Discrete Human Activity Recognition and Fall Detection by Combining FMCW RADAR Data of Heterogeneous Environments for Independent Assistive Living

Umer Saeed 1,*, Syed Yaseen Shah 2, Syed Aziz Shah 1, Jawad Ahmad 3, Abdullah Alhumaidi Alotaibi 4, Turke Althobaiti 5, Naeem Ramzan 6, Akram Alomainy 7 and Qammer H. Abbasi 8

Abstract: Human activity monitoring is essential for a variety of applications in many fields, particularly healthcare. The goal of this research work is to develop a system that can effectively detect fall/collapse and classify other discrete daily living activities such as sitting, standing, walking, drinking, and bending. For this paper, a publicly accessible dataset is employed, which is captured at various geographical locations using a 5.8 GHz Frequency-Modulated Continuous-Wave (FMCW) RADAR. A total of ninety-nine participants, including young and elderly individuals, took part in the experimental campaign. During data acquisition, each aforementioned activity was recorded for 5–10 s. Through the obtained data, we generated the micro-doppler signatures using short-time Fourier transform by exploiting MATLAB tools. Subsequently, the micro-doppler signatures are validated, trained, and tested using a state-of-the-art deep learning algorithm called Residual Neural Network or ResNet. The ResNet classifier is developed in Python, which is utilised to classify six distinct human activities in this study. Furthermore, the metrics used to analyse the trained model’s performance are precision, recall, F1-score, classification accuracy, and confusion matrix. To test the resilience of the proposed method, two separate experiments are carried out. The trained ResNet models are put to the test by subject-independent scenarios and unseen data of the above-mentioned human activities at diverse geographical spaces. The experimental results showed that ResNet detected the falling and rest of the daily living human activities with decent accuracy.

Keywords: radio-frequency; FMCW RADAR; next generation healthcare; contactless monitoring; fall detection; deep learning; ResNet

1. Introduction

Radio Detection and Ranging (RADAR)-based technologies are often associated with military and defence applications, including detecting and tracking of planes and ships. While travelling, many of us have probably seen massive RADARs near airports’ runways spinning to monitor the surrounding area for incoming and departing aircraft [1]. Nevertheless, in recent years, RADAR-based technology has attracted the attention of a wide-
range of disciplines outside of military and air traffic control [2]. Automotive RADAR, for instance, is a relatively recent use of RADAR sensing technology that aids cars in navigating around other modes of transportation and objects [3–5]. Recently, RADAR-based technology has been proposed in the healthcare field to track everyday living activities at home and to monitor patients’ vital signs, including breathing rate and heart rate [6–8]. RADAR technology is also used in the human gesture recognition system to monitor and identify complex movements made by individuals to interact with objects without pressing buttons or touching screens [9–11].

RADAR is no longer just of interest to a limited group of specialists and users but also to a wide variety of students, researchers, and businesses [12]. RADAR sensing spans a broad range of abilities and disciplines, from the design of frequency-dependent microprocessors and components to electromagnetic wave propagation to RADAR signal processing through smart Artificial Intelligence (AI) techniques [13]. As a consequence, researchers will very likely be required to deal with aspects of RADAR sensing as part of the design and development of a bigger system, whether it is smart vehicles, smart homes, smart phones, smart factory, or an array of RADAR sensors for future healthcare applications [14,15].

This article focuses on the healthcare application of RADAR systems, which is one of the most novel and distinctive domains from the conventional defence-oriented technologies. The integration of RADAR sensing and other technological innovations into healthcare services is attributable to the changing social and personal healthcare requirements linked to the global ageing population. The World Health Organisation (WHO) and the United Nations recently projected that by the year 2050, over a third of the world’s population will be the elderly (age 65 and above), and according to the United Kingdom’s Office for National Statistics, the country’s population over the age of 85 is expected to nearly double in the next two decades [16,17]. As a result, due to ageing, people are more likely to develop several chronic health conditions (also known as multimorbidity) and crucial life threatening events such as collapsing or falls [18–20].

Due to the emergence of AI and advanced sensing techniques [21–25], there has been a lot of interest in utilising cutting-edge techniques to offer effective healthcare in private home settings [26–28]. There are two primary aims involved in this. Firstly, maintaining the privacy and autonomy of individuals in their own familiar surroundings as well as preventing hospitalisations and the disruption of established routines and daily habits, which are critical towards individuals well-being. Secondly, encouraging a proactive approach to healthcare in which technology can offer ongoing reliable monitoring systems and early detection of subtle signs linked to deteriorating health problems rather than responding only when there are severe health conditions.

A question that commonly arises is what can RADAR technology provide in the healthcare domain? Recent research regarding RADAR-based approaches has proved that it is a reliable approach for detecting the presence of individuals, tracking their movements in a certain space, and characterisation of several body movements, which may range from large-scale movements (such as limbs or head) to small-scale movements such as motions of the chest and abdomen while breathing [29–31]. In regard to healthcare, research involving RADAR technology has mainly evolved in two ways. Firstly, utilising RADAR systems and acquired data to estimate and monitor vital signs such as heart rate. Secondly, observing daily activity routines and examining the regularities in it. The second point involves the certainty that individuals practice everyday living activities such as food intake, preparation, or personal hygiene. Apart from this, detecting abnormalities in the regular daily pattern and predicting any possible risk, in particular, a fall, which requires an immediate response.

In this paper, we have utilised a Frequency-Modulated Continuous-Wave (FMCW) RADAR system with Micro-Doppler (MD) signatures to capture fall/collapse especially for elderly people [32,33]. Other than this, to discriminate from fall activity, we also captured daily living human activities such as sitting, standing, walking, drinking, and
bending. This research article offers a deep learning-based approach to timely recognise the aforementioned recurring everyday human activities. Figure 1 presents a complete framework of the proposed fall recognition scheme.

![Figure 1](image)

**Figure 1.** Framework of the proposed contactless system model for fall detection.

The remainder of this paper is organised as follows: Section 2 presents the state-of-the-art literature on contactless technology for healthcare domains. The information about the proposed system model is given in Section 3. The experimental findings are presented in Section 4, and lastly, Section 5 provides closing remarks as well as future research studies.
2. Literature Review

This section examines state-of-the-art research in different types of human activity recognition as well as applications of Machine Learning (ML). According to the articles [34,35], a variety of human activities were observed and recorded while the participants wore wearable sensors such as an accelerometer. Once the data were gathered via distinct activities, it was treated with cutting-edge ML algorithms, for instance, Support Vector Machine (SVM), k-Nearest Neighbors (K-NN), and an ensemble learning approach called Random Forest (RF). The SVM classifier produced the leading results with a score of 91.5 per cent.

The research of [36,37] employed the FMCW RADAR scheme to acquire data on fall and related activities such as running, walking, jumping, stepping, and squatting from multiple subjects. The temporal changes, cross-sections, and Doppler were measured using the FMCW RADAR system. Subsequently, the data were processed by a cross-validation technique over the K-NN algorithm, yielding a 95.5 per cent accuracy score. This research reveals how the frequency variations in wireless signals can be utilised to identify distinct human activities. A similar study on the extraction of a multi-channel was published in [38,39].

Authors in [40] utilised a standard dataset of fourteen indoor activities. The data were gathered with the use of a triaxial accelerometer sensor. Distinguishing static and dynamic activities were among the objectives of the study. After data wrangling, the RF algorithm was used to carry out a classification task. The static outcomes received a higher accuracy score of 92.16 per cent, whereas the dynamic actions received a score of 80.0 per cent and an average result of 85.17 per cent.

The study in [41] identified five distinct arm motions using Channel State Information (CSI) on Wi-Fi Orthogonal Frequency-Division Multiplexing (OFDM) signals. Different arm motions were performed by participants while standing between a computer and a Wi-Fi router transmitting wireless signals. The CSI was then recorded and the data were trained through the deep learning algorithm. The Long Short-Term Memory approach was selected since it achieved an accuracy rate of 96 per cent. Authors in [42–44] performed a similar study in terms of healthcare applications.

In [45], the authors utilised the CSI information to identify a particular individual. Different subjects walked through two devices while the data in the form of CSI were sent and stored. Throughout the experiment, the CSI data were acquired when an individual walked across the radio-frequency signals. Afterwards, the data were fed into the ML classifiers: Decision Trees (DT) and RF. The research study discovered that when only two individuals participated in the binary classification task, the algorithms performed better.

In [46], wearable smart watches were utilised to track the motions of table tennis players. The smart watch collected data on how the participants moved the table tennis racket in eight distinct motions, for instance, forehanded flick, forehanded attack, backhanded flick, and so forth. Following that, seven ML algorithms were employed to analyse the data, including DT, RF, SVM, and K-NN. With an accuracy score of 97.80 per cent, the RF approach was determined to be the dominant classifier in this study.

3. System Model

All conventional RADAR signal processing approaches employed for healthcare systems aim to describe the signatures of concern in three primary aspects. First: Time—this exhibits how the subject’s position and location have changed over time. Second: Range—this exhibits the physical space between the target and the RADAR. Third: Velocity—this determines the large/small motions of the target by assessing the induced frequency change and the Doppler effect. The aforementioned three major characteristics of the RADAR technology are often abbreviated as the “RADAR Cube” [47,48]. With the growing adoption of compact RADAR technology with various receiver channels, which is mainly driven by the automotive industry, a fourth related dimension has recently been added:
The angle of arrival or angular orientation. This can be deduced by comparing the received signal by various receiving channels.

3.1. Frequency-Modulated Continuous-Wave (FMCW) RADAR

The FMCW RADAR is a type of sensor that, such as a basic Continuous-Wave (CW) RADAR, emits continuous transmission power. Unlike CW RADAR, FMCW RADAR may vary its operational frequency during observations, which means that the broadcast signal is modulated in phase or in frequency. RADAR observations via runtime measures are only theoretically feasible due to shifts in frequency. Basic CW RADAR systems without frequency modulation have the drawback of being unable to estimate target range since they lack the timing mark required to precisely clock the transmit and receive cycle and transform this to range. This kind of time reference for measuring the distance between stationary objects may be created by modulating the broadcast signal’s frequency. In this technique, a signal is sent that periodically rises or lowers in frequency. When an echoed signal is obtained, the shift in frequency is delayed by $\Delta t$, by a change in runtime, which is referred to as the pulse RADAR approach. Nevertheless, with pulse RADAR, the runtime must be determined directly. Instead, frequency discrepancies between the emitted and received signals are assessed in FMCW RADAR [49].

The following points are the fundamental characteristics of FMCW RADAR. First, the ability to assess both concurrently, the range of the target and its relative velocity. Second, safe against the lack of high-power pulse radiation. Third, targeting ability that can assess minor ranges, compared to the transmitted wavelengths, the minimum measured range is equivalent. Fourth, the signal processing is performed at a lower frequency following mixing, which makes the attainment of processing circuits more simple. Fifth, high precision of range assessment.

The structure of the FMCW RADAR system is illustrated in Figure 2a. This system comprises of a waveform-generator (WG), voltage-controlled oscillator (VCO), transmission antenna (Tx), frequency-mixer (FM), receiving antenna (Rx), low pass filter (LPF), analogue to digital converter (ADC), and digital signal processor (DSP). The WG produces the FMCW RADAR signal with the baseband, which varies frequency over time as presented in Figure 2b. A single waveform here is also known as the “Chirp Signal” in one period.
Figure 2. FMCW RADAR (a) structure (b) signal produced from the waveform-generator [50].

In this setup, the baseband signal transmitted over the ADC can be stated as

$$x[n] = \sum_{l=1}^{L} a_l \exp \left[ j 2\pi \left\{ \frac{2\Delta F}{c \Delta T} d_l + \frac{2f_c}{c} v_l \right\} n T_s + \frac{2f_c}{c} d_l \right]$$

(1)

where \( l \) \((l = 1, 2, 3, ..., L)\) denotes the object’s index and \( a_l \) is the baseband signal’s amplitude. \( v_l \) stands for the \( l \)th object’s relative velocity, and \( d_l \) is the distance to the \( l \)th object. As shown in Figure 2b, \( \Delta T \) is each chirp’s sweep time and \( \Delta F \) represents the bandwidth. Lastly, a single baseband chirp signal is modelled \( N \) times in the ADC with a \( T_s \) sampling period [51].

In the frequency-domain, we can acquire a baseband signal by employing the “Fourier Transform” to the Equation (1), which can be stated as

$$X[k] = \sum_{n=0}^{N-1} x[n] \exp \left( -j 2\pi \frac{n k}{N} \right)$$

(2)

where \( k \) \((k = 0, 1, 2, ..., K-1)\) denotes the index’s frequency. The spectrogram can be obtained by collecting the frequency-domain baseband signal, which depicts the variation in the range of an object with respect to time. Stating alternatively, the outcome in the frequency-domain in the FMCW RADAR scheme can be regarded as that in the range domain. The gathered spectrogram of the baseband signal over \( N_p \) intervals can be stated as

$$X^{(N_p)} = [X^{(1)}, X^{(2)}, ..., X^{(N_p)}]$$

(3)

where \( X^{(i)} = [X^{(i)}[0], X^{(i)}[1], ..., X^{(i)}[K-1]]^T \) denotes the baseband signal’s frequency-domain as a vector, and \( i \) signifies the interval’s index.

Moreover, Figure 3 depicts the FMCW RADAR operating principle, in which the radio-frequency signals are sent and received when it comes into contact with any target within the range. Every motion of the human body generates a unique MD signature that can be used to differentiate between diverse daily activities.
3.2. Residual Neural Network (ResNet) for Classification

In the past, ML-based approaches have been effectively used on a variety of classification problems [52–56]. We used a deep learning-based scheme called ResNet in this work to identify different human activities and detect falling using the generated spectrograms. The usage of skip connections is required to train such a deep neural network. The input used to feed a layer is also used to feed the output of a layer higher up the stack. The goal of training a deep neural network is to get it to model a target function \( h(x) \). If the network’s output is linked to the input, for example, summing a skip connection, the network would be constrained to model \( f(x) = h(x) - x \) instead of \( h(x) \) [57]. This procedure is characterised as residual learning. Moreover, for training the ResNet algorithm, the optimal parameters were obtained utilising a grid search cross-validation technique with \( CV = 5 \). This technique utilises the principles of fit and score to obtain suitable hyperparameters for training the ML models. The procured primary hyperparameters used in this study are provided as follows.

- \( \text{epochs} = 15 \) (experiment 1) / 50 (experiment 2)
- \( \text{activation} = \text{relu} \)
- \( \text{optimizer} = \text{adam} \)
- \( \text{loss} = \text{categorical} - \text{crossentropy} \)

4. Experimental Results and Discussion

4.1. Data Collection

Every body motion creates a distinct pattern on the MD signatures that can be utilised to distinguish between various sorts of human body motions. When the FMCW RADAR encounters any activity or motion such as falling, the radio-frequency signal is broadcast and received within the range. In this paper, we have employed the available datasets from the newly completed project titled “Intelligent Radio-Frequency Sensing for Falls and Health Prediction (INSHEP)” supported by “Engineering and Physical Sciences Research Council (EPSRC)” [58]. http://researchdata.gla.ac.uk/848/ (accessed on 13 June 2021).

Figure 4 shows various spaces/rooms where experiments were conducted to acquire data. As shown in the figure, the datasets were obtained using the FMCW RADAR operating at 5.8 GHz C-band over 400 MHz bandwidth and +18 dBm output power. An antenna with a gain of +17 dBm was attached to the FMCW RADAR. A total of 99 participants, ranging in age from 21 to 99 years, participated in the whole experimental study in different spaces. The data of walking activity were collected for 10 s, and for rest
of the activities, the data were recorded for 5 s. Each activity was repeated three times. Lastly, MD signatures were generated from the acquired data utilising “Short-Time Fourier Transform”.

Figure 4. Diverse spaces used to acquire data on various human activities by employing participants of distinct ages [58].

In this paper, we have performed simulations on six distinct human activities, which were recorded by FMCW RADAR technology in various locations. The recorded activities are: (1) falling on a safe surface, (2) sitting on a chair, (3) standing from a chair, (4) walking back and forth, (5) drinking from a cup while standing, and (5) bending/leaning down to pick up an item. To gather data, each person was instructed to repeat the same activity three times. Following that, using the acquired data, spectrograms of each activity were generated. The samples of spectrograms are shown in Figure 5.

Figure 5. Samples of the acquired micro-Doppler signatures of six different activities.
4.2. Experimental Results

The ResNet method utilised in this study to classify six various human activities was developed in Python, with the TensorFlow and NumPy libraries being used extensively. The performances of the trained models were evaluated by using the metrics: precision, recall, F1-score, classification accuracy, confusion matrix, and model accuracy/loss against number of epochs. Classification accuracy in this research study can be expressed as the fraction of human activities accurately recognised to the total number of human activities.

In this research work, two different experimental studies were performed using the acquired MD signatures of different human activities. In the first experiment, exclusively space 1 data are used for simulations, while in the second experiment, data of various spaces are merged to perform simulations. However, both experiments were carried out by subject-independent scenario, which means that the training and testing datasets are from dissimilar participants. This will reveal the generalisation ability of the ResNet classifier, which is significant for ML-based approaches. Moreover, the reason behind merging data of different spaces in the second experiment is to train an environment-aware ML model that is reliable and more robust. In total, 1026 MD signatures were used to conduct both experimental studies. These experiments are discussed in detail as follows.

4.2.1. Experiment 1

This study comprises six human activities, and to conduct the first experiment, $6 \times 40 \,(\approx 66\%)$ MD signatures were used for training and $6 \times 20 \,(\approx 34\%)$ for testing. Overall, 360 MD signatures were employed in the first experiment. Due to the extent of the acquired dataset, the number of epochs to train the model was set at 15. The performance was assessed using several metrics typically used for ML once the model is trained. Figure 6a presents a confusion matrix of the first experiment. As can be noted, most of the activities including falling have no misclassifications except bending and standing, which have three and two misclassified points, respectively. Figure 6b presents the ResNet classifier accuracy, whereas Figure 6c presents loss against the number of epochs. As the number of epochs enhanced, the ML model was able to achieve an accuracy rate between 0.9 and 1.0 and model loss was recorded between 0.1 and 0.3. Moreover, an entire classification report of the ResNet algorithm is exhibited in Table 1. As can be noted under Experiment 1, the majority of the activity classes unveiled a precision, recall, and F1-score of 100%, which led to procuring an overall accuracy up to 96%.

Table 1. Experiments 1 and 2 classification report of six distinct activities trained by residual neural network.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
</tr>
<tr>
<td>Bending</td>
<td>94%</td>
<td>85%</td>
</tr>
<tr>
<td>Drinking</td>
<td>87%</td>
<td>100%</td>
</tr>
<tr>
<td>Falling</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Sitting</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>Standing</td>
<td>100%</td>
<td>90%</td>
</tr>
<tr>
<td>Walking</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>96%</td>
<td>85%</td>
</tr>
</tbody>
</table>
Figure 6. Experiment 1 results: (a) confusion matrix, (b) ResNet model accuracy, and (c) loss versus number of epochs.

4.2.2. Experiment 2

The second experiment was conducted to further test the robustness of the ResNet classifier. As a result, the MD signatures employed for the first experiment of space 1 were integrated randomly with the MD signatures of various spaces. In this experiment, $6 \times 71$
MD signatures were used for training and $6 \times 40$ ($\approx 36\%$) for testing. Overall, 666 MD signatures were exploited. As the extent of the dataset enhanced for the second experiment, the number of epochs to train the model was increased from 15 to 50. Figure 7a exhibits a confusion matrix of the second experiment. As can be observed, the targeted activity (falling) was detected by the classifier with 100% accuracy, whereas the rest of the activities have few misclassifications. This is due to the fact that in several spaces, data were arbitrarily merged in order to construct an intricate ML model. Furthermore, Figure 7b shows the ResNet model accuracy, and Figure 7c reveals the model’s loss. With several epochs, the model accuracy and loss were recorded consistently between 0.9 and 1.0 and 0.1 and 0.2, respectively. Lastly, Table 1 lists a complete classification report of all the activities trained through the ResNet algorithm. As can be seen in Experiment 2, the falling activity class unveiled a precision rate of 100%, recall 90%, and F1-score 95%. Meanwhile, the walking activity class attained a precision, recall, and F1-score of 100% in both experiments. However, an overall accuracy dropped from 96% to 85% since the rest of the activities, such as bending, drinking, sitting, and standing, revealed slightly lower percentages. The reasons behind this are that these classes have highly similar data points between each other, and the second experiment was signified to be more intricate.

(a)

(b)

Figure 7. Cont.
Figure 7. Experiment 2 results: (a) confusion matrix, (b) ResNet model accuracy, and (c) loss versus number of epochs.

5. Conclusions and Future Work

This research aims to exploit an existing Frequency-Modulated Continuous-Wave (FMCW) RADAR dataset to detect/monitor activities of daily life. The data were collected in various spaces for 99 participants (young and elderly). In this paper, we presented preliminary findings for a scheme that utilises the FMCW RADAR to recognise several human activities, including falling, sitting, standing, walking, drinking, and bending. Participants were asked to perform the six aforementioned daily living activities in various geographic locations as part of the experimental investigation. The micro-Doppler signatures from the RADAR system were employed as primary data and subsequently validated, trained, and tested using ResNet. The ResNet is a prominent deep learning technique that assists in minimising the overfitting/underfitting problem with a unique skip connection approach. Two different experiments were conducted in this work to investigate the robustness of the applied scheme. The trained deep learning models were tested on the unseen data of discrete human activities. The unseen data were intended to be subject-independent as well as geographically diverse. The experimental findings revealed that ResNet detected falling activity with 100% accuracy in both experiments. Moreover, in experiment 1 and experiment 2, the overall accuracy rate was 96% and 85%, respectively. In experiment 2 findings, the accuracy dropped slightly due to the additional complexity of entirely unseen data of various geographical spaces along with subject-independent instances. Nonetheless, in future work, we aim to experiment with different cutting-edge wireless sensing techniques such as software-defined radio as well as advanced deep learning algorithms such as generative adversarial networks to enhance the performance of the proposed scheme.


Funding: This work was supported in part by the Taif University, Taif, Saudi Arabia, through the Taif University Research Grant under Project TURSP-2020/277.

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


