F1	Acc	Recall	F1	Acc	Recall	F1
Raw	--	--	--	0.992	0.993	0.992	0.783	0.766	0.772	0.798	0.783	0.789
Raw + DBSCAN [26]	--	--	--	0.991	0.990	0.991	0.798	0.811	0.805	0.781	0.800	0.794
PointNet [27]	0.815	119	25.331	0.972	0.971	0.971	0.964	0.967	0.967	0.954	0.954	0.955
PCN [28]	4.430	4339	160.390	1.000	1.000	1.000	0.987	0.988	0.989	0.981	0.982	0.983
TopNet [37]	6.193	1916	149.867	0.995	0.996	0.995	0.988	0.987	0.987	0.986	0.985	0.984
GRNet [38]	64.938	10,962	4266.643	1.000	1.000	1.000	0.971	0.972	0.973	0.978	0.978	0.978
RFNet [39]	2.369	6532	434.403	0.996	0.995	0.995	0.991	0.991	0.990	0.989	0.989	0.988
PCE (ours)	0.017	0.303	20.326	0.993	0.992	0.993	0.991	0.992	0.993	0.989	0.988	0.988

However, fall detection based on radar currently uses different data formats, such as time-frequency maps and range-Doppler maps, and there is no public fall detection dataset for radar. It is difficult to find a baseline for a fall detection system; therefore, we compared the reported performance with that of other studies. As shown in Table 6, the proposed system achieves better performance for new people and environments. Although the performances of studies [22,40] was above 0.9, they were tested only on the same people and in the same environment. The performance of these studies may vary for new people and environments in a way similar to the results of the DBSCAN method shown in Table 5. In addition, although CW radar has lower cost, the accuracy of the proposed FMCW fall detection system was 0.989, which was higher than the CW radar system [41,42]. Moreover, even though the 77 GHz 3T4R radar has a higher resolution than our sensor, the proposed fall detection system could still outperform it [22].

**Table 6.** Comparison between the proposed method and other fall detection methods that use radar.

Study	Sensor	Data Format	Test Dataset		Performance
			New Environment	New People	
[22]	FMCW 77 GHz 3T4R	Point clouds	No	No	Accuracy: 0.98
[40]	FMCW 24 GHz 2T4R	Time-frequency map	No	No	Recall: 0.95 F2: 0.92
[41]	CW 25GHz	Time-frequency map	No	No	Accuracy: 0.825
[42]	CW 24GHz	Root mean-squared of signal	No	No	Accuracy: 0.977 Recall: 0.90
ours	FMCW 24 GHz 2T4R	Point Clouds	Yes	Yes	Accuracy: 0.989 Recall: 0.988 F1: 0.988

To investigate the potential of our fall detection system for real-time implementation, we evaluated the computational time of one sample for all the steps in our work. As shown in Table 7, although the computing time for the PCE and classification was 20.349 ms in total, the signal processing cost a significant amount of time because, in this study, we used the super-resolution algorithm MUSIC for DOA, which is considered to be the bottleneck of real-time fall detection systems. However, further research is required to design a system with lower computing costs for real-time detection and mobile devices. Furthermore, it is difficult to collect samples when real falls start among older adults. The limitations of the sample size may cause potential biases. In the future, we will improve our experiment and test it with large-scale experiments in more practical environments and with more subjects.

**Table 7.** Response time for each step in the system.

Step	Computing Time
Signal processing	6204.3 ms
CPE model	20.32 ms
Classification	0.029 ms

## 5. Conclusions

This study demonstrated a fall detection system based on the 3D point clouds of a 24 GHz FMCW radar. Unlike other conventional methods that utilize more costly hardware to improve detection accuracy, we used a low-cost small radar antenna array operating at 24 GHz to maintain high accuracy. By applying the PCE model to a fall detection system, we improved the quality of the point clouds and also the system performance, especially the accuracy, sensitivity, and generalization ability for new users and environments. As a result of our efforts to reduce computing costs in signal processing, the proposed method has the potential for widespread applications in monitoring the health of the elderly in an indoor environment without considering privacy protection.

**Author Contributions:** Conceptualization, T.L.; methodology, T.L.; software, T.L. and R.L.; validation, R.L. and L.Y.; data curation, T.L. and R.L.; writing—original draft preparation, T.L.; writing—review and editing, L.Y.; supervision, H.X.; project administration, C.-J.R.S.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by ICLegend Micro: No. 00000.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article.

**Acknowledgments:** The authors would like to thank all participants at ICLegend Micro for data collection.

**Conflicts of Interest:** Author Lei Yang and Yue Lin were employed by the company ICLegend Micro. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Rajagopalan, R.; Litvan, I.; Jung, T.P. Fall prediction and prevention systems: Recent trends, challenges, and future research directions. *Sensors* **2017**, *17*, 2509. [[CrossRef](#)] [[PubMed](#)]
2. Wang, S.; Wu, J. Patch-Transformer Network: A Wearable-Sensor-Based Fall Detection Method. *Sensors* **2023**, *23*, 6360. [[CrossRef](#)] [[PubMed](#)]
3. Wu, G.; Xue, S. Portable preimpact fall detector with inertial sensors. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2008**, *16*, 178–183. [[CrossRef](#)] [[PubMed](#)]
4. MacLean, M.K.; Rehman, R.Z.U.; Kerse, N.; Taylor, L.; Rochester, L.; Del Din, S. Walking Bout Detection for People Living in Long Residential Care: A Computationally Efficient Algorithm for a 3-Axis Accelerometer on the Lower Back. *Sensors* **2023**, *23*, 8973. [[CrossRef](#)] [[PubMed](#)]
5. Kan, X.; Zhu, S.; Zhang, Y.; Qian, C. A Lightweight Human Fall Detection Network. *Sensors* **2023**, *23*, 9069. [[CrossRef](#)] [[PubMed](#)]

6. Safarov, F.; Akhmedov, F.; Abdusalomov, A.B.; Nasimov, R.; Cho, Y.I. Real-Time Deep Learning-Based Drowsiness Detection: Leveraging Computer-Vision and Eye-Blink Analyses for Enhanced Road Safety. *Sensors* **2023**, *23*, 6459. [[CrossRef](#)] [[PubMed](#)]
7. Li, K.-J.; Wong, N.L.-Y.; Law, M.-C.; Lam, F.M.-H.; Wong, H.-C.; Chan, T.-O.; Wong, K.-N.; Zheng, Y.-P.; Huang, Q.-Y.; Wong, A.Y.-L.; et al. Reliability, Validity, and Identification Ability of a Commercialized Waist-Attached Inertial Measurement Unit (IMU) Sensor-Based System in Fall Risk Assessment of Older People. *Biosensors* **2023**, *13*, 998. [[CrossRef](#)]
8. Cardenas, J.D.; Gutierrez, C.A.; Aguilar-Ponce, R. Deep Learning Multi-Class Approach for Human Fall Detection Based on Doppler Signatures. *Int. J. Environ. Res. Public Health* **2023**, *20*, 1123. [[CrossRef](#)]
9. Ramirez, H.; Velastin, S.A.; Cuellar, S.; Fabregas, E.; Farias, G. BERT for Activity Recognition Using Sequences of Skeleton Features and Data Augmentation with GAN. *Sensors* **2023**, *23*, 1400. [[CrossRef](#)]
10. Balal, Y.; Yarimi, A.; Balal, N. Non-Imaging Fall Detection Based on Spectral Signatures Obtained Using a Micro-Doppler Millimeter-Wave Radar. *Appl. Sci.* **2022**, *12*, 8178. [[CrossRef](#)]
11. Balal, Y.; Balal, N.; Richter, Y.; Pinhasi, Y. Time-Frequency Spectral Signature of Limb Movements and Height Estimation Using Micro-Doppler Millimeter-Wave Radar. *Sensors* **2020**, *20*, 4660. [[CrossRef](#)] [[PubMed](#)]
12. Wang, B.; Zheng, Z.; Guo, Y.-X. Millimeter-Wave Frequency Modulated Continuous Wave Radar-Based Soft Fall Detection Using Pattern Contour-Confined Doppler-Time Maps. *IEEE Sens. J.* **2022**, *22*, 9824–9831. [[CrossRef](#)]
13. Wang, B.; Zhang, H.; Guo, Y.-X. Radar-based soft fall detection using pattern contour vector. *IEEE Internet Things* **2022**, *10*, 2519–2527. [[CrossRef](#)]
14. Liang, T.; Xu, H. A Posture Recognition Based Fall Detection System using a 24 GHz CMOS FMCW Radar SoC. In Proceedings of the 2021 IEEE MTT-S International Wireless Symposium (IWS), Nanjing, China, 23–26 May 2021; pp. 1–3. [[CrossRef](#)]
15. Yoshino, H.; Moshnyaga, V.G.; Hashimoto, K.J.I. Fall Detection on a single Doppler Radar Sensor by using Convolutional Neural Networks. In Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 6–9 October 2019; pp. 171–1724. [[CrossRef](#)]
16. Shankar, Y.; Hazra, S.; Santra, A. Radar-based Non-Intrusive Fall Motion Recognition using Deformable Convolutional Neural Network. In Proceedings of the International Conference of Machine Learning Applications (ICMLA), Boca Raton, FL, USA, 16–19 December 2019; pp. 2889–2892. [[CrossRef](#)]
17. Sadreazami, H.; Bolic, M.; Rajan, S. CapsFall: Fall Detection Using Ultra-Wideband Radar and Capsule Network. *IEEE Access* **2019**, *7*, 55336–55343. [[CrossRef](#)]
18. Wang, B.; Guo, Y. Soft fall detection using frequency modulated continuous wave radar and regional power burst curve. In Proceedings of the Asia-Pacific Microwave Conference (APMC), Yokohama, Japan, 7 December 2022; pp. 240–242. [[CrossRef](#)]
19. Jokanovic, B.; Amin, M. Fall Detection Using Deep Learning in Range-Doppler Radars. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 180–189. [[CrossRef](#)]
20. Tian, Y.; Lee, G.-H.; He, H.; Hsu, C.-Y.; Katabi, D. RF-Based Fall Monitoring Using Convolutional Neural Networks. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–24. [[CrossRef](#)]
21. Jin, F.; Sengupta, A.; Cao, S. mmFall: Fall Detection Using 4-D mmWave Radar and a Hybrid Variational RNN AutoEncoder. *IEEE Trans. Autom. Sci. Eng.* **2022**, *19*, 1245–1257. [[CrossRef](#)]
22. Kim, Y.; Alnujaim, I.; Oh, D. Human Activity Classification Based on Point Clouds Measured by Millimeter Wave MIMO Radar with Deep Recurrent Neural Networks. *IEEE Sens. J.* **2021**, *21*, 13522–13529. [[CrossRef](#)]
23. Guan, J.; Madani, S.; Jog, S.; Gupta, S.; Hassanieh, H. Through Fog High-Resolution Imaging Using Millimeter Wave Radar. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 11461–11470. [[CrossRef](#)]
24. Yang, Y.; Cheng, C. Construction of accurate group portrait of student funding based on Kmeans algorithm. In Proceedings of the 2023 8th International Conference on Information Systems Engineering (ICISE), Dalian, China, 23–25 June 2023; pp. 154–158. [[CrossRef](#)]
25. Wang, C.; Xiong, X.; Yang, H.; Liu, X.; Liu, L.; Sun, S. Application of Improved DBSCAN Clustering Method in Point Cloud Data Segmentation. In Proceedings of the 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), Zhuhai, China, 24–26 September 2021; pp. 140–144. [[CrossRef](#)]
26. Lu, Z.; Zhu, Z.; Bi, J.; Xiong, K.; Wang, J.; Lu, C.; Bao, Z.; Yan, W. Bolt 3D Point Cloud Segmentation and Measurement Based on DBSCAN Clustering. In Proceedings of the 2021 China Automation Congress (CAC), Beijing, China, 22–24 October 2021; pp. 420–425. [[CrossRef](#)]
27. Charles, R.Q.; Su, H.; Kaichun, M.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–16 July 2017; pp. 77–85.
28. Yuan, W.; Khot, T.; Held, D.; Mertz, C.; Hebert, M.J.I. PCN: Point Completion Network. In Proceedings of the 2018 International Conference on 3D Vision (3DV), Verona, Italy, 5–8 September 2018; pp. 728–737.
29. Pistilli, F.; Fracastoro, G.; Valsesia, D.; Magli, E. Learning Robust Graph-Convolutional Representations for Point Cloud Denoising. *IEEE J. Sel. Top. Signal Process.* **2021**, *15*, 402–414. [[CrossRef](#)]
30. Park, K.-E.; Lee, J.-P.; Kim, Y. Deep Learning-Based Indoor Distance Estimation Scheme Using FMCW Radar. *Information* **2021**, *12*, 80. [[CrossRef](#)]

31. Chen, V.V.; Li, F.; HO, S.-S.; Wechler, H. Micro-Doppler effect in radar: Phenomenon, model, and simulation study. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 2–21. [[CrossRef](#)]
32. Patole, S.M.; Torlak, M.; Wang, D.; Ali, M. Automotive radars: A review of signal processing techniques. *IEEE Signal Process. Mag.* **2017**, *34*, 22–35. [[CrossRef](#)]
33. Suzuki, N.; Hirata, K.; Wakayama, T. A fast calculation method of 2-dimensional MUSIC for simultaneous estimation of DOA and frequency. In Proceedings of the 2014 International Symposium on Antennas & Propagation (ISAP), Kaohsiung, Taiwan, 2–5 December 2014. [[CrossRef](#)]
34. Yang, S.; Yu, X.; Zhou, Y. LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example. In Proceedings of the 2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI), Shanghai, China, 12–14 June 2020; pp. 98–101. [[CrossRef](#)]
35. Mohammed, S.; Ab Razak, M.Z.; Abd Rahman, A.H. Using Efficient IoU loss function in PointPillars Network For Detecting 3D Object. In Proceedings of the 2022 Iraqi International Conference on Communication and Information Technologies (IICCIT), Basrah, Iraq, 7–8 September 2022; pp. 361–366. [[CrossRef](#)]
36. Zhou, D.; Fang, J.; Song, X.; Guan, C.; Yin, J.; Dai, Y.; Yang, R. IoU Loss for 2D/3D Object Detection. In Proceedings of the 2019 International Conference on 3D Vision (3DV), Québec City, QC, Canada, 16–19 September 2019; pp. 85–94. [[CrossRef](#)]
37. Tchapmi, L.P.; Kosaraju, V.; Rezatofighi, H.; Reid, I.; Savarese, S. TopNet: Structural Point Cloud Decoder. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–19 June 2019; pp. 383–392.
38. Xie, H.; Yao, H.; Zhou, S.; Mao, J.; Zhang, S.; Sun, W. GRNet: Gridding residual network for dense point cloud completion. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 365–381.
39. Ding, Y.; Yu, X.; Yang, Y. RFNet: Region-aware Fusion Network for Incomplete Multi-modal Brain Tumor Segmentation. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, 11–17 October 2021; pp. 3955–3964. [[CrossRef](#)]
40. Lu, J.; Ye, W.-B. Design of a Multistage Radar-Based Human Fall Detection System. *IEEE Sens. J.* **2022**, *22*, 13177–13187. [[CrossRef](#)]
41. Erol, B.; Gurbuz, S.Z.; Amin, M.G. GAN-based Synthetic Radar Micro-Doppler Augmentations for Improved Human Activity Recognition. In Proceedings of the 2019 IEEE Radar Conference (RadarConf), Boston, MA, USA, 22–26 April 2019; pp. 1–5. [[CrossRef](#)]
42. Hanifi, K.; Karsligil, M.E. Elderly Fall Detection with Vital Signs Monitoring Using CW Doppler Radar. *IEEE Sens. J.* **2021**, *21*, 16969–16978. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.