Angle-insensitive Human Motion and Posture Recognition Based on 4D imaging Radar and Deep Learning Classifiers

Zhao, Yubin; Yarovoy, Alexander; Fioranelli, Francesco

DOI
10.1109/JSEN.2022.3175618

Publication date
2022

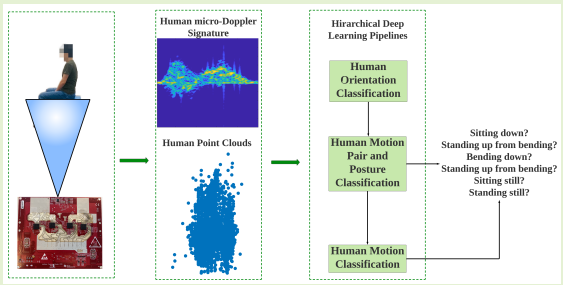
Document Version
Accepted author manuscript

Published in
IEEE Sensors Journal

Angle-insensitive Human Motion and Posture Recognition Based on 4D imaging Radar and Deep Learning Classifiers

Yubin Zhao, Alexander Yarovoy, *Fellow, IEEE*, Francesco Fioranelli, *Senior Member, IEEE*

***Abstract*— The need for technologies for Human Activity Recognition (HAR) in home environments is becoming more and more urgent because of the aging population worldwide. Radar-based HAR is typically using micro-Doppler signatures as one of the main data representations, in conjunction with classification algorithms often inspired from deep learning methods. One of the limitations of this approach is the challenging classification of movements at unfavorable aspect angles (i.e., close to 90°) and of static postures in between continuous sequences of activities. To address this problem, a hierarchical processing and classification pipeline is proposed to fully exploit all the information available from millimeter-wave (mm-wave) 4D imaging radars, specifically the azimuth and elevation information in conjunction to the more conventional range, Doppler, received power, and time features. The proposed pipeline uses the two complementary data representations of Point Cloud (PC) and spectrogram, and its performance is validated using an experimental dataset with 6 activities performed by 8 participants. The results show good performance of the proposed pipeline compared with alternative baseline approaches in the literature, and the effect of key parameters such as the amount of training data, signal-to-noise levels, and virtual aperture size is investigated. Leave-one-subject-out test is also applied to study the impact of body characteristics on the generalizability of the trained classifiers.**



***Index Terms*— Human Activity and Posture Recognition, Imaging Radar, Multiple Angle Classifier, Deep Learning**

I. Introduction

The increasingly aging population and the prevalence of non-communicable diseases have made in recent years more and more urgent the capability to monitor patient health at home. This raises the importance of indoor human activity recognition (HAR) enabling automatic monitoring systems to potentially improve life quality, reduce hospitalization, and most importantly provide timely help in case of emergencies, such as in case of serious fall or stroke events [1], [2].

From a technical perspective, automated HAR was originally mostly based on visual aids [3] or wearable sensors [4], [5]. However, both types of sensors - cameras of various kinds and inertial measurement units - exhibit inherent limitations [6], such as poor functionality in darkness or intense light conditions as well as potential privacy issues for cameras, and end-users' inconvenience to carry or wear sensors that might lead to poor compliance and usage. On the other hand, radar is gaining attention for its potential advantages: it provides consistent sensing quality regardless of light conditions, and works contactless, without end-users needing to carry, wear, or interact with the sensor [7], [8].

The basic principle of radar-based HAR is that each human activity has unique kinematic patterns, and such patterns can be represented by measurable features such as the velocity of different human body parts along with the physical extent of their movements, namely micro-Doppler signatures [9], [10]. Radar-based HAR methods are generally data-driven and can be divided into two categories in terms of the applied data representations and corresponding classification algorithms. To the first category belong, studies where handcrafted features are extracted from the pre-processed radar data and then used together with supervised Machine Learning algorithms. For instance, studies such as [11]–[13] used handcrafted features and machine learning algorithms (e.g. Support Vector Machine, SVM) to classify different human activities, intended as individual motions performed by human subjects (e.g., walking, sitting still, boxing, and so on). To the second category belong, other studies where radar data are represented and treated as image-like or video-like inputs which are treated by means of Deep Learning algorithms. For example, spectrogram images were used as the input to Deep Convolutional Neural Network models [14]–[16], and sequences of spectrogram images were processed by Long-Short Time Memory models, a variant of recurrent neural network [17], [18].

Radar-based HAR approaches that depend strongly on the

This work was in part supported by the NWO KLEIN (M1) *RAD-ART* project awarded to F. Fioranelli.

Y. Zhao, A. Yarovoy and F. Fioranelli are with the Faculty of Electrical Engineering, Mathematics and Computer Science, TU Delft, Delft, The Netherlands.(e-mail: F.Fioranelli@tudelft.nl).

Doppler information suffer from two inherent limitations. Firstly, the classified movements are typically artificially limited to be performed along the line-of-sight direction of the radar, or at a small aspect angle so that the Doppler information remains representative enough. The second limitation is that static postures on their own or in between continuous sequences of movements are rarely investigated, since their micro-Doppler signatures are not easily distinguishable from the static clutter.

To overcome the first limitation, the usage of distributed or even multistatic radar systems has been proposed, in order to simultaneously sample and reconstruct the micro-Doppler signatures from different aspect angles [18]–[22]. While effective, this approach requires the usage of multiple nodes, with an increase in complexity of the overall system and the need to cope with the synchronization of data from the different sensor nodes. Another approach, recently proposed thanks to the availability of millimeter-wave (mm-wave) MIMO (Multiple Input Multiple Output) radars, is based on the use of 4D imaging radar that exploits azimuth and elevation resolution capabilities to attain additional spatial information on the subject's posture [23], [24]. While promising, this approach is still not investigated in detail in the literature and there is a scope to define effective classification processing pipelines that can take the advantages of mm-wave 4D imaging capabilities for HAR.

To address the aforementioned radar-based HAR issues of unfavorable orientations (in terms of aspect angles with respect to radar's line-of-sight) and static human postures, this work proposes a radar-based classification pipeline that exploits the richer information provided by mm-wave 4D imaging radar. The main contributions of the proposed pipeline are as follows.

- Unlike in the other studies such as [23] where the radar point clouds are treated as images, the proposed pipeline is designed with the goal to exploit all the six intrinsic 'features' obtained by imaging radar, namely range, azimuth, elevation, Doppler, received power, and time information, rather than focusing on just one specific data representation.
- The hierarchical structure of the proposed pipeline simplifies the task of HAR in a multi-angle scenario, with the help of the designated neural networks, *T-Net*, and achieves classification of both static postures and dynamic motions.
- The pipeline is designed to be robust to a noisy and limited amount of radar data by replacing the Max Pooling layer in T-Net and PointNet with Average Pooling, and deliberately using lighter-weight neural networks, respectively.

To validate the performance of the proposed pipeline, a custom experimental dataset is collected including 8 human subjects performing 6 in-place activities (including 4 motions and 2 postures) at 5 different orientations. The measurement was conducted in an office-like room to simulate an indoor real-life environment. The proposed pipeline attains the classification accuracy of 87.1%, which is significantly higher than the state-of-the-art alternatives applied to the same dataset. The proposed pipeline also shows robust results in case of low signal-to-noise ratio (SNR), varying dimensions of the virtual apertures (i.e., the number of array channels that worsen or improve angular resolutions), and leave-one-subject-out test to validate performance for unseen individuals.

The rest of the paper is organized as follows. Section II describes the proposed method. Section III presents the measured dataset for validating the performance of the proposed method. Section IV discusses the attained results for the proposed method and its comparison with the state of the art. Finally, conclusions are drawn in Section V.

II. Description of the Proposed Pipeline and Comparative Baselines

Radar conventionally used in HAR (i.e., radar with a single receiver and operating at the relatively low carrier frequency in the 5.8 or 24 GHz ISM bands) generates data from which information related to four intrinsic features of the object can be extracted: range as 1D spatial information, Doppler proportional to the target's radial velocity, received power proportional to the Radar Cross Section of the object, and the temporal relations from the movements of the body parts. The usage of these four features or representations of the data has been thoroughly analyzed in the literature, and it appears that the state-of-the-art research mostly relies on the Doppler information [25]–[28].

To overcome the inherent limitations of Doppler information, additional intrinsic features must be introduced. 4D imaging radar at mm-wave frequencies can provide an estimation of the spatial occupancy of the human body in height and width in different positions. Specifically, the usage of multiple channels in an antenna array allows estimating the angles of arrival of the targets. At mm-wave frequencies, human bodies are perceived as extended targets, with multiple scatterers generated by each moving body part forming the so-called point clouds (PCs) [23]. This scattering behavior, combined with the angular estimation capabilities on both azimuth and elevation, enables a new, broader 'feature space' to explore for radar-based HAR. However, when operating at mm-wave, a disadvantage to account for is that the detection range is shorter than at lower frequencies conventionally used for HAR such as the ISM bands of 2.4, 5.8, and 24 GHz, because of the higher propagation losses. Nevertheless, current mm-wave systems at 60-77 GHz show good detection and classification capabilities at the ranges of interest for HAR applications.

The overview of the proposed method to exploit all these features is given in Figure 1. Specifically, this method exploits the six intrinsic features of range, azimuth, elevation, Doppler, received power, and time by combining both PCs and spectrograms as the input data representations. The hierarchical structure of the classification pipeline includes the so-called '*orientation classification module*' to first classify to which orientation the human subject is facing toward (e.g., 0, 45, 90, 135, 180°). Then, based on the predictions made by this module, the '*PC classification module*' predicts which posture or motion pair the input belongs to (more details about the definition of motion pair is given later in this section). Finally,

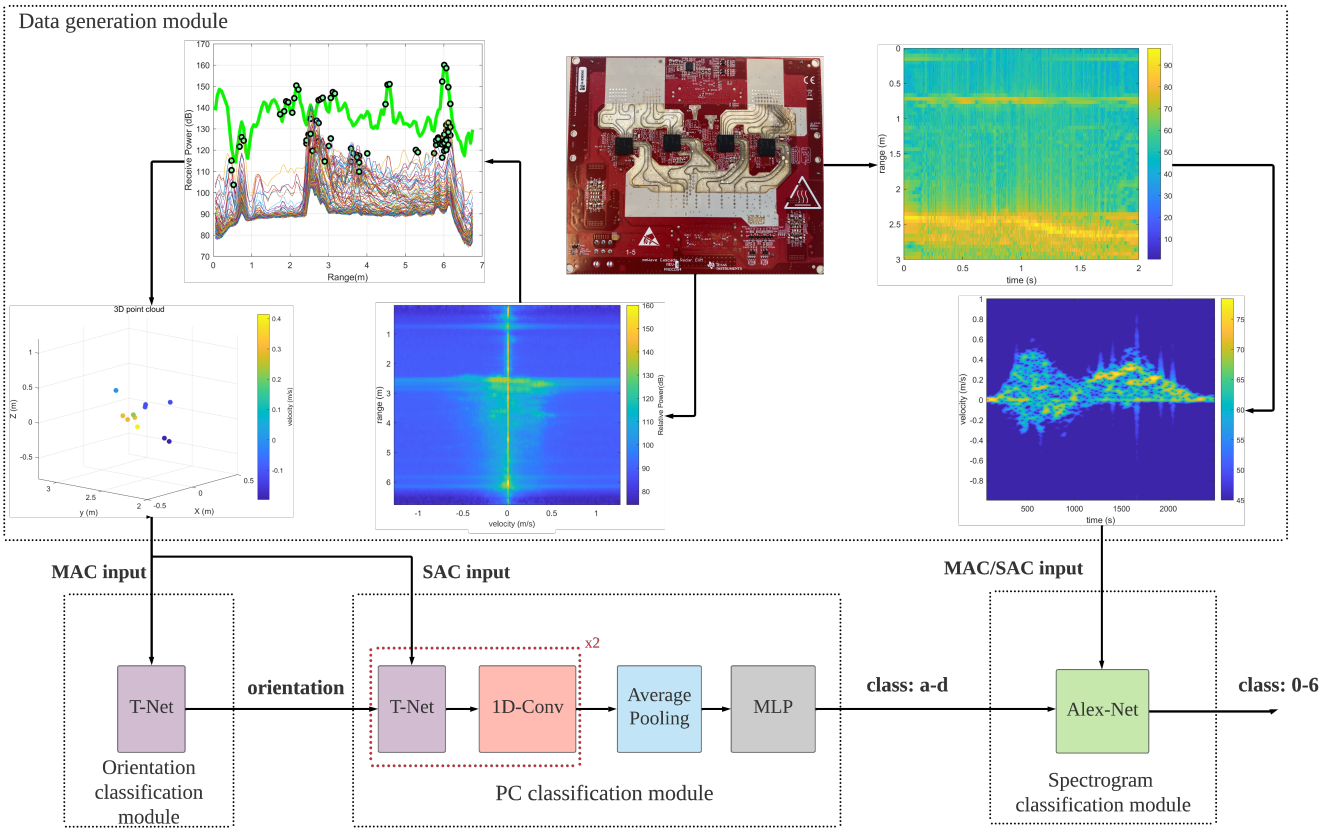


Fig. 1. Overview of the proposed classification pipeline, where the main contributions are the parallel processing and fusion of PCs and spectrograms, and the usage of T-Net to obtain angular orientation insensitivity. Besides the data generation and pre-processing operations (top box), the three main modules include the orientation classification, the PC classification, and the spectrogram classification.

for those inputs predicted as a motion pair, the '*spectrogram classification module*' is then utilized to classify to which specific motion the input sample belongs, e.g., bending over or standing up from bending, or, sitting down or standing up from sitting. The descriptions of the main modules are as follows.

1) The **data generation module** starts with measuring experimental data via the imaging radar. Given the measured raw imaging radar data, the signal processing flow involves two parallel branches: *(a)* 2D Fast Fourier Transform (FFT) is applied on fast-time and slow-time domains to estimate range and Doppler spectra; then 2D FFT is also applied on the virtual channel domain to obtain the azimuth and elevation information. PCs are generated by applying the order-static constant false alarm rate detector in the range-Doppler domain, subsequently detecting local peak values in the azimuth domain and the global peak value for the elevation domain. Each coherent processing interval (frame duration) is 100ms. PCs from 20 frames are aggregated to represent one segment of activity, since the PC generated from a single coherent processing interval is typically not dense enough to represent the shape of a human body [23]. *(b)* spectrograms are generated by applying a Short Time Fourier Transform (STFT) on the slow-time axis aggregated over the range bins where the human subject is present. The specific parameters used for the data generation module are listed in Table I.

TABLE I
PARAMETER SPECIFICS USED IN THE DATA GENERATION MODULE.

Parameter	Value	Parameter	Value
Number of range and Doppler FFT points	256	Number of azimuth FFT points	128
Number of elevation FFT points	64	Number of reference cells in range domain	8
Number of reference cells in Doppler domain	4	Number of guard cells in Doppler domain	12
Number of guard cells in Doppler domain	4	Scaling factor for OS-CFAR algorithm	6.3
Window length for STFT	128	Overlapped window length for STFT	127
Power threshold below the maximum power level per spectrogram	40dB		

2) The **orientation classification module** includes a so-called T-Net architecture modified from [29]. T-Net was originally used to transform the input PC to achieve angle insensitivity. T-Net outputs a 3×3 transformation matrix used to be multiplied with the input PC such that the rotated PCs are transformed to the same aspect angle. Essentially, the feature learning layers within T-Net can learn the geometric characteristics of PCs, which are related to the human orientation. Therefore, in the proposed implementation the multi-perception layers in the original T-Net are modified to be compatible with a classification task, and the max pooling layer is replaced with the average pooling layer to achieve more robustness to noisy data. Based on the predictions on human orientation, the multi-angle human activity

classification problem is simplified to be a uni-angle problem.

3) For each predicted orientation of the human subject, the **PC classification module** is then used. This module is adapted from PointNet [29] by proposing a new global symmetric function average pooling to replace the original max pooling in order to be robust to noise. The PC classification module makes a prediction according to the spatial distribution of PCs aggregated over 20 frames. As previously mentioned, without this aggregation, the sparsity of the radar PC over one or few frames would make the classification task for human movements or postures too challenging. An undesired consequence of this aggregation over time is that dynamic activities and motions that are implicitly shorter than the duration of 20 frames would result in very similar PCs, for example sitting down & standing up from sitting, or bending over & standing up from bending. Therefore, the output of this module may include classes that are possible motion pairs (e.g., *(a)* sitting down on a chair or standing up from sitting, and *(b)* bending over or standing up from bending), or static postures such as *(c)* sitting still and *(d)* standing still. These are summarized in Table II.
4) The last module is the **spectrogram classification module**. This module takes the predicted class from the previous module and uses the spectrogram as the input to the AlexNet [30] to recognize the individual motion class within the motion pair *(a)* or *(b)*. The choice of AlexNet [30] is due to its relatively simple structure and easier convergence compared to deeper neural networks, but it theoretically can be replaced by other spectrogram-based classification approaches from the literature [25], [26] if desired. Therefore, the final classification output of the proposed pipeline is a combination of modules using both PC and spectrograms as input data representations.

As claimed by [31], an angle-insensitive HAR pipeline should be robust to achieve angle-insensitive HAR given training data collected at multiple orientations, or even ideally given training data collected at one orientation only. Thus, two definitions are given for how to use a classification pipeline in terms of different combinations of training data with respect to human orientation. These include training with data from one human orientation and testing with multiple orientations, which is termed as *SAC* (Single Angle Classifier), and *MAC* (Multiple Angle Classifier) if training is performed with data collected at multiple orientations. It should be noted that the orientation classification module becomes insignificant for SAC cases, and thus is bypassed in the proposed pipeline. Angle Sensitivity Matrix (ASM) and Angle Sensitivity Vector (ASV) are used as metrics to evaluate the classification performance of SAC and MAC, respectively, as proposed in [31]. The former has two dimensions of training and test orientations, and the latter is compacted from a matrix into a vector, as data of all orientations are used for training together. The work in [31] is used as a comparative baseline to evaluate the proposed pipeline and is named as *baseline-1*.

Furthermore, additional comparisons are made with other pioneering studies investigating the feasibility of applying imaging radar for radar-based HAR. In particular, [23] used 'snapshots' of PCs aggregated over the interval of activity as the input to their designed deep convolutional neural network. Their method essentially treats imaging radar-based HAR as a 2D image classification problem, unlike the proposed pipeline that processes the 3D coordinates of the detected PCs. It is crucial to determine whether the full use of all the intrinsic features provided by 4D imaging radar helps to achieve better HAR performance, so [23] is used as a replica of *Orientation classification module*, *PC classification module* and the entire pipeline for comparison. This comparative architecture is named *baseline-2*. Furthermore, since the classifier in [23] is not explicitly compared with other image-based classifiers in the literature, four of them - namely, VGG [32], ResNet [33], DenseNet [34] and ViT [35] - are also used for comparison in this paper. They are named *baseline-3* to *baseline-6*.

III. Dataset Description

To validate the performance of the proposed pipeline, a mm-wave MIMO FMCW radar developed by Texas Instruments (cascaded AWR2243 radar) was used to collect an experimental dataset. Mm-wave FMCW MIMO radar has been a popular choice in short-range applications such as indoor HAR thanks to its flexibility, low cost, and small physical size as an off-the-shelf product, mostly driven by the technological development in the automotive sector. Specifically, the cascaded AWR2243 radar operates at 79GHz with 12 transmitters and 16 receivers. Using MIMO configuration, the radar produced a virtual array with an aperture of $43\lambda \times 3\lambda$ (see Figure 2), where λ is the center wavelength of the transmitted signal. For the purpose of comparative studies, we mostly inherited the FMCW waveform parameters from [23], which are given as follows: chirp duration 63 μsec, chirp slope 60 MHz/μsec, chirps per frame 128, frame period 100 ms, frequency bandwidth 2.84 GHz, and A/D sampling rate 2.7 MHz. The radar specifications derived from these parameters are as follows: range resolution of 52.8 mm, azimuth resolution of 1.4 degrees (at broadside), elevation resolution of 18 degrees (at broadside), velocity resolution of $\pm$0.0286 m/s. These resolutions are expected to provide sufficient information on human dynamics as well as body shapes.

TABLE II
LIST OF MOTIONS, MOTION PAIRS, AND POSTURES FOR THE MEASURED DATASET. MOTION PAIRS AND POSTURES REPRESENT THE INTERMEDIATE OUTPUT CLASSES (A-B-C-D) OF THE PC CLASSIFICATION MODULE, WHEREAS INDIVIDUAL MOTIONS AND POSTURES REPRESENT THE FINAL OUTPUT CLASSES OF THE PROPOSED PIPELINE (1-6).

Motion pair	Motion	Posture	
a	1. Sitting down	c	5. Sitting still
	2. Standing up from sitting	d	6. Standing still
b	3. Bending over		
	4. Standing up from bending		

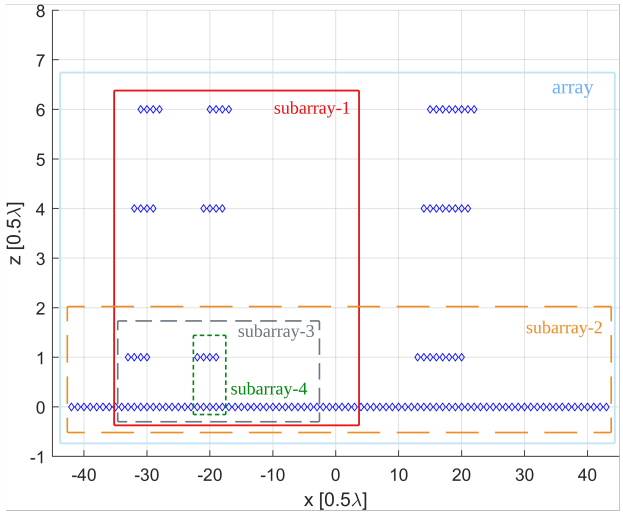


Fig. 2. Visualization of the virtual channels in the used AWR2243 imaging radar, where *array* and *subarrays* are defined according to the number of channels used for the subsequent signal processing.

TABLE III
BODY CHARACTERISTICS OF THE 8 PARTICIPANTS.

Subject index	1	2	3	4	5	6	7	8	Mean±Std
Height (cm)	180	168	170	180	185	178	177	177	176.9±5.5
Weight (kg)	75	70	70	70	95	82	72	72	75.7± 8.8

The dataset was collected in an office-like room with tables, chairs, and cabinets to simulate a real-life indoor environment. The radar was placed at 0.75m height from the ground to illuminate the whole human body in the field-of-view. A chair was placed 2.7m away from radar in the Y-axis direction and participants performed activities around it. Overall, 6 activities were included in the dataset, consisting of 4 most common daily motions, and 2 postures that can be viewed as the transitional states between such motions (Table II). The measurements include eight human subjects, whose body characteristics are very similar, as shown in Table III. During the measurements, postures and motions were recorded separately. Specifically, a complete time interval for each measurement was 2 minutes. During this time, the human subjects were asked to perform either one static posture, e.g. sitting still on the chair, or a motion pair, e.g., sitting down and standing up from sitting for a period of approximately 2 seconds for each individual motion. Labels are then generated manually for these data by visual segmentation. The measured dataset consists of 2,239 samples and is divided into 80% for training and 20% for testing. In addition to the processing chain mentioned in Section II, snapshots of the front view of the aggregated PCs as in [23] are also stored for comparative studies.

To further examine the performance of the proposed pipeline in more challenging conditions, some additional datasets were generated based on the measured one. Specifically, "noisy data" were generated by adding additive white Gaussian noise to the measured raw data (i.e., prior to any processing steps to generate spectrograms or PCs) while assuming the original measured data to be noise-free. Based on the amount of added noise to raw data, the final obtained SNR values are 20dB, 18dB, 15dB, 13dB, 10dB and 8dB. Moreover, "small-aperture datasets" were generated by selecting the raw data of only a subset of virtual channels for subsequent signal processing (essentially the *subarray-1* to *4* shown in Figure 2). This was done to test the effect of reduced angular resolutions on the generated PCs and on the subsequent classification.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Results of the Proposed Pipeline

This section presents the results of the proposed pipeline classifying 6 activities. It should be noted that the training and testing are independently repeated 5 times to increase the reliability of the results, so the results presented in this section express the values averaged from 5 realizations.

The orientation classification module provides very accurate predictions of human subjects' orientation with an accuracy of 97.5%, which will significantly simplify the task for PC and spectrogram classification modules. The PC classification module also makes promising predictions with an average accuracy of 97.9% (defined as the sum of the true positive samples for classes a-d divided by the total number of samples in Table IV). The spectrogram classification module, however, has less favorable performance (83.8%) according to the sum of the number of true positive samples for classes 1-2 and 3-4 divided by the number of true samples for classes a and b, respectively, in Table IV. Specifically, the binary classification accuracy for sitting down/standing up from sitting is 85.7%, and the binary classification accuracy for bending down/standing up from bending is 81.0%. These numbers are lower than the results in some state-of-the-art studies, for example, almost 100% classification accuracy of sitting down and standing up from sitting was achieved in [27] and more than 90% in [28]. However, it should be noted that this difference in performance is related to the fact that these studies did not consider the multi-angle scenario, as more than 65% of the misclassifications of classes 1-4 are actually caused by the samples where the human subjects performed motions at 90° orientation. Considering all the above, each stage of the hierarchical pipeline exhibits satisfactory performance with classification accuracy between 81.0% and 99.0%, leading to an overall average classification accuracy of 87.1% for the proposed pipeline classifying 6 activities.

TABLE IV
CONFUSION MATRIX FOR THE PROPOSED PIPELINE TO CLASSIFY 6 ACTIVITIES, WHERE VALUES WERE AVERAGED FROM 5 INDEPENDENT REALIZATIONS.

True/Pred.		**a**		**b**		**c**	**d**
		1	**2**	**3**	**4**	**5**	**6**
a	**1**	347	73	1	0	0	0
	2	48	413	0	0	0	0
b	**3**	0	2	324	76	0	3
	4	0	1	77	332	0	0
c	**5**	0	5	0	3	432	24
d	**6**	0	5	0	5	9	443

B. Results of Ablation Study

Through the removal of specific modules in the proposed pipeline, this section is expected to study their individual contribution. For example, the PC classification module alone could be adapted to directly predict which activity the subject is performing; or, in case of the orientation classification module being removed, the PC classification module and spectrogram classification module together could accept input samples from all orientations. Table V shows the classification results attained in these ablation studies by mean accuracy/F1-score and their standard deviation from five different realizations, where each realization includes an independent training and testing. It is reasonable to conclude that each classification module makes a crucial contribution to the overall performance of the proposed pipeline. Meanwhile, the results attained with only one classification module on its own are significantly lower than the others, as the accuracy and F1-score in the first two rows are significantly lower than their counterparts in the other rows. This not only demonstrates the advantage of the hierarchical structure, but also reveals the significance of the two complementary data representations that exploit all available features of mm-wave 4D imaging radar.

TABLE V
SUMMARY OF THE ACCURACY AND F1-SCORE RESULTS FROM THE ABLATION STUDY WITH THE DIFFERENT MODULES IN THE PROPOSED PIPELINE.

Used Modules	Accuracy	F1
PC classification	74.3%±1.1%	74.1%±1.2%
spec. classification	73.5%±3.2%	74.4%±1.9%
Ori classification & spec. classification	77.1%±2.5%	76.4%±2.1%
PC classification & spec. classification	80.9%±1.0%	80.7%±1.2%
Ori classification & PC. classification	81.9%±1.8%	81.4%±1.2%
Full Pipeline	**87.1% ±1.2%**	**86.7%±1.3%**

C. Results vs the Comparative Baseline Approaches

This subsection presents the results comparing the proposed pipeline with the alternative baselines from the literature described in Section II, together with the ASM and ASV metrics from [31] as a function of training and testing angles for the different human activities[1].

When using the proposed pipeline as a SAC, the orientation classification module becomes pointless since only the data of one orientation can be used for training. As a result, only the PC and the spectrogram classification modules are utilized from the pipeline. The corresponding ASM and ASV in comparison with those of the *baseline-1* [31] are given in Figures 3 and 4, and Table VI. It should be noted that, since the conventional spectrogram-based classifiers lack the necessary spatial information to classify static postures, only four motions are considered in our implementation of the

[1]Neural networks in [31] and [23] are implemented by the authors and trained from scratch, as the pre-trained models or training data were not made publicly available

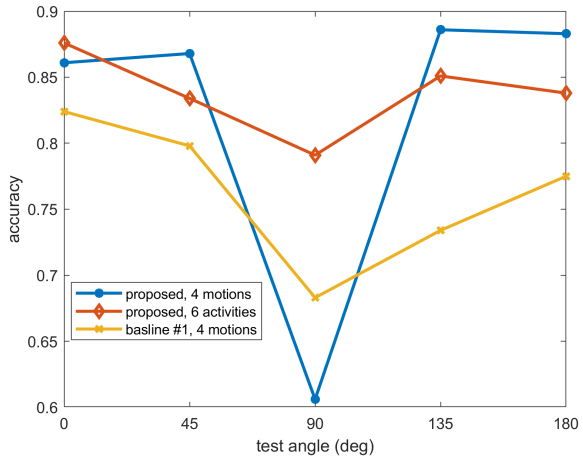


Fig. 3. ASVs of the proposed pipeline and the *baseline-1* [31] as a function of orientation.

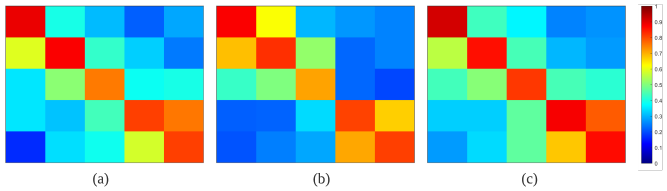


Fig. 4. ASMs of (a) the proposed pipeline for 4 motions, (b) the *baseline-1* [31] for 4 motions, and (c) the proposed pipeline for 6 activities. The horizontal axis is test angle, the vertical axis is training angle, and from top to bottom and from left to right the values are 0, 45, 90, 135 and 180°, respectively.

baseline-1. Three main points are drawn from this analysis as follows:

- Figure 3 shows that the closer 90° the human orientation is toward, the worse the results are, considering all methods. This is presumably due to the fact that movements with a torso orientation of 90° do not generate representative Doppler features for classification.
- Figure 4 shows that the proposed pipeline outperforms the method in [31] for the seen test data, as shown by the cells on the diagonal. Yet, the proposed pipeline appears to be more sensitive to angle variations than Yang's method [31] given test data collected at an orientation close to the training data (e.g., the cells on the non-diagonal upper left or bottom right parts). Meanwhile, for those results obtained from a test angle far from the angle of the training data (e.g., the cells on the upper right or bottom left parts), the proposed pipeline again shows superiority thanks to the use of additional spatial information.
- Following the definitions of quantitative metrics for angle sensitivity in [31], the mean value ($\bar{X}$) and L2-distance ($||v||_2$) of ASV and ASM are reported in Table VI. As it can be seen, the proposed pipeline mean accuracy is significantly higher than that of *baseline-1*, and the L2-distance is smaller. Therefore, quantitatively speaking, the proposed pipeline is better at classifying motions and / or postures than *baseline-1* in terms of both MAC and SAC.

Table VII shows the results in terms of accuracy and F1

TABLE VI
QUANTITATIVE METRICS INSPIRED FROM THE *baseline-1* [31]. COMPARISON OF RESULTS USING THE PROPOSED PIPELINE AND THEIR CLASSIFICATION APPROACH.

Method	Activity	MAC-based results		SAC-based results	
		$\bar{X}$	$\|\|v\|\|_2$	$\bar{X}$	$\|\|v\|\|_2$
proposed	4 motions	82.1%	9.4%	47.7%	10.5%
proposed	6 activities	**83.8%**	**7.4%**	**50.8%**	**9.9%**
baseline-1 [31]	4 motions	76.3%	10.8%	45.6%	10.9%

scores for the different baselines from the literature [23], [32]–[35] and the proposed pipeline. The proposed pipeline shows a lead 12.2% in accuracy and 13.1% in F1-score compared with *baseline-2* to *6* when classifying 6 activities end-to-end. Furthermore, the accuracy advantage of the proposed pipeline is at least 22.4% for classifying human orientations and at least 11.5% when the image classifier baselines are used as a replica of the proposed *PC classification module*. All these results appear to suggest that fully exploiting all the information provided by an imaging radar, e.g. using together spectrograms and the PC coordinates as input to classifiers, is more helpful in achieving good classification results than treating PCs as 'snapshot' images, as the latter approach implicitly causes a loss of information.

TABLE VII
CLASSIFICATION ACCURACY AND F1 SCORE OF THE PROPOSED PIPELINE VS BASELINES IN THE LITERATURE FOR THE TASKS OF ORIENTATION CLASSIFICATION, PC-BASED CLASSIFICATION, AND OVERALL HAR WITH 6 CLASSES. *Diff* INDICATES PERFORMANCE DIFFERENCES COMPARED TO THE PROPOSED PIPELINE.

Method	Task	Acc.	Diff.	F1	Diff.
proposed	HAR	87.1%	0%	86.7%	0%
baseline-2	HAR	74.9%	-12.2%	73.6%	-13.1%
baseline-3	HAR	70.1%	-17.0%	68.9%	-17.8%
baseline-4	HAR	65.3%	-21.8%	64.5%	-22.2%
baseline-5	HAR	62.7%	-24.4%	61.9%	-24.8%
baseline-6	HAR	58.4%	-28.7%	58.2%	-28.5%
proposed	Ori. classification	97.5%	0%	97.3%	0%
baseline-2	Ori. classification	75.1%	-22.4%	73.9%	-23.4%
baseline-3	Ori. classification	53.4%	-44.1%	52.8%	-44.5%
baseline-4	Ori. classification	56.3%	-41.2%	56.4%	40.9%
baseline-5	Ori. classification	62.8%	-34.7%	61.5%	-34.8%
baseline-6	Ori. classification	62.1%	-35.4%	61.6%	-35.7%
proposed	PC classification	99.0%	0%	97.2%	0%
baseline-2	PC classification	86.0%	-13.0%	85.8%	-11.4%
baseline-3	PC classification	87.5%	-11.5%	87.4%	-9.8%
baseline-4	PC classification	83.5%	-15.5%	82.9%	-14.3%
baseline-5	PC classification	80.2%	-18.8%	79.9%	-17.3%
baseline-6	PC classification	81.8%	-17.2%	81.4%	-15.8%

Radar data measurement is generally more complex and time-consuming than using vision-based sensors, thus limiting the typical radar-based HAR dataset to be much smaller than the computer vision datasets. To be more specific, Yang's work [31] and Kim's work [23] included 60 and 288 samples per activity, respectively, whereas our dataset includes roughly 300 training samples per activity. Therefore, an important comparison is about the classification pipeline performance with respect to a limited number of training samples. Figure 5 presents the test accuracy of the baselines and the proposed pipeline trained with only a randomly selected percentage of the training data available from the measurements. The results in Figure 5 show that the performance of the proposed pipeline

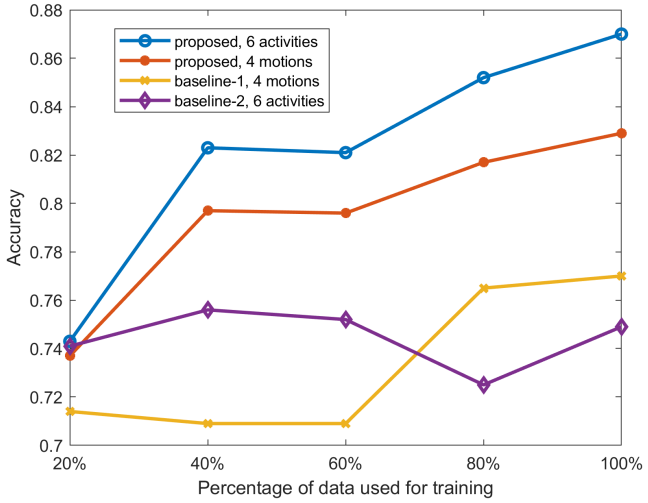


Fig. 5. Classification accuracy of the proposed pipeline and two selected baselines with respect to varying number of training samples.

shrinks sharply from 87.1% to 74.0% accuracy for 6 activities and from 83.0% to 73.5% for 4 motions, as the number of training samples decreases. Nevertheless, the test accuracy remains on average higher than or approximately equal to the state-of-the-art baselines in [23], [31] as the number of training data is reduced. up to the point of using only 20% of the training data (equivalent to 358 samples per activity).

To conclude, in terms of different quantitative metrics presented in this subsection, the proposed pipeline in general has significantly superior performance over the baselines in *1)* angle insensitivity, *2)* classification accuracy, and *3)* robustness against a limited number of training samples.

D. Results for Different SNR

This section focuses on one of the most influential parameters in radar systems, SNR. Because of the randomness nature of additive noise, data generation, training and testing are repeated for five independent realizations and averaged for every value of considered SNR, as described in Section III. Figure 6 shows the average test accuracy and the standard deviation of these five realizations in terms of varying SNR levels. As can be seen, the classification performance decrease almost linearly along with decreasing SNR levels, and the accuracy could drop to nearly 50% for an SNR of 8dB, which is approximately 35% lower than the measured data (assumed to be noise-free in this evaluation). These results suggest that noisy data could significantly undermine the performance of the proposed pipeline. Last and most importantly, the performance gain due to the proposed replacing of max pooling with average pooling is clearly shown by comparing the blue curve with the red in Figure 6. This appears to indicate that average pooling as a symmetric function fits better the task of processing noisy radar data compared to the original max pooling in [29].

For data-driven classification methods, it is also interesting to evaluate whether differences in SNR between training and test data could influence classification performance. In

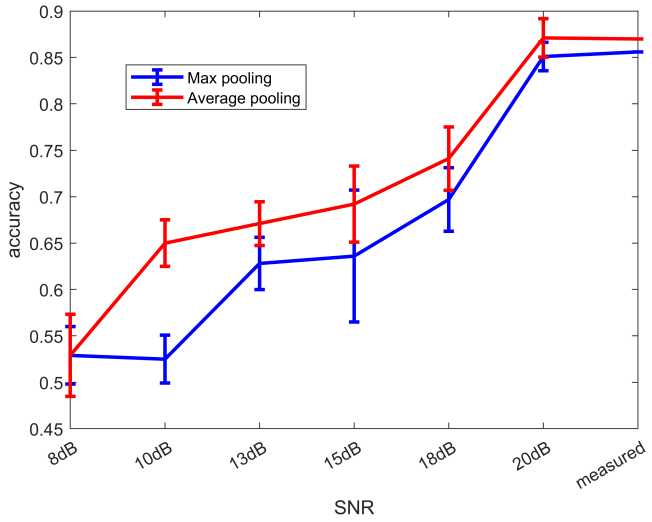


Fig. 6. Classification accuracy with respect to varying SNR levels. Average pooling is used in the proposed pipeline, whereas max pooling refers to the original PointNet and its transformation network T-Net in [29].

this case, we can cross-validate the results of 'training-with-measured-data' and 'testing-with-noisy-data', and vice versa. The results are listed in Table VIII. These results show that radar-based PCs and spectrograms in noisy conditions are unlikely to be sufficiently representative for classifying human body shapes and postures, hence additional work would be needed to make the pipeline more robust to severe degradation in SNR.

TABLE VIII
CROSS-EXAMINATION OF THE ROBUSTNESS OF THE TRAINED PIPELINE GIVEN TESTING DATA OF UNSEEN SNR LEVELS, WHERE *meas.* INDICATES THE ORIGINAL MEASURED DATASET THAT IS ASSUMED TO BE NOISE-FREE.

Train SNR	Test SNR	Acc.	Diff.	F1	Diff.
meas.	*meas.*	87.1%	0%	86.7%	0%
meas.	*20dB*	67.8%	-19.3%	67.4%	-19.3%
meas.	*18dB*	65.3%	-21.8%	64.6%	-22.1%
meas.	*15dB*	60.0%	-21.1%	58.7%	-28.0%
meas.	*13dB*	55.3%	-31.8%	52.8%	-33.9%
meas.	*10dB*	528%	-34.3%	49.8%	-36.9%
meas.	*8dB*	47.8%	-39.3%	44.4%	-42.3%
20dB	*meas.*	69.0%	-18.1%	68.7%	-18.0%
18dB	*meas.*	50.1%	-37.0%	50.4%	-36.3%
15dB	*meas.*	49.6%	-37.5%	50.8%	-35.9%
13dB	*meas.*	41.8%	-45.3%	43.2%	-43.5%
10dB	*meas.*	44.9%	-42.2%	45.1%	-43.6%
8dB	*meas.*	31.4%	-55.7%	30.8%	-55.9%

E. Results for Variations in MIMO Aperture

Using an imaging radar with a relatively smaller MIMO aperture can be economical and power-efficient, whereas this comes at the cost of reduced angular resolutions. The results obtained from four separate pairs of training and test datasets are given in Table IX, where the definitions of *subarray-1* to *4* are visualized in Figure 2. Moreover, we investigate the results when training the proposed pipeline with the data obtained from one of the virtual radar apertures, but testing with the data from a different one. Hence, the cross-examination results of training on the large-aperture dataset and testing on the small-aperture dataset and vice versa are also given in Table IX. The following findings are drawn from this analysis. First, since the largest decrease in performance is as small as 9.2% in accuracy and 9.1% in F1-score, it is reasonable to conclude that the proposed pipeline has promising robustness to be applied on different imaging radars, i.e. with different and often smaller apertures. Secondly, as the classification accuracy continuously drops from the first to the fifth row in Table IX, it is clearly established that the classification results are significantly influenced by the MIMO aperture size. Thirdly, the significance of the consistency between training and test data is highlighted according to the huge accuracy drop as shown in the cross-examination evaluations, suggesting that simply re-using the data collected with different radar may not be an optimal data augmentation method without some manipulation either on the data or on the classification pipeline. Finally, in the cross-examination evaluation, the combination of *array* and *subarray-2* always provides substantially better results. This could be the result of our signal processing chain, i.e., local peak value detection is utilized in the azimuth domain but not in the elevation domain so that PCs are generated by the virtual arrays with the same horizontal (i.e., azimuthal) MIMO aperture are similar.

TABLE IX
PERFORMANCE OF THE PROPOSED PIPELINE WITH DATA WITH VARYING MIMO APERTURE, I.E. DIFFERENT ANGULAR RESOLUTIONS. CROSS-EXAMINATION RESULTS OF TRAINING & TESTING WITH DIFFERENT APERTURES ARE ALSO SHOWN

Train array	Test array	Acc.	Diff.	F1	Diff.
array	*array*	87.1%	0%	86.7%	0%
subarray-1	*subarray-1*	83.6%	-3.5%	83.4%	-3.3%
subarray-2	*subarray-2*	83.3%	-3.8%	83.0%	-3.7%
subarray-3	*subarray-3*	79.4%	-7.7%	79.2%	-7.5%
subarray-4	*subarray-4*	77.8%	-9.3%	77.5%	-9.1%
array	*subarray-1*	26.4%	-60.7%	25.7%	-61.0%
array	*subarray-2*	58.7%	-29.4%	58.8%	-27.9%
array	*subarray-3*	27.4%	-50.7%	26.8%	-59.9%
array	*subarray-4*	31.6%	-55.5%	31.8%	-54.9%
subarray-1	*array*	28.1%	-59.0%	26.3%	-60.4%
subarray-2	*array*	72.8%	-14.3%	67.8%	-18.9%
subarray-3	*array*	26.8%	-60.3%	25.9%	-60.8%
subarray-4	*array*	29.0%	-58.1%	23.6%	-63.1%

F. Results of Leave-One-Subject-Out Test

It is expected that different human participants may have their own body size and height, and thus exhibit specific characteristics in their kinematic patterns when performing daily activities. In the interest of training a generally-applicable classification pipeline, it is significant to learn whether the body physical characteristics and their kinematic patterns can be directly linked by the proposed pipeline. In an ideal situation where human subjects share similar body physical characteristics, like in our dataset (see Table III), assuming that the kinematic patterns are very similar across individuals of similar body shape and size, these could be learnt by observing a small set of individuals. However, as it can be seen in

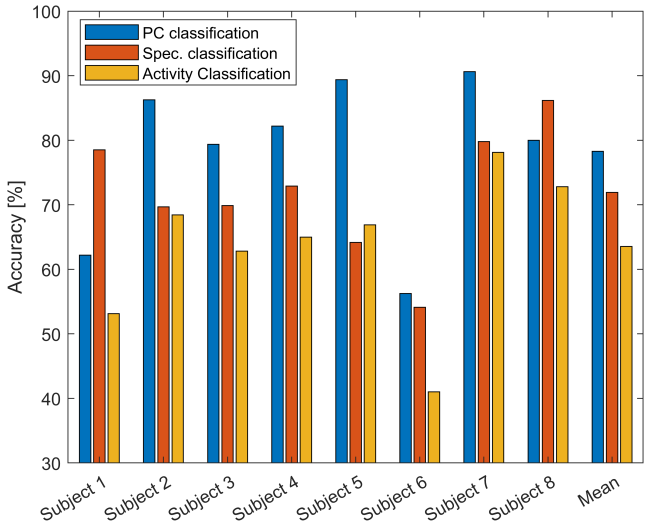


Fig. 7. Classification accuracy of the PC classification module (blue bars), the spectrogram classification module (red bars), and the entire pipeline (orange bars) in the leave-one-subject-out test, where the horizontal axis represents the index of the left-out subject.

Figure 7, in the case of classifying activities of an unseen human subject, the mean accuracy drops to 63.5%, which is approximately 20% lower than the results in Section IV-A. To be more specific, the mean PC classification accuracy and the mean spectrogram classification accuracy are approximately 80% and 71%, showing decreases of more than 17% and 12%, respectively.

Firstly, this finding fits the typical patterns in the leave-one-subject-out tests also observed in other studies with different data for HAR [19], [20]. Secondly, these values suggest that, despite very similar body characteristics, individuals still have distinct kinematic patterns due to their specific way of moving when performing the daily activities. These differences are what cause the performance of the proposed pipeline to suffer when a leave-one-subject-out test is applied. In addition, as shown in Figure 7, the lowest accuracy is attained from subject 6 (height 178 cm, weight 82 kg) instead of subject 5 whose body characteristics (height 185 cm, weight 95 kg) can be seen as an outlier. This again appears to show that the similarity of the kinematic patterns does not necessarily exist even for people who have very similar body characteristics of shape and size, suggesting the importance of having as much diversity as possible in the training data in order to have a more generally-applicable classification pipeline.

V. Conclusion

This work presents a pipeline for recognizing human motions and static postures performed toward multiple orientations using a 4D imaging radar. The proposed pipeline starts with the data generation module, including two parallel processing chains to generate PC and spectrogram as input data representations. Through the combination of PCs and spectrograms to represent the varying human body shapes and kinematic patterns, respectively, the information provided by 4D mm-wave imaging radar in terms of range, Doppler, azimuth, elevation, received power, and time are fully exploited. The proposed hierarchical classification pipeline attains an accuracy of 87.1%, significantly outperforming the state-of-the-start methods that only utilize either point cloud or spectrogram in isolation. It is also demonstrated that the proposed pipeline has substantial robustness against possible unfavorable conditions such as low SNR levels, a limited amount of training data, or relatively poor angular resolutions. The possibility of transferring trained models amongst different people is also demonstrated with a leave one subject out test, even if additional work is needed to increase performance in this case. The data generation module can be further improved by introducing automated segmentation methods such that each data sample contains no redundant information but just the desired activity on its own. Furthermore, architectures such as recurrent neural networks (e.g. GRU or LSTM cells) could be further integrated into the pipeline to model temporal connections between activities performed at different time steps one after the other. Moreover, the pipeline should be validated with additional data in terms of number of participants and types of activities. More participants, where possible with diverse body size and height, will help explore the generalization capabilities of the proposed pipeline to different physical characteristics. More activities, such as for example falling to the ground or other critical activities, will support the application of the proposed pipeline in a more realistic HAR setting.

Acknowledgment

The authors are grateful to all the volunteers who helped with the data collection.

References

[1] S. A. Shah and F. Fioranelli, "RF sensing technologies for assisted daily living in healthcare: A comprehensive review," *IEEE Aerospace and Electronic Systems Magazine*, vol. 34, no. 11, pp. 26–44, 2019.

[2] C. S. Florence, G. Bergen, A. Atherly, E. Burns, J. Stevens, and C. Drake, "Medical costs of fatal and nonfatal falls in older adults," *Journal of the American Geriatrics Society*, vol. 66, no. 4, pp. 693–698, 2018.

[3] N. Lu, Y. Wu, L. Feng, and J. Song, "Deep learning for fall detection: Three-dimensional CNN combined with LSTM on video kinematic data," *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp. 314–323, 2018.

[4] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *IEEE sensors journal*, vol. 15, no. 3, pp. 1321–1330, 2014.

[5] T. R. Bennett, J. Wu, N. Kehtarnavaz, and R. Jafari, "Inertial measurement unit-based wearable computers for assisted living applications: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 28–35, 2016.

[6] K. Chaccour, R. Darazi, A. H. El Hassani, and E. Andres, "From fall detection to fall prevention: A generic classification of fall-related systems," *IEEE Sensors Journal*, vol. 17, no. 3, pp. 812–822, 2016.

[7] C. Li, Z. Peng, T.-Y. Huang, T. Fan, F.-K. Wang, T.-S. Horng, J.-M. Munoz-Ferreras, R. Gomez-Garcia, L. Ran, and J. Lin, "A review on recent progress of portable short-range noncontact microwave radar systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 5, pp. 1692–1706, 2017.

[8] J. A. Nanzer, "A review of microwave wireless techniques for human presence detection and classification," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 5, pp. 1780–1794, 2017.

[9] F. Fioranelli, J. Le Kernec, and S. A. Shah, "Radar for health care: Recognizing human activities and monitoring vital signs," *IEEE Potentials*, vol. 38, no. 4, pp. 16–23, 2019.

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

[11] M. Jia, S. Li, J. Le Kernec, S. Yang, F. Fioranelli, and O. Romain, "Human activity classification with radar signal processing and machine learning," in *2020 International Conference on UK-China Emerging Technologies (UCET)*. IEEE, 2020, pp. 1–5.

[12] S. Zhu, J. Xu, H. Guo, Q. Liu, S. Wu, and H. Wang, "Indoor human activity recognition based on ambient radar with signal processing and machine learning," in *2018 IEEE international conference on communications (ICC)*. IEEE, 2018, pp. 1–6.

[13] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE transactions on geoscience and remote sensing*, vol. 47, no. 5, pp. 1328–1337, 2009.

[14] R. Zhang and S. Cao, "Real-time human motion behavior detection via CNN using mmwave radar," *IEEE Sensors Letters*, vol. 3, no. 2, pp. 1–4, 2018.

[15] B. Erol, S. Z. Gurbuz, and M. G. Amin, "Motion classification using kinematically sifted acgan-synthesized radar micro-Doppler signatures," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 4, pp. 3197–3213, 2020.

[16] L. Tang, Y. Jia, Y. Qian, S. Yi, and P. Yuan, "Human activity recognition based on mixed CNN with radar multi-spectrogram," *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25 950–25 962, 2021.

[17] J. Zhu, H. Chen, and W. Ye, "A hybrid CNN–LSTM network for the classification of human activities based on micro-Doppler radar," *IEEE Access*, vol. 8, pp. 24 713–24 720, 2020.

[18] H. Li, A. Shrestha, H. Heidari, J. Le Kernec, and F. Fioranelli, "Bi-LSTM network for multimodal continuous human activity recognition and fall detection," *IEEE Sensors Journal*, vol. 20, no. 3, pp. 1191–1201, 2019.

[19] R. G. Guendel, M. Unterhorst, E. Gambi, F. Fioranelli, and A. Yarovoy, "Continuous human activity recognition for arbitrary directions with distributed radars," in *2021 IEEE Radar Conference (RadarConf21)*. IEEE, 2021, pp. 1–6.

[20] Y. Zhao, R. Gundel, A. Yarovoy, and F. Fioranelli, "Distributed radar-based human activity recognition using vision transformer and CNNs," in *presented at the 18th European Radar Conference*, London, UK, 2022.

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

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

[23] Y. Kim, I. Alnujaim, and D. Oh, "Human activity classification based on point clouds measured by millimeter wave MIMO radar with deep recurrent neural networks," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13 522–13 529, 2021.

[24] D. Nickalls, J. Wu, and N. Dahnoun, "A real-time and high performance posture estimation system based on millimeter-wave radar," in *2021 10th Mediterranean Conference on Embedded Computing (MECO)*. IEEE, 2021, pp. 1–4.

[25] J. Le Kernec, F. Fioranelli, C. Ding, H. Zhao, L. Sun, H. Hong, J. Lorandel, and O. Romain, "Radar signal processing for sensing in assisted living: The challenges associated with real-time implementation of emerging algorithms," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 29–41, 2019.

[26] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 16–28, 2019.

[27] F. J. Abdu, Y. Zhang, and Z. Deng, "Activity classification based on feature fusion of FMCW radar human motion micro-Doppler signatures," *IEEE Sensors Journal*, vol. 22, no. 9, pp. 8648–8662, 2022.

[28] Q. An, S. Wang, A. Hoorfar, W. Zhang, H. Lv, S. Li, and J. Wang, "Range-max enhanced ultra-wideband micro-Doppler signatures of behind wall indoor human activities," 2020. [Online]. Available: https://arxiv.org/abs/2001.10449

[29] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3D classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.

[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[31] Y. Yang, C. Hou, Y. Lang, T. Sakamoto, Y. He, and W. Xiang, "Omnidirectional motion classification with monostatic radar system using micro-Doppler signatures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3574–3587, 2019.

[32] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[34] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[35] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

Yubin Zhao received the B.S. degree in electrical engineering from University of Electronic Science and Technology of China, Chengdu, China, in 2019 and the M.S. degree in electrical engineering from Delft University of Technology, Delft, The Netherlands, in 2022.

Alexander Yarovoy (F'15) graduated from the Kharkov State University, Ukraine, in 1984 with the Diploma with honor in radiophysics and electronics. He received the Candidate Phys. & Math. Sci. and Doctor Phys. & Math. Sci. degrees in radiophysics, in 1987 and 1994, respectively. In 1987, he joined the Department of Radiophysics at the Kharkov State University as a Researcher and became a Full Professor there, in 1997. From September 1994 through 1996 he was with Technical University of Ilmenau, Ilmenau, Germany as a Visiting Researcher. Since 1999, he is with the Delft University of Technology, Delft, the Netherlands. Since 2009, he leads there a Chair of Microwave Sensing, Systems and Signals. He has authored and coauthored more than 450 scientific or technical papers, six patents and fourteen book chapters. His main research interests are in high-resolution radar, microwave imaging and applied electromagnetics (in particular, UWB antennas). He is the recipient of the European Microwave Week Radar Award for the paper that Best advances the state-of-the-art in radar technology, in 2001 (together with L.P. Ligthart and P. van Genderen) and 2012 (together with T. Savelyev). In 2010, together with D. Caratelli Prof. Yarovoy got the Best Paper Award of the Applied Computational Electromagnetic Society (ACES). He served as an Associated Editor of the International Journal of Microwave and Wireless Technologies from 2011 till 2018 and as a Guest Editor of five special issues of the IEEE Transactions and other journals. He served as the General TPC chair of the 2020 European Microwave Week (EuMW'20), as the Chair and TPC chair of the 5th European Radar Conference (EuRAD'08), as well as the Secretary of the 1st European Radar Conference (EuRAD'04). He served also as the co-chair and TPC chair of the Xth International Conference on GPR (GPR2004). In the period 2008 to 2017, he served as Director of the European Microwave Association (EuMA).

Francesco Fioranelli (Senior Member, IEEE) received the Ph.D. degree with Durham University, Durham, UK, in 2014. He is currently a tenured Assistant Professor at TU Delft, The Netherlands, and was an Assistant Professor with the University of Glasgow (2016–2019), and a Research Associate at University College London (2014–2016). His research interests include the development of radar systems and automatic classification for human signatures analysis in healthcare and security, drones and UAVs detection and classification, automotive radar, wind farm, and sea clutter.