Delft University of Technology

A Survey on Radar-Based Continuous Human Activity Recognition

Ullmann, Ingrid; Guendel, Ronny; Kruse, Nicolas Christian; Fioranelli, Francesco; Yarovoy, Alexander

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

A Survey on Radar-Based Continuous Human Activity Recognition

INGRID ULLMANN [1] (Member, IEEE), RONNY G. GUENDEL[2] (Student Member, IEEE),
NICOLAS CHRISTIAN KRUSE [2] (Student Member, IEEE), FRANCESCO FIORANELLI [2] (Senior Member, IEEE),
AND ALEXANDER YAROVOY[2] (Fellow, IEEE)

(Regular Paper)

[1]Institute of Microwaves and Photonics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany
[2]Microwave Sensing, Signals and Systems Group, Delft University of Technology, 2600 AA Delft, The Netherlands

CORRESPONDING AUTHOR: Ingrid Ullmann (e-mail: ingrid.ullmann@fau.de).

ABSTRACT Radar-based human motion and activity recognition is currently a topic of great research interest, as the aging population increases and older individuals prefer an independent lifestyle. This technology has a wide range of applications, such as fall detection in assisted living, gesture recognition for human-machine interfaces, and many more. Numerous studies exist on various approaches for radar-based activity capture and classification. However, most of these employ rather artificial data, often obtained in laboratory environments, and typically collected under particular conditions. Specifically, most research so far has aimed at distinguishing a predefined set of single activities with a defined start, stop and duration. This paper aims at drawing the attention to a so far less researched issue, one that will be of vital importance for future real-world application of radar-based human activity recognition: continuous activity recognition, i.e. recognizing specific activities in a stream of several sequential activities with unknown duration and arbitrary transitions between different classes of activities. A review on the current state of the art in this relatively new topic is given, followed by a discussion on future research directions.

INDEX TERMS Radar applications, radar signal processing, continuous human activity recognition, activities of daily living.

I. INTRODUCTION

With an aging population, offering ambient assisted living capabilities has become a key societal challenge. Giving older or more vulnerable people the opportunity to live in their own homes for as long as possible and at the same time providing safety in self-determined living is highly desirable. One major risk in such context is falling and related consequences [1]. To detect falls quickly, without the need for another person to be around, appropriate remote sensor systems are required. Such systems must be able to detect falls, but also distinguish them from other human movements related to uncritical daily activities, as well as monitor the general activity pattern of individuals.

There are various sensors capable of capturing human motion [2], [3], [4], wearable sensors among them. However, such sensors might cause discomfort for a person. Furthermore, it must be ensured that they are worn permanently, which could be difficult during daily activities such as bathing or during sleep. Their correct usage and maintenance might also be a problem for cognitively impaired people.

Therefore, contactless sensing is highly desirable. Optical systems such as cameras and lidar enable remote, contactless sensing. However, any optical system requires an unobstructed line of sight to the subject and performance often depends on environmental light conditions. Furthermore, using cameras for surveillance will undoubtedly raise concerns about privacy.

Radar on the other hand does not have the above-named restrictions and has therefore become an interesting alternative for this purpose. As a sensor, radar works in a contactless

fashion but is not restricted to unobstructed line-of-sights, which is why so-called through-wall-radars have been used in search-and-rescue operations and non-line of sight propagation approaches have been demonstrated, for example for detecting around the corner [5], [6], [7]. Another essential advantage of radar is that it is, by nature, capable of measuring motion directly by exploiting the Doppler effect.

Therefore, radar as a sensor for human health monitoring has seen growing interest in recent years [8], [9]. Along with activity recognition and fall detection, typical applications include the monitoring of vital signs such as heartbeat and respiration amongst others [8], [10] or gait analysis [11], [12], [13].

Numerous research works have previously been conducted in the field of radar-based human activity recognition and fall detection [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]. However, there are still open challenges in this area. Most research so far has aimed at distinguishing a set of single activities under pre-defined conditions. In particular, this often includes classifying the individual activities separately, with data sets collected in a way that each recording or sample contains one activity with predefined duration and clear transitions between different activities. This of course is a valuable point to start from, but it is not yet very realistic. In real-world scenarios, human activities take place in a continuous stream, with one motion succeeding another, with variable duration of the single activities. Furthermore, diverse activities in terms of the extent of the body movement are mixed with periods of relatively static postures where only small body movements are present. Therefore, improving our capabilities for continuous human activity recognition will be of vital importance in order to make radar a real candidate for the task of human monitoring in home healthcare.

Furthermore, dealing with continuous data streams is equally relevant for other domains of human monitoring by means of radar. For example, it has been investigated in the context of continuous vital sign monitoring [29], [30], [31], gait dynamics analysis [32], [33], continuous tracking and identification of multiple people [34], and dynamic hand gesture recognition in the context of interaction with smart devices and/or interpretation of sign language [35].

The aim of this review paper is to draw attention to the relatively new research area of continuous human activity recognition, by summarizing the main approaches in the literature. Section II provides a survey of the most relevant techniques and literature in this area. Section III outlines some interesting future research directions to address open challenges, with final conclusions drawn in Section IV.

II. THE HISTORY OF CONTINUOUS HUMAN ACTIVITY RECOGNITION

In this section, a literature review on radar-based human activity recognition is provided. We describe the experimental setups, the radar signal processing strategies, and the classification algorithms or neural networks employed for the
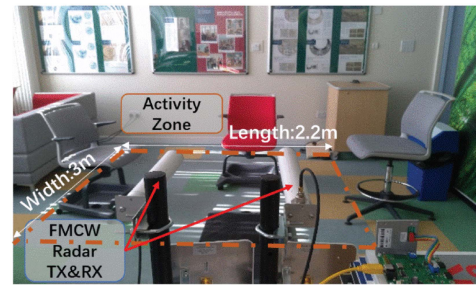
FIGURE 1. Measurement setup at Glasgow University where continuous sequences of activities were recorded with a C-band radar (from [41]).

classification tasks, taking into account the continuity of the signals. A summary of the main information in each reviewed study is provided in Table 1. It is assumed that the reader is familiar with the fundamentals of radar and machine learning, for which a good overview is provided for example in [36], [37], [38].

In 2018, Erol et al. [39] investigated fall detection within a sequence. The power burst curve (i.e., the summation of signal power in a given frequency band) of spectrograms is employed. If this indicator drops below a given threshold, followed by a silence period of five seconds, the algorithm crops the spectrogram segment of 1.5 seconds before the incident. This segment is fed to a pre-trained k-nearest-neighbor (kNN) classifier, in order to discriminate between fall and non-fall. However, no further activity classification was performed in this work yet.

First research efforts on continuous human activity recognition were reported in 2019 by Li et al. [40] at the University of Glasgow. The authors employed a frequency-modulated continuous-wave (FMCW) radar (Ancortek 580-B), operating at 5.8 GHz, with a bandwidth of 400 MHz. The radar was placed 1 m above the ground, facing the experimental scene, which contained typical room furniture. The setup is shown in Fig. 1. 16 participants performed 6 daily living activities. These are: walking, sitting on a chair, standing up from that chair, picking up a pen, drinking water from a glass, and falling. The participants performed the six activities in three predefined sequences. The radar data of these sequences were collected in two fashions, namely in snapshot mode (i.e., isolated recording of each activity) and in continuous mode. In continuous mode, each stream lasted 35 seconds, but the duration of the single activities was unconstrained. Radar signal processing consisted of first applying a notch filter to remove static clutter. Subsequently, a short-time Fourier transform (STFT) with a window size of 0.3 seconds and 95% overlap was performed to obtain the micro-Doppler signatures. For the classification task, 20 features were extracted from the spectrograms and its singular value decomposition (SVD). These included the Doppler centroid and bandwidth, mean and standard deviation and the left and right singular vectors of the singular value decomposition. To tackle the issue of continuity in the data stream, the data were partitioned into windows. Sliding windows with various sizes (3-4-5 s) and

TABLE 1. Summary Table With the Main Information of the Literature Reviewed in This Paper. *Type* Refers to the Classification Approach and Is Discussed at the End of Section **II**. *RT*-Range Time, *RA*-Range Angle, *RD*-Range Doppler, *DT*-Spectrogram(Doppler Time), *PC*-Point Cloud. *5 Activities With 8 Transitions. **6 Activities and 4 ASL Signs. ***3 Activities and 15 ASL Signs. †(bi)GRU,(bi)LSTM. ††(bi)LSTM,(bi)GRU

Year	Authors	Domain	Classifier	Type	Class #	Mode	Centre [GHz]	Bandwidth [GHz]
2018	Erol et al	DT	kNN	Snapshot	fall only	CW	25	N/a
2019	Li et al.	DT	SVM	Time Step	6	FMCW	5.8	0.4
2019	Ding et al.	RD	kNN	Snapshot	6	FMCW	5.8	0.32
2020	Shrestha et al.	DT, RT	(bi-)LSTM	Time Step	6	FMCW	5.8	0.4
2020	Li et al.	DT	SVM, biLSTM	Time Step	6	FMCW	5.8	0.4
2020	Amin et al.	RT, DT	kNN	Snapshot	5	FMCW	25	2
2020	Guendel et al.	RT, DT	Decision Tree	Snapshot	10	PD	4.3	2.2
2020	Vaishnav et al.	DT	LSTM	Time Step	4	FMCW	60	1
2021	Kang et al.	DT	CNN	Snapshot	13*	FMCW	60	6
2021	Kurtoglu et al.	RD, DT	CNN, biLSTM	Snapshot	10**	FMCW	77	4
2021	Guendel et al.	RD	Softmax	Time Step	9	PD	4.3	2.2
2022	Guendel et al.	DT	RNN†	Time Step	5	PD	4.3	2.2
2022	Yang et al.	(Complex) RT, RD, DT	(Complex) CNN	Time Step	9	PD	4.3	2.2
2022	Svenningsson et al.	RD	Bayesian Estimator	Time Step	9	PD	4.3	2.2
2022	Zhu et al.	DT	CNN, RNN††	Time Step	9	PD	4.3	2.2
2022	Kurtoglu et al.	RD, DT, RA	JD-MIMTL	Snapshot	18***	FMCW	77	4
2022	Werthen-Brabants et al.	DT	biRNN	Time Step	9	FMCW	60, 77	1.536
2022	Yu et al.	PC	CNN	Time Step	7	FMCW	62	4

overlaps (30-50-70%) were investigated. For classification, a quadratic kernel support vector machine was used. To emulate realistic testing conditions, the authors followed a "leave-one-person-out" strategy. This means that the data generated by a single participant used for testing was not included in its respective training data. Results gave a maximum accuracy of 84.7% for the 4 s window and 70% window overlap. Remarkably, the overall classification accuracy for this configuration was higher than for the single-snapshot data (80.56%). To improve the classification accuracy further, sequential forward selection (SFS) was investigated for the previous optimal configuration (4 s window and 70% overlap). With this technique, a more compact feature set was obtained, yielding a 2.6% improvement in classification accuracy.

A different approach to the problem was presented by Ding et al. from Nanjing University in 2019 [42]. In this study, an FMCW radar operating at 5.8 GHz with 320 MHz bandwidth was used, which had been designed at the authors' institute. The radar was placed at 1 metre height, in front of the subject, at various distances from 2 to 4 meters. 8 volunteers performed 6 activities. The focus again was on fall detection, and falling as well as 5 similar activities (stepping, jumping, squatting, walking and jogging) were investigated. Sequences of 2 activities were performed one after the other. The data processing in this work was based on range-Doppler maps, thereby using a sliding time window of 0.2 s, to obtain so-called range-Doppler frames. The key idea of the work was then to group a number of consecutive range-Doppler frames into a so-called 'dynamic range-Doppler trajectory'. This trajectory describes the pattern of a specific motion in range, Doppler and radar cross section (RCS) over time. The authors report an optimum number of 6 frames to form an individual trajectory. To obtain the trajectory, a number of points
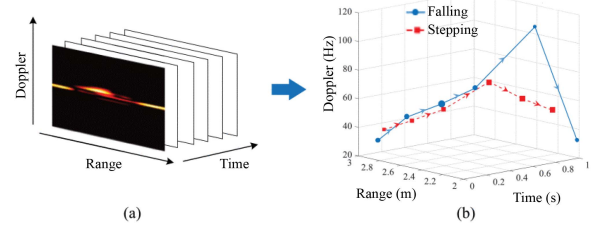
FIGURE 2. Illustration of (a) conventional range-Doppler frame trajectories to process continuous sequences of activities with (b) shown the DRDT of *falling* and *stepping* proposed in [42].

containing most of the image energy (i.e., intensity/RCS) are selected as "points of interest". The weighted average of these points is computed to form the dynamic range-Doppler trajectory map. Each of the investigated activities has its own range-Doppler trajectory map. The process is illustrated in Fig. 2. The concept for recognizing single activities in time is based on the fact that all the activities in the study have a high Doppler component. Therefore, in a continuous stream, peaks along the Doppler dimension are searched for, as they are likely to correspond to some activity. Since the individual motions correspond to six frames in this study, six-point windows are selected around those peaks. These serve as input to the following feature extraction and classification stage. 28 features of 4 comprehensive types were extracted based on the dynamic range-Doppler trajectory maps. The 4 types were dynamic Doppler frequency (Doppler over time), dynamic range change (range over time), dynamic energy change (intensity over time), and dynamic dispersion of range and Doppler (standard deviations). A subspace kNN classifier was used, with one third of the data for training. An average accuracy of 91.9% was obtained. Finally, a close-to-realistic scenario

was investigated with one subject: the volunteer performed a series of all 6 motions at random distances and view angles. It was demonstrated that all activities were classified correctly, albeit this was a relatively simple case.

In 2020, the University of Glasgow team presented more advanced classification techniques [43] and the investigation of a multimodal measurement setup [41]. Shrestha et al. [43] used the data set as in [40], but employed a Long-Short-Term-Memory (LSTM) network as a classifier. LSTM is a type of recurrent neural network (RNN) that interprets the radar data as a temporal sequence. Specifically, it can learn time-dependencies between separated time steps in a stream, which is why the technique is widely used in speech signal processing. Various types of LSTM architectures were investigated:

1) LSTM on spectrograms
2) Bidirectional-LSTM on spectrograms
3) Bidirectional-LSTM on range vs time plots, hence without explicit calculation of (micro-)Doppler signatures

Whereas option 1) only takes into account time dependencies between the current time step and previous ones, the bi-directional LSTM is capable of relating one time-step with previous and future time-steps. It was found that option 2) performed best (mean accuracy of 91% vs 78% for LSTM and 76% for range-bi-LSTM). The performance was also compared to that of a classical support vector machine (SVM). For the SVM, features were extracted from centroid, bandwidth and singular value decomposition of the spectrogram. However, its mean accuracy (66%) was found to be significantly lower than that of the LSTM. Further, it was investigated if prior knowledge about the subject improves performance, but it was found not to. Another finding regarding this setup was that aspect angles up to 30 degrees with respect to the radar line of sight provided acceptable performance results, but the trajectories of the motion remained rather simplistic and limited to a constrained straight line.

In [41], the combination of radar and an inertial measurement unit (IMU) worn on the participant's wrist was investigated. Again, the research was based on the radar and signal processing of [40]. Feature extraction was based on the spectrogram and its SVD, with feature selection performed via sequential backward selection in conjunction with an SVM classifier tailored to the multimodal setup. Various possibilities for the classification of continuous activities were investigated:

1) A sliding window approach with various window sizes and overlapping factors. Here, radar-only reached a maximum of 83.82% accuracy for 4 s window and 90% overlap. Combining radar and IMU yielded an improvement of +6% for the same radar signal processing choices.
2) A bi-directional LSTM, which was found to perform better than the sliding window approach. Again, the "leave-one-person-out" approach was chosen. An accuracy of 88.9% was obtained for radar-only data with this network.
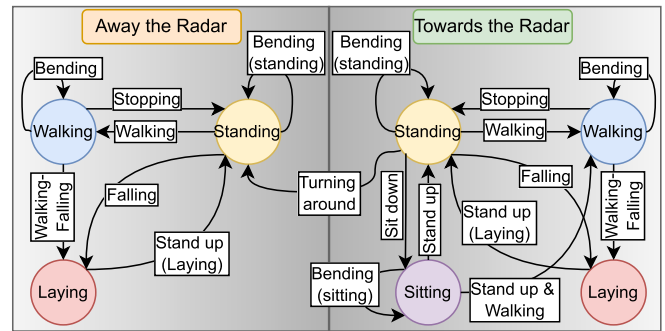
FIGURE 3. A human Ethogram for two-way continuous sequence recognition shows the order in which activities can occur and how they lead to the states of *Standing*, *Walking*, *Sitting*, or *Laying*. Other pathways are mutually exclusive, i.e, a *Laying* state cannot lead into a fall action, or a person cannot sit down twice consequently to reach *Sitting* state, but it can bend and ends in the *Standing* state every time [44].

The paper also investigated various fusion methods for radar and IMU data, which further improved performances but are outside the scope of this review.

Another approach for the problem of continuous activity recognition was introduced by Amin et al. at Villanova University in 2020 [44], [45]. In this work, human activities are regarded as states connected by other activities. States are for instance walking, standing, sitting and lying. A change in state is performed through an activity, e.g., bending, falling, standing up. The idea is based on a so-called 'ethogram', which is a catalogue of possible human motion sequences (see Fig. 3). The ethogram is the basis for the classification since it limits the number of activities which can happen after a certain activity, e.g. "walking cannot be preceded by falling but can be followed by it" [44]. For this reason, the authors also distinguish between forward-time motion sequences and reverse-time motion sequences. The employed radar was the Ancortek SDRKIT 2500B, which is an FMCW radar operating at 25 GHz with a bandwidth of 2 GHz. From the captured data, the range map (range vs time) was computed, as well as the micro-Doppler signature in spectrograms. Separating activities in the continuous stream was performed by employing the Radon transform on the range map for discriminating in-place motions and translation motions. For in-place motions, the power burst curve technique was applied to the micro-Doppler signature, to determine whether there is one or a sequence of in-place motions. A two-dimensional principal component analysis (PCA) was used for feature extraction from the range map and micro-Doppler signature, respectively. The $d = 14$ largest eigenvalues and corresponding eigenvectors were selected from the micro-Doppler signatures, and $d = 4$ largest eigenvalues and corresponding eigenvectors from the range maps. The images were projected onto the d-dimensional subspace to compute the respective principal component matrix. A kNN classifier was used, operating on the fused vectorized

and concatenated micro-Doppler and range-map features. Using the ethogram, the number of possible activities to be classified varied with time, since each state has possible prior and posterior activities. It was shown that this approach yielded a better classification performance than considering all states at all times, including those that are not possible from the ethogram.

Similar to this approach, Guendel et al. from TU Delft proposed in 2020 another technique to separate a stream of activities [46], including the discrimination of translation movement and in-place movement. Different to [44], the discrimination was performed by means of the so-called 'derivative target line' instead of the Radon transform. The derivative target line is the time derivative of the noise-removed range-over-time profile, which corresponds to the target's velocity. By applying a threshold to the derivative target line, the authors distinguish between toward-radar movement, away-from-radar movement, and in-place activity (no movement in range). A change in these movements indicates a change between in-place and translational activities. As in [44], in-place activities are further separated by means of the power-burst-curve applied on the micro-Doppler signature. This information allows for cropping both range map and micro-Doppler into the separated activities. For the experimental verification, a radar of type P410 Humanics was used, which is a pulsed radar with a center frequency of 4.3 GHz, a bandwidth of 2.2 GHz, and a pulse repetition interval of 8.2 ms. Six of these radars, aligned linearly and operated simultaneously, were used to collect data. 20 activity classes were investigated in total, in a multi-activity sequence including a fall. Similar to [44], again, a flow-graph of possible consecutive activities is introduced. This limits the output classes for the classifier in the sense that after one activity is classified, there is only a limited set of possible (i.e., physically sensible) follow-up activity classes. Feature extraction and classification were performed by 2D PCA followed by a decision tree classifier. Classification based on range maps and micro-Doppler individually as well as fusion by concatenating was investigated. It was found that fusion performed slightly better than the individual evaluations.

A combined localization and activity classification technique based on Kalman filtering was introduced by Vaishnav et al. from Infineon Technologies in 2020 [47]. With this, not only position and velocity of a target were incorporated into the state vector of the Kalman filter processing, but also the probabilities of the investigated activity classes. The state transition function between the activity classes was modeled by means of transition weights, making the process non-linear. An unscented Kalman filter (UKF) was therefore used. For the UKF state prediction transform, which was approximated by an unscented transform with sigma points, the aforementioned transition weights were learned by an LSTM. The input to the LSTM were therefore sequences of transitions between different activities in spectrograms. The output of the Kalman filter was not only the target's location, but also its
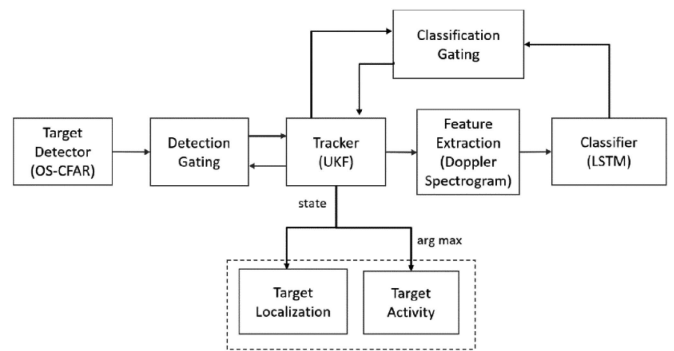
FIGURE 4. Illustration of the combined localization and classification approach proposed in [47].

current activity. In the implementation, the tracker was updated every 16 frames, based on state transition probabilities and current LSTM classification probabilities with adaptive Kalman gain. The process is illustrated in Fig. 4. Infineon's BGT60TR13 C radar chipset was used, which is an FMCW radar at 60 GHz center frequency, with 1 GHz bandwidth. 4 activities (walking, standing, sitting, waving) were classified, obtained from 5 participants with 800 Doppler spectrograms overall. 95.5% classification accuracy were obtained with the proposed approach, which was shown to be higher than using only an LSTM (87.24% accuracy). Furthermore, it was stated that since Kalman filtering also provides uncertainties, this information can be used for reducing false-alarm rates.

In 2021 Kang et al. [48] proposed a two-step procedure for classifying sequences of activities. In the first step, the stream was segmented into single actions, and these were classified in the second step. For the segmentation, the variance of the spectrogram was evaluated. For the subsequent classification, a convolutional neural network (CNN) was used with spectrogram data as input. Two types of outputs were evaluated for the classification of sequential activities: the first simply discriminated between single motion and motion change, whereas the second also classified the type of change (e.g., standing-to-walking). The radar employed in the experimental setup was a 60 GHz millimeter-wave FMCW radar with a bandwidth of 6 GHz, mounted 60 cm above the floor. Experimental results revealed an accuracy of > 97% for the segmentation between activities, and >95% on average for the classification of activities and activity transitions.

An approach to classify continuous activities mixed with signs from the American Sign Language (ASL) was presented by Kurtoglu et al. from the University of Alabama Tuscaloosa in 2021 [49]. The TI AWR1642BOOST radar was used, which is an FMCW radar at 77 GHz with 4 GHz of bandwidth. Subject were located at 1.5 m directly in front of radar for performing the ASL signs; daily activities were recorded at varying distances within 4 m radius. A total of four subjects took part in the study, performing six different daily activities (no movement, walking, sitting, standing up, folding laundry,

ironing) and 4 ASL signs (you, hello, car, push). Three sequences were recorded continuously, yielding 196 samples in total. 80% of the data were used for training, the remaining 20% for testing. The classification was based on three input representations, namely range-Doppler maps, micro-Doppler (μD) spectrograms and spectrogram envelopes, all divided into 0.2 s windows. A two-stage processing strategy was proposed, consisting of motion detection/segmentation plus segment recognition, so that classification was only performed when some motion was detected. Motion detection was performed using range-weighted energy plots with two detection methods: cell-averaging constant-false-alarm-rate (CA-CFAR) and short-term-average-over-long-term-average (STA/LTA). A 3D CNN followed by a Bi-LSTM layer and a time-distributed softmax was then applied to the range-Doppler maps, whereas the spectrograms and envelopes were processed by a 2D and 1D CNN, respectively, followed by a Bi-LSTM layer and softmax as well. While all three data representations were employed for classification individually, fusion was also investigated, namely decision level fusion and feature level fusion. Both achieved > 90% accuracy, with CFAR and STA/LTA. The best performance was achieved using feature fusion and STA/LTA (93.3% accuracy), but the other techniques performed in the same range. Remarkably, the performance of range-Doppler and spectrogram alone yielded a performance similar to the fused data (around 91%). Using the envelopes-alone classification had a slightly lower accuracy (<90%).

In the aforementioned studies, research on continuous activity recognition was restricted to using a single radar. In 2021, Guendel et al. at TU Delft introduced a two-dimensionally distributed radar network for capturing continuous activities [50]. This on the one hand brings a number of benefits, in particular multiple perspectives to view the moving subjects, but at the same time introduces the additional challenge of fusing the network's data in an appropriate fashion. In this case, the radar network consisted of five synchronous, monostatic radars. The Humanics (PulsON) P410 UWB pulsed radars were employed, with a pulse repetition frequency of 122 Hz, aligned in a semicircular baseline, with 45° separation, 1 m above the ground. The experiments were performed in a circular space of 4.38 m diameter, as shown in Fig. 5.

5 subjects performed 7 training data sequences and one test data sequence with activities performed in different order compared to the training set. The signal processing involved first clutter cancellation by subtracting the mean range-Doppler matrix of the training data set from the data. Seven features were then extracted from the range-Doppler map and concatenated into a feature vector for each radar at each sample. For classification a softmax classifier was used. Regarding the fusion of the radar network's data, three approaches were investigated:

1) *Early fusion:* Feature samples from all radars were concatenated into one longer feature vector and classification was performed on that.
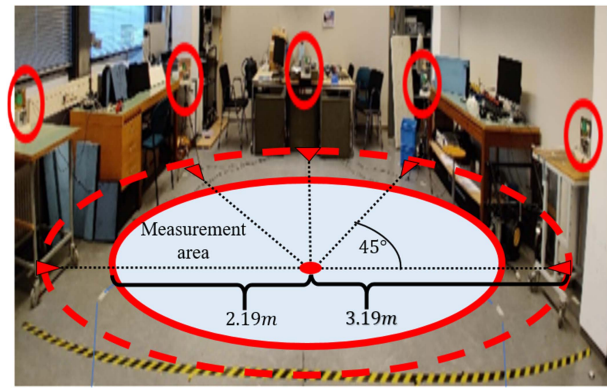
FIGURE 5. View of the TU Delft laboratory of the Microwave Sensing, Signals & Systems Group (MS3) with a radar network setup consisting of five nodes marked by the red circles (figure modified from [51]).

2) *Late fusion by mean:* Classification was performed for all 5 radars separately; afterwards the mean of all probabilities was computed.

3) *Late fusion by median:* Classification was performed for all 5 radars separately; afterwards the median of all probabilities was computed.

Time filtering was used to augment robustness of the classification, mitigating fluctuations over time. As in other work, the "leave-one-person-out" strategy was applied to test the proposed approach. Classification results indicated that all three fusion methods' accuracy values were in same range (about 50%), but feature fusion performed best. It was also demonstrated that the network performed better than any of the five radars alone (< 50% accuracy). In all cases, including the test person in the training set yielded slightly better results than "leave-one-person-out", as expected.

In 2022, Guendel et al. expanded and made their data set publicly available, as well as investigated additional fusion processing for the distributed radar network [52], [53]. The methods proposed in the paper are:

1) *Signal level fusion*, i.e., the simple summation of the range maps of all radar nodes. A spectrogram is then computed from the summation range map and used for classification.

2) *Feature level fusion*, i.e., computing μD spectrograms individually for all nodes and concatenating the data.

3) *Weighted radar selection over time:* In this approach, only one radar node's data is used as input for the classification. For each time step, one radar node out of the five is selected based on the most suitable aspect angle and received power. In order to determine which node is the most suitable, multilateration followed by a tracking filter is implemented to determine position, velocity and acceleration of the target.

4) *Orthogonal radar fusion:* Theoretically, an arbitrary movement in space can be fully captured by two radars with orthogonal line of sight. Two different setups were investigated. The first one was combining two orthogonal radars by means of feature fusion (i.e. processing

each one individually and concatenating the two spectrograms). The second one firstly combined the radar nodes 1 and 5 (which face each other as depicted in Fig. 5) by signal level fusion (i.e. summation), and then fused the result with the orthogonal node 3 (see Fig. 5). Again, fusing the orthogonal nodes was done by feature fusion. Both approaches yield a two-dimensional spectrogram, displaying the x and y components of the two-dimensional velocity vector in space.

The fusion approaches' classification performances were evaluated using a Bi-LSTM. Best results were achieved with the simple signal fusion approach. Orthogonal radar fusion using radars 1, 3 and 5 yielded similar results. All fusion approaches outperformed the use of just one single, fixed radar. As signal level fusion performed best with the Bi-LSTM classifier, this fusion method was further employed to evaluate other types of classification networks. Gated recurrent units (GRU) were tested both in mono- and bi-directional fashion as well as mono- and bi-directional LSTM. All classifiers performed well with an accuracy >90%. However, the authors also investigated other evaluation metrics such as intersection over union and Jaccard index. These might be more suited than accuracy to assess performances for continuous activities with imbalances in the classes, such as much more walking than instances of in-place activities or falls.

Since radar data are typically complex-valued, Yang et al. [54] investigated the use of complex-valued neural networks for classification of the TU Delft data set. Various network architectures were implemented, namely:
1) Multichannel networks operating on magnitude and phase;
2) Multichannel networks operating on real and imaginary part;
3) A deep network with complex-valued layers.

These three types of networks were applied to range-time maps, range-Doppler maps and spectrograms. The use of complex-valued data instead of real data yielded improvement only for some particular cases (e.g., the range-Doppler maps). In the overall performance, no significant improvement could be consistently observed.

Svenningsson et al. [55] proposed another processing approach for the TU Delft data set in 2022. In this paper, a Bayesian network is proposed, where in a recursive filtering algorithm the target's state (position, velocity, heading and turn rate) and motion class were jointly estimated. The state estimation served as a mapping of the points in the range-Doppler map to an estimate of the aspect angle for all sensor nodes. Including the so-found observation conditions was shown to augment classification accuracy. This is because the radars are capable of measuring only the radial velocity component and therefore the data depend highly on the aspect angle. A minimal resource management problem was solved which comprises a selection of sensor nodes to observe future micro-motions. Furthermore, probability calibration methods were introduced. 64.9% classification accuracy was obtained with the proposed classifier.

Another approach for classification of the TU Delft data set was published by Zhu et al. [51] in 2022, with five main steps:
1) Spectrogram computation
2) Spatial feature extraction by a CNN
3) Data fusion from the five radars into one feature map
4) Temporal feature extraction with a RNN
5) Final prediction by a fully-connected neural network

Regarding data fusion, in the paper, a halfway-fusion approach was proposed. It concatenates the feature maps from the CNN into a data cube, then uses a channel-wise maximum pooling to select the most representative features, and finally compresses these to a new feature map. It was demonstrated that compared to early and late fusion, halfway fusion performed best (ca. 87% accuracy, with early fusion having 85%, and late fusion ca. 84%, respectively). Furthermore, compared to single radars, fusion showed better performance. Single radars were only able to achieve ca. 70% accuracy. Regarding the recurrent neural networks, three different types were investigated. These were a simple RNN, an LSTM, and a GRU, each implemented in a mono- and bi-directional fashion. The bi-directional networks performed better in all cases. From the three types of networks, GRU performed best in the study (87.1%). As in other studies, the "leave-one-person-out" strategy was also employed here. Overall, the proposed network was able to achieve 90.7% test accuracy.

Further activities at the University of Alabama were reported by Kurtoglu et al. in 2022 [56]. While the main focus of this work lies on recognition of signs from the American Sign Language and corresponding trigger recognition, activity recognition is also part of the work. As in [49], the TI AWR1642BOOST radar was used. 19 participants performed 5 different sequences, consisting of 15 ASL signs and 3 activities (walking, sitting, standing up). Three representations of the radar data were used: range-Doppler map, µD spectrogram and range-angle map. The latter was achieved via beamforming. This was possible since the employed radar is MIMO (two transmitters and four receivers). In the range-angle maps, the visibility was enhanced using so-called optical flow (i.e., the spatial change in location of pixels from one frame to another). For segmentation of single activities, a variable-window STA/LTA was proposed, which is shown in Fig. 6. The variable length accounts for the variable duration of the single activities. This approach was shown to outperform STA/LTA with fixed-length windows and the so-called dynamic boundary detection technique. For classification, the authors used a joint domain multi-input-multitask learning including all three above named radar data representations. An accuracy of 92% was achieved with that method. It was compared to a CNN followed by Bi-LSTM, operating on the three data representations alone and on feature-level fusion. However, none of these classifiers exceeded 90%, which proved that the joint domain multi-input-multitask learning was more suitable.

For segmentation of single activities in sequences, Kruse et al. [57] proposed to use the Renyi entropy. The segmentation is performed by detecting rapid changes in the entropy of
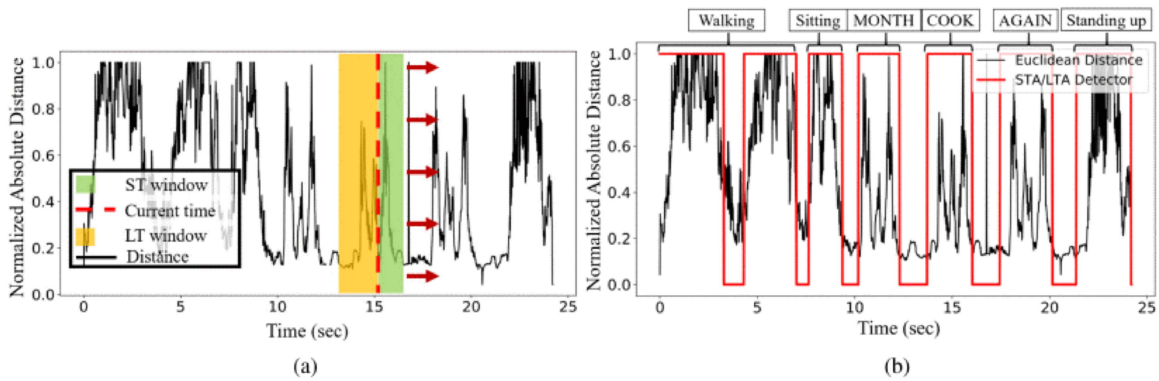
FIGURE 6. Illustration of stream segmentation using STA/LTA with variable window lengths [56].

spectrograms. A different threshold over a fixed time interval is introduced as a discriminator: whenever this threshold is exceeded, a transition between two activities is declared. The proposed method was applied to three different data sets. It was shown that the entropy-based segmentation can outperform the STA/LTA method as described in [49], [56].

To tackle the issue of real-time fall-alerts in hospital environments, Werthen-Brabants et al. [58], [59] proposed the use of a split Bi-RNN. A two-stage classifier was implemented: first, a forward RNN which is computed on an edge device gives an immediate prediction for every time step. Subsequently, a backward RNN is employed, computed on a larger processor (cloud or data center) to improve the prediction of the first step. Micro-Doppler signatures were used as the basis for classification. To reduce the amount of data and thus the computation time, 1D convolutional feature extraction was first applied to the micro-Doppler signatures. 16 features were extracted by a CNN and then used as inputs to the classifiers. Instead of using sliding windows, every single frame was evaluated and classified. Experiments were performed in two hospital-resembling facilities with two radar sensors placed in different locations: the TI xWR14xx radar, an FMCW radar operating at 77 GHz, and the TI xWR68xx radar at 60 GHz. An accuracy of 91% was achieved. Execution times of 1.664 ms for the forward branch and 36.645 ms for the backward branch were reported, which was faster than a standard bi-directional model (81.463 ms) implemented for comparison. A point cloud-based processing and classification methodology is proposed in 2022 by Yu et al. [27]. A point cloud representation of human motion is captured in 12 0.1 s long frames using a TI IWR6843ISK-ODS mmWave system at 62 GHz. The first processing step is to denoise the point clouds by means of the DBSCAN algorithm, which effectively labels points in low density regions as outliers and removes them from the sample. To facilitate the learning task, the dimensions of every classifier sample should be equal, which for mmWave point cloud representations is generally not the case. To circumvent this, the measurement area is divided in $50 \times 50 \times 30$ voxels (length $\times$ width $\times$ height)

and the value of every voxel at each frame is the amount of point cloud points that fall within its boundaries. Finally, sample diversity is enhanced by introducing sample duplicates, randomly rotated in the horizontal plane, to the dataset, with the justification for this method of dataset up-sampling being the rotational invariance of human activities in this plane. For classification, a novel 'Dual-View Convolutional Neural Network' is employed. The network features two CNN-channels operating in parallel on orthogonal projections of the voxelised input, namely the projections on the XZ and YZ planes. In three convolutional/maxPooling blocks the dimensionality of the two projections of each input sample are reduced to 12×576, representing the time and (flattened) spatial components respectively. The outputs of the two parallel blocks are then concatenated and, using three sequential fully connected layers, an activity label is finally computed. For experimental validation, four subjects are recorded individually performing seven activities at various locations in an indoor setting, resulting in a total of 1200 minutes of data before up-sampling. A comparison with several reference classifiers reveals superior accuracy of 97.61% on the recorded data set. On a publicly available dataset [60], an accuracy improvement of 6% (97.8%) is reported.

In terms of classification strategies, two major approaches can be identified in the works discussed in this review. The first, denoted '*Snapshot*' in Table 1, features a motion detection algorithm that extracts an interval of interest from the input sequence, which can subsequently be classified as a whole, often under the assumption that it comprises only a single activity. The second approach entails the classification of each time step in the input sequence, and is labelled correspondingly as '*Time Step*' in Table 1. The latter approach has become more prevalent with the emergence of RNN architectures capable of handling temporal data sequences, but has also successfully been implemented with a sliding window-based method [40]. Advantages of time step classification are the higher temporal resolution, and consequently the lower risk of not detecting an activity of short duration. Conversely, in a snapshot-style approach there exists the possibility of

acquiring an interval of interest that contains multiple activities, but which will be typically assigned only a single label, leading to missed detections. In snapshot style approaches the classifier can often idle in times of no detected activity, leading to potentially lower computational and power requirements; furthermore they benefit from the well established methods developed for single activity classification.

When examining the activity classes considered in the various studies in this review, three categories can be defined.

1) The algorithms that focus on the detection of a single activity such as a fall [39]. These anomaly-detection approaches can potentially aid in a variety of practical applications and have to waste no resources on the distinction among non-anomaly behaviours.

2) Works considering a finite set of activities for the classification task. This category constitutes the majority of research so far and assumes that most human behaviors can be broken up in a smaller subset of activities which can be identified by a suitable classifier. The average number of activity classes under consideration in Table 1 is 6.4±2.5 (not counting transitions, ASL signs, and duplicate activities at different aspect angles.)

3) Studies focusing on a finite set of activities, but placing restrictions on the possible transitions between them. This is accomplished e.g. by means of a human ethogram [44], [45] or Markov chain [55] and is based on basic assumptions about human kinematics.

Finally, a notable trend in the ensemble of works under review here is the strong emphasis on automated feature construction. With the exception of three studies [40], [42], [50], feature extraction is based on PCA/SVD, or a deep learning based alternative.

## III. OPEN ISSUES AND FUTURE RESEARCH DIRECTIONS

As shown by the papers surveyed in the previous section, the topic of continuous human activity recognition is relatively new and offers several open questions yet to be addressed:

- In realistic scenarios there could be situations where there is not only one, but *two or more people inside the field of view of the radar*. Therefore, designing robust classification strategies for superimposed signals of multiple humans is an important topic for realistic applications. There exist a number of strategies to deal with this issue in vital signs estimation for relatively static subjects [10]. Incorporating and possibly adapting such techniques to continuous full body movements is an open question that has not yet been addressed. For example, the merging of tracks of multiple people and the separation of their signatures must be addressed, as well as the temporal segmentation of their activities, which can well be not synchronized (i.e., one person might be performing one activity while the other just starts a new one).

  Even if one single individual is present, the question of how to best resolve multiple activities performed at the same time in a sequence is also important. For example,

in the case where the person is moving and also performing an activity (e.g., walking while carrying an object, speaking over the phone, eating or drinking).

- Separating the signature of multiple people or the contributions of different body parts of a single person can *benefit from additional spatial resolution*, especially in the angular domain. For this purpose, mm-wave multiple-input-multiple-output (MIMO) radars can be suitable, bearing in mind their expected improvement in capabilities in coming years driven, by progress in automotive research. Specifically, larger operational bandwidth of several GHz will be achieved by operating at higher segments of the frequency spectrum (hence finer range resolution), as well as the integration of more MIMO channels in simple radar chips or systems of cascaded chips (hence finer angular resolution in both azimuth and elevation). Using large MIMO radars (e.g. four combined individual radar chips with 3Tx & 4Rx channels each) for single activity classification was investigated in e.g. [17], [21], but not yet for continuous streams of activities. The formulation of the most suitable processing pipeline for such very high resolution data remains an open challenge.

- Even at lower operating frequencies, the question of identifying the *most suitable radar data representation* for continuous HAR remains. As a visual summary, Fig. 7(a) shows data domains extensively used throughout the literature applied on a single-input single-output (SISO) radar, namely: the range-time, range-Doppler, or the µD spectrogram. Less common but potentially interesting can be the phase information of the data, also shown in Fig. 7(a). For example, investigations were made by using the histogram of oriented gradients (HOG) on phase data, a method capable of extracting features from fine lines segments or shapes in images, and then forwarded for classification [61]. Fig. 7(b) provides an overview of additional data domains that can be obtained using multiple-input multiple-output (MIMO) radars, namely the range-angle map, one of the standard domains in the automotive radar sector, as well as an extracted point cloud for HAR [21]. Furthermore, the novel µD spectrogram computed in cross-range direction, known as µω spectrogram, introduced by Aziz et al. [62] is also shown. Whilst Table 1 reveals that the µD spectrogram is still the most prevalent data domain used in classification due to its inherent connection to the subjects' kinematics, it is possible that the introduction of sensors with higher spatial resolution will allow for a more effective extraction of information regarding the subjects' posture. This in turn may make spatial information currently not represented in µD spectrograms more relevant, and prove it to be advantageous in the improvement of classifier performances.

- Once radar-based activity recognition has found its way into real-world assisted living, interference problems might arise when there are several radars concurrently
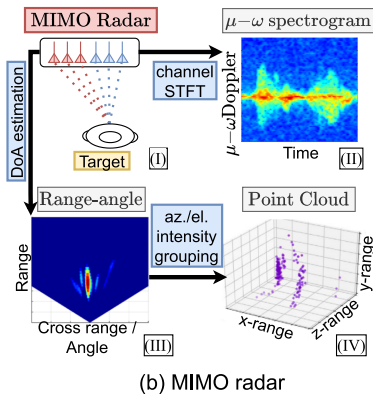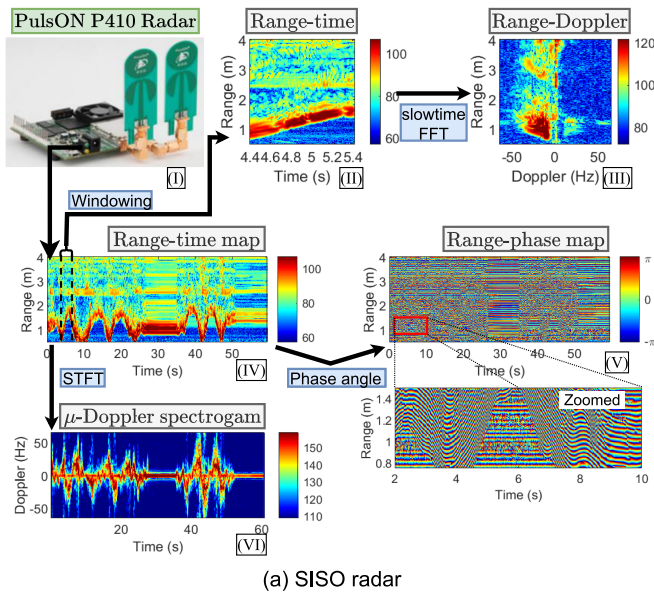
FIGURE 7. (a) common HAR data domains applied on a (I) PulsON P410 SISO radar node with, (II) the range-time obtained by windowing over (IV) the full range-time map, (III) the range-Doppler map by computing the FFT over slow-time of (II), (V) the extracted range-phase map with a magnified view, and (VI) the µD spectrogram obtained using the STFT.
(b) illustrates additional domains that can be obtained using (I) a MIMO radar, such as (II) the µω spectrogram (Images from [62]), (III) the range-angle map, and (IV) the point cloud representation.

present, e.g. in nursing homes. To overcome this issue, robust interference mitigation strategies [63] will have to be developed. In this context, it could be interesting to exploit the variety of already existing HAR sensors. For example, Yang et al. [64] developed a technology-agnostic activity recognition system, which is able to recognize activities using data from three different sensor types: radar, WiFi, and RFID. The availability of multiple sensing technologies for the same purpose could help ease problems of mutual interference.

- An important issue is the *availability and the comparability of labelled data sets* for the evaluation of different approaches for radar-based human activity classification. Very often in the literature, the proposed classification approaches appear to be tailored to a particular setup and

predefined set of classes, with questionable capability to generalize to unseen environments, individuals, and variations in activities. Furthermore, to the best of our knowledge, the data sets typically used in literature in this domain are small, limiting the depth and capabilities of deep learning methods that can be reliably trained on such data. While such relatively small size is understandable given the additional complexity to collect and label radar data with respect to for example video data, the question on how to address this data scarcity problem remains. An opportunity for this is provided by the recent appearance of several data sets that are publicly shared [65], although not all of them contain sequences of truly continuous human activities. Nevertheless, they can promote the benchmarking of algorithms on the same data as well as the development of methods that can perform training and learning of algorithms across diverse data.

IV. CONCLUSION

This paper presented an overview on the state of the art in radar-based continuous human activity recognition. Whereas activity recognition from snapshot-like radar data has been a widely investigated research topic for some time, classifying continuous streams of human activities is a relatively new field. First investigations started no earlier than 2018. Yet, dealing with the continuous nature of realistic data will be essential in order to bring the technology to the market.

FMCW radars as well as pulse radars have successfully been employed for the task. Regarding signal processing, range maps, range-Doppler maps, range-angle maps and µD spectrograms can be used as input for classifiers, as well as numerical features extracted from them. It is shown in the literature that fusing these representations usually yields better results than processing an individual one. Various strategies for segmenting the continuous stream, as well as for feature extraction and classification have been proposed, as could be seen in the survey. For the classification task, it was found that taking time-dependencies into account, e.g. by using LSTM or GRU classifiers, improves classification accuracy. Especially a bi-directional implementation of these networks is highly beneficial. Using a radar network instead of just one radar can provide additional benefit.

Future research will need to address a number of tasks including multiple-subject classification, multi-activity classification as well as more elaborate radar setups for a higher spatial resolution.

REFERENCES

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

[2] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer, "Monitoring activities of daily living in smart homes: Understanding human behavior," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 81–94, Mar. 2016.

[3] 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 Process. Mag.*, vol. 33, no. 2, pp. 28–35, Mar. 2016.

[4] K. Chaccour, R. Darazi, A. H. El Hassani, and E. Andrès, "From fall detection to fall prevention: A generic classification of fall-related systems," *IEEE Sensors J.*, vol. 17, no. 3, pp. 812–822, Feb. 2017.

[5] G. Li, Y. Ge, Y. Wang, Q. Chen, and G. Wang, "Detection of human breathing in non-line-of-sight region by using mmWave FMCW radar," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 8006311.

[6] N. Scheiner et al., "Seeing around street corners: Non-line-of-sight detection and tracking in-the-wild using Doppler radar," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 2065–2074.

[7] S. Vishwakarma, A. Rafiq, and S. S. Ram, "Micro-Doppler signatures of dynamic humans from around the corner radar," in *Proc. IEEE Int. Radar Conf.*, 2020, pp. 169–174.

[8] C. Li, V. M. Lubecke, O. Boric-Lubecke, and J. Lin, "Sensing of life activities at the human-microwave frontier," *IEEE J. Microwaves*, vol. 1, no. 1, pp. 66–78, Jan. 2021.

[9] F. Fioranelli and J. L. Kernec, *Contactless Radar Sensing for Health Monitoring*, Hoboken, NJ, USA: Wiley, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119644316.ch2

[10] S. M. Islam, O. Boric-Lubecke, V. M. Lubecke, A.-K. Moadi, and A. E. Fathy, "Contactless radar-based sensors: Recent advances in vital-signs monitoring of multiple subjects," *IEEE Microw. Mag.*, vol. 23, no. 7, pp. 47–60, Jul. 2022.

[11] A.-K. Seifert, M. G. Amin, and A. M. Zoubir, "Toward unobtrusive in-home gait analysis based on radar micro-Doppler signatures," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 9, pp. 2629–2640, Sep. 2019.

[12] A.-K. Seifert, M. Grimmer, and A. M. Zoubir, "Doppler radar for the extraction of biomechanical parameters in gait analysis," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 2, pp. 547–558, Feb. 2021.

[13] S. Mancini, W. Troy, K. Hall, X. Wu, and H. Wang, "Radar technology as a mechanism for clinical gait analysis: A review," *J. Ann. Bioeng.*, vol. 1, pp. 151–158, 2021.

[14] P. Chen et al., "Multi-view real-time human motion recognition based on ensemble learning," *IEEE Sensors J.*, vol. 21, no. 18, pp. 20335–20347, Sep. 2021.

[15] F. Fioranelli and J. Le Kernec, "Radar sensing for human healthcare: Challenges and results," in *Proc. IEEE Sensors*, 2021, pp. 1–4.

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

[17] 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 J.*, vol. 21, no. 12, pp. 13522–13529, Jun. 2021.

[18] M. Wang, G. Cui, X. Yang, and L. Kong, "Human body and limb motion recognition via stacked gated recurrent units network," *IET Radar, Sonar Navigat.*, vol. 12, no. 9, pp. 1046–1051, 2018.

[19] M. Wang, Y. D. Zhang, and G. Cui, "Human motion recognition exploiting radar with stacked recurrent neural network," *Digit. Signal Process.*, vol. 87, pp. 125–131, 2019.

[20] Y. Yao et al., "Fall detection system using millimeter-wave radar based on neural network and information fusion," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21038–21050, Nov. 2022.

[21] Y. Zhao, A. Yarovoy, and F. Fioranelli, "Angle-insensitive human motion and posture recognition based on 4D imaging radar and deep learning classifiers," *IEEE Sensors J.*, vol. 22, no. 12, pp. 12173–12182, Jun. 2022.

[22] W. Li, R. J. Piechocki, K. Woodbridge, C. Tang, and K. Chetty, "Passive WiFi radar for human sensing using a stand-alone access point," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 1986–1998, Mar. 2021.

[23] S. Vishwakarma, W. Li, C. Tang, K. Woodbridge, R. Adve, and K. Chetty, "Simhumalator: An open-source end-to-end radar simulator for human activity recognition," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 37, no. 3, pp. 6–22, Mar. 2022.

[24] Q. Chen, Y. Liu, B. Tan, K. Woodbridge, and K. Chetty, "Respiration and activity detection based on passive radio sensing in home environments," *IEEE Access*, vol. 8, pp. 12426–12437, 2020.

[25] G. Bhavanasi, L. Werthen-Brabants, T. Dhaene, and I. Couckuyt, "Open-set patient activity recognition with radar sensors and deep learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, 2023, Art. no. 3501005.

[26] Y. Yao, W. Liu, G. Zhang, and W. Hu, "Radar-based human activity recognition using hyperdimensional computing," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 3, pp. 1605–1619, Mar. 2022.

[27] C. Yu, Z. Xu, K. Yan, Y.-R. Chien, S.-H. Fang, and H.-C. Wu, "Noninvasive human activity recognition using millimeter-wave radar," *IEEE Syst. J.*, vol. 16, no. 2, pp. 3036–3047, Jun. 2022.

[28] D. She, X. Lou, and W. Ye, "RadarSpecAugment: A simple data augmentation method for radar-based human activity recognition," *IEEE Sensors Lett.*, vol. 5, no. 4, Apr. 2021, Art. no. 7001004.

[29] S. Schellenberger, K. Shi, F. Michler, F. Lurz, R. Weigel, and A. Koelpin, "Continuous in-bed monitoring of vital signs using a multi radar setup for freely moving patients," *Sensors*, vol. 20, no. 20, 2020, Art. no. 5827.

[30] K. Shi et al., "A contactless system for continuous vital sign monitoring in palliative and intensive care," in *Proc. IEEE Annu. Int. Syst. Conf.*, 2018, pp. 1–8.

[31] T. Koda, T. Sakamoto, S. Okumura, and H. Taki, "Noncontact respiratory measurement for multiple people at arbitrary locations using array radar and respiratory-space clustering," *IEEE Access*, vol. 9, pp. 106895–106906, 2021.

[32] H. Li, A. Mehul, J. Le Kernec, S. Z. Gurbuz, and F. Fioranelli, "Sequential human gait classification with distributed radar sensor fusion," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7590–7603, Mar. 2021.

[33] M. M. Rahman, D. Martelli, and S. Z. Gurbuz, "Gait variability analysis with multi-channel FMCW radar for fall risk assessment," in *Proc. IEEE 12th Sensor Array Multichannel Signal Process. Workshop*, 2022, pp. 345–349.

[34] J. Pegoraro, F. Meneghello, and M. Rossi, "Multiperson continuous tracking and identification from MM-wave micro-Doppler signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 2994–3009, Apr. 2021.

[35] Z. Zhang, Z. Tian, and M. Zhou, "Latern: Dynamic continuous hand gesture recognition using FMCW radar sensor," *IEEE Sensors J.*, vol. 18, no. 8, pp. 3278–3289, Apr. 2018.

[36] S. Z. Gurbuz, *Deep Neural Network Design for Radar Applications*. Stevenage, U.K.: Inst. Eng. Technol., 2020. [Online]. Available: https://digital-library.theiet.org/content/books/ra/sbra529e

[37] M. Amin, *Radar for Indoor Monitoring: Detection, Classification, and Assessment*. Milton, ON, Canada: CRC Press, 2017. [Online]. Available: https://www.routledge.com/Radar-for-Indoor-Monitoring-Detection-Classification-and-Assessment/Amin/p/book/9781138746091

[38] F. Fioranelli, H. Griffiths, M. Ritchie, and A. Balleri, *Micro-Doppler Radar and Its Applications*. Herts, U.K.: IET, 2020. [Online]. Available: https://digital-library.theiet.org/content/books/ra/sbra531e

[39] B. Erol, M. Francisco, A. Ravisankar, and M. Amin, "Realization of radar-based fall detection using spectrograms," in *Compressive Sensing VII: From Diverse Modalities to Big Data Analytics*, vol. 10658, F. Ahmad, Ed., Bellingham, WA, USA: SPIE, 2018, Art. no. 106580B.

[40] H. Li, A. Shrestha, H. Heidari, J. L. Kernec, and F. Fioranelli, "Activities recognition and fall detection in continuous data streams using radar sensor," in *Proc. IEEE MTT-S Int. Microw. Biomed. Conf.*, 2019, vol. 1, pp. 1–4.

[41] 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 J.*, vol. 20, no. 3, pp. 1191–1201, Feb. 2020.

[42] C. Ding et al., "Continuous human motion recognition with a dynamic range-Doppler trajectory method based on FMCW radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6821–6831, Sep. 2019.

[43] A. Shrestha, H. Li, J. Le Kernec, and F. Fioranelli, "Continuous human activity classification from FMCW radar with bi-LSTM networks," *IEEE Sensors J.*, vol. 20, no. 22, pp. 13607–13619, Nov. 2020.

[44] M. G. Amin and R. G. Guendel, "Radar human motion recognition using motion states and two-way classifications," in *Proc. IEEE Int. Radar Conf.*, 2020, pp. 1046–1051.

[45] M. G. Amin and R. G. Guendel, "Radar classifications of consecutive and contiguous human gross-motor activities," *IET Radar, Sonar Navigat.*, vol. 14, no. 9, pp. 1417–1429, 2020.

[46] R. G. Guendel, F. Fioranelli, and A. Yarovoy, "Derivative target line (DTL) for continuous human activity detection and recognition," in *Proc. IEEE Radar Conf.*, 2020, pp. 1–6.

[47] P. Vaishnav and A. Santra, "Continuous human activity classification with unscented Kalman filter tracking using FMCW radar," *IEEE Sensors Lett.*, vol. 4, no. 5, May 2020, Art. no. 7001704.

[48] S.-W. Kang, M.-H. Jang, and S. Lee, "Identification of human motion using radar sensor in an indoor environment," *Sensors*, vol. 21, no. 7, 2021, Art. no. 2305.

[49] E. Kurtoglu, A. C. Gurbuz, E. Malaia, D. Griffin, C. Crawford, and S. Z. Gurbuz, "Sequential classification of ASL signs in the context of daily living using RF sensing," in *Proc. IEEE Radar Conf.*, 2021, pp. 1–6.

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

[51] S. Zhu, R. G. Guendel, A. Yarovoy, and F. Fioranelli, "Continuous human activity recognition with distributed radar sensor networks and CNN–RNN architectures," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, 2022, Art. no. 5115215.

[52] R. G. Guendel, F. Fioranelli, and A. Yarovoy, "Distributed radar fusion and recurrent networks for classification of continuous human activities," *IET Radar, Sonar Navigat.*, vol. 16, no. 7, pp. 1144–1161, 2022.

[53] R. G. Guendel, M. Unterhorst, F. Fioranelli, and A. Yarovoy, "Dataset of continuous human activities performed in arbitrary directions collected with a distributed radar network of five nodes," Nov. 2021. [Online]. Available: https://data.4tu.nl/articles/dataset/Dataset_of_continuous_human_activities_performed_in_arbitrary_directions_collected_with_a_distributed_radar_network_of_five_nodes/16691500

[54] X. Yang, R. G. Guendel, A. Yarovoy, and F. Fioranelli, "Radar-based human activities classification with complex-valued neural networks," in *Proc. IEEE Radar Conf.*, 2022, pp. 1–6.

[55] P. Svenningsson, F. Fioranelli, A. Yarovoy, and A. F. Martone, "A Bayesian network for the classification of human motion as observed by distributed radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 6, pp. 5661–5674, Dec. 2022.

[56] E. Kurtoğlu, A. C. Gurbuz, E. A. Malaia, D. Griffin, C. Crawford, and S. Z. Gurbuz, "ASL trigger recognition in mixed activity/signing sequences for RF sensor-based user interfaces," *IEEE Trans. Human-Mach. Syst.*, vol. 52, no. 4, pp. 699–712, Aug. 2022.

[57] N. Kruse, R. G. Guendel, F. Fioranelli, and A. Yarovoy, "Segmentation of micro-Doppler signatures of human sequential activities using renyi entropy," in *Proc. IET Int. Conf. Radar Syst.*, 2022, pp. 1–6.

[58] L. Werthen-Brabants, G. Bhavanasi, I. Couckuyt, T. Dhaene, and D. Deschrijver, "Split BiRNN for real-time activity recognition using radar and deep learning," *Sci. Rep.*, vol. 12, no. 7436, 2022, Art. no. 7436.

[59] L. Werthen-Brabants, G. Bhavanasi, I. Couckuyt, T. Dhaene, and D. Deschrijver, "Quantifying uncertainty in real time with split BiRNN for radar human activity recognition," in *Proc. IEEE 19th Eur. Radar Conf.*, 2022, pp. 173–176.

[60] A. D. Singh, S. S. Sandha, L. Garcia, and M. Srivastava, "RadHAR: Human activity recognition from point clouds generated through a millimeter-wave radar," in *Proc. 3rd ACM Workshop Millimeter-Wave Netw. Sens. Syst.*, 2019, pp. 51–56. [Online]. Available: https://doi.org/10.1145/3349624.3356768

[61] R. G. Guendel, F. Fioranelli, and A. Yarovoy, "Phase-based classification for arm gesture and gross-motor activities using histogram of oriented gradients," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7918–7927, Mar. 2021.

[62] F. Aziz, O. Metwally, P. Weller, U. Schneider, and M. F. Huber, "A MIMO radar-based metric learning approach for activity recognition," in *Proc. IEEE Radar Conf.*, 2022, pp. 1–6.

[63] R. Amar, M. Alaee-Kerahroodi, P. Babu, and B. S. M. R., "Designing interference-immune Doppler-tolerant waveforms for radar systems," *IEEE Trans. Aerosp. Electron. Syst.*, early access, Oct. 25, 2022, doi: 10.1109/TAES.2022.3215116.

[64] C. Yang, X. Wang, and S. Mao, "TARF: Technology-agnostic RF sensing for human activity recognition," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 2, pp. 636–647, Feb. 2023. [Online]. Available: https://doi.org/10.1109/jbhi.2022.3175912

[65] F. Fioranelli, S. Zhu, and I. Roldan, "Benchmarking classification algorithms for radar-based human activity recognition," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 37, no. 12, pp. 37–40, Dec. 2022.

INGRID ULLMANN (Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree from Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany, in 2016 and 2021, respectively. She is currently a Postdoc and the Head of the research group "Wave-Based Imaging Systems" with the Institute of Microwaves and Photonics, FAU. In 2022, she spent one month as a visiting researcher with Microwave Sensing, Signals and Systems Group, Delft University of Technology, Delft, The Netherlands. Her research interests include radar imaging and radar signal processing for nondestructive testing, security screening, medical radar, and automotive applications. Dr. Ullmann is a reviewer for the European Radar Conference (EuRAD) and various journals in the field of microwaves and is a member of the IEEE CRFID Technical Committee on Motion Capture & Localization. Since 2022, she has been an Associate Editor for IEEE TRANSACTIONS ON RADAR SYSTEMS. She was the recipient of the Argus Science Award (sponsored by Airbus Defense and Space, now Hensoldt) in 2016 and EuRAD Conference Prize in 2019.

RONNY G. GUENDEL (Student Member, IEEE) received the Diploma-Engineering degree (Dipl.-Ing. (FH)) from the University of Applied Sciences Zwickau, Zwickau, Germany, and the Master of Science degree in signal processing and communications in 2019. He has been working toward the Ph.D. degree with the Delft University of Technology, Delft, The Netherlands, since January 2020. He worked with the Fraunhofer Institute for Machine Tools and Forming Technology (IWU). During 2017–2018, he joined Villanova University, Villanova, PA, USA, as a Fulbright scholar in electrical engineering. In 2018, he worked on vehicular wireless communications with Vodafone Chair Dresden, under guidance of Prof. Gerhard Fettweis. In 2018-2019, the Center for Advanced Communications, Villanova University, offered him a research assistantship under Dr. Moeness G. Amin, working on continuous activity classification. He works on monitoring continuous human activities in range and time beyond micro-Doppler by using RF-sensing technology.

NICOLAS CHRISTIAN KRUSE (Student Member, IEEE) received the B.Sc. degree in applied physics from the Delft University of Technology, Delft, The Netherlands, in 2017, and the M.Sc. degree in theoretical physics from the University of Groningen, Groningen, The Netherlands, in 2020. In March 2021, he joined the Microwave Sensing, Systems, and Signals Group, Delft University of Technology, where he is currently researching the classification of continuous human activities.

FRANCESCO FIORANELLI (Senior Member, IEEE) received the Laurea (B.Eng.) (*cum laude*) and Laurea Specialistica (M.Eng.) (*cum laude*) degrees in telecommunication engineering from the Università Politecnica delle Marche, Ancona, Italy, in 2007 and 2010, respectively, and the Ph.D. degree from Durham University, Durham, U.K., in 2014. He was an Assistant Professor with the University of Glasgow, Glasgow, U.K., from 2016 to 2019, and a Research Associate with University College London, London, U.K., from 2014 to 2016. He is currently an Associate Professor with the Delft University of Technology, Delft, The Netherlands. 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. He has authored more than 135 publications between book chapters and journal and conference papers and edited the book *Micro-Doppler Radar and Its Applications* published by IET-Scitech in 2020. He was the recipient of three best paper awards.

ALEXANDER YAROVOY (Fellow, IEEE) received the Diploma (Hons.) in radiophysics and electronics from Kharkov State University, Kharkiv, Ukraine, in 1984, and the 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, Kharkov State University, as a Researcher and became a Full Professor there in 1997. From September 1994 to 1996, he was a Visiting Researcher with the Technical University of Ilmenau, Ilmenau, Germany. Since 1999, he has been with the Delft University of Technology, Delft, The Netherlands, where he has been the Chair of microwave sensing, systems and signals since 2009. He has authored or coauthored more than 450 scientific or technical papers, six patents, and 14 book chapters. His main research interests include high-resolution radar, microwave imaging and applied electromagnetics (in particular, UWB antennas). Prof. Yarovoy was 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 in 2012 (together with T. Savelyev). In 2010, together with D. Caratelli, he received the Best Paper Award of the Applied Computational Electromagnetic Society (ACES). He was the Chair and TPC Chair of the 5th European Radar Conference (EuRA'08), Amsterdam, The Netherlands, and also the Secretary of the 1st European Radar Conference (EuRAD'04), Amsterdam, The Netherlands. He was also the Co-Chair and TPC Chair of the 10th International Conference on GPR (GPR2004) in Delft, The Netherlands. He was an Associated Editor for *International Journal of Microwave and Wireless Technologies* from 2011 to 2018, and a Guest Editor of five special issues of IEEE Transactions and other journals. During 2008–2017, he was the Director of the European Microwave Association (EuMA).