Data portability for activities of daily living and fall detection in different environments using radar micro-doppler

Shah, SA, Tahir, A, Le Kernec, J, Zoha, A & Fioranelli, F
Published PDF deposited in Coventry University's Repository

ORIGINAL ARTICLE

Data portability for activities of daily living and fall detection in different environments using radar micro-doppler

Syed Aziz Shah[1] · Ahsen Tahir[2] · Julien Le Kernec[3] · Ahmed Zoha[3] · Francesco Fioranelli[3]

Abstract

The health status of an older or vulnerable person can be determined by looking into the additive effects of aging as well as any associated diseases. This status can lead the person to a situation of 'unstable incapacity' for normal aging and is determined by the decrease in response to the environment and to specific pathologies with apparent decrease of independence in activities of daily living (ADL). In this paper, we use micro-Doppler images obtained using a frequency-modulated continuous wave radar (FMCW) operating at 5.8 GHz with 400 MHz bandwidth as the sensor to perform assessment of this health status. The core idea is to develop a generalized system where the data obtained for ADL can be portable across different environments and groups of subjects, and critical events such as falls in mature individuals can be detected. In this context, we have conducted comprehensive experimental campaigns at nine different locations including four laboratory environments and five elderly care homes. A total of 99 subjects participated in the experiments where 1453 micro-Doppler signatures were recorded for six activities. Different machine learning, deep learning algorithms and transfer learning technique were used to classify the ADL. The support vector machine (SVM), K-nearest neighbor (KNN) and convolutional neural network (CNN) provided adequate classification accuracies for particular scenarios; however, the autoencoder neural network outperformed the mentioned classifiers by providing classification accuracy of $\sim 88\%$. The proposed system for fall detection in elderly people can be deployed in care centers and is application for any indoor settings with various age group of people. For future work, we would focus on monitoring multiple older adults, concurrently in indoor settings using continuous radar sensor data stream which is limitation of the present system.

Keywords Radar sensing · Fall detection · Activity of daily living · Data classification

1 Introduction

Aging is associated with changes in dynamic biological, environmental, psychological, behavioral, and social processes and is measured by the functional abilities of the person. The health status of an aged or mature person can be determined by looking into the additive effects of aging as well as the associated diseases. This status can lead the mature individual to a situation of 'unstable incapacity' for normal aging and is determined by the decrease in response to the environment and to specific pathologies with apparent decrease of integration and independence in activities of daily living (ADL). Aging affects metabolic, sensory, cardiovascular, respiratory systems, and most noticeable is its effect on the musculoskeletal system for mobility and locomotion [20], a general psychological, physical and functional weakening that ultimately result in increasing the risk of critical accidents such as falls.

Falls are categorized among one of the most hazardous events that a mature person can encounter. According to the reports published by World Health Organization [34], falls are the second leading cause of accidental or unintentional injury deaths worldwide. Each year an estimated 646,000 individuals die from falls globally of which over 80% are in low- and middle-income countries [34]. Fall

✉ Syed Aziz Shah
  syed.shah@coventry.ac.uk

[1] Research Centre for Intelligent Healthcare, Coventry University, Coventry, UK

[2] University of Engineering and Technology, Lahore, Pakistan

[3] James Watt School of Engineering, University of Glasgow, Glasgow, UK

incidences are inevitable but could be managed through timely detection/intervention and prevented as well, if the person under observation is properly supported. These critical and life-threatening events require both technical and human interventions known as the 'gerontechnology.' The outcome of this field is directed toward the autonomy and improvement quality of life of an elderly person.

Timely and reliable detection of fall events is an important strategy to handle the issue and has been studied extensively over a past decade. The advent of wearable sensors, micro-electro-mechanical systems (MEMS)-based miniature inertial sensors (accelerometer, magnetometer and gyroscope) and non-wearable sensors such as radar and Wi-Fi have facilitated this rapid development. The sensors are shrinking in terms of size and weight such that they can be unobtrusively deployed on people's body [16, 26]. Most of the existing systems identify the forward/backward acceleration caused by the human body toward the ground. Brown et al. [3] took only the large acceleration toward the ground into account when detecting falls in mature individuals, and this is considered as one of the first studies on the specific topic. In order to improve the performance of fall detection system, other researchers used data fusion method using complex machine learning algorithms. For instance, the authors in [9, 10] combined a gyroscope and tri-axial accelerometer to detect human activities and fall event. This system involves a three-step process including multiple data acquisition sources such as audio, images and an accelerometer. A large number of researchers have used environmentally mounted sensors in indoor settings to identify falls, such as sensors integrated in floors as in [32], infrared sensors [35], acoustic sensors [6] [11] and camera-based systems [39].

This paper presents an unobtrusive method for a generalized fall detection system using data portability in conjunction with machine learning and deep learning algorithms in elderly care homes using a be frequency-modulated continuous wave (FMCW) radar sensor. The proposed system aims to be independent of age group involved and geometrical location, and can be easily deployed in any indoor settings by importing a trained model in any location for fall detection. The aspect of data portability is emphasized in this work, aiming to develop classification algorithms capable of identifying falls and ADL patterns in different environments and with different groups of subjects, i.e., enabling the system (data processing and machine learning) developed in one environment to work effectively in another one, with different people.

A two-stage method is used to detect falls and ADL in mature people. Initially, the data were collected at nine locations (four laboratory environments and five elderly care centers) using a lightweight FMCW radar. In second stage we used machine learning algorithms (SVM and KNN), deep learning algorithms (CNN and auto-encoder) and transfer learning techniques to classify successfully the acquired data, in view of future real-time implementation of such techniques.

2 Related work

Most of the researchers have used standard imaging sensors to detect falls. Approaches ranging from using single cameras [39, 39] mounted on ceilings or on walls, to multiple cameras deployed in indoor environment generating a 3-dimensional (3-D) objects [2, 10]. Single camera-based fall detection systems use imaging space feature (bounding box ratios) extracted from silhouettes. Multiple camera-based systems rely on extracting features such as the person's speed from 3-D generated objects using back projecting silhouettes. The camera-based system presents several challenges. First and foremost, extracting foreground features consists of background modeling in red–green–blue (RBG) image space that is a challenging task in the context of real-world conditions considering the issues of shadows and light intensity [24]. Second, fall activity in no/low light can only be detected if an infrared (IR) light source is deployed along the cameras. However, the RGB information can be lost when IR is in operation, posing another challenge for background feature extraction. Third, as far as single camera-based system is concerned, feature extraction that can measure the 3-D movements of person and characterize falls [1] can be a challenge. Fourth, in multiple camera-based systems, deployment and calibration in the same video frame introduce several concerns and increase overall computational cost. Due to these challenges, among others, researchers have turned their attention to Microsoft Kinect depth imaging sensors to recognize ADL. The Kinect sensor incorporates several advanced sensing hardware. Most notably, it contains a depth sensor, a color camera, and a four-microphone array that provide full-body 3D motion capture, facial recognition, and voice recognition capabilities [38]. The authors [39] used Kinect motion sensing using 3D images to detect different human activities and fall event. Wichert et al. [33] mounted Kinect sensor 30 cm above the ground to detect falls. Pre-segmentation was made, and areas were identified where a potential fall could happen. Spatial characteristics of objects were used to determine when fall occurred. To summarize, the camera-based systems can be computationally expensive where dedicated devices have to be deployed and above all raise privacy concerns in private homes and private environments.

Wearable devices-based fall detection system attempts to detect fall events using sensors mounted on watches,

belts or on coats. Some of the most widely used wearable sensors are accelerometers, magnetometers, gyroscopes [15], pressure sensors [22] and RFID [30]. These devices only work provided put on by the subjects when the incidence of fall occurs. Smartphone-based fall detection system is one of the promising devices with huge potential thanks to the popularity of sensor-rich smartphones [28]. Even if these fall detectors are appropriate solutions, the always-on-body requirement of the devices may be a limitation, especially in case of non-compliant users, such as people refusing to or forgetting to wear or carry the sensors, particularly at residential care centers or private homes. Recently, non-contact radio-frequency (RF) sensing technologies for ADL and automatic fall detection have been widely used, as they have the ability to monitor disabled or mature people without deploying any device on the subject's body [25, 27][7]. Some of these sensing technologies include channel state information (CSI) extracted using Wi-Fi signals [36][36] or exploiting micro-Doppler information or range-profile obtained using radar technology. However, the CSI data obtained present two major limitations in the context of human activity recognition and fall detection. Firstly, the CSI data suffer from coarse grain resolution and its susceptibility to noise. Secondly, CSI data obtained at one location are largely different than that of another location, limiting the portability of the system. Hence, the Wi-Fi based systems fail in the design of a generic, invariant fault detector that is independent of geometrical location and age group involved. The micro-Doppler signatures obtained using radar technology present very similar patterns for specific human activities in cross-room setup, i.e., in different rooms or environments. These signatures are related to the kinematics of the movements performed by the subjects and less affected by changes in the EM propagation channel and multipath that may hinder Wi-Fi-based performances, and radar signal processing techniques such as moving target indicator (MTI) filtering can remove the presence of static reflectors and clutter. Hence, radar micro-Doppler signatures are promising to design an invariant, generic, and portable fall detection system such as the one proposed in this paper. In the existing literature, researchers have used classification algorithms on micro-Doppler images [25]. However, this work only involved limited number healthy volunteers with high mobility in five different locations. Most of these healthy participants have similar mobility profile for each activity. Eventually, induced identical micro-Doppler signatures for each activity performed by different healthy participant. In order to make a more robust and reliable fall detection monitoring system using radar sensor, we recruited older adults with fast, medium and slow mobility who performed different activities with varying speed. In addition, we increased the complexity of the experimental campaign to make it more realistic and applicable in real-world settings—a total of nine different locations were considered with different geometrical structure, furniture and location of radar sensor.

3 FMCW radar and data processing

Radar systems use electromagnetic waves to detect objects. The typical radar system comprises a transceiver and a signal processing unit. The FMCW radar in operation continuously transmits RF signals, and any object within range reflects the waves that are received by the receiving antenna.

In an FMCW radar, the instantaneous transmission frequency changes linearly across the waveform, providing a widely adopted solution for low-cost, short to medium range sensing applications [4], including ADL and identification of falls. One of the main advantages of FMCW radar is its robustness against external narrow-band interferences from other sources, low peak power, and capability of recording micro-Doppler signatures for target recognition. Therefore, we have used a light-weight, easily deployable FMCW radar sensor manufactured by Ancortek with the specifications listed in Table 1.

The total number of signals transmitted by an FMCW radar in a data recording can be written as:

$$x(t) = \sum_{i=0}^{N_F-1} x^{(i)}(t - iT_F),\qquad(1)$$

Here T_F is the total duration of a frame and N_F represents the total number of transmitted frames. The transmitted FMCW signal comprising L number of chirps at the ith frame can be written as:

$$x^i(t) = \sum_{t=0}^{L-1} x_0(t - lT),\qquad(2)$$

T in Eq. 2 is the duration of an FMCW chirp signal $x_0(t)$ and is written as follows:

Table 1 C-band ancortek radar parameters

Radar model	SDR 580AD2
Waveform	FMCW
Operating frequency	5.6 to 6.0 GHz
Bandwidth	400 MHz
Sweep time	1 ms
Transmitting power	$\sim +20$ dBm

$$x_0(t) = e^{\left(f_0 t + \frac{\mu}{2} t^2\right) \times 2\pi j} \text{ for } 0 \leq t \leq T \qquad (3)$$

Here f_0 is the operating frequency and μ denotes change in instantaneous frequency of an FMCW chirp signal. The value of μ can be determined by bandwidth (B) divided by duration of a chirp (T).

A. Micro-Doppler Signature

If the target within radar range has mechanical/dynamic vibration or rotation in addition to its bulk translation, it induces a frequency modulation on the returned signal that generates sidebands about the target's Doppler frequency shift, that is called the 'micro-Doppler effect' [17]. Any moving person located in P moves with frequency f_v and displacement D_v, having a displacement function, $D(t) = D_v \sin 2\pi f_v t \cos \beta \cos \alpha_p$. Assuming the value of R_0 be the distance between the radar and the person initial position O, the range function changes with respect to time due to the person's micro-motion, represented as: $R(t) + D(t)$. The RF signal received by the radar can be express as follows:

$$s(t) = \rho\left(e^{\left(2\pi f_o t + 2\pi \frac{R(t)}{\lambda}\right) j}\right) = \rho e^{(2\pi f_o t + \Phi(t)) j} \qquad (4)$$

In Eq. 4, f_0 is the carrier frequency, λ is the wavelength of carrier, and ρ denotes the backscattering coefficient. By substituting $R(t)$ in Eq. 4, the signal received by radar can be written as:

$$s(t) = \rho\left(e^{\left(\frac{4\pi R_0}{\lambda}\right) j}\right) \times e^{j 2\pi f_o t + D_v \sin(w_v t) \cos\beta \cos\alpha_p 4\pi/\lambda} \qquad (5)$$

where $w_v = 2\pi f_v$; from Eq. 5, the derivative of the second phase component gives us the expression of micro-Doppler shift that is written as follows:

$$f_{mD} = \frac{w_v D_v}{\pi\lambda} \cos\beta \cos\alpha_p \cos(w_v t) \qquad (6)$$

To summarize, Eq. 6 shows that the micro-Doppler frequency component is directly proportional to the velocity of the movement/displacement of the specific target's part and to the frequency of the radar signal. Furthermore, there is a dependency on the cosine of the aspect angles in azimuth and elevation, so that only the radial component of the velocity vector contributes to the micro-Doppler signature.

With the introduction of micro-Doppler signatures, the extraction of such information from the reflected radar signal can play a vital role in automatic fall detection system. In the next section, the experimental setup and data acquisition using range of subjects are discussed.

4 Experimental setup and data acquisition

The experimental campaign for this work was conducted in three different organizations, namely University of Glasgow, North Glasgow Housing Association Residential Centre and Age UK West Cumbria Daily Centre. For the purpose of training, validation and evaluation, data were collected in 9 rooms as in Fig. 1 (4 at the University of Glasgow, UK, and 5 at two residential and service centers for older people) with a total of 99 participants (with age range from 21 to 98 years) over the course of 10 days. A total of 56 participants took part in experimental campaign at the care centers, where 31 volunteers out of them were aged $50+$.

The six activities of daily living collected were A1—walking back and forth, A2—sitting down on chair, A3—standing up from a chair, A4—bending down to pick up object, A5—drinking water and A6—falling [25]. Activities such as sitting down on a chair, picking up objects from floor, and falling were deliberately chosen, as they are similar in producing a movement of the body toward the floor, in order to introduce classes that may be challenging to classify. These six activities were selected to include common, day-to-day actions performed by people in their home settings, together with the fall, which is an action requiring to be critically identified. The activities of sitting/standing and bending were selected as they present a net acceleration movement toward the radar (or away from it, depending on the geometry of deployment for the radar and the direction of the movement). Such movement often translates into a clear peak in the Doppler spectrum that may resemble the signature of a fall event in some circumstances and is typically included in the set of activities considered for contactless radar-based activity recognition. These very similar activities put the proposed classification approach to test. A reliable identification of fall events is critical with low missed detection and minimum false alarms because lying for longer periods of time after falls can have severe adverse effects on the affected people. The data for each of the six activities were obtained 3 times for each of the 99 subjects, producing a dataset of 1453 observations in total (note that it was not always possible to collect all repetitions of each activity when mature participants were involved). The younger participants we have assumed someone less than 60 years, anyone over 60 is considered as mature. The data collected in each room are summarized in Table 2. The dataset was acquired offline, indicating the activities were performed in a controlled window size. Each activity had duration of 5 s except the first one which was 10 s (walking). The dataset acquired was balanced one considering the time duration. For future work, the unbalanced dataset and unbalanced time-duration

Fig. 1 Nine locations for human activity recognition and fall detection

Table 2 Details of the data acquired during experimental campaign

Room/Location	ID	Number of Observations	Age group	Environment
1	Dataset 1	360	Younger Participants	Laboratory Environment
2	Dataset 2	48	Younger Participants	Laboratory Environment
3	Dataset 3	162	Younger Participants	Laboratory Environment
4	Dataset 4	288	Younger Participants	Laboratory Environment
5	Dataset 5	141	Mature Participants	Old Age Home
6	Dataset 6	105	Mature Participants	Old Age Home
7	Dataset 7	60	Mature Participants	Old Age Home
8	Dataset 8	184	Mature Participants	Old Age Home
9	Dataset 9	105	Mature Participants	Old Age Home

will be considered. The aim of this work was to develop a generalized system independent of geometric location, mobility of elderly people and age limit.

The Ancortek radar transmitted RF signals with 400 MHz bandwidth at 1 kHz pulse repetition frequency (PRF). The transmitted power was approximately + 20dBm; two Yagi antennas were used (one as transmitter, one as receiver as in Fig. 1 of work published in Shah [25] with gain equal to 17 dB. The radar operated in C-band with signal centered at 5.8 GHz. The radar was powered by USB, with its power consumption limited to USB standards. Moreover, upon assessing the lifetime and autonomy of the radar system in realistic deployment scenarios, it would always be connected to a desktop or laptop computer for operational and data acquisition purposes, or to the electricity mains in indoor scenarios.

As far as human activities are concerned, the FMCW radar provides both range and Doppler information. However, in this work, we have mainly focused on the Doppler information since falls and other ADL can be detected merely using this information as well. For data collection, the radar sensor was placed on wooden table in all nine

locations where participants were asked to perform activities within a short range of 1-m to about 3-m range. The two antennas were placed in such a way that it would allow us to keep the torso of the participants in the center of the beam in order to maximize the received signal strength. The data recorded using the FMCW radar sensor were processed using short time Fourier transform (STFT) to characterize spectrograms and generate micro-Doppler signatures. The range-time-intensity plots can initially be obtained by stacking the received radar signals in matrix form and applying fast Fourier transform (FFT) algorithm along the fast-time direction to generate range profiles. STFT is then applied on the range cells containing the signatures of the targets, in this case people moving, to characterize their micro-Doppler signatures. STFT applies a sequence of FFTs with short, overlapping windows along the total duration of the recorded data; the square absolute value of the complex result is the so-called spectrogram, a plot of velocities of moving body parts (measured through the Doppler effect) as a function of time. A notch MTI filter is applied to eliminate the contribution of static targets near 0 Hz such as furniture, walls, ceiling, and floor.

Figure 2a presents examples of micro-Doppler signatures for six human activities, namely walking, sitting down on a chair, standing up, picking up an object, drinking water while standing, and a fall. The positive values of Doppler components correspond with the movements toward the radar sensor, while any movement away from radar generates negative Doppler values. This is evident in Fig. 2 where the person performs walking activity: the main contribution (in red) comes from torso as the subject moves back and forth in front of radar, resulting in alternate values between positive and negative. Examining the fall activity, a strong forward acceleration toward ground is observed. Note that in this figure the spectrograms for each activity were normalized as the distance between radar sensor and subject varies.

Figure 2b shows the spectrograms for various subjects with different age groups retrieved at nine various locations, namely four different rooms in University of Glasgow, three at North Glasgow Housing Association Residential Centre and two at Age UK West Cumbria Daily Centre. Depending on their physical characteristics, some of the participants could move fast while few had limited ability to walk. For instance, the bottom left spectrogram in Fig. 2b was the slowest to move back and forth. We have extracted the average body speed in Fig. 3, as the center of mass of the spectrogram. This can be a proxy for the overall mobility of the subjects, as reduced mobility and risk of falling are usually correlated; plus this shows that our radar-based system can not only measure the velocity of different body parts, but also present the average velocity of the movement.

5 Data classification

Deep learning is a subset of machine learning that has recently experienced very significant growth thanks to increased computation power provided by state-of-the-art GPUs and rapid advancements in algorithms. Deep neural networks are built on artificial neural network concepts, using multiple layers of multiple neurons that increase the overall size of the network and the complexity of the nonlinear input–output relationship. Each neuron in deep network is implemented by linearly convolving multiple inputs plugged into an activation function. Previously, hyperbolic or sigmoid functions were used as activation function; however, the so-called vanishing gradient problem was posing limitation on network size and its proper training. A gradient descent technique is used to train the neural network that helps in minimizing the loss function of the network during the backpropagation process. However, during backpropagation process, the error rate gradually decreases as the input data flow through each layer step by step, resulting in slow training with increase number of layers. This problem was resolved using rectified linear units, commonly known as ReLU working as activation functions. The ReLU primarily has zero output for negative input values and an output of positive for positive input values. The ReLU enables a sparse representation of the data when the network is randomly initialized. In addition, it also drastically reduces the vanishing gradient problem by only considering values between 0 and 1. Consequently, the ReLU has opened doors for the design of state-of-the-art deep networks producing incredible results by classifying massive datasets. The applications of CNN became famous when an eight-layer architecture known as *AlexNet* [21] won Visual Recognition Competition in 2012. Later on, a 16 layer deep
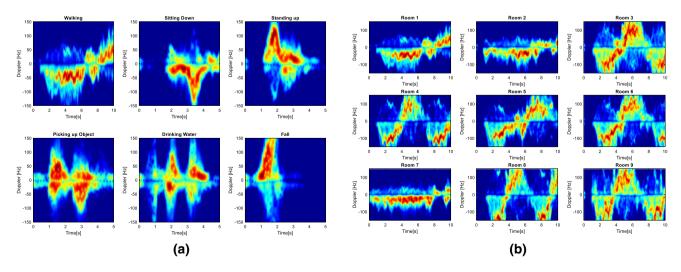

Fig. 2 Spectrograms obtained for human activities. **a** Six human activities, **b** Walking activity in nine different rooms by different participants
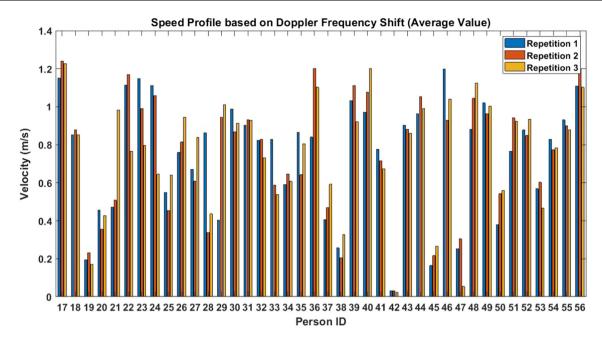
Fig. 3 Speed profile of volunteers extracted from micro-Doppler signatures

network, namely *VGG-Net* [29], and 152 layer network architecture, *ResNet* [8], won the same challenge. Presently, the research work on deep learning algorithms includes processing and classification of millions of images into thousands of classes. This has led academics and researchers within radar and healthcare community to experiment with deep networks to classify datasets obtained from RF signals. However, one of the biggest challenges in applying deep network algorithms to RF signals classification is the availability of limited datasets. The RF data collection is relatively difficult, costly to perform, and time consuming with respect to collecting optical images, especially in the applications of radar systems for monitoring and surveillance purposes. It is impractical and almost impossible to obtain million or micro-Doppler signatures for human activity recognition,- subsequently, those algorithms will be facing overfitting problems by training highly complex deep network architectures using small datasets.

Transfer learning has been proposed to overcome this problem, whereby the deep networks are pre-trained using millions of optical images and then simply fine-tuned with the few available radar data at a subsequent training stage. This approach was shown to outperform random initialization of the weights and biases of the network. In this work, we evaluate and compare different classification approaches in terms of their portability, i.e., performances when the training and testing datasets were collected in different environments with different subjects. In particular, we have CNN to classify directly spectrograms of ADL and falls, compared with other classifiers such as SVM,

KNN, autoencoder-based networks and transfer learning techniques.

A. Transfer Learning

Transfer learning is a pivotal tool in machine learning and deep learning to classify complex problems where insufficient data for training are available. It tries to transfer the knowledge from one trained model to the target model by extracting features with the former and training with latter. This leads to a significant positive impact on the problems where limited datasets are available, such as in our case and in many automatic target classification problems based on radar data.

The architecture of transfer learning is shown in Fig. 4. Transfer learning primarily allows differences between task, distributions of the training and target domains. This
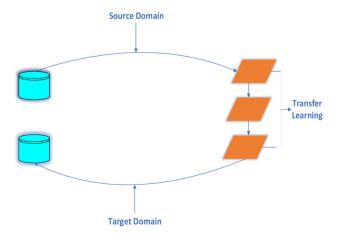

Fig. 4 Transfer learning architecture model

implies that the available training and test datasets might use various distributions such that *P(x)* may follow distinct labeling function *P(y|x)* which might have different set of features for different classes. In the context of transfer learning, the datasets that use similar distributions having same labels with the same feature sets are known as *source*. The core idea of using transfer learning is to learn machine learning algorithms for classifications based on the target datasets that primarily benefits from the existing datasets originated from various sources, for instance, available data that consist of similar patterns but do not have to be exact representative of the target datasets.

We present one example of transfer classifier that uses this same-and different-distribution training data for the neural network part, all followed by SVM and KNN classifiers. This means reusing the weights in one or more layers from a pre-trained network model in a new model, and either keeping the weights fixed, fine tuning them, or adapting the weights entirely when training the model.

B. Autoencoder

Another classifier we have used for classification and then performance comparison is an autoencoder—that is existing classification algorithm previously used in Zhou et al. [37]. It is a feed-forward artificial neural network that reconstructs the input data at the output side of the network under specific circumstances. For instance, for any input dataset x, the autoencoder aims to estimate $h_w(x) = x$. The pre-training algorithm for weight initialization and biases of the autoencoder that provides high performance for small number of available datasets is used. An autoencoder uses unsupervised pre-training by encoding and decoding the input values.

Consider an input vector, x, the encoder maps the input values in a nonlinear manner as follows:

$$e_i = \sigma(Wx_i + b) \tag{10}$$

where σ indicates a nonlinear activation function, W represents weights, and b denotes the biases. The feature vector that is encoded is decoded in next step to reconstruct the specific input values as:

$$z_i = \sigma(W^\sim e_i + b^\sim) \tag{11}$$

In Eq. 11, $W^\sim$ and $b^\sim$ indicate the weights and biases of the autoencoder at decoding side. When unsupervised pre-training process, the neural network minimizes the reconstruction error by tuning weights and biases such that $\theta = b, b^\sim, W, W^\sim$

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N}(x_i - z_i)^2 \tag{12}$$

In order to restrict the neural network, a sparsity parameter is used along cost function that consequently pushes the network to learn the correlation between input values. The cost function, after adding sparsity parameters, thus becomes as follows:

$$\arg\min_\theta J(\theta) = \frac{1}{N}\sum_{i=1}^{N}(x_i - z_i)^2 + \beta\sum_{i=1}^{N}KL(p||p_l) \tag{13}$$

Here β is the sparsity proportion and *KL* represents Kullback–Leibler divergence. The Kullback–Leibler between Bernoulli random values with mean p and p_j is written as:

$$KL(p||p_l) = p\log\left(\frac{p}{p_j}\right) + (1-p)\log\left(\frac{1-p}{1-p_j}\right) \tag{14}$$

Here p_j is the activation value for jth hidden neuron and p is the average activation value, h is the total number of hidden neurons in neural network.

The Kullback–Leibler divergence also known as relative entropy is an exceptional instance of a more extensive divergence and is an asymmetric information theoretic measure of the separation between two likelihood density functions. It is an approximation of how a specific dissemination differs from another, normal likelihood appropriation. Kullback–Leibler divergence has a considerable measure of ongoing applications, specifically in machine learning for healthcare sector. The proposed fall detection system using radar sensing is micro-Doppler signature analysis, using short time Fourier transform and examining it with the help of Kullback–Leibler divergence and autoencoder classifier.

In most applications of data pre-processing and data classification using deep learning algorithms such as autoencoder, we choose the classifier's specific parameters that minimize the mean square approximation error. The same least squares approach has been used in the classical deep neural classification algorithms. However, for deep learning, it turns out that an alternative idea works better minimizing the Kullback–Leibler (KL) divergence. The use of KL divergence is justified if we predict probabilities, but the use of this divergence has been successful in other situations as well. In this paper, we use it for this empirical success. Namely, the least square approach is optimal when the approximation error is normally distributed and can lead to wrong results when the actual distribution is distinct from normal. The need to have a robust criterion, i.e., a criterion that does not depend on the corresponding distribution, naturally leads to the KL divergence.

After pre-training process, the decoder is eliminated from the neural network and encoder values were trained using supervised learning technique by introducing Soft-Max classifier having 6 neurons following the encoder, as we have 6 activities in our case. The SoftMax function

takes as input a vector of K real numbers and normalizes it into a probability distribution consisting of K probabilities, $P(y_k||x_i)$ for $k = 1,2,...K$, proportional to the exponentials of the input numbers [12]. The probability function where input value x_i corresponds to labels class y_k is approximated. The probability class p_k can be mathematically written as:

$$P\left(y = k|x_i = \frac{e^{\theta_k x_i}}{\sum_{k=1}^{K} e^{\theta_k x_i}}\right) \tag{15}$$

The weight and biases, θ, can be optimized by minimizing the cost function as follows:

$$J(\theta) = \sum_{i=1}^{N} \sum_{k=1}^{K} 1\{y^i = k\} log \frac{e^{\theta_k x_i}}{\sum_{k=1}^{K} e^{\theta_k x_i}} \tag{16}$$

Equation 16 is solved using gradient-based technique and is called 'fine-tuning' function where neural network is trained using supervised machine learning algorithm. The autoencoder architecture using spectrograms as images is shown in Fig. 5.

6 Results and discussion

This section discusses the results and analysis of data collected. The classification methods are divided into four main parts. Firstly, two conventional machine learning classifiers, namely support vector machine (SVM) and K-nearest neighbor (KNN), were used to classify activities. Secondly, transfer learning technique was used to extract features with one CNN (AlexNet) and train a machine learning/deep network model on the same features for classification. Thirdly, a convolution neural network was trained from scratch and testing was performed on the specific model. Lastly, an autoencoder neural network was used for training and testing. In all cases, the focus was on using diverse available data in terms of environments and subjects at the training and testing stage.

6.1 Classification results and discussion

This section presents the classification results obtained using different methods. Each of the method is discussed in detail as follows:

(1) Classification using machine learning

(a) Training and test on combined datasets

Initially, we combined all available datasets (healthy individuals' datasets from university laboratory environment and mature people datasets from residential and service care centers) and performed classification tasks to obtain some baseline results. Two conventional machine learning classifiers, namely SVM and KNN, were considered for classification of activities. The SVM uses features in order to produce a hyperplane margin that is based on the distribution of set of features for a particular class. This algorithm is already extensively used for human activity recognition in indoor settings and has been compared with other classifiers [23]. The second classifier, i.e., KNN, is a nonparametric technique used for classification tasks. It compares the distance between an input test sample and the k nearest training samples in its features space, performing a majority vote between the closest neighboring points to assign the test sample to a specific class. The training, validation and testing processes were implemented using MATLAB. Different ways of calculating the distance between points or vectors in features spaces can be used, starting from the simplest Euclidean distance.

The datasets obtained using FMCW radar for all nine locations and 1453 observations were divided into 70% (1017 observations) for training and 30% (436 observations) for testing on the basis of per class. The undesired
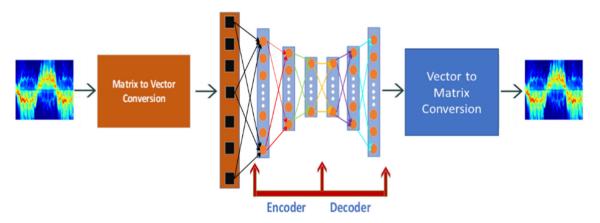

Fig. 5 Three-layer AE, where encoder layers have 200–100–50 neurons and decoder layers have 50–100–200 neurons

biases found in the results are minimized using this deterministic method that would be encountered in imbalance datasets between training and test classes.

Initially, we have selected the optimum features such as mean, root-mean-square, median, skewness, variance, standard deviation and kurtosis of centroid and bandwidth extracted from micro-Doppler signatures for each activity [5].

The KNN classification algorithm with optimized hyperparameters was obtained from estimated objective, distance and total number of neighbor's functions as in Fig. 6. The distance function, namely 'Mahalanobis' and number of nearest neighbors as 10 were used as optimum hyperparameters, where the value of estimated objective function was 0.1971 and estimated function evaluation time was 0.056 s. We have used 'holdout cross-validation' that splits the data into training and test parts where there were no common data points between the two. The Mahalanobis distance is a measure of the distance between a point P and a distribution. It has advantage of using multi-dimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D. This distance is zero for P at the mean of D and grows as P moves away from the mean along each principal component axis. If each of these axes is re-scaled to have unit variance, then the Mahalanobis distance corresponds to standard Euclidean distance in the transformed space. The Mahalanobis distance is thus unitless, scale-invariant, and takes into account the correlations of the dataset.

An example of confusion matrix obtained from the 11th trial for six human activities for all nine locations using
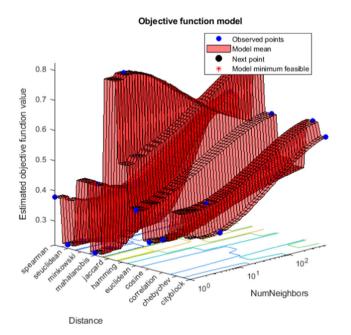

Fig. 6 Optimized k-nearest neighbor fitted classifier

KNN is shown in Table 3. *A*1, *A*2, *A*3, *A*4, *A*5, and *A*6 refer to walking, sitting down, standing up, pick up object. The *test accuracy* in this case is nearly 86% where there was misclassification between picking up object from the ground and pick up glass from the table to drink water. These were similar activities; that is, why the classifier was not able to fully discriminate between the human actions. The average accuracy for all trials was 81.18., drink water and fall event, respectively. The comparable confusion matrix for the KNN is shown in Table 4.

Figure 7 (orange) shows the percentage accuracy of KNN and SVM classification algorithms when training and testing were performed for 30 number of trials. In this scenario, the training and test data (unseen to the classifier) were divided 30 times and both procedures were performed using the same optimized hyperparameters. Table 4 describes the confusion matrix for SVM algorithm. In this case, radial basis kernel function was used as the features were linearly non-separable after they were mapped to a high-dimensional feature space. The maximum accuracy obtained was nearly 80% for trial number 4, and the average accuracy was 77.71% as in Fig. 7 (blue). The misclassification rate between activity 4 (picking up object) and activity 5 (drinking water) was higher than KNN classifier. We have chosen random number of iterations/trials. For instance from iteration number 30 up until 100 and even till 150, the result and graph obtained was almost identical. The only trials number from 1 to 30, we obtained different result as indicated in Figs. 7, 8, 9.

(b)   Training and test on younger people datasets

Figure 8 shows the percentage accuracy for processing data of younger subjects, i.e., employing the data that were obtained inside the University of Glasgow involving mostly students and staff aging 21 to 37 years. A common split technique known as the hold-out method was used where 70% of data for training and the remaining 30% of the data for testing were selected.

Table 5 shows the confusion matrix for KNN classifier. The algorithm was able to classify walking and sitting down activity with no misclassification. On only two instances, standing up was misclassified, and fall event was adequately identified as well. However, the classifier could not properly discriminate between two similar activities such as picking up object from the ground and pick up glass and drink water. The percentage accuracy obtained in this case varied between 76 and 86 with an average accuracy of 83.31 for 30 trials. The SVM classification algorithm performed worse as compared to the KNN, with confusion matrix provided in Table 6. The misidentification rate was much higher; for instance, walking, sitting down, standing up, pick up object drink water and fall were not properly determined 1, 4, 5, 20, 7 and 2 times,
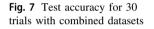
Table 3 Confusion matrix for KNN (combined data, %)

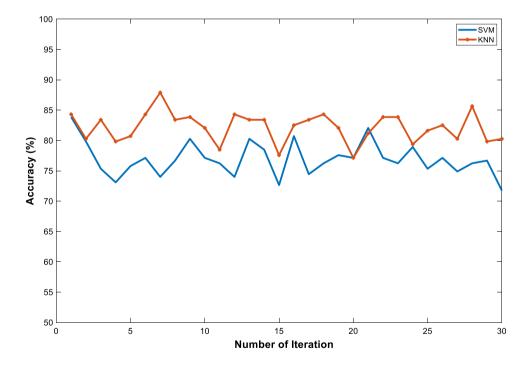| Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)	A6 (%)
A1	**100**	0	0	0	2	0
A2	0	**99**	0	2	4	0
A3	0	0	**97**	2	6	4
A4	4	4	2	**89**	19	0
A5	0	4	9	15	**92**	2
A6	2	0	0	0	4	**61**

The significance of bold indicated correctly classified activities

Table 4 Confusion matrix for SVM (combined data, 100)

Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)	A6 (%)
A1	**50**	0	0	0	1	0
A2	0	**47**	0	1	0	0
A3	0	0	**49**	2	2	0
A4	0	1	1	**25**	9	0
A5	0	2	1	22	**36**	2
A6	0	0	0	0	13	**26**

The significance of bold indicated correctly classified activities

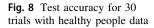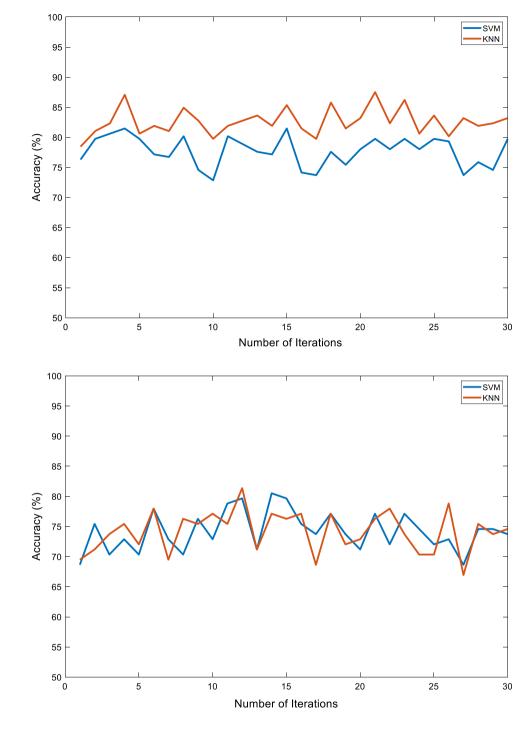**Fig. 7** Test accuracy for 30 trials with combined datasets

respectively, as seen in Table 7 and Fig. 9 (blue). The performance of this classifier fluctuated between 70 and 82% with an average accuracy of 77.57, that decreased by nearly 6% as compared to performance of KNN.

(c) Training and test on mature individuals dataset

The speed at which mature subjects move or perform activities is relatively slower as compared to younger ones as evident from speed profile in Fig. 3. The spectrograms

obtained for the younger individuals are very similar to each other as they move almost in similar fashion, hence there is less misclassification; this is no longer true when data from the older cohort of subjects is considered, where individual mobility issues, health conditions, and presence of walking aids can produce very diverse signatures even for the same activity. This can generate some challenges for the classification algorithm. Figure 9, with Tables 7 and 8, shows the confusion matrices and percentage accuracy

Fig. 8 Test accuracy for 30 trials with healthy people data

Fig. 9 Test accuracy for 30 trials with mature people data

Table 5 Confusion matrix for KNN (younger data)

Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)	A6 (%)
A1	**100**	0	0	0	2	0
A2	0	**100**	0	2	4	0
A3	0	0	**96**	2	6	4
A4	4	4	2	**87**	19	0
A5	0	4	9	15	**91**	2
A6	2	0	0	0	4	**90**

The significance of bold indicated correctly classified activities

Table 6 Confusion matrix for SVM (younger data)

Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)	A6 (%)
A1	**90**	0	0	0	0	0
A2	0	**83**	0	0	0	0
A3	0	0	**80**	0	3	0
A4	0	0	3	**43**	3	0
A5	3	10	10	50	**100**	5
A6	0	0	0	0	13	**82**

The significance of bold indicated correctly classified activities

Table 7 Confusion matrix KNN (mature subjects))

Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)
A1	**100**	0	0	0	0
A2	0	**100**	0	0	0
A3	0	0	**96**	13	9
A4	4	0	4	**39**	26
A5	0	17	4	87	**70**

The significance of bold indicated correctly classified activities

Table 8 Confusion matrix SVM (mature subjects)

Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)
A1	**100**	0	0	0	0
A2	0	**96**	4	0	0
A3	0	0	**96**	13	4
A4	0	0	0	**33**	21
A5	0	0	0	50	**71**

The significance of bold indicated correctly classified activities

for data obtained on mature people at residential and service care homes. The training (70%) and testing (30%) were performed on mature individuals datasets. It should be noted that we have 5 activities, excluding falls that for obvious reasons could not be performed involving mature subjects. The average percentage accuracy obtained using KNN with optimized hyperparameter was 73.27. The SVM classifier performed slightly better (by 1%) with an average accuracy of 74.20 as compared to the KNN.

(2) Results with transfer learning

    (a)   Training on healthy individuals and test on mature people

Transfer learning focuses on exploiting knowledge of one machine learning model from a given classification problem to solve a similar problem but in a different domain. In this work, AlexNet pre-trained on optical images has been used to extract features from the micro-Doppler spectrograms of human ADL, where the spectrograms are used as input images. The deep network can convert the relevant information to discriminate the different classes into features that would not be otherwise easily extracted with simple, handcrafted procedures as the centroid and bandwidth used in previous sections. These new, deeper and less obvious features are then used to train and test the SVM and KNN classifiers.

In this case, a more challenging yet realistic testing approach was implemented, where data for each individual subject were used as the testing set, while all the other subjects' data were used to train the classification algorithm. This 'leave-one-subject-out' test aims to validate the system and classification algorithms in case new, unknown subjects have to be monitored, with the additional challenge of the training data coming from an average much younger cohort of subjects collected in controlled university environment.

The primary aim of using different machine learning classifiers (KNN, SVM, transfer learning, and autoencoder) is to investigate which algorithms would be best fit for data portability, where radar data acquired in different locations can be can trained and testing performed in a separate and different locations. It was concluded that the autoencoder worked best among all the aforementioned in the particular scenario of importing data for generalization purpose.

The accuracies in terms of percentage obtained for each individual are presented in Fig. 10, referring to the data collected at the North Glasgow Housing Association. The person with ID 20 aged 78 had limited ability to move; hence, the challenge in classification accuracy, while person with ID 23, aged 50 had percentage accuracy of nearly 80%. The overall accuracy using transferring learning method for KNN classifier was 62.92. Similar approach was adopted for SVM algorithms as well, that provided the accuracy for each individual and can be seen in Fig. 11. Contrary to KNN, the SVM provided the lowest accuracy for person having ID 23 and 34, while the average accuracy obtained in this case was 56.19%. It should also be noted that three of the individuals (person ID 19, 21, and 29)

Fig. 10 Performance of the mature individuals at North Glasgow Housing Association Residential Centre (KNN) — train on younger people, test on mature ones
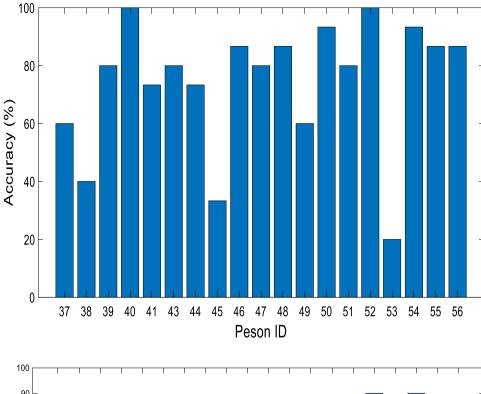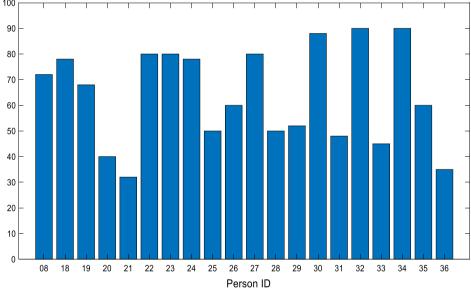

Fig. 11 Performance of the mature individuals at North Glasgow Housing Association Residential Centre (SVM)— train on younger people, test on mature ones
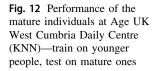
carried walking sticks and one (person ID 20) was using a walker while performed activities.

Figures 12 and 13 show the results obtained for individuals using KNN and SVM, respectively, when the testing subjects' data were collected at the Age UK West Cumbria Daily Centre; the training data were still the same collected at the University of Glasgow with younger subjects. The KNN algorithms classified two persons' activities with 100% accuracy (person ID 40 and 51). The lowest accuracy, less than 20% was recorded for person ID 52. The average accuracy obtained was 74.38%. On the other,

the performance of SVM classifier was for worse than the counterpart, presenting an average accuracy of 57.19%.

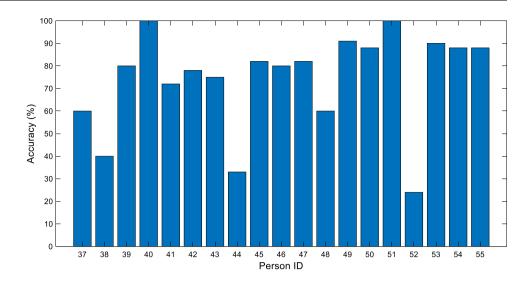(a) Training and test on mature individuals

Contrary to training on healthy people and test on mature people, a great improvement in performance of both classifiers was observed when training and test procedures were performed on mature individuals' datasets. For instance, one dataset of mature people (Age UK West Cumbria) was used for training and another dataset (North Glasgow Homes) for test and vice versa. The KNN and SVM classifiers provided average accuracy of 73.26% and

Fig. 12 Performance of the mature individuals at Age UK West Cumbria Daily Centre (KNN)—train on younger people, test on mature ones
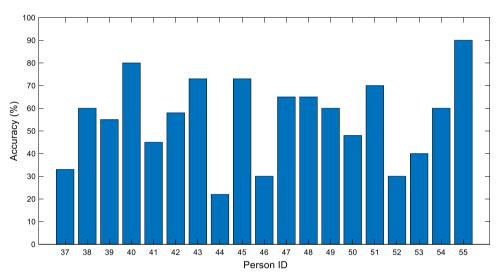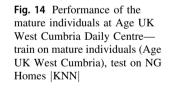

Fig. 13 Performance of the mature individuals at Age UK West Cumbria Daily Centre (SVM)—train on younger people, test on mature one

Fig. 14 Performance of the mature individuals at Age UK West Cumbria Daily Centre—train on mature individuals (Age UK West Cumbria), test on NG Homes |KNN|
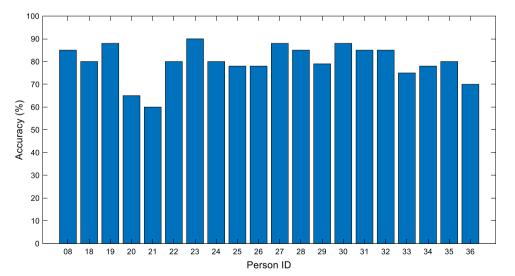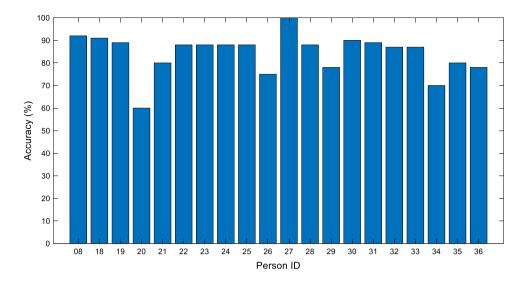
Fig. 15 Performance of the mature individuals at Age UK West Cumbria Daily Centre—train on mature individuals (Age UK West Cumbria), test on NG Homes |SVM|

82.64%, respectively. The improvement in individual classifiers can be seen in Figs. 14 and 15. The average performance of both classification algorithms was almost similar when training was performance on Age UK West Cumbria dataset (mature individuals) and test on NG homes (mature people). The performance of SVM classifier decreased by 8% where both classifier presented average accuracies of 74%.

(3)   Results with convolutional neural network

The convolutional neural network is specially designed to process images (spectrograms in our case) and extract features. The CNNs comprise various layers including image layer, convolutional layer, pooling layer, max pooling, fully connected and output layers. The inputs consist of pixels that are plugged into the input layer where a feature layer is formed that is convolved over the pixels resulting in a convolutional layer. To minimize the total number available features and introduce high correlation between the adjacent pixels, a pooling layer is introduced. Next, a max pooling layer is formed to down sample the elements and to extract the most relevant set of features such as circles and edges. We have used two convolution layers along two max pooling. The 2D convolutional layer

scaled down the original input 582×872 image into 256×256. A rectified linear unit (ReLU) is used to as an activation function. Advantages of rectified linear unit are sparse activation: For example, in a randomly initialized network, only about 50% of hidden units are activated (have a nonzero output). Better gradient propagation: Fewer vanishing gradient problems compared to sigmoidal activation functions that saturate in both directions. Efficient computation: Only comparison, addition and multiplication (Table 9).
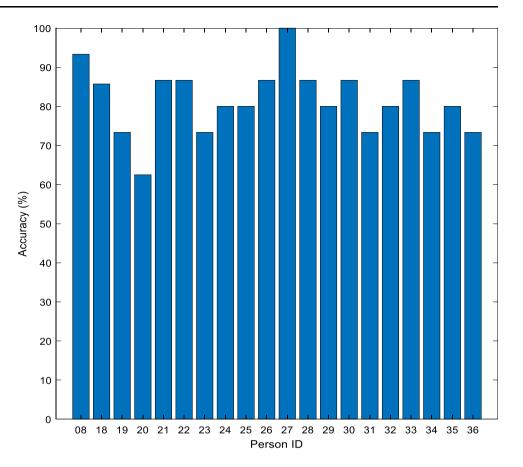
The first convolution layer used 32 size filters with 3 by 3 feature detector matrix. The filter sizer in second convolution layer was increased to 64 with same size of the feature detector. The max pooling layer, right after the first convolution layer used 4 by 4 feature, the one after second convolution layer was reduced to 2 by 2 matrix. The output received from two convolution and two max pooling layers formed a single column matrix that contains all the values and is termed as the flattening layer and is fed into the input to the neural network for classification tasks. The results we have obtained for training and test on mature datasets are shown in Figs. 16 and 17. The average accuracy of 81.41% obtained when training was performed on mature dataset of North Glasgow Homes (mature people data), test on Age UK West Cumbria, while for the test accuracy of North Glasgow Homes of 84.20% was achieved.

Table 10 shows the confusion matrix, when training was performed by Age UK West Cumbria and test was done on North Glasgow homes. The accuracy in this case was 86.2%. On the other hand, when training was performed on North Glasgow homes datasets and test was done on Age UK West Cumbria, an accuracy of 88.3% was obtained.

(4)   Result with autoencoder

This section discusses the classification results using an autoencoder. We have used the fine-tuning algorithm

Table 9 Confusion matrix North Glasgow Homes

Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)
A1	**92**	0	0	0	0
A2	0	**96**	0	0	0
A3	0	0	**84**	20	0
A4	4	8	4	**100**	16
A5	05	4	8	4	**80**

The significance of bold indicated correctly classified activities

Fig. 16 Performance of the mature individuals at on NG Homes—train on mature individuals (NG Homes), test on NG Homes |CNN|

selected cross-entropy instead of mean square error as the loss function. An autoencoder can be piled up hierarchically such that the deepest layers get input values from the output of initial layers. The value of KL divergence sued for sparsity regularization is chosen as 2 and the value of β as 0.1.

The ADAM algorithm, an adaptive moment estimation is used to compute the pretraining and fine tuning as in Kingma and Ba [14] with an initial learning rate of 0.001. We have used grid search to identify the optimal width and depth as in Table 11 and implemented a three-layer autoencoder having 200,100, and 50, respectively. The results show the test results obtained when training was performed on mature individuals at North Glasgow homes and test was performed on Age UK West Cumbria care center. This classifier provided best classification accuracy of 88.1% when width of 40–80-160 and depth of 1 was used. Adam is a replacement optimization algorithm for stochastic gradient descent for training deep learning models. Adam combines the best properties of the CNN, SVM, and KNN to provide an optimization algorithm that can handle sparse gradients on noisy problems. Adam is a method for efficient stochastic optimization that only requires first-order gradients with little memory requirement. The method computes individual adaptive learning rates for different parameters from estimates of first and second moments of the gradients; the name Adam is derived from adaptive moment estimation.

To summarize the performance of training/testing across rooms and environments, we have enlisted all the accuracies for different scenarios in Table 12

7 Conclusion

This paper presented a generalized and portable system to detect fall events in healthy and mature individuals using an FMCW radar exploiting micro-Doppler signatures and images. In this context, an extensive experimental campaign was conducted in nine different locations with 99 volunteers. Machine learning algorithms such as SVM and KNN, transfer learning techniques, CNN and autoencoder were used to classify the data and identify different activities such as walking back and forth, sitting down on chair, standing up from chair, picking up object from the ground, drinking water from a glass and fall events. Different combination of datasets acquired in different environments and with different subjects were used for training and testing, focusing in particular on training on healthy individuals with data collected in university environment, and

Fig. 17 Performance of the mature individuals at Age UK West Cumbria—train on mature individuals (NG Homes), test on Age UK West Cumbria |CNN|
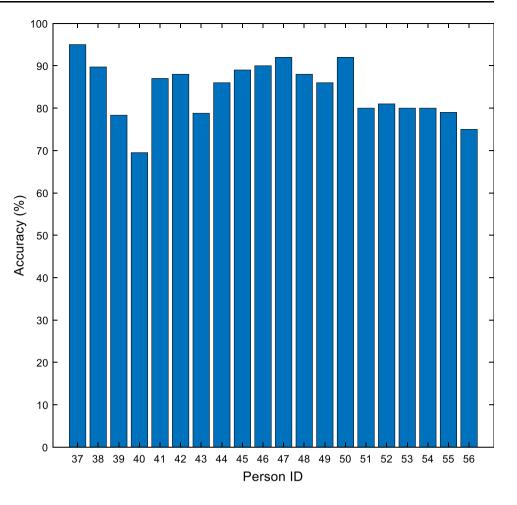
Table 10 Confusion matrix for Age UK West Cumbria

Actual/predicted	A1 (%)	A2 (%)	A3 (%)	A4 (%)	A5 (%)
A1	**100**	0	0	0	0
A2	0	**83**	4	4	4
A3	0	8	**92**	0	4
A4	0	4	4	**88**	13
A5	0	0	0	4	**79**

The significance of bold indicated correctly classified activities

then testing on mature people or vice versa. This was done to develop portable algorithms for ADL classification and fall detection, i.e., algorithms able to maintain good performances when exposed to data and radar signatures for similar activities but performed by different subjects in different environments. Among all classifiers explored in this work, the autoencoder outperformed other classification algorithms by providing an accuracy of 88% for training and test on different data of mature people. The

Table 11 Autoencoder results—training on North Glasgow homes and test on age UK West Cumbria Daily center

No	Width	Depth	Acc. (%)	No	Width	Depth	Acc. (%)
1	20	1	71.2	8	50–100–200	3	82.5
2	50	1	73.0	9	20–50–100–200	3	81.4
3	100	1	75.3	10	40–100–200–400	4	79.2
4	20–50	2	78.0	11	20–50–100–200–400	4	81.3
5	50–100	2	80.2	12	20–50-100–200–400–800	4	82.1
6	20–50-100	2	82.8	13	20–50–100–200–400–800–1600	5	78.2
7	40–80-160	3	88.1	14	40–100–200–400–800–1600	5	80.1

Table 12 Accuracies (%) of different possible combination in cross-room scenarios

Train	Locations	Test	Technique	Accuracy (%)
Combined datasets (healthy + mature people)—70%	1,2,3 and 4	Combined datasets (healthy + mature people)—30%	K-nearest neighbor	81.18
Combined datasets (healthy + mature people)—70%	1,2,3 and 4	Combined datasets (healthy + mature people)—30%	SVM	77.71
Healthy people datasets (University of Glasgow)—70% training	1,2,3 and 4	Healthy people dataset (University of Glasgow)—30% training	KNN	83.31
Healthy people datasets (University of Glasgow)—70% training	5,6,7,8, and 9	Healthy people dataset (University of Glasgow)—30% training	SVM	77.57
Mature people dataset (elderly care homes)—70% training	5,6,7,8, and 9	Mature People Dataset (elderly care homes)—30% training	KNN	73.27
Healthy people datasets (University of Glasgow)—all available data	Train on datasets of 1,2,3 and 4	Test on datasets of 5,6 and 7	Transfer learning—feature extracted using AlexNet—test on KNN	62.92
		Test on datasets of 8 and 9		56.19
Healthy people datasets (University of Glasgow)—all available data	Train on datasets of 1,2,3 and 4	Test on datasets of 5,6 and 7	Transfer Learning—feature extracted using AlexNet—Test on SVM	74.38
		Test on datasets of 8 and 9		57.19
Mature people datasets (location 5, 6, and 7)		Test on datasets of 8 and 9	Transfer learning—feature extracted using AlexNet—test on KNN	73.26
Mature people datasets (location 5, 6, and 7)		Test on datasets of 8 and 9	Transfer Learning—feature extracted using AlexNet—test on SVM	82.64
Mature people datasets (location 8 and 9)		Test on datasets of 5, 6, and 7	Transfer learning—feature extracted using AlexNet—Test on KNN	74
Mature people datasets (location 8 and 9)		Test on datasets of 5,6 and 7	Transfer learning—feature extracted using AlexNet—Test on SVM	74
Mature people datasets (location 5,6 and 7)		Test on datasets of 8 and 9	CNN	81.41
Mature people datasets (location 8 and 9)		Test on datasets of 5,6 and 7	CNN	84.20
Mature people datasets (location 5,6 and 7		Test on datasets of 8 and 9	Autoencoder	86.6
Mature people datasets (location 8 and 9)		Test on datasets of 5, 6, and 7	Autoencoder	88.3

autoencoder is used to reduce the number of feature dimensions under consideration. The main advantage of dimensionality reduction techniques such as autoencoder is to obtain a set of principal variables to improve the performance of the approach as compared to state-of-the-art methods such as SVM, KNN and random forest. The transformation of data from high dimensional to low dimensional can be carried out using autoencoder. This process through densely connected layers extracts the important features in data that can be conveniently utilized for further processes. The high-level representation of data is learned through autoencoder which is normally missed by traditional methods such as SVM, RF and KNN. Other than that, the computational complexity is reduced once the data are transformed to low dimensions. In addition, the autoencoder networks perform better than other classifiers used in this work because they compress large amount of radar data we have obtained from the input layer into a shorter code and then uncompress that code into whatever format best matches the original input. This process learns to encode features for similar activities such as bending down to pick up object and fall activities. The similar human activities are at times misclassified by other algorithms as they only rely on time-domain or frequency-

domain statistical features and do not accurately detect intricate but similar activities. That is why, autoencoder has capability to identify very small difference for similar features/activities as compared to other classifiers.

Declarations

Conflict of interest The authors declare that there is no conflict of interest to report for this paper.

References

1. Akagündüz E et al (2017) Silhouette orientation volumes for efficient fall detection in depth videos. IEEE J Biomed Health Inform 21(3):756–763. https://doi.org/10.1109/JBHI.2016.2570300

2. Auvinet E et al (2011) Fall detection with multiple cameras: an occlusion-resistant method based on 3-D silhouette vertical distribution. IEEE Trans Inf Technol Biomed 15(2):290–300. https://doi.org/10.1109/TITB.2010.2087385

3. Brown G (2005) An accelerometer based fall detector: development, experimentation, and analysis. University of California, Berkeley, pp 1–9

4. Chetty K et al (2017) A low-cost through-the-wall FMCW radar for stand-off operation and activity detection. Radar Sens Technol XXI 10188:1018808. https://doi.org/10.1117/12.2261680

5. Fioranelli F, Ritchie M, Griffiths H (2016) Centroid features for classification of armed/unarmed multiple personnel using multistatic human micro-Doppler. IET Radar Sonar Navig 10(9):120. https://doi.org/10.1049/iet-rsn.2015.0493

6. Guo X et al. (2018) Human activity recognition based on two-dimensional acoustic arrays. In: 2018 IEEE international ultrasonics symposium (IUS), pp 1–4. https://doi.org/10.1109/ULTSYM.2018.8579884.

7. Haider D et al (2019) An efficient monitoring of eclamptic seizures in wireless sensors networks. Comput Electr Eng 75:16–30. https://doi.org/10.1016/J.COMPELECENG.2019.02.011

8. He K et al. (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp 770–778. https://doi.org/10.1109/CVPR.2016.90

9. Hwang JY et al. (2004) Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly. In: The 26th annual international conference of the IEEE engineering in medicine and biology society, pp 2204–2207. https://doi.org/10.1109/IEMBS.2004.1403643

10. Hwang S et al. (2017) Poster abstract: maximizing accuracy of fall detection and alert systems based on 3D convolutional neural network. In: 2017 IEEE/ACM second international conference on internet-of-things design and implementation (IoTDI), pp 343–344

11. Irtaza A et al. (2017) A framework for fall detection of elderly people by analyzing environmental sounds through acoustic local ternary patterns. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC), pp 1558–1563. https://doi.org/10.1109/SMC.2017.8122836

12. Iwata K (2017) Extending the peak bandwidth of parameters for softmax selection in reinforcement learning. IEEE Trans Neural Netw Learn Syst 28(8):1865–1877. https://doi.org/10.1109/TNNLS.2016.2558295

13. Kim B, Kim S, Lee J (2018) A novel DFT-based DOA estimation by a virtual array extension using simple multiplications for FMCW radar. Sensors (Switzerland). https://doi.org/10.3390/s18051560

14. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization, pp 1–15

15. Li H et al. (2017) Multisensor data fusion for human activities classification and fall detection. In: 2017 IEEE SENSORS, pp 1–3. https://doi.org/10.1109/ICSENS.2017.8234179

16. Li H et al (2018) Magnetic and Radar Sensing for Multimodal Remote Health Monitoring. IEEE Sens J. https://doi.org/10.1109/JSEN.2018.2872894

17. Lim H, Park JH (2011) Joint time-frequency analysis of radar engine models. Science 25:1069–1080

18. Liu L et al (2018) Respiration symptoms monitoring in body area networks. Appl Sci 8(4):568. https://doi.org/10.3390/app8040568

19. Lu N et al (2019) Deep learning for fall detection: three-dimensional CNN combined with LSTM on video kinematic data. IEEE J Biomed Health Inform 23(1):314–323. https://doi.org/10.1109/JBHI.2018.2808281

20. Mechling H (2008) Dementia and physical activity. Eur Rev Aging Phys Activity 5(1):1–3. https://doi.org/10.1007/s11556-008-0033-6

21. Monien B, Preis R, Schamberger S (2007) Approximation algorithms for multilevel graph partitioning. In: Handbook of approximation algorithms and Metaheuristics, pp 60-1-60–16. https://doi.org/10.1201/9781420010749

22. Musngi MM et al. (2018) Use of average vertical velocity and difference in altitude for improving automatic fall detection from trunk based inertial and barometric pressure measurements. In: 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 5146–5149. https://doi.org/10.1109/EMBC.2018.8513482

23. Riazul Islam SM et al (2015) The Internet of Things for health care: a comprehensive survey. IEEE Access 3:678–708. https://doi.org/10.1109/access.2015.2437951

24. Roy K et al (2018) Background subtraction using dominant directional pattern. IEEE Access 6:39917–39926. https://doi.org/10.1109/ACCESS.2018.2846749

25. Shah SA, Fioranelli F (xxxx) Human activity recognition: preliminary results for dataset portability using FMCW radar. In: 2019 international radar conference. IEEE, pp 1–4

26. Shah SA et al (2018) Seizure episodes detection via smart medical sensing system. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-018-1142-3

27. Shah SA, Yang X, Abbasi QH (2019) Cognitive health care system and its application in pill-rolling assessment. Int J Numer Modell Electron Netw Dev Fields. https://doi.org/10.1002/jnm.2632

28. Shahzad A, Kim K (2019) FallDroid: an automated smart-phone-based fall detection system using multiple kernel learning. IEEE Trans Ind Inf 15(1):35–44. https://doi.org/10.1109/TII.2018.2839749

29. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. CoRR 1409:1556

30. Toda K, Shinomiya N (2018) Fall detection system for the elderly using RFID tags with sensing capability. In: 2018 IEEE 7th global conference on consumer electronics (GCCE), pp 475–478. https://doi.org/10.1109/GCCE.2018.8574720

31. Wang Z, Yang Z, Dong T (2017) A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time. Sensors (Switzerland). https://doi.org/10.3390/s17020341

32. Werner F et al. (2011) Fall detection with distributed floor-mounted accelerometers: an overview of the development and evaluation of a fall detection system within the project eHome. In: 2011 5th international conference on pervasive computing technologies for healthcare (PervasiveHealth) and workshops, pp 354–361. https://doi.org/10.4108/icst.pervasivehealth.2011.246088

33. Wichert R, Eberhardt B (2012) Ambient assisted living, 5. AAL-kongress. Adv Technol Soc Change

34. World Health Organization (2008) World Health Organization global report on falls prevention in older age, p 47

35. Yang B, Zhang M (2017) Credit-based multiple human location for passive binary pyroelectric infrared sensor tracking system: free from region partition and classifier. IEEE Sens J 17(1):37–45. https://doi.org/10.1109/JSEN.2016.2623325

36. Yang X et al (2019) Diagnosis of the hypopnea syndrome in the early stage. Neural Comput Appl. https://doi.org/10.1007/s00521-019-04037-8

37. Zhou Y, Amimeur A, Jiang C, Dou D, Jin R, Wang P (2018) Density-aware local siamese autoencoder network embedding with autoencoder graph clustering. In: 2018 IEEE international conference on big data (Big Data), Seattle, WA, USA, pp 1162–1167. https://doi.org/10.1109/BigData.2018.8621992

38. Zhang Z (2012) Microsoft kinect sensor and its effect. IEEE Multimedia 19(2):4–10. https://doi.org/10.1109/MMUL.2012.24

39. Zhao F et al (2019) Real-time detection of fall from bed using a single depth camera. IEEE Trans Autom Sci Eng 16(3):1018–1032. https://doi.org/10.1109/TASE.2018.2861382