Human Activity Classification with radar: Optimization and Noise Robustness with Iterative Convolutional Neural Networks followed with Random Forests

Yier Lin, *Student, IEEE*, Julien Le Kernec, *Senior Member, IEEE*, Shufan Yang, *Member, IEEE*, Francesco Fioranelli, *Member, IEEE*, Olivier Romain, *Member, IEEE*, and Zhiqin Zhao, *Senior Member, IEEE*.

***Abstract*—The accurate classification of activity patterns based on radar signatures is still an open problem and is key to detect anomalous behavior for security and health applications. This paper presents a novel iterative convolutional neural networks strategy with an autocorrelation pre-processing instead of the traditional micro-Doppler image pre-processing to classify activities or subjects accurately. The proposed strategy uses an iterative deep learning framework for the automatic definition and extraction of features. This is followed by a traditional supervised learning classifier to label the different activities. Using three human subjects and their real motion captured data, twelve thousand radar signatures were simulated by varying additive white Gaussian noise. Additionally, 6720 experimental radar signatures were captured with a frequency-modulated continuous radar at $5.8GHz$ with $400MHz$ of instantaneous bandwidth from seven activities using one subject and 4800 signatures from five subjects while walking. The simulated and experimental data were both used to validate our proposed method. With SNR varying from -20 to $20dB$ with 88.74% average accuracy at $-10dB$ and 100% peak accuracy at $15dB$. The proposed Iterative Convolutional Neural Networks followed with Random Forests (ICNNRF) does not only outperform the feature-based methods using micro-Doppler images but also the classification methods using other types of supervised classifiers after our proposed iterative convolutional neural network.**

***Index Terms*—Micro-Doppler, Deep Learning, Convolution Neural Networks, Random Forests, Radar.**

I. Introduction

WITH increasing numbers of terror attacks and their diversity, automating the detection and classification of "normal activities"against anomalous behavior has never been more important. Security systems for mass surveillance need to be able to identify and learn what "normal "behavior is, in order to isolate anomalous behaviors or threats from a crowd of people moving in an area or building [1, 2]. Therefore, first and foremost, accurate classification of activities or individuals is a key enabler. Although classification could be performed with video or images, for the sake of protecting privacy, cameras are not allowed in many places, and furthermore might be sensitive to lighting and weather conditions especially outdoors. However, radar is an interesting sensing modality to investigate as an alternative or complementary tool as it can operate night and day and in all weather.

Classically, activities or individuals are distinguished with methods based on micro-Doppler radar signatures. The relative motion of structural components of an object/body generates unique patterns in the time-frequency domain of the radar returns. Therefore, different activities are generating uniquely distinctive features in micro-Doppler signatures (mDs) - a.k.a. spectrograms - that can be used for classification. An overview of mDs is provided in [3, 4]. In general, features are extracted from mDs, followed by supervised machine learning [5, 6]. This kind of technique relies on features that are either formulaic (e.g. centroid, skewness) or handcrafted. The best set of parameters for optimal classification accuracy is usually determined by trial and error, and requires significant effort in fine-tuning the values of the input features. When activities are similar in nature like walk, fast walk and running, the classifier will be faced with "confusers"that drastically reduce accuracy and requires new strategies to deal with those.

Deep learning methods such as Convolutional Neural Networks (CNNs) [7-11] and Deep Belief Network (DBN) [12] have recently revolutionized several applications using hierarchical neural networks, and they have been shown to significantly outperform previous state-of-the-art classifiers relying on domain knowledge-based features. The aforementioned techniques rely on image pre-processing. Most of these methods [7-9, 12] use the square pixel images directly as inputs. mDs images do not always have the right dimensions, so great care is needed in selecting which part of the mDs image to use. A bad selection results in classification errors. In [10, 11], a deep learning method obtained improved accuracy by computing the sparse representation to extract the most salient features. The source of errors is the Principal Component Analysis of the gray-scale mDs image, resulting in 90% classification accuracy for fall, sit, bend, and walk shown in [11]. 90% accuracy is still too low and would result in an unacceptable false alarm rate for practical in-field applications. There is still scope for improvement in automatic classification.

Yier Lin and Zhiqin Zhao are with the School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mails: yierlin@foxmail.com, zqzhao@uestc.edu.cn). Julien Le Kernec, Shufan Yang, and Francesco Fioranelli are with the School of Engineering, University of Glasgow, UK (e-mails: {julien.lekernec, shufan.yang, Francesco.Fioranelli}@glasgow.ac.uk). Olivier Romain is with the ETIS laboratory (Information Processing and System Teams) at University Cergy-Pontoise, France (email: olivier.romain@u-cergy.fr).

In this article, a novel alternative approach for classification of human activities which learns the features directly from the AutoCorrelation Function (ACF) of the radar range information. This contrasts with the classical approach of generating and processing mDs images followed by feature extraction. Our approach outperforms other state-of-the-art classification techniques, especially when activities are very similar in nature and are hardly distinguishable from each other to the human eye or using traditional feature-based classification from mDs. The proposed approach is a two-stage process initiated by an iterative deep learning framework for feature definition and extraction, followed by a traditional classifier. This method bypasses the generation of mDs. Consequently, the trade-off between time and frequency resolution can be avoided and the loss of range information incurred when generating mDs. The deep learning framework based on convolutional neural networks (CNNs) uses the ACF as input from which it defines and extracts features automatically. The CNNs output the selected features it defined, and we will show in the following sections that Random Forests (RF) classifier as second step outperforms other considered classifiers. Furthermore, we will test the robustness and optimization of the method against other conventional classification strategies with respect to signal-to-noise ratio, input size and epoch numbers.

In this paper, scalars will be denoted with lower case symbols, $e.g.\ x$, whereas vectors will be denoted with bold lower case letter, $\boldsymbol{x}$. Matrices will be denoted with bold upper case letter, $\boldsymbol{X}$. Furthermore, $*$ will be utilized to denote convolution operator, The notations $\hat{}$ and $^{-}$ will be used to denote the estimated operator and conjugate transpose operator respectively.

The remainder of this paper is structured as follows. The next section explains the theory and methodology applied for classification. Section III introduces how the radar echoes for human activities are simulated and the classification performance against signal to noise ratio for different classification strategies. In section IV, we apply the proposed method on experimental radar data. The performances of ICNNRF is tested against data input length and epoch numbers and compared to other classification techniques accompanied with discussions and plans for future work. Finally, conclusions are drawn in Section V.

II. Theory and Methods

Classically for classification problems in radar, the mDs - a visual representation of the spectrum (usually in amplitude) of the Doppler modulation as a function of time - displaying the relative motion of structural components of an object/body. These generate unique patterns in the time-frequency domain of the radar returns. Research efforts have mainly been focused on learning features from mDs images directly for classification; however, mDs mean a trade-off between time and Doppler frequency resolutions and a loss of the range information. To avoid information loss, we apply the ACF on the range data directly and apply our proposed iterative CNNs followed with Random Forest.

A. Theoretical development

The radar range information can be considered in of the form

$$sig(t) = a(t) + \sigma(t), t = t_0, \cdots, t_{K-1}, \tag{1}$$

where t denotes the sample times, whereas K denotes the number of available samples, and $a(t)$denotes the radar amplitude range data without the unrelated additive white Gaussian noise (AWGN) $\sigma(t)$. Considering (1), the ACF can be expressed as (2).

$$\begin{aligned} ACF(sig(t)) &= sig(t) * \overline{sig(-t)} \\ &= a(t) * \overline{a(-t)} + a(t) * \overline{\sigma(-t)} \\ &+ \sigma(t) * \overline{a(-t)} + \sigma(t) * \overline{\sigma(-t)}. \\ &= a(t) * \overline{a(-t)} + \sigma(t) * \overline{\sigma(-t)} \\ &= ACF(a(t)) + ACF(\sigma(t)) \end{aligned} \tag{2}$$

Assume that the total signal number is N. All the ACFs can be re-arranged as shown in (3).

$$\sum_{i=1}^{N} ACF(sig_i(t)) = \sum_{i=1}^{N} ACF(a_i(t)) + \sum_{i=1}^{N} ACF(\sigma_i(t)), \tag{3}$$

where $sig_i(t)$, $a_i(t)$, and $\sigma_i(t)$ denote the range data with AWGN, range data without AWGN and AWGN of i^{th} sample respectively.

To make the equations clearly understandable, we assume that

$$\boldsymbol{sig} = \sum\nolimits_{i=1}^{N} ACF(sig_i(t)), \tag{4}$$

$$\boldsymbol{a} = \sum\nolimits_{i=1}^{N} ACF(a_i(t)), \tag{5}$$

$$\boldsymbol{\sigma} = \sum\nolimits_{i=1}^{N} ACF(\sigma_i(t)). \tag{6}$$

Thus (3) can be rewritten as

$$\boldsymbol{sig} = \boldsymbol{a} + \boldsymbol{\sigma} \approx \boldsymbol{a}. \tag{7}$$

To solve (7), the Lagrangian form can be defined as

$$min||\boldsymbol{sig} - \boldsymbol{a}||_2^2 + \lambda_1||\boldsymbol{a}||_1^1. \tag{8}$$

The first term in (8) measures the distance between signal and model, and the second term enforces an overall sparsity between diverse signals and thus the value of λ_1 limits the ACF number. The first term is a ℓ_2-norm function, and the second term is a ℓ_1-norm function. Thus, the solution of (8) is a convex optimization problem [13] as it is a sum of convex functions.

Assume $H(\cdot)$ is an unknown convex feature mapping function, then (4) and (5) can be expressed as: and which yields

$$\boldsymbol{y} = H(\boldsymbol{sig}), \tag{9}$$

$$\boldsymbol{x} = H(\boldsymbol{a}). \tag{10}$$

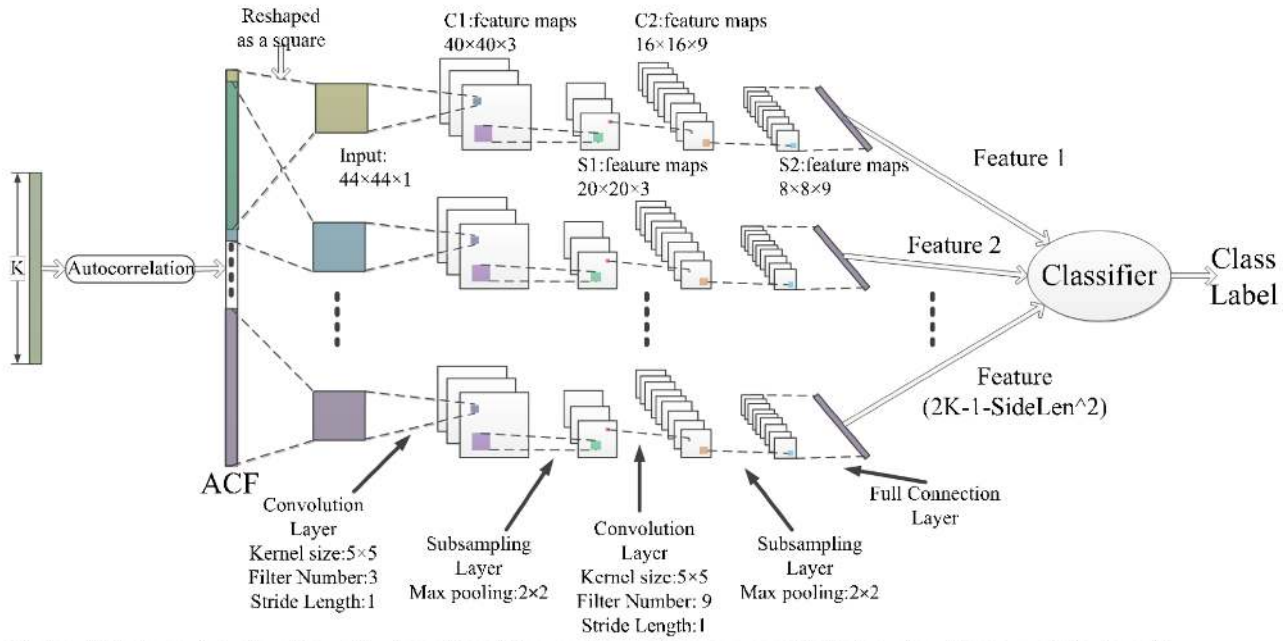

Fig. 1: Structure of the Iterative CNNs followed by a classifier to obtain class labels for an input size of 44×44.

Considering (8-10), we need to solve

$$min||\boldsymbol{y} - \boldsymbol{x}||_2^2 + \lambda||\boldsymbol{x}||_1^1, \tag{11}$$

which is also a convex problem. To obtain a better solution, weights $\boldsymbol{A}$ are added to the ℓ_1-norm:

$$min||\boldsymbol{y} - \boldsymbol{x}||_2^2 + \lambda||\boldsymbol{A}\boldsymbol{x}||_1^1. \tag{12}$$

Assume that

$$\boldsymbol{W} = \boldsymbol{A}\boldsymbol{x}, \tag{13}$$

(12) can be rewritten as:

$$min||\boldsymbol{y} - \boldsymbol{x}||_2^2 + \lambda||\boldsymbol{W}||_1^1. \tag{14}$$

Equation (14) is a simple deep learning form [14], and a fully connected feedforward CNNs method is applied to solve it. In this sort of deep learning method, $\boldsymbol{x}$ represents the desired output at the last layer, as well as $\boldsymbol{y}$ denotes the actual value of the output at the last layer, while $\boldsymbol{W}$ denotes the network weight, and λ denotes the scaling parameter of the ℓ_1-norm function.

B. Method

We propose utilizing LeNet-5 CNNs framework [15, 16] with its "typical"parameters to solve (14), which has two convolutional layers, two subsampling layers, and a fully-connected layer. The CNNs have 5×5 kernels for the convolutional layers with respectively 3 and 9 filters and a scaling of 2 for the max-pooling layers. Assume the CNNs input size is a $B \times B$ matrix, a filter size of $F \times F(F < B)$, $L1$ filters in the first convolutional layer and $L2$ filters in the second. After the first convolution, the output matrix is $(B - F + 1) \times (B - F + 1) \times L1$. The subsampling layer reduces the matrix to $\frac{B-F+1}{2} \times \frac{B-F+1}{2} \times L1$ feature maps. Through second convolution layer, the feature map is shrunk to $(\frac{B-F+1}{2} - F + 1) \times (\frac{B-F+1}{2} - F + 1) \times L2$ followed by a max-pooling layer resulting in $\frac{\frac{B-F+1}{2}-F+1}{2} \times \frac{\frac{B-F+1}{2}-F+1}{2} \times L2$ feature maps before the fully connected layer. The parameters for each layer are show in Table I for varying input sizes. Note that the optimization of the CNNs architecture, filter size and numbers in different layers is beyond the scope of this article and is a future research direction.

TABLE I: Feature Maps Parameters with Various CNNs Input Size

Input Matrix	C1	S1	C2	S2
$44 \times 44 \times 1$	$40 \times 40 \times 3$	$20 \times 20 \times 3$	$16 \times 16 \times 9$	$8 \times 8 \times 9$
$40 \times 40 \times 1$	$36 \times 36 \times 3$	$18 \times 18 \times 3$	$14 \times 14 \times 9$	$7 \times 7 \times 9$
$36 \times 36 \times 1$	$32 \times 32 \times 3$	$16 \times 16 \times 3$	$12 \times 12 \times 9$	$6 \times 6 \times 9$
$32 \times 32 \times 1$	$28 \times 28 \times 3$	$14 \times 14 \times 3$	$10 \times 10 \times 9$	$5 \times 5 \times 9$
$28 \times 28 \times 1$	$24 \times 24 \times 3$	$12 \times 12 \times 3$	$8 \times 8 \times 9$	$4 \times 4 \times 9$
$24 \times 24 \times 1$	$20 \times 20 \times 3$	$10 \times 10 \times 3$	$6 \times 6 \times 9$	$3 \times 3 \times 9$
$20 \times 20 \times 1$	$16 \times 16 \times 3$	$8 \times 8 \times 3$	$4 \times 4 \times 9$	$2 \times 2 \times 9$
$16 \times 16 \times 1$	$12 \times 12 \times 3$	$6 \times 6 \times 3$	$2 \times 2 \times 9$	$1 \times 1 \times 9$

C1: the first convolutional layer; C2: the second convolutional layer; S1: the first subsampling layer; S2: the second subsampling layer.

The solution quality obtained from (14) will solely depend on the CNNs. CNNs only accept square matrices as input. Assuming a length of range radar data is K, the ACF size is $(2K - 1)$. Unfortunately, the ACF length is not always equal to the square of a positive integer. Hence, a single

CNNs processing is resulting in insufficient accuracy. To solve this problem and improve the classification accuracy, we designed an iterative CNNs framework classifier. Figure 1 shows a sample implementation of the iterative CNNs framework classifier with 44×44 as the matrix input size. The classification robustness of our iterative CNNs framework classifier to varying input matrix sizes will be presented in Section IV.

We first select the side length of the CNNs input matrices, which must be a positive integer smaller than the square root of the ACF length. The relation between the iteration number and the side length of the CNNs input square matrices is shown in (15).

$$2K - 1 = IterNum + SideLen^2, \tag{15}$$

where $IterNum$ denotes the iteration number, and $SideLen$ denotes the side length of the CNNs input square matrices.

We apply the CNNs to automatically define and extract features. Different features represent the different outputs of the CNNs from different ACF segments as shown in Fig.1.

The second step aims to classify diverse behaviors [17] and can be formulated by (16).

$$\boldsymbol{z} = G(\boldsymbol{x}), \tag{16}$$

where the function $G(\cdot)$ is unknown, with known $(\boldsymbol{x}, \boldsymbol{z})$ as the training set. The vector $\boldsymbol{z}$ always denotes a class label vector, and the vector $\boldsymbol{x}$ denotes the input features.

In classifying part, we compare several traditional classifier methods. An ensemble learning method called Random Forests (RF) [18] outperforms other methods and obtains the best performances ($e.g.$, accuracy, sensitivity, and specificity), which is shown in the simulation and measurement part of this paper. Hence, we suggest applying the RF as the final classifier, which also has a weighted neighborhood scheme.

Since a forest averages the predictions of a set of m trees with individual weight functions $\boldsymbol{W}_j$, its predictions for unseen samples x' can be expressed as

$$\hat{\boldsymbol{z}} = \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{N} \boldsymbol{W}_j(x_i, x') z_i, \tag{17}$$

where x_i and z_i come from a training set $\{(x_i, z_i)\}_{i=1}^{N}$.

This category of ensemble learning classifier constructs a multitude of decision trees at training time and outputs the class. RF are trained on different parts of the same training set and result in a variance reduction and a boost in classification performances.

Algorithm 1 summarizes our proposed ICNNRF in pseudo-code.

Algorithm 1 The ICNNRF algorithm

Require: sig, $\boldsymbol{x}$, $SideLen$, N
1: **for** ii=1: N **do**
2: Obtain $\boldsymbol{sig}$ via (2);
3: **end for**
4: Obtain $\boldsymbol{IterNum}$ via (15);
5: **for** ii=1: IterNum **do**
6: InputCNNs=$\boldsymbol{sig}$(:,ii:ii+$SideLen^2$);
7: InputCNNs = reshape(InputCNNs',$SideLen$,$SideLen$,N);
8: Obtain $\boldsymbol{y}$ by solving (14) via CNNs;
9: **end for**
10: Solve (16) via Random Forests.

III. Simulation Setup and Numerical Results

One of the objectives of this article is to investigate the reliability of our algorithms by comparing accuracy of other classification method with various signal-to-noise ratio. For this, we need to generate simulated radar returns via motion capture (MoCap) data. Because MoCap-based animations can be used to simulate complex motions, many publications such as [19-23] make use of simulated radar returns via MoCap-based data. Even though the accuracy of the MoCap data cannot be guaranteed as some noise exists in positioning the joints during motion as the markers are considered to be affixed on a rigid structure where in fact they are fitted on skin or clothing that stretch during movements. This being said formulating equations of motion for complex motion such as the Boulic model [24] for walking is very time consuming and is a generalized model of walking. We believe that variability in the performance of an action, noise and different morphology will enhance the classification performance by identifying common trends in the actions performed by the individuals. The generalization is deported from the formulation of equations of motion to the machine learning algorithms. Furthermore, this gives us access to a wealth of complex motions that would be near impossible to model mathematically $e.g.$ salsa dancing.

The simulations were checked manually to verify the quality beforehand to avoid anomalies in motion and outliers. From verified MoCap data, the noise in joint positioning is below

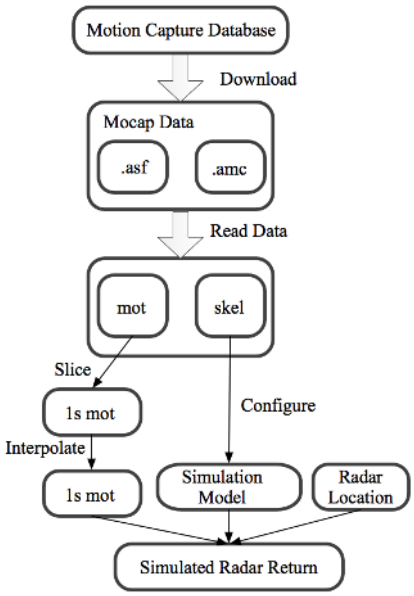

Fig. 2: The flowchart of how the CMU data and basic functions are performed to get the simulated radar returns.

the range resolution of the simulated radar ($2cm$) and therefore does not affect the simulation result as the range is rounded to the nearest integer multiple of the range resolution to simulate radar returns.

The MoCap data in this paper comes from Carnegie Mellon University (CMU) Motion Capture (MoCap) Database [25, 26] and the simulation process is illustrated in Fig. 2 that we developed in [19, 20]. This method is used to assess the performance for different configurations of our method and other classic techniques and their robustness against noise.

In this flowchart, the simulation model is set up as shown in Fig.3.

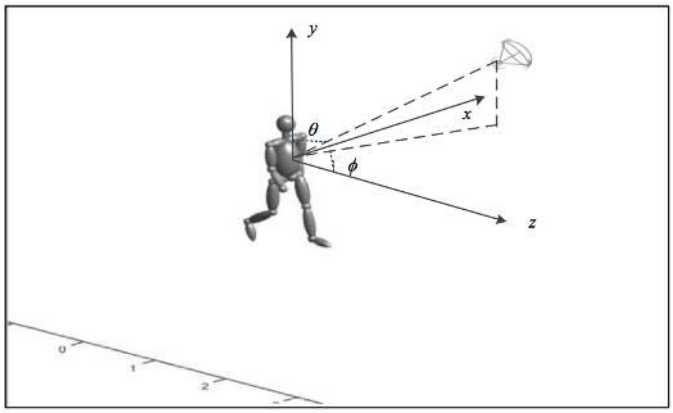

Fig.3: Human walk model along the positive direction of the z-axis extracted based on Motion Capture data from Carnegie Mellon University shown with a sample radar position with respect to the target for simulation of micro-Doppler signatures ($\theta = 78.7°, \phi = 0°$).

The maximum micro-Doppler shift in the radar echo can be written as

$$\{f_d\}_{max} = \frac{f_{carrier} v \cos\theta}{c}, \tag{18}$$

where $\{f_d\}_{max}$ denotes the maximum Doppler shift, and $f_{carrier}$ denotes the carrier frequency of the radar, v and c the target instantaneous motion velocity and speed of light respectively, as well as θ and ϕdenotes the radar inclination and rotation angles, which is illustrated in Fig.3.

The micro-Doppler shift can be expressed with respect to the maximum micro-Doppler shift as

$$\{f_d\} = \{f_d\}_{max} \sin(2\pi f_v + \psi), \tag{19}$$

where $\{f_d\}$ denotes the Doppler shift, and f_v denotes the motion frequency of the body part, as well as ψ denotes the initial phase.

Considering (18) and (19), the radar returns from the whole body can be modeled as the sum of the contributions from the different body parts (legs, arms, torso, etc.) as shown in (20).

$$\sum_{p=1}^{q} f_{dp} = \sum_{p=1}^{q} \frac{f_{carrier} v_p \cos\theta_p}{c} \sin(2\pi f_{vp} + \psi_p), \tag{20}$$

where q denotes the number of body parts considered to model radar signatures, whereas v_p and θ_p denote the instantaneous motion velocity and radar inclination angle of the p-th parts of the body respectively, and meanwhile f_{vp} and ψ_p denote the motion frequency and the initial phase of the p-th parts of the body respectively.

By sampling contributions at various fixed instants in time and integrating the radar cross section of the different body parts, the micro-Doppler signature-spectrogram is reconstructed one slice at a time, which is inspired by Victor. C. Chen's book [24]. However, the classical Boulic model [27, 28] parameter is not used in this case, instead the motion data is directly extracted from the CMU MoCap Database [25] that recorded live human motions. The complex movements and realistic physical interactions can be recreated in a physically/anatomically accurate manner, such as secondary motions, weight and exchange of forces. This ensures more realistic simulations. The Cartesian coordinate extraction from ASF/AMC files from the database was produced via the HDM05 ASF/AMC parser [26]. The radar has a pulse repetition frequency (PRF) of $1kHz$, a carrier frequency of $5.8GHz$ and$[x = 0, y = 2, z = 10]$ meters for the radar position as shown in Fig. 3 giving a velocity ambiguity of $\pm 12.9m \cdot s^{-1}$. The range resolution is $2cm$ in the simulation. The simulation produces range data with the appropriate phase delay for Doppler processing [1]. The details of the algorithm can be found both in [19, 20, 24].

First, it is assessed using simulations from 3 activities, including 2 kinds of walks (normal and fast with long strides) and running. These activities are moving along the radar radial line of sight as shown in Fig. 3. Fig. 4 illustrates the raw range data (a, b and c), ACFs (d, e, and f), spectrograms (g, h, i) of these activities showing snapshots of 1.008 seconds. The subfigures from every column of the Fig. 4 originate from one piece of simulated data. All these three kinds of raw range data simulated from the CMU MoCap database, which are numbered as subject#07(Trial#01), subject#08(Trial#01), and subject#09(Trial#01) in the database.

The spectrograms are produced by employing a Short-Time Fourier Transform (STFT) over the 128 frequency bins of the raw range data accumulated over time. The STFT is executed via the function (tfrstft.m) from the tftb toolbox [29] with a Hamming window over the length of the signal.

In these subfigures, the MoCap data sampling frequency is $120Hz$, and therefore, it is interpolated to obtain $1kHz$ to match the pulse repetition frequency used for the physical radar setup in Section IV. This modality will ensure that our proposed method can extract features autonomously and test its robustness against noise.

From the spectrogram subfigures, discriminating between walking and running can be done visually from the Doppler bandwidth. This can be explained because the micro-Doppler shift mean, maximum micro-Doppler shift and velocity features (analytically extracted) are distinctive between walking and running activities. However, it is harder to evaluate empirically the difference between walking normally and walking fast with long strides. For this reason, features will be extracted from the radar data to perform automatic classification. It should be noted that the features extracted in our proposed method are derived from the ACF rather than the spectrogram.

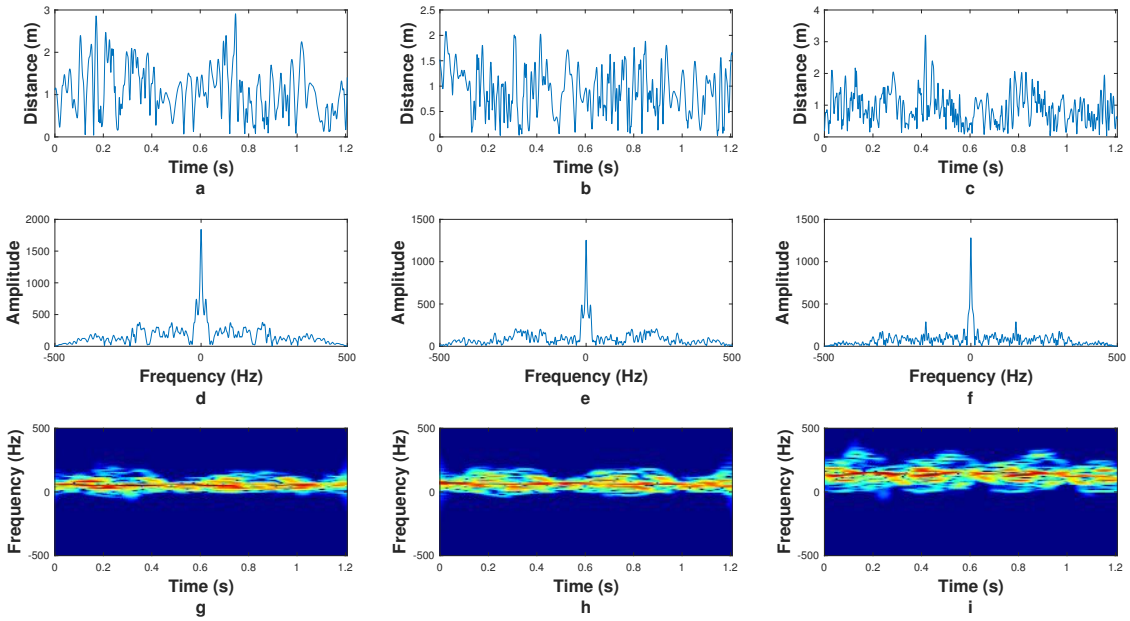

Fig. 4: Raw range data (a), ACF (d) and spectrogram (g) of walk with normal strides from the subject#07(Trial#01) in CMU MoCap Database; raw data (b), ACF (e) and spectrogram (h) of fast walk with long strides from the subject#08(Trial#01) in CMU MoCap Database; and raw data (c), ACF (f) and spectrogram (i) of the run from the subject#09(Trial#01) in CMU MoCap Database, the radar was positioned as described in Fig. 3.

The frequency of original MoCap data is 120 Hz, and the motion data is sliced into small sets of $1.008s$ - giving 121 samples per set. In general, the maximum micro-Doppler frequency shift is produced by the foot swing when walking and running in the radial direction. The resulting Doppler bandwidth may span several hundred Hertz at $5.8GHz$ and more with higher carrier frequencies. To fulfil the Nyquist-Shannon theorem for the micro-Doppler signature, the sampling frequency needs to be greater than twice the maximum Doppler frequency. Hence, the original MoCap data at $120Hz$ is violating Shannon-Nyquist theorem and would create aliasing in the Doppler domain, which would make mDs analysis, rather complicated. To address this issue, the MoCap data was interpolated to obtain a $1kHz$ sampling frequency ($\pm 12.9m \cdot s^{-1}$ in velocity with a carrier frequency of $5.8GHz$), matching the experimental Doppler sampling frequency (see section IV). It should be noted that the features used in our method are derived from the ACF rather than from spectrograms, hence the careful analysis of the interpolation effect on the spectrograms is not performed within the scope of this paper.

After interpolation, each slice now has 1008 samples. 120 datasets data were recorded from three classes with subjects 07, 08, and 09 from the CMU Mocap database. In the 3 activities (normal walk, fast walk and run), the people move along the positive direction of the z-axis, and the radar position was set to $[x = 0, y = 2, z = 10]$ meters to simulate radar returns at $5.8GHz$. From the MoCap data 120 radar returns were generated, this is not large enough to evaluate machine learning techniques. To generate a larger dataset, data augmentation is applied, and 100 different AGWN were added to the radar returns per signal noise ratio (SNR) levels. The different AWGN led to different and independent samples, a total of 12000 different data sets of radar signatures were recorded finally per SNR levels.

To assess the performance of ICNNRF, the numerical results were compared against other classification techniques. These classification techniques can be separated into 3 strategies.

- feature-based classifiers from mDs.
- our iterative CNNs (ICNN) method followed by various classifiers.
- a single CNNs with the ACF segments as input.

All the methods from the first strategy use the micro-Doppler shift mean, maximum micro-Doppler shift and the speed of walk or run as the specified features. The energy threshold was set at 93% to calculate the maximum micro-Doppler shift from the spectrogram. In the first and second strategies, the studied classifiers are LDA, QDA, KNN($K = 3$), Boosting (AdaBoostM2), Bagging, RF, SVM (kernel: 2-order polynomial function with auto kernel scale, and its box constraint is 1 with true standardization). In strategy 3, the different ACF segments selection as the input for deep learning can produce various CNNs accuracies. The mean and maximum of the CNNs accuracies represent the average and best CNNs accuracies respectively for the same ACF signal with different segments selection. The accuracy performance of various techniques is illustrated in Fig. 5. Thereinto, the accuracies in Fig.5 is the average of the accuracies from the 3

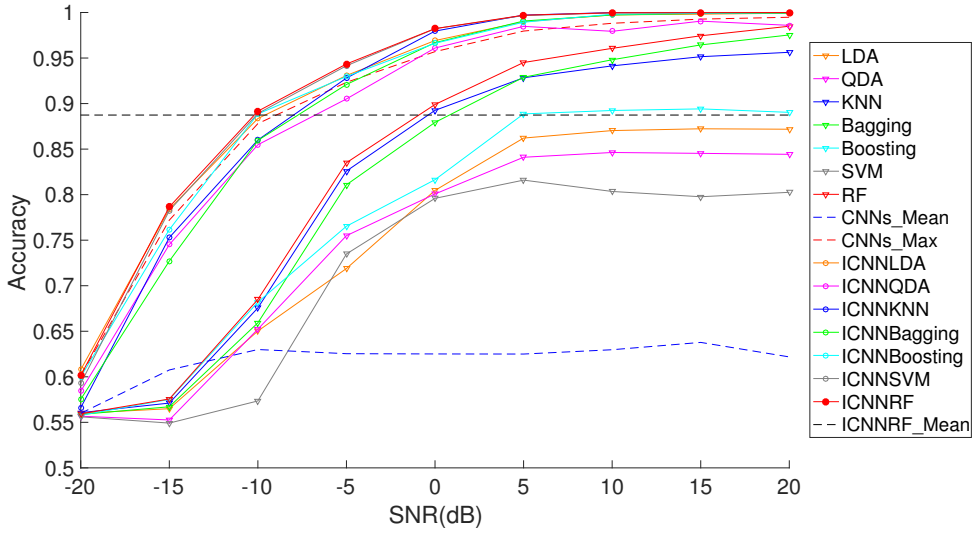

Fig.5: Average classification accuracy of 10 loops of 3 activities (walk, fast walk and run) dependence on SNR of the considered classification techniques. The ICNN and CNNs based methods extract features on their own while the others LDA, QDA, KNN Bagging, Boosting, SVW and RF as standalone techniques use 3 features for classification (micro-Doppler shift mean, maximum micro-Doppler shift and the speed of walk or run as the specified features). The radar echo is simulated by the all trials of the subject#07, subject#08, and subject#09 in the CMU MoCap Database.

activities (walk, fast walk and run) using various classification methods, which are all based on the average of 10 loops, i.e. 10 Monte Carlo Test result. The radar echo is simulated by the all trials of the subject#07, subject#08, and subject#09 in the CMU MoCap Database.

In Fig.5, the SNR levels were tested from -20 to $20dB$. For every SNR level, our ICNN strategy applied$(p = 0.3)$ as the holdout cross validation partition parameters. The training set is selected randomly 70% (8400 samples) of the total sample base (12000 samples) at a given SNR level to learn features, and the test set is 30% (3600 samples) of the total samples. The accuracy at a given SNR level is obtained by averaging the performance via a 9-fold cross validation. For the other strategies, the holdout cross validation (p=0.3) partition is applied in the training and testing partition. All the final results are the average of 10 loops.

Furthermore, based on LeNet-5 CNNs framework inspired from Deep Learn Toolbox [16] which was described in Section II, the same parameters were set to obtain the CNNs label and ICNN features, which include learning rate, batch size, epoch number, input size, and the details from 5-layer fully connected feedforward CNNs framework. The learning rate parameter is set to 1, the batch size is 120 and learning is running for 10 epochs. The sample ($i.e.$, ACF) length is $2015(2K-1)$, because the 1.007-second of data has 1008 samples(K) after interpolation. Herein, we select 44×44 as the input matrix size for the deep learning, and by updating the matrix with one new sample on every iteration 79 features are extracted using the ICNN framework.

It is difficult to obtain the maximum accuracy of the CNNs method because we do not know which segment of the ACF signal can produce the best accuracy via CNNs. Hence, the CNNs performance should mainly be the performance displayed by the mean of the CNNs accuracy. In this article, our ICNN strategy can be considered as a cascade of many CNNs. Owing to (15), the whole ACF was used for the ICNN, and thus all the information of the ACF can be gained via the outputs of the ICNN. Our ICNN strategy outperforms other strategies, because features from the other strategies are incomplete, which is consistent with Fig. 5. The black-dashed line in Fig. 5 demonstrates the average accuracy of ICNNRF with SNR changing between -20 and $20dB$. ICNNRF yielded an average accuracy of 88.74% and peak accuracy of 100% at -10 and $15dB$ respectively.

IV. Measurements and Discussion

In this section, the data were collected using an off-the-shelf frequency-modulated continuous wave (FMCW) radar system in an indoor meeting room at the School of Engineering of the University of Glasgow, where multiple pieces of furniture such as chairs, tables, cupboards, blackboards, and computers were present, as shown in Fig. 6 left. The radar Ancortek SDR-kit 580AD was operated at $5.8GHz$, with an instantaneous bandwidth of $400MHz$ (spatial resolution $37.5cm$) and a chirp duration of $1ms$ yielding an unambiguous Doppler frequency range of $\pm 500Hz(\pm 12.9m \cdot s^{-1})$,which was sufficient to capture the whole human micro-Doppler signature for indoor activities. The transmitted power of the radar was $+19dBm$, and two linearly polarized Yagi antennas with a gain equal to $17dBi$ and beam width of 24° in azimuth and elevation were used. The antennas were located at a height of approximately $1.2m$ to aim at the torso of the human subjects, which provided the strongest contribution to the micro-Doppler signature. The separation between the transmitter and receiver antennas was approximately $40cm$ as shown in Fig. 6 right. The radar was in a corner of the room to have good visibility of the area where

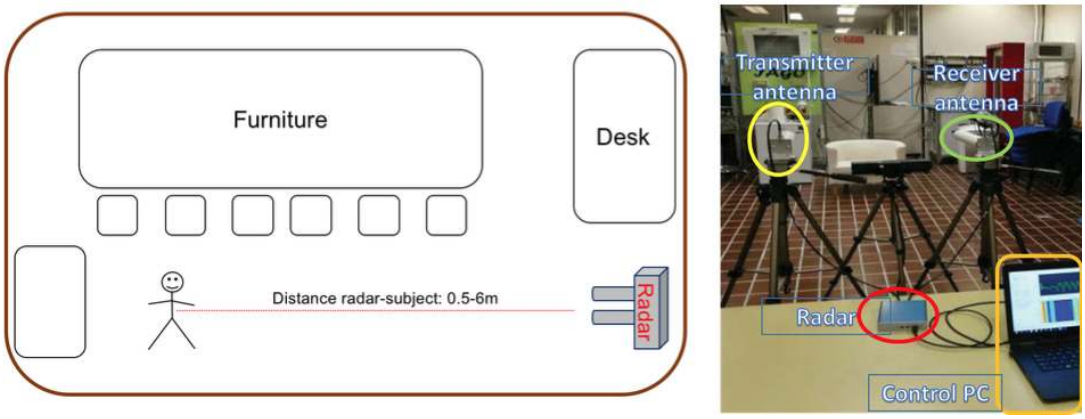

Fig. 6: left) Simplified sketch of the laboratory in room 633 James Watt South Building in University of Glasgow; right) example of radar setup in lab conditions as used to generate data for [20].

the human subjects were moving. In both cases in room 633 and [20], the radar set up was identical.

The experimental data collected from five volunteers and seven different activities are analyzed. The subjects who took part in the experiment included 3 male and 2 female volunteers, with body parameters such as height and weight ranging from $160 - 185cm$ and $55 - 75kg$ respectively. The seven activities were (I) walking; (II) moving arm faster towards radar, slower away; (III) sitting and standing; (IV) circling arm forwards; (V) clapping; (VI) bending to pick up an object and standing back up; and (VII) moving arm slower towards radar, faster away. The walks were performed from just in front of the radar to 6 m, whereas the other movements were performed on the spot at a distance of approximately 5 m from the radar system.

In the first measurement in this paper, seven activities of the same subject are analyzed to investigate the effect of the input matrix size and epoch number on the performance of our ICNN strategy. In the second experiment considered in this paper, data from all seven subjects performing the same type of walking are examined to assess the classification performance.

In Fig. 7, every row shows a different activity; each row presents the range data, the ACF and the mDs for that specific activity. Our ICNN method uses the ACF as input as opposed to traditional feature-based classification using mDs. To assess

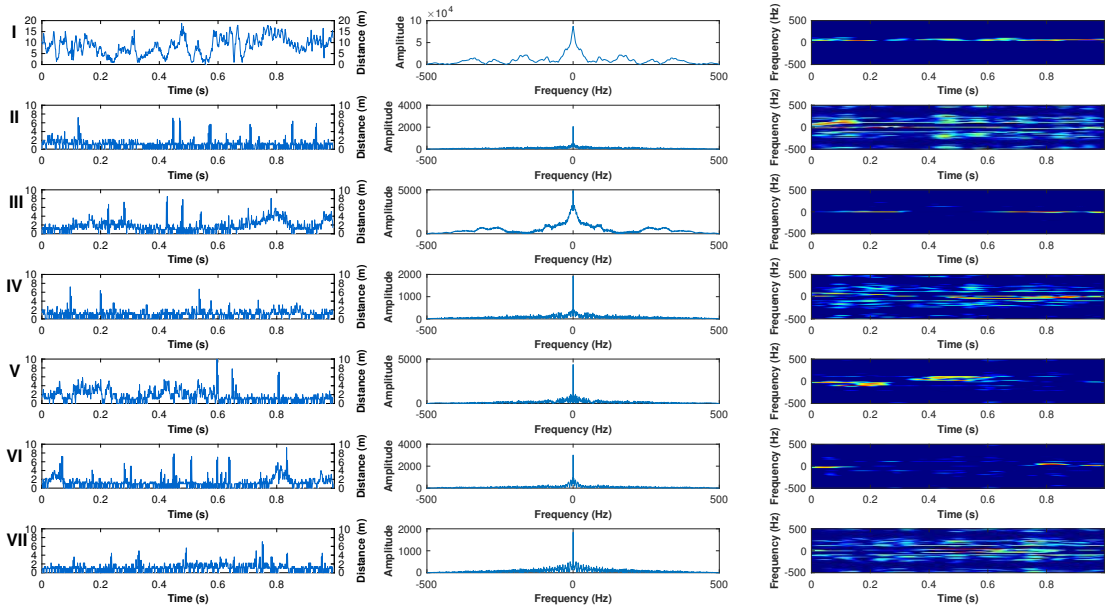

Fig.7: From left to right) The measured range data, ACFs, and mDs of 7 activities, from top to bottom) (I) walking; (II) moving arm faster towards radar, slower away; (III) sitting and standing; (IV) circling arm forwards; (V) clapping; (VI) bending to pick up an object and standing back up; and (VII) moving arm slower towards radar, faster away.

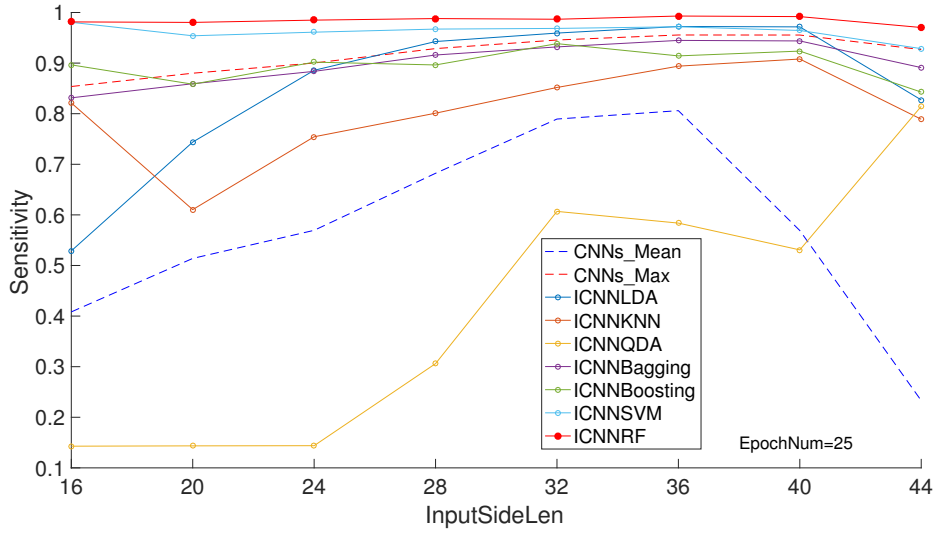

Fig.8:Average classification sensitivity performance of deep learning method based on various side lengths of input matrices over 10 loops.

the classification performance of the ICNNRF, two tests were performed activity and individual/person classifications.

Due to the application of a 4th order moving target indicator, the number of range samples available goes down to 997 samples in length, and hence the ACF signal length is 1993. The spectrogram generation function and parameters are the same as those in Section III, i.e., Matlab function (tfrstft.m) [29]. For the proposed strategy, the holdout cross-validation $(p = 0.3)$ partition is applied by selecting 70% of the data to learn features, and then 9-fold cross-validation partitioning is used for the for final training and testing. The final accuracy is based on the average of every fold accuracy of 9-fold cross validation partition. The confusion matrix is the summary of every fold confusion matrices of 9-fold cross validation partition. For the other strategy, the holdout cross-validation $(p = 0.3)$ partition is applied in selecting 70% of the total samples for training and 30% of those for testing. The results of all trials from both holdout and 9-fold cross validation partitions are the average of 10 loops.

The sensitivity measures the proportion of positives that are correctly identified. The equations of the metrics are described in the appendix. The first objective of the measurement is to investigate the input matrix size and epoch number effect on the sensitivity performance of the ICNN strategy. Seven different activities are measured. Every class has 960 samples, and hence the total sample number is 6720. 4704 samples are chosen as the training set to learn the features, while the other 2016 samples are tested by 9-fold cross-validation.

Figs. 8 shows that the different side lengths of input matrices those are displayed in Table I, effect on the sensitivity performance. Fig. 9 indicates the time it took the ICNN part to learn the features. The batch size is 24 and the network is trained over 25 epochs. The other parameters are the same as for simulations. The time displayed in Fig. 9 is the time required to train all the CNNs in the ICNN architecture. For input size 44×44, the iteration number is 57 resulting in

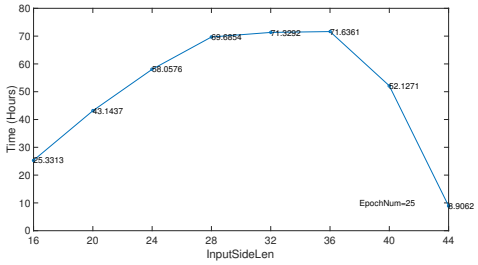

Fig.9: Elapsed time on ICNN training against input side length, i.e. the time to train one CNNs multiplied by the number of iterations.

8.9062 hours of training in total.

From the results, the larger CNNs input matrix, the better CNNs maximum sensitivity (red dashed in Fig.8). Because we do not know which segment of the ACF can produce the maximum sensitivity using CNNs, the mean CNNs sensitivity is a significant parameter. Different input matrix sizes and feature numbers produce different processing times. The processing time is equal to the iteration number ($i.e.$ feature number) multiplied by processing time for each matrix. The matrix processing time increases as the input matrix side length increases, but the iteration number decreases with increased matrix size since the ACF signal length is fixed. Hence, the processing time shown in Fig. 9 does not vary monotonously. The sensitivity of the proposed ICNNRF method is robust against varying input matrix sizes changing between 16×16 and 44×44 with different feature numbers and outperforms the CNNs maximum sensitivity alone. In the following measurements, an input side length of 44×44 is selected as it is the best trade-off between classification performance and training time to compare the different supervised learning classifiers after the ICNN.

The effect of different epoch number on the performance

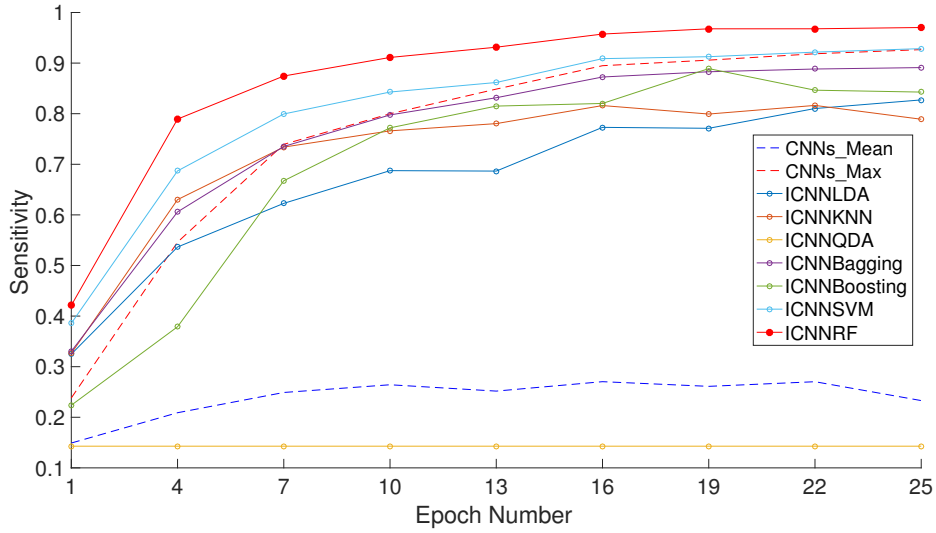

Fig.10: Average classification sensitivity performance of deep learning methods based on the number of epochs over 10 loops.

TABLE II: The Average of Accuracies, Sensitivities and Specificities of 7 Activities from the same Subject

Method	**LDA**	**QDA**	**KNN(K=3)**	**Bagging**	**Boosting**	**SVM**	**RF**
Sensitivity	0.3285	0.3905	0.5471	0.5264	0.3998	0.4644	0.5839
Accuracy	0.8081	0.8259	0.8706	0.8647	0.8285	0.8470	0.8811
Specificity	0.8881	0.8984	0.9245	0.9211	0.9000	0.9107	0.9306
Method	**ICNNLDA**	**ICNNQDA**	**ICNNKNN(K=3)**	**ICNNBagging**	**ICNNBoosting**	**ICNNSVM**	**ICNNRF**
Sensitivity	0.8272	0.8145	0.7890	0.8911	0.8427	0.9286	0.9703
Accuracy	0.9506	0.9470	0.9397	0.9689	0.9551	0.9796	0.9915
Specificity	0.9712	0.9691	0.9648	0.9819	0.9738	0.9881	0.9951

Note: The parameters in the iterative CNNs strategy : Input Matrix size= 44×44, EpochNum=25

TABLE III: The ICNNRF Confusion Matrix of 7 Activities from the same Subject

	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**Class Name**	**Sensitivity**	**Accuracy**	**Specificity**
I	286	1	0	1	0	0	0	walking	0.9931	0.9960	0.9965
II	1	279	1	4	0	0	3	moving arm fast	0.9688	0.9891	0.9925
III	1	0	283	3	0	0	1	sitting and standing	0.9826	0.9945	0.9965
IV	1	1	3	276	2	3	2	circling arm	0.9583	0.9866	0.9913
V	1	1	2	0	283	0	1	clapping	0.9826	0.9960	0.9983
VI	1	1	0	2	0	283	1	bending	0.9826	0.9960	0.9983
VII	1	9	0	5	1	0	272	moving arm slower	0.9444	0.9881	0.9954

Note: Input Matrix size= 44×44, EpochNum=25

is shown in Fig.10. The batch size is 24, and the other parameters are the same as the simulations. It is apparent that an increasing number of epochs yields improved sensitivity. Furthermore, ICNNRF learns faster than other methods, reaching better performances with fewer epochs. Table II presents the mean accuracy, sensitivity, and specificity for different classifiers; they are the average over from seven activities and 10 loops.

Table III displays the confusion matrix of the seven activities, the sensitivity, accuracy and specificity per activity using ICNNRF. The activities II, IV and VII have a lower accuracy than others, because they are all arm movements which are very similar from the radar point of view. If the activities were quite distinct then, the performances would be higher but we made the classification challenging on purpose to test the ICNNRF potential.

To evaluate the subject classification performance, five subjects performing the same activity (normal walking forward and backward) are measured. Every class has 960 samples, and hence the total sample number is 4800. 3360 samples are chosen as the training samples for training the unknown features via $(p = 0.3)$holdout cross-validation, while the other 1440 samples are tested by 9 groups of 9-fold cross-validation. In this case, the input matrix size of the deep learning is still 44×44. The batch size is 24 and the network is trained over 25 epochs. The other parameters are the same as simulations. Because the volunteers perform actions along the radar radial line of sight with the fixed position, the inclination and rotation angles ($i.e., \theta = 90^\circ$ and $\phi = 0^\circ$) are the same for all activities. Hence, the small subject variations produced by the aspect angles can be ignored as the target aspect angle with respect to the radar does not vary as he/she moves towards/away from the radar and the variability of trajectories during the experiment was negligible. Table IV shows the confusion matrix of the classification of 5 subjects and their respective sensitivities, accuracies, specificities obtained by

TABLE IV: The Confusion Matrix for the Classification of 5 Subjects Performing Normal Walking for ICNNRF and Corresponding Accuracies, Sensitivities and Specificities for Each Subject Resulting from the Average of 10 Loops

	Subject I	**Subject II**	**Subject III**	**Subject IV**	**Subject V**	**Sensitivity**	**Accuracy**	**Specificity**
Subject I	280	3	4	1	0	0.9722	0.9868	0.9905
Subject II	6	280	1	1	0	0.9722	0.9882	0.9922
Subject III	5	3	278	2	0	0.9653	0.9889	0.9948
Subject IV	0	2	0	286	0	0.9931	0.9944	0.9948
Subject V	0	1	1	2	284	0.9861	0.9972	1.0000

Note: Input Matrix size= 44×44, EpochNum=25

TABLE V: The Average Accuracies, Sensitivities and Specificities over 10 Loops of the Classification of 5 Subjects Performing Normal Walking using Various Algorithms

Method	**LDA**	**QDA**	**KNN(K=3)**	**Bagging**	**Boosting**	**SVM**	**RF**
Sensitivity	0.2806	0.3313	0.5162	0.5508	0.2833	0.3518	0.6374
Accuracy	0.7122	0.7325	0.8065	0.8203	0.7133	0.7407	0.8550
Specificity	0.8201	0.8328	0.8791	0.8208	0.8208	0.8380	0.9094
Method	**ICNNLDA**	**ICNNQDA**	**ICNNKNN(K=3)**	**ICNNBagging**	**ICNNBoosting**	**ICNNSVM**	**ICNNRF**
Sensitivity	0.8874	0.8652	0.9102	0.9139	0.9267	0.9499	0.9778
Accuracy	0.9549	0.9461	0.9641	0.9656	0.9707	0.9800	0.9911
Specificity	0.9718	0.9663	0.9776	0.9785	0.9817	0.9875	0.9945

Note: The parameters in the iterative CNNs strategy : Input Matrix size= 44×44, EpochNum=25

ICNNRF when they perform normal walking. Table V summarizes the average accuracies, sensitivities, and specificities of the classification of 5 subjects using different algorithms averaged over 10 loops.

In this case, our ICNNRF still obtains the best accuracy (99.11%), sensitivity (97.78%) and specificity (99.45%), and the sensitivities of the feature-specified classifiers have a sensitivity lower than 63.74% (Table V). This table demonstrates that the proposed method has a strong ability to classify the different subjects from the same activity.

V. Conclusion

In this paper, a novel ICNN strategy which uses range data ACF to classify activities or subjects automatically and accurately was presented. It is a two-stage algorithm that uses first ICNN to extract features automatically without image preprocessing or tuning of parameters, and then RF to perform classification. The proposed algorithm does not generate spectrograms through the Short Time Fourier Transform but uses directly the ACF of the complex range data as input to the CNNs section of the processing chain.

The proposed algorithm is validated on both simulated and experimental data. ICNNRF outperforms other methods on accuracy and is shown to be more robust to varying SNR levels over the simulated data set including 3 activities (walk, fast walk and run) based on real motion data stored in the CMU MoCap Database. Experimentally, even for 7 activities from the same subject, or for the same activity walk from 5 subjects, ICNNRF outperforms the other considered classification methods with 97.03% sensitivity for activity classification and 97.78% for subject classification. The performance of the ICNN strategy is better than that of the other alternative strategies with incomplete features gained. Moreover, the robustness of ICNNRF method against varying input matrix size changing between 16×16 and 44×44 has been demonstrated, as well as its faster learning rate and improved performance with increasing epochs. Hence, the proposed ICNNRF method is more versatile, flexible and robust than CNNs alone as little considerations need to be given to the selection of the signal segment size for classification. The ICNNRF performance surpasses that of the other ICNN plus classifier.

In this paper, we have studied the model reliability of our proposed ICNNRF algorithm with respect to SNR, input size and training time. The classification robustness is a very important to characterize, especially with respect to operational parameters such as the aspect angle to the line of sight of the radar and the range resolution. This will be performed in future work along the line of the interesting paper [30] which looks at robustness in simulation is studying the accuracy performance dependence on transmit frequency, range and Doppler resolution, antenna-target geometry, signal-to-noise ratio, and dwell time. That work shows a gradual degradation of performance from 0° to 90° aspect angle, however we need to consider that the antenna beam width is directive and that the energy received also decreases as the aspect angle changes [8]. A more thorough investigation on the aspect angle considering inclination and rotation is warranted in order to gauge how it affects the choice of classifiers, features and how different actions might be better classified with different antenna placements in order to maximize classification performance.

The influence of range measurement accuracy on performed classification should be considered. In this paper, in simulation we have used 2cm and 37.5cm in experiments obtaining over 99% accuracy in both cases. In [31], preliminary results on a novel implementation of the radar return simulations for Wind turbines shows richer information in the resulting spectrograms using a simulation that emulates radar pulses to obtain multi-domain information. Furthermore, in [32], two radar systems a continuous wave (CW) radar operating in K-band and an FMCW radar operating in C-band are used for the classification of ten indoor human activities. The classification accuracy of the CW radar was 10% lower than that of the FMCW radar. This would indicate that varying range resolution (or lack thereof) has an influence on classification accuracy and richer information in spectrograms can enhance classification.

However, further investigation is necessary to determine the exact dependence on range resolution and confirm the initial findings.

ACKNOWLEDGMENT

The authors are grateful to the volunteers who helped in the experimental data collection, in particular A. Angelov who coordinated the data collection. This work has been partly financed by EPSRC Engineering for Prosperous Nation - INSHEP, Intelligent RF Sensing for Falls and Health Prediction (EP/R041679/1), PHC Cai Yuanpei - 41457UK, CSC Cai Yuanpei Program - No. 201806070002.

APPENDIX A
STATISTICAL MEASURES OF THE PERFORMANCE

To analyze the performance of classification methods, 3 kinds of statistical measures is applying in this paper, including the accuracy, sensitivity, and specificity [33, 34], which can be expressed as

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}, \tag{21}$$

$$Sensitivity = \frac{TP}{TP + FN}, \tag{22}$$

$$Specificity = \frac{TN}{FP + TN}, \tag{23}$$

which yield

$$P = TP + FP, \tag{24}$$

$$N = TN + FN, \tag{25}$$

where P denotes the number of positive samples, and N denotes the number of negative samples. TP denotes number of true positive (labeled correctly). FP denotes number of false positive, (other activity labeled as the activity under test a.k.a false alarm). Furthermore, TN denotes number of true negative, (correct rejection), and FN denotes number of false negative (missed detections).

REFERENCES

[1] Chen,V. C., *et al.*, "Micro-Doppler Effect in Radar: Phenomenon, Model, and Simulation Study", *IEEE Trans. Aerosp. Electron. Syst.,* vol.42, no.1, pp.221, 2006.

[2] Cippitelli, E., *et al.*, "Radar and RGB-Depth Sensors for Fall Detection: A Review", *IEEE Sensors J.,* no.17, vol.12, pp.3585-3604, 2017.

[3] Chen, V. C., *et al.*, "Analysis of Radar Micro-Doppler Signature with Time-Frequency Transform", *Proc. of the IEEE Workshop on Statistical Signal and Array Processing (SSAP)* , Pocono, PA, 2000, pp.463-466.

[4] Chen, V. C., *et al.*, "Time-Frequency Transforms for Radar Imaging and Signal Analysis", Norwood, MA: Artech House, 2002.

[5] Fioranelli,F., *et al.*, "Multistatic Human Micro-Doppler Classification of Armed/Unarmed Personnel", *IET Radar, Sonar & Navigation*, vol.9, no.7, pp.857-865, 2015.

[6] Fioranelli, F., *et al.*, "Performance Analysis of Centroid and SVD Features for Personnel Recognition using Multistatic Micro-Doppler", *IEEE Geosci. Remote Sens. Lett.*, vol.13, no.5, pp.725-729, 2016.

[7] Kim Y., *et al.*, "Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks", *IEEE Geosci. Remote Sens. Lett.*, vol.13, no.1, pp.8-12, 2016.

[8] Kim, Y., *et al.*, "Hand Gesture Recognition Using Micro-Doppler Signatures with Convolutional Neural Network", *IEEE Access*, vol.4, pp.7125-7130, 2016.

[9] Kim, B. K., *et al.*, "Drone Classification using Convolutional Neural Networks with Merged Doppler Images", *IEEE Geosci. Remote Sens. Lett.*, vol.14, no.1, pp.38-42, 2016.

[10] Seyfioğlu M. S., *et al.*, "Deep learning of micro-Doppler features for aided and unaided gait recognition",*in IEEE Radar Conf., Seattle, USA,* pp.1125-1130, May 8-12, 2017.

[11] Jokanović, B.,*et al.* , "Fall Detection Using Deep Learning in Range-Doppler Radars", *IEEE Trans. Aerosp. Electron. Syst.*, vol.54, no.1, pp. 180-189, 2017.

[12] Mendis, G. J.,*et al.* , "Deep Learning based Doppler Radar for Micro UAS Detection and Classification", *Military Communications Conference, IEEE*, pp.924-929, 2016.

[13] Boyd, S., *et al.*, "Convex Optimization", Cambridge University Press, 2004.

[14] Bouvrie J., "Notes on Convolutional Neural Networks", *Neural Nets*, 2006.

[15] O'Neill, M., "Neural Network for Recognition of Handwritten Digits ", Dec.5, 2006. Online: http://www.codeproject.com/KB/library/NeuralNetRecognition.aspx

[16] https://github.com/rasmusbergpalm/DeepLearnToolbox

[17] Sugiyama M., "Graphic Machine Learning ", Chinese Edition, Kodansha Beijing Culture Ltd.

[18] Chan, C. W., *et al.*, "Evaluation of Random Forest and Adaboost Tree-based Ensemble Classification and Spectral Band Selection for Ecotope Mapping using Airborne Hyperspectral Imagery", *Remote Sensing of Environment*, 2008, vol.112,no.6, pp.2999-3011, 2008.

[19] Lin, Y., et al. , *et al.*, "Performance Analysis of Classification Algorithms for Activity Recognition Using Micro-Doppler Feature", *International Conference on Computational Intelligence and Security. IEEE Computer Society*, pp.480-483, 2017.

[20] Shrestha, A., *et al.*, "Feature Diversity for Fall Detection and Human Indoor Activities Classification Using Radar Systems", *In: RADAR 2017: International Conference on Radar Systems*, Belfast, UK, 23-26 Oct 2017.

[21] Seyfioğlu M. S., *et al.*, "Diversified Radar Micro-Doppler Simulations as Training Data for Deep Residual Neural Networks", *Radar Conference (RadarConf18)*, 2018.

[22] Karabacak C., *et al.*,"Knowledge Exploitation for Human Micro-Doppler Classification", *IEEE Geosci. Remote Sens. Lett.*, vol.12, no.10, pp.2125-2129, 2015.

[23] Karabacak C., *et al.*,"Automatic Human Activity Classification using Radar", *IEEE Signal Processing and Communications Applications Conference*, pp.1051-1054, 2014.

[24] Chen,V. C., "Micro-Doppler Effect in Radar", Artech House, 2011.

[25] http://mocap.cs.cmu.edu

[26] M. Müller., *et al.*, "Documentation Mocap Database HDM05", *Technical report*, No. CG-2007-2, ISSN 1610-8892, Universität Bonn, June 2007.

[27] Boulic, R., *et al.*,"A Global Walking Model with Real-time Kinematic Personication", *Visual Computing*, vol.6, pp.344-358, 1990.

[28] Erol, B., *et al.*,"A Kinect-based Human Micro-doppler Simulator", *IEEE Aeros. Electron. Syst. Mag.*, vol.30, no.5, pp.6-17, 2015.

[29] http://tftb.nongnu.org/

[30] Gürbüz, S. Z., *et al.*, "Operational Assessment and Adaptive Selection of Micro-Doppler Features", *IET Radar Sonar & Navigation*, vol.9, no.9, pp.1196-1204, 2005.

[31] Ochieng, F. X., *et al.*, "Deflection Characterisation of Rotary Systems Using a Ground-based Radar", *IET International Radar Conference*, Nanjing, China, 17-19 Oct 2018.

[32] Shrestha, A., *et al.*, "Activity Recognition with Cooperative Radar Systems at C and K band", *IET International Radar Conference*, Nanjing, China, 17-19 Oct 2018.

[33] Neovius M. G., *et al.*,"Sensitivity and Specificity of Classification Systems for Fatness in Adolescents", *American Journal of Clinical Nutrition*, vol. 80, no.3, pp.597-603, 2004.

[34] Zhu W., *et al.*,"Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS Implementations", *NESUG Proceedings: Health Care and Life Sciences*, Baltimore, Maryland, 2010.