Fall Detection Using Deep Learning in Range-Doppler Radars

BRANKA JOKANOVIĆ, Student Member, IEEE
MOENESS AMIN, Fellow, IEEE
Villanova University, Villanova, PA USA

In this paper, we propose an approach that uses deep learning to detect a human fall. The proposed approach automatically captures the intricate properties of the radar returns. In order to minimize false alarms, we fuse information from both the time-frequency and range domains. Experimental data is used to demonstrate the superiority of the deep learning based approach in comparison with the principal component analysis method and those methods incorporating predefined physically interpreted features.

Manuscript received January 27, 2017; revised May 22, 2017; released for publication July 18, 2017. Date of publication August 15, 2017; date of current version February 7, 2018.

DOI. No. 10.1109/TAES.2017.2740098

Refereeing of this contribution was handled by R. M. Narayanan.

This work was supported by NPRP under Grant # NPRP 6-680-2-282 from the Qatar National Research Fund (a member of Qatar Foundation).

Authors' addresses: B. Jokanović and M. Amin are with the Center for Advanced Communications, Villanova University, Villanova, PA 19085 USA, E-mail: (branka.jokanovic@villanova.edu; moeness.amin@villanova.edu). *(Corresponding author: Branka Jokanović.)*

I. INTRODUCTION

Falls are a major public health problem worldwide. According to the Centers for Disease Control and Prevention, falls were the leading cause of unintentional injury death for people aged 65 and older in 2013 [1]. World Health Organization (WHO) defines fall as "an event which results in a person coming to rest inadvertently on the ground or other lower level" [2]. Falls can cause fatal or nonfatal injuries. WHO estimates that each year around 420 000 falls have fatal outcome. This number makes fall the second leading cause of unintentional injury death [1]–[4].

Even though various fall detectors have been introduced in the literature, and some are available on the market, none has been tagged as a superior approach [5]–[20]. It is estimated that almost half of elderly who experience a fall does not report it to their health care provider [1]. This fact promotes the use of nonwearable devices that can provide remote monitoring. Radar has been proven successful in monitoring human motions and, unlike camera and infrared-based solutions, its function is not impeded by lighting or temperature conditions [21]–[29].

Radar signal returns, corresponding to human gross-motor activities, are nonstationary in nature. For these signals, the time-frequency (TF) domain is typically used, attributing to its ability to reveal velocities, accelerations, and higher-order Doppler terms of limbs and various human body parts in motion. The range is another important information, which can be obtained from multifrequency or wideband radar returns [30], [31]. It may be used to reveal human location versus time, and with fine range resolution, it can tag the body main scatterers to their respective Doppler signatures. In general, velocity and range are considered fundamental to human motion classification.

Traditional radar approaches for motion classification and detection predicate on three steps: Data processing, feature extraction, and classification. Extracting a predefined feature set in single and joint-variable domains has been a common practice in feature extractions for motion classification [32]–[35]. However, deciding on relevant features can be a tedious task. Additionally, these features exhibit large variations, depending on the individuals being monitored in terms of size, habits, and health conditions. One way to overcome these issues is to devise an automated approach that learns and captures the intricate properties of the human motion signatures in different domains [36], [37]. This can be accomplished via deep learning which provides a different paradigm to motion classification and can work independently of or in tandem with other feature selection methods. Learning of the most prominent characteristics is typically done over several layers.

In this paper, we propose a deep learning based fall detection using frequency modulated CW (FMCW) radars. A FMCW radar, also called a range-Doppler radar, allows us to analyze target returns in multiple joint-variable domains and monitor both the Doppler shift and the range. The work in [36] reported initial results when deep learning is applied only in the time-frequency domain. In this

work, we focus on two domains for analyzing radar returns, namely, the time-frequency domain and the range domain, and propose different combination methods for integrating the information present in these domains. The deep learning scheme includes two stacked sparse autoencoders and a logistic regression classifier. Experimental results demonstrate that deep learning improves fall detection over traditional approaches, and that incorporating the range along with the Doppler signatures, provided by range-Doppler radars, yields higher success classification rates than the case when using a single domain.

The paper is organized as follows. Section II briefly describes the radar system and the domains used to represent the input data. Additionally, a general scheme for fall detection is described. The deep learning based approach for fall detection is discussed in Section III. Section IV presents different combinations for incorporating the range and TF domain information. Experimental results are shown in Section V, while the conclusion is given in Section VI.

II. DATA PREPARATION

In this section, we provide background on the radar system and describe the domains that are employed. Traditional radar approaches for motion detection are also discussed.

A. Radar System

Continuous wave (CW) radar has been a radar of choice for many years when observing human motions. It represents one of the simplest hardware implementations in which a fixed frequency signal is transmitted and reflections of that signal from objects in the field of view are received and mixed with the transmitted carrier. A CW radar detects the radial velocity of a moving object, which alters the frequency of signal it reflects, known as the Doppler shift. Additional frequency modulations occur due to the vibrations or rotations of target parts. These modulations generate sidebands around the main Doppler shift and they represent the micro-Doppler effect. Objects moving closer to or away from a CW radar cause the reflected signals to move up and down in frequency, respectively, compared to the signal transmitted by the radar. The Doppler component could be only a few Hz shift on top of a multi-GHz carrier signal. The Doppler shifts carry information about the velocity of the human torso, limbs, and other parts of human body.

Unmodulated CW radar can monitor velocity, but cannot measure distance of the target. Wideband radar, including FMCW, can overcome this shortcoming.

B. Data Representation Domains

Since the human body in motion exhibits different velocities that change over time, it is necessary to represent the data in such a way that the time-varying frequency characteristics of these radar returns can be clearly manifested. Due to the relation between velocity and Doppler frequency, the TF domain has been traditionally used to represent the backscattering signals from human subjects [38]–[41]. One

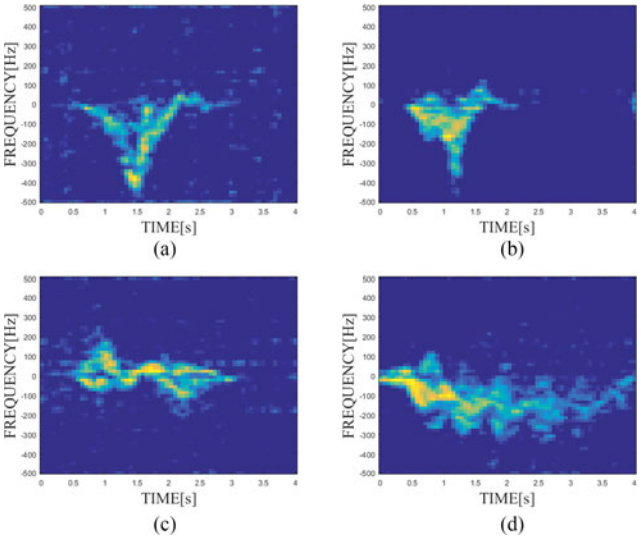


Fig. 1. Spectrograms of four human motions: (a) Fall, (b) sit, (c) bend, and (d) walk.

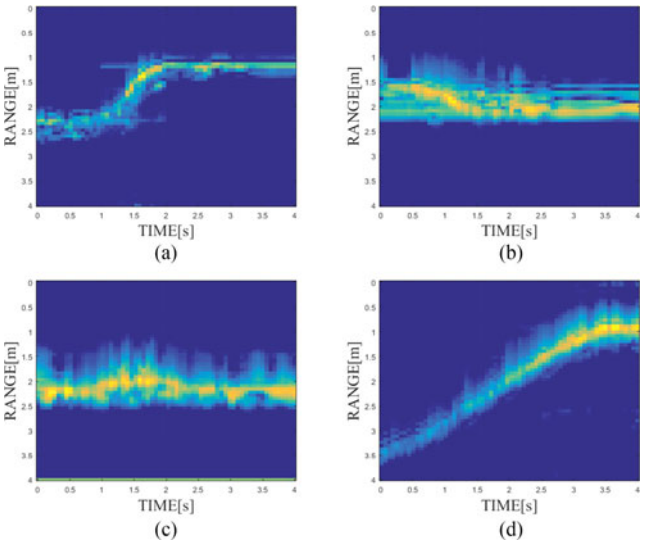


Fig. 2. Range maps of four human motions: (a) Fall, (b) sit, (c) bend, and (d) walk.

of the most commonly used TF signal representation is the spectrogram which depicts the distribution of the signal power over time and frequency. The spectrogram of a discrete signal $s(n), n = 0 \ldots N-1$ is defined as

$$SPEC(n,k) = \left| \sum_{m=0}^{N-1} w(m) s(n-m) e^{-j2\pi km/N} \right|^2 \quad (1)$$

where $w(m)$ is a window function. High resolution quadratic time-frequency distributions (TFDS) can be used in lieu of spectrograms and can improve performance at the cost of increased complexity [42].

Target range and target motion translation extent are also considered useful parameters for motion classifications [43]. The target joint slow-time range representation is referred to as the range map. Figs. 1 and 2 show the spectrograms and range maps of a fall and three most

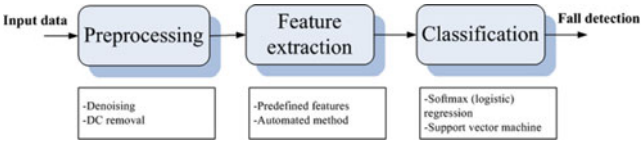


Fig. 3. General scheme for fall detection.

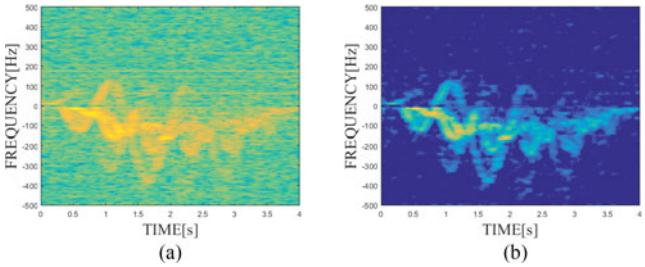


Fig. 4. Spectrogram of a walk before and after preprocessing.

common motions: Sitting, bending, and walking. These motions were also chosen because they are a common cause of fall false alarms. It is evident that each motion professes different levels of distinction in the observed domains. That is, the difference between two motions can be significant in one representation domain and insignificant in another. The following remarks are in order.

1) Time-frequency domain: Walking is most distinguishable in the TF domain and does not appear similar to any of the other motions. On the other hand, other motions, i.e., falling, sitting, and bending can be confused due to the resemblance of their respective TF signatures.
2) Range map: Sitting and bending range maps bear little similarity to those of falling and walking. However, distinguishing sitting from bending or walking from falling can be a difficult task for a classifier if only the range maps are employed.

C. General Scheme for Fall Detection

Fig. 3 depicts a general scheme for fall detection. The first step is to preprocess the data. Typical preprocessing techniques are denoising and DC removal. Fig. 4 shows the spectrogram of a human walk before and after preprocessing. The latter is used for feature extraction, which can be divided into three categories.

1) *Manual Feature Extraction:* For fall detection based on spectrograms, many predefined features with clear physical motion interpretations have been proposed. We recognize three candidate features: Extreme frequency magnitude, extreme frequency ratio, and time-span of the event. These features have been shown to provide a good description of a fall [27]. The extreme frequency magnitude is defined as

$$F = \max(f_{+\max}, -f_{-\min}) \tag{2}$$

where $f_{+\max}$ and $f_{-\min}$, respectively, denote the maximum frequency in the positive frequency range and the minimum frequency in the negative frequency range. The extreme frequency ratio is defined as

$$R = \max(|f_{+\max}/f_{-\min}|, |f_{-\min}/f_{+\max}|). \tag{3}$$

Selecting predefined features is a cumbersome process and requires significant tuning of parameters. In contrast, principal component analysis (PCA) and deep learning build on features that are not encapsulated in few parameters, and as such, are considered alternative viable motion classification approaches.

2) *Principal Component Analysis:* The PCA-based motion classification comprises the following steps [44]. First, each preprocessed gray-scale spectrogram image in the training set is vectorized and stored as a column of the "training matrix." Next, the average of the vectorized training images is subtracted from each column of the training matrix. The resulting training matrix is used for generating the eigenvalues and the respective eigenvectors (eigen images). The training set is projected onto the space spanned by the selected eigen images corresponding to the dominant eigenvalues. These projections are then used in the classification process. When the test data becomes available, it is projected onto the eigen space and the resulting projection is compared with all projections obtained for the training images. The minimum Euclidean distance determines the class with the closest match to the observed test motion.

3) *Deep Learning:* Deep learning is a nonlinear generalization of the PCA [45] and is the focus of this paper. Its application to fall detection is described in the next section.

Once the features are extracted, they are fed to the classifier. It should be noted that the choice of features has greater impact on the classification results than the choice of a classifier [46].

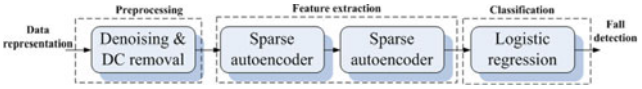


Fig. 5. Deep learning based architecture for fall detection.

III. DEEP LEARNING BASED FALL DETECTION USING RANGE-DOPPLER RADAR

The deep learning architecture for fall detection is shown in Fig. 5. In the general scheme, this method follows three steps, namely, preprocessing, feature extraction, and classification. Data representation can take the form of spectrogram or range map of the observed motion. The preprocessed representation is used as input to stacked autoencoders that perform feature extraction.

A. Stacked Autoencoders

An autoencoder is defined as a neural network which learns nonlinear approximation of its input [45]. In other words, it attempts to reconstruct its input at its output. The learning is achieved via a single hidden layer that may have different dimensionality than the input and output layers. Typically, the hidden layer has fewer units than the input layer. In this case, the network tries to learn a compressed version of the input data, i.e., a sparse representation.

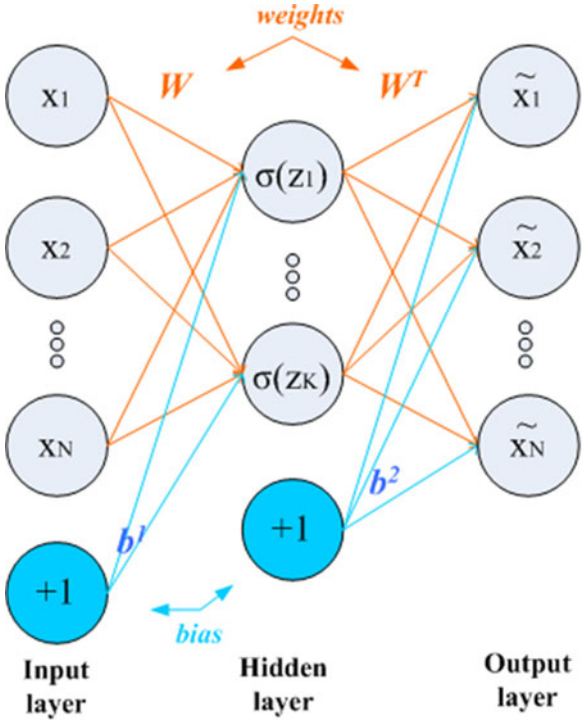


Fig. 6. Example of an autoencoder with N input units and K hidden units.

Fig. 6 depicts a sparse autoencoder with N and K units in the input and hidden layer, respectively. Connections between layers are established by the weight matrix W and bias vectors b^1 and b^2, which are typically represented by units with the "+1" label.

Each hidden unit applies a nonlinear function $f\{\bullet\}$ to the weighted and biased input data x, i.e., the output of hidden layer unit is $a_i = f(z_i)$ where z_i is given by

$$z_i = \sum_{j=1}^{N} W_{i,j} x_j + b_i^1, \qquad i = 1, \ldots, K. \tag{4}$$

The sigmoid function is used as the nonlinear function:

$$f(z) = \sigma(z) = \frac{1}{1 + \exp^{-z}}. \tag{5}$$

The values of output layer units are obtained in a similar way, by applying the nonlinear function $f\{\bullet\}$ to weighed and biased hidden layer units a_i.

The weights and biases of neurons are learned in such manner which minimizes the reconstruction error and promotes sparsity. Suppose that we are given a set of M unlabeled training samples $\{x^{(1)}, x^{(2)}, \ldots, x^{(M)}\}$, where $x^{(m)} \in R^N$. The cost function for sparse autoencoder can be formulated as

$$J(W, b) = \sum_{m=1}^{M} E(x^{(m)}) + \lambda \sum_{i=1}^{N} \sum_{j=1}^{K} (W_{j,i})^2 + \beta \sum_{j=1}^{K} D_{KL}(\rho, \hat{\rho}_j) \tag{6}$$

where $E(x^{(m)})$ is the reconstruction error for single example, i.e., the error between the input data and the autoencoder output:

$$E(x^m) = \left\| x^{(m)} - \tilde{x}^{(m)} \right\|_2^2 = \left\| x^{(m)} - [\sigma(W^T \sigma(W x^{(m)} + b^1) + b^2)] \right\|_2^2. \tag{7}$$

$D_{KL}(\rho, \hat{\rho}_j)$ is Kullback–Leibler divergence that is defined as

$$D_{KL}(\rho, \hat{\rho}_j) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j}. \tag{8}$$

The second term in (6) is a regularization term which prevents the weights from assuming high values. The last term in the cost function, the divergence $D_{KL}(\rho, \hat{\rho})$, is responsible for obtaining sparse representation. The importance of this term is determined by parameter β. Kullback–Leibler divergence measures the difference between two probability distributions. In our case, ρ is a sparsity parameter that we specify, whereas $\hat{\rho}_j$ is the average activation of each hidden neuron. By setting ρ to a small value close to zero, we enforce the average activation of the hidden unit to be approximately equal to that value, leading to sparsity. Minimization of the cost function is typically done using any gradient descent method. The backpropagation algorithm is used to efficiently compute partial derivatives that are required for the gradient descent.

By computing the sparse representation, we are, in essence, extracting the most prominent features. Since images of human motions contain significant amount of useful information, it is prudent to extract this information in several layers where each layer represents a different concept of the input data. Stacked autoencoders are useful for capturing hierarchical structure. For example, one layer can learn the edges, while the next layer can learn the shapes which contain these edges. The learning of the input representation in multiple levels can be achieved using stacked autoencoders where the output of one autoencoder is the input to the next one. In this paper, feature extraction is performed using two sparse autoencoders. This number was chosen empirically. Employing more layers did not provide any significant improvement in results, while it considerably increased computation time and the network complexity.

It is interesting to visualize what sparse autoencoder learns as features of the input data. Fig. 7 shows few images that represent the learned features using one autoencoder. The network was trained using spectrograms of four human motions: Falling, sitting, bending, and walking. Hidden layer with 300 nodes attempts to learn the most important information from 4096 input layer units that correspond to 64×64 image. We can notice that the autoencoder captures signatures of human motions in the time-frequency domain. These images resemble the eigen images that would be obtained by PCA [47], which is expected since PCA is linear version of an autoencoder.

B. Classification

In general, the classification problem observes multiple motion classes. In this case, softmax regression (or

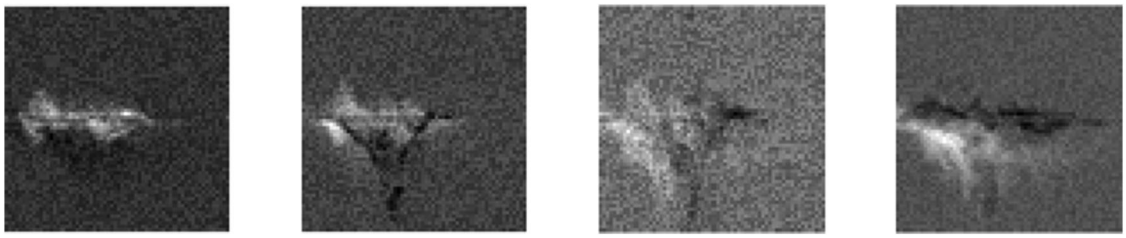


Fig. 7. Images of few learned features when sparse autoencoder is applied to spectrograms.

multinomial logistic regression) can be used to classify the data. Multinomial logistic regression is a supervised learning algorithm, which uses feature data z and their corresponding labels. The output of this classifier is defined as L-dimensional vector where L denotes the number of classes. The elements of this vector represent the estimated probabilities that the data z belongs to the class label y_l, $l = 1, \ldots, L$. These probabilities p_l, $l = 1, \ldots, L$, are defined by parameter θ, i.e.

$$p_l = \frac{1}{\sum_{l=1}^{L} e^{\theta_l^T z}} e^{\theta_l^T z}. \tag{9}$$

Given a set of M labeled samples $\{(z^{(1)}, y^{(1)}),..., (z^{(M)}, y^{(M)})\}$, the parameters θ_l are obtained by minimizing the following cost function:

$$J(\theta) = -\sum_{m=1}^{M}\sum_{l=1}^{L} 1\{y^{(m)} = l\} log \frac{e^{\theta_l^T z^{(m)}}}{\sum_{l=1}^{L} e^{\theta_l^T z^{(m)}}} \tag{10}$$

where $1\{\cdot\}$ is the indicator function which is defined as

$$1\{\text{statement}\} = \begin{cases} 1, & \text{if statement is true,} \\ 0, & \text{if statement is false.} \end{cases}$$

Since we observe only two classes (fall versus nonfall), the problem is binary and softmax regression amounts to logistic regression, i.e., $L = 2$ with two probabilities at the output.

IV. INTEGRATION OF VELOCITY AND RANGE INFORMATION

In this section, we integrate both spectrogram and range map information prior to classification. Below, we consider two approaches for information fusion. Fusion methods are typically used to enhance classifier performance of individual classifiers.

A. Parallel Processing of Spectrograms and Range Maps

Fig. 8 shows two possible fusion architectures for the underlying fall detection problem. Fusion has been successfully used in applications when multiple sources of information are available [48], [49]. In general, fusion can occur at early stages (domain level, feature extraction level) or at the classifier level. In this paper, we focus on the early integration of both domains. The late fusion, i.e., fusing the outputs post individual classifications, is not considered since we have only two outputs and the classical late

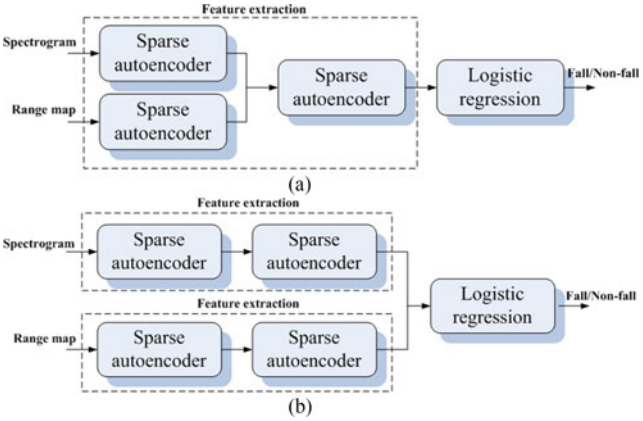


Fig. 8. Different architectures for fall detection using spectrograms and range maps in parallel manner.

fusion classification methods, such as the majority voting, require at least three outcomes in order to make the final decision. Since we are dealing with stacked autoencoders, there is an option of performing the fusion after each layer. Fig. 8(a) performs fusion after the first layer, whereas in Fig. 8(b), fusion is performed just before classification. In both cases, fusion is performed using features which correspond to different aspect of the input data. Since features of both domains provide complimentary information, the fusion of features at any level outperforms the classification results obtained without fusion.

B. Sequential Processing of Spectrograms and Range Maps

Both architectures shown in Fig. 8 depict the parallel use of time-frequency and range signatures before making the final decision. However, as mentioned in Section II, different motions can exhibit variable distinctions in different domains. This suggests an implementation of sequential tree-like classification where the decision proceeding from one level to another is based on a different signal representation. This type of sequential processing has been proven successful in motion classification [21]. The proposed scheme is given in Fig. 9. First, the range map is used to differentiate between motions with significant range translations, e.g., falling and walking, and those which almost occur in place, like sitting and bending, without considerable range translations. In essence, we create two classes, based on range information, where a fall is a

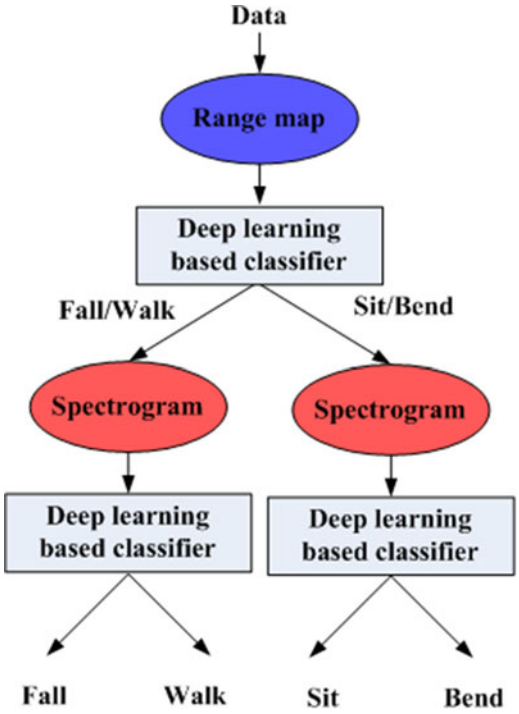
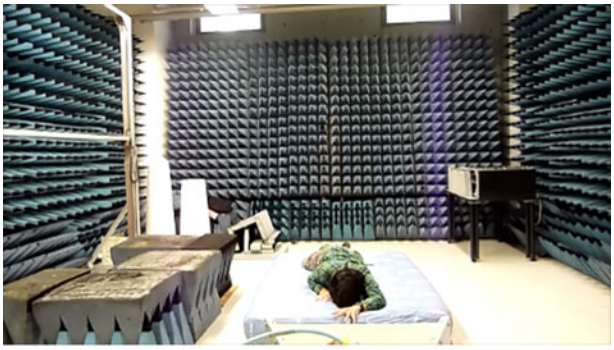


Fig. 9. Sequential processing of the spectrogram and range map. The block "Deep learning based classifier" corresponds to the scheme in Fig. 5.



(a)

(b)

Fig. 10. Semicontrolled(a) and uncontrolled environments(b) where experiments were performed.

member of one class. To discriminate fall from the other members of the same class, we use the distinctive features associated with the member TF signatures. The classification at each level is performed using deep learning based architecture depicted in Fig. 5.

V. EXPERIMENTAL RESULTS

The FMCW experiments were performed in both semicontrolled and uncontrolled environments mimicking a room. The former is shown in Fig. 10(a), and amounts to a room with covered back and side walls. An office with typical furniture (chairs, tables, and bookshelves) was used as the uncontrolled environment [Fig. 10 (b)]. A mattress was placed on the floor to prevent subjects from injuries during the fall experiments. The experiments were conducted at the Center for Advanced Communications, Villanova University. The system used in the experiments, named SDRKIT 2500B, is developed by Ancortek, Inc. Center frequency is 25 GHz, whereas the bandwidth is 2 GHz which provides 0.075 m range resolution. Three male subjects, all aged 26 years, participated in the experiments. Their physical characteristics are given in Table I.

TABLE I
Physical Characteristics of Subjects

	Subject A	Subject B	Subject C
Height	5'9"	5'9"	5'10"
Weight (lbs)	171	207	194

The dataset contained four human motions: falling, sitting, bending, and walking. Each motion was observed during a time span of 4 s and images of grid size 64×64 were generated corresponding to the spectrograms and range maps, and used as inputs. We consider different direction angles. Each subject performed five trials of each type of motion at 22.5°, 30°, and 45°, resulting in 180 trials performed at a nonzero angle. The rest of trials were performed at 0°. In this paper, the interest is in detecting different motion classes, i.e., interclass variability. Intraclass variability, which can be attributed to direction angles, is not considered. Different number of features are used for spectrograms and range maps since the two representations carry different information. In general, range maps are simpler and they embed fewer features. After learning the features, the final stage is a logistic regression classifier which determines the probability that the input data belongs to one of two possible classes (fall versus nonfall).

The dataset contains 408 signals: 117 falls and 291 nonfalls (111 sit, 115 bend, and 65 walk signals). The training set consisted of 100 falls and 100 nonfalls(30 sit, 30 bend, and 40 walk signals), while the rest of the dataset was used for testing. Cross-validation was performed using part of the training dataset to set the hyperparameters such as the number of units in the hidden layer and regularization parameters. In the case of spectrograms, the number of units in the hidden layer for the first autoencoder was set to 300, meaning that the network would attempt to compress 4096 coefficients into 300. The 300 outputs were further

TABLE II
Confusion Matrix for the Architecture Shown in Fig. 5—Spectrograms are Used as Inputs

Predicted/Actual class	Fall	Nonfall
Fall	16	21
Nonfall	1	170

Success rate is 89.4%.

TABLE III
Confusion Matrix for the Architecture Shown in Fig. 5—Range Maps are Used as Inputs

Predicted/Actual class	Fall	Nonfall
Fall	16	32
Nonfall	1	159

Success rate is 84.1%.

TABLE IV
Confusion Matrix for the Architecture Shown in Fig. 8(a)—Fusion at the First Layer

Predicted/Actual class	Fall	Nonfall
Fall	17	12
Nonfall	0	179

Success rate is 94.2%.

TABLE V
Confusion Matrix for the Architecture Shown in Fig. 8(b)—Fusion at the Second Layer

Predicted/Actual class	Fall	Nonfall
Fall	16	8
Nonfall	1	183

Success rate is 95.7%.

compressed using only 100 units in the second hidden layer. For range maps, we used 100 units in the first autoencoder, while the second autoencoder utilized 50 units. The maximum number of epochs was set to 100. In order to properly verify our approaches, 50 different combinations of training and testing sets are chosen and the results are averaged. The confusion matrices for different architectures are given in Tables II–V. Tables II and III provide results when the spectrograms and range maps are used separately. Their success rates are lower than those obtained with fusion-based methods (Tables IV and V). We can notice that fusion at the second level slightly outperforms fusion at the first level. This could be explained by the fact that it is better to extract features separately for each domain in order to preserve all relevant information. If the fusion is performed in earlier stages, pertinent features of a single domain representation could be lost and this would impact the final result. The sequential approach results are given in Table VI. This approach slightly outperforms the previous approaches, thus, promoting the notion of domain selectivity when representing human motions. Based on the results, we can notice the fluctuations in the false alarm and missed detection rates. Reducing these rates may further require fine tuning the network parameters and number of layers. We are primarily concerned in this paper with the demonstration of improved classification accuracy when choosing proper data representation domains for the same network.

TABLE VI
Confusion Matrix for the Architecture Shown in Fig. 9

Predicted/Actual class	Fall	Nonfall
Fall	15	4
Nonfall	2	187

Success rate is 97.1%.

TABLE VII
Confusion Matrix for the Architecture Shown in Fig. 5—Integrated Slow-Time Range-Doppler Maps are Used as Inputs

Predicted/Actual class	Fall	Nonfall
Fall	10	11
Nonfall	7	180

Success rate is 91.3%.

It should be noted that more sophisticated time-frequency representations could be used instead of spectrograms [42]. However, even though using higher-resolution or a higher-order time-frequency distribution may increase classification rate, this improvement comes at a higher computational complexity. It is also possible to use other joint-variable domains provided by the FMCW radar [30]. In addition to the two domains depicting range and Doppler versus slow-time, we considered the integrated time range-Doppler map which combines velocity and range information [31]. This map has been shown to capture different features for improved motion classification. Table VII contains results when integrated slow-time range-Doppler domain is used for fall detection. Results demonstrate that even though this domain offers high classification rate, it is still less effective compared to the combination of the other two domains.

We have also investigated the effect of subject's physical characteristics on the classifier performance. The classifier is trained on the spectrograms from two subjects, while the testing is performed on the spectrograms from a third subject. Simulation results given in Table VIII does not show significant difference in the classification rate compared to the training on a generic dataset (see Table II).

Another important part of deep learning is the dataset size. We have investigated classifier performance for different size of training set. Fig. 11 shows the learning curve. It can be noticed that even though the curve has not reached the saturation stage, the increments in success rates become significantly smaller after the curve reaches 80%.

In order to compare deep learning with traditional approaches, the confusion matrices for the classifiers based

TABLE VIII
Confusion Matrix for the Approach When the Classifier is Trained on Samples Belonging to Subjects A and B, While the Testing is Performed on the Samples Belonging to the Subject C

Predicted/Actual class	Fall	Nonfall
Fall	14	24
Nonfall	3	167

Success rate is 87%.

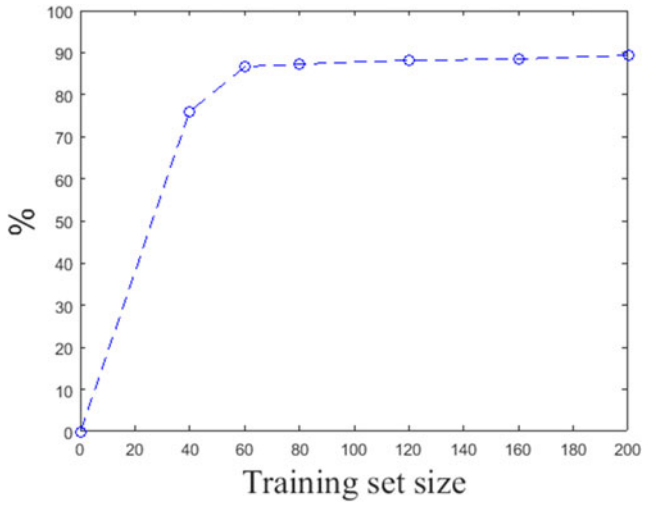


Fig. 11. Learning curve. Success rates for different training set sizes when spectrograms are used as inputs.

TABLE IX
Confusion Matrix for the Manual Feature Extraction Approach

Predicted/Actual class	Fall	Nonfall
Fall	10	34
Nonfall	7	157

Success rate is 80.2%.

TABLE X
Confusion Matrix for the PCA Approach

Predicted/Actual class	Fall	Nonfall
Fall	16	30
Nonfall	1	161

Success rate is 85%.

on manual feature extraction and PCA, both implemented in the TF domain, are given in Tables IX and X, respectively. It should be noted that in the case of predefined feature extraction, significant amount of tuning and image processing was performed in order to ensure that meaningful values are extracted. The number of dominant components for PCA was set to 20 in order to capture 90% of variance in the data. There was no improvement in the results by using higher number of components. The corresponding tables show inferior success rates to the deep learning approach. This demonstrates that limited number of predefined features or the use of dominant eigenvectors is not capable of representing the complexity of human motions.

VI. CONCLUSION

In this paper, we applied deep learning for range-Doppler radar-based fall motion detection. The deep learner consisted of two stacked autoencoders and a logistic regression classifier. Different architectures using spectrograms and range maps were presented. In order to verify our approaches, four human motions were considered, namely, walking, falling, bending/straightening, and sitting. The experimental results demonstrated the superiority of the deep learning based approach over conventional and PCA-based methods in detecting a fall. Among the deep learning methods, incorporation of both spectrograms and range maps outperformed the methods where individual data domain representations are used separately.

ACKNOWLEDGMENT

The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] G. Bergen, M. R. Stevens, and E. R. Burns
Falls and Fall Injuries Among Adults Aged 65 Years United States, 2016. [Online]. Available: http://dx.doi.org/10.15585/mmwr.mm6537a2

[2] World Health Organization and Life Course Unit
WHO Global Report on Falls Prevention in Older Age. World Health Organization, Geneva, Switzerland, 2008.

[3] Administration for Community Living *Aging Statistics*, ACL, Washington, DC, USA, 2016.

[4] E. R. Burns, J. A. Stevens, and R. Lee
The direct costs of fatal and non-fatal falls among older adult-sunited states
J. Safety Res., vol. 58, pp. 99–103, 2016.

[5] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers
A review of wearable sensors and systems with application in rehabilitation
J. Neuroeng. Rehabil., vol. 9, no. 1, pp. 1–17, 2012.

[6] J. Chen, K. Kwong, D. Chang, J. Luk, and R. Bajcsy
Wearable sensors for reliable fall detection
In *Proc. 27th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2005, pp. 3551–3554.

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

[8] C.-C. Yang and Y.-L. Hsu
A review of accelerometry-based wearable motion detectors for physical activity monitoring
Sensors, vol. 10, no. 8, pp. 7772–7788, 2010.

[9] S. J. Preece, J. Y. Goulermas, L. P. Kenney, D. Howard, K. Meijer, and R. Crompton
Activity identification using body-mounted sensors—A review of classification techniques
Physiol. Meas., vol. 30, no. 4, p. R1–33, 2009.

[10] Q. Li, J. A. Stankovic, M. A. Hanson, A. T. Barth, J. Lach, and G. Zhou
Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information
In *Proc. IEEE 6th Int. Workshop Wearable Implantable Body Sensor Netw.*, Jun. 2009., pp. 138–143.

[11] A. K. Bourke and G. M. Lyons
A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor
Med. Eng. Phys., vol. 30, no. 1, pp. 84–90, 2008.

[12] M. Nyan, F. E. Tay, and E. Murugasu
A wearable system for pre-impact fall detection
J. Biomech., vol. 41, no. 16, pp. 3475–3481, 2008.

[13] J. Dai, X. Bai, Z. Yang, Z. Shen, and D. Xuan
Perfalld: A pervasive fall detection system using mobile phones
In *Proc. IEEE 8th IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, 2010, pp. 292–297.

[14] M. Chen, Y. Ma, J. Song, C.-F. Lai, and B. Hu
Smart clothing: connecting human with clouds and big data for sustainable health monitoring
Mob. Netw. Appl., vol. 21, no. 5, pp. 825–845, 2016.

[15] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau
Fall detection from human shape and motion history using video surveillance
In *Proc. IEEE 21st Int. Conf. Adv. Inf. Netw. Appl. Workshops*, 2007, vol. 2, pp. 875–880.

[16] R. Cucchiara, A. Prati, and R. Vezzani
A multi-camera vision system for fall detection and alarm generation
Expert Syst., vol. 24, no. 5, pp. 334–345, 2007.

[17] E. Auvinet, F. Multon, A. Saint-Arnaud, J. Rousseau, and J. Meunier
Fall detection with multiple cameras: An occlusion-resistant method based on 3-d silhouette vertical distribution
IEEE Trans. Inf. Technol. Biomed, vol. 15, no. 2, pp. 290–300, Mar. 2011.

[18] G. Mastorakis and D. Makris
Fall detection system using kinects infrared sensor
J. Real-Time Image Process., vol. 9, no. 4, pp. 635–646, 2014.

[19] Y. Zigel, D. Litvak, and I. Gannot
A method for automatic fall detection of elderly people using floor vibrations and soundproof of concept on human mimicking doll falls
IEEE Trans. Biomed. Eng., vol. 56, no. 12, pp. 2858–2867, Dec. 2009.

[20] J. Rodrigues, S. Misra, H. Wang, and Z. E. Zhu
Special section on ambient assisted living communications
IEEE Comm. Mag., vol. 53, no. 1, pp. 24–87, Jan. 2015.

[21] Y. Kim and H. Ling
Human activity classification based on micro-Doppler signatures using a support vector machine
IEEE Trans. Geosci. Remote Sens., vol. 47, no. 5, pp. 1328–1337, May 2009.

[22] P. van Dorp and F. Groen
Human walking estimation with radar
IET Radar, Sonar Navig., vol. 150, no. 5, pp. 356–365, 2003.

[23] B. G. Mobasseri and M. G. Amin
A time-frequency classifier for human gait recognition
Proc. SPIE, vol. 7306, 2009, Art. no. 730628.

[24] D. Tahmoush and J. Silvious
Time-integrated range-doppler maps for visualizing and classifying radar data
In *Proc. IEEE RadarCon*, 2011, pp. 372–374.

[25] M. Mercuri, D. Schreurs, and P. Leroux
SFCW microwave radar for in-door fall detection
In *Proc. IEEE Topical Conf. Biomed. Wireless Technol., Netw., Sensing Syst.*, 2012, pp. 53–56.

[26] L. Liu, M. Popescu, M. Skubic, M. Rantz, T. Yardibi, and P. Cuddihy
Automatic fall detection based on Doppler radar motion signature
In *Proc. 5th Int. Conf. Pervasive Comput. Technol. Healthcare*, 2011, pp. 222–225.

[27] L. R. Rivera, E. Ulmer, Y. D. Zhang, W. Tao, and M. G. Amin
Radar-based fall detection exploiting time-frequency features
In *Proc. IEEE China Summit Int. Conf. Signal Inf. Process.*, Xi'an, China, Jul. 2014, pp. 713–717.

[28] M. G. Amin, Ed.
Radar for Indoor Monitoring. Boca Raton, FL, USA: CRC Press, 2017.

[29] C. Garripoli *et al.*
Embedded DSP-based telehealth radar system for remote in-door fall detection
IEEE J. Biomed. Health Informat., vol. 19, no. 1, pp. 92–101, Jan. 2015.

[30] Z. Peng, J.-M. Muñoz-Ferreras, R. Gómez-García, and C. Li
FMCW radar fall detection based on ISAR processing utilizing the properties of RCS, range, and Doppler
In *Proc. IEEE Int. Microwave Symp.*, May 2016, pp. 1–3.

[31] B. Erol and M. G. Amin
Fall motion detection using combined range and Doppler features
In *Proc. 24th Eur. Signal Process. Conf.*, 2016, pp. 2075–2080.

[32] Y. Kim and H. Ling
Human activity classification based on micro-doppler signatures using an artificial neural network
In *Prov. IEEE Antennas Propag. Soc. Int. Symp.*, 2008, pp. 1–4.

[33] Q. Wu, Y. Zhang, W. Tao, and M. Amin
Radar-based fall detection based on doppler time frequency signatures for assisted living
IET Radar, Sonar Navig., vol. 9, no. 2, pp. 164–172, 2015.

[34] 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 Navig., vol. 9, no. 2, pp. 173–183, 2015.

[35] B. Y. Su, K. Ho, M. J. Rantz, and M. Skubic
Doppler radar fall activity detection using the wavelet transform
IEEE Trans. Biomed. Eng., vol. 62, no. 3, pp. 865–875, Mar. 2015.

[36] B. Jokanovic, M. Amin, and F. Ahmad
Radar fall motion detection using deep learning
In *IEEE RadarCon*, May 2016, pp. 1–6.

[37] Y. Kim and T. Moon
Classification of human activity on water through micro-dopplers using deep convolutional neural networks
Proc. SPIE, vol. 9829, 2016, Art. no. 982917.

[38] V. C. Chen
Analysis of radar micro-doppler with time-frequency transform
In *Proc. IEEE Workshop Stat. Signal Array Process.*, Pocono Manor, PA, USA, Aug. 2000, pp. 463–466.

[39] V. C. Chen
The MicroDoppler Effect in Radar. Norwood, MA, USA: Artech House, 2011.

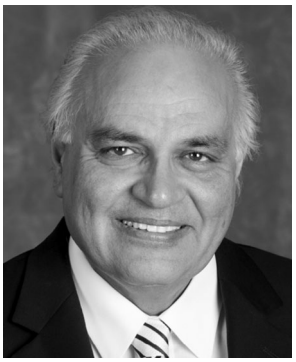
[40] C. Clemente, A. Balleri, K. Woodbridge, and J. J. Soraghan
Developments in target micro-Doppler signatures analysis: Radar imaging, ultrasound and through-the-wall radar
EURASIP J. Adv. Signal Process., vol. 2013, no. 1, pp. 1–18, 2013.

[41] C. Clemente, L. Pallotta, A. Maio, J. Soraghan, and A. Farina
A novel algorithm for radar classification based on Doppler characteristics exploiting orthogonal pseudo-Zernike polynomials
IEEE Trans. Aerosp. Electron. Syst., vol. 51, no. 1, pp. 417–430, Jan. 2015.

[42] B. Jokanovic, M. G. Amin, and F. Ahmad
Effect of data representations on deep learning in fall detection
In *IEEE Sensor Array Multichannel Signal Process. Workshop*, 2016, pp. 1–5.
[43] B. Erol and M. Amin
Effects of range spread and aspect angle on radar fall detection
In *IEEE Sensor Array Multichannel Signal Process. Workshop*, 2016, pp. 1–5.
[44] M. Turk and A. Pentland
Eigenfaces for recognition
J. Cognit. Neurosci., vol. 3, no. 1, pp. 71–86, 1991.
[45] G. E. Hinton and R. R. Salakhutdinov
Reducing the dimensionality of data with neural networks
Science, vol. 313, no. 5786, pp. 504–507, 2006.
[46] S. Gurbuz, B. Tekeli, M. Yuksel, C. Karabacak, A. Gurbuz, and M. Guldogan
Importance ranking of features for human micro-doppler classification with a radar network
In *Proc. 16th Int. Conf. Inf. Fusion*, Jul. 2013, pp. 610–616.
[47] B. Jokanovic, M. Amin, F. Ahmad, and B. Boashash
Radar fall detection using principal component analysis
Proc. SPIE, vol. 9829, 2016, Art. no. 982919.
[48] B. Xie and H. Minn
Real-time sleep apnea detection by classifier combination
IEEE Trans. Inf. Technol. Biomed., vol. 16, no. 3, pp. 469–477, May 2012.
[49] S. Rasheed, D. W. Stashuk, and M. S. Kamel
Integrating heterogeneous classifier ensembles for EMG signal decomposition based on classifier agreement
IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 3, pp. 866–882, May 2010.

Branka Jokanović (S'12) received the B.Sc. and M.Sc. degrees in electrical engineering from the Faculty of Electrical Engineering, University of Montenegro, Podgorica, Montenegro, in 2010 and 2012, respectively. She is currently working toward the Ph.D. degree in electrical engineering at Villanova University, Villanova, PA, USA.

She is currently a Research Assistant at the Villanova University, Villanova, PA, USA. Her research interests include time-frequency analysis, compressive sensing, and their application to radars.



Moeness Amin (S'82–F'01) received the Ph.D. degree in electrical engineering from the University of Colorado, Boulder, CO, USA, in 1984. He is the Director of the Center for Advanced Communications, Villanova University, Villanova, PA, USA. He is the Fellow of the International Society of Optical Engineering, the Fellow of the Institute of Engineering and Technology, and the Fellow of the European Association for Signal Processing. He has more than 700 journal and conference publications in signal processing theory and applications. He co-authored 18 book chapters and is the Editor of the books *Through the Wall Radar Imaging* (CRC Press, 2011), *Compressive Sensing for Urban Radar* (CRC Press, 2014), and *Radar for Indoor Monitoring* (CRC Press, 2017).

Dr. Amin was a Distinguished Lecturer of the IEEE Signal Processing Society, 2003–2004, and was the Chair of the Electrical Cluster of the Franklin Institute Committee on Science and the Arts (2012–2016). He received the 2014 IEEE Signal Processing Society Technical Achievement Award, the 2009 Individual Technical Achievement Award from the European Association for Signal Processing, the IEEE Warren D White Award for Excellence in Radar Engineering, the IEEE Third Millennium Medal, the 2010 NATO Scientific Achievement Award, the 2010 Chief of Naval Research Challenge Award, the Villanova University Outstanding Faculty Research Award, 1997, and the IEEE Philadelphia Section Award, 1997.