



# Technical Note Millimeter-Wave Radar Monitoring for Elder's Fall Based on Multi-View Parameter Fusion Estimation and Recognition

Xiang Feng, Zhengliang Shan, Zhanfeng Zhao \*, Zirui Xu, Tianpeng Zhang, Zihe Zhou, Bo Deng and Zirui Guan

School of Information Science and Engineering, Harbin Institute of Technology, Weihai 264209, China; fengxiang@hit.edu.cn (X.F.)

\* Correspondence: zhaozhanfeng@hitwh.edu.cn

Abstract: Human activity recognition plays a vital role in many applications, such as body falling surveillance and healthcare for elder's in-home monitoring. Instead of using traditional micro-Doppler signals based on time-frequency distribution, we turn to another way and use the Relax algorithm to process the radar echo so as to obtain the required parameters. In this paper, we aim at the multi-view idea in which two radars at different views work synchronously and fuse the features extracted from each radar, respectively. Furthermore, we discuss the common estimated time-frequency features and time-varying spatial features of multi-view radar-echo and then formulate the parameters matrix via principal component analysis, and finally transform them into the machine learning classifiers to make further comparisons. Simulations and results show that our proposed multi-view parameter fusion idea could lead to relative-high accuracy and robust recognition performance, which would provide a feasible application for future human–computer monitoring scenarios.

**Keywords:** multi-view feature fusion; millimeter-wave radar; parameters estimation; Relax algorithm; neural network



Citation: Feng, X.; Shan, Z.; Zhao, Z.; Xu, Z.; Zhang, T.; Zhou, Z.; Deng, B.; Guan, Z. Millimeter-Wave Radar Monitoring for Elder's Fall Based on Multi-View Parameter Fusion Estimation and Recognition. *Remote Sens.* 2023, *15*, 2101. https:// doi.org/10.3390/rs15082101

Academic Editor: Piotr Samczynski

Received: 26 March 2023 Revised: 10 April 2023 Accepted: 13 April 2023 Published: 16 April 2023



**Copyright:** © 2023 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

Human activity recognition plays a vital role in many applications, such as surveillance, and healthcare [1]. Especially, with the aging population worldwide, intelligent equipment for eldercare and healthcare monitoring has been badly needed. As discussed in [2], about one-third of elderly people suffer severe lesions and half of them suffer fall events repeatedly. How can sensors be used to detect or monitor fall events? The employed sensors can be categorized into wearable and contactless ones [3]. Wearable ones, such as bracelets and ankle sensors, must be worn or carried constantly, which might be inconvenient and easily broken or forgotten with high false alarm rates [4]. Given these limitations, contactless ones have gained wide research interests, which include cameras, microphones, and micro-radar systems. Cameras are vulnerable to lighting conditions and blind spots, and microphones are sensitive to ambient noise. Furthermore, these sensors might all infringe on privacy especially when deployed in private homes.

Based on the rapid development of low-cost frequency-modulated continuous-wave (FMCW) radars (such as TI 1443 mm-wave radar), the exploitation and extraction of radarecho data are utilized to detect, identify, and recognize motion [5–7]. Authors in [8] address the problem of radar-based human gait recognition by the dual-channel deep convolutional neural network (DC-DCNN). Meanwhile, a multidimensional principal component analysis (MPCA) was proposed to combine features of time, Doppler, and range information to achieve fall detection [9]. Authors in [10] tried to extract phase information contained in the complex high-resolution range profile to derive instantaneous velocity, acceleration, and jerking of the human body for fall detection and monitoring. Gurbuz and Amin [11] introduced deep learning (DL)-based data-driven approach for motion classification in indoor monitoring areas. Le Kernec et al. [12] proposed radar signal processing approaches for assisted living through three typical applications, i.e., human daily activity recognition, respiratory disorders, and sleep stages classification.

Typically, Chaccour et al. [13] summarized the existing fall-related systems and divided them into three categories, i.e., the wearable method, the nonwearable method, and the fusion method. Among these nonwearable methods, ones based on micro-Doppler signatures have been the hot-topic [14,15]. Authors in [16] focus on the feasibility of classifying different human activities based on micro-Doppler signatures where six features are extracted from the Doppler spectrogram and a support vector machine (SVM) is then trained using the measurement features to classify the activities. Other than the classical machine learning mechanisms (such as SVM), deep convolutional neural networks (DC-NNs) have also been proposed to tackle the micro-Doppler signatures. One issue must be considered wherein the design of handcrafted features might affect the conventional supervised learning paradigm, and also limit the scalability of the proposed schemes. Authors in [17] try to apply the DCNN directly to a raw micro-Doppler spectrogram for both human detection and activity classification problems, which can jointly learn the necessary features and classification boundaries without employing any explicit features on the micro-Doppler signals. Obviously, most researches focus on either the use of micro-Doppler signatures or different CNN models and might ignore the fact that practical applications often incorporate multi-view monitoring conditions. However, almost all these methods above (supervised or unsupervised ones) do not make full use of the spatially distributed radars that are realistically available. In addition, the non-ideal quality of radar echo would also affect the micro-frequency extraction even with much more computation.

Instead of using micro-Doppler features based on time-frequency distribution (TFD) [18], we turn to another way and use the Relax algorithm to analyze the radar echo. The Relax algorithm is an effective super-resolution spectrum estimation method, which estimates the signal parameters by minimizing the nonlinear least squares with strong adaptability. Most importantly, it does not make any restrictive assumptions about noise or clutter and achieves instantaneous frequency estimation in an iterative manner. Compared with traditional TFD features, radar-echo features by the Relax algorithm are equipped with higher time-frequency resolution and relative-low noise sensitivity. In this paper, as shown in Figure 1, we mainly focus on chirp-parameter estimation of the multi-view framework, which could separate the radar-echo contributions of each individual activity in a continuous recording way, and further process these parameters into a fusion matrix. Therein, we also discuss both the common estimated time-frequency features and time-varying spatial features, formulate the feature image after principal component analysis, and finally input them into the machine learning classifiers to make further comparisons. Our proposed multi-view parameter fusion estimation idea could lead to relative-high accuracy and robust recognition performance even in various conditions of distance, view angle, direction, and individual diversity.

The rest of this paper is organized as follows: Section 2 introduces the theory and algorithm of the parameter estimation method. In Section 3, the neural network structure has been discussed. Section 4 introduces the experimental methods and presents analyses and discussion of the recognition results. Section 5 is the conclusion.



Deep Learning

Figure 1. Overall schematic diagram of a multi-view monitoring system.

#### 2. Materials and Methods

#### 2.1. Radar Echo Analysis and Discussion

For human activity recognition using a prevalent 77 GHz Mm-wave radar, the micro-Doppler characteristic could reflect the target's geometric structure and motion properties and has been widely used in the classification and recognition of human activity [19,20]. Existing methods for instantaneous Doppler frequency exploitation can be broadly classified into two categories, i.e., non-parametric methods and parametric methods [21], but most ignore the multi-view effect. Therein, non-parametric methods do not need any parameterized model or equation, they only need to calculate the time-frequency distribution of echo signal, and extract the peak value of time-frequency distribution over time. Note that the Time-Frequency Distribution (TFD) usually equips low time-frequency resolution and noise sensitivity. To tackle this, a series of instantaneous Doppler frequency estimations based on parametric models are presented [22]. Parametric methods need to establish a multi-parameter model by using prior information and then using the original data to estimate the model parameters.

In this paper, we mainly adopt the parametric idea rather than TFD to extract obvious spatial-time-Doppler features of human activity, especially the falling of elders. When using the LFMCW radar, for any moment of activity monitoring, the echo signal of the body could be seen as the linear superposition of several parts (such as legs or arms). Assuming that the radar-echo signal is a linear superposition of *M* Linear Frequency Modulation (LFM) signal components, the echo model can be expressed as

$$S(t) = \sum_{i=1}^{M} A_i \exp[j \cdot 2\pi (f_{0_i}t + \frac{1}{2}m_it^2)]$$
(1)

where  $A_i$  denotes the complex amplitude of the *i*-th LFM signal component, and  $f_{0_i}$ ,  $m_i$ , M denote its initial frequency, the chirp rate, and the number of components, respectively.

Different from the conventional dechirp methods [23], here we introduce the FM Relax algorithm to estimate the initial frequency, chirp rate, and amplitude of each LFM signal component, where the Relax algorithm tries to estimate the signal parameters by minimizing the nonlinear least-squares (NLS) idea and has some strong adaptability [24,25]. Suppose that the radar echo of the human body could be constituted by *M* parts (such as

torso, legs, or arms), which corresponds to the *M* signal component of (1), then the discrete echo signal model could be expressed as

$$S(n) = \sum_{i=1}^{M} A_i \exp[j \cdot 2\pi (f_{0_i}n + \frac{1}{2}m_in^2)] + e(n), n = 0, 1 \dots N - 1$$
(2)

where *N* is the sample number, and  $A_i \exp[j \cdot 2\pi (f_{0_i}n + \frac{1}{2}m_in^2)]$  denotes the *i*-th LFM component, e(n) is the additive noise. Furthermore, if the noise part follows the Gaussian version, we can minimize the following nonlinear least square problem to obtain  $\{\hat{A}_i, \hat{f}_{0_i}, \hat{m}_i\}$ ,

$$C(\{f_{0_i}, A_i, m_i\}_{i=1}^M) = \left\| S - \sum_{i=1}^M \omega(f_{0_i}, m_i) A_i \right\|^2$$
(3)

$$S = [S(1) \ S(2) \ \dots \ S(N-1)]^T$$
 (4)

$$\omega(f_{0_i}, m_i) = \left[1 \exp[j \cdot 2\pi (f_{0_i} + \frac{1}{2}m_i)] \dots \exp[j \cdot 2\pi (f_{0_i}(N-1) + \frac{1}{2}m_i(N-1)^2)]\right]^T$$
(5)

where  $C(\{f_{0_i}, A_i, m_i\}_{i=1}^M)$  represent the cost function of nonlinear least square problem. Let  $A = [A_1 \dots A_M]^T$ ,  $\Omega = [\omega(f_{0_i}, m_i) \dots \omega(f_{0M}, m_M)]$ , then the problem in (3) can

be rewritten  $\left\{\hat{A}_{i}, \hat{f}_{0_{i}}, \hat{m}_{i}\right\} = \arg \min_{A_{i}, f_{0_{i}}, m_{i}} \|S - \Omega A\|^{2}$ (6)

To obtain the estimation value of  $\{\hat{A}_i, \hat{f}_{0_i}, \hat{m}_i\}$ , we first define the *k*-th LFM signal component as

$$S_{k} = S - \sum_{i=1, i \neq k}^{M} \hat{A}_{i} \omega(\hat{f}_{0_{i}}, \hat{m}_{i})$$
(7)

and assuming that the estimation values of  $\{\hat{f}_{0_i}, \hat{A}_i, \hat{m}_i\}$   $(i = 1, 2, ..., M, i \neq k)$  have been obtained, we can use Equations (8) and (9) to obtain  $\{\hat{A}_k, \hat{f}_{0_k}, \hat{m}_k\}$  [26],

$$(\hat{f}_{0_{k}}, \hat{m}_{k}) = \arg\min_{f_{0_{k}}, m_{k}} \left\| \left[ I - \frac{\omega(f_{0_{k}}, m_{k})\omega^{H}(f_{0_{k}}, m_{k})}{N} \right] S_{k} \right\|^{2}$$

$$= \arg\max_{f_{0_{k}}, m_{k}} \left| \omega^{H}(f_{0_{k}}, m_{k})S_{k} \right|^{2}$$
(8)

$$\hat{A}_{k} = \left. \frac{\omega^{H}(f_{0_{k}}, m_{k})S_{k}}{N} \right|_{f_{0_{k}} = \hat{f}_{0_{k}}}$$
(9)

The parameter estimation mechanism based on Relax algorithm has been listed below: Step 1: Suppose the number of pre-estimated signal component M = 1, we could use the dechirp method to get  $\hat{A}_1$ ,  $\hat{f}_{0_1}$ ,  $\hat{m}_1$ ;

Step 2: Suppose the number of pre-estimated signal component M = 2, based on the results of step 1, we could obtain  $S_2$  by (7);

Step 3: Once  $S_2$  has been obtained, we could obtain the  $\hat{A}_2$ ,  $\hat{f}_{0_2}$ ,  $\hat{m}_2$  by using (8) and (9); Step 4: Use  $\hat{A}_2$ ,  $\hat{f}_{0_2}$ ,  $\hat{m}_2$  to compute  $S_1$  by (7) again, and recompute  $\hat{A}_1$ ,  $\hat{f}_{0_1}$ ,  $\hat{m}_1$ ;

Step 5: Repeat steps 2–4, until the convergence threshold has been satisfied, which means the difference between two adjacent iterations is less than a certain threshold.

Step 6: Reiterate these steps until LFM signal component number *M* is equal to the total number of signal components.

Note that when solving the problem in (8), the step size of the modulation frequency should be sufficiently small. Typically, when using FFT to carry out the computation of the Fourier transform, the signal component would be supplemented by a sufficiently-long zero. In order to reduce the amount of calculation, a rough estimation can be obtained by using a larger frequency step size, and the smaller step size can be further used to optimize around the estimated value. As shown in Figure 1, in this paper, the Relax algorithm is used to process and estimate parameters of multi-view of radar A and B, respectively. To evaluate the effect of different signal components, for example, we could select 3 LFM components to make an analysis. Here, the estimated parameters of Radar A can be defined as Radar A-1~3, and Radar B can be defined as Radar B-1~3. Thus, these parameters vs. duration times of 6 human activities have been collected in Figure 2, where the horizontal axis represents the duration time of action, and the vertical axis represents the instantaneous frequency estimation at each moment.



Figure 2. Parameters varying with time for six different activities.

## 2.2. Feature Image Formulation Based on Relax Algorithm

Radar echo data can be divided into many time slices to perform further data processing. Radar echo in each time slice can be approximated as the superposition of M LFM signal components. Here, the Relax algorithm is first used to estimate the initial frequency of LFM components in each time slice, which corresponds to the instantaneous Doppler frequency. The index moment of the corresponding time slice and its estimated initial frequency are recorded in the form of a special matrix. In this matrix, each column represents the corresponding time slice of radar echo, and each row records their instantaneous Doppler frequencies for the *sub*-LFM component. Namely, the element in *i*-th row and *h*-th column of the matrix represents the instantaneous estimated Doppler frequency of the *i*-th LFM component at the *h*-th time slice. In this way, the estimated micro-Doppler information can be extracted and stored in this matrix. Typically, as the human body falls, different parts of the body have different characteristics. When only one radar is used to detect the fall action, the characteristics of the human body might be occluded, and the limited perspective would also lead to incomplete fall characteristics and affect the sequential classification. To tackle this, the dual-view radar detection idea is introduced in this paper, and two radars in different directions would be used to achieve detection synchronously. Therein, the echo data from two radars are processed, respectively, and their characteristic matrices are recorded simultaneously and spliced together as shown in Figure 3.



Figure 3. Characteristic matrices of two radars after matrix splicing.

Furthermore, after matrix splicing, we also use the principal component analysis (PCA) method to refine the characteristic matrices according to [27]. This is because the data from the two radars are not independent of each other, that is, there must be duplicate or closely related parts between the features extracted by the forward radar and the lateral radar. Moreover, simple matrix splicing doubles the input dimension of the neural network, greatly increasing the computational complexity of the network. PCA is a data dimensionality reduction method that aims to replace more variables with fewer variables and can reflect the majority of information from multiple variables. Using PCA can integrate data features from the two radars, preserving the most significant features and filtering out repetitive and interfering parts, which can improve the accuracy of fall monitoring.

For a sample  $X_{n \times p}$  containing p variables and n data, the covariance matrix is  $\sum_{p \times p}$ . The p eigenvalues of the covariance matrix ranging from large to small are  $\lambda_1, \lambda_2 \dots \lambda_p$  and the corresponding eigenvectors are  $T_1, T_2 \dots T_p$ . Then the *i*-th principal component can be obtained by using the following formula:

$$Y_i = XT_i , \ 1 \le i \le p \tag{10}$$

The cumulative contribution rate of the first *m* principal components is:

$$\psi_m = \frac{\sum\limits_{k=1}^m \lambda_k}{\sum\limits_{k=1}^p \lambda_k}$$
(11)

If the cumulative contribution rate reaches over 90%, we believe that selecting m principal components can well preserve the information of the original sample. Then, update the original characteristic matrix with the newly obtained m principal components as shown in Figure 4. The updated matrix is called a dual-view fusion matrix.



Figure 4. Refine the characteristic matrices using PCA. (PCm represents the *m*-th principal component).

Our proposed characteristic refining mechanism based on Relax algorithm has been listed below:

Step 1: For multi-view monitoring (such as two radars), once the radar-echo data are received, simultaneously divide them into 64 time slices, save each moment of the time slice, and name single frames.

Step 2: Use the Relax estimating method to obtain the initial frequencies of each time slice of these two radars, and store the frequency parameters that correspond to different LFM components.

Step 3: Formulate the characteristic matrix of Figure 3 and align each row and each column.

Step 4: Utilize the PCA method to refine the representative components.

Step 5: Transform the fusion feature matrix into a version of the color map, i.e., feature images, which would be further compared with TFD images.

The flow of the whole fusion process can be seen in Figure 5.



Figure 5. Fusion process of dual-view feature image.

Since these principal components might have some relations mapping to the torso, legs, arms, or head, the fusion matrix containing a series of information about human activity would be fed into a neural network that effectively judges whether it is a fall or not.

## 3. Neural Network Structure

For the neural network structure, as shown in Figure 6, the basic convolutional neural network (CNN) provides an end-to-end processing model in which their weights can be trained by the gradient descent method, and the trained CNN could learn the features of the image, and obtain the relative-high classification results.



Figure 6. Basic structure of convolution neural network.

Furthermore, we try to discuss three CNN models and evaluate their performance for our feature images. The typical VGG model focuses on stacking multiple  $3 \times 3$  convolution kernels to replace the large-scale convolution kernels so as to ensure the same receptive fields while reducing network parameters and increasing the network nonlinear expression capability [5,15]. The structure of VGG is shown in Figure 7.



Figure 7. Basic structure of VGG.

The Inception model replaces the fully connected layer with GAP and uses auxiliary classifiers to accelerate the convergence of the network [28]. The Inception structure uses the convolution layer of three convolution cores for parallel feature extraction, which can increase the width of the network model. The structure of Inception is shown in Figure 8.





The ResNet model proposes "skip connection" to solve the problem of model degradation [29]. The structure of Resnet is shown in Figure 9.



Figure 9. Basic structure of Resnet.

In this paper, we mainly propose the multi-view parameter estimating mechanism to achieve human activity classification. Here, we use these three different kinds of CNN above to train the dataset and further evaluate their training performance.

## 4. Experimental Setup and Result Analysis

## 4.1. Experimental Environment Construction and Scene Setting

In this section, we use the TI's IWR1443-BOOST mm-wave radar sensor, which has 2 transmitters and 4 receivers and a DCA1000 board to collect data. Here, we only used 1 transmitter and 4 receivers of them. The PC configuration is based on Windows 10, NVIDIA Quadro P620 graphics, and 12-core Intel i7-8700 CPU. Table 1 has listed the detailed parameters of radar configuration.

Table 1. Parameters of two radar configuration.

<b>Radar Parameters</b>	Value
Start Frequency	77 GHz
Frequency Slope	33 MHz/µs
Idle Time	100 µs
Bandwidth	1.981 MHz
ADC Start Time	6 µs
ADC Samples	256
Sample Rate	5 MHz
Number of Chirps	128
Number of Frames	64

The experimental scenarios have been shown in Figures 10 and 11, where radar A is the forward radar and radar B is the lateral radar. The angles of the two radars are 90 degrees and 45 degrees, respectively. Both radars are placed at the same level. The height of them from the ground is about 80 cm, and these radars are placed on the edge of the test area to ensure that the radars' beam is not occluded. Moreover, a cushion is placed in the experimental area directly in front of radar A which acts as a buffer in case the volunteer falls. The size of the cushion was 1.9 m  $\times$  0.9 m  $\times$  0.2 m, and the volunteer is about 2.5 m away from the forward radar A. There is no dynamic interference near the experimental area, and two radars start to collect data synchronously.



**Figure 10.** Schematic Diagram of Radar Experiment (**a**) The angle formed by two radars is 90 degrees; (**b**) The angle formed by two radars is 45 degrees.



Figure 11. Experiment scene diagram.

### 4.2. Dataset Production

Here, our experiment mainly collects six different action data including falls, bending, turning, etc., where falls have been further divided into fast falls and slow ones. Obviously, bending, turning and other actions are not truly falls. The detailed actions are shown in Table 2. To enlarge the data's scope and improve the generalization, 10 volunteers with different body shapes are selected whose ages are ranging from 19 to 40 years old, their heights are 160–185 cm and their weights are 55–85 kg. Each volunteer has performed all 6 actions in Table 2, and each action has been repeated 30 times where the admission time of each action is 6.4 s. In the end, a total dataset of 1800 action files has been stored.

Table 2. Different actions performed by volunteers.

Actions	Classification
Fall Fast	Fall
Fall Slowly	Fall
Bend	Not Fall
Turn Around	Not Fall
Walk	Not Fall
Step in Situ	Not Fall

It is worth noting that it would be very dangerous to have the elderly themselves finish the fall actions. Therefore, we use young volunteers aged 19–40 to replace them in the experiment. Naturally, the production of the dataset should eliminate the bias between the fall of the elderly and the fall of young volunteers. Therefore, we carefully study the real scene of the fall of the elderly and request all volunteers to strictly imitate the fall of the elderly to eliminate this bias. Moreover, from the perspective of radar echo, there are common physical characteristics of fall actions for both elderly and young subjects.

### 4.3. Dataset Processing and Feature Image Generation

Once the echo data of the two radars are obtained, the Relax algorithm is used to extract and estimate the parameters of each time slice. The estimated instantaneous frequency values are arranged by the time order to formulate the feature matrices of two perspectives, i.e., dual-view. The fusion feature matrix after PCA refining corresponding to different action files is transformed into a color mapping version, thus the feature images of different actions can be shown in Figure 12.



Figure 12. Characteristic images obtained by the Relax estimation for different actions.

Additionally, to carry out further comparison and analysis, we also present the traditional Micro-Doppler characteristic spectra of the TFD idea. As discussed in [30], the micro-motion of the target, such as vibration, rotation, and acceleration, will modulate the echo frequency and reveal some micro-Doppler characteristics. By using one-dimensional FFT and two-dimensional FFT on radar-echo data, respectively, micro-Doppler characteristic spectra of different actions were obtained and shown in Figure 13.



Figure 13. Traditional Micro-Doppler characteristic spectra of different actions.

## 4.4. Experimental Results and Analysis

Once the feature images were generated, these images were input into different neural networks for training. The dimensions of the input images are all  $224 \times 224 \times 3$ . The framework used for training is Pytorch. The train-test ratio of the dataset was set to 7:3. The other detailed training parameters set during the training process are shown in Table 3.

Training Parameters	Value
Optimizer	Adam
Train-test Ratio	7:3
Learning Rate	0.0001
Batchsize	32
Iterations	300
Epoch	60

Table 3. Training parameters of three CNN.

The loss function used for training is Cross Entropy Loss. Assume that the approximate probability distribution of sample x is Q(x), the true probability distribution is P(x), and the number of categories output by the model is n. The cross entropy can be defined as follows:

$$H(P,Q) = -\sum_{i=1}^{n} P(x_i) \log(Q(x_i))$$
(12)

If batchsize is *m*, the Cross Entropy Loss function is defined as follows:

$$loss = -\frac{1}{m} \sum_{i=1}^{m} \sum_{i=1}^{n} P(x_{ij}) \log(Q(x_{ij}))$$
(13)

The training results obtained through different methods and different networks are shown in Table 4.

Method	Network	Accuracy	<b>Prediction Time</b>	Average Accuracy
Single-view-based	Resnet	81.90%	0.3424 ms	
Relax mechanism	VGG	76.15%	0.2004 ms	78.83%
(90°)	Inception	78.45%	0.5313 ms	
Dual size has d	Resnet	93.36%	0.3543 ms	
Dual-view-based	VGG	90.36%	0.2122 ms	92.34%
Relax mechanism	Inception	93.30%	0.5591 ms	
Traditional	Resnet	77.66%	0.3562 ms	
micro-Doppler	VGG	71.56%	0.2141 ms	74.75%
spectrum mechanism	Inception	75.04%	0.5869 ms	

 Table 4. Comparison of different ideas for elder fall monitoring.

In Table 4, for different CNN models, such as Resnet, VGG, and Inception model, the dual-view-based Relax estimation mechanism achieved the best performance in average accuracy. The Resnet and Inception models outperformed VGG, irrespective of it being the dual-view case, singe-view case, or micro-Doppler spectrum case. From the second and third rows of Table 4, it can be seen that the accuracy of the dual-view Relax method was improved by 15.7%, 18.8%, and 18.26%, respectively, compared with the traditional micro-Doppler spectrum method. Meanwhile, for the metric of prediction time, the proposed dual-view characteristic matrix incorporated less data to be processed, so it consumes less time than the traditional micro-Doppler spectrum manner. Furthermore, comparing the first and third rows of Table 4, applying the Relax algorithm to extract fall features is significantly better than the traditional micro-Doppler spectrum one. Meanwhile, the accuracy of the single-view Relax method is higher than that of the traditional micro-Doppler method, with an average improvement of 4.08%. In terms of prediction time, the Relax method is significantly lower than that of the traditional micro-Doppler method, enabling a faster judgment of fall. Comparing the first and second rows of Table 4, the method with dual-view monitoring significantly improved the accuracy of fall detection compared to single-view monitoring. The accuracy of three CNNs improved by 15.46%, 12.21%, and 18.85%, respectively, and the average accuracy improved by 15.51%. In

addition, it can be seen in Figure 14 that the training loss of the dual-view data input network clearly converges to a lower value than that of the single-view. This indicates that the dual-view point could extract the feature information of the falling process more effectively.



Figure 14. Variation of loss during network training.

Confusion matrices are widely used in the assessment of classification problems. To further illustrate the effectiveness and advancement of the dual-view fusion Relax parameter estimation method proposed in this paper, a two-dimensional confusion matrix is first established, as shown in Figure 15.



**Figure 15.** 2D confusion matrix of fall. (TP, FP, FN, and TN represent True Positives, False Positives, False Negatives, and True Negatives, respectively.).

Based on the two-dimensional confusion matrix in Figure 15, we establish the following evaluation metrics and evaluate the experimental results accordingly. The evaluation results are shown in Table 5.

• Recall (*R*):

$$R = \frac{TP}{TP + FN} \tag{14}$$

• Precision (*P*):

$$P = \frac{TP}{TP + FP} \tag{15}$$

• F1-score (F1):

$$F1 = \frac{2 \times P \times R}{P + R} \tag{16}$$

• Matthews correlation coefficient (*Mcc*):

$$Mcc = \frac{TP \times TN - TP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}$$
(17)

• True Positive Rate (*TPR*):

$$TPR = \frac{TP}{TP + FN} \tag{18}$$

• False Positive Rate (*FPR*):

$$FPR = \frac{FP}{FP + TN} \tag{19}$$

Table 5. Comparison of classification performance of different ideas.

Method	Network	F1	Мсс
Single-view-based Relax mechanism (90°)	Resnet VGG Inception	0.8021 0.7663 0.7286	0.7988 0.8864 0.7766
Dual-view-based Relax mechanism	Resnet VGG Inception	0.8842 0.9247 0.8654	0.9769 0.9654 0.9887
Micro-Doppler spectrum mechanism	Resnet VGG Inception	0.5889 0.6077 0.6285	0.6672 0.6474 0.6988

In Table 5, compared with the traditional micro-Doppler spectrum method, the values of F1 and Mcc of the proposed dual-view Relax mechanism are obviously higher. Both the Relax parameter estimation method and the introduction of the dual-view idea greatly improve the accuracy of the model prediction. Furthermore, based on two metrics of TPR and FPR, we can also evaluate them by using the Receiver Operating Characteristic curve (ROC) and Area Under roc Curve (AUC). Generally, the ROC curve of an effective classifier should be above the line (0,0) and (1,1), and a larger AUC represents better performance. ROCs and its corresponding AUC of three different mechanisms can be seen in Figure 16 and Table 6.

Table 6. AUC obtained from three different mechanisms.

Method	AUC
Single-view-based Relax mechanism (90°)	0.7071
Dual-view-based Relax mechanism	0.8163
Micro-Doppler spectrum mechanism	0.6438



Figure 16. ROCs of three different mechanisms.

In Figure 16 and Table 6, the ROC of the dual-view Relax mechanism can completely wrap around the other two ROCs and thus has the largest AUC, showing optimal classification performance. In addition, we also listed the confusion matrix of classification results from different actions as shown in Figure 17. We compared the confusion matrix of the traditional micro-Doppler method and the dual-view Relax method. It is obvious that the classification effect of the confusion matrix obtained by the latter is noticeably better.



**Figure 17.** Confusion matrix of classification results from different actions: (**a**) Dual view based Relax mechanism; (**b**) Traditional micro-Doppler spectrum mechanism.

In Figure 17b, typically, we could see that for the traditional micro-Doppler case, "fall fast" and "fall slowly" have led to some obvious confusion while the dual-view-based case in Figure 17a has achieved this discrimination absolutely. Although the actions of "fall fast" and "fall slowly" have similar motion characteristics, there also exist spatial differences when using the dual-view idea. Compared with Figure 17a,b, the dual-view case has demonstrated better performance to judge "fall fast" and "fall slowly". Especially, by analyzing the confusion matrix of Figure 17b, we can conclude that the traditional micro-Doppler case might lose its performance with a relative-high alarm leakage rate.

Furthermore, we discuss the effect of a pre-defined number of LFM components, which affects the performance of the Relax mechanism. As we know, the pre-defined number also decides the radar-echo components of the human body. We made comparisons as shown in Figure 18.



Figure 18. Average accuracy when different numbers of LFM components are predefined.

In Figure 18, we list the average accuracy when different numbers of LFM components are predefined. The introduction of the dual-view idea makes a significant improvement in accuracy. As we are mainly focusing on the multi-view (i.e., dual-view) case, we omit the comparison with the traditional micro-Doppler spectrum case. With the number *n* increasing, the phenomenon would demonstrate that some typical components have affected the final average accuracy, namely, the predefined appropriate number corresponding to different body parts such as legs, head, and torso would also decide the final classification results.

## 5. Conclusions

This paper proposes a novel fall detection method based on the dual-view fusion idea, where the dual-view-based Relax mechanism combined with different CNN has achieved better classification results for fall detection than traditional ways. Therein, the instantaneous frequency estimation of time slices for dual-view radar-echo data has been estimated by the FM Relax algorithm, which makes full use of the characteristics of radar original data. In experiments, volunteers of different ages and body types are selected to perform a series of actions and further recorded into the training dataset where results and analysis have demonstrated some obvious performance. In addition, considering that the actual work scenarios of fall detection are more sophisticated, the anti-interference ability should be strengthened in our future work, so as to further improve the practical application for the elderly at home.

Author Contributions: Conceptualization, X.F. and Z.S.; Data curation, Z.S. and Z.Z. (Zhanfeng Zhao); Formal analysis, X.F. and Z.S.; Funding acquisition, X.F. and Z.Z. (Zhanfeng Zhao); Investigation, T.Z., X.F. and Z.S.; Methodology, X.F., Z.S., Z.Z. (Zhanfeng Zhao) and Z.X.; Project administration, X.F. and Z.S.; Resources, Z.S., Z.X., Z.Z. (Zihe Zhou), Z.G. and B.D.; Supervision, X.F.; Validation, Z.Z. (Zihe Zhou), B.D. and Z.G.; Writing—original draft, X.F. and Z.S.; Writing—review and editing, X.F. and Z.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant No. 42127804), the Major Scientific and Technological Innovation Projects of Shandong Province (Grant No. 2021ZLGX05 and 2022ZLGX04).

Data Availability Statement: Not applicable.

**Acknowledgments:** This research was supported by the National Natural Science Foundation of China (Grant No. 42127804), the Major Scientific and Technological Innovation Projects of Shandong Province (Grant No. 2021ZLGX05 and 2022ZLGX04).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- 1. Li, X.; He, Y.; Fioranelli, F.; Jing, X. Semisupervised human activity recognition with radar micro-Doppler signatures. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5103112. [CrossRef]
- 2. Gutierrez-Madronal, L.; La Blunda, L.; Wagner, M.F.; Medina-Bulo, I. Test event generation for a fall-detection IoT system. *IEEE Internet Things J.* **2019**, *6*, 6642–6651. [CrossRef]
- 3. Ding, C.; Hong, H.; Zou, Y.; Chu, H.; Li, C. Continuous human motion recognition with a dynamic range-Doppler trajectory method based on FMCW radar. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6821–6831. [CrossRef]
- 4. Amin, M.G.; Zhang, Y.D.; Ahmad, F.; Ho, K.C.D. Radar signal processing for elderly fall detection the future for in-home monitoring. *IEEE Signal Process. Mag.* 2016, 33, 71–80. [CrossRef]
- 5. Abdu, F.J.; Zhang, Y.; Deng, Z. Activity classification based on feature fusion of FMCW radar human motion micro-Doppler signatures. *IEEE Sens. J.* 2022, 22, 8648–8662. [CrossRef]
- Ding, C.; Zhang, L.; Chen, H.; Hong, H.; Zhu, X.; Li, C. Human motion recognition with spatial-temporal-convLSTM network using dynamic range-doppler frames based on portable FMCW radar. *IEEE Trans. Microw. Theory Tech.* 2022, 70, 5029–5038. [CrossRef]
- 7. 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]
- 8. Bai, X.; Hui, Y.; Wang, L.; Zhou, F. Radar-based human gait recognition using dual-channel deep convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9767–9778. [CrossRef]
- 9. Erol, B.; Amin, M.G. Radar data cube analysis for fall detection. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 2446–2450.
- Tran, N.; Kilic, O.; Nahar, S.; Ren, L.; Wang, H.; Fathy, A.E. Contactless monitoring and classification of human motion activities by using SFCW radar. In Proceedings of the 2016 IEEE International Symposium on Antennas and Propagation (APSURSI), Fajardo, PR, USA, 25 June–2 July 2016; pp. 883–884.
- 11. Gurbuz, S.Z.; Amin, M.G. Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring. *IEEE Signal Process. Mag.* 2019, *36*, 16–28. [CrossRef]
- 12. Kernec, J.L.; Fioranelli, F.; Ding, C.; Zhao, H.; Romain, O. Radar signal processing for sensing in assisted living: The challenges associated with real-time implementation of emerging algorithms. *IEEE Signal Process. Mag.* **2019**, *36*, 29–41. [CrossRef]
- 13. Chaccour, K.; Darazi, R.; Hajjam, A.; Andres, E. From fall detection to fall prevention: A generic classification of fall-related systems. *IEEE Sensors J.* 2017, *17*, 812–822. [CrossRef]
- 14. Narayanan, R.M.; Zenaldin, M. Radar micro-Doppler signatures of various human activities. *IET Radar Sonar Navig.* 2015, 9, 1205–1215. [CrossRef]
- 15. Alnujaim, I.; Oh, D.; Kim, Y. Generative adversarial networks for classification of micro-Doppler signatures of human activity. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 396–400. [CrossRef]
- 16. Kim, Y.; Ling, H. Human activity classification based on microDoppler signatures using a support vector machine. *IEEE Trans. Geosci. Remote Sens.* **2009**, 47, 1328–1337.
- 17. Kim, Y.; Moon, T. Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks. *IEEE Geosci. Remote Sens. Lett.* **2015**, *13*, 8–12. [CrossRef]
- Zhao, R.; Ma, X.; Liu, X.; Li, F. Continuous human motion recognition using micro-Doppler signatures in the scenario with micro motion interference. *IEEE Sens. J.* 2020, 21, 5022–5034. [CrossRef]
- 19. Luo, F.; Poslad, S.; Bodanese, E. Human activity detection and coarse localization outdoors using micro-Doppler signatures. *IEEE Sens. J.* **2019**, *19*, 8079–8094. [CrossRef]
- 20. Qiao, X.; Amin, M.G.; Shan, T.; Zeng, Z.; Tao, R. Human activity classification based on micro-Doppler signatures separation. *IEEE Trans. Geosci. Remote Sens.* 2021, 60, 5105014. [CrossRef]
- 21. Li, Y.; Xia, W.; Dong, S. Time-based multi-component irregular FM micro-Doppler signals decomposition via STVMD. *IET Radar* Sonar Navig. 2020, 14, 1502–1511. [CrossRef]
- 22. Wang, Y.; Yang, W.; Li, D.; Zhang, J. A novel time-frequency model, analysis and parameter estimation approach: Towards multiple close and crossed chirp modes. *Signal Process.* **2022**, 201, 108692. [CrossRef]
- Li, W.; Yang, J.; Zhang, Y.; Lu, J. Robust wideband beamforming method for linear frequency modulation signals based on digital dechirp processing. *IET Radar Sonar Navig.* 2019, 13, 283–289. [CrossRef]
- 24. Ren, J.; Zhang, T.; Li, J.; Stoica, P. Sinusoidal parameter estimation from signed measurements via majorization–minimization based RELAX. *IEEE Trans. Signal Process.* **2019**, *67*, 2173–2186. [CrossRef]

- 25. Serbes, A.; Qaraqe, K. A fast method for estimating frequencies of multiple sinusoidals. *IEEE Signal Process. Lett.* **2020**, *27*, 386–390. [CrossRef]
- 26. Shao, S.; Zhang, L.; Wei, J.; Liu, H. Two-dimension joint super-resolution ISAR imaging with joint motion compensation and azimuth scaling. *IEEE Geosci. Remote Sens. Lett.* 2020, *18*, 1411–1415. [CrossRef]
- 27. Erol, B.; Amin, M.G. Radar data cube processing for human activity recognition using multisubspace learning. *IEEE Trans. Aerosp. Electron. Syst.* 2019, *55*, 3617–3628. [CrossRef]
- 28. Xu, S.; Zhang, L.; Huang, W.; Wu, H.; Song, A. Deformable convolutional networks for multimodal human activity recognition using wearable sensors. *IEEE Trans. Instrum. Meas.* 2022, *71*, 2505414. [CrossRef]
- 29. Mansour, R.F.; Aljehane, N.O. An optimal segmentation with deep learning based inception network model for intracranial hemorrhage diagnosis. *Neural Comput. Appl.* 2021, 33, 13831–13843. [CrossRef]
- 30. Hanif, A.; Muaz, M.; Hasan, A.; Adeel, M. Micro-Doppler based target recognition with radars: A review. *IEEE Sens. J.* 2022, 22, 2948–2961. [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.