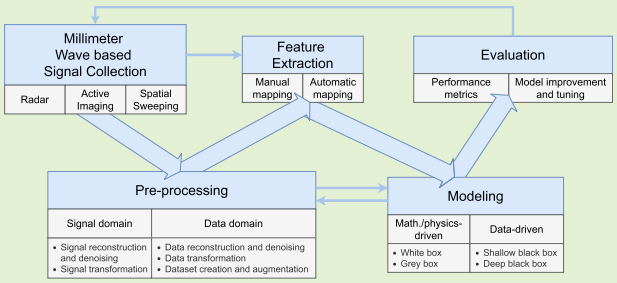
Millimeter Wave Sensing: A Review of Application Pipelines and Building Blocks

Bram van Berlo, Amany Elkelany, Tanir Ozcelebi, and Nirvana Meratnia, *Member, IEEE*

***Abstract*—The increasing bandwidth requirement of new wireless applications has lead to standardization of the millimeter wave spectrum for high-speed wireless communication. The millimeter wave spectrum is part of 5G and covers frequencies between 30 and 300 GHz that correspond to wavelengths ranging from 10 to 1 mm. Although millimeter wave is often considered as a communication medium, it has also proved to be an excellent 'sensor', thanks to its narrow beams, operation across a wide bandwidth, and interaction with atmospheric constituents. In this paper, which is to the best of our knowledge the first review that completely covers millimeter wave sensing application pipelines, we provide a comprehensive overview and analysis of different basic application pipeline building blocks, including hardware, algorithms, analytical models, and model evaluation techniques. The review also provides a taxonomy that highlights different millimeter wave sensing application domains. By performing a thorough analysis, complying with the systematic literature review methodology and reviewing 165 papers, we not only extend previous investigations focused only on communication aspects of the millimeter wave technology and using millimeter wave technology for active imaging, but also highlight scientific and technological challenges and trends, and provide a future perspective for applications of millimeter wave as a sensing technology.**



***Index Terms*—5G, analytical modeling, millimeter wave, millimeter wave sensing application pipeline, radar, systematic literature review.**

I. Introduction

INTRODUCTION of new wireless applications, their higher service quality requirements, and a significant increase of application users have put continuous demand on digital wireless communication bandwidth. To address this, Netflix and YouTube, for instance, have reduced video streaming quality in Europe to mitigate data traffic peaks in the 2.4 and 5 GHz WiFi bands. These peaks are caused by an increased amount of video and movie streams due to home confinement measures taken during the COVID-19 pandemic [1]. In the future, 8K video streaming will require a minimum data rate of 50 Mbps per TV. In case multiple simultaneous streams are launched on the same network, gigabit connections will approach their limit [2]. Maintaining the required network bandwidth and processing speed remains challenging for cloud gaming as a service [3].

The increasing bandwidth requirement has lead to standardization of the millimeter wave spectrum for high-speed wireless communication in the IEEE 802.11ad, 802.11aj, and 802.11ay amendments [4]–[6]. The millimeter wave spectrum is also part of the fifth-generation mobile communication technology standard (5G) [7], [8]. Although use of the millimeter wave spectrum for communication is often directly associated with 5G, there are key differences. The millimeter wave spectrum is just one part of what 5G networks use to provide higher data rates. It covers frequencies between 30 and 300 GHz that correspond to wavelengths ranging from 10 to 1 mm [9], [10] and is located between the centimeter wave and terahertz wave spectra.

Thanks to covering a wide bandwidth and utilizing a short wavelength, the millimeter wave spectrum is able to provide higher data rates compared to other widely used wireless technologies such as WiFi. Compared to the spectrum in which 5 GHz WiFi operates, the bandwidth of the millimeter wave spectrum is 10x higher, i.e., 270 GHz versus 27 GHz. Under the assumption that environmental and legislative limitations do not exist, this means that the number of channels that can be created in the millimeter

Manuscript received November 9, 2020; revised January 19, 2021; accepted January 24, 2021. Date of publication February 5, 2021; date of current version April 5, 2021. The associate editor coordinating the review of this article and approving it for publication was Prof. Piotr J. Samczynski. *(Corresponding author: Bram van Berlo.)*

The authors are with the Interconnected Resource-Aware Intelligent Systems Cluster, Department of Mathematics and Computer Science, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (e-mail: b.r.d.v.berlo@tue.nl; a.m.a.elkelany@tue.nl; t.ozcelebi@tue.nl; n.meratnia@tue.nl).

Digital Object Identifier 10.1109/JSEN.2021.3057450

wave spectrum is also 10x greater. Because millimeter waves operate at much higher frequencies compared to the WiFi bands, the wavelengths are much shorter. Consequently, the size of electronic components can be reduced. This, in turn, causes the beam emitted by electronic components to be much narrower. The narrow beams, in combination with greater signal attenuation compared to the WiFi bands, allow increased communication density [10], i.e., allow an increased number of messages communicated in parallel over separate links with the same carrier frequency in a limited area.

Although millimeter wave is often considered as a communication medium, it has also proved to be an excellent *'sensor'* for humans, objects, and environmental sensing [9], [10] thanks to its narrow beams, operation across a wide bandwidth, and interaction with atmospheric constituents. Narrow beams result in greater sensing resolution and directivity. A wide bandwidth and specific penetration, reflection, and attenuation reactions to different materials allow distinguishing different objects and humans [11], [12].

However, using millimeter waves in sensing applications also has disadvantages. Signals at extremely high frequencies suffer from significant attenuation. Millimeter waves can therefore hardly be used for long-distance applications [13]. The financial cost of deploying a millimeter wave sensing system, e.g., on a vehicle, is still high even though this cost is projected to decrease in the future [14]. Millimeter waves also suffer significant penetration loss through solid materials like concrete. As rain-drops are comparable in size to the wavelength of millimeter waves, heavy rain can cause considerable attenuation due to scattering [9], [10].

The applications that use millimeter waves utilize a wide variety of hardware and algorithms for data collection, data pre-processing, feature extraction, feature analysis, analytical modeling, and modeling evaluation. This paper provides a comprehensive overview and analysis of millimeter wave sensing application pipelines and basic building blocks, including hardware, algorithms, analytical models, and model evaluation techniques. It also offers an insight into challenges and trends, and provides a future perspective for applications of millimeter wave as a sensing technology. Therefore, the review has struck a balance between the amount of details that are provided for each millimeter wave sensing application pipeline building block and does not elaborate on details of each application domain and algorithm independently. Due to the multidisciplinary nature of millimeter wave sensing application pipelines, the review offers technical details for each reader depending on familiarity with a certain field. For example, data collection and a part of data pre-processing are topics of electrical engineering, while analytical modeling and modeling evaluation are topics of data science and artificial intelligence. The main contributions of this paper are:

- Extending previous investigations focused *only on* communication aspects of the millimeter wave technology and using millimeter wave technology for active imaging. To the best of our knowledge, this is the first review that completely covers millimeter wave sensing application pipelines and pipeline building blocks.

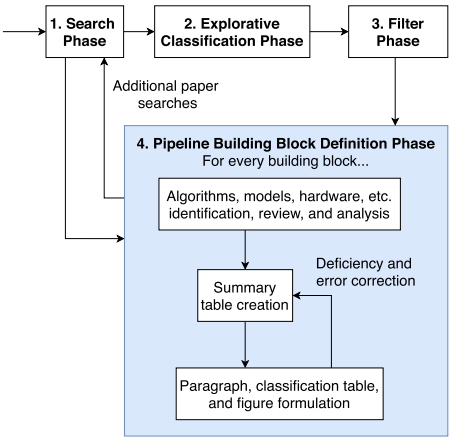


Fig. 1. Review process block diagram.

- Providing a taxonomy that highlights different millimeter wave sensing application domains.
- Making a comprehensive review and analysis of hardware, algorithms, analytical models, and model evaluation techniques for each of the identified millimeter wave sensing application pipeline building blocks.
- Identifying commonalities, gaps, and shortcomings of current studies and solutions focusing on the use of millimeter wave as a sensing technology.
- Highlighting scientific and technological challenges and trends, and providing a future perspective for applications of millimeter wave as a sensing technology.

The review is organized as follows: Section II explains the review methodology. Section III establishes an application taxonomy based on the research papers that report on millimeter wave sensing. Section IV identifies the common building blocks that make up the application pipelines presented in the millimeter wave sensing papers. It also presents a thorough analysis of the wide variety of hardware, algorithms, analytical models, and model evaluation techniques across these papers in the respective building blocks. Section V identifies and explains the challenges and trends, and provides a future perspective for applications of millimeter wave as a sensing technology. Finally, Section VI concludes the paper.

II. Review Methodology

We have conducted this review with a specific process involving four sequential phases. These phases are depicted in Figure 1. The process is based on the iterative Bioinformatic and Systematic Literature Review (BiSLR) spiral model developed by Mariano *et al.* [15]. The differences compared to the iterative BiSLR spiral model are explained below. Afterwards, the four sequential process phases are explained.

The iterative BiSLR spiral model starts with a protocol definition phase. In this phase, main and specific research questions, research objectives, and inclusion and exclusion criteria are defined [15]. We omit the definition of main and specific research questions. The defined research objectives can be found in Section I. Inclusion and exclusion criteria are defined in a later paper filter phase. During the reference collection phase, the iterative BiSLR spiral model suggests to

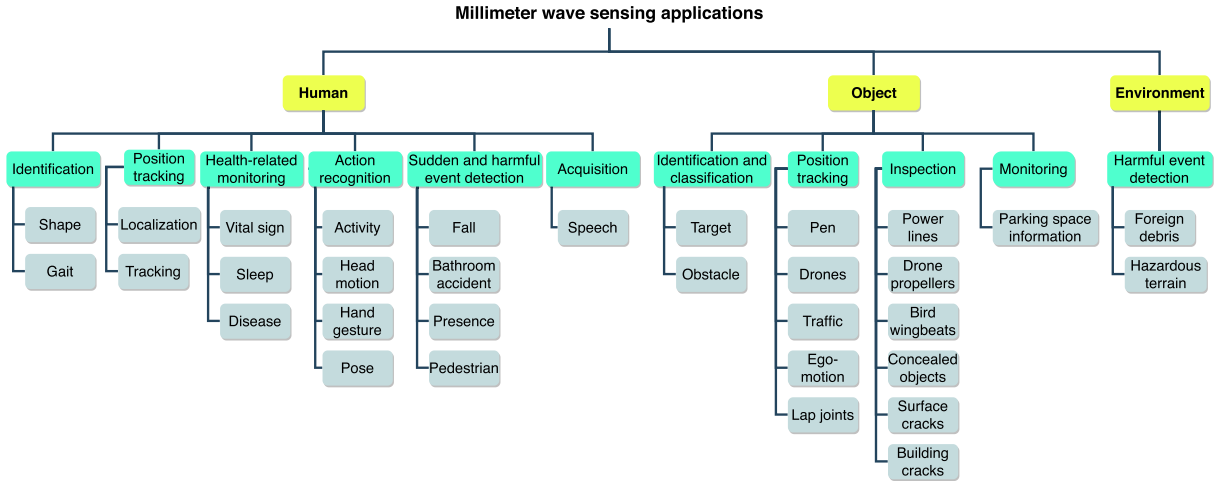


Fig. 2. Taxonomy of millimeter wave sensing applications.

select scientific databases, develop and evaluate search keywords, and iteratively repeat the phase with different scientific databases and search keywords in case the research objectives cannot be reached [15]. We initially performed rigorous reference collection during a search phase. Later during the pipeline building block definition phase, several small search phases were performed iteratively as well. In the data evaluation phase, the iterative BiSLR spiral model performs title, abstract, diagonal, and full-text reading to determine if papers should be included in the list of papers used for data collection. The iterative BiSLR spiral model finishes with data collection and narrative synthesis [15]. We used title, abstract, and diagonal reading for, and collected data during, the explorative paper classification phase and used full-text reading, collected data, and performed narrative synthesis during the application pipeline building block definition phase.

In the first phase, we searched for relevant papers published and indexed in digital libraries such as IEEE Xplore, ACM, Elsevier, and SPIE. Because millimeter wave technology is used across different sensing applications that use different terminologies, we used diverse search terminologies. Therefore, finding relevant papers required using synonyms of terms, which we categorized as exact, similar, narrower, and broader terms. Similar terms for 'millimeter wave' include 'mmwave', 'mm-wave', and 'mmw'. Narrower terms include 'v-band', 'w-band', 'Ka-band', 'Ka band', 'v band', and 'w band'. Broader terms include 'microwave', 'micro-Doppler', 'Doppler', and 'radar'. Several gesture recognition papers only refer to the specific technology used for tracking in the title. The narrower term for gesture recognition used in various papers is 'Soli'. Term combinations used for searching were made with exact, similar, and narrower terms and consisted of a millimeter wave term and application type term. Based on the application domains addressed in the papers, we made an application taxonomy that is shown in Figure 2. Application type terms were derived from this taxonomy. Term combination examples include 'millimeter wave tracking', 'mm-wave gesture', and 'w-band detection'. Broader millimeter wave terms such as 'microwave' and 'radar' resulted in too many hits that mostly included papers outside the scope of this review. Several millimeter wave sensing papers used in the review identify their content with broader millimeter wave terms. These papers were cited in the reference lists of millimeter wave papers found during the paper search.

In the second phase we classified over 140 papers found during the first phase to identify relevant information. The classification was based on the most commonly identified elements, such as application domain addressed, objectives, target group, dataset size, modeling technique, methodology, deployment environment, evaluation parameters, and opportunities and challenges. The research papers featured varying publication dates from the year 1994 up until 2020. Based on the results obtained in the explorative phase, we made an application taxonomy (see Section III) as well as an application pipeline consisting of several generic pipeline building blocks used in these papers.

The third phase was used to determine inclusion criteria for this review. After studying over 140 papers, we discovered that some papers claim to cover millimeter wave systems while they do not. In some of these papers, the carrier frequency utilized falls outside the 30 - 300 GHz band [16], [17]. We also excluded papers written in languages other than English [18]. Papers focusing on radiometry (passive sensing and imaging) [19]–[34], spectroscopy [35], [36], interferometry [37], and utilization of waveguide technology [38]–[40] were also excluded. Radiometry and spectroscopy do not focus on actively emitting and measuring effects on millimeter waves for sensing. The waveguide technology has been used to design electric probes for interacting with membrane systems, corrosion, sintering processes, etc., [38]–[40]. The technology has not been used explicitly for guiding millimeter waves to and from an antenna. Interferometry, rather than measuring effects on millimeter waves, measures wave interference using a multi-radar setup with special radar configurations [37] for sensing.

Radiometry measures electromagnetic radiation originating from humans or objects with receiver setups [41]. One group of radiometers creates passive millimeter wave images. These

images are bi-dimensional radiation maps of a scene. Several studies focus on the creation of these systems [23], [34]. In addition, several studies perform multiple object detection and tracking [20], hidden object detection [19], [21], [22], [25], [28], [31], and military target detection [26] with analytical models that take passive millimeter wave images as input data. Another group of radiometers generates one dimensional radiation profiles. In [30], a passive radiometric temperature profile is a one dimensional output voltage signal which can be generated in either power detection mode or correlation mode. In power detection mode, the output signal depends on antenna temperature which is proportional to the radiometric temperature of humans or objects within the radiometer's Field of View (FOV). In correlation mode, output signals from two receivers in power detection mode are used in a correlator to produce an output signal which does not contain objects that mostly reflect and scatter radiation rather than radiate it. The detection capability of detection algorithms that can be used with radiometric temperature profiles has also been tested [30]. Millimeter wave hardware and a prediction algorithm that can generate human body emitted energy output traces and detect weapons and explosives [24], [33] have also been created. Yujiri *et al.* [27] measure radiation temperature profiles of buried mines. Other uses of radiometry include analysis of sun brightness temperature and precipitating cloud extinction by means of analytical modeling and a sun-tracking radiometer [29], and creation of an analytical model for predicting relative humidity profiles from clear-air radiances [32].

Spectroscopy is a kind of radiometry in which the interaction between radiation and matter is measured [36]. Schmalz *et al.* [35] use gas spectroscopy to perform breath analysis. Schmalz *et al.* explain that "a typical gas spectrometer consists of a radiation source, an absorption cell, a detector, and optical elements. The radiation is transmitted through an absorption cell, which is filled with a gas at a particular pressure and impinges on a detector, which generates an output voltage or current" [35]. Spectroscopy has also been used to analyze solar system objects by probing temperature and molecular abundance in planetary atmospheres [36].

In the fourth phase, we defined the application pipeline building blocks. For every building block, we identified, reviewed, and analyzed the designed (or used) algorithms, models, and hardware, and summarized findings in tables that map to the individual papers. Using the summary tables, comparative tables and figures were formulated (see Section IV) that explain the application pipeline building blocks. Occasionally, during comparative table and figure formulation, summary table deficiencies and errors in the summary tables were discovered. These deficiencies and errors were corrected by revising the summary tables and afterwards updating the associated comparative tables and figures. This was an iterative process, during which extra papers were searched (phase 1) to cover the field as much as possible and to include all relevant papers. As a result, 25 additional papers were found (on top of the first 140 papers). These papers were inspected with the inclusion criteria (phase 2) and used during application pipeline building block definition. Therefore, 165 papers in total were analyzed during the literature review.

III. Application Taxonomy

Millimeter wave sensing applications can be classified based on the entity (i.e. human, object, or the environment) being sensed, application goal, and application context. Figure 2 represents the application domains that we have identified based on these criteria. Application types and goals in sensing humans with millimeter waves include: identification, position tracking, action recognition, health-related monitoring, sudden and harmful event detection, and acquiring speech data. Application types and goals in sensing objects include: object identification and classification, position tracking, object inspection, and monitoring information derived from object characteristics. The environment is sensed in applications whose types and goals are mainly focused on detecting harmful events related to space landing and airport runways.

IV. Application Pipeline

This section presents the five common building blocks found in application pipelines of the reviewed papers: data collection, pre-processing, feature extraction, analytical modeling, and modeling evaluation. For each building block, we review and analyze relevant papers and provide a comparative table highlighting their focus. We use these tables to identify commonalities and gaps of employed techniques and methodologies, and to highlight challenges.

A. Collection Systems

The first building block in millimeter wave sensing applications is data collection, in which millimeter wave related measurements are collected using a variety of different measurement systems. In this section, we first briefly explain different data collection approaches and their fundamentals, including suitable antenna types and designs, and then describe the variables that are measured. Hardware calibration [65], [90], [94], [96], [141], [142] and physical noise reduction [109] are considered to be outside the scope of this paper and will therefore not be addressed.

1) Suitable Antenna Types and Designs: Among the papers reporting on measurement systems, a variety of different suitable antenna types and designs have been identified that are used in these measurement systems. To prevent confusion in later sections, we decouple the concept of transmitters and receivers from transmit and receive antennas in the elaboration on suitable antenna types and designs. This means that any transmitter or receiver has a certain amount (at least one) of transmit or receive antennas depending on the antenna type and design. Suitable antenna types include, but are not limited to, horn antenna [64], [65], [68], [74], [102], [109], [116]–[119], [142]–[155], yagi antenna [149], lens antenna [59], [105], reflectarray antenna [85], [86], cassegrain antenna [88], [113], (on-chip integrated) patch antenna [49], [54], [82], [127], [128], [139], [151], [156], [157], and parabolic antenna [110], [112]. Suitable antenna designs include, but are not limited to, phased array antenna design [144], [145], [147], [158]–[162], frequency scanned antenna design [87], fan and pencil beam antenna design [119], gaussian beam antenna design [100], and omni-directional antenna design [146].

By far most measurement systems identified in the papers either use on-chip integrated patch, i.e. microstrip or printed, antennas or series fed patch antenna arrays. Even though this is not evident when looking at the suitable antenna types and designs that are actually reported, many papers list commercial measurement systems, without listing the used antenna type and design, that use on-chip integrated patch antennas or series fed patch antenna arrays. A patch antenna is advantageous compared to other antenna types because it is low cost, can easily be integrated in printed circuit boards of various sizes, and is easy to mass produce [10]. A series fed patch antenna array is an array of multiple patch antennas connected via a single feed line in series and is used to create a beam pattern with certain characteristics that cannot be created with a single patch antenna [163], [164].

The phased array antenna design is an antenna array design in which densely packed small gain antennas are phase shifted by a separate analog or digital phase shift module to create an overall high gain beam that can be steered without mechanically rotating the antennas [163]. This antenna array design is very important in millimeter wave communication systems since, due to experiencing more penentration loss than centimeter wave communication systems and solely relying on Line of Sight (LOS) propagation, transmitter and receiver beams constantly have to be re-aligned [10]. This antenna array design can also be used in ubiquitous sensing systems that rely on rotation for measuring certain variables to steer the sensing system to a sensing position of interest without experiencing noise caused by mechanical rotation. More information regarding the other antenna types and designs can be found in [10], [163].

2) Radar: Radars emit electromagnetic radiation signals, which are either reflected or scattered from targets with a smooth or rough surface, respectively. The differences between the emitted and received signals are of interest to the sensing application [163]. To understand the basic principles of the radar approaches, we consider in this section situations in which (i) there is a single object or human in the radar's FOV and (ii) no noise artifacts exist. We will deal with issues related to multiple targets, noise artifacts, increasing sensing resolution, and isolating phase shift components in Section IV-B. Table I presents an overview of papers that used radar for data collection in millimeter wave sensing applications. These papers are compared based on their carrier frequency, modulation scheme, and measured variables, each of which is explained below:

- The carrier frequency indicates what part of the millimeter wave frequency band has been explored and to what extent. It is also strongly correlated to the attainable measurement distance in two ways. Firstly, the higher the carrier frequency, the faster the attenuation of the transmitted signal. Secondly, there are exceptions to the first correlation in the form of several attenuation peaks. For example, applications operating at 60-69 GHz do so under signal absorption of the oxygen in the atmosphere. This allows, for example, mobile phone hand gesture recognition in compact populated areas. Human position tracking applications typically operate at around 70-79 GHz. This frequency interval, compared to 60-69 GHz, allows measurements at greater distances [165].
- The modulation scheme dictates which variables the radar can measure and how. Certain modulation schemes have been used to test joint communication and sensing [140].
- Measured variables strongly depend on an application type and its objective. Identification, position tracking, action recognition, etc. rely on range, velocity and angle information [46], [47]. Health monitoring and speech acquisition rely on phase and Doppler shift information in time [62], [109], [113]. Object identification and classification employ Intermediate Frequency (IF) signal variations across different IF channels [137], [139].

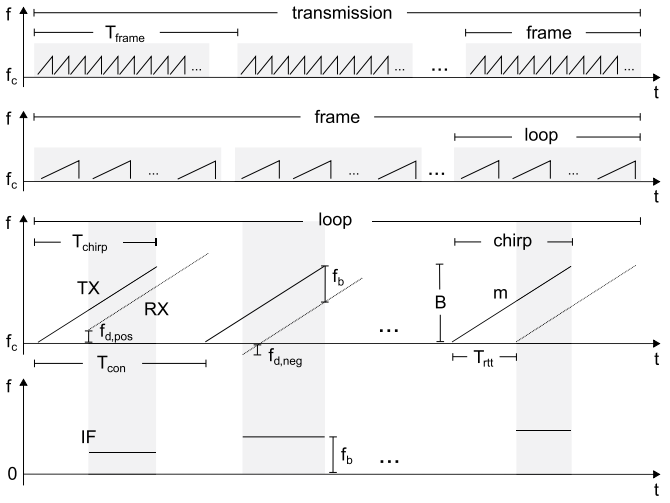


Fig. 3. FMCW radar frequency sawtooth modulation across time. Recreated from [168]. Changes include addition, removal, and alteration of variables, and removal of the figure part referred to by the authors as the 'lower half'.

The most widely used radar type in millimeter wave sensing applications is the Frequency Modulated Continuous Wave (FMCW) radar, which was commercialized recently with utilization of the Integrated Circuit (IC) technology [45], [166], [167]. Almost all papers found that use this radar modulate the signal frequency according to a sawtooth pattern, which converts the signal into a continuous stream of chirps [168]. A minority of papers modulate the signal frequency according to a triangular pattern [85], [100]. An explanation of this frequency modulation pattern can be found in [163]. Sawtooth frequency modulation is depicted in Figure 3.

A chirp is part of a trigonometric function across a limited time window with length T_{chirp}. In this time window, the signal frequency is increased linearly across a bandwidth B with slope m using a voltage-controlled oscillator. The transmitted chirp at transmitter TX is then reflected or scattered from a target and received at receiver RX after a round trip time T_{rtt}. This causes the received chirp to show a frequency deviation f_b compared to the transmitted chirp at a specific time instant. An IF signal (a.k.a. beat or baseband signal) operating at constant frequency deviation f_b can be obtained using a down-conversion mixer and low-pass filter in sequence

TABLE I

SUMMARY OF DATA COLLECTION DEVICES USING RADAR TECHNOLOGY. THE MODULATION SCHEME AND MEASURED VARIABLE ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-A.2

	Carrier frequency (GHz)										Modulation scheme					Measured variables									
	30-39	40-49	50-59	60-69	70-79	80-89	90-99	120-129	160-169	220-229	P	RN	BPSK	CW	FMCW	R	v_r	θ	φ	$\Delta\phi_t / f_d$ across time	Reflective intensity	Attenuation power	IF amplitude	Local min/max received signal amplitudes	RSS
[42]				X											X	X	X								
[43]					X										X	X	X	X							
[44]					X										X	X	X		X						
[45]				X									X		X	X	X								
[46]					X										X	X	X	X	X						
[47]					X										X	X	X	X	X						
[48]					X										X	X	X	X		X					
[49]				X											X	X	X								
[50]					X										X	X	X	X							
[51]											X					X	X	X							
[52]					X						X					X	X	X							
[53]					X										X	X	X	X		X					
[54]														X				X		X					
[55]				X											X	X	X	X							
[56]				X											X	X		X							
[57]					X										X	X				X					
[58]					X										X	X		X							
[59]							X													X					
[60]										X				X						X					
[61]					X						X					X	X	X							
[62]					X										X	X		X		X					
[63]				X											X	X	X								
[64]	X													X						X					
[65]							X							X						X					
[66]																X	X								
[67]																X	X		X						
[68]							X							X							X				
[69]					X										X										
[70]					X										X	X									
[71]					X										X	X	X	X							
[72]					X										X	X			X						
[73]					X										X	X									
[74]					X		X								X	X									
[75]					X										X					X					
[76]					X										X	X	X	X	X						
[77]					X										X	X	X	X							
[78]					X										X		X								
[79]														X						X					
[80]				X																					
[81]				X											X	X	X			X					
[82]			X																	X					
[83]					X										X	X		X							
[84]					X											X									
[85]					X										X	X		X	X						
[86]					X										X	X		X							
[87]									X		X						X								
[88]							X				X					X									
[89]					X										X	X	X	X	X						
[90]					X										X	X	X	X							X
[91]				X												X	X								
[92]					X										X	X				X					
[93]					X										X	X	X	X							
[94]	X			X											X	X		X							
[95]					X										X	X	X	X							
[96]				X											X	X	X	X							
[97]					X										X	X	X	X	X						
[98]					X										X	X		X							
[99]					X										X	X		X							
[100]							X								X	X									
[101]							X				X					X		X							
[102]												X				X				X					
[103]					X										X	X	X	X	X						
[104]	X										X														
[105]							X							X						X					
[106]					X										X	X		X		X					
[107]						X									X	X				X					
[108]					X										X	X				X					
[109]				X										X						X					
[110]	X													X						X					
[111]		X												X						X					
[112]	X																			X					
[113]							X							X						X					
[114]					X										X	X	X	X							
[115]					X										X	X	X	X							
[116]							X							X				X							
[117]							X								X	X									
[118]	X															X		X							
[119]							X								X	X	X								
[120]					X										X	X	X	X	X						
[121]					X										X	X	X	X							
[122]								X							X	X	X					X			
[123]					X										X	X	X								
[124]					X										X	X	X	X	X			X			
[125]				X												X	X								
[126]				X											X	X	X	X							
[127]				X											X	X	X								
[128]				X											X	X		X	X						
[129]					X										X	X				X					
[130]					X										X	X	X	X	X						
[131]					X										X	X	X	X							
[132]															X	X		X							
[133]					X												X	X							
[134]					X											X	X	X							
[135]					X						X					X	X	X					X		
[136]					X										X	X		X					X		
[137]				X											X								X		
[138]				X											X	X	X	X	X						
[139]				X											X								X		
[140]					X								X			X									

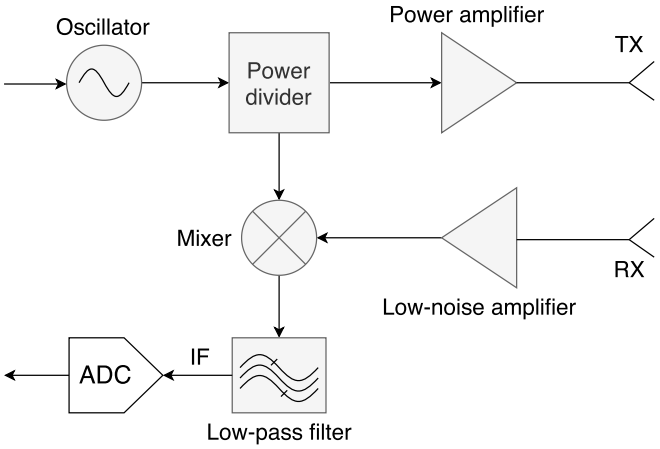


Fig. 4. Block diagram of fundamental components of CW radar. Recreated from [169].

as shown in Figure 4. This operation returns a correct output in time window $T_{chirp} - T_{rtt}$, where the transmitted and received chirps overlap [166], [169]. Given a trigonometric function $x = A\cos(2\pi f t + \phi)$ for the transmitted chirp x_{tx} and the received chirp x_{rx}, the IF signal $x_{if} = [(A_{rx} \cdot A_{tx})/2]\cos[2\pi(f_{tx} - f_{rx})t + (\phi_{tx} - \phi_{rx})]$ [170]. Using frequency deviation $f_b = f_{tx} - f_{rx}$, the range R between the radar and the target is computed using Equation 1 [169]. The symbol c denotes a constant representing the speed of light. Due to its unmodulated constant signal frequency, a Continuous Wave (CW) radar is unable to measure range.

$$R = \frac{f_b \cdot c}{2 \cdot \frac{B}{T_{chirp}}} \tag{1}$$

Radial velocity v_r of the target can be obtained by using consecutive chirps separated by time window T_{con} emitted across a loop. In case the target is in motion, the IF signal resulting from the chirps will experience a significant phase difference $\Delta\phi_t$ relative to the previous loop. The resulting frequency difference is indiscernible for small motion changes. The radial velocity with one chirp per loop is computed using Equation 2 (left). The symbol λ_c refers to the radar carrier wavelength. Another approach that can be used to obtain radial velocity is to measure a frequency difference between the transmitted and the received chirps, which is commonly known as the Doppler frequency shift f_d. The radial velocity is then computed using Equation 2 (right). The symbol f_c refers to the radar carrier frequency. In case of FMCW modulation, the carrier frequency and the wavelength refer to the starting frequency and the wavelength of a transmitted chirp [166], [169]. Other modulation techniques can measure the phase difference, the Doppler shift, and the radial velocity by using the same techniques and associated equations. For example, Pulse (P) modulation can measure a Doppler shift between a pair of transmitted and received pulses [171]. The authors in [64] recognize that when using a bistatic radar configuration, in which the transmitter and the receiver are placed perpendicular to each other, Doppler shifts due to yaw, pitch, and roll movements of the head are more clearly

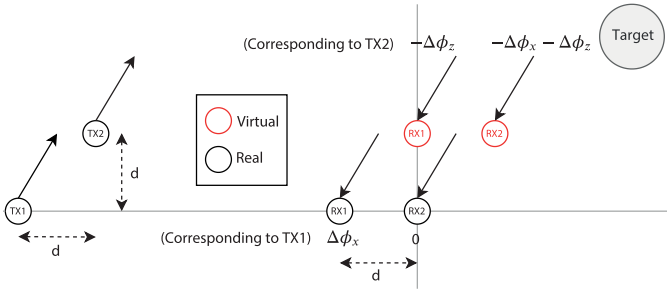


Fig. 5. Uniform linear MIMO radar layout including two transmitters and receivers. Recreated from [172].

distinguishable. For radars using an In/Quadrature (I/Q)-phase mixer, the phase difference $\Delta\phi_{t,\mathrm{I/Q}} = \tan^{-1}(x_q/x_i)$. Symbols x_q, x_i denote the quadrature-phase and in-phase IF signals coming from the I/Q mixer [59], [65], [113].

$$v_r = \frac{\lambda_c \Delta\phi_t}{4\pi T_{con}} \quad v_r = \frac{f_d \cdot c}{2 \cdot f_c} \tag{2}$$

The Angle of Arrival (AoA) of the reflected or scattered signal in both the elevation dimension and azimuth dimension (bearing angle) is calculated based on Multiple Input Multiple Output (MIMO) radar principles [130], [172], [173]. Rather than having one transmitter and one receiver, multiple transmitters and receivers are used. Almost all radars explained in the papers use a uniform linear layout. An example can be found in Figure 5. A minimum redundancy layout [58] and a uniform circular [118] layout have also been explored in the literature. Real and virtual receivers, spaced according to a matrix structure on a horizontal surface, observe signal reflections coming from a target with a distance d apart from one another along the front and or right direction. The signal reflection observed at a given receiver has to travel a certain distance further or shorter along a vertical and or horizontal direction in relation to one specific receiver in the array to arrive at the receiver. This distance in the vertical and or horizontal direction is unique to every receiver. The difference in distance between two neighboring receivers in the vertical or horizontal direction directly corresponds to a phase shift component denoted by $\Delta\phi_{\in\{x,z\}} = [2\pi d \sin(\theta)]/\lambda_c$. A phase shift component overview is presented in Figure 5. The azimuth or elevation angle is calculated by using Equation 3 [166] with the phase shift component in the horizontal or vertical direction. Some approaches use multiple radars. For example, [85], [103] use azimuth angles from two different radars, in which one was rotated 90 degrees counter-clockwise compared to the other. The rotation makes the azimuth angle correspond to the elevation angle. The data collection systems in [85], [86], [132] rotate the radar to obtain angle measurements.

$$\theta = \sin^{-1}\left(\frac{\lambda_c \Delta\phi_x}{2\pi d}\right) \quad \varphi = \sin^{-1}\left(\frac{\lambda_c \Delta\phi_z}{2\pi d}\right) \tag{3}$$

A Random Noise (RN) radar transmits noise signals. Binary Phase-shift Keying (BPSK) radars are normally used for communicating a bit stream. The bit value determines the phase offset ϕ (0 or π) of the transmitted signal x_{tx}. Both radar types use cross-correlation between the transmitted and received

TABLE II
SUMMARY OF DATA COLLECTION DEVICES USING ACTIVE IMAGING APPROACH. THE MODULATION SCHEME, MEASURED VARIABLE AND SCANNING METHOD ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-A.3

	Carrier frequency (GHz)							Scanning method				Modulation scheme			Measured variables									
	30-39	50-59	60-69	70-79	80-89	90-99	260-269	Mechanical line array	Mechanical plane	Mechanical, single radar	Free space	P	AM	FMCW	Scattering signal power distribution	Scattering coefficient	Scattering signal	Visibility function	IF signal	v	mean elevation	roughness property	Reflection intensity
[141]	X							X						X		X								
[174]											X	X						X						
[156]						X			X									X						
[175]						X				X				X					X					
[176]				X							X	X							X					
[177]						X					X	X							X					
[161]	X					X					X	X									X	X	X	
[149]			X					X					X		X									
[151]				X				X					X											X
[142]		X								X				X		X								
[178]					X					X				X					X					
[152]				X				X						X					X					
[154]							X			X		X					X							
[179]				X						X				X					X					

signals. This cross-correlation relates to a time delayed version (delayed by $2R/c$) of the transmitted signal to obtain an auto-correlation function. When plotted, this function shows a peak that corresponds to a target's range [102], [140]. The noise radar in [102] measures a Doppler frequency shift based on a tone embedded into the noise signal.

Pulse radars emit a powerful high gain signal pulse for a short time period. This period is called the pulse width. Afterwards, the radar waits for a given rest time to receive reflections before emitting another pulse. The range can be determined based on the round trip time between the emitted pulse and received reflection of this pulse in the rest time: $(c \cdot T_{rtt})/2$. In addition to MIMO radar principles, the azimuth and elevation angles can also be retrieved from the radar's own directivity compared to a baseline direction denoting north or the ground respectively [163]. The authors in [87] measure mean velocity based on the correlation between two pulses under an unambiguous pulse interval for a sloped terrain. Pulse radars normally work well in long distance applications [163]. One exception to this observation was found in [104], in which breast cancer detection was performed successfully. The authors in [45] observe that pulse radar does not provide enough resolution for tasks requiring high resolution measurements such as gesture recognition.

3) Active Imaging: Active imaging is a specific type of radar-based data collection approach that obtains measurement results in the form of map-like images [163]. The radar scans a given area. For every position, the radar emits a signal, measures the reflections returned, and based on a measured variable assigns, for example, a color or gray-scale value to the pixel corresponding to that position. Most active imaging research between 2002 and 2014 was focused on data collection devices for monitoring cracks in civil infrastructures, detecting concealed objects on the human body, and measuring hazardous landing terrain [87], [161]. Review papers on active imaging used for detecting explosives and monitoring civil infrastructures include [41], [180]. Active imaging recently received attention again in the context of map-like image deep learning and applying existing techniques using a cost effective method in commercial applications [156], [176], [177], [181]. Commercial applications include parking space monitoring and analyzing objects in enclosed packaging for food quality control and non-invasive fault detection. Table II is an extension of previous reviews [41], [180], providing an overview of active imaging approaches used for data collection in millimeter wave sensing applications. In addition to the carrier frequency, the modulation scheme, and the measured variables, we consider the scanning method here as well.

The scanning method defines how the measurement system moves from one position to another to obtain, for example, a color or gray-scale value for every pixel in the map-like image. One such method is mechanical scanning where radar devices are mechanically moved in fixed directions using a rail system. The literature covers three types of mechanical scanning. In line array scanning, a 1-dimensional array of antenna's on a straight line scan in a single direction back and forth (i.e., up/down or right/left) [141], [149], [151], [152]. In plane scanning, antennas arranged into a 2-dimensional plane scan the target from different distances by moving perpendicular to the plane [156]. Some papers employ scanning a target by moving a small radar system in two directions on a virtual plane [142], [154], [175], [178], [179]. In the so called free space scanning method, the forward motion of a radar connected to a moving device, such as a drone or planetary lander, is used to construct images. Sensing application pipelines using this type of scanning either use a large antenna array to cover big areas [161] or a limited number of ingrained low-cost radar device antennas [176], [177].

The free space scanning methods found in the literature are commonly referred to as Synthetic Aperture Radar (SAR). A few millimeter wave application papers present application pipelines containing data collection devices that do not belong to active imaging by means of free space scanning and mention use of SAR [152] or SAR pre-processing methods [154], [178], [179]. We consider these data collection devices not part of SAR since SAR originates from Side Looking Airborne Radar (SLAR) (a free space scanning method). SAR provides a solution to the impractically long antenna or use of extremely short wavelengths, resulting in severe atmospheric attenuation, required for sufficient azimuth resolution in SLAR images. SAR essentially *synthesizes* a very long antenna to obtain high resolution images [182]. More information regarding SAR can

TABLE III
SUMMARY OF DATA COLLECTION DEVICES USING SPATIAL SWEEPING APPROACH. THE MODULATION SCHEME, MEASURED VARIABLE AND SWEEPING METHOD ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-A.4

	Carrier frequency (GHz)			Sweeping method					Measured variables							
	30-39	60-69	70-79	Fixed	Rotational	Radio Tomography	V2I	Base Station Exchange	CIR	RSS	Phase	Reflection loss	f_d across time	ToF	AoD	AoA
[158]		X		X					X							
[159]		X		X						X	X					
[143]		X		X	X					X						
[160]		X			X					X						
[184]			X					X						X	X	
[144]		X		X	X					X		X				
[145]		X						X								X
[146]		X					X						X			
[147]		X		X	X					X		X				
[148]		X			X					X						
[162]		X						X						X		
[150]		X				X				X						
[185]	X						X							X	X	X
[153]		X						X		X						
[186]	X							X		X						X
[155]		X			X				X							

be found in [182], [183]. Raw SAR images are severely out of focus since they do not represent spatial information correctly yet and therefore need additional pre-processing which is elaborated on in Section IV-B.2.

The IF signal is introduced in Section IV-A.2. Radars using P modulation obtain an IF signal through a concept called pulse compression [177], [183], during which the pulse frequency is modulated in a similar pattern as FMCW modulation. Pulse compression allows transmission powers comparable to a pulse with a long pulse width while simultaneously retaining the range resolution which is only attainable with a short pulse width in the context of SAR. Low transmission power results in low receiver signal detectability and measurement precision [187]. The IF signal is either used directly in pre-processing [177] or its amplitude [176]–[178], magnitude [175], [179], and or phase values [175]–[178] are used to color the image or for further pre-processing. The authors in [152] measure a lap-joint position with an IF signal component. The component accounts for amplitude variation based on radiation characteristics and the target's shape and a phase variation based on the target's position at the center frequency.

Other less frequently measured variables include a visibility function and a scattering coefficient. Visibility is defined as a correlation value between two non-directive receivers. Both mechanical [156] and free space [174] approaches have been used for measuring these variables. The scattering coefficient can be considered as an energy ratio between transmitted and received energy [142]. When a signal scatters due to a rough surface, more energy will be observed by the receiver compared to when most energy is reflected away from the receiver due to a smooth reflection surface [183].

4) Spatial Sweeping: Spatial sweeping is an approach where signal response metrics measured in the context of the millimeter wave radio communication between sender and receiver systems are analyzed. Changes in the measured metrics are caused by the interaction between the target of interest to be sensed and the communication signals. Spatial sweeping refers to the observation that almost all collection systems in Table III have transmitter and or receiver components that spatially move and rotate during communication. Movement and rotation naturally increase the system's sensing FOV. Table III provides an overview of spatial sweeping approaches used for data collection in millimeter wave sensing applications. In addition to the carrier frequency and measured variables, we consider the used sweeping method here as well.

The sweeping method defines how senders and receivers move and rotate during communication with each other while at the same time being used for measurement procedures. Table III shows that several sweeping methods are quite generic while others are more specific. This is because certain systems are tested for specific real-life contexts while other systems are tested for determining feasibility of, for example, rotational and fixed communication for sensing. A variety of sweeping methods exist, such as rotational sweeping, fixed sweeping, and radio tomography. Most rotational sweeping methods rely on a transmitter rotating itself to transmit a signal at beam angles within a given area. During this time, receivers make a variable measurement at one angle of arrival if the angle is not in line with the current beam angle. This process is repeated sequentially for a number of angles of arrival to form a measurement matrix [143], [144], [147], [148]. The authors in [160] use a robot that moves and rotates. The authors in [155] use a fixed position where both the transmitter and receiver are located. Measurements are taken at a given location and associated angles [155], [160]. In fixed sweeping, position and direction of both transmitters and receivers are fixed. The authors in [158] put a transmitter and receiver in direct line-of-sight, while the authors in [159] use a quasi-omni transmit antenna to cover a measurement area. Certain rotational sweeping methods fall back to fixed sweeping by rotating a transmitter and receivers towards a fixed angle to perform additional measurements [143], [144], [147] while the angle remains fixed. Radio tomography uses a grid network of transceivers to cover a measurement area [150]. Vehicle to Infrastructure (V2I) communication relies on fast Dedicated Short Range Communication (DSRC) between a stationary Roadside Unit (RSU) and On-Board Unit (OBU) inside a moving vehicle [188]. Base station exchange covers a measurement area via a set of moving receivers [145], [153], [184], [186].

5) Gaps and Challenges: From Tables I, II, and III, a number of interesting observations can be made. Firstly, most data collection systems operate in the 60-79 GHz frequency

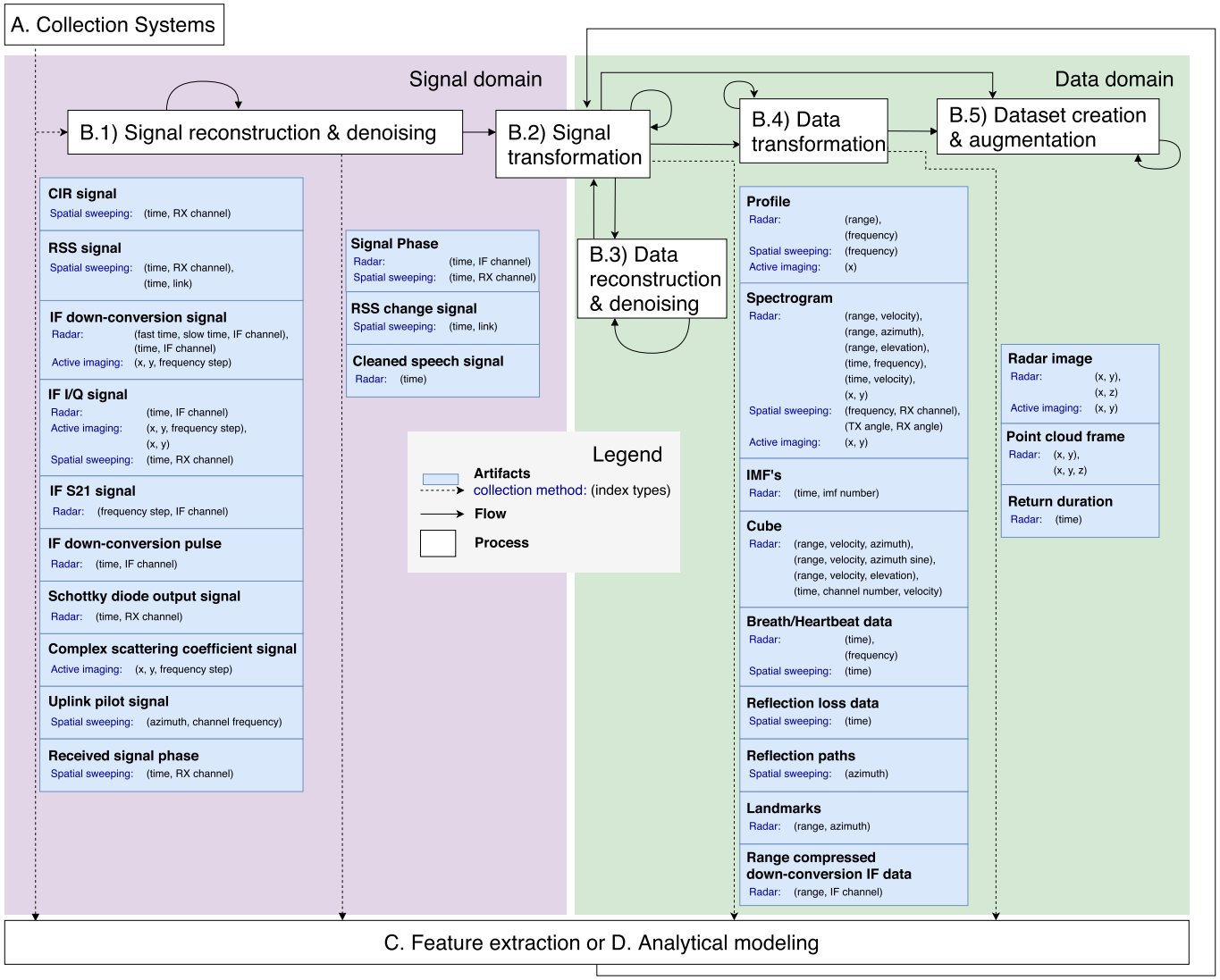


Fig. 6. Abstract pre-processing pipeline diagram. It does not show spatial sweeping across time (see Section IV-A.4). It also omits artifact parts extracted with pre-processing and additional artifact features that are normally not extracted together with the artifact. Artifact part examples include profile, spectrogram, cube bins and radar image parts [189]. Feature examples include velocity [124], [138] and intensity [124] with point cloud frames and vehicle location [176], [177], empty parking site [176], [177], obstacles [177], length-width ratio [176], barycenter location [176], and intensity [103] with radar images.

band. This leaves ample room for performing research and data collection on the 30-59 GHz and 80-299 GHz bands. Secondly, most radar systems stick to variables that are delivered out of the box by commercialized systems. These include range, radial velocity, and AoAs. It remains to be seen and investigated whether additional variables such as IF signal amplitude variation across time and Received Signal Strength (RSS) provide increased sensing accuracy to a wide variety of applications.

B. Pre-Processing

By using the collection systems mentioned in Section IV-A, discrete signal vectors in the time domain are sampled via an Analog-to-Digital Converter (ADC). For the majority of sensing applications, these raw signal vectors cannot be used to infer information of interest to the sensing application. Therefore, raw signal vectors are passed through a pre-processing pipeline to extract data types that can be used to infer information. An abstract pipeline overview can be found in Figure 6. This section reports on the pre-processing methods that are mentioned in the papers. Simulation with a ray tracing tool [145], [162], [184], [186], single-radar simulation [140], multi-radar simulation [101], imaging simulation [156], and radar data modeling [174], [201], [202] are considered to be outside the scope of this review. The pipeline overview does not accurately represent the exact flow that every methodology follows and position in the pipeline where artifacts are extracted since parts are skipped, not reported, etc. It covers the predominant processes, flows, and places where artifacts are extracted. The pre-processing space has been divided into two domains: the signal domain and data domain. The signal domain encompasses all pre-processing methods that are applied directly on the ADC sampled signal vectors. Once signal vectors are passed through a signal transformation method for the first time, the artifacts resulting from these methods are referred to as data and are subsequently processed in the

TABLE IV
APPLICATION PIPELINES THAT DIRECTLY RETRIEVE ARTIFACTS FROM THE COLLECTION SYSTEM HARDWARE OR USE A PRE-PROCESSING PIPELINE BUILDING BLOCK RETRIEVED FROM A PAPER

	Artifact(s)	(index types)	Collection method/Paper
[42]	Spectrogram	(range, velocity)	Radar
[47]	Point cloud frame	(x, y, z)	Radar
[48]	Spectrogram	(range, velocity)	Radar
[49]	IF I/Q signal, Spectrogram	(time, IF channel), (range, velocity)	Radar
[51]	R, θ, and v_r	Undefined	Radar
[52]	Long. and lat. R and v	Undefined	Radar
[56]	Profile θ	(range) Undefined	Radar
[61]	Long. and lat. R and v	Undefined	Radar
[146]	Doppler shift data	(time, TX channel)	[190]
[63]	Spectrogram	(range, velocity)	Radar
[181]	Radar image	(x, z)	Active imaging
[191]	Radar image	(x, z)	Active imaging
[67]	R, φ, and v_r	Undefined	Radar
[192]	Radar image	(x, z)	Active imaging
[193]	Radar image	(x, z)	Active imaging
[69]	Heartbeat data	(time)	Radar
[189]	Radar image	(x, z)	Active imaging
[72]	Spectrogram	(x, z)	[194]
[176]	Radar image	(x, y)	Active imaging
[177]	Radar image	(x, y)	Active imaging
[77]	Point cloud frame R, v_r, θ, and power	(x, y) Undefined	Radar
[81]	Spectrogram	(range, velocity)	[195]
[83]	Spectrogram	(range, azimuth)	Radar
[84]	Profile	(range)	Radar
[87]	Range compressed down-conversion IF data	(range, IF channel)	Radar
[89]	Point cloud frame	(x, y, z)	Radar
[91]	Spectrogram	(range, velocity)	Radar
[95]	Point cloud frame v_r	(x, y, z) Undefined	Radar
[97]	Point cloud frame	(x, y, z)	Radar
[99]	R and θ	Undefined	Radar
[196]	Reflection intensity	Undefined	Radar
[102]	Profile Spectrogram	(range) (time, frequency)	[197]
[104]	Radar trace	Undefined	Radar
[178]	Spectrogram	(x, y)	Active imaging
[105]	Heartbeat data	(time)	[198]
[112]	Spectrogram	(time, frequency)	[199]
[185]	Positioning data	Undefined	Spatial sweeping
[115]	R and θ	Undefined	Radar
[153]	RSS signal	(time, RX channel)	Spatial sweeping
[117]	Profile Waterfall chart	(range) (time, range)	Mono-pulse pipeline
[118]	Dynamic (time-varying) signal components	Undefined	Radar
[120]	Point cloud frame v_r	(x, y, z) Undefined	[200]
[124]	Point cloud frame v_r and intensity	(x, y, z) Undefined	[166]
[125]	Spectrogram	(range, velocity)	Radar
[131]	Spectrogram	(x, y)	Radar
[133]	v_r and θ	Undefined	Radar
[134]	R, v_r, and θ	Undefined	Radar
[135]	R, v_r, θ, and power amplitude	Undefined	Radar
[136]	R, θ, and power amplitude	Undefined	Radar

data domain. Several pre-processing methods reported in this section are based on analytical modeling. Because these models are used to execute pre-processing tasks, i.e., not used to extract higher level information relevant to the millimeter wave application, they are considered to be pre-processing methods. Several application pipelines directly retrieve artifacts from the collection system hardware or use a pre-processing pipeline retrieved from a paper. The papers associated to these pipelines do not elaborate on the pre-processing methods that are executed prior to retrieving the artifacts. These pipelines and associated papers are summarized in Table IV. Several papers elaborate on additional pre-processing methods that are executed after retrieving the artifacts. These methods are explained in the respective pre-processing subsections.

1) Signal Reconstruction and Denoising: Signal reconstruction and denoising resolve signal corruption and spectral leakage. Signal corruption refers to observing a sampled signal that is drastically different from its theoretical definition. Changes are caused by superimposed noise from unwanted stationary and systematic reflections from a target and nearby objects [59], [70], [104], [110], [112], [113], [118], [144], [148], [150], [157], [159], high frequency static noise (noise that is concentrated around, and remains in, a high frequency range) [116], [144], [148], phase wrapping around a certain value that causes signal jumps [59], [65], [70], DC offset [116], [144], [148], hardware noise [60], [110], [112], [113], harmonic noise [110], [113], and channel noise [113]. Spectral leakage refers to non-zero values that show up in a signal's frequency profile at frequencies other than the frequency components actually present in the signal after signal transformation. Signal reconstruction and denoising resolve signal corruption and spectral leakage differently.

Signal reconstruction artificially creates or alters a signal with help from a construction algorithm. Phase regeneration exploits the fact that the received signal phase exhibits a periodic pattern, even though it is not linear with respect to the target's moving distance. Phase regeneration creates an artificial phase shift signal by counting phase shifts of the measured signal, creating a predefined phase shift on every count, and linking two neighboring counts with a linear increasing or decreasing trend [159]. Phase unwrapping is not explained in [59], [60], [65].

Signal denoising encompasses all methods that remove signal components from the sampled signal. A signal component, in the context of Fourier series, refers to a wave with a less complex waveform in a set of waves with a less complex waveform that reconstitute the sampled signal if summed together. Mean centering [59] is done by subtracting the geometric mean vector of a complex output trace from each sampled signal point. A moving average filter averages a number of the sampled signal points to produce a new signal point [203]. The high, low, and bandpass filters attenuate signal components from the sampled signal that have a certain frequency. The highpass filter attenuates everything below a certain cutoff frequency, the lowpass filter attenuates everything above a certain cutoff frequency, and the bandpass filter attentuates everything outside of a predefined frequency band. In the signal domain, signal component attenuation based on a certain cutoff frequency or frequency band is performed by convolution of the sampled signal with a filter's impulse response. The impulse responses of both ideal highpass and lowpass filters can be found in Equation 4 [204]. The symbol ω_0 is the angular cutoff frequency and $\delta(n)$ the delta function. In the context of convolution, the delta function can be regarded as the identity of convolution (convolution of a sampled signal with the delta function does not change the sampled signal) [203]. Line fitting estimates a demodulated IF signal phase by fitting a straight line to the phase of a lowpass filtered demodulated IF signal version and obtaining the y-intercept of the demodulated IF signal with the straight line [70]. Dual-differential Background Removal. (DDBR) takes two differentials. The first differential is computed with signal points at time $t-1$ and t and the second differential with signal points at time t and $t+1$. DDBR then adds the differentials to obtain a background canceled signal point. A standard signal window function used in combination with a signal transformation operation such as Short-time Fourier

TABLE V
SUMMARY OF SIGNAL RECONSTRUCTION & DENOISING PRE-PROCESSING METHODS DEPLOYED IN MILLIMETER WAVE SENSING PIPELINES. THE DENOISING ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-B.1

	Reconstruction		Denoising									
	Phase regeneration	Phase unwrapping	Mean centering	Moving average filter	Highpass filter	Lowpass filter	Line fitting	Bandpass filter	DDBR	Windowing	Wavelet Packet Noise Reduction	Manual background subtraction
[159]	X								X			
[55]										X		
[144]				X				X				
[59]		X	X		X	X		X				
[60]		X						X				
[62]										X		
[64]										X		
[65]		X										
[70]						X	X					
[157]					X							
[147]				X				X				
[148]								X				X
[85]										X		
[86]										X		
[96]										X		
[98]										X		
[150]												X
[104]												X
[142]										X		
[110]											X	
[112]												X
[114]										X		
[116]								X		X		
[118]												X
[119]										X		
[121]										X		
[127]										X		
[130]										X		

Transformation (STFT) can be understood as a brick-wall filter, i.e., an ideal filter [64]. Cosine-sum and adjustable windows such as Hann [85], Hamming [127], [142], Dolph-Chebyshev [86], and Kaiser-Bessel [127] windows reduce spectral leakage in a signal transformation output [205]. Wavelet packet noise reduction [110] removes signal noise by utilizing a signal enhancement scheme on a fast wavelet transformed signal representation and afterwards transforming the signal back to the time domain. The main idea behind manual background subtraction is to first obtain a sampled signal from an environment without any target and afterwards subtracting the sampled signal from a sampled signal obtained when a target is present in the environment [104], [148]. Manual background subtraction in [112], [150] averages multiple signals obtained from an environment without a target. Exponential averaging can also be used to obtain a sampled signal from an environment without a target [118].

$$h_{lpf}(n) = \frac{\sin(\omega_0 n)}{\pi n} \quad h_{hpf}(n) = \delta(n) - h_{lpf}(n) \tag{4}$$

2) Signal Transformation: Signal transformation is the most important pre-processing operation of almost every millimeter wave sensing application pipeline. Without signal transformation, many of the artifacts depicted in Figure 6 would not exist. Signal transformation refers to transforming a time domain signal, through use of mathematical operations, into a data structure that represents certain aspects of the signal in either the time or frequency domain.

The most widely used signal transformation operation is the Fast Fourier Transformation (FFT). The FFT is an algorithm that computes the discrete Fourier transformation with a computation complexity of $O(n \log n)$ instead of $O(n^2)$, where n is the input data size, and was popularized by Cooley and Tukey in 1965 [206]. In Section IV-A.2, it was explained that an IF signal operating at a certain beat frequency can be obtained when a single object or person is standing in a FMCW radar's FOV. When multiple objects or persons are standing in the radar's FOV, the sampled IF signal will exhibit a more complex waveform. The IF signal is a constitution of multiple IF components because multiple reflections are obtained by the receiver channels. These components will have different beat frequencies if objects or persons are standing at different distances. When running the sampled IF signal through a FFT operation across the fast time dimension, the resulting data structure will be a one dimensional set of complex numbers, i.e., a profile X_{if}. The explanation ignores the complex conjugates in the set of complex numbers [166]. In-phase and quadrature-phase IF signals can be combined into a single complex signal [82] and used in the FFT operation. The magnitude of several complex numbers will show a peak compared to the other complex numbers. The associated frequency index of these complex numbers directly corresponds to a range value as indicated in Equation 1. Range resolution, i.e., minimum required distance separation between persons or objects, has a direct relation with the bandwidth size across which the chirp frequency is modulated [166].

A mathematical definition of the discrete Fourier transformation can be found in Equation 5. The exponential term in the definition is the exponential part of Euler's formula which describes rotation along a circle with a radius of 1 in the complex plane. Multiplying IF signal x_{if} with the exponential part of Euler's formula results in a rotating set of complex numbers of which the radius is scaled according to the IF signal. The IF signal is essentially 'wrapped' around a circle. The minus sign allows rotation in the clockwise direction and 2π indicates rotation along a circle's full circumference. The symbol n serves as an index symbol and k is the wrapping frequency. To explain the wrapping frequency, an IF signal

TABLE VI
SUMMARY OF SIGNAL TRANSFORMATION PRE-PROCESSING METHODS DEPLOYED IN MILLIMETER WAVE SENSING PIPELINES. THE ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-B.2

	Time domain								Frequency domain									
	Reflection path extraction	Impulse response convolution	EMD	Synchronization	Surface normal calculation	Peak detection	Reflection loss extraction	Range compression	FFT	STFT	Beamformer	MLE	MSER	Visual saliency detection	Landmark extraction	Pulse integration	POSP	Frequency compression
[158]									X									
[141]									X								X	
[43]									X									
[44]									X									
[45]									X									
[46]									X									
[50]									X		X							
[53]									X									
[160]	X																	
[55]									X		X							
[57]									X									
[58]									X		X							
[144]						X	X		X									
[60]										X								
[62]									X		X							
[64]										X								
[65]										X								
[66]									X									
[68]						X												
[70]									X									
[71]									X									
[73]		X																
[147]						X	X		X									
[176]														X				
[177]													X	X				
[74]									X									
[75]									X	X								
[76]									X									
[78]									X	X								
[79]			X															
[80]									X									
[81]									X									
[82]									X									
[161]								X										
[85]									X									
[86]									X									
[88]									X							X		X
[90]									X									
[92]									X	X								
[93]									X	X								
[94]									X									
[96]									X		X							
[98]									X		X							
[100]									X									
[102]										X								
[103]									X									
[151]				X														
[142]									X									
[178]					X													
[106]			X						X							X		
[107]									X									
[108]									X									
[109]									X									
[113]										X								
[152]									X									
[114]									X		X							
[116]									X	X								
[118]												X						
[119]									X	X								
[121]									X									
[154]									X									
[122]									X									
[123]									X									
[126]									X									
[127]									X									
[128]									X							X		
[129]									X	X						X		
[130]									X							X		
[179]									X							X		
[131]											X							
[132]															X			
[138]									X									

with a finite length of 2 seconds is assumed. A wrapping frequency of 2 Hz indicates that the IF signal is wrapped around a circle with 0.5 seconds per rotation. Therefore, an IF signal with a finite length of 2 seconds makes 4 full rotations around a circle in total. The sum is used to compute the center of mass of the wrapped IF signal. When the wrapping frequency matches the frequency of an IF signal component, the center of mass deviates far more from the origin compared to a situation in which the wrapping frequency does not match any IF signal component. More information can be found in an explanation given by Sanderson [207].

$$X_{if}(k) = \sum_{n=0}^{N-1} x_{if}(n) e^{-\frac{2\pi j}{N} kn} \tag{5}$$

Radial velocity, according to Section IV-A.2, can be determined with multiple chirps emitted across a loop. The profiles originating from the chirps are stacked on top of each other to form a two dimensional data structure made up of complex numbers. The newly introduced dimension is called slow time (a.k.a. Coherent Processing Interval (CPI)). Across slow time, the complex number's phase, due to the motion of persons or objects, rotates at a constant rate. The phase of every complex number is a sum of phases corresponding to different objects or persons [166]. After FFT operations have been performed across slow time at every frequency index, a spectrogram (a.k.a. heatmap) indexed by frequency and phase difference is obtained. The phase difference corresponds to a radial velocity value according to Equation 2. Magnitude peaks in the spectrogram correspond to objects or persons with a certain range travelling at a given radial velocity. Radial velocity resolution, i.e., minimum required radial velocity separation between persons or objects, has a direct relation with the number of chirps per loop [166].

AoA, as explained in Section IV-A.2, is determined through MIMO radar principles. The explanation assumes that spectrograms originating from the real and virtual receiver channels in Figure 5 are available. Transmit techniques required to obtain spectrograms from real and virtual receivers channels are elaborated on in Section IV-B.3. Azimuth and elevation angles are computed in separate sets of FFT operations. To determine the azimuth angle, the spectrograms originating from real receiver channels RX1 and RX2 are stacked to form a three dimensional data structure. The newly introduced dimension is called IF channel. Across IF channel, like with determining radial velocity, the complex number's phase rotates. This is due to extra distance which has to be traversed by a signal reflection. After performing FFT operations for every element in the spectrogram across IF channel, a data cube is obtained. Magnitude peaks in the cube correspond to objects or persons with a certain range, radial velocity, and phase shift in the horizontal direction. The phase shift corresponds to azimuth angle according to Equation 3 (left) [166]. When looking at Figure 5, one may assume that the same FFT operations are valid for elevation angle on spectrograms from real receiver channel RX2 and virtual receiver channel RX1. In bigger uniform linear layouts, more signal transformation steps are required to retrieve the elevation angle. A phase shift isolation technique for a bigger uniform linear layout is presented in [208]. The main idea is that after several FFT operation steps the elevation phase difference can be isolated by phasor (another way to represent a complex number) multiplication. AoA resolution, i.e., minimum required angle separation between persons or objects, has a direct relation with the number of transmitter and receiver channels used in the MIMO radar layout [166].

Other uses of the FFT operation include retrieving Channel State Information (CSI) frequency domain profile [158], conversion of complex scattering coefficient signal to frequency domain [141], aid in applying bandpass filters to isolate signal components [107], [144], [147], conversion of phase data to frequency domain [62], [108], conversion of RSS variance to frequency domain [147], inverse FFT on IF S21 or complex scattering coefficient signal to retrieve range profile [74], [142], FFT on IF signal coming from CW radar to determine breath/heartbeat data [82], [109], dimension reduction [152], filtered frequency index dominant frequency determination [116], FFT and inverse FFT for image reconstruction [154], [179], and retrieving range profile from uplink pilot signal [154]. The main difference between FFT and STFT is that STFT separates FFT operations in chunks across time. STFT has been used to retrieve phase information across time [60], [65], for creation of time and frequency spectrograms [64], [113], [116], and to create range and micro-velocity or micro-Doppler spectrograms from a stack of range profiles [75], [78], [92], [93], [129] or directly from a signal containing Doppler information [102], [119]. Micro-velocity spectrograms are used to analyze fine-grained velocity features. These features can be attributed to for example arm swinging while a person is walking or presence of something in the context of strong background noise. These fine-grained features help to compute new types of information such as drone blade rotation [116] or allow analytical models to better predict information of interest. Raja *et al.* [64] denote that STFT is well suited for identification of movement direction, time of occurrence, and duration from a RSS signal.

The beamformer is an estimation technique that can be used to calculate a one dimensional set of complex numbers associated to a user-defined AoA. It is used as an alternative to the AoA FFT. The sensing application pipelines either use the Minimum Variance Distortionless Response (MVDR), i.e., the Capon beamformer [50], [55], [96], [98], [131] or do not specify the beamformer type [58], [62], [114]. The Capon beamformer computes the set with a steering vector and covariance matrix. The covariance matrix is a model describing spatial and frequency domain IF signal characteristics. The steering vector represents the phase rotation due to extra distance, which has to be traversed by a signal reflection at receiver channels in an array in phasor notation. Reflection path extraction is used to iteratively extract RSS signal components from the signal at the receiver associated to a path between reflector and receiver [160]. Empirical Mode Decomposition (EMD) decomposes a signal into several Intrinsic Mode Functions (IMFs) [79]. The output from both methods stays in the time domain. Matsuguma and Kajiwara [73] retrieve a range profile based on a convolution in time between an IF pulse, i.e., single chirp, and impulse echo response. Oka *et al.* [151] synchronize a Schottky diode output signal with an encoder distance signal to obtain an intensity image. Pawliczek *et al.* [178] perform surface normal calculation with a phase data spectrogram to improve defect visualization. Peak detection is used to count breath/heartrate from a filtered time-series RSS signal [144], [147] or to determine if a Schottky diode output signal indicates presence of metallic or non-metallic objects. Reflection loss extraction isolates reflection loss from the total RSS loss with a set of equations [144], [147]. Häfner *et al.* [118] measure azimuth angle and time of arrival by means of a maximum-likelihood based parameter estimator. Landmark extraction [132], based on a profile measured at a given azimuth angle, returns a

set of landmarks, i.e., set of range and azimuth value tuples. Landmark extraction performs a set of filtering operations, after which magnitude values are scaled according to the probability that the magnitude value indicates a landmark. Continuous peaks at certain ranges indicate a landmark. Pulse integration uses the integration operation with the purpose of improving signal-to-noise ratio [128], [179], help discover movement in a spectrogram [106], and to deduce a spectrogram representing micro-doppler signatures [129], [130]. Pulse integration types include incoherent integration [106], coherent integration [128], spectrogram integration across range [129], [130], and a wideband signal filter operation [179]. POSP is used in [141] to calculate an integral. Further details, including the long version of POSP, are not present. Frequency compression is not explained in [88].

After using pulse compression on a given set of positions, while scanning an area by means of free space scanning with a moving device, a two dimensional structure of complex values, representing combined IF I/Q signals, also known as a SAR image is obtained. When using the magnitude of these complex values to create a grayscale image, the resulting raw SAR grayscale image is severely out of focus because it does not represent spatial information correctly yet. In most SAR pre-processing pipelines, range and azimuth reference functions are generated and convoluted with the SAR image in sequence to increase the focus. As a result, the SAR image correctly represents spatial information. These techniques are known as range [161] and azimuth compression [182], [183]. More information can be found in [182], [183]. Maximally Stable Extremal Region (MSER) exploits the fact that reflections originating from the metal structures of vehicles, in contrast to reflections from the road surface, show up as bright stable area's in a SAR grayscale image. Candidate regions are selected that stay below a grayscale area variation rate [177]. Visual saliency detection, after a set of pre-processing steps performed on a SAR image, returns a binary image in which white area's indicate presence of objects [177] or parked vehicles [176].

3) Data Reconstruction and Denoising: Several sensing application pipelines deploy reconstruction and denoising techniques in the data domain. The techniques try to resolve data corruption, which refers to observing values in a data structure that are drastically different from the expected values in a certain context. In addition to filtering the types of corruption already mentioned in Section IV-B.1, data reconstruction and denoising also filter Doppler components caused by transmitter time multiplexing, remove redundant data parts, and filter data transients. Data transients are caused by persons that become stationary after walking into a room [81].

Under a sampling and signal change assumption, one dimensional phase unwrapping is performed by [57]. The phases of complex numbers in the slow time direction are analyzed from a two dimensional data structure prior to radial velocity FFT operations. A wrapped phase change greater than π in a pair of consecutive complex numbers indicates that the phase of the second complex number should be corrected by adding or subtracting 2π through means of phasor multiplication. One dimensional phase unwrapping can also be solved with a null range measurement obtained at the first chirp of the first frame. Subsequent chirp and frame processing is combined with multiplying the IF signal with the null range measurement [70]. Two dimensional phase unwrapping through a path following algorithm is considered in the sensing application pipeline explained in [178].

Constant False Alarm Rate (CFAR) is an adaptive thresholding technique that is used to extract, i.e., reduce a spectrogram containing magnitude values to, spectrogram parts, based on a sliding window, that indicate presence of targets against data corruption present in the spectrogram. Several CFAR types have been considered. One dimensional types include cell averaging [93], [114], [121], [123], [124], cell averaging smallest of [55], [71], [96], [114], cell averaging greatest of [71], clutter map [83], [84], and ordered statistics [92]. A two dimensional custom CFAR type is considered in [131]. The type indicates how the threshold is determined. More information on CFAR types can be found in [163], [209], [210].

Background subtraction is performed by removing point cloud data in a frame with zero Doppler velocity [43], [115] and Cartesian coordinates that fall outside certain boundaries [43], removing measured profile and spectrogram (part) averages from profiles and spectrograms [50], [81], [102], [138], deleting a profile belonging to an empty FOV from target measurement profiles [86], and statically removing the 0 Hz component from maximum Doppler frequency data across time [116].

Moving Target Indication (MTI) is not explained in [53], [58]. The papers suggest that it is a more general term used to indicate that denoising is performed. Static thresholding is used to extract spectrogram parts from which phase variation over time [62] or magnitude [127] exceeds a threshold or create binary radar images [67], [142] (unknown purpose [67] or to make material faults visible [142]). Linear Least Squares Estimation (LLSE) is used to estimate a shift caused by DC offset in complex numbers' imaginary and real parts of a 2D data structure, prior to radial velocity FFT (used to estimate chest vibration). The shift is used afterwards to adjust the complex number's imaginary and real parts [57]. 2D quadratic function fitting is used [178] to fit a function to phase data, which is later subtracted from the phase data to eliminate phase curvature. A per-pixel Gaussian model is used to subtract noise coming from unwanted background reflections from spectrograms [125]. The Otsu algorithm is an automatic threshold selection method supported by image segmentation. It is used for noise reduction in radar images [189]. Noise reduction with help from Singular Value Decomposition (SVD) can be applied to extract desired spectrogram parts indicating presence of target reflection [142]. A feed-forward neural network can be used to filter ghost targets (i.e., superimposed noise from unwanted background and target reflections) in automotive radar sensing [95]. The bandpass and highpass filters reported in this section are applied in the frequency domain. Profiles are multiplied with a transfer function that represents the bandpass or highpass filter's frequency response [211]. The transfer function can be custom made or determined

TABLE VII
SUMMARY OF DATA RECONSTRUCTION & DENOISING PRE-PROCESSING METHODS DEPLOYED IN MILLIMETER WAVE SENSING PIPELINES. THE DENOISING ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-B.3

	Reconstructon	Denoising																				
	Phase unwrapping	CFAR	MTI	Thresh-olding	LLSE	Function fitting	Gaussian model	Manual background subtraction	Otsu algo-rithm	SVD	Feed-forward NN	Band-pass filter	High-pass filter	Gaussian filter	Moving average filter	Offset version recombination	Doppler compen-sation	Deboun-cing	Mirror-ing	Distance correction	Window-ing
[43]								X													
[45]		X																			
[50]								X													
[53]		X	X																		
[55]		X																			
[57]	X				X																
[58]			X																		
[144]												X									
[62]	X	X		X								X									X
[67]				X																	
[70]	X																				
[189]									X												
[71]		X																			
[147]												X									
[78]													X								
[81]								X				X						X			
[83]		X																			
[84]		X																			
[85]				X				X													
[86]								X													
[92]		X											X								
[93]		X																			
[95]											X										
[96]		X						X													
[98]								X													
[102]								X					X							X	
[142]				X						X											
[178]	X					X															
[106]	X											X									
[107]												X									
[108]	X											X									
[113]																X					
[114]		X															X				X
[115]								X													
[116]								X													
[121]		X																			
[154]																			X		
[123]		X						X					X	X							
[124]		X																			
[125]							X														
[126]		X																			
[127]				X											X						
[128]				X											X						
[130]		X																			
[131]		X																			
[138]								X													

by the Laplace or Z domain (discrete-time equivalent of the Laplace domain). The explanation further assumes a linear, time-invariant, and single-input single-output filter. In the Z domain, the transfer function represents the ratio of the filter's Z transformed output to Z transformed input as a function of a complex number $z = Ae^{j\phi}$. The symbol A is the radius and the symbol ϕ the phase. The transfer function is determined from filter behavior by Z transforming an n^{th} order linear difference equation (both the equation and order are dependent on the filter type) and solving the equation for the Z transformed output to Z transformed input ratio. The Z domain is a more general form of the frequency domain and therefore z can be substituted with $e^{j\omega}$ to convert a transfer function to the frequency domain [212]. The transfer function of an ideal highpass filter in the frequency domain can be found in Equation 6 [211]. A 2D Gaussian filter is convolved over a spectrogram for extra noise reduction in addition to background subtraction [123]. Moving average filtering is performed in time by first subtracting an empty background spectrogram from a measurement spectrogram. Afterwards, the background spectrogram is updated with the measurement spectrogram [127]. In offset version recombination, a noisy set of complex numbers created with STFT is offset with a simple addition of a frequency dependent anti-symmetric function [113]. Spectrograms originating from real and virtual receiver channels cannot be retrieved simultaneously at real receivers. Radar collection systems employ multiplexing strategies to retrieve the spectrograms [172]. When time multiplexing is used, spectrograms from real and virtual receivers will experience an unwanted Doppler induced phase shift. This shift is compensated for by means of phasor multiplication [208]. To mitigate analytical model performance degradation caused by data transients, a debouncing logic can be implemented [81]. Radar images acquired through imaging do not properly represent the geometry of an environment due to multipath propagation. Image correction can be applied by means of mirroring techniques [154]. Distance correction [102] scales correlation profiles while taking into consideration that sample strength falls off according to a certain pattern. To remove motion corrupted segments from heartbeat data, the heartbeat data is segmented and segments are removed based on the outcome of a thresholding procedure [62]. Two dimensional windowing prior to performing a FFT operation to determine velocity [114] can be considered

a brick-wall filter [64].

$$H_{hpf}(\omega) = 1 - H_{lpf}(\omega)$$
$$H_{lpf}(\omega) = \text{rect}(\omega) = \begin{cases} 0, & \text{if } |\omega| > \omega_0 \\ \frac{1}{2}, & \text{if } |\omega| = \omega_0 \\ 1, & \text{if } |\omega| < \omega_0 \end{cases} \quad (6)$$

4) Data Transformation: Data transformation refers to using mathematical operations on data structures that either cause them to change into new, higher-level data types or cause data structure aspects, e.g., value range or axis range considered, to change. There is a wide variety in goals that sensing application pipelines try to achieve through use of data transformation. For example, value range changes can be used to unify the value range of different variables. This will omit bias towards variables that have a bigger value range compared to other variables during analytical model training [213]. Data type changes allow higher-level information to be extracted from lower-level information. For example, position information in Cartesian coordinates can be extracted from lower-level range and AoA information [138], voxels created from position information encapsulate higher-level body shape information [46], etc.

Peak detection is used to retrieve velocity information from a range, azimuth, and velocity cube [55], retrieve breath/heartbeat data by means of average inter-peak distance in a phase data sequence in time [62], and determine magnitude peak position in a range profile [107]. Thresholding is either applied on a range-compressed down-conversion IF signal to determine vertical range between the collection system and the ground [161] or used to determine if space debris is detected or not [88]. Spectral analysis is a general term used to describe the fact that IMFs are converted to a frequency domain representation to determine if they can be attributed to heartbeat or breathing data based on a given frequency range [106]. Dimension reduction is performed by transforming a two dimensional structure of complex values into a structure containing magnitude values and determining the position that corresponds to the dominant reflection of a lap joint. Data evaluation is performed at the position that yields a one dimensional structure [152]. Voxelization can be thought of as creating a three dimensional grid. The grid consists of voxels, i.e., custom values (or sets of custom values). Values found in the sensing application pipelines include number of points present [47] and accumulated velocity information of every point [138] in bounded coordinate regions of a point cloud frame. Voxelization is used to turn a variable number of point information rows [138], associated to a point cloud frame, into a data structure with fixed dimensions [47], [138]. Binarization refers to the creation of a binary image from a radar image. Binary images contain black and white values assigned through use of a threshold [191]. Mathematical morphology is a term used to describe a set of matrix operations that are used to transform binary images [176], [191]. These matrix operations are called erosion, dilation, opening, and closing, and involve an input image and a structuring element. The structuring element is similar to a kernel in the context of two dimensional convolution. Opening has been used to transform a binary saliency map into connected regions [176]. Dilation has been used to improve image visualization [189]. Low Rank Matrix Factorization (LRMF) is used to factorize a radar image into two: a foreground, i.e., concealed object, and background to detect concealed objects. Factorization is performed with a patch-based, i.e., segment-based, Gaussian mixture model [189]. Laplacian of Gaussian and Canny edge detection are used on a gray scale converted spectrogram containing magnitude values to discover Doppler frequency information [116]. Once range and AoA values retrieved from spectrograms or cubes are known, the values can be transformed from range and AoA information, through polar or spherical to two or three dimensional Cartesian coordinate conversion equations, to position information in Cartesian coordinates [43], [44], [46], [53], [55], [71], [93], [96], [103], [114], [121], [138]. The new data structure is referred to as either a point cloud frame or point scan. The data structures differ in context in which measurements were conducted. Stationary radar collection systems are used to retrieve point cloud frames while moving radar collection systems on a robot or vehicle are used to retrieve point scans. The active imaging application pipeline presented in [141] transforms coordinates by means of integration. Point transformation is concerned with converting a point cloud frame based on fixed coordinate ranges into a radar image. To every pixel, a RGB value is assigned. The R value corresponds to a x-coordinate, the G value to a y/z-coordinate (separate xy/xz images), and B value to a magnitude value. Coordinates that contain no points are assigned a black RGB value [103]. The application pipeline in [120] uses a similar approach in which points converted to camera coordinates are used to compute RGB values. Fast Wavelet Transformation (FWT) is used to transform a radar image into the wavelet domain with the intent to denoise it in the wavelet domain [191]. Greyscaling refers to converting a spectrogram consisting of magnitude values to a grayscale image [116]. Once uplink pilot signal AoA and time of arrival, and a radar image consisting of reflective shapes are known, casting a ray, i.e., millimeter wave, at the estimated AoA from a base station will eventually lead to a mobile user's range with respect to the base station under the assumption of reflection at a specular angle [154] with a computation on time of arrival.

Normalization refers to mapping real value ranges of several data structures. Spectrograms with magnitude values are normalized to unit scale [45], between 0 and 1 and centered around the mean value [92], by dividing every magnitude value with the sum of all magnitude values in the spectrogram [123] or by scaling magnitudes logarithmically and performing a max-min truncation [125]. RGB values assigned to a radar image with help from a point cloud frame with fixed coordinate ranges are based on normalized range, AoA, and magnitude information [103]. Point cloud frame coordinates are max-min normalized based on fixed coordinate ranges [138]. Region Of Interest (ROI) selection significantly reduces search dimensionality over a spectrogram during feature extraction computations [45]. To isolate periodic chest movement, among all spectrogram columns corresponding to range in an acceptable FOV, the range column

TABLE VIII
SUMMARY OF DATA TRANSFORMATION PRE-PROCESSING METHODS DEPLOYED IN MILLIMETER WAVE SENSING PIPELINES. THE ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-B.4

	Artifact (type) change														No Artifact (type) change							
	Peak detection	Auto-correl-ation	Thresh-olding	Spectral analysis	Dimension reduction	Voxeli-zation	Binari-zation	Math. morph-ology	LRMF	Edge detection	Coord-inate trans-form	Point trans-form	FWT	Grey-scaling	Ray casting	Normali-zation	ROI extraction	Edge region discovery	Windowing/ Segmentation	MRC	Manual smoothing	Conv-olution mask	Down-sampling
[158]																							X
[141]											X												
[43]											X										X		
[44]											X												
[45]																X	X						
[46]						X					X								X				
[47]						X													X				
[53]											X												
[55]	X										X												
[57]																	X						
[58]																							
[62]	X																						
[191]							X	X					X					X	X				
[189]								X	X														
[71]											X												
[176]								X															
[77]																			X				
[81]																							
[161]			X																				
[88]			X																				
[92]																X							
[93]											X												
[96]											X												
[100]																							
[103]											X	X				X							
[142]																						X	
[106]				X																			
[107]	X																X						
[108]	X	X																					
[152]					X																		
[114]											X												
[116]										X				X							X		
[120]												X											
[121]											X												
[154]															X								
[123]																X			X				
[125]																X							
[126]																				X			
[128]																	X						
[138]						X					X					X							

with the maximum average magnitude value is selected [57]. With background information on test person distance, a frequency domain ROI can be selected [107]. After thresholding, the closest detected range profile part in range is selected for each radar in case of multiple detected target peaks [128]. Through use of a pixel condition argument, person edges are discovered in image segments if a pixel value in the condition argument is greater than a certain threshold. The edge values are set to zero afterwards [191]. Windows containing a time dependent sequence of voxel grids are created and used to form a dataset [46], [47]. An image segmentation procedure is presented in [191]. Dynamic sized temporal input data can be formatted by using the Markov Frame method [77]. Windowing can also be employed to track target position across spectrograms [123]. Maximal Ration Combining (MRC) refers to the creation of mean range and velocity and range and azimuth spectrograms from a range, velocity, and azimuth cube containing magnitude values by computing a weighted mean [126]. Discontinuous range, azimuth sine, and velocity sequence values are smoothed. At sequence window ends, a discontinuous point is set to the value of its nearest neighbor. When eligible values are at both sides of the discontinuous point, the discontinuous point becomes the average of these values [43]. Frequency index smoothing is used in the pipeline explained in [116]. A convolution mask can be applied to an active imaging spectrogram to make a concealed tile image and the crack within it more visible [142]. Downsampling is used to convert a CSI frequency profile into a smaller feature vector [158].

5) Dataset Creation and Augmentation: After datasets, consisting of data samples, have been generated by both the signal and data transformation processes, several sensing application pipelines augment these datasets with artificially created data samples. In case of profiles, spectrograms, and cubes, it is assumed that real value structures consisting of magnitude values are generated. Reasons for data augmentation include improving analytical modeling performance on unseen data during inference [63], [127], [128], [138] and to increase the dataset size to a size required during model training to achieve good analytical model performance [138], [192] without the need for extra sampling. For each original spectrogram in the dataset or for an average spectrogram per class, artificial spectrograms can be generated by sampling pixel values according to a normal distribution [63], [127]. Radar images can be fused with image segments at a random location [192]. Radar images and point cloud frames have also been rotated, scaled, and content inside the images and point cloud frames has been translated [128], [138]. To prepare datasets for analytical model training in several sensing application pipelines, data samples are labeled with class labels and part of the dataset is designated for model training while other parts are designated for model validation and testing.

TABLE IX
SUMMARY OF DATASET AUGMENTATION & CREATION PRE-PROCESSING METHODS DEPLOYED IN MILLIMETER WAVE SENSING PIPELINES

	Augmentation			Creation	
	Normal distribution	Object at random location fusion	Rotation, scaling, skew, and translation	Labeling	Splitting
[47]					X
[50]					X
[53]				X	
[63]	X				X
[64]				X	X
[181]					X
[67]				X	
[192]		X			X
[193]				X	X
[78]					X
[79]					X
[84]					X
[92]					X
[149]					X
[93]				X	X
[103]					X
[123]					X
[125]				X	X
[126]					X
[127]	X				X
[128]			X	X	X
[129]					X
[130]					X
[137]					X
[138]			X		
[139]				X	X

6) Gaps and Challenges: During the analysis of pre-processing methods, it was discovered that there is a lot of variety in the definition of what is considered to be pre-processing and way that pre-processing pipelines are constructed. We consider signal processing methods and typical data science pre-processing methods to be pre-processing methods. Other papers consider signal processing methods to be part of the signal collection system and only consider typical data science pre-processing methods to be pre-processing methods. Sensing application pipelines perform the processes mentioned in Figure 6 in different orders, skip processes, revert back to pre-processing after feature extraction or analytical modeling, etc.

When looking at Tables V, VI, VII, VIII, and IX, there are a few observations that raise interest. Many of the table columns contain a few crosses while a large portion of the crosses is located in just a few columns. There are two explanations for this. Many sensing application pipelines rely on pre-processing methods that are well known. In signal transformation, many sensing application pipelines rely on FFT and STFT for signal conversion. However, estimation techniques such as Capon beamforming, Bartlett beamforming, MUltiple SIgnal Classification (MUSIC), Sparse Asymptotic Minimum Variance (SAMV) beamforming, Maximum Likelihood Estimation (MLE), etc. exist that can be used in the pre-processing pipeline too. In [72], [93], test results have been reported for the MUSIC algorithm. Many of the pre-processing methods are sensing application pipeline specific. For example, landmark extraction [132] works with a specific radar collection system that collects a signal at specific azimuth angles, Doppler compensation [114] is only used when transmit time multiplexing is used for measuring AoA information, etc.

Another observation is that many table rows contain multiple crosses. This indicates use of multiple reconstruction, denoising, and transformation pre-processing methods. In [81], multiple pre-processing pipeline paths are used in parallel to obtain different artifacts and artifact types that all use their own denoising and transformation methods. In [62], there are two consecutive pre-processing stages. In the first stage, phase data is obtained in a pre-processing pipeline where CFAR is used to denoise spectrogram data. In the second stage, phase data is bandpass filtered before being used to extract breath/heartbeat data. Pre-processing methods are also used consecutively. In [159], it is explained that signal reconstruction is specifically used to omit limitations encountered with denoising of the received signal phase through DDBR.

C. Feature Extraction

The end result of the pre-processing phase in the application pipeline is a set of raw data samples in the form of, for example, sequence windows [137], [139], [158], (voxelized) point cloud frames [46], [47], [138], spectrograms [48], [76], [93], radar images [103], [120], [174], etc. This raw form of data, however, is not always used directly by analytical models. Degradation in modeling performance is sometimes caused by noise and redundancy in raw data samples. Extraction of relevant and informative data features from raw data samples is performed to remove the noise and redundancy [214]. Reducing the raw dataset to a limited number of features also makes it easier to visualize raw data samples for better understanding and gaining knowledge about the process that led to the generated raw data samples [215].

The feature extraction methods used by millimeter wave sensing applications take in the raw dataset and map the entire dataset to a new feature space [214]. The feature extraction phase is, however, not a required step. Recent deep learning techniques [47], [103], [104], [123], [124], [149], [192] and several modeling algorithms [56], [159], [184] perform well on raw data samples and, therefore, do not require the feature extraction phase. The border between feature extraction and analytical modeling has become blurry in recent papers on millimeter wave sensing applications. Several deep learning models contain layers designated for feature extraction [16], [46], [47], [92], [125], [126], [130], [138], [158]. Certain positioning and environment mapping algorithms resemble feature extraction methodologies [160], [184]. Transfer learning [120], [181], [193] is not included in this section since the methodologies transfer analytical model parameters to another analytical model rather than extracted data features. Feature analysis [43], [77], [91], [129] was only addressed by a minority of papers and is therefore excluded from this paper as well.

Table X presents our analysis of feature extraction approaches used in millimeter wave sensing applications. In what follows, we explain the main feature extraction methods used in millimeter wave sensing applications.

1) Manual Feature Mapping: Manual feature mapping relies heavily on experience and knowledge of domain experts to extract features [215]. The main feature categories found in the pipelines include statistics, calculus, geometry, vision,

TABLE X
SUMMARY OF FEATURE EXTRACTION METHODS DEPLOYED IN MILLIMETER WAVE SENSING PIPELINES. THE AUTOMATIC MAPPING ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-C

	Manual mapping		Automatic mapping					
	Time dependent	Frequency dependent	Clustering	A-priori estimation	Meta learning	t-SNE	Representation learning	PCA
[42]	X	X						
[43]	X	X						
[159]	X							
[45]	X	X						
[46]			X					
[47]								X
[143]	X							
[49]	X	X						
[50]								X
[174]							X	
[55]				X				
[57]		X						
[145]	X							
[60]		X						
[146]	X	X						
[62]		X	X					
[64]	X	X						X
[189]	X							
[72]		X						
[73]	X							
[147]	X	X						
[176]								X
[76]								X
[78]	X	X						
[79]	X							
[81]			X					
[161]	X							
[83]	X							
[84]		X						
[91]	X	X						
[93]			X					
[94]				X				
[96]				X				
[98]		X						
[114]				X				
[116]		X						
[127]					X	X		
[128]		X						
[129]	X	X						
[137]	X							
[139]	X							

and gesture recognition. Statistical features include (weighted) mean [45], [49], [64], [72], [91], [98], [137], [139], [143], [147], median [129], total value [42], [45], [49], [91], standard deviation and variance [45], [143], [147], [159], root mean square [45], [49], [137], [139], moving average [73], [98], sequence window centroid [45], [91], [161], range [43], [129], [147], [157], quartiles [147], higher order cumulants [83], coefficients [98], [129], value normalization [43], [84], [129], value distribution [42], [49], [79], [91], [189], probability [161], and histogram features [189]. Calculus features include maximum [43], [49], [57], [60], [64], [72], [84], [116], [137], [139], [145], [147], minimum [49], [64], [128], [137], [139], absolute value [49], [73], [137], [139], a signal pattern across time [159], discrete differentiation [42], [45], [49], [73], [91], discrete integration [45], derivative sum [45], integrated sum [45], and integrated delta [45]. Geometrical features include intensity areas [72], roundness [72], eccentricity [72], perimeter [72], and shape slope [72]. A specific visual feature used is the so called local binary pattern [189]. The authors in [78] use a feature extraction algorithm comprised of steps that involve some of the features mentioned above.

In 2016, Lien *et al.* [45] conducted extensive research into the development of range-doppler spectrogram specific features in the context of gesture recognition. The goal was to reduce computational overhead on resource constrained devices. After ROIs have been selected, matrix calculations are applied to these regions to obtain multi-channel integration, multi-channel derivative, and temporal derivative matrices. In 2018, Flintoff *et al.* [91] introduced a new gesture recognition feature in the form of a sonar value.

2) Automatic Feature Mapping: Automatic feature mapping methods used in the millimeter wave sensing applications mainly aim at dimensionality reduction. Several tracking applications that utilize point cloud frames either use clustering [46], [93] or the Kalman filter's prediction process [55], [94], [96], [114] to reduce point clouds into point cloud cluster centroid values. Since the number of point clouds per frame is unknown, a clustering algorithm that does not require a number of clusters to be defined a priori is required [46]. More information about the prediction process can be found in [216]. Clustering is also used to reduce people counts corresponding to the same person [81] or to obtain a median heart rate measurement from a noisy spectrum [62]. Principal Component Analysis (PCA) is used for linearly reducing an input data vector v into principle component vector $p = W^T v$ containing decorrelated principal components suitable for use with Support Vector Machines (SVMs) and decision trees. The symbol $W \in \mathbb{R}^{(N,L)}$ denotes a projection matrix where N is the input vector size and L the number of desired principal features. The projection matrix elements are calculated such that the variance included in input data vectors used for training is maximized [217]. To reduce a dataset to a low feature dimension for visualization purposes, t-distributed Stochastic Neighbor Embedding (t-SNE) can be used [127].

Other automatic feature mapping methods include representation and meta learning [127], [217]. Even though these methods also involve dimensionality reduction, it is considered to be a positive side effect rather than a goal. The main idea behind representation learning is to create a machine learning model that has the ability to extract features from a dataset. The features can be fed into a variety of downstream models used for solving a task that is loosely related to the task associated to the upstream model [217]. Meta learning differs from representation learning because it has the goal of providing features that can be used for a variety of different task types rather than loosely related tasks [127].

3) Gaps and Challenges: Table X shows that the majority of approaches that utilize feature extraction rely on manual feature mapping. In contrast to reducing computational overhead, manual feature mapping costs a lot of time and human resources. Automatic feature mapping in millimeter wave sensing applications is an unexplored field of research. Only one methodology found [127] uses a feature extraction technique for feature visualization to gain an understanding of extracted features. Lastly, meta and representation learning are severely underexploited in millimeter wave pipelines.

D. Analytical Modeling

Most millimeter wave sensing applications cannot reach their goal by simply executing a set of pre-processing steps and extracting information relevant to the application. They require a model to translate information from pre-processed datasets or extracted feature sets to the application goal. Examples of

these models include classification models [46], [47], [79], filter models [46], [89], [185], and measurement models [133], [159], [184], [186]. Classification models assign class labels to a set of input data. The class can indicate performed activities, gestures or events/failures, presence of a specific person/object, etc. Filter models estimate trajectories of tracked entities, filter out false positive class predictions, etc. in an environment where noise causes trajectory measurements and class predictions to be stochastic in nature. Measurement models calculate the value of information such as position, yaw rate, or absolute velocity using mathematical models. In some cases, measurement models use a cost minimization technique for value calculation. The costs are computed with a function that uses input data as a function parameter. We classified the models used in millimeter wave sensing applications based on the model class (data driven vs. model driven) and the model type (black, grey, white models). Black, grey and white model types do not refer to the degree of model explainability. The model types refer to model determinism and use of physical knowledge and or data for model construction. The model types were adopted from a physiological model identification study presented by Duun-Henriksen *et al.* [218]. Apart from occasionally explaining that a certain model fuses millimeter wave data with data originating from other sensing domains in Section IV-D.3 to avoid hybrid model explanation ambiguity, further discussion of models that fuse millimeter wave data with data originating from other sensing domains is outside the scope of this review. In what follows, we analyze the analytical model classes and model types used in millimeter wave sensing applications.

1) Model Driven Modeling: Model driven modeling relies on physical and or mathematical knowledge about the technology or the environment to construct a model. The knowledge refers to information, a set of rules, and or a set of equations obtained from physical phenomena or mathematical proofs. The model driven approaches are classified either as white box or grey box models. White box models vary in determinism and solely rely on physical and or mathematical knowledge for model construction. For example, phase tracking [159] deploys position initialization, resulting in an extra dependency on a random variable in addition to phase inputs for a certain position output. In contrary, instantaneous ego-motion estimation [133] relies on Monte-Carlo simulation to test a model. This indicates that the computations performed by the model are deterministic in principle. Grey box models are non-deterministic and also rely on physical and or mathematical knowledge for model construction. They differ from white box models because parts of the model are continuously altered across time or completed with information extracted from input data and data features [218]. Examples include hidden state updates and use of cost and similarity parameters.

Table XI presents our analysis of model driven approaches used in millimeter wave sensing application pipelines. Next, we explain different white and grey box model approaches.

a) White box models: The phase tracking model [159] calculates a two dimensional position in Cartesian coordinates based on successive phase shifts. The two dimensional position is calculated with a combination of distance equation sets and triangulation. The initial position is measured with a special acquisition module that is independent of the input data. The velocity and the Ackerman model [133] first calculates a radar velocity vector containing longitudinal and lateral velocities based on the sinusoidal progress of measured radial velocities over the azimuth angle using a least-square approach. Afterwards, an object's absolute velocity and yaw rate are calculated using a model based on the Ackerman condition involving a velocity and yaw rate equation. The angle/time model [184] calculates mobile object positions based on a configuration involving multi-path component Time of Flight (ToF) and AoA information or a configuration involving multi-path component ToF and Angle of Departure (AoD) information. The information is retrieved from a ray tracing tool. This information, combined with meta-data such as a map layout, results in object locations. The eventual location is determined based on majority voting. Relative motion between two radar scans can be measured in a least-square sense by using SVD on a set of landmark matches [132]. Double Time Difference Of Arrival (DTDOA) is a positioning model consisting of a set of equations that use pseudo time of arrival to compute a position estimate [162].

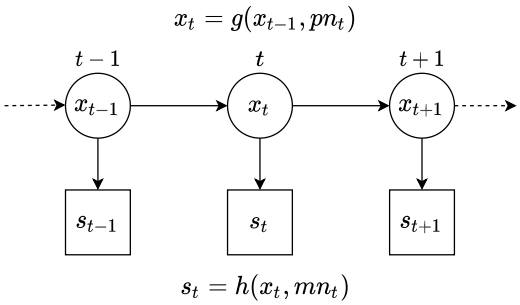


Fig. 7. Hidden markov model containing observed measurements *s* and a hidden state *x*, describing information of interest to the application, at time instant *t*. Symbols *g* and *h* denote a state transition and observation function. Symbols *pn* and *mn* denote process and measurement noise. Recreated from [219].

b) Grey box models: Bayesian filters originate from stochastic filtering theory and Bayesian statistics. As indicated in Section IV-B, millimeter wave sensing applications obtain measurements as discrete samples. Therefore, the explanation assumes a discrete Bayesian filter modeled as a Hidden Markov Model (HMM). The model contains one hidden state that varies with time. The state describes information that is of interest to the millimeter wave sensing application. The exact information is unobservable and thus hidden. Furthermore, action variables are omitted. An overview can be found in Figure 7. The hidden state x_t is dependent on one previous state x_{t-1}. Measurements s_t in the form of raw data samples or data features are observed at, i.e., sampled from, the pre-processing or feature extraction phase. These measurements are dependent on the current hidden state. The state transition function includes noise to account for a distribution of different outputs the function can yield at a given time instant. The measurement observation function accounts for noise to take care of measurement inaccuracies [219], [220].

TABLE XI
SUMMARY OF MODEL DRIVEN ANALYTICAL MODELS USED IN MILLIMETER WAVE SENSING PIPELINES. THE ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-D

	White box					Grey box									
	Phase tracking	SVD	Velocity/Acker-man model	Angle/Time model	DTDOA	Shadow RTI	Bayesian filter	Association/ Allocation	Scan matching	MLE	Triangulation	Angle/Time model	RSA model	Rician model	Velocity model
[44]							X								
[159]	X														
[46]							X	X							
[51]							X								
[52]							X								
[55]							X	X							
[184]				X											
[58]							X								
[145]											X	X			
[61]							X								
[146]							X								
[201]								X							
[161]															X
[87]															X
[89]							X	X	X						
[90]														X	
[94]							X	X							
[162]					X										
[96]							X	X							
[150]						X									
[152]										X					
[185]							X	X							
[114]							X	X							
[115]							X								
[153]													X		
[132]		X						X							
[133]			X												
[134]									X	X					
[135]							X		X						

The state transition and the state observation functions $x_t = g(x_{t-1}, pn_t)$ and $s_t = h(x_t, mn_t)$ are modeled as Probability Density Functions (PDFs) $p(x_t|x_{t-1})$ and $p(s_t|x_t)$ [220]. PDFs explicitly represent the uncertainty in variables taking on a particular value in a range of values at a specific time instant. The Bayesian filter continuously executes two functions called the predict and update functions with a recursive algorithm. The role of the predict function is to estimate a new hidden state PDF $p(x_t|s_{1:t-1}) = \sum_{x_{t-1}} p(x_t|x_{t-1})p(x_{t-1}|s_{1:t-1})$ based on the previous hidden state PDF and state transition PDF. The new hidden state PDF can be used to derive hidden state values [221]. The role of the update function is to adjust the hidden state PDF when measurements are observed. The update function is based on Bayes' theorem and can be noted in Equation 7. The update function ensures that the hidden state PDF can always be used to derive hidden state values that closely represent, and do not drift away from, the exact hidden state values. Calculation and or modeling of $p(x_t|s_{1:t-1})$, $p(s_t|x_t)$, and $p(x_t|x_{t-1})$ is a core task of Bayesian filtering. More information can be found in a comprehensive review authored by Chen [219].

$$p(x_t|s_{1:t}) = \frac{p(s_t|x_t)\,p(x_t|s_{1:t-1})}{p(s_t|s_{1:t-1})}$$
$$p(s_t|s_{1:t-1}) = \sum_{x_t} p(s_t|x_t)\,p(x_t|s_{1:t-1}) \tag{7}$$

Several Bayesian filters exist, such as the custom Bayesian [45], [185], particle [44], [136], $\alpha - \beta$ [128], Kalman [46], [53], [58], extended Kalman [55], [94], [96], [114], [115], [135], [186], fusion extended Kalman [95], [146], unscented Kalman [89], fusion adaptive Kalman [51], and adaptive Sage-Husa Kalman [52], [61] filter. The Kalman filter [219], under linear, quadratic, and Gaussian assumptions, can represent the state transition and observation functions $x_t = g(x_{t-1}, pn_t)$ and $s_t = h(x_t, mn_t)$ as a set of linear equations. All other Kalman filters loosen these assumptions. The extended Kalman filter, for example, allows the functions to be approximated as a Jacobian matrix [219]. Particle filters [219], [220] approximate $p(x_t|s_{1:t-1})$ as a set of particles $\{< x_t^{(i)}, w_t^{(i)} >\}_{i=1}^{l}$ where $w_t^{(i)} = p(s_t|x_t^{(i)})$. The $\alpha - \beta$ filter [128] is a simplified Bayesian filter that limits the number of hidden states to two. Predictions and updates are executed with a simple set of equations. Symbols α and β refer to manually set correction gains used during the update process.

Scan matching models are concerned with finding the best rotation and translation operations that have the ability to align two point scans or a point scan to an existing area map. The Iterative Closest Point (ICP) model involves two iterative steps until convergence. In the first step, points from two are matched based on closeness to one another in a given space. Closeness is measured by means of a distance metric. In the second step, to find the optimal rotational angle and translation, the sum of squared distances between the points is used [135]. Normal Distribution Transformation (NDT) scan matching [89], [134] requires a grid of probability functions created from a map for matching.

Association/allocation models either use combinatorial optimization [46], [185], a gating function [55], [94], [96], [114], template model matching [201] or landmark association [89], [132]. These models associate input data to an existing entity or create new entities based on an optimization criterium. This criterium is dependent on a cost metric involving the input data. For example in tracking, the models determine if incoming input data belong to a certain track that is already known or indicate a new track.

Other grey box models include Shadow RTI [150], triangulation, angle/time model [145], RSS Series Analysis (RSA) model [153], velocity model [161], Rician model [90], and

TABLE XII
SUMMARY OF DATA DRIVEN ANALYTICAL MODELS USED IN MILLIMETER WAVE SENSING PIPELINES. THE ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-D

	Shallow black box							Deep black box							
	SVM	SVDD	MPM	SOM	Decision tree	Clustering	k-NN	LVQ	One-versus-one SVM	AdaBoost SVM	Random forest	Feed-forward NN	CNN	LSTM	CNN+LSTM
[158]													X		
[42]											X				
[141]													X		
[43]					X										
[159]					X										
[46]														X	
[47]									X			X		X	X
[143]											X			X	
[49]											X				
[50]											X				
[174]													X		
[56]				X				X							
[63]															X
[16]															X
[64]									X						
[181]													X		
[192]													X		
[193]													X		
[189]										X					
[71]						X									
[72]							X							X	
[73]						X									
[147]					X										
[176]	X														
[76]														X	
[77]															X
[78]									X						
[79]									X						
[83]		X													
[84]			X												
[91]											X				
[149]												X			
[93]													X		
[103]													X		
[104]												X			
[120]													X		
[123]													X		
[124]													X		
[125]															X
[126]															X
[127]						X	X								
[129]							X								
[130]													X		
[137]											X				
[138]						X									X
[139]									X						

MLE [134], [152]. Shadow RTI [150] measures link RSS attenuation changes across a network of millimeter wave transceivers. The knowledge utilized for model construction is that when people walk by, the RSS attenuation changes. Every link is made of a set of 'pixels'. Link RSS attenuation contribution is computed for every pixel by solving a least square problem based on a measured RSS change vector. The triangulation and angle/time models in [145] use AoA spectra coming from multiple access points to determine a mobile client's position from a set of measured anchor positions based on an associated cost metric. Additional input data include meta-data that are not measured such as room boundaries and a permanent obstacle set. The RSA model [153] measures object characteristics based on separate models. The object characteristics include surface curvature, surface boundary, and material. The surface boundary is determined with a surface reflection model that predicts RSS series of a reflection surface with fixed-size surface boundaries. The surface boundary of a measured RSS series is determined by matching the RSS series to a set of predicted RSS series produced with the surface reflection model based on a similarity metric. The Rician model in [90] was derived from the Rician model in [222]. The model compares a measured radar cross section with pre-simulated radar cross sections to identify a class based on log-likelihood computations. The measured cross section receives the class of the pre-simulated radar cross section that results in the largest computed log-likelihood value. MLE is used to estimate data of interest from noisy measurement signals. The model in [152] uses reference data and a likelihood function to return a data estimate that maximizes the likelihood function value. Joint spatial and Doppler-based ego-motion estimation in [134] computes state PDFs for a vehicle's yaw rate, longitudinal velocity, and lateral velocity by using a joint optimization problem. Mean velocity [87], [161] is measured with a set of equations and pulse-pair phase estimation, which relies on a correlation product that depends on the input data.

2) Data Driven Modeling: This type of modeling solely relies on information contained in the input data itself as well as the feature sets to construct a model [218]. Model construction typically involves choosing a model type and associated hyperparameters, fitting the model parameters to a dataset designated for training and testing the model performance on a dataset designated for validation. Fitting and validation are performed continuously in an iterative manner based on a cost or reward function by changing the model type or hyperparameters until the model exhibits the desired performance on the validation dataset [223]–[225]. Data driven models are classified either as a shallow black box or deep black box model. Both types are non-deterministic and solely rely on data for model construction. The difference between the two is that in shallow black box models, the input only passes through a single model structure before obtaining a model output while in a deep black box model, the input

goes through multiple submodels such as layers, decision trees, or SVMs.

Table XII presents our analysis of data driven approaches used in millimeter wave sensing applications. Next, we explain different shallow and deep black box model approaches.

a) Shallow black box models: Supervised shallow black box models, such as SVMs and decision trees, used for classification can perform well with limited datasets of small size [226]. However, this comes at the cost of requiring carefully crafted features. The features are either crafted manually [43], [78], [79], [147], [159] or automatically in a linear fashion [47], [50], [64], [176]. This process is labor intensive and or can limit the eventual model performance for more complicated tasks during model validation.

The pipelines deploying SVMs use it to learn and execute a C-Support Vector Classification (C-SVC) task. SVMs are applicable to binary classification tasks in which an optimal hyperplane is learned from training data. The hyperplane divides the data sample space into two sections, each corresponding to a given class. The classification is based on a decision function, which uses a set of support vector samples and a kernel function. The kernel function introduces non-linearity to the optimal hyperplane. Additional information on how SVMs learn can be found in [224].

Decision trees learn a tree structure containing nodes, branches, and leaves from a training dataset. Nodes represent binary input attribute tests, branches represent test outputs, and leaves represent a class that can be assigned. An unseen input dataset goes throughout the entire tree and is ultimately assigned to a single class label [43]. More information regarding decision tree construction can be found in [225].

Most approaches using a SVM or decision tree have been classified as an ensemble of several classifiers. These ensembles are formed with special multi-class and ensemble learning strategies such as one-versus-one [47], [64], [78], [79], [139], random forest [42], [45], [49], [50], [91], [137], [143], and AdaBoost [189]. Commonly, by letting several weak classifiers assign class labels to a given input, the final classification output is determined via majority voting.

Other models include clustering [53], [73], [136], [160], Support Vector Data Description (SVDD) [83], Minimax Probability Machine (MPM) [84], Self-Organizing Map (SOM) [56], Learning Vector Quantization (LVQ) [56], and k-Nearest Neighbors (k-NN) [72], [127], [129]. Clustering is an unsupervised learning algorithm to divide data into several clusters of data points that are similar to one another. Different techniques such as K-means [73], [160], DBSCAN [53], [138], Mean shift [71], or DenStream [136] clustering have been used. SVDD and MPM are similar to a SVM. They differ from a SVM because both learn a hypersphere rather than a hyperplane that separates all data and feature samples into two classes [83], [84]. A SOM is an unsupervised one layer neural network that reduces n-dimensional input vectors into a 2D feature map. A class is assigned to an unseen input sample based on how close the input sample corresponds to a feature map weight vector based on a distance metric. The LVQ is a supervised two layer neural network containing a competitive and fully connected layer. The competitive layer is similar to a SOM apart from having a limited and predefined number of outputs instead of having an output for every feature map weight vector and containing a transfer function. The output of the transfer function is fed into a fully connected layer to return a classification result for a set of user defined classes [56]. A k-NN model assigns a class to an input based on a majority vote of labels from k-nearest samples in a given data space.

b) Deep black box models: Most deep black box models belong to the machine learning paradigm called deep learning. The basic idea of deep learning is that by concatenating a number of submodels (layers), increasing the number of computation nodes in a submodel (neurons), and combining every submodel with an activation function, a non-linear vector mapping function that varies in complexity is learned from training data. The varying complexity comes at the cost of requiring large training datasets and a long training time. In their most basic form, deep learning models consist of a number of fully connected layers. This means that every neuron from the previous layer is involved in the computation of a neuron in the current layer during a forward pass of input data samples. This has been represented mathematically via matrix multiplication in Equation 8. The symbol x is the input vector, y the output vector, b the bias vector, σ the activation function (element-wise function), and W the weight matrix. Weights determine the degree of involvement of a certain neuron from the previous layer in computing a neuron in the current layer. The superscript i is the layer index. The input $x^{(i)}$ of the current layer is the output of the previous layer $y^{(i-1)}$.

Training a deep learning model happens with data batches and epochs. During a single epoch, a complete dataset is passed through the model with so called batches (smaller subsets of a dataset). After a batch has been passed through the model, the model output vectors are compared to label vectors, expected values, original input batch, etc. via a cost function C depending on the type of learning and model task type. The cost function is minimized via gradient descent which involves iterative weight and bias updates based on an average batch cost function value. A single weight update is depicted in Equation 9. The symbol η is a hyperparameter that controls the amount of a partial derivative that is used to update weights and biases. Because the forward pass of input data samples involves layer concatenations, the partial derivative is calculated via a backpropagation algorithm that has been derived with help of the Calculus chain rule. Additional information on deep learning can be found in [223].

$$y^{(i)} = \sigma(W^{(i)}x^{(i)} + b^{(i)}) \tag{8}$$

$$W_n^{(i)} = W_n^{(i)} - \eta \frac{\partial C}{\partial W_n^{(i)}} \tag{9}$$

A Convolutional Neural Network (CNN) is a deep learning model of which at least one layer involves convolution operations. CNNs found in millimeter wave sensing application pipelines are mainly combined with 2D spectrograms [48],

TABLE XIII
SUMMARY OF HYBRID ANALYTICAL MODELS USED IN MILLIMETER WAVE SENSING PIPELINES. THE ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-D

	Model driven								Data driven				
	White box			Grey box					Shallow black box	Deep black box			
	Path geometry	Triangulateration	Trilateration	Bayesian filter	Scan matching	Association/ Allocation	Angle/Time model	RSS model	Clustering	Random forest	CNN	LSTM	CNN+LSTM
[45]				X						X			
[48]						X			X		X		
[53]				X					X		X		
[160]	X								X				
[92]				X							X		X
[95]				X					X				
[186]		X		X			X	X	X				
[128]			X	X							X	X	
[136]				X	X				X				

[53], [92], [93], [123], [127], [130] and 2D radar images [103], [120], [174], [181], [192], [193]. CNNs have been used extensively for images originating from vision and object detection domains in the past. Spectrograms and radar images share similar characteristics with these images. The data types are 2-dimensional and features relevant to the mapping function are made up of values local to one another in the value matrix. Another observation is that many approaches base their model on existing vision and object detection models such as VGG [92], [93], ResNet [92], [123], ZFnet [193], Faster R-CNN [130], [181], YOLO [192] and FCOS [120]. The pipelines therefore rely on experience gathered with images in the vision and object detection domains. In addition, a few approaches have tried using a temporal CNN with 1D profile [158] data and a feature distributed CNN with a single point cloud frame [124] as input.

An approach that can be combined with time sequence data is the Recurrent Neural Network (RNN). RNNs are very good at discovering time dependencies among these sequences because they take into account historical information (i.e. information from previous input data samples) via hidden state vectors. Most approaches adopt the Long Short Term Memory (LSTM) network which is a special type of RNN that deals with the vanishing gradient problem noted when training traditional RNNs [128]. The authors in [77] use a Gated Recurrent Unit (GRU) layer. This layer is comparable to a LSTM layer apart from not having separate memory cells. The LSTM network is mainly combined with a flattened (voxelized) point cloud frame sequence [46], [47], Manually extracted 2D spectrogram sequence features [72], RSS fingerprint matrix [143], PCA extracted feature sequence [76], and alpha-beta filtered trajectory [128]. An important observation here is that LSTM requires some kind of feature extraction prior to using the time sequences for training.

CNN/LSTM combinations harness the strengths from both models. Comparable to local spectrogram and radar image features, point cloud frames contain clouds that are located very sparsely. In contrast, the positioning of points inside a given cloud is very compact. Therefore, time distributed CNNs can be used to obtain features from a single spectrogram, radar image, and point cloud frame in a sequence while LSTM models the time dependency in the sequence afterwards in an end-to-end fashion [77], [125], [138]. The authors in [46] denote that bi-directional LSTMs converge faster than CNN/LSTM combinations for a person identification task. However, this does not mean that they are better since the authors in [47] note that a CNN/LSTM combination outperforms a bi-directional LSTMs network in a human activity detection task. Therefore, model testing remains the norm when creating new models.

3) Hybrid Modeling: In Tables XI and XII several rows include multiple crosses, indicating use of multiple models. There are reasons for combining models. The authors in [47], [56], [145] compare several self-created models to each other. Authors in [46], [55], [72], [89], [94], [96], [114], [127], [132], [134], [135], [143], [185] incorporate different processes in their millimeter wave sensing application pipeline that use a distinct model. Several papers use pipelines that include a tree or sequential set of multiple models coming from both the data and model driven paradigms. These sets are considered to be hybrid models. Table XIII presents our analysis of hybrid models used in millimeter wave sensing application pipelines.

In principle, hybrid models are constructed by putting the mathematical and data driven models together in a sequential or tree-based manner. Hybrid models constructed in a sequential manner were found in [45], [48], [53], [92], [160]. In [45], a Bayesian filter was combined with a random forest output to reduce sporadic false-positive errors. A patient behavior detection task first tracks the patient with a combination of clustering and a Bayesian filter. Tracking information is used to construct two dimensional Doppler spectrograms that can be used with a CNN for behavior detection. This limits computational complexity since additional pre-processing with STFT or FWT is omitted [53]. Environmental mapping [160] is performed with a specific set of steps. First, spatial channel profiles consisting of AoA, AoD and RSS information are associated to a potential reflector through clustering. Afterwards, reflector points are retrieved through elementary geometry. An issue experienced in moving target classification is having micro-Doppler signatures spread across many different range bins in profiles across time. This issue is solved by tracking the position of these bins with a Bayesian filter and association/allocation combination. Afterwards, the bins are passed through STFT and fed into different neural networks for classification [92]. Motion behavior detection in [48] is performed by first applying a clustering algorithm on point

TABLE XIV
SUMMARY OF EVALUATION METRICS USED FOR MEASURING PERFORMANCE OF THE ANALYTICAL MODELS MENTIONED IN SECTION IV-D. THE DATA DRIVEN EVALUATION METRIC ABBREVIATIONS ARE EXPLAINED THROUGHOUT SECTION IV-E.1

	Data driven										Model driven		
	Accuracy	Invalid gesture rate	Precision	Recall	Specificity	F1 score	Confusion matrix	ROC	Error	Visual inspection	Accuracy	Error	Visual inspection
[158]	X									X			
[42]	X												
[43]	X	X											
[44]												X	X
[159]	X											X	
[45]	X										X		
[46]	X						X					X	
[47]	X						X						
[48]	X												
[143]	X						X						
[49]	X						X						
[50]							X						
[51]											X	X	
[52]												X	
[53]	X												
[174]	X												
[160]	X										X	X	
[55]											X	X	X
[184]												X	
[56]	X						X						
[58]												X	
[144]	X						X						
[145]												X	
[61]												X	
[146]												X	
[63]	X						X						
[16]	X												
[64]	X					X	X						
[181]			X	X		X		X					
[191]			X	X									
[67]													X
[192]			X										
[201]												X	
[193]			X										
[189]			X	X									
[72]	X		X	X		X	X						
[73]	X												
[147]	X						X						
[176]				X									
[76]			X	X		X	X						
[77]	X												
[78]	X						X						
[79]	X						X						
[161]													X
[83]										X			
[84]										X			
[89]												X	
[90]											X		
[91]	X												
[92]	X												
[149]				X									
[93]			X	X									
[94]											X	X	
[95]												X	
[96]													X
[150]												X	X
[103]									X				
[104]	X			X	X								
[152]												X	
[185]											X	X	
[114]											X	X	X
[115]												X	
[153]												X	
[120]			X	X									
[186]												X	
[123]	X		X	X		X	X						
[124]	X						X						
[125]	X						X						
[126]							X						
[127]	X					X	X			X			
[128]	X						X						
[129]	X						X						
[130]			X	X			X						
[132]												X	
[133]												X	
[134]												X	
[135]												X	
[136]												X	
[137]	X						X						
[138]									X				
[139]	X						X						

cloud data to form micro-Doppler signature data. The signature data is afterwards fed into a CNN to predict motion behavior.

There are two versions of tree-based hybrid models as addressed in [95], [128], [136], [186]. In the first version, the tree-based hybrid model either branches off to perform multiple processes simultaneously or fuses the output of several simultaneously performed processes to perform a single consecutive process. In addition to vision data fusion with a Bayesian filter, millimeter wave data is first clustered [95]. Simultaneous Localization And Mapping (SLAM) is solved by first clustering a raw radar scan to make it more sparse. Secondly, the scan is fused with odometry sensor data in a particle filter to estimate an odometry pose. Meanwhile, scan

matching with a reference scan is implemented on the sparse radar scan for retrieving new parameters used for updating the cluster model [136]. In the second tree-based hybrid model version, a model decision is taken based on a particular context at particular points in the tree. An air writing pipeline [128] tracks consecutive three dimensional locations with an $\alpha - \beta$ filter. The locations are determined via trilateration with range estimates coming from several radars. Afterwards, both a CNN and LSTM are tested on a character recognition task. Another model pipeline used for solving the SLAM problem [186] first chooses a triangulateration, angle/time or RSS model for both client and anchor node localization based on whether the environment map is known or not. Secondly, a decision between a RSS or angle/time model is made for estimating obstacle surfaces. Thirdly, the surface limits are determined either by analyzing the difference in RSS at AoAs situated at the limits or interaction of obstacle sides if the shape is two dimensional. Fourthly, clustering is used to reduce measurement bias due to noise. Lastly, an extended Kalman filter is deployed to improve the results.

4) Gaps and Challenges: An interesting observation derived from Tables XI and XII is that use of white box and shallow black box models in the context of millimeter wave sensing applications is restricted. White box models solely rely on physical and or mathematical knowledge for model construction. Several shallow black box models, such as SVMs and decision trees, rely on manually crafted data features for model construction [43], [78], [79], [147], [159]. Crafting features manually relies heavily on experience and knowledge of domain experts. The observation and information regarding white box and shallow black box model types suggest that there is ample room for research geared towards better understanding the millimeter wave sensing environment, i.e., ample room for analyzing millimeter wave sensing environment dynamics and uncertainties. Secondly, Table XII indicates that most research with deep black box models is restricted to CNNs and LSTMs. This leaves room for exploring the use of Temporal Convolutional Networks (TCNs) for modeling time dependency in raw data sample or data feature sequences. Lastly, Tables XI and XII indicate that the variety of models used in the millimeter wave sensing application pipelines is limited. Many application pipelines use CNNs, LSTMs, Bayesian filters, and or association/allocation. Therefore, increasing model variety and thus the knowledge on how analytical modeling in the millimeter wave sensing environment can be approached is a challenge that remains to be addressed.

E. Modeling Evaluation

In this section, we review evaluation metrics used to measure model performance of models employed in millimeter wave sensing applications as mentioned in Section IV-D. We also describe techniques used to improve data driven and grey box model training performance and evaluation.

1) Performance Evaluation Metrics: Table XIV summarizes performance metrics that are used in the reviewed papers. It can be seen that confusion matrix and accuracy are the most often used metrics in data driven model performance assessment. In training and validation of data driven models, the validation dataset is used to make choices about the model, including its hyperparameters [223]. After iteratively training and validating the model, the final model performance in relation to the application goal is evaluated on a held out test dataset consisting of samples that the model has not seen before. The outcome of the testing is compared to data labels and results are presented in the form of a confusion matrix and or metrics that summarize the content of a confusion matrix. The confusion matrix [78], [123] shows model classification performance for every class in the test dataset. The rows and columns represent ground truth and predicted classes respectively. Diagonal elements in the matrix denote the number, or fraction, of true positives for every class. When the diagonal elements are removed, the remaining row elements denote the false positives and remaining column elements the false negatives for every class. Positive samples are those that have been predicted to belong to a certain class. Negative samples are those that have been predicted to not belong to a certain class. Accuracy is defined as the overall proportion of predicted test dataset labels that match with the ground truth test dataset labels. Precision is a ratio of the number of true positive predictions to the number of all positive predictions made for a certain class. Recall is a ratio of the the number of true positive predictions to the number of all samples with a positive label in the ground truth dataset for a certain class [223]. Sometimes the inverse of accuracy and precision, the so called misclassification rate [43], [51], [94] and false discovery rate [189], [191], are also used. F1 score is the harmonic mean of precision and recall [123]. Specificity is a ratio of the number of true negative predictions to the number of all samples with a negative label in the ground truth dataset for a certain class [104]. The area under the Receiver Operating Characteristic (ROC) curve has also been used, which shows the degree of output separability [181]. A paper on gesture recognition used invalid gesture rate in conjunction with accuracy and misclassification rate to test data driven modeling performance for a set of gesture classes. Invalid gesture rate is defined as a ratio of samples classified as invalid gesture compared to the total number of samples in the test dataset [43]. Several model driven model results are also reported with the accuracy metric.

Model driven model performance is measured based on output comparison with a baseline, i.e., via an error metric with a well-established measurement technique. Error is defined as output deviation compared to the baseline and can be measured through various parameters such as Root Mean Square (RMS), Cramér-Rao Lower Bound (CRLB), mean, and standard deviation. CRLB is defined as the minimum achievable variance [152] of a certain parameter. A few data driven models were also tested with an error metric. Both data and model driven models sometimes rely on visual inspection to determine model performance. Visual inspection can be used to assess feature separability [83], [84], [127], [158] or model behavior [44], [55], [67], [96], [114], [150], [161].

2) Model Improvement and Evaluation Techniques: After developing mathematical proofs or conducting numerous quantitative experiments, physical and or mathematical

TABLE XV
SUMMARY OF MODEL IMPROVEMENT, OPTIMIZATION, AND EVALUATION TECHNIQUES USED TO IMPROVE AND EVALUATE DATA DRIVEN AND GREY BOX MODEL PERFORMANCE

	Regularization					Hyperparameter tuning			Cross validation			Training stabilization			
	Loss constraint	Dropout	Noise training	Weight decay	Early stopping	Manual	Grid search	Adaptive	K fold	Monte carlo	Leave one out	Batch normalization	Weight initialization	Architecture addition	Momentum
[158]	X					X		X							
[42]						X									
[43]						X									
[159]						X									
[45]						X									
[46]		X				X		X							
[47]	X	X				X	X	X	X						
[48]		X				X								X	
[50]							X		X						
[53]		X				X		X						X	
[174]						X		X							
[56]						X									
[63]		X				X		X				X	X		
[16]												X			
[181]				X		X		X							X
[192]						X		X				X	X		
[193]	X					X		X							X
[72]						X		X							
[76]						X		X							
[77]		X	X			X								X	
[78]	X					X	X		X						
[79]	X					X	X		X						
[91]						X			X						
[92]		X		X		X		X	X			X		X	
[93]						X		X							
[150]	X														
[103]		X			X	X		X							
[104]						X									
[120]				X		X		X					X	X	X
[123]						X		X		X		X		X	
[124]						X		X				X		X	
[125]		X				X		X	X						X
[126]		X				X		X					X		
[127]		X				X		X					X		
[128]		X		X	X	X		X				X	X		
[129]						X			X						
[130]		X				X		X							
[137]									X		X				
[138]		X										X		X	
[139]									X		X				

information required for solving a particular problem becomes known. Once this information is known, the white box model construction process for solving the problem is straightforward and the resulting model will exhibit good performance [227]. This is not the case for grey and black box models. Grey box models involve hidden state updates and or use of cost and similarity functions that depend on input data or features. This requires grey box model evaluation through numerous experiments every time the sensing application context, input data or feature distribution changes. As explained in Section IV-D.2, black box models are constructed through an iterative process in which a model is eventually created that performs well on a test dataset or feature set. This section elaborates on several problems that occur during grey box and black box model evaluation and black box model creation, and how these problems can be solved with special model improvement, evaluation, and or optimization techniques. These techniques are summarized in Table XV. Because new model creation is explained in Section IV-D, we do not focus on it in this section. Elements causing training instability are not further explained. Prediction denoising through Non-Maximum Suppression (NMS) or fusion performed by [181] is out of context in this section and is therefore not addressed. Next we explain main model improvement, optimization, and evaluation techniques used with data driven and grey box models in millimeter wave sensing applications.

a) Regularization: An issue frequently encountered during model training is model overfitting, in which model parameters and mapping function completely adapt to the training data or feature set. This will cause the model to exhibit sub-optimal performance on held-out (unseen) data. Various regularization techniques try to avoid overfitting and help the model to perform well on a wider variety of data it might encounter when deployed in specific applications [223]. For example, the ability of a cost function to minimize or a reward function to maximize itself can be limited with a special constraint [47], [78], [79], [150], [158], [193] and weight decay [92], [120], [128], [181] parameters added to the function. It can be decided to occasionally stop a parameter update for a limited number of model parameters in a particular training cycle through means of dropout [46]–[48], [53], [63], [77], [92], [103], [125]–[128], [130], [138]. In early stopping, the validation data or feature set can be used to inspect the model performance after every training cycle. In case the performance starts to diminish compared to previous training cycles, model training is stopped [103], [128]. A certain percentage of the training set can include noise samples to make the eventual model robust to input noise [77].

b) Hyperparameter tuning: Most data driven models involve some kind of manual hyperparameter tuning as shown in Table XV. Hyperparameters include number of layers, neurons, forest size, forest depth, number of training cycles,

number of data or feature inputs before an update is applied to model parameters, etc. Sometimes it is difficult to set the right initial hyperparameters due to a lack of a priori knowledge and experience. In this case, an exhaustive hyperparameter search using a grid search [47], [50], [78], [79] can be implemented, through which a complete grid of hyperparameter values is constructed. Afterwards, for every possible hyperparameter combination, a model is created and its performance is evaluated. The set of final hyperparameters is selected as the one that gives the best model performance. Not adapting the learning rate, i.e., a hyperparameter that controls how much parameter update is used to change the model parameters, during training cycles can result in long execution time for model training or training instability. To combat these problems, adaptive learning rate techniques such as Adaptive Moment (ADAM) estimation [46], [47], [53], [63], [76], [92], [103], [123], [124], [126]–[128], [130], [158] or manual learning rate decay [125], [174] can be employed.

c) Cross validation: Cross validation refers to training a model and evaluating model performance multiple times on a variety of different data or feature set splits. Splitting refers to partitioning a data or feature set into a subset designated for model training, a subset designated for model validation, and a subset designated for model testing. If model performance results are retrieved from a single training, validation, and test split, the model performance results will have a split specific bias. This means that the performance result will strongly deviate from the mean performance result that could be expected based on a given data or feature set. To combat this, model training, validation, and testing can be executed on a variety of splits for more robust performance result evaluation. This can be implemented very exhaustively with leave-one-out cross validation in which the train-validate-test split is repeated for every possible combination [137], [139]. Instead, methods such as Monte Carlo [123] and k-fold cross validation [47], [50], [78], [79], [91], [92], [125], [129], [137], [139] can be used in which the data or feature set splits are limited to a certain number.

d) Training stabilization: Elements like gradient exploding, gradient vanishing, internal covariate shift, noisy parameter updates, and bad parameter initialization have a negative impact on training stability. Training stability refers to an analytical model whose cost or reward function gradually minimizes or maximizes over time when new training samples are encountered during model training. Training instability examples include the cost or reward function prematurely becoming constant, a cost function that suddenly starts to increase with continuous acceleration, etc. Batch normalization [16], [63], [92], [123], [124], [128], [138], [192], initializing the model parameters prior to training with Xavier initialization [126]–[128], [192], making changes to existing model types with for example the ResNet layer [92], [123] or LeakyRELU activation function [48], [53], [77], [138], and parameter update averaging over multiple training cycles with momentum [120], [125], [181], [193] have been proposed to solve the problem of training instability.

3) Gaps and Challenges: A challenge often encountered in model evaluation is determining a set of metrics that can completely and objectively assess model performance. From Table XIV, it can be concluded that there is no such universal metric and none of the existing works covered all or even a majority of the given performance metrics. This, in combination with model variation, makes benchmarking between models difficult. A widely used metric that does not portray a complete performance picture is the accuracy metric. Accuracy cannot discriminate between a good or bad performing model in case data or feature sets are heavily skewed in terms of class distribution. More information regarding the selection of a suitable set of performance metrics can be found in [228]. Designing models that perform well in both lab and real-life contexts is still a major challenge. Model simulation [133], [184], [186], [201], controlled experiments [47], [83], [84], [104], and system challenges related to generalization [142] and impact [107] are frequently observed in scientific studies. Examples of controlled experiments and restricted experimental environment include test subjects performing activities directly in front of a radar [47] and object detection with very limited and pre-defined objects placed at specific positions in a radar's FOV [83], [84]. Table XV shows that many papers do not harness the strength of multiple regularization methods. It is reasonable to think that combining loss constraint, dropout, and or weight decay is not required or it is undesired because a single one of these methods may already result in good model performance on unseen data or they might negatively influence the training process [223]. Therefore, combining loss constraint, dropout, and or weight decay to improve model performance on unseen data is always subject to extensive evaluation. However, using early stopping in combination with other regularization methods can omit the need for exhaustively tuning the required number of training cycles to achieve optimal model performance while not negatively influencing the training process [223].

V. Challenges, Trends, and Future Perspective

This section presents and explains identified scientific and technological challenges and trends for applications of millimeter wave as a sensing technology. The challenges and trends have been categorized into several challenge and trend categories: hardware, unsupervised representation learning, support from other sensing domains, application integration, and crowd analysis. In addition, this section also provides a future perspective for applications of millimeter wave as a sensing technology.

A. Hardware

Several challenges related to the hardware used to collect data have an effect on how well a millimeter wave sensing methodology performs. Human vital sign monitoring pipelines are currently not reliable enough in the context of multiple humans that are all located at varying distances from each other and at various positions in reference to the data collection system. Issues include different relative error results compared to a baseline in the context of changing measurement positions (measuring in front or at a side of a test subject) [107], [147], impact of random body movement [62], [105], [144],

and measurement occlusion issues when humans are standing too close to each other [144], [147]. One study focused on finding the most optimal vital sign sensing position [202]. Several pipelines experience performance deterioration at large distances from the data collection system [79], [86], [97], [102], [113]–[115], [124], in the presence of solids such as occluding objects [93], [100], [102], [114] and brick/concrete wall barriers [79], [112], [113], not using 3D printed placement constraints [139], and in the context of hard to distinguish entities [78], [91], [103], [116], [149], [153]. A challenge in concealed object detection and pedestrian detection is the variable reflection intensity caused by different types of clothing people wear [189], [196]. One gesture recognition methodology in a car analyzed the impact of measurement position [42]. The optimal sensor placement was found to be on the center console (for use by front seat passenger and the driver) or in between the backs of the front seats (for use by the back-seat passengers). Performance deterioration was encountered when the sensor was placed too close to detectable objects such as the gear shift. Robustness against environmental effect analysis is limited to ego-motion estimation in [132]. Other challenges for manufacturers are to bring the cost of millimeter wave systems down to allow large deployments on a budget (e.g., for crowd analytics) and to improve the design and the form factor such that these systems do not cause architectural and aesthetic concerns for environments where they are deployed.

B. Unsupervised Representation Learning

Unsupervised representation learning has been used to pre-train a denoising autoencoder [174] in millimeter wave application pipelines. The network can be used afterwards to extract features from a dataset that generalize to a wide variety of end tasks [217]. Features are transferred from a pre-trained neural network to a supervised end learning task network. The networks used during pre-training and the end task share no learning task relation [229]. More information can be found in a review written by Bengio *et al.* [217]. In 2014, a new breakthrough in unsupervised pre-training was realized by Dosovitskiy *et al.* [230]. A concept called self-supervised learning was envisioned in which a supervised neural network tries to classify data transformations applied to unlabeled data in the pre-training stage. Since millimeter wave data is unlabeled too when it is sampled from the underlying hardware, it remains to be determined if unsupervised representation learning can reduce the amount of labeled data required for training end task deep learning models by using either unsupervised or self-supervised learning in a given pre-training stage. Millimeter wave datasets that can be used for experimentation include the UWCR radar mini, NuScenes, RadHAR, and Solinteraction datasets [47], [93], [139], [200]. However, the ability to explore is currently hindered by a lack of more datasets as indicated in [103] for pose estimation.

C. Support From Other Sensing Domains

Some pipelines apply sensor fusion of millimeter wave data with data originating from other sensing domains. Sensing domains include vision [51], [61], [66], [67], [71], [95], [105], [120], [146], depth [44], [105], lidar [71], inertial measurements [89], and exerted forces [91]. Data originating from these domains complement millimeter wave data and therefore cause more accurate analytical modeling performance in several situations. Analytical modeling performance in on-road detection and tracking suffers from limited spatial resolution of, and noise in, millimeter wave data [51], [61], [67], [95], [120]. Odometry information model performance also suffers from noise in millimeter wave data [89]. Data originating from other sensing domains have been used to guide a radar collection system to the most accurate vital sign sensing position [105] and have lead to increased analytical model performance [44], [66] and model state accuracy in object detection for visually impaired people [44]. Millimeter wave data also provide benefits to other sensing domains in return. The data are used for example to differentiate objects based on material, density or volume where exerted forces only measure similar object geometries [91]. In the future, fusion with other sensing domains can be extended. For example, when measuring crowd density in a limited space inside a building, fusion with heat energy obtained through a temperature sensor can be explored.

D. Application Integration

A variety of different sensing applications integrate different application types. Examples include combining communication and sensing [80], [140], combining vital sign, activity, and gesture sensing for more robust occupancy detection [81], SLAM [136], [154], [185], [186], detection/tracking and identification [46], [123], [143], and detection and activity recognition [77]. Detection and tracking make identification and activity recognition more robust in unknown environments. Research combining communication and sensing is limited to vision [80] and range simulations [140]. Due to performance deficiencies with communication-only access point sensing [80], sensing capabilities built into special access points are envisioned to enable better elderly monitoring and building analytics without extra costs related to installation of dedicated sensor hardware. The combination of communication and sensing in Vehicle to Vehicle (V2V) scenario's is envisioned to enable robust driver assistance systems [231]. Yassin *et al.* [186] denote that SLAM research with millimeter waves is still at its infancy.

E. Crowd Analysis

Throughout the review process no papers have been found that explore crowd analysis with millimeter wave sensing. The non-image based crowd counting review by Kouyoumdjieva *et al.* [232] is recommended as an introductory read since millimeter wave crowd analysis applications in the future will belong to the non-image based crowd analysis application category. Future millimeter wave crowd analysis research will include density/size, flow/trajectory/movement, and activity/behavior analysis. Density/size analysis refers to counting the number of people and their distribution in a given area. Kouyoumdjieva *et al.* [232] denote that for disaster management and city-wide public transportation crowd counting

tier models should be developed that have the ability to provide a macro crowd count based on local estimates calculated over different micro or meso areas. Flow/trajectory/movement analysis refers to determining major movement flows and directions in a given area, as well as identifying minimum and maximum bounds in terms of target numbers and dynamicity. Activity/behavior analysis refers to determining meso and macro level crowd activities, as well as identifying the correlation between location and activities and minimum and maximum bounds in terms of target numbers and granularity of activities. Major challenges for millimeter wave crowd analysis include methodology scalability to major events including thousands of people and operation security. The cost to deploy a grid of millimeter wave data collection systems is currently too high and deployed grids, in case the grid is not concealed in the environment, will interfere with decor design. Kouyoumdjieva *et al.* [232] denote that no non-image based crowd counting methodologies explore security in the form of robustness to output manipulation and malicious users trying to disable methodology functionality.

F. Future Perspective

We notice that most millimeter wave sensing research is currently limited to a personal lab scope. This means that research is most often performed with one person, a few people, one object or a few objects in simulated or controlled experiment environments. We believe that there is a lot of potential for research geared towards exploring millimeter wave sensing in large scale active and dynamic industrial and urban area's involving numerous people, other mammals, and or objects in the future.

Millimeter wave sensing infrastructures are unobtrusive in nature and will be deployed ubiquitously. Future millimeter wave infrastructures will also co-exist and collaborate with other sensing and communication infrastructures, and serve multiple sensing applications in parallel. For example, sensing capabilities will be integrated into special communication access points in the future to enable better elderly monitoring and building analytics without costs related to installation of dedicated sensor hardware [80]. Integration of communication and sensing in V2V scenario's will enable robust driver assistance systems [231].

Analytical modeling will become more prominent in millimeter wave sensing application pipelines. Jiang *et al.* [158] have presented an analytical model that has the ability to extract input data features that are environment and user specific information (i.e. domain) independent in the context of human activity detection. We believe that the ability to learn extraction of domain independent input data features is an important goal for future research in the context of analytical models that work with millimeter wave data. This ability will allow analytical models to be robust against influences from a variety of domains and therefore to perform well in real-life scenario's in the future. Analytical modeling research should be conducted to determine which domain influences can or cannot be mitigated, which domains can or cannot be integrated in an analytical model, how an analytical model can learn latent domains from the input data, and which analytical model types work best in a wide variety of different application types. For example, several hardware challenges can be regarded as a domain, rather than trying to mitigate hardware challenge influence and application robustness deterioration with cancellation methods [233] per challenge, using sensor fusion, or integrating application types. A variety of different models that extract domain independent input data features can be found in papers which do not consider millimeter wave sensing. For example, papers that consider WiFi CSI [234]–[236]. An important step in analytical modeling, due to the labor intensive nature of labelling millimeter wave data, is the development of unsupervised models that achieve performance that is on par with supervised counterparts. We believe that it is also important to investigate how learning domain independent input data feature extraction can be integrated with creation of analytical models through means of unsupervised or self-supervised representation learning.

Stimuli such as sound, light, stress, anxiety, overtraining, temperature, humidity, etc. have an effect on our health and vitality [237]. The COVID-19 pandemic and confinement measures taken during the pandemic result in stress and anxiety in a wide variety of different people [238]. We believe that this will spark debate in the general public about the effect of stress and anxiety and how we can reduce stress and anxiety in our daily lives in the near future. We also believe that this will stimulate research regarding the effect of stimuli on health and well-being with millimeter wave sensing systems. Vital sign sensing with millimeter wave systems can be used in niches where so called contact sensors, i.e., sensors involving electrodes, air analysis with nasal cannula or mask, a strap-on system, smart watch, smart phone, etc., [233], [239], cannot be used. Examples include, but are not limited to, skin irritation and allergic reactions, damaged skin (rashes, burns, hives, etc.), in the presence of clothes and obstacles [240], multiple beings, and humans that have certain behavioral conditions (e.g. severe autism, dementia, etc.).

VI. Conclusion

This is the first review that completely covers millimeter wave sensing application pipelines and pipeline building blocks in the form of a systematic literature review to the best of our knowledge. The millimeter wave technology covers a wide bandwidth and its short wavelength gives limited range. The limited range in turn provides low signal interference. This means transceivers can be packed very densely in an area without disrupting each others' communication signals. These properties of the technology not only yield high communication rates, but also provide a great opportunity for sub-millimeter accuracy level sensing of the surroundings, easily penetrating through simple obstacles like plastic and fabric.

Our analysis of the literature showed that there are indeed a variety of application types for millimeter sensing that we group into three domains; namely, human, object, and environment sensing. The application pipelines in the literature are made up of (a subset of) five common building blocks: data collection, pre-processing, feature extraction, analytical

modeling, and modeling evaluation. There is naming confusion in the literature in terms of which models and techniques take part in which building blocks. In this paper, we provided sharp descriptions of millimeter wave sensing building blocks; i.e., the hardware, algorithms, analytical models, and or model evaluation techniques that are covered by each block.

For different applications in the literature, each building block may select from a variety of models and techniques. A close look into many application instances reveals that a large majority of them stick to a combination of a few common solutions in their sensing pipelines. The rest of the models and techniques remain application-specific and their usage is not explored widely in other applications. This is not surprising as this field of research and development is still young and it is safer to rely on widely accepted solutions. For example, it is very common to utilize deep black box models employing CNNs and LSTM networks in the more frequently used data driven modeling, whereas Bayesian filters and association/allocation seem to be the first choices among model driven approaches. In terms of feature extraction, manual feature mapping is the predominant choice of researchers. We therefore encourage the researchers entering, or planning to conduct new research in, the millimeter wave sensing environment to try out new models. This will increase the model variety and thus the knowledge on how analytical modeling in the millimeter wave sensing environment can be approached.

REFERENCES

[1] H. Gold. (Mar. 2020). *Netflix and YouTube are Slowing Down in Europe to Keep the Internet From Breaking*. [Online]. Available: https://edition.cnn.com/2020/03/19/tech/netflix-internet-overload-eu/index.html

[2] J. Hendrickson. (Jan. 2020). *8K TV has Arrived Here's What You Need to Know*. [Online]. Available: https://www.howtogeek.com/397365/8k-tv-has-arrived.-heres-what-you-need-to-know/

[3] O. Soliman, A. Rezgui, H. Soliman, and N. Manea, "Mobile cloud gaming: Issues and challenges," in *Proc. Int. Conf. Mobile Web Inf. Sys.*, F. Daniel, G. A. Papadopoulos, and P. Thiran, Eds. Cham, Switzerland: Springer, 2013, pp. 121–128.

[4] Wireless LAN Working Group. (Dec. 2012). *802.11ad-2012 IEEE Standards Association*. [Online]. Available: https://standards.ieee.org/standard/802_11ad-2012.html

[5] Wireless LAN Working Group. (Oct. 2016). *802.11aj-2018 IEEE Standards Association*. [Online]. Available: https://standards.ieee.org/standard/802_11aj-2018.html

[6] Wireless LAN Working Group. (Feb. 2019). *P802.11ay IEEE Standards Association*. [Online]. Available: https://standards.ieee.org/project/802_11ay.html

[7] M. Elkashlan, T. Q. Duong, and H.-H. Chen, "Millimeter-wave communications for 5G: Fundamentals: Part I [Guest Editorial]," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 52–54, Sep. 2014.

[8] M. Elkashlan, T. Q. Duong, and H.-H. Chen, "Millimeter-wave communications for 5G—Part 2: Applications [Guest Editorial]," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 166–167, Jan. 2015.

[9] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, 2nd Quart., 2018.

[10] X. Wang *et al.*, "Millimeter wave communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1616–1653, 3rd Quart., 2018.

[11] Texas Instruments. (2010). *MmWave Radar Sensors—What is MmWave*. [Online]. Available: http://www.ti.com/sensors/mmwave/what-is-mmwave.html

[12] F. Khan and Z. Pi, "MmWave mobile broadband (MMB): Unleashing the 3–300GHz spectrum," in *Proc. 34th IEEE Sarnoff Symp.*, May 2011, pp. 1–6.

[13] I. F. Akyildiz, C. Han, and S. Nie, "Combating the distance problem in the millimeter wave and terahertz frequency bands," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 102–108, Jun. 2018.

[14] J. Lin and H. Hu. (Sep. 2017). *79GHz to Replace 24GHz for Automotive Millimeter-wave Radar Sensors*. [Online]. Available: https://www.digitimes.com/news/a20170906PD208.html

[15] D. C. B. Mariano, C. Leite, L. H. S. Santos, R. E. O. Rocha, and R. C. de Melo-Minardi, "A guide to performing systematic literature reviews in bioinformatics," 2017, *arXiv:1707.05813*. [Online]. Available: http://arxiv.org/abs/1707.05813

[16] Z. Zhang, Z. Tian, M. Zhou, W. Nie, and Z. Li, "Riddle: Real-time interacting with hand description via millimeter-wave sensor," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[17] Z. Li, Z. Yang, C. Song, C. Li, Z. Peng, and W. Xu, "E-eye: Hidden electronics recognition through mmWave nonlinear effects," in *Proc. 16th ACM Conf. Embedded Netw. Sensor Syst.*, Nov. 2018, pp. 68–81.

[18] B. Ozen, S. Baykut, O. Tulgar, A. U. Belgul, I. K. Yalcin, and D. S. Armagan Sahinkaya, "Foreign object detection on airport runways by mm-Wave FMCW radar," in *Proc. 25th Signal Process. Commun. Appl. Conf. (SIU)*, May 2017, pp. 1–4.

[19] S. López-Tapia, R. Molina, and N. Pérez de la Blanca, "Using machine learning to detect and localize concealed objects in passive millimeter-wave images," *Eng. Appl. Artif. Intell.*, vol. 67, pp. 81–90, Jan. 2018.

[20] V. K. Klochko, V. V. Strotov, and S. A. Smirnov, "Multiple objects detection and tracking in passive scanning millimeter-wave imaging systems," in *Proc. SPIE*, vol. 11164, Oct. 2019, Art. no. 111640E.

[21] S. Yeom, D.-S. Lee, J.-Y. Son, and S.-H. Kim, "Concealed object detection using passive millimeter wave imaging," in *Proc. 4th Int. Universal Commun. Symp.*, Oct. 2010, pp. 383–386.

[22] L. Guo and S. Qin, "High-performance detection of concealed forbidden objects on human body with deep neural networks based on passive millimeter wave and visible imagery," *J. Infr., Millim., THz Waves*, vol. 40, no. 3, pp. 314–347, Mar. 2019.

[23] S. E. Clark, J. A. Lovberg, C. A. Martin, and J. A. Galliano, Jr., "Passive millimeter-wave imaging for concealed object detection," *Proc. SPIE Sensors, Command, Control, Commun. Intell. Technol. Homeland Defense Law Enforcement*, vol. 4708, pp. 128–133, Aug. 2002.

[24] B. Kapilevich, B. Litvak, A. Shulzinger, and M. Einat, "Portable passive millimeter-wave sensor for detecting concealed weapons and explosives hidden on a human body," *IEEE Sensors J.*, vol. 13, no. 11, pp. 4224–4228, Nov. 2013.

[25] S. Yeom, D.-S. Lee, Y. Jang, M.-K. Lee, and S.-W. Jung, "Real-time concealed-object detection and recognition with passive millimeter wave imaging," *Opt. Exp.*, vol. 20, pp. 9371–9381, Apr. 2012.

[26] L. Li, J. Yang, G. Cui, Z. Jiang, and X. Zheng, "Method of passive MMW image detection and identification for close target," *J. Infr., Millim., THz Waves*, vol. 32, no. 1, pp. 102–115, Jan. 2011.

[27] L. Yujiri, B. I. Hauss, and M. Shoucri, "Passive millimeter wave sensors for detection of buried mines," *Detection Technol. Mines Minelike Targets*, vol. 2496, pp. 2–6, Jun. 1995.

[28] H. Isiker, S. Demirci, B. Yilmaz, S. Gokkan, and C. Ozdemir, "Detection of small and large hidden metallic objects via passive millimeter wave imaging system with an auto-segmentation routine," in *Proc. Prog. Electromagn. Res. Symp. (PIERS-Toyama)*, Aug. 2018, pp. 1362–1365.

[29] V. Mattioli, L. Milani, K. M. Magde, G. A. Brost, and F. S. Marzano, "Retrieval of sun brightness temperature and precipitating cloud extinction using ground-based sun-tracking microwave radiometry," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 7, pp. 3134–3147, Jul. 2017.

[30] J. A. Nanzer, E. Popova, and R. L. Rogers, "Analysis of the detection modes of a human presence detection millimeter-wave radiometer," in *Proc. IEEE Antennas Propag. Soc. Int. Symp.*, Jul. 2010, pp. 1–4.

[31] H. Zong, L. Bao, B. Liu, and J. Qiu, "Application of convolutional neural network in target detection of millimeter wave imaging," in *Proc. IEEE Int. Symp. Antennas Propag. USNC/URSI Nat. Radio Sci. Meeting*, Jul. 2018, pp. 1217–1218.

[32] C. R. Cabrera-Mercader and D. H. Staelin, "Passive microwave humidity profile retrievals using neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Aug. 1994, pp. 2057–2059.

[33] B. Kapilevich, B. Litvak, and A. Shulzinger, "Passive non-imaging mm-Wave sensor for detecting hidden objects," in *Proc. IEEE Int. Conf. Microw., Commun., Antennas Electron. Syst. (COMCAS)*, Oct. 2013, pp. 1–5.

[34] Y. Meng, A. Qing, C. Lin, J. Zang, Y. Zhao, and C. Zhang, "Passive millimeter wave imaging system based on helical scanning," *Sci. Rep.*, vol. 8, no. 1, May 2018, Art. no. 7852.

[35] K. Schmalz, N. Rothbart, P. F.-X. Neumaier, J. Borngraber, H.-W. Hubers, and D. Kissinger, "Gas spectroscopy system for breath analysis at mm-Wave/THz using SiGe BiCMOS circuits," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 5, pp. 1807–1818, May 2017.

[36] F. P. Schloerb, "Millimeter-wave spectroscopy of solar system objects: Present and future," in *Proc. Eur. Southern Observatory Conf. Workshop*, vol. 22, P. A. Shaver and K. Kjar, Eds., Jan. 1985, pp. 603–615.

[37] E. S. Gonçalves, F. C. Teixeira, D. F. Albuquerque, and E. F. Pedrosa, "Asynchronous mmWave radar interference for indoor intrusion detection," in *Proc. 4th Iberian Robot. Conf.*, M. F. Silva, J. L. Lima, L. P. Reis, A. Sanfeliu, and D. Tardioli, Eds. Cham, Switzerland: Springer, 2020, pp. 367–378.

[38] H. Rodilla, A. A. Kim, G. D. M. Jeffries, J. Vukusic, A. Jesorka, and J. Stake, "Millimeter-wave sensor based on a $\lambda/2$-line resonator for identification and dielectric characterization of non-ionic surfactants," *Sci. Rep.*, vol. 6, no. 1, Jan. 2016, Art. no. 19523.

[39] S. Sano, A. Tsuzuki, J. Li, A. Gotou, Y. Makino, and S. Miyake, "Millimeter-wave dielectric measurement of sic powders as a basis of millimeter-wave sintering of ceramics," in *Proc. Int. Symp. Novel Mater. Process. Adv. Electromagn. Energy Sour.* Amsterdam, The Netherlands: Elsevier, 2005, pp. 151–154.

[40] M. Ghasr, S. Kharkovsky, R. Zoughi, and R. Austin, "Comparison of near-field millimeter wave probes for detecting corrosion pit under paint," in *Proc. 21st IEEE Instrum. Meas. Technol. Conf.*, May 2004, pp. 2240–2244.

[41] D. M. Sheen, D. L. McMakin, and T. E. Hall, *Chapter 9—Detection Explosives by Millimeter-Wave Imaging*. Amsterdam, The Netherlands: Elsevier, 2007, pp. 237–277.

[42] K. A. Smith, C. Csech, D. Murdoch, and G. Shaker, "Gesture recognition using mm-Wave sensor for human-car interface," *IEEE Sensors Lett.*, vol. 2, no. 2, pp. 1–4, Jun. 2018.

[43] C. Liu, Y. Li, D. Ao, and H. Tian, "Spectrum-based hand gesture recognition using millimeter-wave radar parameter measurements," *IEEE Access*, vol. 7, pp. 79147–79158, 2019.

[44] N. Long, K. Wang, R. Cheng, W. Hu, and K. Yang, "Unifying obstacle detection, recognition, and fusion based on millimeter wave radar and RGB-depth sensors for the visually impaired," *Rev. Sci. Instrum.*, vol. 90, no. 4, Apr. 2019, Art. no. 044102.

[45] J. Lien *et al.*, "Soli : Ubiquitous gesture sensing with millimeter wave radar," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–19, Jul. 2016.

[46] P. Zhao *et al.*, "MID: Tracking and identifying people with millimeter wave radar," in *Proc. 15th Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, May 2019, pp. 33–40.

[47] 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 Millim. Netw. Sens. Syst.*, 2019, pp. 51–56.

[48] R. Zhang and S. Cao, "Real-time human motion behavior detection via CNN using mmWave radar," *IEEE Sensors Lett.*, vol. 3, no. 2, Feb. 2019, Art. no. 3500104.

[49] K. Diederichs, A. Qiu, and G. Shaker, "Wireless biometric individual identification utilizing millimeter waves," *IEEE Sensors Lett.*, vol. 1, no. 1, Feb. 2017, Art. no. 3500104.

[50] M. Alizadeh, H. Abedi, and G. Shaker, "Low-cost low-power in-vehicle occupant detection with mm-Wave FMCW radar," in *Proc. IEEE SENSORS*, Oct. 2019, pp. 2–5.

[51] X. Wang, L. Xu, H. Sun, J. Xin, and N. Zheng, "On-road vehicle detection and tracking using MMW radar and monovision fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 7, pp. 2075–2084, Jul. 2016.

[52] G. Zhai, C. Wu, and Y. Wang, "Millimeter wave radar target tracking based on adaptive Kalman filter," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 453–458.

[53] F. Jin *et al.*, "Multiple patients behavior detection in real-time using mmWave radar and deep CNNs," in *Proc. IEEE Radar Conf. (RadarConf)*, Apr. 2019, pp. 1–6.

[54] O. Boric-Lubecke, J. Lin, V. M. Lubecke, A. Host-Madsen, and T. Sizer, "Microwave and millimeter-wave Doppler radar heart sensing," *Proc. SPIE on Radar Sensor Technol.*, vol. 6547, May 2007, Art. no. 65470C.

[55] Texas Instruments. (2020). *People Tracking and Counting Reference Design Using mmWave Radar Sensor*. [Online]. Available: https://www.ti.com/tool/TIDEP-01000

[56] A. Patra, P. Geuer, A. Munari, and P. Mähönen, "Mm-Wave radar based gesture recognition: Development and evaluation of a low-power, low-complexity system," in *Proc. 2nd ACM Workshop Millim. Wave Netw. Sens. Syst.*, Oct. 2018, pp. 51–56.

[57] M. Alizadeh, G. Shaker, J. C. M. D. Almeida, P. P. Morita, and S. Safavi-Naeini, "Remote monitoring of human vital signs using mm-Wave FMCW radar," *IEEE Access*, vol. 7, pp. 54958–54968, 2019.

[58] T. Horiuchi, J. Konishi, H. Yamada, and S. Muramatsu, "Indoor human tracking with millimeter-wave minimum redundancy MIMO radar," in *Proc. Int. Symp. Antennas Propag.*, 2019, pp. 2–3.

[59] S. Bakhtiari *et al.*, "Compact millimeter-wave sensor for remote monitoring of vital signs," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, pp. 830–841, Mar. 2012.

[60] D. T. Petkie, C. Benton, and E. Bryan, "Millimeter-wave radar for vital signs sensing," *Proc. SPIE Radar Sensor Technol.*, vol. 7308, Apr. 2009, Art. no. 73080A.

[61] G. Zhai, C. Wu, and Y. Wang, "Target tracking based on millimeter wave radar in complex scenes," *Int. J. Performability Eng.*, vol. 14, no. 2, pp. 232–244, 2018.

[62] A. Ahmad, J. C. Roh, D. Wang, and A. Dubey, "Vital signs monitoring of multiple people using a FMCW millimeter-wave sensor," in *Proc. IEEE Radar Conf.*, Apr. 2018, pp. 1450–1455.

[63] S. Hazra and A. Santra, "Robust gesture recognition using millimetric-wave radar system," *IEEE Sensors Lett.*, vol. 2, no. 4, Dec. 2018, Art. no. 7001804.

[64] M. Raja, Z. Vali, S. Palipana, D. G. Michelson, and S. Sigg, "3D head motion detection using millimeter-wave Doppler radar," *IEEE Access*, vol. 8, pp. 32321–32331, 2020.

[65] S. Diebold *et al.*, "A W-band MMIC radar system for remote detection of vital signs," *J. Infr., Millim., THz Waves*, vol. 33, no. 12, pp. 1250–1267, Dec. 2012.

[66] S. M. Kwon *et al.*, "Hands-free human activity recognition using millimeter-wave sensors," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Nov. 2019, pp. 1–2.

[67] W. Huang, Z. Zhang, W. Li, and J. Tian, "Moving object tracking based on millimeter-wave radar and vision sensor," *Appl. Sci. Eng.*, vol. 21, no. 4, pp. 609–614, 2018.

[68] B. Kapilevich and M. Einat, "Detecting hidden objects on human body using active millimeter wave sensor," *IEEE Sensors J.*, vol. 10, no. 11, pp. 1746–1752, Nov. 2010.

[69] E. Al-Masri and M. Momin, "Detecting heart rate variability using millimeter-wave radar technology," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2018, pp. 5282–5284.

[70] M. Z. Ikram, A. Ahmad, and D. Wang, "High-accuracy distance measurement using millimeter-wave radar," in *Proc. IEEE Radar Conf.*, Apr. 2018, pp. 1296–1300.

[71] A. A. Belyaev, I. O. Frolov, T. A. Suanov, and D. O. Trots, "Object detection in an urban environment using 77GHz radar," in *Proc. Radiat. Scattering Electromagn. Waves (RSEMW)*, Jun. 2019, pp. 436–439.

[72] T. Akita and S. Mita, "Object tracking and classification using millimeter-wave radar based on LSTM," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 1110–1115.

[73] S. Matsuguma and A. Kajiwara, "Bathroom accident detection with 79GHz-band millimeter wave sensor," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Mar. 2019, pp. 1–5.

[74] I. Ünal and S. Eker, "Investigations on millimeter wave detection of power lines from a safe distance," in *Proc. 10th Int. Conf. Electr. Electron. Eng.*, 2018, pp. 964–967.

[75] S. Björklund, H. Petersson, A. Nezirovic, M. B. Guldogan, and F. Gustafsson, "Millimeter-wave radar micro-Doppler signatures of human motion," in *Proc. 12th Int. Radar Symp.*, 2011, pp. 167–174.

[76] Y. Sun, R. Hang, Z. Li, M. Jin, and K. Xu, "Privacy-preserving fall detection with deep learning on mmWave radar signal," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2019.

[77] P. Kaushik, "Radar as a security measure–real time neural model based human detection and behaviour classification," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2019, pp. 1–5.

[78] S. Bjorklund, T. Johansson, and H. Petersson, "Evaluation of a micro-Doppler classification method on mm-Wave data," in *Proc. IEEE Radar Conf.*, May 2012, pp. 934–939.

[79] D. P. Fairchild and R. M. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 8, no. 5, pp. 425–434, Jun. 2014.

[80] M. Alloulah and H. Huang, "Future millimeter-wave indoor systems: A blueprint for joint communication and sensing," *Computer*, vol. 52, no. 7, pp. 16–24, Jul. 2019.

[81] A. Santra, R. V. Ulaganathan, and T. Finke, "Short-range millimetric-wave radar system for occupancy sensing application," *IEEE Sensors Lett.*, vol. 2, no. 3, Sep. 2018, Art. no. 7000704.

[82] T.-Y.-J. Kao and J. Lin, "Vital sign detection using 60-GHz Doppler radar system," in *Proc. IEEE Int. Wireless Symp. (IWS)*, Apr. 2013, pp. 1–4.

[83] W. Baoshuai and Z. Wei, "FOD detection based on millimeter wave radar using higher order statistics," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC)*, Oct. 2017, pp. 1–4.

[84] W. Baoshuai, L. Jianghong, Z. Xiaoliang, and H. Minjue, "A novel hierarchical foreign object debris detection method for millimeter wave radar," in *Proc. Int. Appl. Comput. Electromagn. Soc. Symp.*, 2017, pp. 1–2.

[85] G. Mehdi and J. Miao, "Millimeter wave FMCW radar for foreign object debris (FOD) detection at airport runways," in *Proc. 9th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2012, pp. 222–225.

[86] P. Feil, W. Menzel, T. P. Nguyen, C. Pichot, and C. Migliaccio, "Foreign objects debris detection (FOD) on airport runways using a broadband 78 GHz sensor," in *Proc. 38th Eur. Microw. Conf.*, Oct. 2008, pp. 1608–1611.

[87] B. D. Pollard and G. Sadowy, "Next generation millimeter-wave radar for safe planetary landing," in *Proc. IEEE Aerosp. Conf.*, Mar. 2005, pp. 1213–1219.

[88] R. Battaglia, M. Ferri, V. Dainelli, F. Sarullo, and M. Demeo, "Warden: W-band advanced radar for debris early notification form ISS," in *Proc. IEEE Aerosp. Conf.*, vol. 1, Aug. 2003, pp. 1–81.

[89] Y. Almalioglu, M. Turan, C. Xiaoxuan Lu, N. Trigoni, and A. Markham, "Milli-RIO: Ego-motion estimation with low-cost millimetre-wave radar," 2019, *arXiv:1909.05774*. [Online]. Available: http://arxiv.org/abs/1909.05774

[90] X. Tang, X. Wu, S. B. Yeap, R. Luo, T. Dai, and L. Huang, "Experimental results of target classification using mm wave corner radar sensors," in *Proc. Asia–Pacific Microw. Conf. (APMC)*, Nov. 2018, pp. 842–844.

[91] Z. Flintoff, B. Johnston, and M. Liarokapis, "Single-grasp, model-free object classification using a hyper-adaptive hand, Google soli, and tactile sensors," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 1943–1950.

[92] A. Angelov, A. Robertson, R. Murray-Smith, and F. Fioranelli, "Practical classification of different moving targets using automotive radar and deep neural networks," *IET Radar, Sonar Navigat.*, vol. 12, no. 10, pp. 1082–1089, Oct. 2018.

[93] X. Gao, G. Xing, S. Roy, and H. Liu, "Experiments with mmWave automotive radar test-bed," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, Nov. 2019, pp. 1–6.

[94] A. Antonucci *et al.*, "Performance analysis of a 60-GHz radar for indoor positioning and tracking," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2019, pp. 1–7.

[95] R. Zhang and S. Cao, "Extending reliability of mmWave radar tracking and detection via fusion with camera," *IEEE Access*, vol. 7, pp. 137065–137079, 2019.

[96] Texas Instruments. (May 2019). *TIDEP-01018: Automated Doors Reference Design Using mmWave Sensors*. [Online]. Available: http://www.ti.com/lit/pdf/tiduer1

[97] Texas Instruments. (Aug. 2018). *TIDEP-01003: Zone Occupancy Detection Reference Design Using mmWave Sensor*. [Online]. Available: http://www.ti.com/lit/pdf/tiduea7

[98] Texas Instruments. (Apr. 2018). *TIDEP-01001: Vehicle Occupant Detection Reference Design*. [Online]. Available: http://www.ti.com/lit/pdf/tidue95

[99] M. Kishida, K. Ohguchi, and M. Shono, "79 GHz-band high-resolution millimeter-wave radar," *FUJITSU Sci. Tech. J.*, vol. 51, no. 4, pp. 55–59, Oct. 2015.

[100] A. Etinger, N. Balal, B. Litvak, M. Einat, B. Kapilevich, and Y. Pinhasi, "Non-imaging mm-Wave FMCW sensor for pedestrian detection," *IEEE Sensors J.*, vol. 14, no. 4, pp. 1232–1237, Apr. 2014.

[101] M. Leonardi, E. G. Piracci, and V. Fastella, "W-band multi-radar processing for airport foreign object debris and humans detection," in *Proc. Eur. Microw. Conf. Central Eur.*, vol. 2019, pp. 285–288.

[102] K. A. Gallagher and R. M. Narayanan, "Human detection and ranging at long range and through light foliage using a W-band noise radar with an embedded tone," in *Proc. SPIE Radar Sensor Technol.*, vol. 8714, K. I. Ranney and A. Doerry, Eds., 2013, pp. 1–12.

[103] A. Sengupta, F. Jin, R. Zhang, and S. Cao, "Mm-pose: Real-time human skeletal posture estimation using mmWave radars and CNNs," *IEEE Sensors J.*, vol. 20, no. 17, pp. 10032–10044, Sep. 2020.

[104] S. Caorsi and C. Lenzi, "Can a mm-Wave ultra-wideband ANN-based radar data processing approach be used for breast cancer detection?" in *Proc. Int. Conf. Electromagn. Adv. Appl. (ICEAA)*, Sep. 2017, pp. 1236–1239.

[105] I. V. Mikhelson, P. Lee, S. Bakhtiari, T. W. Elmer, A. K. Katsaggelos, and A. V. Sahakian, "Noncontact millimeter-wave real-time detection and tracking of heart rate on an ambulatory subject," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 5, pp. 927–934, Sep. 2012.

[106] I. Walterscheid, O. Biallawons, and P. Berens, "Contactless respiration and heartbeat monitoring of multiple people using a 2-D imaging radar," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 3720–3725.

[107] S. Wang *et al.*, "A novel ultra-wideband 80 GHz FMCW radar system for contactless monitoring of vital signs," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 4978–4981.

[108] M. Alizadeh, G. Shaker, and S. Safavi-Naeini, "Remote heart rate sensing with mm-Wave radar," in *Proc. 18th Int. Symp. Antenna Technol. Appl. Electromagn. (ANTEM)*, Aug. 2018, pp. 1–2.

[109] H.-R. Chuang, H.-C. Kuo, F.-L. Lin, T.-H. Huang, C.-S. Kuo, and Y.-W. Ou, "60-GHz millimeter-wave life detection system (MLDS) for noncontact human vital-signal monitoring," *IEEE Sensors J.*, vol. 12, no. 3, pp. 602–609, Mar. 2012.

[110] S. Li *et al.*, "A new kind of non-acoustic speech acquisition method based on millimeter waveradar," *Prog. Electromagn. Res.*, vol. 130, pp. 17–40, 2012.

[111] Z.-W. Li, "Millimeter wave radar for detecting the speech signal applications," *Int. J. Infr. Millim. Waves*, vol. 17, no. 12, pp. 2175–2183, Dec. 1996.

[112] M. Jiao, G. Lu, X. Jing, S. Li, Y. Li, and J. Wang, "A novel radar sensor for the non-contact detection of speech signals," *Sensors*, vol. 10, no. 5, pp. 4622–4633, May 2010.

[113] S. Li *et al.*, "A 94-GHz millimeter-wave sensor for speech signal acquisition," *Sensors*, vol. 13, no. 11, pp. 14248–14260, Oct. 2013.

[114] Texas Instruments. (Sep. 2019). *TIDEP-0090: Traffic Monitoring Object Detection and Tracking Reference Design Using Single-Chip mmWave Radar Sensor*. [Online]. Available: http://www.ti.com/lit/pdf/tidud31

[115] S. Dogru, R. Baptista, and L. Marques, "Tracking drones with drones using millimeter wave radar," in *Proc. 4th Iberian Robot. Conf.*, M. F. Silva, J. Luís Lima, L. P. Reis, A. Sanfeliu, and D. Tardioli, Eds. Cham, Switzerland: Springer, 2020, pp. 392–402.

[116] A. K. Singh and Y. H. Kim, "Accurate measurement of drone's blade length and rotation rate using pattern analysis with W-band radar," *Electron. Lett.*, vol. 54, no. 8, pp. 523–525, Apr. 2018.

[117] M. Caris, W. Johannes, S. Sieger, V. Port, and S. Stanko, "Detection of small UAS with W-band radar," in *Proc. 18th Int. Radar Symp. (IRS)*, Jun. 2017, pp. 1–6.

[118] S. Haefner *et al.*, "Contribution to drone detection by exploiting parameter estimation for a prototype mm-Wave radar system," in *Proc. 22nd Int. ITG Workshop Smart Antennas*, 2018, pp. 1–8.

[119] S. Rahman and D. A. Robertson, "Radar micro-Doppler signatures of drones and birds at K-band and W-band," *Sci. Rep.*, vol. 8, no. 1, Nov. 2018, Art. no. 17396.

[120] S. Chang *et al.*, "Spatial attention fusion for obstacle detection using mmWave radar and vision sensor," *Sensors*, vol. 20, no. 4, p. 956, Feb. 2020.

[121] Texas Instruments. (Jun. 2017). *TIDEP-0094: 80-m Range Object Detection With IWR1642 mmWave Sensor Reference Design*. [Online]. Available: http://www.ti.com/lit/pdf/tidud93

[122] K. Stasiak, M. Ciesielski, P. Samczynski, D. Gromek, and K. Kulpa, "Preliminary results of Drone's propellers detection using K-band and mm-Wave FMCW radar," in *Proc. 20th Int. Radar Symp. (IRS)*, Jun. 2019, pp. 1–7.

[123] P. Janakaraj, K. Jakkala, A. Bhuyan, Z. Sun, P. Wang, and M. Lee, "STAR: Simultaneous tracking and recognition through millimeter waves and deep learning," in *Proc. 12th IFIP Wireless Mobile Netw. Conf. (WMNC)*, Sep. 2019, pp. 211–218.

[124] Y. Liu, Y. Wang, H. Liu, A. Zhou, J. Liu, and N. Yang, "Long-range gesture recognition using millimeter wave radar," 2020, *arXiv:2002.02591*. [Online]. Available: http://arxiv.org/abs/2002.02591

[125] S. Wang, J. Song, J. Lien, I. Poupyrev, and O. Hilliges, "Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum," in *Proc. 29th Annu. Symp. User Interface Softw. Technol.*, Oct. 2016, pp. 851–860.

[126] S. Hazra and A. Santra, "Radar gesture recognition system in presence of interference using self-attention neural network," in *Proc. 18th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2019, pp. 1409–1414.

[127] S. Hazra and A. Santra, "Short-range radar-based gesture recognition system using 3D CNN with triplet loss," *IEEE Access*, vol. 7, pp. 125623–125633, 2019.

[128] M. Arsalan and A. Santra, "Character recognition in air-writing based on network of radars for human-machine interface," *IEEE Sensors J.*, vol. 19, no. 19, pp. 8855–8864, Oct. 2019.

[129] Y. Sun, T. Fei, F. Schliep, and N. Pohl, "Gesture classification with handcrafted micro-Doppler features using a FMCW radar," in *Proc. IEEE MTT-S Int. Conf. Microw. Intell. Mobility (ICMIM)*, Apr. 2018, pp. 1–4.

[130] Y. Sun, T. Fei, S. Gao, and N. Pohl, "Automatic radar-based gesture detection and classification via a region-based deep convolutional neural network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 4300–4304.

[131] J. Wang, X. Geng, and S. Wei, "Airport runway FOD detection system based on 77GHz millimeter wave radar sensor," in *Proc. IEEE Int. Conf. Integr. Circuits, Technol. Appl. (ICTA)*, Nov. 2019, pp. 140–143.

[132] S. H. Cen and P. Newman, "Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 6045–6052.

[133] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using Doppler radar," in *Proc. 16th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2013, pp. 869–874.

[134] M. Rapp, M. Barjenbruch, M. Hahn, J. Dickmann, and K. Dietmayer, "Probabilistic ego-motion estimation using multiple automotive radar sensors," *Robot. Auto. Syst.*, vol. 89, pp. 136–146, Mar. 2017.

[135] E. Ward and J. Folkesson, "Vehicle localization with low cost radar sensors," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 864–870.

[136] F. Schuster, M. Worner, C. G. Keller, M. Haueis, and C. Curio, "Robust localization based on radar signal clustering," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 839–844.

[137] H.-S. Yeo, G. Flamich, P. Schrempf, D. Harris-Birtill, and A. Quigley, "RadarCat: Radar categorization for input & interaction," in *Proc. 29th Annu. Symp. User Interface Softw. Technol.*, Oct. 2016, pp. 833–841.

[138] Z. Li, Z. Lei, A. Yan, E. Solovey, and K. Pahlavan, "ThuMouse: A micro-gesture cursor input through mmWave radar-based interaction," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2020, pp. 1–9.

[139] H.-S. Yeo, R. Minami, K. Rodriguez, G. Shaker, and A. Quigley, "Exploring tangible interactions with radar sensing," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 4, Dec. 2018, Art. no. 200.

[140] K. Ammar, O. Ben Haj Belkacem, and R. Bouallegue, "DSSS transmission technique to joint radar sensing and wireless communications in V2V system," in *Proc. Workshop Web, Artif. Intell. Netw. Appl.*, L. Barolli, F. Amato, F. Moscato, T. Enokido, and M. Takizawa, Eds. Cham, Switzerland: Springer, 2020, pp. 376–385.

[141] L. Shaobei and L. Shiyong, "Target detection and recognition based on active millimeter-wave imaging system," in *Proc. IEEE 2nd Int. Conf. Electron. Technol. (ICET)*, May 2019, pp. 74–77.

[142] S. Agarwal, D. Singh, and N. P. Pathak, "Active millimeter wave radar system for non-destructive, non-invasive underline fault detection and multilayer material analysis," in *Proc. IEEE Int. Microw. RF Conf. (IMaRC)*, Dec. 2014, pp. 369–372.

[143] T. Gu, Z. Fang, Z. Yang, P. Hu, and P. Mohapatra, "MmSense: Multi-person detection and identification via mmWave sensing," in *Proc. 3rd ACM Workshop Millim.-Wave Netw. Sens. Syst.*, 2019, pp. 45–50.

[144] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra, "Monitoring vital signs using millimeter wave," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jul. 2016, pp. 211–220.

[145] J. Palacios, G. Bielsa, P. Casari, and J. Widmer, "Single- and multiple-access point indoor localization for millimeter-wave networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1927–1942, Mar. 2019.

[146] M. T. Ortiz, H. Groll, E. Zochmann, and C. F. Mecklenbraucker, "Vehicle tracking through vision-millimeter wave Doppler shift fusion," in *Proc. IEEE-APS Topical Conf. Antennas Propag. Wireless Commun. (APWC)*, Sep. 2019, pp. 359–362.

[147] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra, "Vital sign and sleep monitoring using millimeter wave," *ACM Trans. Sensor Netw.*, vol. 13, no. 2, pp. 1–32, Jun. 2017.

[148] Y. Zeng, P. H. Pathak, Z. Yang, and P. Mohapatra, "Poster abstract: Human tracking and activity monitoring using 60 GHz mmWave," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2016, pp. 1–2.

[149] K. Watabe, K. Shimizu, K. Mizuno, and M. Yoneyama, "Millimeter-wave imaging using neural networks for object recognition," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1996, pp. 1135–1138.

[150] A. Patra, L. Simic, and M. Petrova, "MmRTI: Radio tomographic imaging using highly-directional millimeter-wave devices for accurate and robust indoor localization," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–7.

[151] S. Oka, S. Mochizuki, H. Togo, and N. Kukutsu, "Inspection of concrete structures using millimeter-wave imaging technology," *NTT Tech. Rev.*, vol. 7, no. 3, 2009. [Online]. Available: https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200903sf4.pdf

[152] J. O. Schrattenecker, S. Schuster, A. Haderer, G. Reinthaler, and A. Stelzer, "Accuracy limits of a seam-tracking algorithm for microwave systems at mm-Wav frequencies," in *Proc. 21st Eur. Signal Process. Conf.*, 2013, pp. 1–5.

[153] Y. Zhu, Y. Zhu, B. Y. Zhao, and H. Zheng, "Reusing 60GHz radios for mobile radar imaging," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2015, pp. 103–116.

[154] M. Aladsani, A. Alkhateeb, and G. C. Trichopoulos, "Leveraging mmWave imaging and communications for simultaneous localization and mapping," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 4539–4543.

[155] H. Ajorloo, C. J. Sreenan, A. Loch, and J. Widmer, "On the feasibility of using IEEE 802.11ad mmWave for accurate object detection," in *Proc. 34th ACM/SIGAPP Symp. Appl. Comput.*, Apr. 2019, pp. 2406–2413.

[156] L. Wang, "3D holographic millimeter-wave imaging for concealed metallic forging objects detection," in *Emerging Microwave Technologies in Industrial, Agricultural, Medical and Food Processing*. London, U.K.: IntechOpen, Jul. 2018, pp. 125–139.

[157] C. Li, J. Wang, D. Rodriguez, A. Mishra, Z. Peng, and Y. Li, "Portable Doppler/FSK/FMCW radar systems for life activity sensing and human localization," in *Proc. 14th Int. Conf. Adv. Technol., Syst. Services Telecommun. (TELSIKS)*, Oct. 2019, pp. 83–93.

[158] W. Jiang *et al.*, "Towards environment independent device free human activity recognition," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw.*, Oct. 2018, pp. 289–304.

[159] T. Wei and X. Zhang, "MTrack: High-precision passive tracking using millimeter wave radios," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2015, pp. 117–129.

[160] A. Zhou, S. Yang, Y. Yang, Y. Fan, and H. Ma, "Autonomous environment mapping using commodity millimeter-wave network device," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2019, pp. 1126–1134.

[161] B. D. Pollard, G. Sadowy, D. Moller, and E. Rodriguez, "A millimeter-wave phased array radar for hazard detection and avoidance on planetary landers," in *Proc. IEEE Aerosp. Conf. Process.*, May 2003, pp. 1115–1122.

[162] A. P. Toda and F. De Flaviis, "Mm-Wave motion tracking system using beamforming antennas," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Jul. 2014, pp. 105–106.

[163] C. Wolff. (Nov. 1998). *Radar Basics*. [Online]. Available: https://www.radartutorial.eu/index.en.html

[164] H.-C. Chen, T. Chiu, and C.-L. Hsu, "Design of series-fed bandwidth-enhanced microstrip antenna array for millimetre-wave beamforming applications," *Int. J. Antennas Propag.*, vol. 2019, pp. 1–10, Jun. 2019.

[165] L. J. Ippolito, "Radio propagation for space communications systems," *Proc. IEEE*, vol. 69, no. 6, pp. 697–727, Jun. 1981.

[166] S. Rao. *Introduction to mmWave Sensing: FMCW Radars*. Dallas, TX, USA: Texas Instruments, 2017. [Online]. Available: https://training.ti.com/sites/default/files/docs/mmwaveSensing-FMCW-offlineviewing_2.pdf

[167] NXP Semiconductors. (2019). *TEF810X 77GHz Automotive Radar Transceiver*. [Online]. Available: https://www.nxp.com/docs/en/fact-sheet/TEF810XFS.pdf

[168] M. Constapel, M. Cimdins, and H. Hellbrück, "A practical toolbox for getting started with mmWave FMCW radar sensors," in *Proc. 4th KuVS/GI Expert Talk Localization*, 2019, pp. 2–4.

[169] S. Suleymanov, "Design and implementation of an FMCW radar signal processing module for automotive applications," M.S. thesis, Dept. Elect. Eng., Math. Comput. Sci., Comput. Archit. Embedded Syst. (CAES) Res. Group, Univ. Twente, Enschede, The Netherlands, Aug. 2016. [Online]. Available: http://essay.utwente.nl/70986/

[170] A. M. Niknejad. (2005). *Introduction to Mixers*. [Online]. Available: http://rfic.eecs.berkeley.edu/ niknejad/ee142_fa05lects/pdf/lect15.pdf

[171] M. Parker. (May 2011). *Radar Basics—Part 2: Pulse Doppler Radar*. [Online]. Available: https://www.eetimes.com/radar-basics-part-2-pulse-doppler-radar/

[172] S. Rao. (Jul. 2018). *MIMO Radar*. [Online]. Available: http://www.ti.com/lit/an/swra554a/swra554a.pdf

[173] S. Sharenson, "Angle estimation accuracy with a monopulse radar in the search mode," *IRE Trans. Aeronaut. Navigational Electron.*, vols. ANE–9, no. 3, pp. 175–179, Sep. 1962.

[174] Y. Ma and Y. Li, "Millimeter-wave InSAR target recognition with deep convolutional neural network," *IEICE Trans. Inf. Syst.*, vol. E102.D, no. 3, pp. 655–658, 2019.

[175] C. Zech *et al.*, "Active millimeter-wave imaging system for material analysis and object detection," *Proc. SPIE Millim. Wave THz Sensors Technol.*, vol. 8188, Sep. 2011, Art. no. 81880D.

[176] C. Wang, J. Pei, M. Li, Y. Zhang, Y. Huang, and J. Yang, "Parking information perception based on automotive millimeter wave SAR," in *Proc. IEEE Radar Conf.*, Apr. 2019, pp. 1–6.

[177] M. Li *et al.*, "Parking space information monitoring by millimeter wave SAR based on unmanned aerial vehicle," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 9216–9219.

[178] S. Pawliczek, R. Herschel, and N. Pohl, "3D millimeter wave screening for metallic surface defect detection," in *Proc. 16th Eur. Radar Conf.*, 2019, pp. 113–116.

[179] M. E. Yanik and M. Torlak, "Near-field 2-D SAR imaging by millimeter-wave radar for concealed item detection," in *Proc. IEEE Radio Wireless Symp. (RWS)*, Jan. 2019, pp. 1–4.

[180] D. Huston and D. Busuioc, *Chapter 8—Radar Technology: Radio Frequency, Interferometric, Millimeter wave and Terahertz Sensors for Assessing and Monitoring Civil Infrastructures*, vol. 55. Chicago, IL, USA: Woodhead, 2014, pp. 201–237.

[181] T. Liu, Y. Zhao, Y. Wei, Y. Zhao, and S. Wei, "Concealed object detection for activate millimeter wave image," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9909–9917, Jan. 2019.

[182] Y. K. Chan and V. C. Koo, "An introduction to synthetic aperture radar (SAR)," *Prog. Electromagn. Res. B*, vol. 2, pp. 27–60, Sep. 2008.

[183] A. Moreira. (2013). *Synthetic Aperture Radar (SAR): Principles and Applications*. [Online]. Available: https://earth.esa.int/documents/10174/642943/6-LTC2013-SAR-Moreira.pdf

[184] O. Kanhere, S. Ju, Y. Xing, and T. S. Rappaport, "Map-assisted millimeter wave localization for accurate position location," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[185] H. Kim, H. Wymeersch, N. Garcia, G. Seco-Granados, and S. Kim, "5G mmWave vehicular tracking," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 541–547.

[186] A. Yassin, Y. Nasser, A. Y. Al-Dubai, and M. Awad, "MOSAIC: Simultaneous localization and environment mapping using mmWave without A-Priori knowledge," *IEEE Access*, vol. 6, pp. 68932–68947, 2018.

[187] B. R. Mahafza, *Introduction to Radar Analysis*. Boca Raton, FL, USA: CRC Press, 1998.

[188] G. Păter and S. Zsolt, *Chapter 9—Vehicle to Infrastructure Interaction (V2I)*. Budapest, Hungary: BME MOGI, 2014.

[189] X. Wang, S. Gou, X. Wang, Y. Zhao, and L. Zhang, "Patch-based Gaussian mixture model for concealed object detection in millimeter-wave images," in *Proc. IEEE Region Conf.*, Oct. 2018, pp. 2522–2527.

[190] H. Groll *et al.*, "Sparsity in the delay-Doppler domain for measured 60 GHz vehicle-to-infrastructure communication channels," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–6.

[191] K. Du, L. Zhang, W. Chen, G. Wan, and R. Fu, "Concealed objects detection based on FWT in active millimeter-wave images," in *Proc. 7th Int. Conf. Electron. Inf. Eng.*, Jan. 2017.

[192] S. Wang, Z. Ye, and Y. Wang, "Real-time dangerous objects detection in millimeter wave images," in *Proc. 10th Int. Conf. Digit. Image Process. (ICDIP)*, Aug. 2018, Art. no. 108060Z.

[193] W. Xing, J. Zhang, and L. Guo, "A fast detection method based on deep learning of millimeter wave human image," in *Proc. Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, 2018, pp. 67–71.

[194] Y. Amano, *Recognition Ability for Millimeter-Wave Radar Hyper Resolution by New Principle at Once, Romance Three Kingdoms Vehicular Sensor Field 2nd Division*. Tokyo, Japan: Nikkei Business Publishing (BP), Mar. 2018, pp. 25–30. [Online]. Available: https://en.wikipedia.org/wiki/Nikkei_Business_Publications

[195] A. Santra *et al.*, "Short-range multi-mode continuous-wave radar for vital sign measurement and imaging," in *Proc. IEEE Radar Conf.*, Apr. 2018, pp. 0946–0950.

[196] N. Yamada, Y. Tanaka, and K. Nishikawa, "Radar cross section for pedestrian in 76GHz band," in *Proc. Eur. Microw. Conf.*, Oct. 2005, p. 4.

[197] K. A. Gallagher, "Simultaneous human detection and ranging using a millimeter-wave radar system transmitting wideband noise with an embedded tone," M.S. thesis, Dept. Elect. Eng., College Eng., School Elect. Eng. Comput. Sci., PA State Univ., Pennsylvania, PA, USA, May 2013. [Online]. Available: https://etda.libraries.psu.edu/files/final_submissions/8576

[198] I. V. Mikhelson, S. Bakhtiari, T. W. Elmer, and A. V. Sahakian, "Remote sensing of patterns of cardiac activity on an ambulatory subject using millimeter-wave interferometry and statistical methods," *Med. Biol. Eng. Comput.*, vol. 51, nos. 1–2, pp. 135–142, Feb. 2013.

[199] H. Ding, I. Y. Soon, S. N. Koh, and C. K. Yeo, "A spectral filtering method based on hybrid Wiener filters for speech enhancement," *Speech Commun.*, vol. 51, no. 3, pp. 259–267, Mar. 2009.

[200] H. Caesar *et al.*, "NuScenes: A multimodal dataset for autonomous driving," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11621–11631.

[201] Y. Suzuki *et al.*, "Model-based vehicle position estimation using millimeter wave radar," *Int. J. Future Comput. Commun.*, vol. 8, no. 3, pp. 94–98, Sep. 2019.

[202] K. Konishi and T. Sakamoto, "Automatic tracking of human body using millimeter-wave adaptive array radar for noncontact heart rate measurement," in *Proc. Asia–Pacific Microw. Conf. (APMC)*, Nov. 2018, pp. 836–838.

[203] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*. Riverside, CA, USA: California Technical, 1997.

[204] P. Saini, A. Kumar, and N. Singh, "Fpga implementation of 2D and 3D image enhancement chip in hdl environment," *Int. J. Comput. Appl.*, vol. 62, no. 21, pp. 24–31, Jan. 2013.

[205] *Instrument Fundamentals*, Nature Instrument, Austin, TX, USA, 2017. [Online]. Available: https://www.ni.com/en-ca/innovations/white-papers/06/understanding-ffts-and-windowing.html

[206] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, no. 90, pp. 297–301, 1965.

[207] G. Sanderson. (Mar. 2015). *3Blue1Brown*. [Online]. Available: https://www.youtube.com/c/3blue1brown

[208] Texas Instruments. (Oct. 2017). *MMWAVE_SDK 01_01_00_02*. [Online]. Available: http://software-dl.ti.com/ra-processors/esd/MM.WAVE-SDK/01_01_00_02/index_FDS.html

[209] H. Rohling, "Ordered statistic CFAR technique—An overview," in *Proc. 12th Int. Radar Symp. (IRS)*, 2011, pp. 631–638.

[210] R. Nitzberg, "Clutter map CFAR analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 4, pp. 419–421, Jul. 1986.

[211] T. Bower. (Feb. 2008). *Frequency Domain Filters*. [Online]. Available: http://faculty.salina.k-state.edu/tim/mVision/freq-domain/freq_filters.html

[212] J. Semmlow, "Chapter 8—Analysis of discrete linear systems—The Ztransform and applications to filters," in *Circuits, Signals and Systems for Bioengineers*, (Biomedical Engineering). New York, NY, USA: Academic, 2018, pp. 345–397.

[213] S. Raschka. (Jul. 2014). *About Feature Scaling and Normalization*. [Online]. Available: https://sebastianraschka.com/Articles/2014_about_feature_scaling.html
[214] D. Storcheus, A. Rostamizadeh, and S. Kumar, "A Survey of Modern Questions and Challenges in Feature Extraction," in *Proc. 1st Workshop Feature Extraction: Modern Questions Challenges*, vol. 44, 2015, pp. 1–18.
[215] I. Guyon and A. Elisseeff, *An Introduction to Feature Extraction*. Cham, Switzerland: Springer, 2006, pp. 1–25.
[216] M. Livshitz. (Mar. 2018). *Tracking Radar Targets with Multiple Reflection Points*. [Online]. Available: https://e2e.ti.com/cfs-file/key/communityserver-discussions-components-files/1023/Tracking-radar-targets-with-multiple-reflection-points.pdf
[217] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
[218] A. K. Duun-Henriksen *et al.*, "Model identification using stochastic differential equation grey-box models in diabetes," *Diabetes Sci. Technol.*, vol. 7, no. 2, pp. 431–440, 2013.
[219] Z. Chen. (2003). *Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond*. [Online]. Available: https://www.researchgate.net/publication/238689222_Bayesian_Filtering_From_Kalman_Filters_to_Particle_Filters_and_Beyond
[220] C. Yardim, Z.-H. Michalopoulou, and P. Gerstoft, "An overview of sequential Bayesian filtering in ocean acoustics," *IEEE J. Ocean. Eng.*, vol. 36, no. 1, pp. 71–89, Jan. 2011.
[221] M. Rubinstein. (2009). *Introduction to Recursive Bayesian Filtering*. [Online]. Available: https://people.csail.mit.edu/mrub/talks/filtering.pdf
[222] L. M. Ehrman and A. D. Lanterman, "Automated target recognition using passive radar and coordinated flight models," in *Proc. SPIE Autom. Target Recognit.*, vol. 5094, F. A. Sadjadi, Ed., 2003, pp. 196–207.
[223] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org
[224] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Sys. and Tech.*, vol. 2, no. 3, May 2011, article 27.
[225] S. K. Murthy, "Automatic construction of decision trees from data: A multi-disciplinary survey," *Data Mining Knowl. Discovery*, vol. 2, no. 4, pp. 345–389, Dec. 1998.
[226] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Mach. Other Kernel-based Learn. Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[227] Mathematical Model. (2005). *Wikipedia*. [Online]. Available: https://en.wikipedia.org/wiki/Mathematical_model
[228] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *Int. J. Data Mining Knowl. Manage. Process*, vol. 5, no. 2, pp. 1–11, 2015.
[229] R. Raina, A. Battle, H. Lee, B. Packer, and A. Y. Ng, "Self-taught learning: Transfer learning from unlabeled data," in *Proc. 24th Int. Conf. Mach. Learn. (ICML)*, 2007, pp. 759–766.
[230] A. Dosovitskiy, J. T. Springenberg, M. Riedmiller, and T. Brox, "Discriminative unsupervised feature learning with convolutional neural networks," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, vol. 1. Cambridge, MA, USA: MIT Press, 2014, pp. 766–774.
[231] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 845–860, Mar. 2012.
[232] S. T. Kouyoumdjieva, P. Danielis, and G. Karlsson, "Survey of non-image-based approaches for counting people," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1305–1336, 2nd Quart., 2020.
[233] M. Kebe, R. Gadhafi, B. Mohammad, M. Sanduleanu, H. Saleh, and M. Al-Qutayri, "Human vital signs detection methods and potential using radars: A review," *Sensors*, vol. 20, no. 5, p. 1454, Mar. 2020.
[234] C. Xiao, D. Han, Y. Ma, and Z. Qin, "CsiGAN: Robust channel state information-based activity recognition with GANs," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10191–10204, Dec. 2019.
[235] Y. Zheng *et al.*, "Zero-effort cross-domain gesture recognition with Wi-Fi," in *Proc. 17th Annu. Int. Conf. Mobile Syst., Appl., Services*, Jun. 2019, pp. 313–325.
[236] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–25, Jan. 2018.
[237] W. Ross. (Jul. 2017). *The Autonomic Nervous System and its Effects on Health and Vitality*. [Online]. Available: https://www.crossroadsapothecary.com/blog/2017/7/17/integrative-medicine-at-crossroads-the-autonomic-nervous-system-and-its-effects-on-health-and-vitality-author-dr-warren-ross
[238] Centers for Disease Control and Prevention. (Jul. 2020). *Mental Health and Coping During COVID-19*. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html
[239] D. J. Plews, B. Scott, M. Altini, M. Wood, A. E. Kilding, and P. B. Laursen, "Comparison of heart-rate-variability recording with smartphone photoplethysmography, polar H7 chest strap, and electrocardiography," *Int. J. Sports Physiol. Perform.*, vol. 12, no. 10, pp. 1324–1328, Nov. 2017.
[240] V. L. Petrovic, M. M. Jankovic, A. V. Lupsic, V. R. Mihajlovic, and J. S. Popovic-Bozovic, "High-accuracy real-time monitoring of heart rate variability using 24 GHz continuous-wave Doppler radar," *IEEE Access*, vol. 7, pp. 74721–74733, 2019.

Bram van Berlo received the EIT Digital dual MSc. degree in embedded systems engineering from the KTH Royal Institute of Technology and the Eindhoven University of Technology (TU/e) with a specialization in entrepreneurship and embedded networking in 2019, where he is currently pursuing the Ph.D. degree with the Interconnected Resource-Aware Intelligent Systems (IRIS) Research Cluster, Department of Mathematics and Computer Science. He is also working on quantifying the effect of stimuli in mammals with millimeter wave radar. His research interests include pervasive sensing, ambient computing, and cyber-physical systems.

Amany Elkelany received the M.Sc. degree in information technology from Alexandria University, Egypt. She joined the Department of Mathematics and Computer Science, Eindhoven University of Technology (TU/e) in 2020. She is currently a Guest Ph.D. Student with the Interconnected Resource-Aware Intelligent Systems (IRIS) Research Cluster. Her research interests include the area of machine learning, artificial intelligence-based decision-making systems, human–computer interaction, and the Internet of Things (IoT) systems.

Tanir Ozcelebi received the Ph.D. degree in electrical engineering from Koc University, Istanbul. He joined the Department of Mathematics and Computer Science, Eindhoven University of Technology (TU/e), as a Postdoctoral Researcher in 2006, and later as a Faculty Member in 2008. He is currently an Associate Professor of intelligent Internet of Things (IoT) systems with the Interconnected Resource-Aware Intelligent Systems (IRIS) Research Cluster, TU/e Computer Science.

Nirvana Meratnia (Member, IEEE) is currently a Full Professor and the Chair of the Interconnected Resource-Aware Intelligent Systems (IRIS) Research Cluster, Eindhoven University of Technology (TU/e), The Netherlands. Her research interests include the area of pervasive computing, edge AI, and intelligent networked (embedded) systems.