Radar and RGB-Depth Sensors for Fall Detection: A Review

Enea Cippitelli, Francesco Fioranelli, *Member, IEEE*, Ennio Gambi, *Member, IEEE,* and Susanna Spinsante, *Senior Member, IEEE*

Abstract— **This paper reviews recent works in the literature on the use of systems based on Radar and RGB-Depth sensors for fall detection, and discusses outstanding research challenges and trends related to this research field. Systems to detect reliably fall events and promptly alert carers and first responders have gained significant interest in the past few years in order to address the societal issue of an increasing number of elderly people living alone, with the associated risk of them falling and the consequences in terms of health treatments, reduced well-being, and costs. The interest in radar and RGB-D sensors is related to their capability to enable contactless and non-intrusive monitoring, which is an advantage for practical deployment and users' acceptance and compliance, compared with other sensor technologies such as video-cameras, or wearables. Furthermore, the possibility of combining and fusing information from heterogeneous types of sensors is expected to improve the overall performance of practical fall detection systems. Researchers from different fields can benefit from multidisciplinary knowledge and awareness of the latest developments in radar and RGB-D sensors that this paper is discussing.**

Keywords—**Radar sensors, RGB-D sensors, micro-Doppler, fall detection, human movements analysis, ambient assisting living, feature extraction and classification.**

I. Introduction

T HE PROPORTION of people aged over 65 years is increasing worldwide, with different estimations for different countries projecting the percentage of over 65 to 30% in 2050 in the European Union and in China, and to 20.2% in 2050 in the United States [1]. This aging population is pushing towards a different healthcare delivery model, evolving from the conventional hospital-centric approach where patients are diagnosed and treated for acute conditions in specialized hospitals, to a more home-centric model where care is delivered to the patient in his/her own home as long as possible, supported by the use of new technologies [1]. This home-centric approach to long-term care improves the quality of life of the patients, who are able to live longer in a familiar environment without major changes to their habits, and also reduces the public costs for providing healthcare. In this context where many elderly people live alone at their home, the risk posed by fall events and subsequent injuries is an important issue to tackle. The World

Health Organization defines falls as 'an event which results in a person coming to rest inadvertently on the ground or floor or other lower level, and estimates the proportion of people aged over 65 falling each year to be approximately 28-35%, and 32-42% for people aged over 70 [2]. Fall events lead to immediate physical injuries such as cuts, abrasions, and fractures of bones, as well as to a psychological impact leading to the fear of falling again and, in general, to reduced confidence and diminished level of physical activity [3]. Furthermore, research also reports a reduction in life expectancy for those who experience a long-lie period after the fall event, i.e. an involuntary rest on the ground for an hour or longer (it is for instance estimated in [4] that half of the elderly people who experienced a long-lie following a fall died within 6 months).

To address this issue, in the past few years significant research activity focused on developing solutions for secure and reliable systems to monitor elderly people in their daily activities and promptly detect fall events. These systems would directly benefit elderly people by allowing them to continue having an independent lifestyle, without the need to move to institutionalized care, enabling timely and effective intervention in case of need, and ultimately reducing the emotional and financial burden for the elderly and their families. This has also a clear societal and economical effect by reducing the costs and resources needed to treat the consequences of fall events, especially in case of complications following the long-lie period on the floor.

It should be also noted that reliable monitoring systems can be beneficial not only for fall detection, but also to evaluate the pattern of life of an individual. This includes for instance how active the person is, how often he/she moves in different parts of the house and what activities are performed, in particular fundamental activities (the so-called Activities of Daily Living – ADL) such as food intake and personal hygiene. Irregularities with respect to the normal pattern of life of a person can be used for early detection of deteriorating health conditions (for instance initial symptoms of dementia), providing the opportunity for timely and more effective treatment [5].

Many different technologies have been proposed in the literature for people monitoring and specifically for fall detection [6]–[8]. These include wearable devices such as

E. Cippitelli, E. Gambi, and S. Spinsante are with the Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy (e-mail: e.cippitelli@univpm.it, s.spinsante@staff.univpm.it, e.gambi@staff.univpm.it).

F. Fioranelli is with the School of Engineering, University of Glasgow, Glasgow, UK (e-mail: francesco.fioranelli@glasgow.ac.uk).

accelerometers, gyroscopes, and panic push buttons, inertial sensors such as those within smartphones, infrared, vibration, acoustic, magnetic sensors, video cameras, RGB-Depth (RGB-D) sensors, and radar sensors, or a combination of these systems, whereby their information is used jointly and fused to optimize the overall performance.

The choice of reviewing these two types of sensing technologies (radar and RGB-D) is related to the fact that similar reviews have been conducted for wearable sensors [9]–[13], and video-camera based sensors [7], and so a critical review of alternative technologies can offer opportunities to compare and complement the available results in the literature. Additional reviews surveyed some of these different technologies for fall detection and proposed to classify them in several ways. The work in [14], published in 2012, classified the existing fall detection methods in three categories, namely wearables (tri-axial accelerometers mostly), ambient sensors (e.g. sensing pressure, vibrations, audio), and vision-based sensors (mostly cameras to detect posture and changes in body shape and activities). The 2015 survey on vision-based fall detection systems in [15] distinguished works that used single RGB cameras, multiple RGB cameras, and 3D methods using depth cameras, and mentioned three publicly available datasets recorded using Microsoft Kinect cameras (described later in section III of this manuscript). The work in [16], published in 2015, proposed a slightly different classification of sensors for fall detection, distinguishing between sensors worn by the users and context-aware sensors (i.e. infrared, acoustic, pressure, vibration, camera-based sensors sensing the presence and the activities in a certain area). An important contribution of this paper is the list of works that operate by combining and fusing information from multiple sensors, either belonging to the same category (e.g. two different types of wearables or two different types of context-aware sensors), or mixing wearables with context-aware sensors. Another, more articulate classification of systems for fall detection and fall prevention is provided in the recent work in [17]. Here the different sensors and systems are classified in wearables (further divided into those worn on the body and those worn on feet or shoes), non-wearables (further divided into ambient sensors, vision sensors, and radio-frequency sensors), and fusion or hybrid systems.

None of the aforementioned works surveyed radar-based systems for fall detection and activity monitoring, apart from [17] which mentioned pulse-Doppler radar only, but without discussing existing works in this domain or alternative radar technologies (e.g. frequency modulated techniques as opposed to pulsed systems). Regarding systems based on depth cameras, only reference [15] discussed to some extent the works in the literature, whereas references [16]-[17] only mentioned a few of them and reference [14] only considered normal cameras.

There is therefore scope to propose a joint survey of radar-based and RGB-D sensors for fall detection and daily activities monitoring. These sensors provide attractive advantages compared with other technologies, particularly in terms of privacy preservation and non-cooperative monitoring capabilities. Ambient sensors, especially vision-based sensors, can raise sensitive issues in terms of the confidentiality of the

data and privacy of the patients, which may not be an issue for wearable sensors [16]. Wearables sensors, however, require users' cooperation and compliance to be worn or carried, which could be potentially problematic and uncomfortable as highlighted in the introduction of [15]. Radar and depth sensors can address these issues. For RGB-D sensors, if the data processing algorithm for fall detection relies on depth data only, no direct optical images of the monitored people are collected. RGB images could be used only in case of dangerous events, with the user's agreement. For radar sensors, this is also true as no images of the monitored people are collected. Furthermore, there is also an element of non-stigmatizing the subjects to be monitored and their specific needs, as with these technologies there is no need to alter one's usual behavior because of the introduction of the sensor at home, or to wear unusual devices. All these aspects can help address some of the key users' acceptance issues highlighted for wearables, smartphones, and video-cameras [6], making radar and RGB-D interesting technologies to evaluate in the assisted living context. It has also been highlighted how radar systems are not affected by low or bad lighting conditions as opposed to video-cameras, and both radar and RGB-D systems are more resilient than acoustic systems to water flowing interference and degrading interference by multiple echoes [18]. This can be an advantage for practical deployment of fall detection systems in environments such as toilets and bathrooms, where the risk of falling may be significantly higher because of slippery wet surfaces and there are obvious privacy constraints to be taken into account.

Finally, the review of the state-of-the-art for radar and RGB-D sensors in the context of fall detection and human activity recognition is important to identify gaps and future research directions, in particular the possibility of having "multi-sensing systems" that combine these two technologies. It is believed this is a significant research direction, for example to develop systems where the radar part may provide longer detection ranges and insensitivity to light conditions, and the RGB-D part provides depth information useful when the Doppler information normally obtained from the radar is not sufficient for good classification. Additional discussion on gaps and complementarity of radar and RGB-D sensors is provided later on in this paper. Interesting examples of multi-sensing systems are provided in [16], [17], for example Kinect and accelerometers, or cameras with microphones plus accelerometers, but examples of radar and RGB-D systems are not mentioned in those review papers.

The rest of the manuscript is organized as follows. Section II presents a review of works that used radar systems, whereas section III is focused on RGB-D systems. Section IV provides a discussion on advantages and disadvantages of the different solutions described in the previous sections, and proposes future research challenges and trends. Section V finally draws the conclusions of this paper.

II. LITERATURE REVIEW ON RADAR SYSTEMS

Radar systems have been proposed only recently to address the problem of fall detection, with preliminary works such as

[19] starting to appear in the literature around 2011, and a growing trend of research works proposed from then up to now. The interest in this topic is demonstrated by two special issues published in February 2015 and March 2016, respectively the special issue on Application of Radar to Remote Patient Monitoring and Eldercare by *IET Radar, Sonar & Navigation* [20], and the special issue on Signal Processing for Assisted Living: Developments and Open Problems [21] by the *IEEE Signal Processing Magazine*.

There are many ways in which a person can fall, for instance falling forward rather than backward or towards the side, falling after tripping on some items or obstacles ,or as a consequence of the loss of balance or consciousness, or falling while attempting to reach a chair or a sofa to sit on [18]. However, the research work in the literature identifies the common characteristic of fall events to be *a quick and sudden acceleration during the actual fall followed by a slow deceleration while the person is lying on the floor* [22]. The proposed techniques to detect fall events aim to identify this fast acceleration from the radar data and to develop robust classification algorithms to reject false alarms caused by other movements that may cause comparable fast accelerations, such as bending to pick up an object or sitting down on a chair or sofa. The majority of the work presented on radar sensors for fall detection exploits the analysis of micro-Doppler signatures of people performing different activities, in order to extract the information on the velocity (proportional to the Doppler shift), and then identify the specific signature of a fall event. Other works discuss the use of range information obtained by Ultra Wide Band (UWB) radar systems, as well as the information on target velocity/acceleration that can be extracted from the phase of the received signal.

The general processing approach for fall detection is summarized in Figure 1. The starting point is always a dataset of experimental data, or data simulated with kinematics models in order to increase the amount of available samples to improve the classification performance. Then features have to be extracted from the data, i.e. numerical parameters that an algorithm implemented on a computer can understand. As mentioned before, these can be extracted from the data in the range domain, in the phase domain, or in the Doppler domain, with different approaches specific to each domain. In the case of micro-Doppler information, a suitable time-frequency transformation, such as the popular Short Time Fourier Transform (STFT) or Wavelet Transform (WT) and Extended Modified Beta Distribution (EMBD), is applied to the data before the feature extraction step in order to characterize the Time-Doppler pattern of the movement under test. The feature extraction step can be combined with a pre-screening step aimed at selecting the specific amount of data to be used for feature extraction in order to reject false alarms. For example, this step can identify the beginning and end of a potential fall looking at the velocity and acceleration, and only the amount of data between these two instants will be used for feature extraction. The final step is using the extracted feature samples as inputs to a classifier based on machine learning (ML) methods, where part of the data has been used to train the classifier, and the remainder is used for testing to assess the performance. Principal Component Analysis (PCA) can also be used to reduce the dimensionality of the available feature space, concentrating the relevant information for classification in a smaller number of features [23], and to automatize the feature extraction and selection procedure by reducing the influence of human operators' choices [24]. Many different types of classifiers featuring different computational complexity have been suggested in the literature, such as simple heuristic thresholds on parameters, Naïve Bayes, Nearest Neighbor with k elements (kNN), Support Vector Machines (SVM), and random forests. Most of these classical ML methods are based on the assumption that the input feature samples are independent and identically distributed, but this is not always true for human behavior data, whereby the actions that someone is doing at some point depend on previous actions and influence future actions [8]. Other ML approaches have been suggested to approach the case when the independent and identically distributed data assumption is not considered, such as Hidden Markov Models (HMMs), Conditional Random Fields (CRFs), and certain type of neural networks such as recurrent networks [8]. Methods to select a certain number of features for the classifier among those available can also be used at this stage, in particular wrapper methods that test all the possible feature combinations to find the optimal solution in terms of best classification performance, or filters methods that rank the available features according to a certain metric (e.g. the T-test or the mutual information) and then select the best N features [25]–[27]. It should be noted that similar processing steps are followed also for different types of sensors, such as wearable inertial measurement units (IMUs) as detailed in [13]. In the remainder of this section, details about the proposed techniques, their experimental validation, and their advantages and disadvantages will be discussed.

Table 1 at the end of the section summarizes the different approaches proposed by the different papers in terms of feature extraction, i.e. where and how one can process the valuable information capable of discriminating fall events from other movements. The choice of highlighting the different feature extraction methods is justified, as the selection of a suitable set of features is expected to have a more significant effect on the overall performance than the choice of a specific type of classifier [22], [28]. Table 2 provides a summary of the different types of radar sensors used in the reviewed works to collect data for fall detection. Most of the sensors appear to operate in the range of frequencies including C-band and X-band, specifically around the 5.8 GHz Wi-Fi band and 8 GHz, as well as a few systems operating at higher frequency in K-band (24 GHz), to exploit the larger bandwidth and related range resolution achievable at that frequency. Only few systems are reported working at lower frequencies, probably because the antennas tend to become rather large and unfeasible for indoor monitoring scenarios, and because the micro-Doppler signature is less suitable for feature extraction and classification, as the Doppler shift is proportional to the carrier frequency of the signal. Sensors using higher frequencies in the W-band region, for example in the 77 GHz or 90 GHz region where automotive

radar operate, were also not reported. This could be an interesting direction of research as very large bandwidth values are achievable at these frequencies enabling fine range resolution as well as large Doppler shifts, and the path loss suffered by the signal can still be suitable for indoor applications.

The use of radar sensors may raise some questions on possible hazards posed by the electromagnetic radiations. The transmitted power levels of the systems listed in Table II are of the order of a few dBm (e.g. +3 dBm reported for the network analyzer output), or limited to below +20 dBm for the commercial sensors. These levels appear to be comparable with the power transmitted by conventional Wi-Fi routers and smartphones (e.g. transmitted power up to +20 dBm EIRP for Wi-Fi access points by Cisco and other manufacturers), and microwave signals at these frequencies do not pose any risk of ionizing radiations. Additionally, the perceived risks associated to the use of radar devices featuring the transmitted power levels discussed above, shall always be traded off with the advantages obtainable by the continuous monitoring of a subject affected by physical or cognitive impairments, or disabilities that could expose him/her to potentially life-threatening conditions and dangers.

For practical deployment of radar systems, the actual transmitted power could be carefully specified based on the operational conditions, given the carrier frequency of the waveform, the maximum detection range, and the minimum SNR for the radar receiver to work properly. The well-known radar equation can be used for this power budget [29]. If we assume for a rule of thumb calculation a transmitted power equal to 100 mW (20 dBm), each antenna gain equal to 10 dB, 5.8 GHz carrier frequency (equal to a 5.2 cm wavelength), maximum range of 10 m for a fairly large room, and Radar Cross Section for an average human of 1 m^2 [30], this yields approximately -58 dBm of received power. This result is well within the receiver sensitivity to operate with satisfactory Signal-to-Noise Ratio (SNR), also taking into account the additional gain at the radar receiver chain (e.g. low noise amplification stages). For example, the Vector Network Analyzer (VNA) used as Doppler radar in some works such as [4] has a dynamic range of between approximately 90 and 110 dB depending on the frequency range. A precise calculation of the SNR of the radar signature will depend on the specific hardware used, on the transmitted power, and target reflectivity, but the above rule of thumb calculation and the works in the literature show that detection ranges of a few meters for indoor fall detection can be obtained with reasonable radar transmitted power levels. When considering detection and classification of actions through radar Doppler signatures, the contributions to the noise are given by clutter and thermal noise. The majority of the clutter is expected to be static clutter and can be filtered out with digital filters, whereas thermal noise can be represented by the noise figure of the specific radar receiver used in each case.

A. Fall detection using micro-Doppler

Micro-Doppler is defined as the additional frequency

modulations added to the main Doppler shift of a moving target, and in the case of human signatures these modulations are related to the swinging movements of limbs, torso, and head of the person [31]–[33]. Human micro-Doppler signatures have been extensively investigated for a variety of applications, including the recognition of humans versus vehicles or animals such as dogs and horses [34]–[37], the discrimination between different activities performed by different people, such as walking, running, crawling, and carrying objects [23], [25], [38]–[45], and the identification of specific individuals based on their particular walking gait [28], [46]. As an example, Figure 2 shows micro-Doppler signatures for four actions performed by the same subject and recorded by an off-the-shelf C-band radar system at the University of Glasgow. The four actions were sitting and standing, bending to pick up an object from the floor and then standing up, pushing towards the radar (i.e. moving one arm and hand quickly towards the radar and slowly away from it), and pulling away from the radar (i.e. moving one arm and hand slowly towards the radar and then quickly away from it).
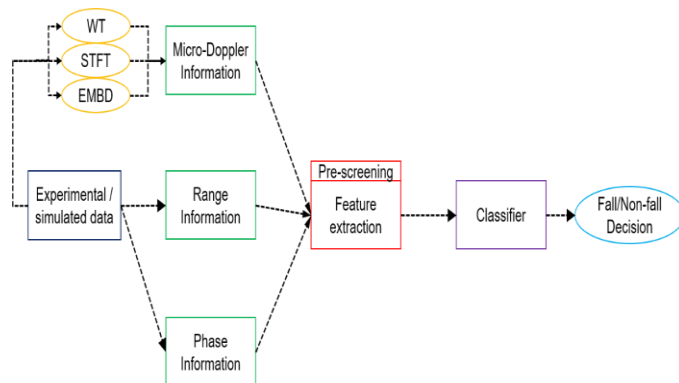

Fig. 1 Block diagram for data processing in radar-based fall detection
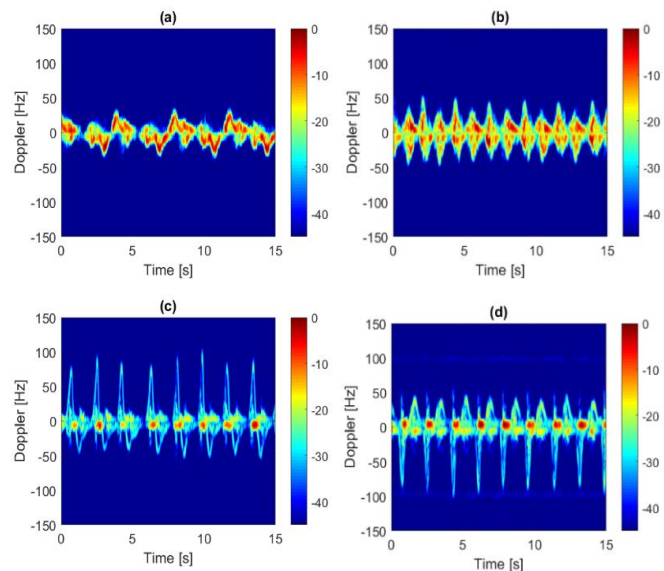

Fig. 2 Micro-Doppler signature of four actions performed by one subject: (a) sitting and standing, (b) bending to pick up an object and coming back up, (c) pushing towards the radar, and (d) pulling away from the radar

The research on fall detection draws extensively from the

aforementioned references, sometimes employing similar methods to characterize the time-frequency Doppler signature (e.g. STFT) and similar features to identify fall events (e.g. Mel-Frequency Cepstral Coefficients, MFCC), sometimes proposing specific features for the fall detection issue.

The work in [19] was one of the first published examples on the use of a commercial off-the-shelf pulse-Doppler radar system at 5.8 GHz for fall detection. The experimental data considered examples of various types of fall (forward, backward, towards the sides) and non-fall movements (such as bending, walking, kneeling, tying shoes, sitting) which were performed by eight subjects, including a professional stunt actor who tried to mimic an elderly person. The movements were performed in a laboratory. The data processing consisted of the calculation of the spectrograms via STFT, followed by a pre-screening step to identify the fall event looking at the energy burst curves (i.e. the time instant when significant acceleration is detected) and feature extraction using MFCCs coefficients. Simple kNN classifier with 3 neighbors and SVM classifier were used. The same data processing is applied in [47] to data collected in the realistic scenario of apartments specifically designed for assisted living and related to movements and actual falls of elderly people, as well as of stunt actors who helped provide fall data to train the classifiers. Besides this significant element of realism, this work is relevant to show the joint use of the radar system with a network of simple infrared motion detection sensors to reduce false alarms. The simple proposed fusion scheme assumes that a possible fall event recorded by the radar is likely to be a false alarm if a movement is detected by the infrared sensor shortly after it in the same room, which basically means that the subject is still able to walk and move around. A discussion on the suitable settings of the sensors and data processing parameters is provided, in order to achieve a good trade-off between rejecting false alarms but still being able to detect all the actual falls. The same group presented in [18] the use of the Wavelet Transform to extract suitable features for fall detection, and investigated different types of wavelet functions at dyadic scale to obtain the best classification accuracy. Three large datasets were analyzed, including both data collected in laboratory conditions and in actual apartments with elderly people living in them, and again professional stunt actors were also employed to try and perform the movements with the best realism as possible. The datasets are comprehensive of many different types of fall and non-fall movements, such as trip and fall, slip and fall, lose balance and fall, lose consciousness and fall, reach a chair or sofa and fall. The feature samples were processed by a simple NN classifier with a single threshold yielding good results. SVM classifier with linear and Gaussian kernel was also tested, but the results were not as good as for the NN classifier.

A different type of Wavelet Transform was also used in [48], where the continuous Morlet wavelet function was employed to process the data, followed by the extraction of three features from the resulting Doppler vs time pattern, referred to as 'scalogram'. These features are the lowest scale or corresponding highest frequency component, the ratio of the energy, and the rate of change of scale, and the classification

algorithm used a simple Mahalanobis distance metric between the feature vectors. This approach yielded good results, but was only preliminary tested in laboratory conditions on a small dataset, which included falling backward movements versus stand and sit movements performed by two subjects. The radar system was a Vector Network Analyser (VNA) operating as a Doppler radar at 8 GHz.

The use of Fractional Fourier Transform (FrFT) and Fractional Short Time Fourier Transform (FrSTFT) was proposed in [49] to achieve higher signal energy concentration and improve classification results for scenarios with low signal-to-noise ratio. The processing scheme assumed to use the FrFT on the data to compare the results with a threshold, and then initiate the classification routine based on the FrSTFT if this threshold is exceeded. Experimental data collected using an off-the-shelf radar system operating at 5.8 GHz with 100 MHz bandwidth were used to validate the proposed idea. The data included fall movements and non-fall movements such as walking, running, and performing deep squats, as carried out by several subjects. The feature samples were processed using simple Bayesian classification.

The work presented in [50] is also interesting for the combined use of two different Radio-Frequency (RF) sensors, namely a Doppler radar sensor operating at 24 GHz and a receiver array sensors operating at 800 MHz. These were previously investigated independently by the same authors in [51] and [52], respectively. The former system is expected to perform best in line-of-sight (LOS) conditions, whereas the latter can exploit multipath and non-line-of-sight (NLOS) propagation phenomena by looking at the signal subspace spanned by eigenvectors as a feature for movement classification. The features for the Doppler radar are MFCC coefficients, whereas those for the receiver array are the received signal strength, but also the correlation between the eigenvector under test and one collected when nobody was moving in the room, as well as a metric based on the eigenvalue. The feature samples for the two sensors are processed separately by two SVM classifiers based on radial basis functions kernel, and the results combined by a straightforward OR function to have a fall or non-fall decision. Promising results are shown also in NLOS scenarios thanks to the receiver array sensor, which can compensate for the non-optimal performance of the Doppler radar sensor in these scenarios.

The work in [53] proposed three features extracted from the spectrograms of the radar data for fall detection. The first step of the processing consisted on calculating the STFT of the data and extracting power burst curves to identify a possible fall event based on a threshold, i.e. time bins in the spectrograms with a sudden increase in velocity, which could be related to fall events. This was followed by the application of segmentation and morphological processing (i.e. image processing methods) on the identified portion of spectrogram to obtain binary black-white images. Suitable features were then extracted from the images, namely the extreme frequency magnitude, the extreme frequency ratio, and the length of the event, and used as inputs for a classifier, and the Mahalanobis distance metric between the feature vectors was used for

classification. The proposed method was validated with experimental data collected using a VNA triggered as Doppler radar at 8 GHz in laboratory conditions. The data included non-fall movements (such as sitting and standing or bending and standing) performed at normal and fast speed, and falling forward and backward with and without waving arms during the movement. In [54] this feature extraction procedure was applied in conjunction with sparse Bayesian learning based on the Relevance Vector Machine (RVM), showing promising results.

In [55] the same authors presented a hybrid approach combining compressive sensing and multi-window analysis based on Hermite and Slepian functions in order to restore time-frequency signatures to be used for fall detection. The concept of multi-window spectrograms involves having the weighted sum of K spectrograms, each calculated with K different window functions, among which the authors considered Slepian and Hermite functions, but also an ad-hoc kernel functions developed to reduce cross-terms components in the final signal. The concept of sparsity involves the capability of reconstructing the time-frequency signature in case some time domain samples are randomly missing and the resulting spectrograms look noisy as a result of that. The authors tested three different methods for reconstruction, using single and multiple measurement vectors with and without multi-windowing. In terms of feature extraction and classification, the authors used a similar approach to [53], by localizing the possible fall event using power burst curves and center of gravity metrics, and then extracting the extreme frequency magnitude. The classifier used was SVM, with data collected at 8 GHz for two subjects performing five types of movements such as bending over, sitting and standing, falling backward and forward, as well as falling with a 45° aspect angle with respect to the line-of-sight of the radar.

The work in [56] presented an interesting comparison of using different types of features on the same dataset. The data were collected using a commercial Doppler radar at 24 GHz and processed using SVM classifier with radial basis functions. The movements considered were falling, sitting, walking, and picking up an object, and involved four subjects who took part to the experiment. The different types of features were the three empirical features extracted by the spectrograms as in [53], the power burst curves used as a whole vector for classification, the energy between the start and the end of the fall event calculated on the result of Wavelet Transform as in [18], and MFCC coefficients as in [19]. The first set of features extracted from the spectrograms appeared to outperform the other features.

Although not directly related to fall detection, the work presented in [57] is interesting to characterize the walking gait when walking assistive devices are used, for example a walking cane in this case. Any practical fall detection system needs to be able to differentiate between fall events and normal movements, but it is likely that actual elderly people will use walking devices such as canes or walkers in their daily activities, hence their effect on the overall human micro-Doppler signature has to be characterized [58]. Another element of interest in [57] is the use of a different time-frequency

distribution, the Extended Modified Beta Distribution (EMBD), as an alternative to the STFT to overcome the problem of the trade-off in frequency and time resolution.

The recent work presented in [4], [59] introduced the use of deep learning algorithms and convolutional neural networks (CNN) to perform fall detection. The interest in this approach is that the feature extraction step is bypassed and the spectrograms and related class labels are directly provided as inputs to the CNN, after a pre-processing step to reduce the noise and apply a grey scale to the images. The task of identifying common patterns within a class and discriminant features between different classes is left to the CNN itself, reducing the possibility of discarding useful information when extracting features with procedures designed by human operators. CNNs will imply larger computational costs to be trained, but the technological trend of increasing computational power at reduce cost is likely to make this approach more and more common. The paper provided a preliminary validation of the idea, showing good classification results for four movements (falling, walking, sitting, and bending and getting up) and improved results with respect to the conventional method of using the features in [53] and SVM classifier.

Finally, it should be pointed out the use of PCA in the context of radar-based fall detection [24], [60], and more in general for micro-Doppler based classification of human activities and movements [61], [62] and for automatic target classification [23], [63], [64]. PCA and its variant Robust PCA have been investigated in the literature as a way to reduce the dimensionality of the feature vectors in order to select the most relevant features, and ease the computational burden and the amount of training data needed for the classifier.

B. Fall detection using range information

The recent work presented in [65] introduced the idea of exploiting also range information generated by UWB radar systems to improve the classification performance for fall detection. The authors argued that fall events not only present a high velocity component and sudden acceleration in the Doppler domain, but also a larger spread in the range domain caused by the simultaneous and not-coordinated movement of the whole body while falling. The use of these range-based features together with more conventional Doppler features was preliminary demonstrated on data containing three actions (falling, sitting, and bending over) performed by four subjects and processed by SVM classifier. A UWB radar operating at 24 GHz with 2 GHz bandwidth (corresponding to 7.5 cm range resolution) was employed to collect these data. The research question on how much range resolution and therefore waveform bandwidth is actually required to extract effective range-based features is still open.

C. Fall detection using phase and velocity information

The work in [66] showed a different approach for fall detection rather than using the micro-Doppler signatures. The authors proposed to use the phase information of the complex high resolution range profiles obtained from the inverse Fourier Transform of the response of a Stepped Frequency Continuous

Wave (SFCW) radar. The velocity and the acceleration of the moving target were then extracted from these profiles. The proposed radar system had 1 GHz bandwidth between 2.5 and 3.5 GHz, with a Pulse Repetition Frequency of 2 kHz. Although only preliminary results for sitting on a chair and falling down were shown, this type of features can be interesting to be explored as complementary information to improve micro-Doppler based classification.

A similar idea is proposed in [67], where the authors developed a radar system capable of generating a hybrid waveform which can alternate between a sinusoidal tone at 5.8 GHz and a SFCW waveform sweeping between 6 and 7 GHz. It should be noted that this system was developed to be compliant with the Federal Communications Commission (FCC) UWB specifications, so its practical deployment together with other electronic systems would not incur electromagnetic compatibility issues. The data processing used the I and Q components generated by the radar to extract the velocity of the target, with the assumption that fall events present quickly increasing velocity followed by a sudden stop when the person touches the ground, whereas normal movements present more controlled transitions of the velocity. The speed signals were processed either using Fast Fourier Transform (FFT) with SVM classifier with linear or radial basis function kernel, or using STFT and SVM classifier with global alignment kernel. The authors also tested the use of the Dynamic Time Warping method for classification together with the Euclidean distance metric. They showed that SVM with global alignment kernel provided the best result. The data used to validate these methods and the proposed system contained falling and walking movements performed by two subjects in laboratory conditions, but trying to have furniture in the environment to mimic a realistic indoor room. In [67] the radar system was mounted on the wall at approximately 150 cm height, but in [68] the authors briefly discussed the possibility of having the sensor mounted on the ceiling and showed that the classification results can be improved by exploiting this different aspect angle to the person to monitor. The system presented in [67] had an integrated microcontroller and a Zigbee module to enable communication with a base station in the perspective of realizing a complete monitoring system, beyond the sensor component. This idea was expanded in [69], where the authors considered a whole telemedicine system whereby the board with radar sensor and Zigbee communication module interacts with a base station performing on-line signal processing on the data. This base station had in memory a classifier model that was calculated off-line in MATLAB at the training step, and performed on-line testing on the incoming data to provide a fall or non-fall decision. The data processing was similar to [67], with the SVM classifier with global alignment kernel operating on the STFT of the velocity signals extracted from the I and Q radar data. The experimental validation used data from 3 subjects to train the classifier and data from 16 subjects for testing. Non-fall movements such as walking, dropping objects, sitting and standing, and performing daily actions such as eating, drinking, talking on a mobile phone were considered, as well as different types of falls either directly on the floor or with attempts of grabbing objects while falling. The results were promising, with reported accuracy close to 100% with no false positives and real-time operations. The authors highlighted also limitations of their approach, for instance the fact that the subject can be obstructed by furniture, outside of the radar antenna beam-width, or at an unfavorable aspect angle for the radar to detect the velocity signal. The use of multiple cooperating sensors was indicated as possible solution to address these issues.

TABLE 1 SUMMARY OF DIFFERENT APPROACHES PROPOSED IN THE LITERATURE FOR FEATURE EXTRACTION AIMED AT FALL DETECTION

References	Approach for feature extraction
[19], [47], [50], [56] [51]	MFCC coefficients
[53]–[56]	Empirical features extracted from spectrograms after STFT
[49]	Features extracted from fractional FFT and STFT
[18], [48], [56]	Features from Wavelet Transform
[57]	Empirical features after EMBD
[4][59]	Feature extraction bypassed by using CNNs
[65]	Additional features extracted from range information
[66]	Velocity information from phase of radar range profiles
[67]–[69]	Velocity information from I-Q radar data
[50], [52]	Features from subspace decomposition of signals received at array

TABLE 2 SUMMARY OF DIFFERENT RADAR SENSORS PROPOSED IN THE LITERATURE FOR FALL DETECTION

References	Type of radar sensor
[18], [19], [47]	Commercial pulse-Doppler radar operating at 5.8 GHz
[49]	Commercial Frequency Modulated Continuous Wave (FMCW) radar operating at 5.8 GHz (100 MHz bandwidth)
[47], [52]–[55]	Vector Network Analyser operating as a Doppler radar at 8 GHz (1 kHz sampling rate)
[4]	Vector Network Analyser operating as a Doppler radar at 6 GHz (1 kHz sampling rate)
[50], [51]	CW Doppler radar operating at 24 GHz (1024 Hz sampling rate) or receiver array operating at 2.457 GHz. Alternative is a combination of these two sensors, with the array operating at 800 MHz
[56]	Commercial CW Doppler radar operating at 24 GHz (1 kHz sampling rate)
[65]	Ultra Wide Band radar with centre frequency at 24 GHz with 2 GHz bandwidth

[67]–[69]	Hybrid radar: Stepped Frequency Continuous Wave (SFCW) radar operating at 6-7 GHz plus CW radar with sinusoidal tone at 5.8 GHz
[66]	SFCW radar operating at 2.5-3.5 GHz with Pulse Repetition Frequency equal to 2 kHz

III. Literature review on RGB-D systems

The availability in the market of inexpensive RGB-D sensors fostered researchers working in the computer vision area to combine depth data and RGB images. The depth information, previously available using expensive Time-of-Flight (TOF) cameras or multiple calibrated cameras, has brought many advantages in the development of vision-based solutions. Depth data provide 3D information which can simplify many challenging tasks, such as people segmentation and tracking, body part recognition, or motion estimation [70]. These tasks are fundamental when the aim is to monitor people and detect dangerous events, such as falls. Microsoft Kinect allows also the extraction of skeleton joints, which provide a compact and informative representation of the human body [71], as shown in Figure 3. Two different versions of the Kinect sensor, namely v1 and v2, have been released in the past years, with different characteristics summarized in Table 3. Both versions provide RGB, depth, and IR raw data with different resolution, and audio. Microsoft SDK enables the extraction of these data and the evaluation of the skeleton joints of a human. The algorithms for depth sensing exploited by Kinect, i.e. structured light for Kinect v1 and TOF for Kinect v2, are based on IR signals. These algorithms may be affected by errors if the monitored area is characterized by a reflective surface [72], and this uncertainty in the evaluation of depth data can generate an error in the joint estimation process [73]. Even when some corrections may be required, depth data extracted from Kinect can be used to design algorithms with performance comparable to gold-standard systems, as discussed in [74].

TABLE 3 CHARACTERISTICS OF THE MOST USED RGB-D SENSORS: KINECT V1 AND V2

Feature	Kinect v1	Kinect v2
Depth sensing technology	Structured light	Time of flight
RGB image resolution	640x480 @ 15/30 fps 1280x960 @ 12 fps	1920x1080 @ 30 fps (15 fps with low light)
IR image resolution	640x480 @ 30 fps	512x424 @ 30 fps
Depth sensing resolution	640x480 @ 30 fps 320x240 @ 30 fps 80x60 @ 30 fps	512x424 @ 30 fps
Depth sensing range	0.4-3 m (near mode)	0.5-4.5 m
	0.8-4 m (normal mode)	
Field of View	57° horizontal, 43° vertical	70° horizontal, 60° vertical
Skeleton tracking	Skeleton with 20 joints Up to 2 subjects	Skeleton with 25 joints Up to 6 subjects
Audio	multi-array microphone	multi-array microphone

Many solutions based on depth data processing have been proposed to detect falls, with different setups of the RGB-D device and different types of data used as source of information. Similar to radar-based approaches, algorithms for fall detection exploiting vision-based devices process data acquired from a dataset including multiple repetitions of different classes, usually organized in fall or non-fall classes. As shown in Figure 4, the first step consists in the computation of features from one or multiple types of data (RGB, depth, skeleton, or two of them). Then, the algorithms available in the literature may exploit a rule-based approach or a ML approach. In the former case, the algorithm does not need to be trained and some rules are derived empirically. Such rules can be often related to distances between the human and the floor plane, and/or to the trajectories of the skeleton joints. ML approaches can be also considered when a sufficiently large training dataset is available, including different fall and non-fall sequences. These techniques can be also adopted to discriminate between different actions (bending, lying, sitting, etc) considering each one as a different class. Classic ML algorithms such as kNN, AdaBoost, SVM, HMM have been tested with good results.
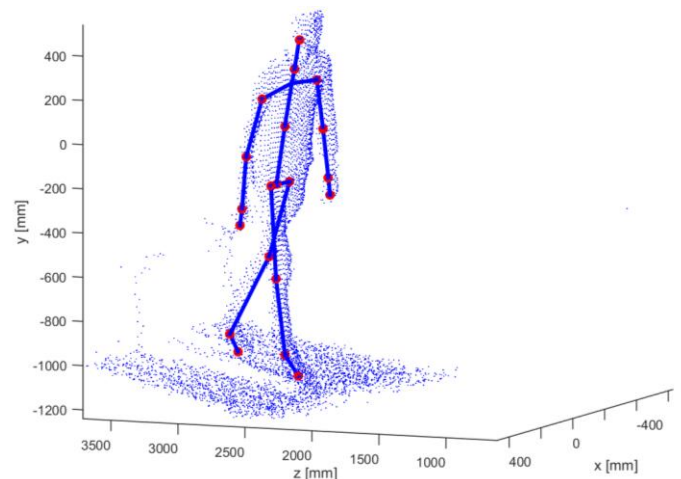

Fig. 3 Point cloud and skeleton joints of a human extracted from Kinect depth data
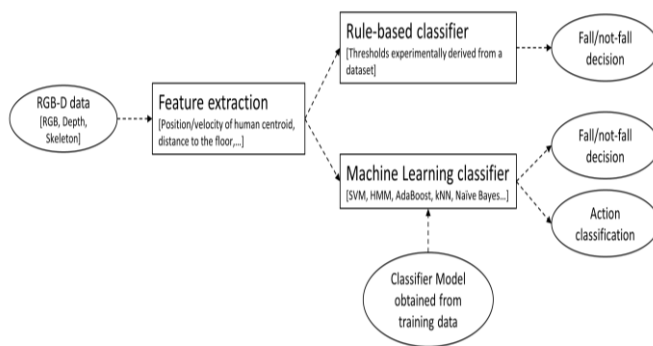
Fig. 4 Block diagram of a fall detection algorithm based on RGB-D data

In this section, the different works are classified into two groups based on the data exploited for fall detection, namely approaches exploiting only depth data in the process of features computation, and approaches based on multiple information fusion. The availability of depth data is the main reason of using an RGB-D sensor, thus most of the approaches are using this type of information. The feature extraction process may apply background subtraction algorithms to extract the human silhouette, and the computation of some features from the shape of the silhouette. In addition to the identification of the person, many algorithms consider the automatic identification of the floor, which can be efficiently carried out by exploiting depth information. In addition to depth data, many algorithms exploited also the human skeleton provided by Kinect. The skeleton simplifies the process of features extraction, since the joint coordinates can be directly considered as features related to the position of the human body and its parts. Depth information is still processed to extract the floor plane, which is often considered when the classification is simply binary (fall or non-fall). All the reviewed methods are summarized in Table 4.

A. Fall detection using depth data

Rougier et al. [75] proposed a solution which is robust against occlusions, exploiting two features, namely the human centroid height relative to the ground, and the body velocity. The former is a simple and efficient quantity to detect falls, whereas the latter helps overcome the problem of occlusions, allowing the detection of falls even when the subject is completely occluded behind furniture. They used the V-disparity image method [76] to detect the ground plane, assuming that the floor is a large part of the scene. A threshold-based approach with a background image is used to extract the foreground depth map, which is the person's silhouette. The impact of occlusions can be limited by considering a different setup for the RGB-D sensor. Gasparrini et al. [77] proposed a system where the device is placed in a "on-ceiling" configuration, to provide a complete top view of the scene. The pre-processing and segmentation phase is exploited to extract the foreground depth map from a static background frame and to segment the objects on the scene. Some features extracted from object shapes allow to detect the presence of a person, and it is possible to track the movements and detect a fall considering the distance from the floor. The algorithm has been

evaluated on a dataset recorded from 4 actors performing 20 tests, 10 of which involve the presence of several subjects in the area. The same sensor configuration has been used by Liciotti et al. [78]. A background subtraction algorithm based on Gaussian Mixture Model returns a foreground image that contains the people and, possibly, moving objects. The segmentation algorithm exploits a set of parameters to distinguish people from objects, namely the height of each person, the size of each head and the head-shoulders distance. Falls are detected when the depth value of the head is close to the floor level. Other methods based on the shape of subjects have been proposed, exploiting for example a set of moment functions which approximate the human shape to an ellipse, whose coefficients are calculated to determine the direction and position of the individual [79]. In order to have a more accurate system, the centroid of human body and the angle between this and the floor plane have been calculated. Empirically calculated thresholds allow the fall event detection. A threshold-based method allows to obtain the human silhouette with background subtraction, while the floor detection is initially achieved by a V-disparity map with the adoption of least squares method for the estimation of the floor plane equation. Mastorakis and Makris [80] measured the velocity based on the contraction or expansion of a 3D bounding box enclosing the human silhouette. A training dataset is considered to optimize, using random search, the velocity thresholds for the height and the width–depth composite vector of the bounding box, and for the number of frames constituting a fall. The exceeding of some thresholds on height and width-depth of 3D bounding box detects the starting of a fall, while the final step is the monitoring of the subject for some time after the fall to detect any motion. This algorithm does not require the computation of the floor plane, and the human shape is extracted using features of OpenNI library [81]. A similar approach, using the 3D bounding box of the human, is adopted also in [82], where the y-coordinate of the top left vertex of the box is monitored to reduce false alarms. Nghiem et al. [83] proposed an algorithm to extract the human head position from depth information, and its application to fall detection. Specifically, the fall event is detected by considering the vertical speed of head and body centroid, together with their distance from the floor. A dataset consisting of 30 fall, 18 crouch, and 13 sit down actions has been used for evaluation, which resulted on the correct classification of 29 falls out of 30.

Fall/non-fall decisions can be achieved also by using statistical methods, instead of using threshold-based approaches. A two-stage fall detection system, based on depth data, is presented in [84]. The first step is the characterization of the vertical state of a segmented 3-D object and the identification of the so-called "on-ground events". The second stage extracts five features, including velocity and acceleration, from these events and computes a fall confidence index considering an ensemble of decision trees. The method proposed by Zhang et al. [85] processed the depth frame to extract the head region from the human body and the floor level. The system always considers an interval of frames, extracts five features, and makes a decision for the whole interval using the

fall and non-fall distributions computed at training. The performance was evaluated on a dataset consisting of two viewpoints including 12 real falls in scene 1 and 14 real falls in scene 2, together with other fall-like activities (sitting down on the floor, picking up an object from the floor, etc). Kepski and Kwolek [86] extracted the ground plane automatically using the V-disparity images, Hough transform and the RANSAC algorithm. The human shape is detected using a depth reference image, which is periodically updated considering pixels from the current depth image. A virtual box surrounding the person is computed, and features based on shape and distance are extracted and used as input to different classifiers. Thirty-five young volunteers were involved in the recording of the dataset used to evaluate the system, collected considering two Kinect sensors. Bilski et al. [87] proposed the use of two synchronized Kinect sensors and an algorithm based on kNN to detect falls. The depth frame is initially transformed into the absolute representation based on global space coordinates. From the human silhouette, the x-y-z coordinates of its center of mass and its magnitude, which is the effective reflection area, are extracted for each frame, constituting a four-dimensional trajectory considering the whole sequence. A set of characteristic points, consisting in the minimum, maximum, and differences at some specific times in the sequence, has to be computed from each pattern and used for classification. The solution has been evaluated on a dataset of 18 fall scenarios and 18 scenarios corresponding to other actions. All actions were performed by two actors and recorded by two Kinects.

More complex algorithms are able to distinguish falls and other actions. Ma et al. [88] proposed an approach based on CSS features [89] computed from human silhouette which are invariant to human translation, rotation, scaling, and action length. The improved Extreme Learning Machine (ELM) algorithm is less sensitive to tuning parameters and allows the classification of five more actions in addition to falls: walking, sitting, squatting, bending, and lying, all included in the *SDUFall dataset*. Aslan et al. [90] proposed the use of Fisher Vector (FV) representation to build the vocabulary, instead of classical BoW (Bag of Words) approaches based on k-means for clustering. Then, they used binary SVM to distinguish fall actions from other actions. The CSS features adopted in this work are obtained from the human silhouette extracted by using Canny edge detector after foreground segmentation on depth map. An accuracy classification of 89.84% was obtained on *SDUFall dataset* [88], considering five actions (bending, lying, sitting, squatting, and walking) out of six as non-fall activities. The method proposed in [91] is based on the real-time detection of the center of mass of any moving object. A dynamic background subtraction technique is adopted to extract, from each depth frame, the mobile points, that feature a different depth value if compared to the background. After the extraction of the center of mass for each person, a tracking procedure has been implemented, and the recognition process is performed considering a HMM with a number of states corresponding to the number of classes in the dataset. The system calculates the probability of being in one of the states and associates that probability to the corresponding action. Considering a dataset constituted by 8 activities, the system achieves a sensitivity of 90% and a specificity of 100% for the falling events. A shape sequence descriptor, namely the Silhouette Orientation Volume (SOV), has been proposed in [92]. This descriptor has been associated to BoW models, for which the codebook has been built considering k-medoids clustering technique, and then Naïve Bayes classifier can be used to recognize fall related actions of the *SDUFall dataset* and also more general actions of the Weizmann dataset [93]. Following a static background subtraction process, morphological operations are used to remove the noise close to the segmented human shape. Edge detection is then adopted to obtain the edges of the silhouette. A SOV descriptor is a sequence of SOIs (Silhouette Orientation Images) or a volume of silhouette orientations. SOIs are robust to scale, planar rotations, and starting point, and can provide a global definition of the silhouette. Considering fall vs. non-fall classification, the proposed solution achieves a performance of 91.89%.

B. Fall detection using multiple information fusion

Planinc and Kampel [94] proposed the computation of the major axis of a human shape using skeleton data obtained by Kinect. In particular, they exploit the 3D coordinates of the head, shoulder center, spine, hip and knees to extract the human's axis, and process depth data to obtain the ground floor. A fall is detected when the person is parallel to the ground floor and the distance between spine joint and the ground floor is small. The evaluation has been carried out on 72 different sequences, with 40 falls and 32 non-falls performed by 2 actors twice. In [95], the same authors introduced a pose estimation algorithm based on fuzzy logic to define the similarity to the ground plane and the distance to this. A skeleton based algorithm has been proposed in [96], where a first evaluation is performed considering position and velocity of the user's center of mass. A reduction of false alarms is then achieved by a postural recognition algorithm which analyses the relative positions of lower body joints. Kawatsu et al. [97] considered taking a fall/non-fall decision every frame. The distance between all the skeleton joints that are in the "tracked" state and the ground floor must be lower than a threshold to detect a fall. A more robust algorithm considers data from multiple frames to distinguish between falls and people lying on the floor. Finally, once a fall is detected, the event has to be confirmed by the user through a voice recognition system and the Kinect microphone array. A method exploiting depth shape analysis and RGB images is proposed in [98]. A threshold-based background subtraction is adopted to extract the objects, and the human is detected considering skin colored pixels. The human is tracked through the coordinates of its centroid and, if a large vertical motion event is revealed, another mechanism based on the orientation of the main axis of the human shape is adopted to discriminate between fall or squat event.

Bian et al. [99] proposed a method with low computational cost to extract useful joints from depth maps. Specifically, an improved randomized decision tree (RDT) algorithm can extract head and hip joints. After the extraction of the floor plane, the trajectory of the distance between the joint and the

floor is computed and considered as input feature vector to a SVM classifier. The performance has been evaluated considering the scenario proposed in [100], where 4 categories of falls and a set of non-fall actions are evaluated. Each scenario has been simulated in a real bedroom several times by 4 young people, having a total number of 380 samples classified with an accuracy of 97.6%. Amini et al. [101] proposed a comparison between heuristic and ML algorithms for fall detection with Kinect. The heuristic method is based on skeletal data, and the 3D coordinates of head joint are tracked. A fall is detected by setting a threshold on the velocity and acceleration of head joint, together with a small distance between the head and the floor. The ML approach is based on an AdaBoost algorithm that combines a series of weighted weak classifiers to have a final boosted classifier. Only the velocity and the subject's head distance to the floor have been considered for the ML approach. For both heuristic and ML algorithm, the dataset was captured considering 11 young subjects. Each subject performed six true positive and six false positive fall incidents, which included laying down or sitting on the floor. The rule-based approach reached an accuracy of 95.42% of falls detected, while the machine learning one is less accurate (88.33%) due to the limited number of subject's samples. Dubey et al. [102] proposed to use RGB and depth data to extract the motion from the data using Three-Dimensional Motion History Images (3D-MHIs) [103] and then to compute features, represented by 7 Hu moments [104], from the 3D-MHIs. The 3D-MHIs can detect change in motion in x-y-z direction, increasing the classification capability with respect to MHI [105]. The features considered for classification are the Hu moments for each of the 3D-MHIs, which are then used to train a SVM to recognize between falls and non-falls. The method presented in [106] uses only the extracted skeleton data and is optimized to detect falls related to weight shifting problem. Features calculated from skeletons considers height, vertical speed of upper body, body orientation and its variations, projection of the center of mass on the ground. A linear SVM takes the extracted features and classifies fall events from non-fall events, where non-fall ADLs activities include walking standing, sitting and sleeping.

Zhang et al. [107] proposed to use the joints of head and torso, which are correctly detected if a person is standing or sitting, and wrongly estimated if a person falls. They defined a kinematic feature vector considering the angles between couples of joints on different skeletons, and minimum and maximum values of the height of the person within a sequence of frames. They can detect 5 fall related actions exploiting also RGB information to extract the human shape if the skeleton is not available. Dai et al. [108] have chosen HMMs to model temporal sequences of postures which constitute an action. The 60 dimensional vector with the coordinates of 20 skeleton joints are first reduced using Principal Component Analysis (PCA). All the sequences from 7 actions (6 ADLs and 1 fall) are partitioned into clusters to extract relevant postures. Finally, a HMM model is trained for each action, and the motion class corresponding to the model achieving the highest likelihood is the recognized class. Alazrai et al. [109] proposed a view-invariant Motion-Pose Geometric Descriptor (MPGD)

computed from skeleton joint positions, capable of capturing the motion and poses of human body-parts while preserving the temporal ordering of the moving body-parts. The fall detection framework consists of two classification layers. The first one is a set of SVMs which describes the state of the person on each frame. At the second layer the constraint dynamic time warping (cDTW) technique is used to classify the whole sequence of states into falling or non-falling events. This method achieved good results in the classification of four activities.

C. RGB-D datasets for fall detection

The growing interest on RGB-D data fostered some researchers to collect datasets and to provide them to the community. A fair comparison among different algorithms can be performed considering common data that are made available to the research community. In the past years, many researchers have recorded and shared several datasets containing RGB-D data, reviewed in [110] [111]. Only a few of these datasets are suitable for fall detection, and their characteristics are summarized in Table 5.

The *TST Fall Detection v2* [112] is the most recent dataset, recorded using Microsoft Kinect v2 and 2 accelerometers placed on the wrist and waist of the subjects. Each of them performed 4 different ADLs and 4 types of falls 3 times in laboratory environment, generating a total number of 264 sequences. The ADLs are: Sit on a chair, Walk and pick up an object from the floor, Walk back and forth, Lie down on the mattress. The 4 types of falls are: Fall from the front ending up lying, Fall backward ending up lying, Fall to the side ending up lying, Fall backward ending up sitting.

The *UR Fall Detection* [113] is the only other dataset providing acceleration samples. It has been collected from 5 subjects and 2 cameras, one parallel to the floor and another one mounted on the ceiling. Some additional features, e.g. those characterizing the bounding box around the person, are also provided. The dataset consists of 70 images sequences with 40 ADLs and 30 falls belonging to two categories, falls from standing position and falls from sitting on the chair.

The *SDUFall dataset* [88] includes data captured from 20 people performing 6 different actions, and is the largest available dataset, with each subject repeating each action 10 times. The considered actions are falling down, bending, squatting, sitting, lying and walking, and they are different in each repetition as the actors may carry or not carry large objecst, turn the light on or off, change direction and position relative to the camera.

The *Falling Detection dataset* [85] has been collected in a laboratory environment, with two Kinects mounted at two upper corners of the room. The actions performed by 6 subjects include 26 real falls and other fall-like actions, such as picking up something from floor, tying shoelaces, sleeping down on the bed, sitting on the floor, opening drawers close to the floor, jumping on the floor and sleeping down on the floor.

The *ACT42 dataset* [114] mainly focuses on the ADLs, including 14 actions such as Collapse, Drink, Make Phone Call, Mop Floor, Pick Up, Put On, Read Book, Sit Down, Sit Up, Stumble, Take Off, Throw Away, Twist Open and Wipe Clean.

Two categories of falls are considered, namely Collapse (fall due to internal factors) and Stumble (fall due to external obstacles). All the actions were performed multiple times by 24 people.

The *Falling Event dataset* [115] provides only skeleton data of 5 activities including falls and non-fall events, such as standing, fall from standing, fall from sitting, sit on a chair, and sit on floor. The dataset has been recorded considering actions performed by 5 people under two different environmental conditions, sufficient and insufficient illumination. The webpage was no longer available at the time of writing.

The *EDF dataset* [15] was collected at the University of Texas, where a simulated apartment has been set up. Two Kinects have been installed to cover with different direction of

falling. The falls are then repeated for each viewpoint, leading to a total number of 320 sequences. In addition to falls, a number of 100 sequences of 5 different actions that could be associated to falls are recorded. The additional actions are pick up an object, sit on the floor, lie down on the floor, tie shoelaces, and do plank exercise.

The *OCCU dataset* [113], as the previous one, includes data from two Kinects placed at two corners of a simulated apartment. The main feature of this dataset is the presence of occluded falls for which the end of the action is completely occluded by an object. Five subjects simulated 12 falls, 6 for each viewpoint. Similarly to the *EDF dataset*, 80 sequences of actions that can be confused with falls are also provided.

TABLE 4 SUMMARY OF ALGORITHMS FOR FALL DETECTION BASED ON RGB-D DATA. THE USAGE OF RGB, DEPTH DATA, AND SKELETON JOINTS IS INDICATED BY RGB, D AND S RESPECTIVELY (S* DENOTES THE USAGE OF A SKELETON MODEL ALTERNATIVE TO THE ONE PROVIDED BY KINECT)

Algorithms	Data	Actions	Features	Classification
Rougier et al. [75]	D	Fall/non-fall	- Human centroid height relative to the ground - Body velocity	Rule-based
Gasparrini et al. [77]	D	Fall/non-fall	Human centroid height relative to the ground	Rule-based
Liciotti et al. [78]	D	Fall/non-fall	Head height relative to the ground	Rule-based
Yang et al. [79]	D	Fall/non-fall	- Human centroid height relative to the ground - Orientation of human body relative to the ground	Rule-based
Mastorakis and Makris [80]	D	Fall/non-fall	Velocity of contraction or expansion of a 3D bounding box enclosing the human silhouette	Rule-based
Bevilacqua et al. [82]	D	Fall/non-fall	- Velocity of contraction or expansion of a 3D bounding box enclosing the human silhouette - Real world y-coordinate of 3D bounding box vertex	Rule-based
Nghiem et al. [83]	D	Fall/non-fall	- Vertical speed of head and body centroid - Human centroid and head distance to the floor	Rule-based
Stone and Skubic [84]	D	Fall/non-fall	- Identification of "on-ground" events using vertical state estimation time series - Velocity-based features to extract falls from "on-ground" events	Ensemble of decision trees
Zhang et al. [85]	D	Fall/non-fall	Features related to position/velocity of human head to the ground	Gaussian model for falls Histogram model for non-falls
Kepski and Kwolek [86]	D	Fall/non-fall	- Shape features from 3D bounding box enclosing the human - Human centroid height relative to the ground	KStar, AdaBoost, SVM, MLP, Naïve Bayes, kNN
Bilski et al. [87]	D	Fall/non-fall	Features from the four-dimensional trajectory of silhouette "mass center" and its magnitude	kNN
Ma et al. [88]	D	6 actions (*SDUFall*)	CSS features from human silhouette	VPSO-ELM
Aslan et al. [90]	D	6 actions (*SDUFall*)	CSS features from human silhouette	FV-ELM
Charpillet Dubois [91]	D	8 actions	Tracking of human center of mass	HMM

Akagunduz et al. [92]	D	6 actions (*SDUFall*)	Silhouette Orientation Volume (SOV) from human silhouette	Naïve Bayes
Planinc and Kampel [94]	D, S	Fall/non-fall	- Major orientation of the person - Spine distance to the ground	Rule-based
Planinc and Kampel [95]	D, S	Fall/non-fall	- Major orientation of the person - Spine distance to the ground	Fuzzy logic
Lee and Lee [96]	S	Fall/non-fall	- Position and velocity of human center of mass - Relative position of lower joints	Rule-based
Kawatsu et al. [97]	D, S, Audio	Fall/non-fall	- Vertical velocity of skeleton joints - Voice recognition to validate the fall	Rule-based
Yang and Lin [98]	D, RGB	Fall/non-fall	- Large vertical motion event of human centroid - Main axis orientation of the human shape	Rule-based
Bian et al. [99]	D, S*	Fall/non-fall	Trajectory of distance between head/hip joints and the floor	SVM
Amini et al. [101]	D, S	Fall/non-fall	- Velocity of head joint - Head distance to the floor	AdaBoost
Dubey et al. [102]	D, RGB	Fall/non-fall	Hu moments for each of the 3D-MHIs	SVM
Zhang et al. [107][115]	D, RGB, S (*Falling Event*)	5 actions	- Angles between couple of joints and min/max values of height of the human - RGB adopted if the skeleton is not available	SVM
Dai et al. [108]	S	7 actions	3D skeleton coordinates	HMM
Alazrai et al. [109]	S	4 actions	Motion-Pose Geometric Descriptor (MPGD) computed from skeleton joint positions	SVM + cDTW

TABLE 5 RGB-D DATASETS FOR FALL DETECTION. IN THE COLUMN RELATED TO DATA EACH LABEL REPRESENTS THE AVAILABILITY OF A DIFFERENT TYPE OF DATA: RGB (R), DEPTH (D), SKELETON (S), ACCELERATION (A).

Name	Actions	Actors	Samples	Data	Cameras	Year
TST Fall Detection v2 [112] http://www.tlc.dii.univpm.it/blog/databases4kinect	5	11	264	D, S, A	1	2015
UR Fall Detection [113] http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html	2	5	70	R, D, A	2	2014
SDUFall [88] http://www.sucro.org/homepage/wanghaibo/SDUFall.html	6	20	26400	R, D, S	1	2014
Falling Detection [85] http://vlm1.uta.edu/~zhangzhong/fall_detection/	8	6	87	D	2	2012
ACT4[2] [114] *https://sites.google.com/site/qinleisite/Home/dataset*	14	24	6844	R, D	4	2012
Falling Event [115]	5	5	150	S	1	2012
EDF [15] *https://sites.google.com/site/kinectfalldetection/*	6	10	420	R, D, S	2	2015
OCCU [116] *https://sites.google.com/site/occlusiondataset/*	5	5	110	R, D, S	2	2014

IV. FURTHER DISCUSSION AND RESEARCH TRENDS

A. Accuracy and error rates

An important figure of merit to compare the different radar and RGB-D systems described in the previous sections is the rate of correct fall detection (i.e. how many fall events are actually detected out of all the events), and similarly the false alarm rate (how many non-fall events are mistakenly classified as falls). However, a fair comparison of the different methods proposed in different studies is not easy, as the methods are often evaluated on different datasets, acquired with different configurations of the sensors and involving different numbers of actions and subjects.

The most common evaluation method for both radar sensors and RGB-D sensors consists in the acquisition of ad-hoc datasets representing some activities of daily living and one or different types of falls, followed by the subsequent analysis of

fall vs non-fall classification. The performance reported by the authors is usually quite high, sometimes higher than 95% in terms of accuracy, but this is generally achieved over a limited number of subjects and activities considered. Activities often include sitting, lying, bending and picking up an object vs fall events. For RGB-D sensors, the number of subjects is generally around 5 people, but more subjects have been reported in [80] (8 people) and in [101] (11 people), and in particular in [86] (35 people) and [91] (26 people). Accuracy close to 100% is reported in [83], [85], and [87], where a limited amount of sequences, respectively 64, 87 and 36, have been tested. Regarding radar sensors, the reported accuracy rate can vary between approximately 80% up to a claimed 100%, and the corresponding false alarm rates can vary between 20% down to close to 0%. For example, in [59] accuracy and false alarm rates between 80-100% and 13-20% respectively are reported when using neural networks to perform the classification and different signal processing techniques to extract the micro-Doppler signatures. In [56] the rates are 89-92% and 6-13% respectively when testing different feature extraction techniques on the same group of data, and the same authors propose an improvement in [65] with claimed accuracy up to 95% with 0% false alarm rate by using range as well as Doppler information, but with validation performed on different data. In [67] the authors reported accuracy between 80 and 94% with 0-15% false alarm rates, and then an improvement of their system in [69] with accuracy close to 100% and 0% false alarm rates (although the authors highlighted some limitation of their study, e.g. falls performed only with favorable aspect angles with respect of the radar beam and the use of young subjects rather than actual elderly people). In general, the validation approach on the ad-hoc datasets mentioned above makes it hard to compare between different works, even because some authors do not fully specify the details about actions and subjects used to validate their algorithms, or they do not provide an objective index of their performance, for example in terms of accuracy, or sensitivity and specificity.

The public release of fall detection datasets allows researchers to work with already available data, reducing the time to develop and test algorithms, as the data collection work was already performed by someone else, and providing a common dataset for a fair comparison of different approaches. The largest RGB-D dataset for fall detection is the *SDUFall*, including six actions performed multiple times by 20 subjects. Two other works have used this dataset after Ma et al. [88], who originally collected the data. In particular, the accuracy originally obtained on this dataset was around 87%, and it has been improved to 89% in [90] and almost 92% in [92]. Regarding radar sensors, to authors' best knowledge, openly accessible datasets of radar signatures for fall vs non-fall detection are at the moment not available to the research community. This is an important challenge to be addressed to improve the reliability of the proposed radar-based methods and allow fair comparisons of these methods among each other [22].

Finally, there is a small selection of studies which are closer to real world scenarios, as they involve long-term data acquisition in realistic living environments and hence a better evaluation of the performance. Stone et al. [84] equipped homes for elderly residents with Kinect and acquired data for several days. The dataset includes activities of residents and some falls simulated by actors (a total number of 445). In addition to this data, 9 real falls also occurred and were captured, and 7 out of 9 have been correctly detected by their algorithm. These figures can give an idea of the effectiveness of their method, even if the authors state that it is possible that undetected falls exist in the data. For radar work, the studies presented in [18], [47] stand out for the use of data collected in residential care homes with actual elderly people, although some of the fall events were performed by professional stunt actors. A detailed analysis of the accuracy, sensitivity, and specificity of different classifiers using Wavelet and MFFCs is reported in [18], showing that the accuracy may change significantly between 68-93% with the different approaches, and that the use of Wavelet based classifiers can reduce false alarms for the same detection rate in comparison with MFCCs based classifiers.

B. Outstanding challenges

For both radar and RGB-D systems, open challenges remain to be addressed in order to deploy and employ these systems in practical scenarios. Specific issues related to the use of radar systems arise both from the deployment perspective and from the signal processing perspective [22]. These challenges include:

- the presence of strong scatterers and clutter in indoor environments which may generate multipath and ghost targets, or simply obscure the person to be monitored from the sensor, which can also be a problem for RGB-D sensors;
- the possibility of having pets or other people (e.g. visitors, multiple elderly) moving inside the monitored area, thus complicating the signature and generating false alarms. Again this could potentially be a problem for RGB-D sensors as well;
- the compliance of the selected radar waveforms with directives from the telecommunication regulatory bodies, with potential constraints in terms of the achievable bandwidth and transmitted power, hence limiting the range resolution and the Signal to Noise Ratio (SNR);
- the dependence of the micro-Doppler signature on the cosine of the aspect angle between the velocity vector of the movement and the line-of-sight of the radar, which in some cases can significantly attenuate the signatures and make them unsuitable for feature extraction aimed at fall detection. This is related to the issue of establishing the best location to deploy the radar sensor to avoid this attenuation of the radar Doppler signature, for instance on the ceiling rather than on the wall;
- the possibility to reliably detect a fall, irrespective of the type of movement or activity performed before, and of the dynamics of the fall itself (falling forward or backward, tripping rather than losing balance or consciousness, falling while sitting or standing up from chairs or sofas). This would imply developing fall detection procedures that can

take into account the actual dynamics of elderly people moving, for instance the effects on the radar signatures of using walking assistive devices [57], [58]. However, collecting data from actual elderly in their environments is challenging for practical and ethical reasons, and most of the studies in the literature are based on experiments with younger subjects, apart from references [18], [47]. The question on how far methods and results generated from data collecceted from younger subjects can be actually applied to elderly people remains open for further investigation, and the collection of more valuable data from true patients or elderly is definitely of great interest;

- the difficulty of developing a well-performing general system, capable to take into account the specificity of the person under care and his/her context, such as any physical or cognitive impairment or any specific scenario constraint. This complicates the possibility of training effectively the classifier without using long observation times and large amount of data directly related to the specific person and environment, and makes the case for using unsupervised learning methods, which may provide an advantage at the expense of higher computational complexity.

Other issues specific to the adoption of RGB-D sensors in fall detection applications are those listed below:

- Coverage area and depth sensing range. Differently from wearable devices, vision-based sensors have a limited coverage area, and many sensors may be required to monitor the whole apartment, leading to higher costs of installation. Depth sensors have a limited range, which is usually around 4 meters. Some devices (among which Kinect) can provide data up to 8 meters, but they become quite unreliable beyond 4/5 meters and the skeleton information is not available. Again this may force to use many sensors in a big room or corridor, even if there are no issues directly related to the coverage area, but only because the depth information is not available for the areas located further away than a certain given threshold. Moreover, if RGB data can be obtained with the usage of omnidirectional cameras that can monitor a whole room [117], depth sensors usually have a limited field of view, which is for example 70° x 60° for Kinect v2;

- Occlusions. Vision-based sensors suffer from occlusions, for example from pieces of furniture. The coverage area may be also limited by the presence of some occluding objects, which are temporarily interposed between the subject to be monitored and the sensor. If permanent occlusions can be overcome considering many devices, time-limited occlusions cannot be avoided in principle and this is a limiting factor for these sensors;

- Skeleton data reliability. Many algorithms based on Kinect sensors rely on skeleton data, which can be used to extract the position and posture of the human. However, for the skeleton information to be correctly estimated, the person should be facing the sensor. Some errors may affect the estimation when the aspect angle is different. Moreover, the estimation algorithm can detect some spurious skeletons, that are actually objects. Some techniques to remove noisy skeleton data should be developed and included in the fall detection algorithm to design a reliable fall detection solution.

Crucial for both radar and RGB-D sensors, and more in general for all the technologies investigated for fall detection, is the issue of users' acceptance and compliance. One aspect to consider are privacy concerns [6], [8] as sensors providing the most informative data are usually perceived as very privacy-invasive, with the case limit of video-cameras, which can obviously provide excellent recognition of human activities but are unacceptable in most rooms with high risk of falling, such as bathrooms or bedrooms. As in Figure 5 from [8], simpler sensors are not perceived as a risk for privacy and one could think of using many of them to generate more information up to comparable level with more informative sensors. However, this would increase the complexity and the cost of the installation and deployment of these sensors and make the whole system more complicated to maintain. Another aspect to consider is that elderly people may not be familiar with electronic devices or willing to engage with new technologies, perceived as a disruption of their normal habits and behaviour [6]. The design and development of solutions for fall detection will need to take this into account, and consider the inputs from social sciences, psychology, and primary care disciplines to inform more effective technical choices.

Besides the selection and characterization of a particular type of sensor, another challenge is designing and implementing the overall monitoring system to provide fall detection capabilities. This should be integrated and interconnected with other devices in the indoor environments, e.g. landline phones, smartphones, computers, and various home appliances in an Internet of Things perspective, and offer capabilities for connection towards external entities (e.g. healthcare professionals and first responders in case of an actual fall) and from external entities (e.g. the possibility for relatives to connect to the system to check on the elderly person). A few examples of proposed architectures can be found in [1], [118], [119], and for example the work in [69] shows some effort in presenting the sensor (radar) in its wider application system.

As mentioned in section IV-A, the different systems and algorithms have been often validated on data related to a relatively small number of young and healthy subjects, so a population sample which may be not representative of the final beneficiaries of these systems, and potentially not enough statistically significant. This is also a problem for the validation of other technologies. For example, the work in [9] has reported that only a minority of studies on wearable sensors for human motion analysis used more than 6-8 subjects, and a considerable amount of studies validated their findings only on one subject. Furthermore, the range of age of the subjects was often not reported, or limited within the 20-44 years range, most likely the average age of the researchers and academics themselves. Similar observations can be made for the works described in this paper, where the majority of the studies used a limited number of subjects, with rather limited details provided about their age. It is true that involving elderly people in experimental campaigns increases the complexity of the logistics and

requires the necessary ethical approval, making even more difficult and time consuming the process of collecting large amount of experimental annotated data. A collective effort of the research community to create a large, shared database of signatures could help address this challenge, offer the possibility of thorough validation of proposed algorithms, and foster better collaborations between researchers with different and often complementary expertise [6], [9], [22]. The studies in [18], [19], [47] for radar and in [84] for RGB-D Kinect sensor are an exception, as they used data from actual elderly people, as well as data from professional stunt actors mimicking elderly people.
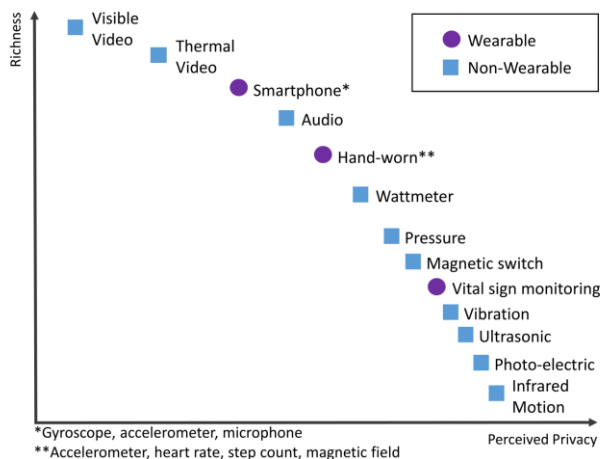

Fig. 5 Perceived user privacy and richness in information for different types of sensors used for indoor monitoring and fall detection [8]

C.   Future trends

Some of the open challenges mentioned above can be addressed and mitigated by the use of multiple cooperating sensors, and this is likely to be investigated as future research work. This could exploit the complementarity of one technology with another, or more simply to extend coverage or exploit multi-perspective views on the area of interest, compared with the use of a single sensor. For radar this approach could involve the use of multistatic systems where different nodes have spatially distributed transmitter and receiver capabilities and can illuminate the area of interest from different aspect angles. This has shown promising results for human micro-Doppler characterization in outdoor scenarios [28], [42]–[45], and it is expected to provide a useful contribution also in indoor scenarios for fall detection [51]. A simpler approach could just use a combination of independent radar sensors (i.e. multiple monostatic sensors rather than a network of multistatic sensors), and then develop algorithms to use their information jointly [51]. The generalization of this concept is the use of multiple heterogeneous sensors, some of them mentioned in the previous sections, such as radar system plus motion sensors [47], radar system plus an array sensor exploiting multipath and NLOS propagation [50], and RGB-D system plus wearable devices [112], [120]. When multiple sensors are used, it is important to make sure that the information generated by them is relevant and not redundant, and to characterize the algorithms to achieve information fusion

[9], [51]. Referring back to Figure 1, this information fusion can happen for example at feature extraction level by using feature samples from all sensors at a centralized classifier, or at decision level by combining the decisions of separate, independent classifiers based on data from each sensor. The best approach to synchronize the behavior of different sensors and to fuse their heterogeneous information (e.g. micro-Doppler radar signatures and skeleton joints from Kinect sensors) remains an open research question, as well as how to select the best type of sensor and its location for a specific scenario. This is also influenced by non-technical aspects regarding the user acceptance of these sensors as mentioned in section IV-B, hence it is somewhat expected that inputs from psychological and behavioral science will also inform more the engineering development and decision process in the future.

The effectiveness of the solutions based on radar and RGB-D sensors for fall detection has been discussed in section IV-A in terms of correct detection and false alarm rates. Besides investigating novel algorithms and data processing techniques to improve this effectiveness, it is also important to consider the easiness of extracting the required features and the related computational power required. These aspects are often not much discussed in the works examined in this paper, as most of the times the data are analyzed off-line, after their collection, and with general purpose platforms (desktop computers) and software (MATLAB or other high level processing software). An interesting exception is the work presented in [69], where the sensor can relay data wirelessly via Zigbee to a separated signal processing station, and the actual fall detection and classification can happen in real-time on this DSP platform (apart from the offline training of the classifier). Several open questions remain, for example how to transmit information from sensors to the processing station (wired or wireless, and what wireless protocol), whether the processing station is separated from the sensors but still co-located where the person to be monitored lives or remotely located on a cloud-platform with information exchanged over the Internet, and whether some pre-processing can be performed locally within each sensor to reduce the amount of data to be transferred at the price of increased complexity of the sensors. The current trends of increasing computational resources available in smaller and cheaper hardware are likely to offer different alternatives to address the aforementioned issues. Specifically for radar sensors, the innovations in the automotive radar industry can be rather significant, with the integration of more and more hardware blocks (including the whole analogue chain and digitization) and processing functionalities (including optimized FFTs to perform range and Doppler estimation for FMCW radar) in single chips, allowing smaller sensors but capable of providing more information. Furthermore, the development in the field of deep learning can deliver a step change in radar signal processing, with the possibility of using convolutional neural networks to bypass the feature extraction step (as shown in the very preliminary results in [4], [59]), and achieve systems capable of continuous learning and adaptation to changes in the operational scenario (e.g. changes of the person's habits, new furniture position, moving in a different

house or room). There is also ongoing research on efficient implementation of deep neural networks on relatively simple and inexpensive hardware, for example using binary networks or low precision weights and activation functions on FPGA boards rather than on computer clusters [121], [122]. The possiblity that such powerful classification tools can be simply implemented will contribute to better integration of the sensors and their data processing and to the development of improved fall detection and classification algorithms.

V. CONCLUSION

This paper presented a comprehensive review of recent works in the field of fall detection systems based on radar and RGB-D sensors. Fall detection has become a progressively relevant research topic in the past few years, as the number of elderly people living alone and at risk of falling is increasing, posing a significant societal issue with related health hazards and economic costs. Radar and RGB-D sensors offer the advantage of providing contactless and non-invasive monitoring capabilities, whereby the sensors may be simply deployed in the area to be monitored (e.g. in the corner of an indoor environment), with no need for the people to wear or carry any device, or change their normal habits and behavior, and no privacy concerns which could be raised by for example video based systems. These sensors will provide a significant contribution to the development of reliable fall detection systems, complementing other sensing technologies such as wearable devices to provide overall improved monitoring performance.

Details on the different sensors' configurations, algorithms, and performance evaluation have been provided in the previous sections, as well as the analysis of the outstanding challenges to be addressed for practical deployment and use of these systems in realistic environments. Bringing together multidisciplinary expertise is expected to be an important step to go beyond the proof of concept validation of the different methods and algorithms on a small set of subjects, in more or less controlled conditions. Expertise should include the designing and development of the different types of sensors, the integration of sensors with the wider home network of devices in an Internet of Things perspective, as well as with external stakeholders such as carers or first responders, the competence from medical professionals to infer health information from the activity patterns extracted from sensors information, and inputs from social sciences and psychology experts to address issues of users' acceptance of fall detection systems and sensors.

REFERENCES

[1] Z. Pang, L. Zheng, J. Tian, S. Kao-Walter, E. Dubrova, and Q. Chen, "Design of a terminal solution for integration of in-home health care devices and services towards the Internet-of-Things," *Enterp. Inf. Syst.*, vol. 9, no. 1, pp. 86–116, Jan. 2015.

[2] World Health Organisation, "WHO Global Report on Falls Prevention in Older Age.," 2008.

[3] M. Terroso, N. Rosa, A. Torres Marques, and R. Simoes, "Physical consequences of falls in the elderly: a literature review from 1995 to 2010," *Eur. Rev. Aging Phys. Act.*, vol. 11, no. 1, pp. 51–59, 2014.

[4] B. Jokanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," *2016 IEEE Radar Conference (RadarConf)*. pp.

[5] 1–6, 2016.

[5] A. Koenig, C. F. Crispim Junior, A. Derreumaux, G. Bensadoun, P. D. Petit, F. Bremond, R. David, F. Verhey, P. Aalten, and P. Robert, "Validation of an Automatic Video Monitoring System for the Detection of Instrumental Activities of Daily Living in Dementia Patients," *J. Alzheimer's Dis.*, vol. 44, no. 2, pp. 675–685, 2015.

[6] R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," *Biomed. Eng. Online*, vol. 12, no. 1, pp. 1–24, 2013.

[7] F. Erden, S. Velipasalar, A. Z. Alkar, and A. E. Cetin, "Sensors in Assisted Living: A survey of signal and image processing methods," *IEEE Signal Processing Magazine*, vol. 33, no. 2. pp. 36–44, 2016.

[8] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer, "Monitoring Activities of Daily Living in Smart Homes: Understanding human behavior," *IEEE Signal Processing Magazine*, vol. 33, no. 2. pp. 81–94, 2016.

[9] I. H. Lopez-Nava and M. M. Angelica, "Wearable Inertial Sensors for Human Motion Analysis: A review," *IEEE Sensors Journal*, vol. PP, no. 99. p. 1, 2016.

[10] S. C. Mukhopadhyay, "Wearable Sensors for Human Activity Monitoring: A Review," *IEEE Sensors Journal*, vol. 15, no. 3. pp. 1321–1330, 2015.

[11] O. Aziz and S. N. Robinovitch, "An Analysis of the Accuracy of Wearable Sensors for Classifying the Causes of Falls in Humans," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 6. pp. 670–676, 2011.

[12] T. Shany, S. J. Redmond, M. R. Narayanan, and N. H. Lovell, "Sensors-Based Wearable Systems for Monitoring of Human Movement and Falls," *IEEE Sensors Journal*, vol. 12, no. 3. pp. 658–670, 2012.

[13] T. R. Bennett, J. Wu, N. Kehtarnavaz, and R. Jafari, "Inertial Measurement Unit-Based Wearable Computers for Assisted Living Applications: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 33, no. 2. pp. 28–35, 2016.

[14] M. Mubashir, L. Shao, and L. Seed, "A survey on fall detection: Principles and approaches," *Neurocomputing*, vol. 100, pp. 144–152, Jan. 2013.

[15] Z. Zhang, C. Conly, and V. Athitsos, "A Survey on Vision-based Fall Detection," in *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2015, p. 46:1--46:7.

[16] G. Koshmak, A. Loufti, and M. Linden, "Challenges and Issues in Multisensor Fusion Approach for Fall Detection: Review Paper," *Hindawi J. Sensors*, vol. 2016, no. ID 6931789, p. 12, 2016.

[17] K. Chaccour, R. Darazi, A. H. El Hassani, and E. Andrès, "From Fall Detection to Fall Prevention: A Generic Classification of Fall-Related Systems," *IEEE Sensors Journal*, vol. 17, no. 3. pp. 812–822, 2017.

[18] B. Y. Su, K. C. Ho, M. J. Rantz, and M. Skubic, "Doppler Radar Fall Activity Detection Using the Wavelet Transform," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 3. pp. 865–875, 2015.

[19] L. Liu, M. Popescu, M. Skubic, M. Rantz, T. Yardibi, and P. Cuddihy, "Automatic fall detection based on Doppler radar motion signature," *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. pp. 222–225, 2011.

[20] F. Ahmad, R. Narayanan, and D. Schreurs, "Application of Radar to Remote Patient Monitoring and Eldercare," *IET radar, Sonar Navig.*, vol. 9, no. 2, p. 115, 2015.

[21] F. Ahmad, A. E. Cetin, K. C. D. Ho, and J. Nelson, "Signal Processing for Assisted Living: Developments and Open Problems [From the Guest Editors]," *IEEE Signal Processing Magazine*, vol. 33, no. 2. pp. 25–26, 2016.

[22] M. G. Amin, Y. D. Zhang, F. Ahmad, and K. C. D. Ho, "Radar Signal Processing for Elderly Fall Detection: The future for in-home monitoring," *IEEE Signal Processing Magazine*, vol. 33, no. 2. pp. 71–80, 2016.

[23] J. Zabalza, C. Clemente, G. Di Caterina, J. Ren, J. J. Soraghan, and S. Marshall, "Robust PCA micro-doppler classification using SVM on embedded systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 3, pp. 2304–2312, 2014.

[24] B. Jokanovic, M. Amin, F. Ahmad, and B. Boashash, " Radar fall detection using principal component analysis ," 2016, vol. 9829, pp. 982916–982919.

[25] S. Z. Gürbüz, B. Erol, B. Cağliyan, and B. Tekeli, "Operational

[26] B. Tekeli, S. Z. Gurbuz, and M. Yuksel, "Information-Theoretic Feature Selection for Human Micro-Doppler Signature Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5. pp. 2749–2762, 2016.

[27] B. Frénay, G. Doquire, and M. Verleysen, "Theoretical and empirical study on the potential inadequacy of mutual information for feature selection in classification," *Neurocomputing*, vol. 112, pp. 64–78, Jul. 2013.

[28] F. Fioranelli, M. Ritchie, and H. Griffiths, "Performance Analysis of Centroid and SVD Features for Personnel Recognition Using Multistatic Micro-Doppler," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5. pp. 725–729, 2016.

[29] M. Skolnik, *Radar Handbook, Third Edition*. McGraw-Hill Education, 2008.

[30] E. Piuzzi, P. D'Atanasio, S. Pisa, E. Pittella, and A. Zambotti, "Complex Radar Cross Section Measurements of the Human Body for Breath-Activity Monitoring Applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 8. pp. 2247–2258, 2015.

[31] V. C. Chen, "Doppler signatures of radar backscattering from objects with micro-motions," *IET Signal Processing*, vol. 2, no. 3. pp. 291–300, 2008.

[32] V. C. Chen, D. Tahmoush, and W. J. Miceli, Eds., *Radar Micro-Doppler Signatures: Processing and Applications*. Institution of Engineering and Technology, 2014.

[33] D. Tahmoush, "Review of micro-Doppler signatures," *IET Radar, Sonar & Navigation*, vol. 9, no. 9. pp. 1140–1146, 2015.

[34] Y. Kim and T. Moon, "Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1. pp. 8–12, 2016.

[35] D. Tahmoush and J. Silvious, "Remote detection of humans and animals," *2009 IEEE Applied Imagery Pattern Recognition Workshop (AIPR 2009)*. pp. 1–8, 2009.

[36] X. Shi, F. Zhou, L. Liu, B. Zhao, and Z. Zhang, "Textural feature extraction based on time-frequency spectrograms of humans and vehicles," *IET Radar, Sonar & Navigation*, vol. 9, no. 9. pp. 1251–1259, 2015.

[37] Y. Kim, S. Ha, and J. Kwon, "Human Detection Using Doppler Radar Based on Physical Characteristics of Targets," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 2. pp. 289–293, 2015.

[38] Y. Kim and H. Ling, "Human Activity Classification Based on Micro-Doppler Signatures Using a Support Vector Machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5. pp. 1328–1337, 2009.

[39] R. M. Narayanan and M. Zenaldin, "Radar micro-Doppler signatures of various human activities," *IET Radar, Sonar & Navigation*, vol. 9, no. 9. pp. 1205–1215, 2015.

[40] D. P. Fairchild and R. M. Narayanan, "Multistatic micro-doppler radar for determining target orientation and activity classification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 1. pp. 512–521, 2016.

[41] S. Björklund, H. Petersson, and G. Hendeby, "Features for micro-Doppler based activity classification," *IET Radar, Sonar & Navigation*, vol. 9, no. 9. pp. 1181–1187, 2015.

[42] F. Fioranelli, M. Ritchie, and H. Griffiths, "Centroid features for classification of armed/unarmed multiple personnel using multistatic human micro-Doppler," *IET Radar, Sonar & Navigation*. Apr-2016.

[43] F. Fioranelli, M. Ritchie, and H. Griffiths, "Multistatic human micro-Doppler classification of armed/unarmed personnel," *IET Radar, Sonar & Navigation*, vol. 9, no. 7. pp. 857–865, 2015.

[44] F. Fioranelli, M. Ritchie, and H. Griffiths, "Classification of Unarmed/Armed Personnel Using the NetRAD Multistatic Radar for Micro-Doppler and Singular Value Decomposition Features," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 9, pp. 1933–1937, 2015.

[45] F. Fioranelli, M. Ritchie, and H. Griffiths, "Aspect angle dependence and multistatic data fusion for micro-Doppler classification of armed/unarmed personnel," *IET Radar, Sonar & Navigation*, vol. 9, no. 9. Institution of Engineering and Technology, pp. 1231–1239, 2015.

[46] R. Ricci and A. Balleri, "Recognition of humans based on radar micro-Doppler shape spectrum features," *IET Radar, Sonar & Navigation*, vol. 9, no. 9. pp. 1216–1223, 2015.

[47] L. Liu, M. Popescu, M. Skubic, and M. Rantz, "An automatic fall detection framework using data fusion of Doppler radar and motion sensor network," *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. pp. 5940–5943, 2014.

[48] A. Gadde, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Fall detection and classifications based on time-scale radar signal characteristics," in *Proc. SPIE 9077, radar Sensor Technology XVIII, 907712*, 2014, vol. 9077, pp. 907712–907719.

[49] S. Liu, Z. Zeng, Y. D. Zhang, T. Fan, T. Shan, and R. Tao, "Automatic human fall detection in fractional fourier domain for assisted living," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 799–803, 2016.

[50] J. Hong, S. Tomii, and T. Ohtsuki, "Cooperative fall detection using Doppler radar and array sensor," *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. pp. 3492–3496, 2013.

[51] S. Tomii and T. Ohtsuki, "Falling detection using multiple Doppler sensors," in *2012 IEEE 14th International Conference on e-Health Networking, Applications and Services, Healthcom 2012*, 2012, pp. 196–201.

[52] J. Hong and T. Ohtsuki, "A state classification method based on space-time signal processing using SVM for wireless monitoring systems," *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*. pp. 2229–2233, 2011.

[53] L. R. Rivera, E. Ulmer, Y. D. Zhang, W. Tao, and M. G. Amin, "Radar-based fall detection exploiting time-frequency features," *Signal and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on*. pp. 713–717, 2014.

[54] Q. Wu, Y. D. Zhang, W. Tao, and M. G. Amin, "Radar-based fall detection based on Doppler time-frequency signatures for assisted living," *IET Radar, Sonar & Navigation*, vol. 9, no. 2. pp. 164–172, 2015.

[55] B. Jokanovic, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Multi-window time-frequency signature reconstruction from undersampled continuous-wave radar measurements for fall detection," *IET Radar, Sonar & Navigation*, vol. 9, no. 2. pp. 173–183, 2015.

[56] B. Erol, M. Amin, F. Ahmad, and B. Boashash, " Radar fall detectors: a comparison ," 2016, vol. 9829, pp. 982918–982919.

[57] M. G. Amin, F. Ahmad, Y. D. Zhang, and B. Boashash, "Human gait recognition with cane assistive device using quadratic time-frequency distributions," *IET Radar, Sonar & Navigation*, vol. 9, no. 9. pp. 1224–1230, 2015.

[58] S. Z. Gürbüz, C. Clemente, A. Balleri, and J. J. Soraghan, "Micro-Doppler-based in-home aided and unaided walking recognition with multiple radar and sonar systems," *IET Radar, Sonar Navig.*, Jun. 2016.

[59] B. Jokanovic, M. G. Amin, and F. Ahmad, "Effect of data representations on deep learning in fall detection," *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. pp. 1–5, 2016.

[60] M. Wu, X. Dai, Y. D. Zhang, B. Davidson, M. G. Amin, and J. Zhang, "Fall Detection Based on Sequential Modeling of Radar Signal Time-Frequency Features," *2013 IEEE International Conference on Healthcare Informatics*. pp. 169–174, 2013.

[61] B. G. Mobasseri and M. G. Amin, "A time-frequency classifier for human gait recognition," 2009, vol. 7306, pp. 730628–730629.

[62] Z. A. Khan and W. Sohn, "Feature extraction and dimensions reduction using R transform and Principal Component Analysis for abnormal human activity recognition," *2010 6th International Conference on Advanced Information Management and Service (IMS)*. pp. 253–258, 2010.

[63] C. Clemente, A. W. Miller, and J. J. Soraghan, "Robust principal component analysis for micro-doppler based automatic target recognition," in *3rd IMA conference on Mathematics in Defence*, 2013.

[64] A. G. Stove, "A Doppler-Based Target Classifier Using Linear Discriminants and Principal Components," *2006 IET Seminar on High Resolution Imaging and Target Classification*. pp. 107–125, 2006.

[65] B. Erol, M. Amin, Z. Zhou, and J. Zhang, "Range information for reducing fall false alarms in assisted living," *2016 IEEE Radar Conference (RadarConf)*. pp. 1–6, 2016.

[66] H. Wang, L. Ren, E. Mao, and A. E. Fathy, "Phase based motion characteristics measurement for fall detection by using stepped-

frequency continuous wave radar," *2016 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireleSS)*. pp. 43–45, 2016.

[67] M. Mercuri, P. J. Soh, G. Pandey, P. Karsmakers, G. A. E. Vandenbosch, P. Leroux, and D. Schreurs, "Analysis of an Indoor Biomedical Radar-Based System for Health Monitoring," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 5. pp. 2061–2068, 2013.

[68] M. Mercuri, P. J. Soh, X. Zheng, P. Karsmakers, G. A. E. Vandenbosch, P. Leroux, and D. Schreurs, "Analysis of a fall detection radar placed on the ceiling and wall," *2014 Asia-Pacific Microwave Conference*. pp. 947–949, 2014.

[69] C. Garripoli, M. Mercuri, P. Karsmakers, P. J. Soh, G. Crupi, G. A. E. Vandenbosch, C. Pace, P. Leroux, and D. Schreurs, "Embedded DSP-Based Telehealth Radar System for Remote In-Door Fall Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1. pp. 92–101, 2015.

[70] T. D'Orazio, R. Marani, V. Renò, and G. Cicirelli, "Recent trends in gesture recognition: how depth data has improved classical approaches," *Image Vis. Comput.*, vol. 52, pp. 56–72, 2016.

[71] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," *Stud. Comput. Intell.*, vol. 411, pp. 119–135, 2013.

[72] A. Corti, S. Giancola, G. Mainetti, and R. Sala, "A metrological characterization of the Kinect V2 time-of-flight camera," *Rob. Auton. Syst.*, vol. 75, pp. 584–594, Jan. 2016.

[73] E. Cippitelli, S. Gasparrini, E. Gambi, and S. Spinsante, "Quality of Kinect Depth Information for Passive Posture Monitoring," in *Ambient Assisted Living: Italian Forum 2013*, S. Longhi, P. Siciliano, M. Germani, and A. Monteriù, Eds. Cham: Springer International Publishing, 2014, pp. 107–116.

[74] E. Cippitelli, S. Gasparrini, S. Spinsante, E. Gambi, F. Verdini, L. Burattini, F. Di Nardo, and S. Fioretti, "Validation of an optimized algorithm to use Kinect in a non-structured environment for Sit-to-Stand analysis," *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. pp. 5078–5081, 2015.

[75] C. Rougier, E. Auvinet, J. Rousseau, M. Mignotte, and J. Meunier, "Fall Detection from Depth Map Video Sequences," in *Toward Useful Services for Elderly and People with Disabilities: 9th International Conference on Smart Homes and Health Telematics, ICOST 2011, Montreal, Canada, June 20-22, 2011. Proceedings*, B. Abdulrazak, S. Giroux, B. Bouchard, H. Pigot, and M. Mokhtari, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 121–128.

[76] J. Zhao, J. Katupitiya, and J. Ward, "Global Correlation Based Ground Plane Estimation Using V-Disparity Image," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 2007, pp. 529–534.

[77] S. Gasparrini, E. Cippitelli, S. Spinsante, and E. Gambi, "A depth-based fall detection system using a Kinect® sensor," *Sensors (Switzerland)*, vol. 14, no. 2, 2014.

[78] D. Liciotti, G. Massi, E. Frontoni, A. Mancini, and P. Zingaretti, "Human activity analysis for in-home fall risk assessment," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, 2015, pp. 284–289.

[79] L. Yang, Y. Ren, and W. Zhang, "3D Depth Image Analysis for Indoor Fall Detection of Elderly People," *Digit. Commun. Networks*, vol. 2, no. 1, pp. 24–34, 2016.

[80] G. Mastorakis and D. Makris, "Fall detection system using Kinect's infrared sensor," *J. Real-Time Image Process.*, vol. 9, no. 4, pp. 635–646, 2012.

[81] "OpenNI." http://structure.io/openni.

[82] V. Bevilacqua, N. Nuzzolese, D. Barone, M. Pantaleo, M. Suma, D. D'Ambruoso, A. Volpe, C. Loconsole, and F. Stroppa, "Fall detection in indoor environment with kinect sensor," in *Innovations in Intelligent Systems and Applications (INISTA) Proceedings, 2014 IEEE International Symposium on*, 2014, pp. 319–324.

[83] A. T. Nghiem, E. Auvinet, and J. Meunier, "Head detection using Kinect camera and its application to fall detection," in *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*, 2012, pp. 164–169.

[84] E. E. Stone and M. Skubic, "Fall Detection in Homes of Older Adults Using the Microsoft Kinect," *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 1, pp. 290–301, 2015.

[85] Z. Zhang, W. Liu, V. Metsis, and V. Athitsos, "A Viewpoint-Independent Statistical Method for Fall Detection," *Int. Conf. Pattern Recognit.*, no. Icpr, pp. 3626–3630, 2012.

[86] M. Kepski and B. Kwolek, "Unobtrusive Fall Detection at Home Using Kinect Sensor," in *Computer Analysis of Images and Patterns: 15th International Conference, CAIP 2013, York, UK, August 27-29, 2013, Proceedings, Part I*, R. Wilson, E. Hancock, A. Bors, and W. Smith, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 457–464.

[87] P. Bilski, P. Mazurek, and J. Wagner, "Application of k Nearest Neighbors Approach to the fall detection of elderly people using depth-based sensors," in *Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2015 IEEE 8th International Conference on*, 2015, vol. 2, pp. 733–739.

[88] X. Ma, H. Wang, B. Xue, M. Zhou, B. Ji, and Y. Li, "Depth-based human fall detection via shape features and improved extreme learning machine," *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 6, pp. 1915–1922, 2014.

[89] F. Mokhtarian, "Silhouette-based isolated object recognition through curvature scale space," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 5, pp. 539–544, 1995.

[90] M. Aslan, A. Sengur, Y. Xiao, H. Wang, M. C. Ince, and X. Ma, "Shape feature encoding via Fisher Vector for efficient fall detection in depth-videos," *Appl. Soft Comput.*, vol. 37, pp. 1023–1028, 2015.

[91] A. Dubois and F. Charpillet, "Human activities recognition with RGB-Depth camera using HMM," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 4666–4669.

[92] E. Akagunduz, M. Aslan, A. Sengur, H. Wang, and M. Ince, "Silhouette Orientation Volumes for Efficient Fall Detection in Depth Videos," *IEEE J. Biomed. Heal. Informatics*, vol. PP, no. 99, p. 1, 2016.

[93] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri, "Actions as Space-Time Shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2247–2253, 2007.

[94] R. Planinc and M. Kampel, "Introducing the use of depth data for fall detection," *Pers. Ubiquitous Comput.*, vol. 17, no. 6, pp. 1063–1072, 2013.

[95] R. Planinc and M. Kampel, "Robust Fall Detection by Combining 3D Data and Fuzzy Logic," in *Computer Vision - ACCV 2012 Workshops: ACCV 2012 International Workshops, Daejeon, Korea, November 5-6, 2012, Revised Selected Papers, Part II*, J.-I. Park and J. Kim, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 121–132.

[96] C. K. Lee and V. Y. Lee, "Fall Detection System Based on Kinect Sensor Using Novel Detection and Posture Recognition Algorithm," in *Inclusive Society: Health and Wellbeing in the Community, and Care at Home: 11th International Conference on Smart Homes and Health Telematics, ICOST 2013, Singapore, June 19-21, 2013. Proceedings*, J. Biswas, H. Kobayashi, L. Wong, B. Abdulrazak, and M. Mokhtari, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 238–244.

[97] C. Kawatsu, J. Li, and C. J. Chung, "Development of a fall detection system with microsoft kinect," *Adv. Intell. Syst. Comput.*, vol. 208 AISC, pp. 623–630, 2013.

[98] S.-W. Yang and S.-K. Lin, "Fall detection for multiple pedestrians using depth image processing technique," *Comput. Methods Programs Biomed.*, vol. 114, no. 2, pp. 172–182, 2014.

[99] Z.-P. Bian, J. Hou, L.-P. Chau, and N. Magnenat-Thalmann, "Fall detection based on body part tracking using a depth camera.," *IEEE J. Biomed. Heal. informatics*, vol. 19, no. 2, pp. 430–9, 2015.

[100] N. Noury, A. Fleury, P. Rumeau, A. K. Bourke, G. O. Laighin, V. Rialle, and J. E. Lundy, "Fall detection - Principles and Methods," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 1663–1666.

[101] A. Amini, K. Banitsas, and J. Cosmas, "A comparison between heuristic and machine learning techniques in fall detection using Kinect v2," in *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2016, pp. 1–6.

[102] R. Dubey, B. Ni, and P. Moulin, "A Depth Camera Based Fall Recognition System for the Elderly," in *Image Analysis and Recognition: 9th International Conference, ICIAR 2012, Aveiro, Portugal, June 25-27, 2012. Proceedings, Part II*, A. Campilho and M. Kamel, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 106–113.

[103] B. Ni, G. Wang, and P. Moulin, "RGBD-HuDaAct: A color-depth video database for human daily activity recognition," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, 2011, pp. 1147–1153.

[104] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[105] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 3, pp. 257–267, 2001.

[106] A. Davari, T. Aydin, and T. Erdem, "Automatic fall detection for elderly by using features extracted from skeletal data," in *Electronics, Computer and Computation (ICECCO), 2013 International Conference on*, 2013, pp. 127–130.

[107] C. Zhang, Y. Tian, and E. Capezuti, "Privacy Preserving Automatic Fall Detection for Elderly Using RGBD Cameras," *Int. Conf. Comput. Help. People with Spec. Needs*, pp. 625–633, 2012.

[108] X. Dai, M. Wu, B. Davidson, M. Mahoor, and J. Zhang, "Image-Based Fall Detection with Human Posture Sequence Modeling," in *Healthcare Informatics (ICHI), 2013 IEEE International Conference on*, 2013, pp. 376–381.

[109] R. Alazrai, A. Zmily, and Y. Mowafi, "Fall detection for elderly using anatomical-plane-based representation," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 5916–5919.

[110] M. Firman, "RGBD Datasets: Past, Present and Future," *CoRR*, vol. abs/1604.0, 2016.

[111] J. Zhang, W. Li, P. O. Ogunbona, P. Wang, and C. Tang, "RGB-D-based action recognition datasets: A survey," *Pattern Recognit.*, vol. 60, pp. 86–105, 2016.

[112] S. Gasparrini, E. Cippitelli, E. Gambi, S. Spinsante, J. Wåhslén, I. Orhan, and T. Lindh, "Proposal and Experimental Evaluation of Fall Detection Solution Based on Wearable and Depth Data Fusion," in *ICT Innovations 2015 : Emerging Technologies for Better Living*, vol. 399, S. Loshkovska and S. Koceski, Eds. Cham: Springer International Publishing, 2016, pp. 99–108.

[113] B. Kwolek and M. Kepski, "Human fall detection on embedded platform using depth maps and wireless accelerometer," *Comput. Methods Programs Biomed.*, vol. 117, no. 3, pp. 489–501, 2014.

[114] Z. Cheng, L. Qin, Y. Ye, Q. Huang, and Q. Tian, "Human Daily Action Analysis with Multi-view and Color-Depth Data," in *Computer Vision -- ECCV 2012. Workshops and Demonstrations: Florence, Italy, October 7-13, 2012, Proceedings, Part II*, A. Fusiello, V. Murino, and R. Cucchiara, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 52–61.

[115] C. Zhang and Y. Tian, "RGB-D Camera-based Daily Living Activity Recognition," *J. Comput. Vis. Image Process.*, vol. 2, p. 12, 2012.

[116] Z. Zhang, C. Conly, and V. Athitsos, "Evaluating Depth-Based Computer Vision Methods for Fall Detection under Occlusions," in *Advances in Visual Computing: 10th International Symposium, ISVC 2014, Las Vegas, NV, USA, December 8-10, 2014, Proceedings, Part II*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, R. McMahan, J. Jerald, H. Zhang, S. M. Drucker, C. Kambhamettu, M. El Choubassi, Z. Deng, and M. Carlson, Eds. Cham: Springer International Publishing, 2014, pp. 196–207.

[117] M. l. Wang, C. c. Huang, and H. y. Lin, "An Intelligent Surveillance System Based on an Omnidirectional Vision Sensor," in *2006 IEEE Conference on Cybernetics and Intelligent Systems*, 2006, pp. 1–6.

[118] J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield, S. T. C. Wong, and G. Z. Yang, "Big Data for Health," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4. pp. 1193–1208, 2015.

[119] C. Arvindkumar, N. Jaiswal, J. Tailor, Z. Liu, C. Parihar, J. Zhou, K. Holmes, S. Dean, and M. Islam, "Non-intrusive user-oriented interactive fall-detection system for seniors," *Humanitarian Technology Conference (IHTC2015), 2015 IEEE Canada International*. pp. 1–4, 2015.

[120] M. Kepski and B. Kwolek, "Fall detection using ceiling-mounted 3d depth camera," *Int. Conf. VISAPP, pages II*, pp. 640–647, 2014.

[121] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio, "Quantized Neural Networks: Training Neural Networks with Low Precision Weights and Activations," *CoRR*, vol. abs/1609.0, 2016.

[122] G. Lacey, G. W. Taylor, and S. Areibi, "Deep Learning on FPGAs: Past, Present, and Future," *CoRR*, vol. abs/1602.0, 2016.