A convolution neural network approach for fall detection based on adaptive channel selection of UWB radar signals

Ping Wang^{1,2} · Qimeng Li^{1,3} · Peng Yin¹ · Zhonghao Wang¹ · Yu Ling² · Raffaele Gravina³ · Ye Li¹

Received: 12 June 2021 / Accepted: 25 November 2021
© The Author(s) 2021

Abstract

According to the World Health Organization and other authorities, falls are one of the main causes of accidental injuries among the elderly population. Therefore, it is essential to detect and predict the fall activities of older persons in indoor environments such as homes, nursing, senior residential centers, and care facilities. Due to non-contact and signal confidentiality characteristics, radar equipment is widely used in indoor care, detection, and rescue. This paper proposes an adaptive channel selection algorithm to separate the activity signals from the background using an ultra-wideband radar and to generalize fused features of frequency- and time-domain images which will be sent to a lightweight convolutional neural network to detect and recognize fall activities. The experimental results show that the method is able to distinguish three types of fall activities (i.e., stand to fall, bow to fall, and squat to fall) and obtain a high recognition accuracy up to 95.7%.

Keywords Ultra-wideband Radar · Fall detection · Healthcare · Channel Selection · Convolutional Neural Network

Ping Wang, Qimeng Li have contributed equally to this work.

✉ Peng Yin
peng.yin@siat.ac.cn

Ping Wang
wangping@siat.ac.cn

Qimeng Li
qimeng.li@unical.it; qm.li@siat.ac.cn

Zhonghao Wang
wang.zh@siat.ac.cn

Yu Ling
lingyu@fhxy.net.cn

Raffaele Gravina
r.gravina@dimes.unical.it

Ye Li
ye.li@siat.ac.cn

¹ Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

² Fiberhome Technologies College, Wuhan Research Institute of Post and Telecommunications, Wuhan, China

³ University of Calabria, Rende, CS 87036, Italy

1 Introduction

In today's world, the population is growing and aging, resulting in an unbalanced population structure. The United Nations and Social Affairs reported that by 2050 the world's population will reach 10 billion from the current 7.7 billion [1]. The population over the age of 65 accounts for about 9.1% (i.e., 701 million) and will reach 16% in 2050. At that time, elderly care will be one of the most prominent issues in the world.

Moreover, according to the World Health Organization and other authorities, falls account for 50.96% of the causes of accidental injuries [2] and even deaths among older people. Therefore, timely fall detection and treatment are essential to protect the health of the elderly.

Researchers have adopted different solution to detect and identify falls in elderly care [3–5]. Wearable sensors such as accelerometers, gyroscopes [6, 7], electrocardiogram (ECG) are firstly used in monitoring senior's physical and physiological status for fall detection and recognition. However, wearable devices are easy to forgotten or lose [8]. Therefore, non-contact methods have been introduced to capture fall actions. One of the techniques is to

use computer vision-based methods [9, 10], but the drawback is that it faces privacy issues.

Other methods are also used, such as ambient sensors [11], which mainly include pressure sensors, infrared sensors, and ultrasonic sensors. Ambient sensors are characterized by diverse data, however, based on the size and complexity of the environment, several devices may need to be deployed [12].

In this work, we adopted an ultra-wideband (UWB) radar to obtain raw data and used an adaptive channel select method [13] to separate the background from the useful signal. Then, a fused feature set of frequency- and time-domain images are used to train the model for fall recognition.

Our approach can capture human activities even through an obstacle such as a wall and track human movements. The main contributions of the work are the following:

- continuous monitoring and recognition of the most common acute events (e.g., fall) in the home life of the elderly;
- an adaptive channel selection algorithm for distinguishing background with fall activity;
- a feature fusion method based on frequency- and time-domain images for increasing recognition accuracy.

The rest of the article is organized as follows: Sect. 2 discusses the related work on sensing methods of fall detection. Section 3 describes the proposed radar-based system and introduces the experiment setup. Section 4 describes in detail the algorithm in our paper. Section 5 explains the experimental results and discussion. Finally, Sect. 6 concludes the paper and outlines planned future work.

2 Related work

As introduced in the previous section, there are several solutions to automatically detect human falls. We can identify three main approaches:

1. computer-vision based method (based on cameras);
2. wearable technologies such as smartwatches, smartphones, and smart belts;
3. non-contact sensors such as passive infrared sensors, magnetic contact sensors, pressure mats, or radio-frequency sensors.

2.1 Computer vision-based fall detection

Computer vision-based fall detection methods generally use a camera, which is usually fixed at a specific position and obtains continuous data frames for detecting and

recognizing activities. In [14], the authors proposed a 3-Dimensional convolutional neural network (CNN) based method for fall detection, which only uses video kinematic data to train an automatic feature extractor and could circumvent the requirement of a large sample size of data. In [15], the authors proposed a method to detect falls by analyzing human shape deformation during a video sequence. The experiments have been conducted on an actual data set of daily activities and simulated falls and give promising results compared with other standard image processing methods. An interesting study has been conducted by [16], where a lightweight neural network, namely You Only Look Once third edition (YOLOv3), was proposed to improve accuracy and responsiveness of fall detection.

However, privacy concerns cannot be ignored for the camera-based approach. Therefore, Kinect depth images are being used to capture shadow-like images of the patient and their room to resolve concerns about patients' privacy in [17]. As a result of previous research, a fall detection system has been developed and installed in hospital rooms, and alarms are generated upon the detection of fall events. Then nurses would go through the stored depth videos to investigate for possible injury as well as the causes that may have caused the patient's fall to prevent future occurrences.

The data from computer vision-based sensors is intuitive and easy to analyze but lacks the ability to protect privacy. Even though depth cameras exhibit low-privacy concerns, the cost of such devices is typically high.

2.2 Wearable sensor-based fall detection

Wearable-based fall detection is currently the most popular approach due to the thriving development of sensor technologies and pervasive computing. Indeed, this approach is mainly focused on using the motion data from sensors, such as accelerometers, and gyroscopes. These sensors are directly integrated into devices with microcontrollers, such as smartphones or smartwatches, and can be worn by the residents [18, 19]. Ballı et al. used a smartwatch machine learning approach to recognize the fall action [20]. Zhao et al. proposed a method based on a tri-axial gyroscope for fall events recognition. A tri-axial gyroscope is placed at the user's waist to collect tri-axial angular velocity information [21]. Also, De Araújo et al. [22] present a smartwatch-based accelerometer to detect falls. Although wearable sensors are usually small, lightweight and easy to deploy, wireless communication is sometimes unstable; meanwhile, the device is easily forgotten or lost and requires frequent charging.

2.3 Ambient sensor-based fall detection

Ambient devices that monitor falls mainly include pressure sensors, infrared sensors, radar systems, and radio-frequency equipment. In [23], a novel double pressure sensors-based system was proposed. The result with random forest yielded the best fall detection model with 100% accuracy. Ogawa et al. [24] proposed a fall detection method using infrared radiation sensor array. Miawarni et al. [25] proposed 2-dimensional lidar as the main sensor in a fall detection system, with high recognition accuracy.

The application of environmental sensors in fall detection usually uses infrared sensors or pressure sensors, etc. These sensors require higher requirements of the environment, for example, the infrared sensor requires an environment without any obstructing objects, and the pressure sensors need to be deployed on a large area in some scenes.

2.3.1 UWB radar-based fall detection

Compared with previous methods, the radio-frequency method is a better choice for home care monitoring and to detect accidents that occur (e.g., fall), thanks to the non-contact data collection and intrinsic privacy preservation. Li et al. [26] used UWB radar and three inertial sensors on the wrist, waist, and ankle, and rely on a bidirectional Long Short-Term Memory (bi-LSTM) network with multi-information fusion, and has achieved high accuracy in detecting falls. However, using multiple heterogeneous sensors leads to complex operations and increases the complexity of the algorithm. Julien et al. [27] extracted fall features based on weight joint distance time-frequency transformation and uses bagged decision tree and k-Nearest Neighbor (kNN) to obtain an accuracy of 91.5% and 88.6%, respectively. Sadreazami et al. [28] presented a radar-based fall detection method using compressed features of the radar signals. The compressed features are obtained by using deterministic row and column sensing. The time-frequency analysis is first performed on the radar time series and resulting spectrogram is projected onto a binary image representation. The binary images are then compressed using a 2D deterministic sensing technique by preserving the aspect ratio of the images in the compressed domain. The performance of the proposed method is evaluated using several classifiers. It is shown that the proposed compressive sensing based method can improve fall versus non-fall activities recognition. Khawaja et al. [29] used multiple UWB radar transceivers introducing a fall detection, localization and tracking technique for people in need of assistance. The proposed method allows for precise monitoring of people with special needs without any tags or wearables. To enhance ranging

precision, the authors introduced a novel method of fall detection based on the residual co-variance from extended Kalman filter. Computer simulations demonstrated the effectiveness of the proposed technique for fall detection applications.

As shown in Table 1, we can clearly understand the advantages and disadvantages of various solutions. In this paper, we select UWB radar for its properties of non-contact monitoring, easy deployment, high-resolution, and low-privacy concerns for detecting falls of the elderly. Meanwhile, to improve the accuracy and stability of fall monitoring, we analyze the signal to accurately identify the localization information of the fall activity and extract its detailed action and using a fusion method with a deep convolutional neural network to archive a high accuracy recognition.

3 Proposed system and experimental setup

3.1 System prototype

As shown in Fig. 1a, the UWB radar equipment chip used for data acquisition is the NVA-R631 in the NVA-R6X1 Novella series development board produced by Novelda; two patch antennas for transmission and reception are placed parallel. Universal Serial Bus-Serial Peripheral Interface (USB-SPI) bus conversion interface supports the data transmission with a transmission rate of 480Mb/s.

In a radar system, the sampling data consists of two-dimensional time: slow time sampling, which refers to the actual time; fast time sampling, which refers to the distance of an object from the radar. Figure 1b shows the data transmitted via a transmitting antenna to a receiving antenna, then passed by ADC (analog-to-digital converter) to generate the raw data in which the horizontal axis is fast time, and the vertical axis is slow time.

3.2 System model for detecting human motion

UWB radar transmits a first-order Gaussian pulse signal $p(t)$, which can be roughly expressed by Formula 1, where $A(t)$ represents the pulse waveform, and T_p represents the pulse interval.

$$p(t) = \begin{cases} A(t), & 0 \leq t < T_p \\ 0, & \text{others} \end{cases} \quad (1)$$

The transmitted signal $p_{tr}(t)$ of the system can be expressed by Formula 2, where M represents the number of transmitted pulses, and T_{pr} represents the pulse period.

Table 1 Comparison of the Advantages and Disadvantages of common fall detection approaches

Solutions	Advantage	Disadvantage
Computer Vision [14–17]	Visual Intuitive Easy to analyze Large amount of information	Poor privacy protection
Wearable Sensors [18–22]	Small volume Lightweight Easy to deploy	Easily forgotten and lost Frequent charging May have communication problem
Ambient Sensors [12, 23–25]	Diversified data types Low privacy concern	Higher requirements of the space environment
UWB radar [27–29]	Wide bandwidth Low privacy concern	Fine signal processing methods required

$$p_{tr}(t) = \sum_{M=1}^M p(t - (M-1)T_{pr}) \quad (2)$$

After high-speed sampling, the radar system receives the signal frames, see Formula 3, where m represents the m -th frame of the received radar signal r_m , with $n \in N$ represents the number of the channel. All continuous frames can be formed as the radar signal $\mathbf{R}$ (see Formula 4), where T_s and T_f represent fast time and slow time sampling interval, respectively.

$$\mathbf{r}_m = [\mathbf{r}_{m,1} \quad \dots \quad \mathbf{r}_{m,n} \quad \dots \quad \mathbf{r}_{m,N}]^T \quad (3)$$

$$\mathbf{R}[m, n] = s(t = mT_s, \tau = nT_f) \quad (4)$$

The received signals $s(t, \tau)$ can be expressed by Formula 5, where a_j is the shock signal amplitude of a stationary object in the surrounding environment; a_v is the signal amplitude of the human body; c is the electromagnetic wave speed; τ and t are fast time and slow time at a certain moment, respectively; r_0 , $\Delta r(t)$ represent the average distance with its change of the human body from the radar, respectively; f_j and A_j represent the frequency and amplitude of each channel, respectively; s_{noise} represents all the noise signals of the radar.

$$\begin{cases} s(t, \tau) = \sum_j a_j p(\tau - \tau_j) + a_v p(\tau - \tau_v(t)) + s_{noise} \\ \tau_v(t) = \frac{2(r_0 + \Delta r(t))}{c} = \frac{2(r_0 + \sum_j A_j \sin(2\pi f_j t))}{c} \end{cases} \quad (5)$$

3.3 Experiment setup

The radar system is assembled in an indoor environment and deployed on a desk with a height of 1 m to the floor. The data collection area is 1.2 m–1.5 m away from the radar in the longitudinal direction as shown in Fig. 2a and b.

The collected data samples came from nine male volunteers aged 24–40. The average height and weight of the volunteers are 172.5 ± 4.6 m and 69.6 ± 7.9 kg. The

volunteers simulated three fall actions, which frequently occur in senior population. We collected a total of 400 sets containing three types of fall data. The radar data relies on Matlab scripts for data collection. Our experiment sets the fast time channel and the slow time sampling frequency of the radar are 512 and 10 Hz, respectively.

The protocol of the experiments is summarized in the following:

- *Stand to Fall*—The subject was asked to stand in front of the mat and, after holding for two seconds, to fall down on the mat;
- *Bow to Fall*—The subject was asked to bend over in front of the mat and, after holding for two seconds, to fall down on the mat;
- *Squat to Fall*—The subject was asked to squat down in front of the mat and, after holding for two seconds, to fall down on the mat.

Each participant was asked to repeat all the above simulated falls 15 times, and each activity lasted 4 seconds on average.

4 Method

4.1 Data preprocessing

The collected data contains various noise sources including low-frequency noise reflected by the surrounding environment and high-frequency noise inside and outside the radar, which will seriously affect the detection of falls. Therefore, it is necessary to remove low-frequency and high-frequency noise for obtaining good detection. Wang et al. [30] used a single stage canceller to filter out low-frequency noise, but high-frequency noise and clutter will still exist and would affect the detection accuracy. In [31], the stationary and non-stationary clutters were removed by employing the singular value decomposition (SVD) algorithm when the signal-to-noise ratio (SNR) is low.

Considering the computational time cost and complexity, we decided to send raw data in parallel to both Fast

Fig. 1 Diagram of the radar system

Fourier transform (FFT) and SVD filters. The FFT filter will filter out the direct-current (DC) component and some parts of the low frequency of the signal to obtain the FFT image features (i.e., frequency-domain feature); and the SVD filter will remove high-frequency clutters and low-frequency background noise to generalize SVD image features (i.e., time-domain feature).

4.1.1 FFT filter and frequency-domain feature image extraction

FFT decomposes a function of time (a signal) into the frequencies. As we can observe in the Fig. 3, the raw data from the radar sensor is a discrete signal. The FFT, therefore, is used to preprocess the signal (see Formula 6) to filter the DC component and a part of low frequency noises of the signal.

(a) View of the Experimental Setup.

(b) Radar Deployment Profile.

Fig. 2 Deployment of the device with the experiment environment

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{\frac{-j2n\pi}{N}} \quad (k = 0, 1, 2, \dots, N-1) \quad (6)$$

After filtering out the noise, the activity performed by the subject can be observed in blue dash boxes (see Fig. 3). The frequency-domain feature image is the center-shifted FFT image.

4.1.2 SVD filter and time-domain feature image extraction

The SVD algorithm is widely used for dimension reduction and noise filtering in signal processing. In our method, the image can be seen as a matrix $\mathbf{A}$ (i.e., an $m \times n$ -order), all elements of which belong to the domain K , i.e., the real numbers space or the complex numbers space. According to the SVD, $\mathbf{A}$ can be decomposed as in Formula 7, where

Fig. 3 Raw data processed with FFT filter

U is a unitary matrix of order $m \times m$; Σ is a positive semi-definite $m \times n$ order diagonal matrix; and V^T , the conjugate transpose of V , is a unitary matrix of order $n \times n$. U and V are both orthogonal matrices, such that $UU^T = I$ and $VV^T = I$.

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T \quad (7)$$

The left singular vector is the eigenvector of AA^T , and the right singular vector is the eigenvector of $A^T A$, as shown in Formula 8, 9, where λ_i and ζ_i are eigenvalues corresponding to the eigenvectors u_i and v_i , respectively. Obviously, we can get U , V from u_i, v_i .

$$(AA^T)u_i = \lambda_i u_i \quad (8)$$

$$(A^T A)v_i = \zeta_i v_i \quad (9)$$

Finally, Σ can be calculated by the Formula 10, where σ_i is the singular value and used to compose the Σ .

$$\begin{aligned} A_{m \times n} &= U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T \\ \Rightarrow AV &= U \Sigma V^T V \\ \Rightarrow AV &= U \Sigma \\ \Rightarrow Av_i &= \sigma_i u_i \\ \Rightarrow \sigma_i &= u_i^T Av_i \end{aligned} \quad (10)$$

After the above calculation, we can obtain $A = \sum_{i=1}^r \sigma_i u_i v_i^T$. Generally, the larger σ is, the more significant its contribution to matrix A . Therefore, we filter according to this principle to remove some values with smaller σ . After SVD filtering and reconstruction, the time-domain feature image is obtained as shown in Fig. 4, and the activity performed by the subject is also clearly observed in blue dash boxes.

Fig. 4 Raw data processed with SVD filter

Fig. 5 Energy spectrum with energy threshold selection

4.2 Adaptive channels selection algorithm

After filtering, we can observe the activity that occurs in specific channels (see Fig. 5), and the energy of these channels is higher than that of the channels related to the background; the channel energy is calculated by Formula 11.

$$E(k) = |X(w)|^2 \quad (11)$$

Therefore, in this work, we propose an adaptive channel selection Algorithm 1 to distinguish the background from the activity performed by the subject.

Algorithm 1 Adaptive Channel Selection Algorithm

Require: X : data sample contains fall activity;
 x_i : data from i_{th} channel;
 j : the length of sliding window with $j = 2n$ ($n \in \mathbb{N}^*$);
 $f[]$:empty array; $l[]$:empty array.
Ensure: i_{first} : the index of the selected start channel;
 i_{last} : the index of the selected end channel;
1: Initial: $i = 1$;
2: Calculate energy of i_{th} channel, $E(x_i)$; energy threshold, $E(th)$; energy difference of i_{th} channel, $\Delta E(x_i)$
3: **while** $i < i_{max}$ **do**
4: **if** $\Delta E(x_i) < 0$ **then**
5: $\Delta E(x_i) = 0$;
6: **else**
7: $\Delta E(x_i) = 1$;
8: **end if**
9: $i += 1$
10: **end while**
11: **for** $i = 1$; $i < i_{max} - j$; $i += j$ **do**
12: Count the number of '1' in the sequence $\Delta E(x_i : x_{i+\frac{j}{2}+1})$ and record it as c ; Count the number of '1' in the sequence $\Delta E(x_{i+\frac{j}{2}-1} : x_{i+j-1})$ and record it as d ;
13: **if** $c > \frac{j}{2}$ **then**
14: $f[]$.append(i)
15: **end if**
16: **if** $d > \frac{j}{2}$ **then**
17: $l[]$.append(i)
18: **end if**
19: **end for**
20: $i_{first} =$ the index of first element in f , $i_{last} =$ the index of last element in $l + \frac{j}{2}$;
21: **return** i_{first} ; i_{last} .

In the algorithm, channels are automatically selected by an energy threshold $E(th)$. Here, as shown in Table 2, we have listed the parameters, and in the following Sect. 5.1, we will discuss the performance evaluation.

4.3 Data normalization

The data of fall action collected by the radar is preprocessed to generate a standardized data set. The standardized data set is divided into a training set and a test set.

4.4 Deep convolutional neural network

In recent years, convolutional neural networks (CNN) have shined in the field of image processing and image recognition with their particularly outstanding performances. The internal convolution calculation of the CNN can automatically extract the features of the data so that the

feature selection no longer requires a lot of time and effort and dramatically improves the recognition accuracy.

The main proposal of this work is to fuse the frequency- and time-domain images as inputs for fall detection and recognition from the radar signal. Table 3 depicts the network architecture setup.

Since our sample size is small, we design a relatively low-depth network and reduce the number of parameters in each layer of the network. To prevent over-fitting, our proposed method introduces two measures:

- **Add L2 regularization Term**—In deep learning, small samples for deep networks easily cause over-fitting;

adding regularization to the network is a way to solve over-fitting. Therefore, we add the L2 regularization term to the network to prevent over-fitting of the model results.

Table 2 Parameters for channel selection

Parameters	Value	Description
Mean	0.411	Mean energy of channels
Max	1.827	Maximum energy of channels
Min	0.065	Minimum energy of channels
Std	0.386	Standard deviation of the energy of channels
Median	0.217	Median energy of the channels
Mode	0.064	Energy mode of the channel
Var	0.149	Variance energy of the channel
Q3 value	0.653	The larger value of the quartile
Q1 value	0.138	The smaller of the quartile

Table 3 Parameter setup of our deep CNN

Layer	Description	Activation	Learnable parameters
Image Input	Image input with zero center normalization	200*400*2	–
Convolution	64*5*5 Padding 2	200*400*64	Weights 5*5*2*64, Bias 1*1*64
Dropout	drop 0.2	200*400*64	–
ReLu	ReLu	200*400*64	–
MaxPooling	3*3 Stride 2	100*200*64	–
Batch Normalization	Normalization	100*200*64	Offset 1*1*64, Scale 1*1*64
Convolution	256*3*3 Padding 1	100*200*256	Weights 3*3*64*256, Bias 1*1*256
ReLu	ReLu	100*200*256	–
MaxPooling	3*3 Stride 2	50*100*256	–
Convolution	256*3*3 Padding 1	50*100*256	Weights 3*3*256*256, Bias 1*1*256
ReLu	ReLu	50*100*256	–
MaxPooling	3*3 Stride 2	25*100*256	–
Full Connected	256 full connect	1*1*256	Weights 256*320000 Bias 256*1
Dropout	drop 0.2	1*1*256	–
ReLu	ReLu	1*1*256	–
Full Connected	3 full connect	1*1*3	Weights 3*256, Bias 3*1
Softmax	Softmax	1*1*3	–

- **Introducing dropout layer**—In the deep learning network, the more complex the model, the more parameters it needs to learn. Therefore, a dropout layer is introduced to randomly reduce parameters by 20% to eliminate less important information. This allows the model to obtain good results in the training set, and makes the model easier to generalize and improves robustness.

In the deep neural network training process, the network usually sends each batch to train to have a different distribution. In addition, the data distribution will also change during the training process. It brings difficulties to the next layer of network learning. Therefore, batch normalization is to force the data back to the normal distribution with a mean of 0 and a variance of 1, making the data distribution consistent and avoiding gradient disappearance.

5 Experiment result and discussion

In our process, the proposed model automatically learns a large number of parameters by extracting characteristics of the fall signal. Finally, these parameters will be tested with the test set to verify the learning effect of the model.

For ensuring the independence of data distribution, the test set and the train set are divided randomly at a ratio of 1:1, and both parts are independent of each other. Five experiments are repeated, and each experiment needs to re-divide the train set and the test set.

5.1 Threshold parameter selection of the adaptive algorithm

The Adaptive Algorithm requests a threshold to divide the background and the activity. As shown in Table 2, we considered different parameters to evaluate its performance in recognizing various fall activities. The values fall into the range of $0.0648 \rightarrow 1.827$, which belongs to one of the input images.

We selected the *mean* as a measure of central tendency, and calculate $\frac{1}{4}k, k \in (1, 2, \dots, 7)$. Figure 6 shows that the value of the threshold is less than *mean*, the recognition accuracy, recall, and precision are slowly increasing. When the threshold value reaches the mean value, the recognition accuracy, recall, and precision reach the peak. Above the *mean* value, the recognition accuracy decreased quickly; therefore, we can avoid selecting parameters values greater than 0.411. On the left of Fig. 6, the data values are very close, i.e., the accuracy values of $\frac{1}{2}$ *mean*, $\frac{3}{4}$ *mean*, *Std*, and *mean* are 94.7, 95.2, 94.3, and 94.92, respectively. And, from Table 4, it can be observed that choosing $\frac{3}{4}$ *mean* as the filtering channel threshold produces the best effect. However, the actual difference between the results obtained by $\frac{3}{4}$ *mean* with the adjacent parameters is very small. Later, as the threshold becomes higher and higher, the classification effect will drop sharply; in other words, using the *mean* value as a threshold for screening channels is already effective enough to eliminate interference and

Fig. 6 Performance with different thresholds

filter out worthless channels. To facilitate calculations, therefore, we here selected using *mean* as the threshold.

5.2 Performance evaluation of the training and loss

The loss function reflects degree of convergence of the model in the deep learning network. When the loss function converges to a small value and no more extended changes, it indicates that the model has converged. In this paper, cross-entropy is used as the loss function. The model accuracy reaches 100%, and the loss function drops to about 0.1513 on the train set. Since the loss function fluctuates slightly near a smaller value, the model has converged. The training process and loss function changes are shown in Fig. 7. Therefore, we consider the loss function changes to set up our network parameters to the right epoch. In this paper, we set up our epoch as 10.

5.3 Performance evaluation of fusion features

Tables 5 and 6 show the frequency-domain and the time-domain feature used to recognize the fall activities. As we can observe, the maximum accuracy is obtained by using a single feature of FFT images in five times of testing is 91.3%, and the average accuracy is 90.64%. It is much similar as the case of a single feature of SVD images; the maximum accuracy is 91.4% with an average accuracy of 90.46%. The FFT or SVD image features provide almost

the same effect for classifying the fall activities in our experiments.

While we adopted FFT and SVD images as a fused feature, as shown in Table 7, the maximum accuracy raised to 95.7% with an average accuracy of 94.92%. As a consequence, we can conclude that each feature may contain information that the other does not possess; therefore, combining two different features is a better choice than using a single feature.

5.4 Comparison with other algorithms

It can be seen from Table 8 that the performance of other algorithms is not as good as the proposed method using the same inputs. With traditional machine learning techniques (i.e. kNN, SVM, Naive Bayes, AdaBoost, and Random Forest), the maximum accuracy we could achieve was 92.6% with the SVM classifier.

The above results show that the proposed adaptive channel selection algorithm with the deep neural network is more suitable for fall detection using the UWB radar sensor.

5.5 Discussion

The purpose of our study is to detect and identify the fall event, so as to provide medical services as soon as possible. However, falling is a dangerous situation for people of any

Table 4 Parameter selection with different values

Parameters	$\frac{1}{4}$ mean	$\frac{1}{2}$ mean	$\frac{3}{4}$ mean	mean	$\frac{5}{4}$ mean	$\frac{3}{2}$ mean	$\frac{7}{4}$ mean
Accuracy (%)	93.7	94.7	95.2	94.92	86.3	84.7	67.3
Recall (%)	93.9	95.1	94.7	95.21	81.6	83.4	69.1
Precision (%)	87.7	89.8	90.1	89.46	85.7	86.2	63.2
F1 score	0.907	0.923	0.923	0.918	0.836	0.848	0.660
AUC	0.989	0.990	0.998	0.991	0.894	0.873	0.688

Fig. 7 Model training process and loss function changes

Table 5 Performance of the proposed algorithm using FFT image feature

Number	1	2	3	4	5	Average
Accuracy (%)	91.3	90.4	89.9	90.7	90.9	90.64
Recall (%)	89.3	90.2	89.7	90.5	90.3	90.00
Precision (%)	84.5	83.41	81.2	82.7	84.3	83.17
F1 score	0.869	0.867	0.852	0.864	0.872	0.865
AUC	0.922	0.913	0.902	0.914	0.917	0.914

Table 6 Performance of the proposed algorithm using SVD image feature

Number	1	2	3	4	5	Average
Accuracy (%)	89.7	91.4	89.3	90.7	91.2	90.46
Recall (%)	89.3	91.4	89.1	90.7	91.3	90.36
Precision (%)	83.7	84.6	81.0	83.5	82.3	83.03
F1 score	0.864	0.877	0.849	0.870	0.866	0.865
AUC	0.913	0.920	0.901	0.912	0.924	0.914

Table 7 Performance of the proposed algorithm using fused features

Number	1	2	3	4	5	Average
Accuracy (%)	95.3	94.4	94.9	95.7	94.3	94.92
Recall (%)	95.3	95.2	94.9	95.7	94.5	95.12
Precision (%)	91.2	88.9	87.2	91.7	88.3	89.46
F1 score	0.932	0.919	0.909	0.937	0.913	0.918
AUC	0.996	0.980	0.980	0.998	0.993	0.991

age. Therefore, the data we used in the research phase are all from the simulation of falls performed by young people in the laboratory.

The experiment took place in a big office with a relatively complex environment, differently from the usual laboratory with absorbing walls. This is because the actual application of this monitoring method will be used in environments with objects such as furniture, plants, etc. For simplifying the experiment, the radar was just set on a desk with a distance of 1.5m to the subject.

During the experiment, we found that the characteristics of the radar feedback images obtained are also very different due to the different types of actions. Therefore, we

Table 8 Comparison with other machine learning algorithms

Algorithm	kNN	SVM	NB	AdaBoost	RF	Proposed
Accuracy (%)	88.7	92.6	71.4	90.3	89.3	95.7
Recall (%)	83.3	91.2	74.9	89.9	89.5	95.7
Precision (%)	76.8	88.9	77.2	91.1	90.4	91.7
F1 score	0.799	0.900	0.760	0.905	0.899	0.919
AUC	0.808	0.936	0.756	0.920	0.903	0.990

KNN - k-nearest neighbors; SVM - support vector machine; NB - naive bayes

AdaBoost - Adaptive Boosting; RF - Random Forest

only detect and recognize three common falls in this article. The three falling behaviors we proposed correspond to three situations, namely:

1. Tripping by obstacles;
2. Bending over to pick up things;
3. Squat down and lacing up a shoe.

As we can observe from Fig 3, although their radar feedback chart looks different, it is hard to distinguish them by naked eyes. At the same time, we also found that the power changes in the channels are closely related to human activities. Therefore, we used an adaptive channel selection method to obtain the useful channels. Experiments have proved that the recognition accuracy of inputting complete channels related to the activity is higher than that obtained by a single channel input. Since the human body is a whole, the distance change between each part of the human body to the radar is related to different actions. Simultaneously, channels correspond to the distance of radar to subject. Therefore, the relationship between channels (corresponding in the area of subject exist) reflects the relationship of motion status between different parts of the human body. This is the reason for using an adaptive selector to find out the edge channel of the activity.

6 Conclusion

This paper proposed an adaptive channel selection algorithm to reduce the data dimensions and used fusion features images of FFT and SVD via deep neural networks. By calculating the radar signals' energy change, a threshold is utilized to adaptively select the area of signals that is most likely to be the fall activity. Through the miniaturization of the network and the optimal configuration of parameters, the network can be adapted to small sample data to detect and identify three types of falls, i.e., stand to fall, bow to

fall, and squat to fall. Results showed that using selected channels with fused features compared to single input frequency- and time-domain data will significantly increase the recognition accuracy from 90.64% and 90.46% to 95.7%, respectively.

In future work, we plan to conduct more experiments on expanding sample size; we will also define an algorithm that can detect and recognize activities in multi-residential environments.

Acknowledgements The authors express greatest gratitude to all the volunteers who participated in the experiments.

Funding This research was supported by the National Natural Science Foundation of China (Grant Number 11801542, U1913210), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDB 38040200), the Shenzhen Science and Technology Projects (Grant No. JCYJ20170818163445670 and JCYJ20180703145002040), and Guangdong Province Science and Technology Project (2017A040405027). Ping Wang, Qimeng Li: These authors contributed equally to this work.

Declarations

Conflict of Interest The authors declared that they have no conflicts of interest to this work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. United Nations (2019) World Population Prospects 2019, [Online]. Available: <https://www.un.org/development/desa/pd/node/1114>
2. World Health Organization (2018) WHO Global Report on Falls Prevention in Older Age
3. Mubashir M, Shao L, Seed L (2013) A survey on fall detection: Principles and approaches. *Neurocomputing* 100:144–152, special issue: Behaviours in video. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0925231212003153>
4. Li Q, Gravina R, Li Y, Alsamhi SH, Sun F, Fortino G (2020) Multi-user activity recognition: challenges and opportunities. *Inf Fusion* 63:121–135
5. Wang X, Ellul J, Azzopardi G (2020) Elderly fall detection systems: a literature survey. *Front Robot AI*, vol. 7(71)
6. Gibson RM, Amira A, Ramzan N, de-la Higuera PC, Pervez Z (2016) Multiple comparator classifier framework for accelerometer-based fall detection and diagnostic. *Appl Soft Comput*,

- 39:94–103, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1568494615007061>
7. Zhu C, Sheng W (2011) Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living. *IEEE Trans Syst Man Cybern Part A Syst Humans* 41(3):569–573
 8. Mukhopadhyay SC (2015) Wearable sensors for human activity monitoring: A review. *IEEE Sens J* 15(3):1321–1330
 9. Cippitelli E, Fioranelli F, Gambi E, Spinsante S (2017) Radar and RGB-depth sensors for fall detection: a review. *IEEE Sens J* 17(12):3585–3604
 10. Wu D, Pigou L, Kindermans P-J, Le ND-H, Shao L, Dambre J, Odobez J-M (2016) Deep dynamic neural networks for multi-modal gesture segmentation and recognition. *IEEE Trans Pattern Anal Mach Intell* 38(8):1583–1597
 11. Chaccour K, Darazi R, El Hassani AH, Andrès E (2017) From fall detection to fall prevention: a generic classification of fall-related systems. *IEEE Sens J* 17(3):812–822
 12. Li Z, Li W, Lv H, Zhang Y, Jing X, Wang J (2013) A novel method for respiration-like clutter cancellation in life detection by dual-frequency IR-UWB radar. *IEEE Trans Microw Theory Tech* 61(5):2086–2092
 13. Xin L, Qiao D, Ye L, Dai H (2013) A novel through-wall respiration detection algorithm using uwb radar. In: *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*
 14. Lu N, Wu Y, Feng L, Song J (2019) Deep learning for fall detection: three-dimensional CNN combined with LSTM on video kinematic data. *IEEE J Biomed Health Inf* 23(1):314–323
 15. Rougier C, Meunier J, St-Arnaud A, Rousseau J (2011) Robust video surveillance for fall detection based on human shape deformation. *IEEE Trans Circuits Syst Video Technol* 21(5):611–622
 16. Wang X, Jia K (2020) Human fall detection algorithm based on yolov3. In: *2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC)*, pp 50–54
 17. Enayati M, Banerjee T, Popescu M, Skubic M, Rantz M (2014) A novel web-based depth video rewind approach toward fall preventive interventions in hospitals. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp 4511–4514
 18. Harris A, True H, Hu Z, Cho J, Fell N, Sartipi M (2016) Fall recognition using wearable technologies and machine learning algorithms. In: *IEEE International Conference on Big Data (Big Data) 2016*:3974–3976
 19. Desai K, Mane P, Dsilva M, Zare A, Shingala P, Ambawade D (2020) A novel machine learning based wearable belt for fall detection. In: *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, pp 502–505
 20. Ballı S, Sagbaşı E. A, Korukoglu S (2018) Design of smartwatch-assisted fall detection system via smartphone. In: *2018 26th Signal Processing and Communications Applications Conference (SIU)*, pp 1–4
 21. Zhao S, Li W, Niu W, Gravina R, Fortino G (2018) Recognition of human fall events based on single tri-axial gyroscope. In: *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, pp 1–6
 22. de Araujo IL, Dourado L, Fernandes L, Andrade RMDC, Aguilar PAC (2018) An algorithm for fall detection using data from smartwatch. In: *2018 13th Annual Conference on System of Systems Engineering (SoSE)*, pp 124–131
 23. Youngkong P, Panpanyatep W (2021) A novel double pressure sensors-based monitoring and alarming system for fall detection. In: *2021 Second International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP)*, pp 1–5
 24. Ogawa Y, Naito K (2020) Fall detection scheme based on temperature distribution with ir array sensor. In: *IEEE International Conference on Consumer Electronics (ICCE) 2020*:1–5
 25. Miawarni H, Sardjono TA, Setijadi E, Arraziqi WD, Gumelar AB, Purnomo MH (2020) Fall detection system for elderly based on 2d lidar: a preliminary study of fall incident and activities of daily living (ADL) detection. In: *2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, pp 1–5
 26. Li H, Shrestha A, Heidari H, Le Kernec J, Fioranelli F (2020) Bi-LSTM network for multimodal continuous human activity recognition and fall detection. *IEEE Sens J* 20(3):1191–1201
 27. Li H, Le Kernec J, Mehul A, Gurbuz SZ, Fioranelli F (2020) Distributed radar information fusion for gait recognition and fall detection. In: *2020 IEEE Radar Conference (RadarConf20)*, pp 1–6
 28. Sadreazami H, Mitra D, Bolic M, Rajan S (2020) Compressed domain contactless fall incident detection using uwb radar signals. In: *2020 18th IEEE International New Circuits and Systems Conference (NEWCAS)*, pp 90–93
 29. Khawaja W, Koohifar F, Guvenc I (2017) Uwb radar based beyond wall sensing and tracking for ambient assisted living. In: *2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC)*, pp 142–147
 30. Wang B, Guo L, Zhang H, Guo Y-X (2020) A millimetre-wave radar-based fall detection method using line kernel convolutional neural network. *IEEE Sens J* 20(22):13364–13370
 31. Nezirovic A, Yarovsky AG, Ligthart LP (2010) Signal processing for improved detection of trapped victims using UWB radar. *IEEE Trans Geosci Remote Sens* 48(4):2005–2014

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.