SimHumaLator: An Open Source Passive WiFi Radar Based Human Simulator For Activity Recognition

Shelly Vishwakarma*, Wenda Li*, Chong Tang*, Karl Woodbridge†, Raviraj Adve‡, Kevin Chetty*
*Department of Security and Crime Science, University College London, UK
†Department of Electronic and Electrical Engineering, University College London, UK
‡ Department of Electrical and Computer Engineering, University of Toronto, Canada
{s.vishwakarma, wenda.li,chong.tang.18, k.woodbridge, k.chetty}@ucl.ac.uk, rsadve@comm.utoronto.ca

***Abstract*—We present a simulation framework to generate human micro-Dopplers in passive WiFi radar scenarios, wherein we simulate the IEEE 802.11g standard complaint WiFi transmissions using MATLAB's WLAN toolbox and human animation models using a marker-based motion capture system. We integrate WiFi transmissions with the human animation data to generate the micro-Doppler features capturing the diversity of human motion characteristics, and the sensor parameters. We consider five human activities for the study. We uniformly benchmark the classification performances of multiple machine learning and deep learning models against a common dataset. Further, we validate the performances using the real radar data captured simultaneously with the motion capture system. Experimental results, simulations, and measurements demonstrate good classification accuracy of ≥ 95% and ≈ 90%, respectively.**

***Index Terms*—Passive WiFi Sensing, micro-Dopplers, activity recognition, deep learning, simulator**

I. Introduction

Humans are non-rigid bodies whose motion gives rise to frequency modulations, popularly known as micro-Dopplers. Over the last decade, the radar sensors have used these micro-Doppler signatures to detect and track human activities for various applications. These include applications that range from law enforcement, security, and surveillance purposes [1]–[5] to various ubiquitous sensing applications such as assisted living (fall detection in older people) [6]–[10], bio-medical applications for non-intrusively monitoring patients [11]–[14], and smart home applications such as occupancy detection [15], [16] and hand gesture recognition [17], [18]. Micro-Dopplers have been observed with active and passive radar sensors [19]–[23]. However, due to the sudden rise in transmitters of opportunity, passive sensing has attracted significant attention for indoor monitoring applications in recent years, [21], [24]–[27]. Passive sensing leverages the existing communication signals and infrastructure because of its receives only nature, thus leading to low power consumption and lighter construction.

The passive WiFi radar (PWR) data have been gathered through actual measurements in laboratory conditions using various radar hardware platforms [14], [25], [26], [28]. Measurement data is excellent for a thorough evaluation of various signal processing and machine learning algorithms' performances in realistic scenarios. The performances of these algorithms are generally tied to the large volumes of high-quality training data. Unlike the vision and image processing communities, we have a limited number of real radar databases. Therefore, there should be some means of simulating radar returns in passive WiFi sensing scenarios. The simulation data can be used for preliminary evaluation of different algorithms, studying effects of radar phenomenology, and would serve as an excellent means of generating large volumes of training data.

There exist multiple methods for simulating human micro-Doppler data in active radar scenarios. The earliest method modeled the human leg as a double pendulum structure [29]. However, this model does not simulate radar returns from other human body parts such as torso and arms, which also majorly contribute to the micro-Doppler returns. The second method developed a human walking model based on extensive biomechanical experiments [23], [30]. Here, twelve analytical expressions govern the motion trajectories of 17 reference points on the human body as a function of the human's height and the relative velocity. This model is a constant velocity model. Therefore, it cannot capture variations in more complex motions such as falling, sitting, jumping. The third technique uses animation data from motion capture systems to model more realistic and complex human motions. There are two types of motion capture technology available- marker-based and marker-less. Several markers are placed on the live actor's body parts such as head, torso, arms, and legs to capture their three-dimensional time-varying positions in space in a marker-based motion capture system. Authors in [31], [32], first developed a complete end-to-end active radar simulator of humans using a marker-based motion capture technique. The radar scatterings were simulated by integrating the animation data of humans with primitive based electromagnetic modeling. Alternatively, authors in [33], [34], gathered animation data using a cheaper marker-less motion capture technology based on Microsoft's Kinect. We adopt the same simulation methodology presented in [31], [34] but, for passive WiFi scenarios.

In this work, we present a simulation framework for generating human micro-Doppler signatures in passive WiFi scenarios. Our passive WiFi sensing exploits target reflections through cross-correlation based processing to determine range and Doppler information. We simulate the IEEE 802.11g standard WiFi transmissions using MATLAB's WLAN toolbox [35], and human animations using a marker-based motion capture system called Phase-Space [36]. The simulator generates the micro-Doppler radar returns as a function of- **target motion characteristics** (aspect angle, initial position in space, different motions type), **sensor parameters** (different PWR radar configurations-monostatic, bistatic-in line, and bistatic circular, waveforms) and **radar signal processing parameters** (such as coherent processing interval (CPI), and pulse repetition interval (PRI)). By varying these parameters and PWR radar operating conditions, we generate a vast simulation database. We uniformly benchmark the performances of different machine learning and deep learning classification algorithms against this common dataset. There is currently no open-source simulation tool for generating human micro-Doppler radar data to the best of our knowledge. Therefore, we publically release the simulator and hope it will be useful for benchmarking future algorithms and generating large volumes of high quality and diverse radar datasets. We also believe it will reduce the expense and labor involved in data acquisition by other researchers. The simulator is available for the interested users on https://uwsl.co.uk/.

To validate the performance of our simulator, we did simultaneous real measurements. We considered five human motion classes- human standing at a place and rotating body (HBR), human kicking (HK), human punching (HP), human grabbing an object (HG), and a human walking back and forth in front of the radar (HW). The micro-Doppler signatures corresponding to each of these activities share standard features because of the motion similarity. All these motions are periodic and thus have alternating positive and negative micro-Doppler features. Therefore, it becomes challenging for any classifier to discern the correct motion class. We evaluate different classical machine learning (handpicked features, cadence velocity features, sparse features) [18], [37]–[40], and deep learning-based classifiers (deep convolutional neural network, AlexNet, GoogLeNet, and ResNet18) [41]–[43], using both the simulation and measurement data for these five motion categories. We observe average classification accuracy ≥ 90 in almost all the deep learning frameworks.

To summarize, our contributions in this paper are the following:

1) Public release of a PWR human simulator that can simulate radar returns as a function of target motion characteristics, sensor parameters, and radar signal processing parameters. Simulator involves five human motion classes- human body rotating (HBR), human Kicking (HK), human Punching (HP), human grabbing an object (HG), and human Walking (HW).
2) Uniform benchmarking across different classifiers: traditional machine learning (handpicked features + support vector machines (SVM), automatic feature extraction + SVM), deep neural networks (both pre-trained and untrained)
3) Uniform benchmarking across different PWR geometries: monostatic with a single aspect angle, monostatic with aspect angles variation, bistatic with bistatic angle variation.
4) Performance validation using measured PWR data

Our paper is organized as follows. Section II describes our simulation framework, and simulation database generation. Next, Section III presents the classification results of different classification algorithms in different PWR scenarios. We describe our measurement data collection and validation of all the algorithms' performance in Section IV. We finally conclude our paper in SectionV.

II. A PWR SiHumLator

In this section, we first describe the simulation framework used for developing the PWR human simulator. We give a brief introduction of the simulator's capability in generating a diverse set of human micro-Doppler signatures for different- radar parameters, target parameters, and different radar signal processing parameters. Finally, we present the simulation dataset generated using the tool.

A. Simulation Framework

A typical PWR sensing setup is shown in Fig. 1. It comprises of two antennas- reference and surveillance antenna and a signal processing unit. The reference antenna is a directional antenna that captures the direct signal from the WiFi access point (AP). On the other hand, the surveillance antenna is Omni-directional to capture the reflected signals of the targets present anywhere in the sensing area. The signals reflected off the targets are attenuated, time-delayed, and Doppler-shifted direct signals. The time delay is directly proportional to the target range, Doppler shift to the target's velocity, and the complex reflectivity to the target's size, shape, and material. The radar signal processing unit aims at estimating these parameters using both the direct and the reflected signals. It employs match-filtering in the digital domain wherein the direct and the reflected signals are cross-correlated in the delay-Doppler plane to generate cross ambiguity function (CAF) plots. The match-filtering is adopted to maximize the signal-to-noise ratios.

We simulate a standard IEEE 802.11g WiFi signal using MATLAB's WLAN toolbox and human animation data using a marker-based motion capture technology. We hybridize both together to generate the radar scatterings off the humans, as shown in Fig. **??**. We describe these steps in greater detail in the following sections.

1) A PWR Signal Model

We use MATLAB's WLAN toolbox to generate the IEEE 802.11g standard-compliant orthogonal frequency-division

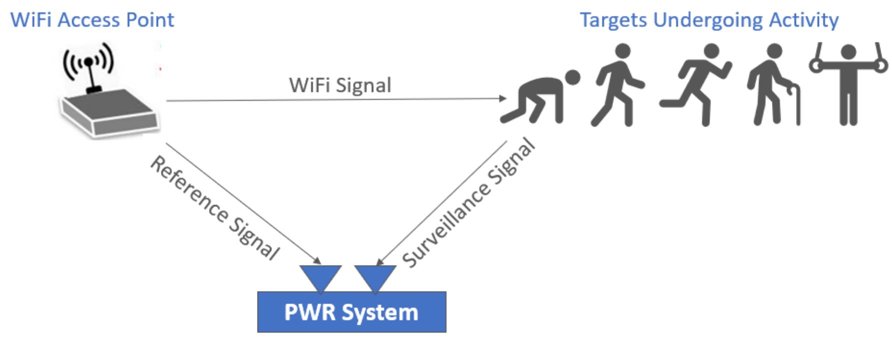

Fig. 1: A typical passive WiFi radar scenario comprising of transmissions from WiFi access points and targets undergoing motion in the same propagation environment

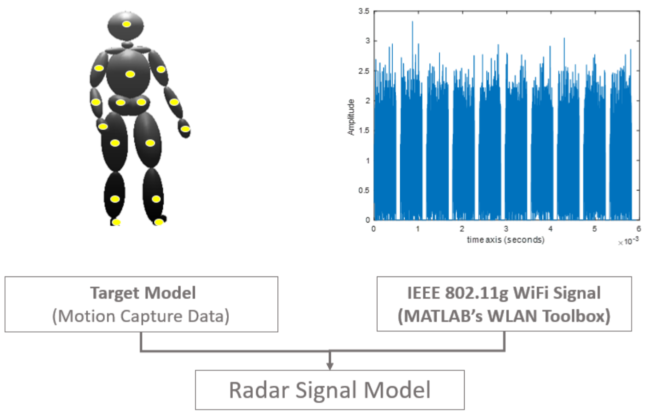

Fig. 2: Radar signal model after integration of the target model with WiFi transmissions

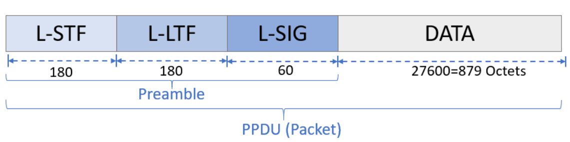

Fig. 3: IEEE 802.11g standard complaint OFDM transmission packet structure

multiplexing (OFDM) waveform [35]. The other standards that can be synthesized with the tool are 802.11a, b, g, n, ad, ac, ah, and ax. However, we restrict our discussion to 802.11g packet structures to mimic the real WiFi transmission formats at the 2.4GHz band with a channel bandwidth of $BW = 20MHz$. The physical layer of IEEE 802.11g standards use a packet-based protocol. Each transmission packet (a physical layer conformance procedure (PLCP) protocol data unit (PPDU)) comprise of a preamble part and the data part, as shown in Fig. 3. The preamble field is embedded with three subfields- 180 bit legacy short training field (L-STF), 180 bit legacy long training field (L-LTF), and 60 bit legacy signal field (L-SIG). L-STF possesses excellent correlation properties and is therefore used for detecting the start of the packet, L-LTF field for communication channel estimation, and the third preamble field L-SIG for indicating the amount of data transmitted (in octets). On the other hand, the data field contains information such as user payload, medium access control (MAC) headers, and the cyclic redundancy check (CRC) bits. The data bits together with the preamble bits form a discrete-time sequence $x_T[n]$. The sequence $x_T[n]$ undergoes some steps before being up-converted for transmission into the air. $x_T[n]$, is first passed through a digital-to-analog-converter as shown in 1.

$$x_T(t) = \sum_{n=1}^{N} x_T[n]\delta(t - nt_s) \tag{1}$$

Here N, is the total number of bits with each bit of duration $t_s = (1/BW) = 50n$sec. The signal is then amplified, and passed through a pulse shaping filter $h_T(t)$. The resulting output signal becomes

$$y_T(t) = \sqrt{(P_t)}(x_T \times h_T)(t) = \sum_{n=1}^{N} x_T[n]h_T(t - nt_s) \tag{2}$$

Where, P_t, is the transmit signal power amplification factor. The baseband signal is finally up-converted for transmission at a carrier frequency of $f_c = 2.4GHz$, as shown in 3.

$$y_T(t) = y_T(t)e^{-j2\pi f_c t} \tag{3}$$

We synthesized multiple such sequences to form a continuous stream of WiFi transmission signals. The sequences differ by a delay T_p equivalent to one sequence's transmission time and a short idle time between sequences. The resulting transmission signal is shown in 4.

$$y_T(t) = \frac{1}{\sqrt{(P)}} \sum_{p=1}^{P} y_T(t - pT_p)e^{-j2\pi f_c(t - pT_p)} \tag{4}$$

Where P is the number of transmission packets and T_p corresponds to the pulse repetition interval (PRI).

2) Dynamic Target Model

We present a realistic human simulation model in Fig. 4. The first step is to capture the animation data of dynamic humans. We use an active tracking Phase-Space system to gather the 3-dimensional time-varying location of several LED markers placed on the live actor's bodysuit [36]. The

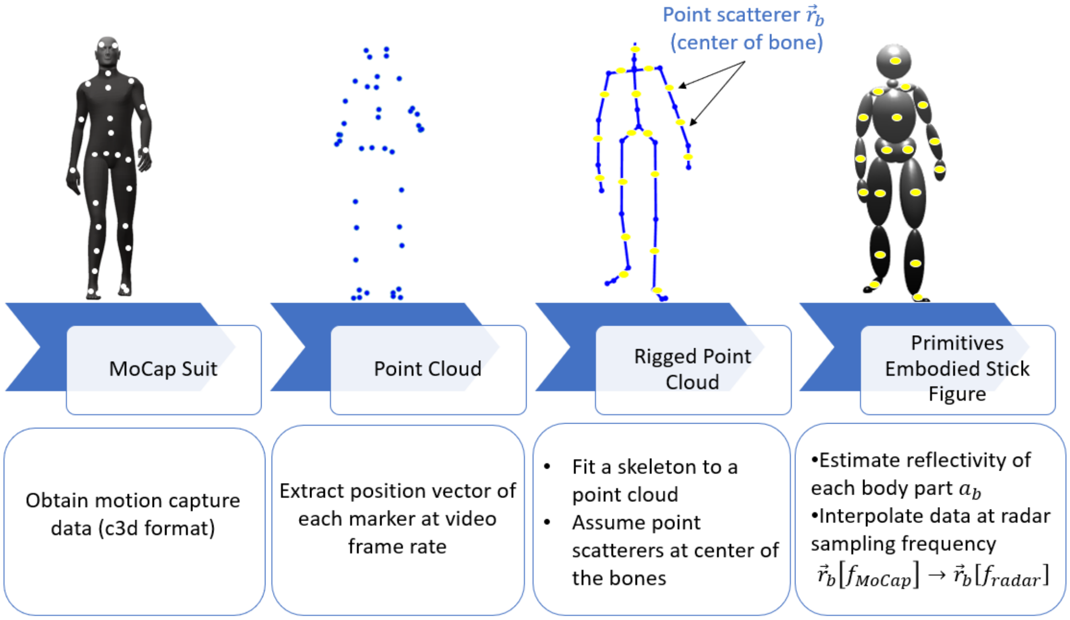

Fig. 4: Simulation framework for generating electromagnetics based human primitive model from motion capture animation data

Phase-Space system consists of 8 cameras that can track 71 markers at a frame rate of 960 frames per second (FPS). However, we use only 25 markers to extract information from 25 joints on the human body. There are several standard formats for saving an animation data file, such as C3D, BVH, ASF/AMC. Our Phase-Space system export animation data as a set of three-dimensional points into the C3D files. The C3D file format is not a hierarchical format like BVH and ASF/AMC; therefore, we rig the point cloud data with a skeleton model to specify the bones' hierarchical distribution in the human body. We finally integrate the human animation model with an electromagnetic scattering center model.

We embody the human skeleton with B elementary shapes to model different parts of the body, such as - torso, arms, and legs using ellipsoids while the head using a sphere. We assume the radar scattering centers to be lying approximately at the center of these primitive shapes. The complex reflectivities $a_b(t)$ of each of these B primitive shapes depends on various factors such as the material properties, aspect angle $\theta_b(t)$, and the position $r_b(t)$, of the scattering center on the primitive shape with respect to the radar [**?**], [19]. The reflectivity of a primitive at any time instant t is given by

$$a_b(t) = \frac{\zeta(t)\sqrt{\sigma_b(t)}}{r_b^2(t)} \tag{5}$$

Here, $\zeta(t)$ subsumes propagation effects such as attenuation, antenna directivity, processing gains, $\sigma_b(t)$ is the radar cross section of primitives. The RCS of primitive shapes are well characterised at microwave frequencies. The RCS of an ellipsoid of length L_b and radius R_b is given by

$$\sigma_b(t) = \sqrt{\Gamma}\frac{\frac{\pi}{4}R_b^4 L_b^2}{R_b^2\sin^2\theta_b(t) + \frac{1}{4}L_b^2\cos^2\theta_b(t)} \tag{6}$$

We incorporate the effect of the dielectric properties of human skin into the RCS estimation through Fresnel reflection coefficient Γ. We assume human to be a single layer dielectric with dielectric constant of 80 and conductivity of 2 S/m.

3) Hybrid electromagnetic radar scattering from dynamic humans

Assume WiFi AP transmits P packets in the propagation channel comprising human target with B number of point scatterers. The received signal $y_R(t)$ comprises both the direct signal and the complex sum of time-varying reflections from each of these B point scatterers. The target reflections are simply attenuated, time-delayed $\tau_b = 2r_b/c$, and Doppler-shifted $f_{Db} = 2v_b f_c/c$, version of the transmitted signal. Ignoring the multipath, the baseband received signal that is the signal after down-conversion can be represented as

$$y_R(t) = \sum_{p=1}^{P}\sum_{b=1}^{B} a_b(t) y_T(t - \tau_b - pT_p)e^{-j2\pi f_{Db} p T_p} + z(t) \tag{7}$$

Here, $c = 3e^8$m/s is the speed of light and $n(t)$ is the additive circular-symmetric white noise.

Since the WiFi AP transmissions are a continuous stream of signals, the received signal is also a long sequence of data spanning T_{Total} duration. Cross-correlation and Fourier processing over this enormous data is a computationally expensive task. Therefore, we process received data in M batches each of duration equal to one coherent processing

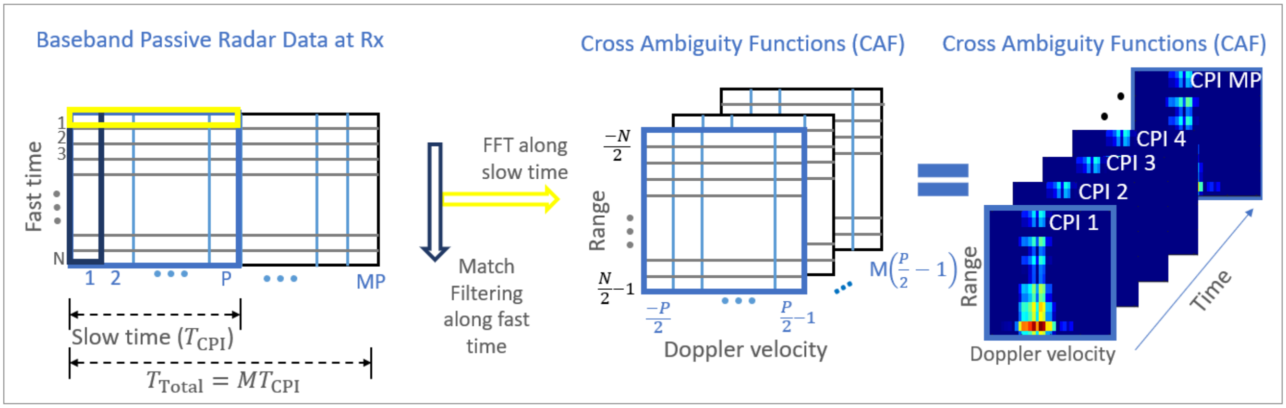

Fig. 5: cross ambiguity function generation through cross-correlation between the direct WiFi transmission and the reflected signals off the targets

interval of $T_{Total} = MT_{CPI}$. We assume the target Doppler frequencies to remain constant within one CPI T_{CPI}. The approximation follows from the fact that maximum Doppler shift f_{Dmax} is always $f_{Dmax} \ll 1/(T_p)$; therefore, the phase rotation can be approximated as a constant over one CPI. The baseband digitized signal for one CPI is shown in 8.

$$y_R[n,p] = \sum_{p=1}^{P}\sum_{b=1}^{B} a_b[n]y_T(nt_s - \tau_b - pT_p)e^{-j2\pi f_{Db}pT_p} + z[n] \tag{8}$$

Here, $n = 1:N$ is the number of fast time samples within one p^th PRI, and $p = 1:P$ are the number of PRIs (slow time samples) in one CPI.

4) CAF processing

We implement cross ambiguity function processing over two-dimensional received data $y_R[n,p]$ and the direct reference signal data (that is without target reflections) to compute the delay τ_b and Doppler information f_{Db} of the target. The adopted CAF processing is shown in Fig. 5. We perform match-filtering along the fast time samples and FFT along the slow time samples to generate CAFs for each m^{th} CPI. The regular CAF processing is shown in (9).

$$\chi_m^{RD}[\tau, f_D] = \frac{1}{NP}\sum_{p=1}^{P} x_T(nt_s)y_R^*(nt_s - \tau - pT_p)e^{j2\pi f_D pT_p} \tag{9}$$

Multiple CAFs spanning T_{Total} duration, are processed to generate the Doppler-time spectrogram as shown in Fig.6. Here, for each m^{th} CPI, the peaks along the range axis are coherently added for each Doppler bin. Mathematically it can be represented as

$$\chi^{DT}[f_D, m] = max(\sum_{n=1}^{N}\chi_m^{RD}[\tau, f_D]) \quad \forall \; f_D \; at \; each \; m^{th} \; CPI \tag{10}$$

B. Simulation Database Generation

The PWR signal parameters we use in this work are provided in TABLE I. We fix the CPI to 0.3s, which is sufficient to capture time-varying micro-Doppler features in joint time-frequency space.

TABLE I: Simulated PWR Parameters

Radar Parameters	Values
Carrier frequency (f_c)	$2.4GHz$
Bandwidth (f_s)	$20MHz$
Pulse Repetition Frequency ($PRF = 1/T_p$)	$500Hz$
Coherent Processing Interval	$0.3s$
Maximum Doppler ($f_{Dmax} = \pm PRF/2$)	$\pm 250Hz$
Doppler resolution ($\Delta f_D = 1/CPI$)	$3.3Hz$

TABLE II summarizes our entire data simulated using the described methodology and the radar signal model described in TABLE I. We simulate the human micro-Doppler data for five complicated motion classes- human body rotating (HBR), human kicking (HK), human punching (HP), human grabbing an object (HG) and a human walking (HW). The number of animation data files in each of these categories HBR, HK, HP, HG, and HW are 10, 20, 20, 20, and 19. The number of files differs because some of the animation data files had many missing markers positions, which resulted in insufficient marker information. Therefore, we decided to drop these files and continued with the remaining ones. Please note that the human motions in these repeated measurements were unrestricted, and therefore, the micro-Doppler signatures vary due to differences in gait patterns every time. The duration of each measurement is 4.5sec. We use a sliding window of duration 1.5sec with an overlapping time of 0.5sec over the entire signature of duration 4.5sec. It results in 9 spectrograms, each of duration 1.5sec from every motion capture file.

We generate the spectrograms in three different PWR radar configurations:

1) **PWR Monostatic Configuration For Fixed Aspect Angle of Target**:
In this radar configuration, the WiFi AP and radar receiver are co-located, and the target is supposed to be moving at 0° aspect angle with respect to the radar receiver. The micro-Doppler signatures for this configuration are shown in Fig. 7. Each spectrogram is 1.5sec long. Fig. 7(a), presents spectrogram of a

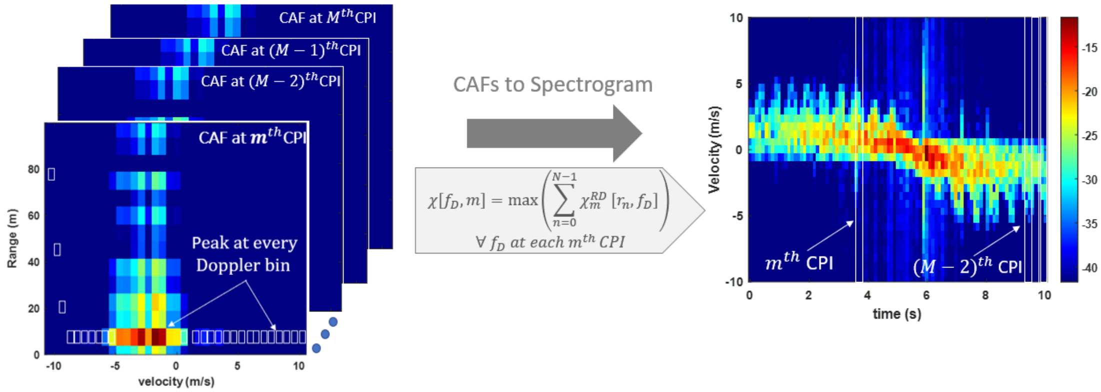

Fig. 6: Simulation methodology to generate micro-Doppler spectrograms using multiple CAFs spanning the entire duration of motion

TABLE II: Simulation Dataset Description

Target Class	Number of MoCap Files	Duration of Data (in each file)	Number of Spectrograms (1.5sec each) In Radar Configuration		
			Monostatic		Bistatic
			Aspect Angle 0 (Overlapping Window 0.5sec)	Varying Aspect Angle 0:5:360 (No Overlapping)	Varying Bistatic Angle 0:5:360 (No Overlapping)
HBR	10	4.5 sec	90	2190	2190
HK	20	4.5 sec	180	4380	4380
HP	20	4.5 sec	180	4380	4380
HG	19	4.5 sec	180	4161	4161
HW	20	4.5 sec	171	4380	4380
Total Data in Each Configuration			**801**	**19491**	**19491**

human undergoing a body rotation motion. We can observe both positive and negative Dopplers due to the rotational motion of the body. Since the human is standing at a place while performing the motion, and there is no bulk body translational motion, the body Doppler is mostly centered around zero. It holds even for three other motion categories- human kicking, human punching, and human grabbing, as shown in Fig. 7(b)(c) and (d), respectively. We notice that there are only minor differences in the micro-Doppler patterns within these motion classes due to the similarity of motions. The fifth target class shown in Fig.7(e) corresponds to a human walking in front of the radar. The human is always walking in the direction of the radar, thus have positive micro-Dopplers mostly. The Doppler resolution is limited because, in passive sensing scenarios, Doppler resolution is dictated by the CPI (fixed at 0.3sec) and the carrier frequency, which is 2.4GHz for the IEEE 802.11g standard. We obtain 90, 180, 180, 180, and 171 spectrograms from HBR, HK, HP, HG, and HW motion category.

2) **PWR Monostatic Configuration For Varying Aspect Angle of Target**
In most realistic scenarios, the human motions might not be restricted to a single aspect with respect to the radar. In such scenarios, the spectrograms might differ significantly. It could be due to the shadowing of some part of the human body if captured at different angles. Fig.8(a)-(d), shows the spectrograms of a human walking at four different aspect angles- 0°, 60°, 120° and 180° with respect to the radar respectively. The signatures at 60° shown in Fig.8(b), mostly have reduced positive Dopplers compared to 0° angle. It is because the target is still approaching but with reduced radial component in the direction of the radar. The Dopplers become negative when the target aspect angle is 120°. Here, at this angle, the target is supposed to be moving away from the radar. Finally, the signatures at 180° represent a human walking away from the radar, thus have all the negative Dopplers.
We gather micro-Doppler data for aspect angles varying from 0° to 360° with an interval of 5° resulting in 73 unique spectrograms of duration 4.5sec each for every file. We further divide the resulting spectrograms into three of the duration 1.5sec each. We repeat the process for all of the files and all the target classes. The resulting number of spectrograms are shown in TABLE II.

3) **PWR Bistatic Configuration For Varying Bistatic Angle And Fixed Target Aspect Angle**
In most passive WiFi sensing scenarios, a certain distance separates the WiFi AP and the radar receiver. Therefore, this motivates us to do simulations that can capture the effect of varying bistatic angles. In our simulations, we varied the bistatic angle from 0° to 360° with an interval of 5° like the previous case and obtained a humongous amount of radar data. The

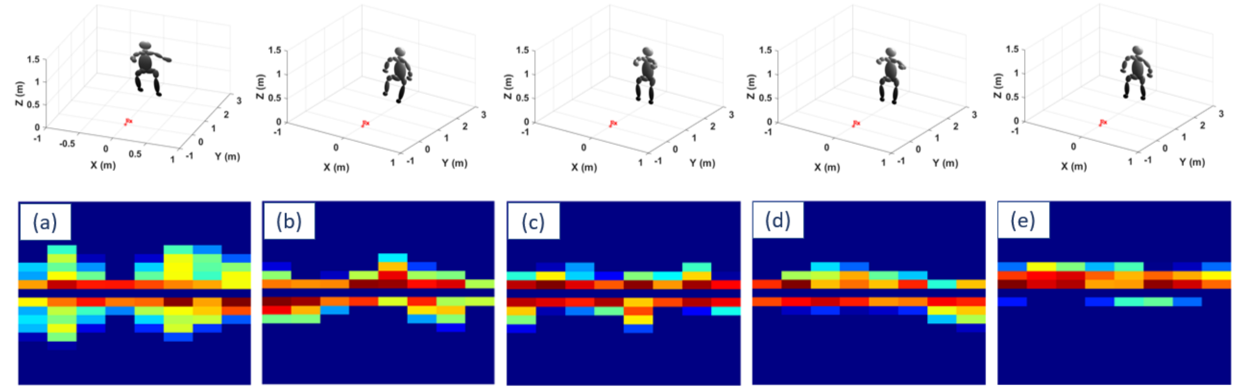

Fig. 7: Radar micro-Doppler signatures for a human undergoing (a) a body rotation motion, (b) kicking motion, (c) punching motion, (d) grabbing an object motion and (e) walking in the direction of the monostatic configuration of PWR radar.

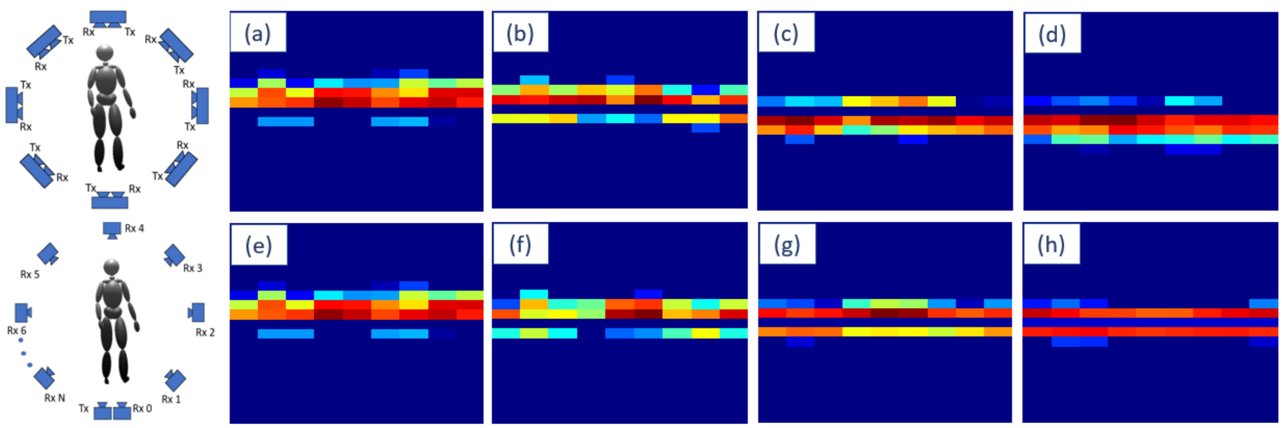

Fig. 8: (a)-(d) Radar micro-Doppler signatures of human walking at four aspect angles 0°, 60°, 120°, and 180° with respect to the radar respectively. (e)-(h) Radar micro-Doppler signatures of human walking at aspect angles 0° to the radar at four bistatic radar configurations with following bistatic angle 0°, 60°, 120° and 180° respectively.

dataset size is shown in TABLE II.The micro-Doppler signatures of a walking human at four bistatic angles 0°, 60°, 120° and 180° are shown in Fig.8(e)-(h) respectively. In bistatic scenarios, the Doppler frequency $f_D = \frac{2v f_c}{c} cos(\theta) cos(\beta/2)$, is governed by aspect angle θ and the bistatic angle β. For simplicity, we keep the aspect angle of the target fixed at 0°; therefore, the Dopplers only depend upon the bistatic angles $f_D = \frac{2v f_c}{c} cos(\beta/2)$. When $\beta = 180°$, the Doppler should be zero ($f_D = 0$). It is known as a forward scatterer position. However, since humans are the extended targets, we get some micro-Doppler returns due to the swinging motion of arms and legs. This is evident in Fig.8(h).

III. Simulation Results

In this section, we study the different classification algorithms' robustness to classify micro-Dopplers in more complex scenarios, such as varying aspect angles and varying bistatic angles. We use handpicked features [37], Cadence velocity features (CVD) [38], and automatically extracted sparse features [18], [40] from the micro-Doppler signatures to test the performance of classical machine learning-based support vector machine (SVM) classifier [44]. We then compare their performances with the recent deep convolutional neural network (DCNN), which has a joint feature extraction and classification framework within the same network. We designed a 24-layered deep network comprising three components (convolution layer, pooling layer, and activation functions). We also test some of the pre-trained deep neural networks such as AlexNet, GoogLeNet, and ResNet18. We used 70% of each target's spectrograms as the training data set, 15% as the validation set, and the remaining 15% as the test data set. The algorithms are run on an Intel(R) Core(TM) i7-5500U CPU running at 2:40 GHz; 16-GB RAM,Windows 10 (64 b).

To give readers a better understanding of the sensitivity of the algorithm's performance to simulation database; we considered the following three classification scenarios.

Case 1a: Train using data from a fixed zero aspect angle: We trained and tested the algorithms' performances using a simulation database generated for a fixed $\circ$ aspect angle of the target. The resulting classification accuracies are

TABLE III: Classification accuracies of multiple algorithms for a simulation database (captured for a fixed aspect angle of the target)

Target Class/ Algorithm	Handpicked Features	CVD Features	Sparse Features	Neural Networks			
				DCNN	AlexNet	GoogLeNet	ResNet18
HBR	100	98.9	99.1	100	100	100	100
HK	83.9	83	93.1	100	100	100	96.3
HP	73.8	92.3	92.4	100	96.3	96.3	100
HG	82.5	98.3	97.6	96.3	100	100	100
HW	100	94.5	100	100	100	100	100
Overall Accuracy (%)	**88**	**93.4**	**96.4**	**99.3**	**99.3**	**99.3**	**99.3**

presented in TABLE III. We observe that all deep neural networks outperform the classical machine learning-based methods and achieve an average classification accuracy of ≈ 99%. The five target classes share common features in micro-Doppler feature space because of the proximity between different motion categories, resulting in poor performance of the classical feature extraction methods. The performance using the handpicked features is 88%, while the CVD features' performance is 93.4%. The performance using sparse features shows that the sparsity-based algorithms can extract underlying features in different hyperplanes, resulting in good classification accuracy of 96.4% even when the motion classes are similar. We also observe that HW class is hardly getting confused with other target classes since HW shares less similarity with other motion classes such as HBR, HK, HP, and HG.

Case 1b: Train using data from multiple aspect angle: Next, we analyse the performance when the algorithms are trained using a simulation database comprising micro-Doppler signatures captured at multiple aspect angles. It is a significantly more challenging and realistic scenario since no aspect angle information is available during the test phase. Please note that the spectrograms used at the test time have not been used during training. TABLE IV presents the resulting classification accuracies across different algorithms. Here, we present the classification accuracy in the form of confusion matrices to draw a more intuitive sense of the results. The average classification accuracies for handpicked features, CVD features, are 69.2% and 80.9%, respectively. The reason for poor performance is that the Doppler spectrogram for a particular motion class at certain aspect angles might look similar to the spectrogram for other motion classes at the same angle. Therefore, the handpicked and the CVD features, are not discriminative enough and result in poor performance. There is degradation of 5% in the performance of sparsity based algorithm. This could be due to sharing of same subspace in sparse domain between different classes for different aspect angles. HW remains the best recognised class amongst all the classes considered in the study as it has more distinctive spectrograms from rest of the classes. On the other hand, other algorithms based on deep networks perform exceptionally well even under diverse training and test datasets. It indicates that these algorithms are specifically suited for problems dealing with a great deal of diversity in the radar data. The best performing network is ResNet18, with an average classification accuracy of 97.8%.

Case 1c: Train using data from multiple bistatic angles: Finally, we train the algorithms with a simulation dataset comprising micro-Doppler data captured at different bistatic angles. It is a more practical scenario one can encounter. Here, we believe the performance to be lower compared to the previous two cases. It is because some micro-Doppler signatures would be captured in a forward scatterer position, resulting in micro-Dopplers centered mostly around 0 Doppler for almost all the motion classes. It can lead to a significant reduction in the classification accuracies. We present the classification results for this case in TABLE V. We note that HW's performance is good across all the algorithms (≥ 90). Compared to the previous case, there is a drop of ≈ 3% in classification accuracy for AlexNet, GoogLeNet, and DCNN. ResNet seems to be working best amongst all classifiers with an average classification accuracy of up to 95.8% (2% reduction compared to the previous case).

A. Performance Under Noisy Conditions

The results presented so far have been computed for high signal to noise ratios ($SNR \approx$ dB). Most realistic scenarios have environmental factors that significantly affect the resulting SNRs. Therefore, to assess SNR's impact on the classification performances, we introduced additive Gaussian noise of varying levels of SNR (-2 to 10dB) to the simulated micro-Doppler signatures. The simulation database used for this study is captured for a fixed ∘ aspect angle of the target. We carried out ten repetitive trials using randomly selected training and test noisy micro-Doppler signatures. Fig. 9 shows the variation of average classification performances of different algorithms as a function of SNR. As the noise level is increased, we observe an expected drop in different algorithms' overall performance. However, the drop in performance is higher for classical learning-based methods (represented with black color) than deep learning models (represented with red color). Deep networks can extract features that are more robust to noise and would be preferable over other possible features.

IV. Measurements

In this section, we do a more thorough evaluation of the algorithms' performance in real-world scenarios.

A. Measurement Data Collection

We deployed the real PWR measurement setup and the motion capture system to simultaneously gather the simulation animation data and the measurement PWR data

TABLE IV: Classification accuracies for a simulation database (captured for varying aspect angle of the target with respect to the radar receiver)

Algorithm		True/Pred Class	HBR	HK	HP	HG	HW
Handpicked Features		**HBR**	**95.4**	3.4	0.9	0.2	0
		HK	2.4	**70.9**	8.7	14.5	3.5
		HP	2.7	32.2	**20.1**	43	1.9
		HG	0.6	18.4	11.9	**68.9**	0.1
		HW	0.6	6.2	1.3	1.4	**90.3**
CVD Features		**HBR**	**93.4**	6.2	0.2	0.2	0
		HK	3.4	**57.4**	9.7	20.9	8.6
		HP	1.4	7.1	**87.4**	1	2.9
		HG	0.1	5.7	3.8	**79.7**	10.7
		HW	0.8	0.8	4.4	11.3	**82.6**
Sparse Features		**HBR**	**98.6**	0.9	0	0.5	0
		HK	0.1	**81.8**	10	5.2	2.9
		HP	0.6	8.1	**88.1**	3.2	0
		HG	0	3.3	6.5	**90.2**	0
		HW	0	2.8	0.6	0.1	**96.5**
Neural Networks	**DCNN**	**HBR**	**100**	0	0	0	0
		HK	0.3	**94.2**	4.4	0.8	0.3
		HP	0.9	3.3	**95.7**	0	0
		HG	0.3	2.9	3.4	**92.5**	0.9
		HW	0.3	11.9	0	0	**87.8**
	AlexNet	**HBR**	**100**	0	0	0	0
		HK	0.3	**91.8**	3.9	0.5	3.5
		HP	0.6	0.8	**98.3**	0	0.3
		HG	0.2	1.2	8.5	**89.3**	0.8
		HW	0	1.4	0	0	**98.6**
	GoogLeNet	**HBR**	**99.4**	0	0	0.6	0
		HK	0.2	**88.6**	1.5	1	8.7
		HP	0.3	4.9	**93.6**	1.1	0.1
		HG	0	1.2	0.3	**98.5**	0
		HW	0	0.5	0	0.1	**99.4**
	ResNet18	**HBR**	**100**	0	0	0	0
		HK	0	**93.6**	1.7	1	3.7
		HP	0	0.6	**97.6**	1.7	0.1
		HG	0	0.5	0.1	**99.4**	0
		HW	0	0.2	0	0	**99.8**

TABLE V: Classification accuracies for a simulation database (captured under varying bistatic circular configurations)

Target Class/ Algorithm	Handpicked Features	CVD Features	Sparse Features	Neural Networks			
				DCNN	AlexNet	GoogLeNet	ResNet18
HBR	93.6	91.8	99.1	80.2	82.6	87.2	91.2
HK	17.2	35.5	81.8	86.9	83.9	85.8	91.6
HP	21.2	88.7	89.7	97	95.9	91.3	99.5
HG	64.4	60.3	89.3	94.8	97.9	93.6	95.1
HW	90.9	73.8	95.8	98.6	99.2	100	99.4
Overall Accuracy (%)	**57.46**	**70.2**	**91.14**	**91.5**	**91.9**	**91.58**	**95.79**

from the same motion classes HBR, HK, HP, HG, and HW. The measurement setup is depicted in Fig.1. In real setup, we used a Raspberry Pi to attempt handshakes with a WiFi access point ('AP') to generate WiFi transmissions. By constantly transmitting probe requests, the WiFi AP continuously emitted probe response signals. We captured these transmissions at two antennas- reference antennas and surveillance antenna. The reference channel receives direct transmissions from the WiFi AP, while the surveillance channel gathers signals reflected off the targets moving in the same propagation channel. Fig.10(a)-(e), shows measured spectrograms corresponding to five motion classes HBR, HK, HP, HG and HW respectively. We can see from the spectrograms that four target classes HBR, HK, HP, and HG, have a periodic motion with positive and negative Dopplers. HW spectrogram, shown in Fig.10(e), can be discerned from other motion classes as it has a Doppler shift due to bulk body motion and additional micro-Dopplers. However, due to several environmental factors such as multipath, shadowing, path-loss, the measured spectrograms are a bit noisy.

TABLE VI summarizes our entire measurement data.

B. Measurement Results and Analyses

We validated the performance of different algorithms across a common measurement dataset, and the resulting accuracies are reported in TABLE VII. Since the real spectrograms subsume environment effects such as noise, multipath, and shadowing, we expect the performances to lower than simulations. We can already see from the results, handpicked features and CVD features completely fail to classify motion classes in 60% of the cases. Because now the real spectrograms have noise and multipath components,

TABLE VI: Measurement Dataset Description

Target Class	Number of Measurement Files	Duration of Data (in each file)	Number of Spectrograms (1.5sec each) In Monostatic PWR Configuration Scenario (Overlapping Window of 0.5sec)
HBR	20	4.5 sec	180
HK	20	4.5 sec	180
HP	20	4.5 sec	180
HG	20	4.5 sec	180
HW	20	4.5 sec	180
Total Data in Each Configuration			**900**

TABLE VII: Classification accuracies of different algorithms using measurement data (captured using monostatic configuration of PWR radar)

Target Class/ Algorithm	Handpicked Features	CVD Features	Sparse Features	Neural Networks			
				DCNN	AlexNet	GoogLeNet	ResNet18
HBR	31.3	45.5	48.7	100	81.5	85.2	88.9
HK	3.1	64.7	90.9	85.2	85.2	85.2	100
HP	73.7	85.3	71.9	74.1	96.8	100	85.2
HG	21.1	12.5	39.5	92.6	93.1	77.8	92.6
HW	77.5	9.7	81.8	96.3	96.3	96.3	96.3
Overall Accuracy (%)	**41.3**	**43.5**	**66.6**	**89.6**	**90.6**	**88.9**	**92.6**

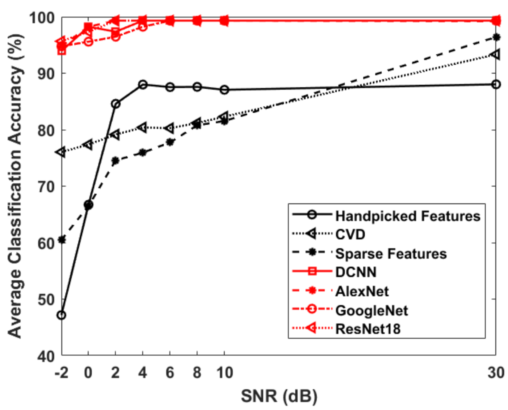

Fig. 9: Average classification accuracies as a function of SNR

the spectrograms can no longer be well represented through the sparse features. Therefore, the average classification accuracy sparsity-based algorithm is reduced to 66.6% only. Hence, we believe that the classical machine learning algorithms are not suited for dealing with a great deal of diversity in the radar data.

On the other hand, the neural network's results indicate that the deep networks are more robust to environmental factors such as noise and can perform well even in diverse operating conditions. We were able to obtain classification accuracies $\approx$ 90% for all the neural networks. The ResNet performs best with an average classification accuracy of 92.6%.

V. Conclusion

In this work, we first presented a bespoke simulator that can simulate human micro-Doppler radar returns as a function of a diverse set of target parameters, radar parameters, and radar signal processing parameters. We used three different simulation databases with different parameter variations to study and evaluate machine learning and deep learning algorithms' classification performances in more complex scenarios. We achieved average classification accuracies $\geq$ 90% for almost all the deep learning models. The classification was challenging due to the close similarity of motion classes considered in the study. To validate the performances of the algorithms, we gathered measurement data. Please note that this data was captured simultaneously with the motion capture animation data to maintain consistency across both simulations and measurements. We were able to get decent classification accuracies (up to 90%) even with the real dataset compared to those obtained from simulations.

The study demonstrates the feasibility of passive WiFi sensing for activity recognition applications. Current experiments presented in the paper are limited due to the small number of participants and the diversity of experimental scenarios (considered only line-of-sight conditions). Therefore, we plan to extend our simulation framework to incorporate more participants and more realistic indoor through-wall scenarios, including environmental factors such as multipath and shadowing. We also plan to add more motion categories like falling, hand gestures, bending down to our existing simulator. Interested researchers can download the simulator from https://uwsl.co.uk/. We believe the simulator will pave the way for benchmarking future algorithms and generate large volumes of simulation data.

Acknowledgments

This work is part of the OPERA project funded by the UK Engineering and Physical Sciences Research Council (EPSRC), Grant No: EP/R018677/1.

References

[1] F. Fioranelli, M. Ritchie, and H. Griffiths, "Classification of unarmed/armed personnel using the netrad multistatic radar for micro-doppler and singular value decomposition features," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1933–1937, 2015.

[2] ——, "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.

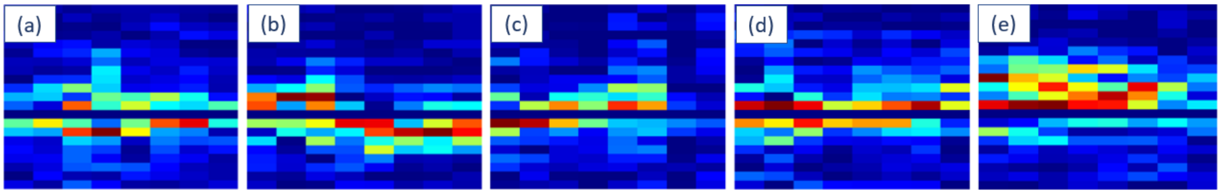

Fig. 10: Real micro-Doppler signatures for a human undergoing (a) a body rotation motion, (b) kicking motion, (c) punching motion, (d) grabbing an object motion and (e) a human walking in the direction of the monostatic configuration of PWR radar.

[3] Q. Chen, Y. Liu, F. Fioranelli, M. Ritchie, B. Tan, and K. Chetty, "Dopnet: A deep convolutional neural network to recognize armed and unarmed human targets," *IEEE Sensors Journal*, vol. 19, no. 11, pp. 4160–4172, 2019.

[4] D. P. Fairchild and R. M. Narayanan, "Micro-doppler radar classification of human motions under various training scenarios," in *Active and Passive Signatures IV*, vol. 8734. International Society for Optics and Photonics, 2013, p. 873407.

[5] ——, "Classification of human motions using empirical mode decomposition of human micro-doppler signatures," *IET Radar, Sonar & Navigation*, vol. 8, no. 5, pp. 425–434, 2014.

[6] B. Erol and M. G. Amin, "Radar data cube analysis for fall detection," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 2446–2450.

[7] A.-K. Seifert, M. G. Amin, and A. M. Zoubir, "Toward unobtrusive in-home gait analysis based on radar micro-doppler signatures," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2629–2640, 2019.

[8] 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.

[9] D. L. Hall, T. D. Ridder, and R. M. Narayanan, "Abnormal gait detection and classification using micro-doppler radar signatures," in *Radar Sensor Technology XXIII*, vol. 11003. International Society for Optics and Photonics, 2019, p. 110030Q.

[10] 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*, 2020.

[11] Q. Chen, B. Tan, K. Chetty, and K. Woodbridge, "Activity recognition based on micro-doppler signature with in-home wi-fi," in *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE, 2016, pp. 1–6.

[12] M. S. Seyfioğlu, S. Z. Gürbüz, A. M. Özbayoğlu, and M. Yüksel, "Deep learning of micro-doppler features for aided and unaided gait recognition," in *2017 IEEE Radar Conference (RadarConf)*. IEEE, 2017, pp. 1125–1130.

[13] F. Fioranelli, H. Li, J. Le Kernec, V. Busin, N. Jonsson, G. King, M. Tomlinson, and L. Viora, "Radar-based evaluation of lameness detection in ruminants: preliminary results," in *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, vol. 1. IEEE, 2019, pp. 1–4.

[14] 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.

[15] M. G. Amin, A. Ravisankar, and R. G. Guendel, "Rf sensing for continuous monitoring of human activities for home consumer applications," in *Big Data: Learning, Analytics, and Applications*, vol. 10989. International Society for Optics and Photonics, 2019, p. 1098907.

[16] S. Zubeyde Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," in *Signal Processing Magazine*. IEEE, 2019.

[17] Y. Kim and B. Toomajian, "Hand gesture recognition using micro-doppler signatures with convolutional neural network," *IEEE Access*, vol. 4, pp. 7125–7130, 2016.

[18] G. Li, R. Zhang, M. Ritchie, and H. Griffiths, "Sparsity-driven micro-doppler feature extraction for dynamic hand gesture recognition," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 2, pp. 655–665, 2017.

[19] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Analysis of micro-doppler signatures," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 271–276, 2003.

[20] ——, "Micro-doppler effect in radar: phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and electronic systems*, vol. 42, no. 1, pp. 2–21, 2006.

[21] K. Chetty, G. E. Smith, and K. Woodbridge, "Through-the-wall sensing of personnel using passive bistatic wifi radar at standoff distances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1218–1226, 2011.

[22] S. Vishwakarma and S. S. Ram, "Detection of multiple movers based on single channel source separation of their micro-dopplers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, pp. 159–169, 2017 .

[23] V. C. Chen, *The micro-Doppler effect in radar*. Artech House, 2019.

[24] W. Li, B. Tan, and R. J. Piechocki, "Wifi-based passive sensing system for human presence and activity event classification," *IET Wireless Sensor Systems*, vol. 8, no. 6, pp. 276–283, 2018.

[25] B. Tan, Q. Chen, K. Chetty, K. Woodbridge, W. Li, and R. Piechocki, "Exploiting wifi channel state information for residential healthcare informatics," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 130–137, 2018.

[26] 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 Transactions on Geoscience and Remote Sensing*, 2020.

[27] 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. 12 426–12 437, 2020.

[28] W. Li, Y. Xu, B. Tan, and R. J. Piechocki, "Passive wireless sensing for unsupervised human activity recognition in healthcare," in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2017, pp. 1528–1533.

[29] G. E. Smith, K. Woodbridge, and C. J. Baker, "Multistatic micro-doppler signature of personnel," in *Radar Conference, 2008. RADAR'08. IEEE*. IEEE, 2008, pp. 1–6.

[30] R. Boulic, N. M. Thalmann, and D. Thalmann, "A global human walking model with real-time kinematic personification," *The visual computer*, vol. 6, no. 6, pp. 344–358, 1990.

[31] S. S. Ram and H. Ling, "Simulation of human microdopplers using computer animation data," in *2008 IEEE Radar Conference*. IEEE, 2008, pp. 1–6.

[32] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, "Simulation and analysis of human micro-dopplers in through-wall environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 2015–2023, April 2010.

[33] B. Erol and S. Z. Gurbuz, "A kinect-based human micro-doppler simulator," *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 5, pp. 6–17, 2015.

[34] A. D. Singh, S. S. Ram, and S. Vishwakarma, "Simulation of the radar cross-section of dynamic human motions using virtual reality data and ray tracing," in *2018 IEEE Radar Conference (RadarConf18)*. IEEE, 2018, pp. 1555–1560 .

[35] W. Matlab, "Toolbox. 2018," 2018.

[36] "Phase-space motion capture system," online accessed on 10 August 2020. [Online]. Available: https://www.phasespace.com/

[37] Y. Kim and H. Ling, "Human activity classification based on micro-doppler signatures using a support vector machine," *IEEE Transac-

tions on Geoscience and Remote Sensing, vol. 47, no. 5, pp. 1328–1337, 2009.

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

[39] S. Vishwakarma and S. S. Ram, "Dictionary learning for classification of indoor micro-doppler signatures across multiple carriers," in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 0992–0997.

[40] ——, "Dictionary learning with low computational complexity for classification of human micro-dopplers across multiple carrier frequencies," *IEEE Access*, vol. 6, pp. 29 793–29 805, 2018.

[41] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and< 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.

[42] Z. Zhong, L. Jin, and Z. Xie, "High performance offline handwritten chinese character recognition using googlenet and directional feature maps," in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, 2015, pp. 846–850.

[43] S. Ayyachamy, V. Alex, M. Khened, and G. Krishnamurthi, "Medical image retrieval using resnet-18," in *Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications*, vol. 10954. International Society for Optics and Photonics, 2019, p. 1095410.

[44] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. Murthy, "A fast iterative nearest point algorithm for support vector machine classifier design," *IEEE transactions on neural networks*, vol. 11, no. 1, pp. 124–136, 2000.