

RadarSense: Accurate Recognition of Mid-Air Hand Gestures with Radar Sensing and Few Training Examples

ARTHUR SLUYTERS, Université catholique de Louvain, LouRIM, Belgium

SÉBASTIEN LAMBOT, Université catholique de Louvain, Belgium

JEAN VANDERDONCKT, Université catholique de Louvain, Belgium

RADU-DANIEL VATAVU, Ștefan cel Mare University of Suceava, Romania

Microwave radars bring many benefits to mid-air gesture sensing due to their large field of view and independence from environmental conditions, such as ambient light and occlusion. However, radar signals are highly dimensional and usually require complex deep learning approaches. To understand this landscape, we report results from a systematic literature review of ($N=118$) scientific papers on radar sensing, unveiling a large variety of radar technology of different operating frequencies and bandwidths, antenna configurations, but also various gesture recognition techniques. Although highly accurate, these techniques require a large amount of training data that depend on the type of radar. Therefore, the training results cannot be easily transferred to other radars. To address this aspect, we introduce a new gesture recognition pipeline that implements advanced full-wave electromagnetic modeling and inversion to retrieve physical characteristics of gestures that are radar independent, *i.e.*, independent of the source, antennas, and radar-hand interactions. Inversion of radar signals further reduces the size of the dataset by several orders of magnitude, while preserving the essential information. This approach is compatible with conventional gesture recognizers, such as those based on template matching, which only need a few training examples to deliver high recognition accuracy rates. To evaluate our gesture recognition pipeline, we conducted user-dependent and user-independent evaluations on a dataset of 16 gesture types collected with the Walabot, a low-cost off-the-shelf array radar. We contrast these results with those obtained for the same gesture types collected with an ultra-wideband radar made of a vector network analyzer with a single horn antenna and with a computer vision sensor, respectively. Based on our findings, we suggest some design implications to support future development in radar-based gesture recognition.

CCS Concepts: • **Human-centered computing** → **Gestural input**; **Graphical user interfaces**; *Interactive systems and tools*; • **Computing methodologies** → *Model development and analysis*; • **Software and its engineering** → Runtime environments; • **Hardware** → Radio frequency and wireless interconnect.

Additional Key Words and Phrases: Dimension reduction, Hand gesture recognition, Radar-based interaction.

ACM Reference Format:

Arthur Sluyters, Sébastien Lambot, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2023. RadarSense: Accurate Recognition of Mid-Air Hand Gestures with Radar Sensing and Few Training Examples. *ACM Trans. Interact. Intell. Syst.* 13, 1, Article 111 (June 2023), 45 pages. <https://doi.org/10.1145/3589645>

Authors' addresses: Arthur Sluyters, arthur.sluyters@uclouvain.be, Université catholique de Louvain, LouRIM, Place des Doyens, 1, Louvain-la-Neuve, 1348, Belgium; Sébastien Lambot, sebastien.Lambot@uclouvain.be, Université catholique de Louvain, Earth and Life Institute, Croix du Sud, 2, Louvain-la-Neuve, 1348, Belgium; Jean Vanderdonckt, jean.vanderdonckt@uclouvain.be, Université catholique de Louvain, Louvain Research Institute in Management and Organizations, Institute for Information and Communication Technologies, Electronics and Applied Mathematics, Place des Doyens, 1, Louvain-la-Neuve, 1348, Belgium; Radu-Daniel Vatavu, Ștefan cel Mare University of Suceava, MintViz Lab, MANSiD, 13 Universitatii, Suceava, 720229, Romania, radu.vatavu@usm.ro.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM

1 INTRODUCTION

Gesture-based user interfaces promise natural [171] and intuitive interaction [57] for users, in addition to many other desirable and convenient characteristics [78] for interacting with computer systems. According to the number of dimensions on which gestures are acquired by sensors and/or articulated by users, respectively, two main gesture categories have been frequently examined in the scientific community, as follows:

- *Two-dimensional (2D) touch and stroke-gestures.* These gestures are performed on a touch-sensitive surface [186], such as a trackpad (e.g., touch gestures on Windows PC touchpads), a smartphone or smartwatch touchscreen [154], the surface of a tablet [14, 32] or of an interactive tabletop [127], including flexible and elastic surfaces [77, 150] and gestures performed on tangibles [12, 13]. Gestures from this category typically require a contact-based sensor and are prevalent on mobile and wearable devices with touchscreens.
- *Three-dimensional (3D) mid-air gestures.* These gestures specify the movement of the body or a part of the body that involves manipulating a physical object in space [13], adopting a specific pose [90], or performing a motion in 3D [152, 164]. Vision-based sensors, such as video cameras, Leap Motion Controller (see QUANTUMLEAP [137] for example), Kinect, or RealSense, have traditionally been used to capture gestures from this category. These gestures can also be captured by inertial sensors [6, 95] that report e.g., the acceleration and orientation of the body part to which the sensor is attached and are already integrated in many devices, such as smartphones, smartwatches, and electronic rings.

Recognition of gestures from the first category can be achieved with high precision by employing algorithms that have received repeated validation in the community [100], and that need few resources for training, e.g., just a few training examples per gesture type [153, 154]. Gestures of the second category need dedicated recognition approaches because of their large diversity, for which computer vision techniques [28, 34, 103] have been proposed. The diversity of gestures that can be detected with wearable sensors, including inertial sensing, has also been documented in the scientific literature [46]. Although effective techniques have been proposed to address some of this diversity, the well-known limitations of vision- and inertial-based gesture sensing, among which the sensitivity to environmental conditions of the former [19, 20, 52, 178, 180] and obtrusiveness [19] of the latter, demand exploration of alternative gesture sensing approaches for interactive systems that implement 3D mid-air gestures.

Radar sensing [178] has recently been examined as one of such alternative technologies with applications in VR [58], activity recognition [19, 20], object and material classification [30, 44, 64, 178], 2D and 3D scene reconstruction [108], motion sensing [50], and others [179, 180]. In addition to the extrinsic motivation to use radars to address the limitations of other gesture sensing approaches, several characteristics of radar sensing make this technology particularly attractive to Human-Computer Interaction (HCI) researchers and practitioners [180]. However, prior research on radar-based gesture sensing and recognition [18, 20, 52, 112] has employed custom-built radars that are difficult to reproduce by non-experts. In addition, Machine Learning (ML) and Deep Learning (DL) models require large training sets to address high-dimensionality data represented by radar signals. With the exception of Google Soli [92], a proprietary radar chip design integrated into a commercial smartphone model, the results reported in previous work on custom radar technologies have not yet been transposed, to the best of our knowledge, to standard portable radar sensors.

Consequently, many practical aspects of gesture recognition, such as the effect of dimensionality and cardinality examined for other gesture types [152], are still to be addressed for existing ML/DL radar-based gesture recognition techniques. Furthermore, despite the rich literature on radar-based gesture recognition, only small gesture sets composed of simple gesture types have been used so far. For instance, Soli [92] was originally demonstrated with four gestures

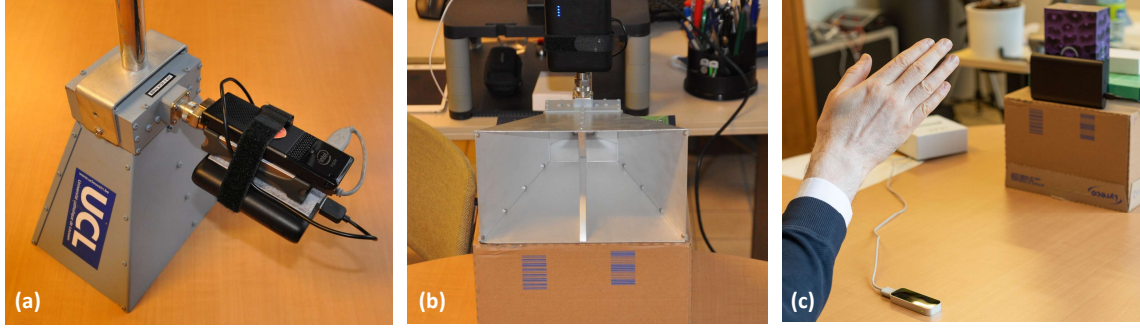


Fig. 1. The devices used in our experiments: (a) the “horn”, our single transmitting and receiving, ultra wideband horn antenna, top view and (b) front view; (c) the Walabot (top) and the Leap Motion Controller (bottom).

only: pointing the finger up and down and left and right, respectively, to control home appliances, such as a lamp, a radio, or a TV. In this context, we make the following contributions in this paper:

- We review the state-of-the-art in radar sensing technologies with a focus on interaction with computer systems. We briefly review results reported in the scientific literature for the Walabot, a low-cost, off-the-shelf, portable radar that we are using in our work to capture and recognize mid-air gestures.
- We perform a Systematic Literature Review (SLR) of the scientific literature on radar-based interaction with 118 peer-reviewed papers. We report on the large diversity of radar sensors operating in various bandwidths and physical configurations that have been used in the scientific literature, as well as on the algorithms developed for radar gesture recognition and the gestures that they support.
- We introduce a new gesture recognition pipeline for mid-air hand gesture recognition using radars that employs an advanced formalism of full-wave electromagnetic modeling and inversion [71, 72] to address several challenges, including the high dimensionality of radar signals [148]. Our processing pipeline reduces the raw radar data to a normalized 2D space specified by two measures: δ , the distance between the user’s hand and the radar, and ϵ_r , the apparent relative permittivity of the hand characterizing its reflectivity. This last quantity is thereby independent of the distance and characterizes the configuration of the hand with respect to the incident field.
- We evaluate our gesture recognition pipeline under both user-dependent and independent conditions with a dataset of 16 gesture types acquired with the Walabot (Fig. 1-c), a low-cost off-the-shelf radar sensor. We report the recognition accuracy rates obtained after each processing step of our pipeline. We compare these results to the those obtained for gestures acquired with a Leap Motion Controller as a baseline and to the results of a previous evaluation with a radar setup using a Vector Network Analyzer (VNA) and a single transmitting and receiving, ultra-wideband horn antenna (Fig. 1-a and b).
- Based on our theoretical and experimental findings, we propose design implications for radar-based gesture sensing and recognition.

This paper extends our previous work [136] by introducing a new SLR on radar-based interaction, an expanded radar gesture dataset with two additional users and twice as many gesture recordings per user, an improved procedure for gesture acquisition (which drops the horn radar antenna), a user-independent testing scenario, and an expanded list of design implications for radar-based gesture sensing and recognition.

2 RELATED WORK

After reviewing selected previous work on the recognition of mid-air hand gestures, we highlight in this section key moments of the development and evolution of radar technology, including its applications in Computer Science overall and in HCI in particular. We also present the Walabot device that is used throughout our paper as a reference for a portable, off-the-shelf, low-cost radar sensor.

2.1 Mid-Air Hand Gesture Recognition

Several vision-based sensors have been commercially available at affordable prices, such as the Intel RealSense, Microsoft Kinect Azure, and the Leap Motion controller (LMC), to name a few. LMC, a two-camera vision-based sensor that tracks both hands at a cost of approximately 100 USD was reported to be precise enough for interactive applications that implement hand tracking [170] and recognition [183]. However, LMC is affected by the common problems of vision-based approaches, such as high sensitivity to environmental conditions [180], limited field of view [52], sensing perturbations including occlusion [25], and privacy concerns [19]. Other sensors that address these issues have been explored, such as Aili [90], a custom device that reports hand skeleton data without privacy concerns, and electromyography (EMG) and inertial measurement unit (IMU) devices employed for gesture recognition [56, 95, 99, 146]. For example, Akl and Valaee [6] used Dynamic Time Warping (DTW) and affinity properties to recognize gestures from acceleration signals. Laput and Harisson [76] used accelerometer and bio-acoustic data from an off-the-shelf smartwatch to recognize a set of 25 activities performed with the hands. De Smedt et al. [34] used depth and skeletal models. Li et al. [83] recognized finger gestures with high precision using an entropy-weight allocation k Nearest-Neighbor (k-NN) approach. In addition, 3D gesture recognition has become invariant to translation and rotation [15] to become insensitive to variations of gesture articulations performed in space. Such 3D gestures have been also incorporated in some design tools, such as MAGIC 2.0 [69], and development environments, such as CODESPACE [24].

2.2 Evolution of Radar Technology

The range of radar applications is diverse and the potential for further development is growing due to miniaturization and cost decrease. From this perspective, a useful technology is passive radar, which has the advantage of not needing generation and transmission of its own radio frequency signals. For example, Passive Coherent Location (PCL) is based on locating objects with Electromagnetic (EM) transmissions already existing in the environment. A key component of radar technology is the information content of radar signals. Fundamentally, each operating frequency and transmitter-receiver couple represent a piece of information with some degrees of independence. Therefore, complex radar applications, such as subsurface imaging that uses ground penetration radar (GPR) [135], resort to ultra-wide bandwidths and antenna arrays to maximize image resolution as well as the well-posedness of inverse scattering problems [33]. In that respect, GPR constitutes one of the first radar technologies, from 1904 (Patent DE 165 546 by Christian Hlsmeier), after the results of Heinrich Hertz on electromagnetic radiation from the 1880s. One of the first GPR surveys was conducted by Stern in 1929 [139] to determine the depth of a glacier. Since then, radar technology has found many applications, object detection in air or within materials, navigation, earth and space observation, medical imaging, speed determination, etc.

There are two main classes of radars: *frequency-domain radars*, which are based on the transmission of continuous waves that span a specified frequency range, and *time-domain radars*, which are based on the transmission of electromagnetic pulses. The latter category spans a wide bandwidth and dominated the market in the twentieth century, given lower production cost at that time and their acquisition rates of several orders of magnitude larger compared to the frequency-domain radars. The frequencies range from a few MHz for deep subsurface investigations to the order of THz

for high-resolution non-destructive testing (*e.g.*, inspection of airplane wing layers), and even larger for light detection and ranging systems (LiDAR). Depending on the specific application, the type of radar, its operating frequencies, and antenna configurations are chosen to collect the required information with an optimal resolution, penetration depth and at an acceptable cost. Recent advances in electronics, electromagnetic modelling, and miniaturization have opened new avenues for radar applications. For example, Wu et al. [174] used a handheld VNA and a home-made hybrid horn-dipole antenna to build a lightweight radar for soil moisture mapping using drones. Radars and antennas can be printed on circuit boards when the size and type of antennas allow for such a manufacturing approach, such as for millimeter-wave on-chip radars. In that respect, frequency-modulated continuous-wave (FMCW) radars are becoming more widespread, with production costs as low as a few US dollars for the simpler models. When a sufficiently wide bandwidth is available, an inverse Fourier transform of the radar signal provides the time-domain counterpart, similarly to a pulse radar, and the range information is directly accessible. The Walabot device, used for the acquisition of mid-air hand gestures in this paper, is such a device with 3D positioning and geometrical reconstruction capabilities based on an array of 18 bowtie antennas for multi-offset measurements.

2.3 Radar-based Interaction

Radar detection allows a variety of interactions, including interactions performed below surfaces, such as a desk [19], walls, or any material that does not significantly affect recognition [112]. These materials are, for instance, *e.g.*, wallpaper, cardboard, and wood, which have relatively low permittivity and negligible electrical conductivity. Radars are also insensitive to many environmental and lighting conditions [180].

The history of radar-based interaction starts with MAGIC CARPET [113], a Doppler radar used to detect coarse 3D body motions. Since then, radars have become increasingly popular in HCI. For example, RADARCAT [178] was introduced to recognize physical objects and materials with a processing technique based on a Random Forest classifier. In an evaluation study, RADARCAT was effective in identifying 26 types of materials, 16 transparent materials, and 10 body parts from 6 participants.

In addition, materials with relatively low permittivity are quite transparent to radars and do not significantly alter recognition, especially if the material is accounted for in the processing pipeline. Yeo et al. [179] used a radar to implement a tangible interaction involving counting, ordering, and identifying objects. Radars have also been used for human detection [8], activity recognition [19], position estimation [20], and motion detection and classification [112].

GESTUREVLAD [22] is a Doppler radar hand pose recognition framework that supports slight variations in gesture execution while accurately differentiating between gesture classes. PANTOMIME [112] used a fixed feet-based radar with high-frequency *i.e.*, 76-81GHz equipped with a 4GHz continuous bandwidth and composed of four receiving and three transmitting antennas. A deep learning approach was used to accurately recognize twenty-one gestures acquired by forty-five participants from 3D point clouds. That radar works at a frequency about ten times larger than that of the Walabot, resulting in a wavelength ten times smaller and a resolution that is about ten times finer. Wang et al. [160] used only two antennas to recognize 2D stroke gestures due to their low dimensionality. Short-range radar-based gestures could also be recognized using 3D Convolutional Neural Networks (CNNs) with a triplet loss [55].

Most of these previous works have exploited custom radars built to expose specific features and rely on specific architectures. This approach results in systems that are difficult to reproduce, which causes a barrier to their direct adoption for radar-based interaction. On the contrary, Google Soli [92, 161], one of the few radars on the market specifically designed for interactive applications, was introduced for micro gesture recognition, such as finger wiggle, hand tilt, check mark, and thumb slides. This set was expanded with Convolutional and Recurring Neural Networks

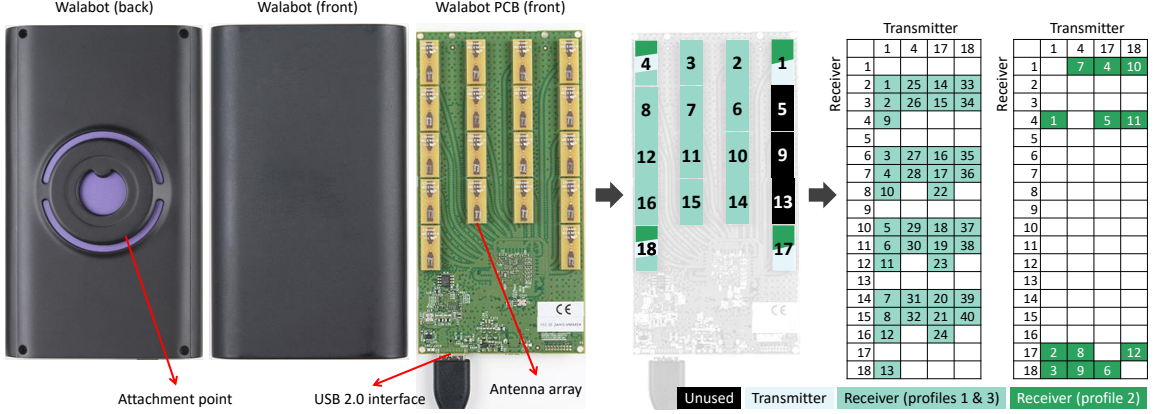


Fig. 2. The Walabot Developer: device, antenna array, antenna IDs, and antenna pairs IDs used in profiles 1, 2, and 3 [136].

(RNNs) to eleven gesture classes. Also, Choi et al. [29] proposed a gesture recognition system for Soli that recognizes a set of ten hand gestures in real time. Soli has also been employed with RadarNet [52], an efficient gesture recognizer for five gesture classes (*i.e.*, left, right, up, down swipes, and an omnidirectional swipe). Google Soli can be embedded in mobile devices, such as a smartwatch or smartphone, or in physical objects. For example, Flintoff et al. [44] integrated Soli into a robot, Attygalle et al. [18] also exploited a Google Soli sensor via a 3D CNN (Conv3D) and a spectrogram-based ConvNet to recognize on-object gestures (94% for a five-gesture set) and to investigate to what extent different materials would impact recognition accuracy [30]. Unlike embedded radars in objects [179], an external radar enables the tracking of physical objects. Ren et al. [120] explored the possibility of combining radar gesture sensing and wireless communication on a smartphone equipped with a 802.11ad 60 GHz WiFi chip, and reported high accuracy for a set of five gesture types.

Radar-based datasets [5, 112] are rarely available. When available, their size is in the order of several GigaBytes, which makes them challenging to process, especially for mobile applications. Moreover, the datasets available in the scientific community are radar-dependent.

2.4 Interactions with the Walabot Radar

The Walabot consists of a 3D imaging sensor represented by an ultra-wideband (UWB) frequency modulated continuous-wave (FMCW) radar. Its dimensions are 144 mm×85 mm×18 mm (5.67 in.×3.35 in.×0.71 in), which makes it easy to use with a smartphone, tablet, or PC in both mobile and stationary contexts of use. Walabot comes in two versions: a US/FCC version operating in the 3.3-10 GHz frequency range and an EU/CE version in the narrower 6.3-8 GHz range. The latter is the most restrictive and challenging for gesture recognition as the resolution is inherently lower. We used that version in the study reported in this paper. The Walabot Developer contains an array of 18 antennas, of which four are used as transmitters (depicted in orange in Fig. 2) and the rest are receivers (depicted in green in Fig. 2) and three are unused (gray in Fig. 2). The Walabot SDK provides three different profiles for distant scanning, which define the frequency range, the set of antenna pairs (right part of Fig. 2), the number of fast-time samples per frame, and the frame rate, as follows:

- Profile 1 (PROF_SENSOR): 6.3-8 GHz range, 40 antenna pairs, 8192 fast-time samples / frame, ~20 fps.
- Profile 2 (PROF_SENSOR_NARROW): 6.3-8 GHz range, 12 antenna pairs, 4096 fast-time samples / frame, ~41 fps.

- Profile 3 (PROF_WIDE_BAND): 3.3-10 GHz range, 40 antenna pairs, 8192 fast time samples / frame, ~15 fps. This profile is only available with the US/FCC version of the Walabot.

We selected Walabot in our work because it has been proven to be effective and efficient in various application domains, including material identification [1] and activity recognition [19, 202]. As a representative example, Avrahami et al. [19] used it to recognize human activities at the checkout counter of a convenience store and a typical office desk. Despite its many benefits, such as its relatively low price (~500 USD), stock availability, and small size, only one paper used Walabot for hand gesture recognition. Zhang et al. [187] proposed a deep neural network for continuous gesture recognition and evaluated it on a dataset of eight hand gestures with 150 samples per gesture class performed very close to the radar.

3 A SYSTEMATIC LITERATURE REVIEW OF RADAR-BASED INTERACTION

We conducted a systematic review of the literature (SLR) [68] to identify the types of radar, algorithms, and gesture sets used for radar-based gesture recognition (Fig. 3). The query RQ="radar" AND "gesture recognition" was run in five digital libraries: (1) ACM Digital Library, (2) IEEE Xplore, (3) MDPI (Sensors), (4) SpringerLink, and (5) Elsevier ScienceDirect and resulted in 1,515 articles, from which $N = 118$ were identified as relevant. Table 1 groups these papers based on their year of publication. 91.5% of the papers were published between 2018 and 2020, indicating that this field of research has recently become quite active. All papers provide some evaluation of the system performance (Fig. 4), while 14 papers ($\frac{14}{118}=12\%$) demonstrate potential use cases of their system with a prototype and only one paper ($\frac{1}{118}=1\%$) evaluates the user experience of the proposed system.

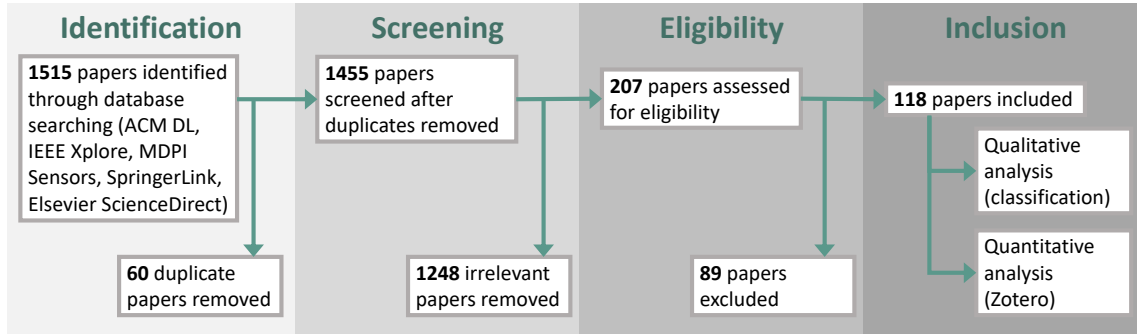


Fig. 3. PRISMA diagram [111] of our systematic literature review.

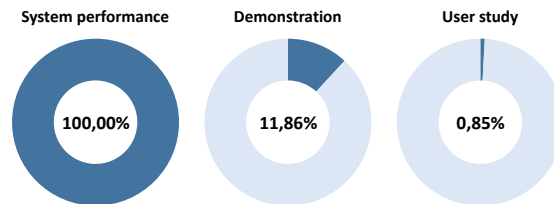


Fig. 4. Types of validation.

3.1 Radars

We identified 123 radar systems and extracted their parameters, including the radar type, the frequency band in which they operate, the model and their position. We identified four types of radars (Fig. 5a):

- *Frequency-Modulated Continuous-Wave* (FMCW) radars ($\frac{65}{123}=53\%$), which continuously transmit a signal that varies in frequency, *i.e.*, modulated in frequency. These radars can resolve both the range and Doppler [7], enabling the recognition of static poses and dynamic gestures.
- *Unmodulated Continuous-Wave* (CW) radars ($\frac{36}{123}=29\%$), which continuously transmit a signal at a constant frequency and amplitude [109]. These radars are simple and inexpensive. They can measure the Doppler shift of the return signal caused by moving objects but cannot provide range information or identify stationary targets, as the latter do not cause a Doppler shift. Consequently, radars of this type are not suitable for detecting static poses of the hand or body.
- *Pulse* radars ($\frac{21}{123}=17\%$) transmit short radar pulses and listen for the returning echo. These radars can measure the range of targets, stationary and moving, making them suitable for the recognition of static poses and dynamic gestures. The pulse length affects the radar detection range (a longer pulse length increases the maximum range but may prevent the detection of very close targets) as well as its range resolution (a shorter pulse length allows for better discrimination between two spatially separated targets) [23]. The smaller the pulse length in the time-domain, the larger the bandwidth in the frequency domain counterpart.
- *Direct-Sequence Spread Spectrum* (DSSS) radars ($\frac{1}{123}=1\%$) transmit signals modulated by a random bit sequence. These radars can measure the range of stationary and moving targets, and their architecture is simpler than FMCW radars [145].

Most of the papers ($\frac{98}{118}=83\%$) relied on a single radar sensor to capture gestures, while the rest ($\frac{20}{118}=17\%$) used two or more radars. Among the latter, four papers introduced different radar models. Some papers introduced other types of sensors, such as [129], where gestures were collected using a radar, a Kinect, and an RGB camera for comparison purposes. Gigie et al. [47] synthesized micro-Doppler radar signatures of gestures from the skeleton provided by a Kinect. Other papers performed data fusion with RGB cameras [106, 149], depth cameras [106, 147], and thermal cameras [134], respectively.

Fig. 6 summarizes the position of the radar systems identified in the user environment. Almost half of the identified papers ($\frac{49}{118} = 42\%$) did not provide enough information about the position of the radar(s) in the user's environment.

Year	#References	Reference(s)
2009	1	[89]
2014	1	[157]
2015	1	[106]
2016	6	[66, 92, 101, 114, 161, 193]
2017	11	[35, 62, 65, 67, 74, 84, 102, 122, 125, 156, 190]
2018	21	[53, 59, 73, 75, 85, 86, 98, 107, 107, 116, 124, 126, 138, 140, 144, 177, 194, 195, 197, 199, 201]
2019	35	[4, 9, 11, 16, 17, 22, 27, 29, 31, 38, 39, 41–43, 45, 47, 49, 54, 55, 60, 70, 93, 133, 141, 159, 162, 164–166, 168, 189, 191, 196, 198, 200]
2020	41	[3, 10, 21, 37, 40, 51, 61, 63, 79, 81, 82, 87, 88, 91, 94, 96, 104, 115, 119, 121, 123, 128, 129, 132, 134, 142, 143, 147, 149, 158, 160, 163, 167, 175, 176, 181, 182, 184, 185, 188, 192]
2021	1	[2]

Table 1. Distribution of papers from our systematic literature review, based on their year of publication.

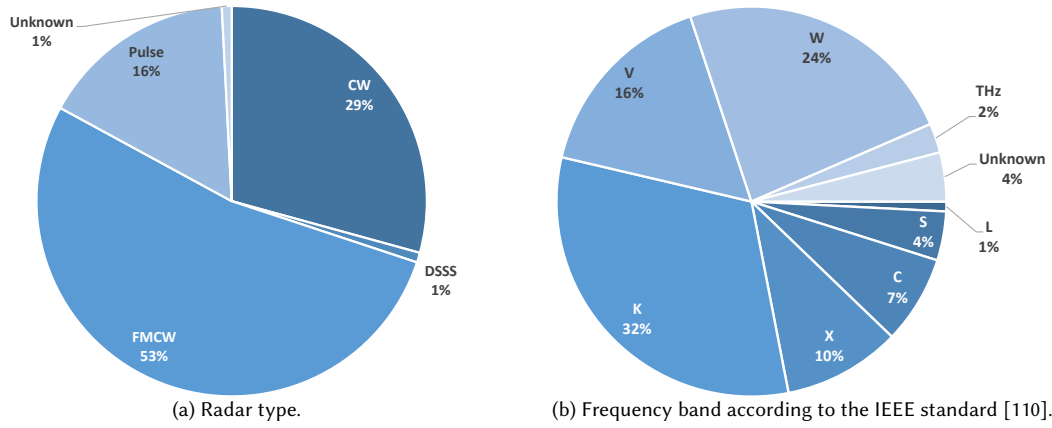


Fig. 5. Distribution of the radars identified in our systematic literature review.

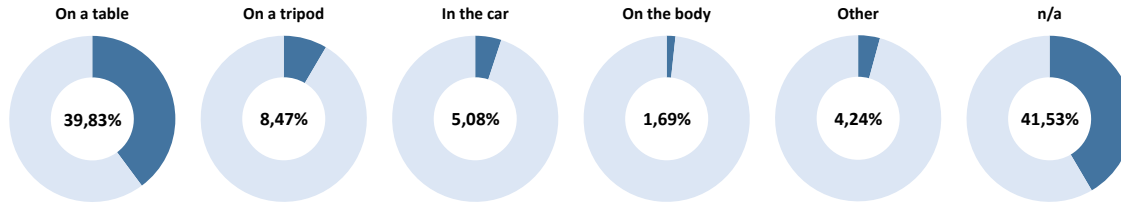


Fig. 6. Positions of the radar.

From the papers that provided this information, 47 ($\frac{47}{118} = 40\%$) placed the radar(s) on a table, while 10 ($\frac{10}{118} = 8\%$) placed them on a tripod, and six ($\frac{6}{118} = 5\%$) in the car (e.g., on the center console [106], near the steering wheel [4, 144], or on the roof console [141]). Other positions include on the body [31, 104], on fabric [175], and on a robot [89].

We also classified radars according to their frequency band according to the IEEE Standard Letter Designations for Radar-Frequency Bands [110] (Fig. 5b). The larger the frequency band, the better resolution the radar provides, and, in principle, the finer the gestures can be acquired. Fig. 5b shows that most of the radars ($\frac{88}{123} = 72\%$) operate in the K (18 to 27 GHz, 32%), V (40 to 75 GHz, 16%), and W (75 to 110 GHz, 24%) bands. The V band encompasses the Google Soli [92], which operates at 60 GHz and was used in five papers [22, 29, 31, 92, 161] identified by our SLR. Three papers used very high frequency THz band radars (300 to 1000 GHz) to achieve better resolution [158, 159, 201]. Only $\frac{27}{123} = 22\%$ of the radars operated at frequencies below 18 GHz: 10% in the X band (8 to 12 GHz), 7% in the C band (4 to 8 GHz), 4% in the S band (2 to 4 GHz), and 1% in the L band (1 to 2 GHz). The benefit of lower frequencies is the larger detection range. For a given application, a trade-off has therefore to be chosen between resolution and range. The detection range can also be increased by increasing the transmitted power, but this increases energy consumption and can lead to adverse health effects. Fig. 8 summarizes the distribution of frequency bands by radar type.

Finally, we extracted the radar models used in the papers that provided this information; see Fig. 7. More than a third of the papers ($\frac{46}{123} = 37\%$) did not provide enough information to identify the exact models. Of the 77 remaining publications, we identified 29 unique models with the most popular being the Texas Instruments AWR1642 ($\frac{13}{123} = 11\%$) and IWR1443 ($\frac{5}{123} = 4\%$), the Novelda XeThru X4 ($\frac{9}{123} = 7\%$), the RFBeam K-LC2 ($\frac{8}{123} = 7\%$), and Google Soli ($\frac{6}{123} = 5\%$), respectively. Except for Google Soli, which was designed for integration into smartphones and smart speakers, most of the radar systems from the SLR were not meant for everyday use. However, some of the papers introduce prototypes of radars that could be integrated into furniture [102] or even stitched into clothing [175].

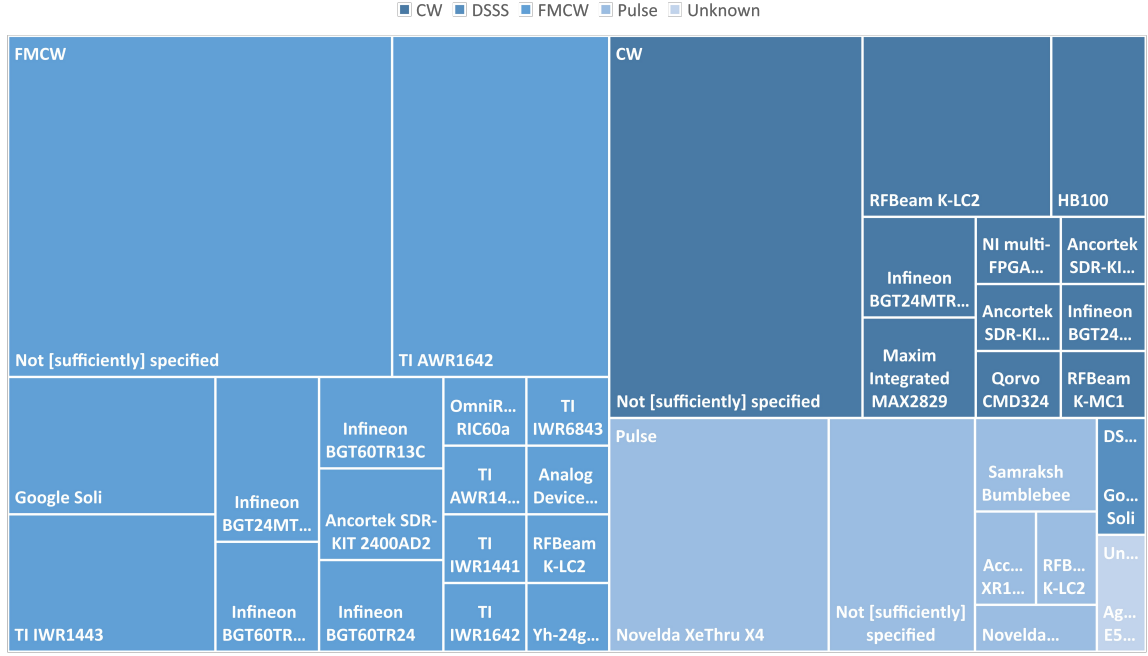


Fig. 7. Tree map [130] of the radar models identified in our systematic literature review. The color of a cell indicates the type of radar (CW, DSSS, FMCW, Pulse, or unknown) while the size of a cell is proportional to the number of papers featuring this model of radar.

From the above analysis, we make the following observations:

- O₁. *Predominance of FMCW over other types of radar* (Fig. 5a and 8): these radars are the most frequently exploited due to their ability to process both static and dynamic gestures, mostly covered by the K and W ranges, but also because several models are commercially available. The frequency band covered by these two categories are high enough to acquire raw data with high resolution.
- O₂. *Limited coverage of very low/high frequency bands* (Fig. 9): lower frequencies result in worse range resolution, making it more difficult for the radar to distinguish between two targets. Such radars are thus less interesting for smaller scale gestures, such as hand and finger gestures, which explains the lower number of works covering these systems. On the contrary, higher frequencies enable better range resolution and thus are better suited for small-scale hand and finger gestures (e.g., Google Soli [92, 161]), at the cost of higher complexity and lower range as the frequency increases. Most systems operate at frequencies between 4 GHz and 100 GHz, which can provide sufficient range resolution with relatively low complexity and power consumption. In addition, some off-the-shelf radar systems used in HCI research were originally developed for other applications and operate specific frequencies, such as collision avoidance systems in cars, which often operate in the K band.

The Walabot sensor used in this paper is an off-the-shelf FMCW radar. Thus, it can serve for recognizing both static poses and dynamic gestures. The US/FCC version of the Walabot Developer covers three IEEE ranges, i.e., S, C, and K (Fig. 9), from 3.3 to 10GHz, while the EU/CE version is limited to only one IEEE range, i.e., K, from 6.3 to 8GHz. Both versions of the Walabot provide a good compromise for capturing relatively fine-grained gestures performed at a short distance. However, the limited frequency range of the EU/CE Walabot imposes a significant constraint on gesture

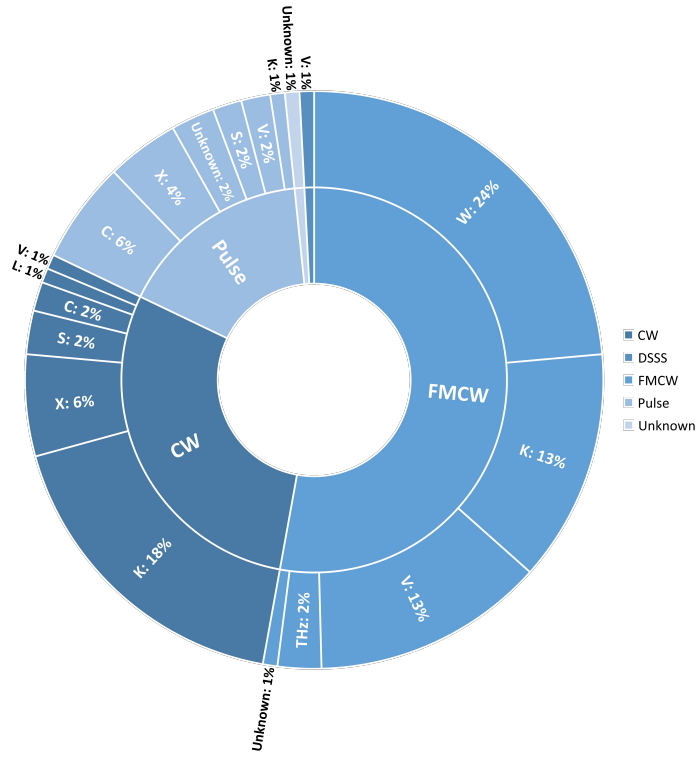


Fig. 8. Sunburst chart of the radars identified in the SLR. The inner ring and color show the proportion of each radar type, while the outer ring indicates the proportion of each frequency band for each radar type according to the IEEE standard [110].

recognition because of its lower range resolution. Nevertheless, it is worth noting that the use of full-wave modeling and inversion pushes forward the resolution frontiers compared to classical time-domain, straight-ray propagation approaches. Indeed, with full wave modeling, close targets can be resolved, to some extent, although only one reflection is visible in the time domain [105, 117]. This constitutes one of the benefits of the approach proposed in this paper.

3.2 Algorithms for Gesture Recognition

Fig. 10 presents the algorithms used in our corpus: 151 algorithms for radar-based gesture recognition were identified, the majority representing deep learning approaches ($\frac{81}{151} = 54\%$), of which Convolutional Neural Networks (CNNs) were the most frequently used. CNNs were often combined with other models, such as Long Short-Term Memory (LSTM) or Support Vector Machines (SVM), in $\frac{23}{158} = 15\%$ of the cases. Classical machine learning algorithms were still used, as follows: K Nearest Neighbors ($\frac{23}{151} = 15\%$), Support Vector Machines ($\frac{16}{151} = 11\%$), ensemble learning ($\frac{8}{151} = 5\%$), Decision Trees ($\frac{6}{151} = 4\%$), Hidden Markov Models ($\frac{3}{151} = 2\%$), and Bayesian networks ($\frac{2}{151} = 1\%$).

From the above analysis, we make the following observations

- O₄. *Predominance of CNN algorithms*: CNNs, often combined with another model like LSTM, are predominant for recognizing radar-based signals. They are particularly appropriate for image processing, but also they benefit from their intrinsic capability to select the discriminant radar features without any human supervision, in

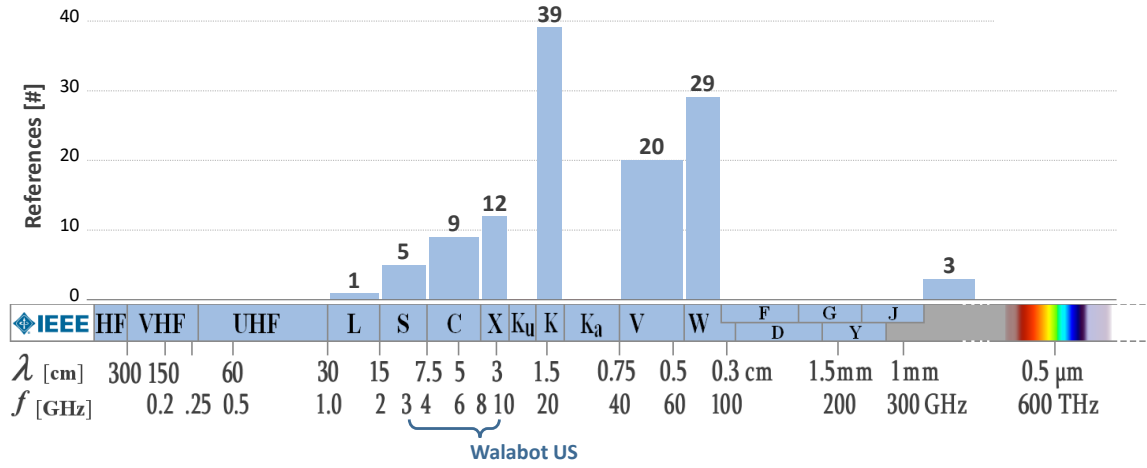


Fig. 9. Number of radar systems according to their frequency band [110] distributed along the electromagnetic continuum (Bottom image of waves and frequency ranges used by radar used with permission from Christian Wolff [172]), from a total of 118 radars. Five radars did not mention their frequency band.

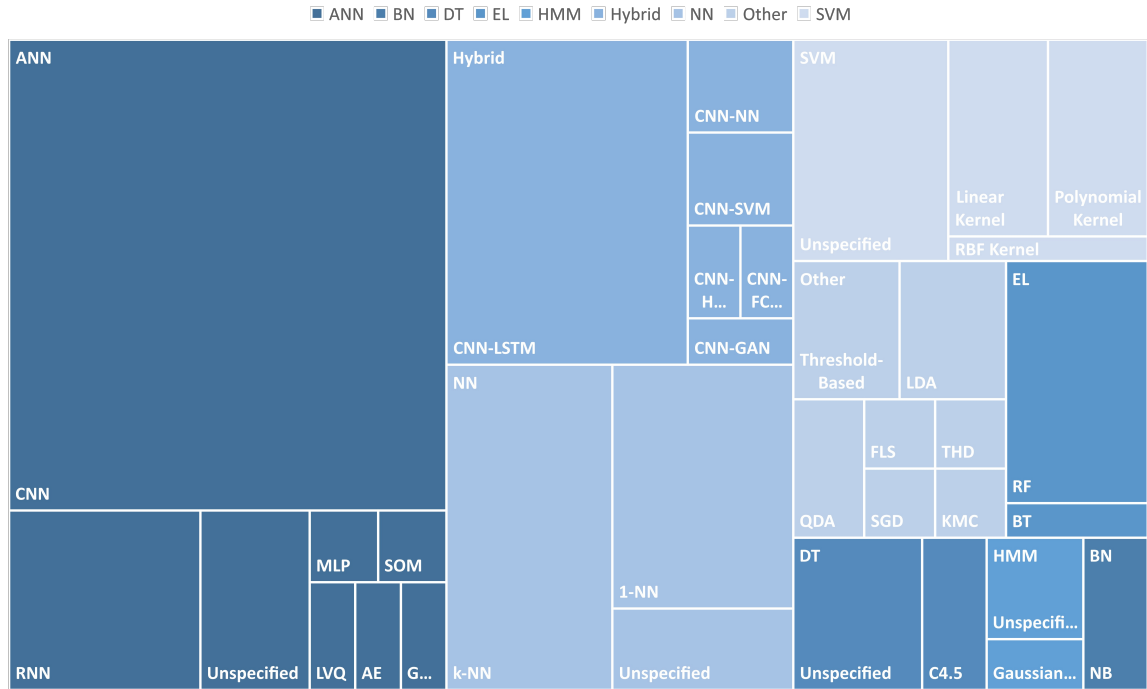


Fig. 10. Tree map [130] of the algorithms for radar-based gesture recognition identified in the corpus. The color of a cell indicates the type of algorithm (e.g., ANN, HMM, or SVM) while the size of a cell is proportional to the number of papers featuring this algorithm.

particular when dealing with gesture sets that are very large in size and complexity (radar images are very challenging to discriminate by a non-expert).

- O₅. *Dependence of CNN with respect to the radar*: since CNNs automatically select features directly from radar images, these key features are not pretrained and only learned when the gesture set is examined, thereby making the CNN very appropriate, but also very specific to the radar images produced. The CNNs therefore suffer from a high dependence on radar type and possibly radar unit for a given model.
- O₆. *No modeling of the radar functioning*: since CNNs ensure most of the work for a specific radar, we did not identify any work in gesture recognition aimed at modeling the radar functioning in some way to ensure some radar independence or some mechanism to reason about its processing.

In this paper, we rely on a preexisting 1-NN algorithm for gesture recognition, namely, Taranta II et al. [146]’s Jackknife, as the goal of this paper is not to create new algorithm for gesture recognition, but to investigate techniques for processing radar signal for data normalization and dimension reduction. This gesture recognition algorithm performs relatively well on many types of data, requires very few training templates, is fast to train and execute, and is easy to understand. Our pipeline is not restrictive and can thus be combined with any algorithm for gesture recognition, including deep learning approaches such as CNNs, to improve their accuracy.

3.3 Datasets

Fig. 11 gives an overview of the types of gestures used in the identified papers. Most papers ($\frac{107}{118} = 91\%$) focused on mid-air arm and hand gestures, which are relatively easy to recognize thanks to their scale that is neither too small, thus not requiring extremely fine range resolution, nor too large, which makes it possible for the radar to focus exclusively on the current user, thus simplifying signal processing. Almost half of the papers ($\frac{54}{118} = 46\%$) featured mid-air finger gestures, which require higher frequency radars (capable of finer range resolution) to accurately capture the fine-grained finger motion. Other papers featured whole body gestures [88, 89, 149], touch input [31, 175], or even head gestures [157]. We identified only seven publicly available datasets out of the 118 papers from our SLR; see Table 2.

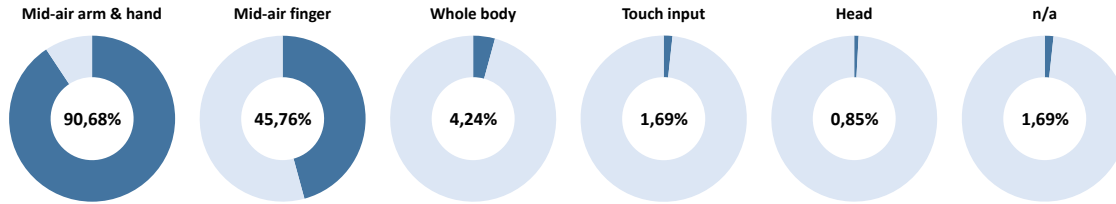


Fig. 11. Types of gestures.

Name	#Classes	#Users	#Samples	Sensor	Original paper	Other papers
mHomeGes	10	25	>22000	TI IWR1443	[94]	/
mmASL (wake-words)	2	3	3700	NI multi-FPGA platform (w/ NI PXIe7902, 7976, 3610/3630) and SiBeam phased antenna array	[129]	/
mmASL (signs)	50	15	12236	NI multi-FPGA platform (w/ NI PXIe7902, 7976, 3610/3630) and SiBeam phased antenna array + Kinect + RGB camera	[129]	/
deep-soli	11	11	5500	Google Soli	[161]	[22]
dop-net	4	6	3052	Ancortek SDR-KIT 2400AD2	[121]	[21]
HARrad (gestures)	6	9	2347	77GHz FMCW radar	[149]	/
HARrad (events)	6	9	1505	77GHz FMCW radar	[149]	/

Table 2. Summary of the publicly available datasets identified in our systematic literature review.

The mHomeGes dataset [94] comprises 10 types of arm gestures collected from 25 participants: arm up/down, push, pull, draw a circle, draw a zigzag, clap hands, mimic knocking on a table, yawn, and lift both arms up. Each gesture was recorded in seven environments, including a bedroom and a parlor. Santhalingam et al. [129] collected 50 different ASL signs from 15 participants in three environments (classroom, laboratory, and conference room) and five different scenarios for single and multiple users. Data were recorded simultaneously with a radar, a Kinect, and an RGB camera. Deep-soli is a dataset introduced by Wang et al. [161] that comprises 11 hand gestures performed by 10 users above Google Soli: pinch index/pinky, finger slide, finger rub, slow/fast swipe, push, pull, palm tilt, circle, and palm hold. Deep-Soli was also used by Berenguer et al. [22] to evaluate GestureVLAD, their proposed framework for radar gesture recognition. Ritchie et al. [121] introduce Dop-net, a dataset of four gestures recorded with two different radar sensors (CW and FMCW). Only the FMCW dataset was released in the form of a challenge. Dop-net was used in another paper to compare the performance of CW and FMCW radars for gesture recognition [21]. Finally, HARrad consists of two datasets of six gestures of arm and hand gestures collected from nine participants with a 77GHz radar [149]: drumming, shaking, swiping left/right, and thumb up/down. HARrad events contains six different actions, including entering or leaving a room, sitting down, standing up, clothe, and unclothe.

From the above analysis, we make the following observations:

- O₈. *Limited availability of datasets*: very little datasets are available for benchmarking.
- O₉. *Huge size of datasets*: the size of the raw data captured for a single gesture is very high, making the entire dataset huge in size and challenging to process. Little or no dimension reduction has been observed.

In this paper, we acquire a gesture dataset with a wide range of mid-air finger, hand, and arm gestures, using an off-the-shelf sensor. The dataset is publicly available, as opposed to most gesture sets that we identified in the literature, and its recording process can be replicated by following the procedure described in Section 5.

4 RADARSENSE: A PIPELINE FOR RECOGNIZING RADAR-BASED GESTURES

4.1 Overview of the Pipeline

We present a new radar data processing pipeline for mid-air hand gesture recognition that reduces the high dimensionality of raw radar signals to only two physically meaningful features independent of the radar. The pipeline is composed of eight steps, according to a principle-based approach, as follows (Fig. 12):

- (1) **Raw data capture**. Raw data is captured from each radar antenna (Fig. 13a). Depending on the radar type, data may be in the time or frequency domain. For instance, the Walabot provides data in the time domain.
- (2) **Fast Fourier Transform (FFT)**. The radar signal is transformed from the time domain to the frequency domain if necessary. The minimum and maximum frequencies are then set, depending on the observed frequency spectrum, as a trade-off between bandwidth and signal-to-noise ratio. This operation is required by the next step.
- (3) **Removal of radar source and antenna effects**. The radar equation [71, 72] is applied to the raw frequency-domain signal to remove the radar source and antenna effects (e.g., internal reflections and transmissions) and antenna-target interactions (Fig. 13b). For this purpose, a model of each pair of radar antennas is constructed using a series of radar measurements taken at increasing distances from a large copper plate (radar calibration) [72]. By applying the radar equation, the resulting filtered signal is a Green's function that is independent of the radar: the signal becomes normalized and free from undesired echos, thereby making the user's reflections stand out. It is worth noting that the Green's function used in the radar equation of Lambot and André [71], Lambot et al. [72] represents the backscattered x -component of the electric field for a unit x -directed electric source for wave

propagation in 3D planar layered media. Hence, the function obtained in our application with the user as the reflector becomes an equivalent Green function. For instance, this equivalence concept was also used by Lopera et al. [97] for landmine detection. Complex targets can be modeled using 3D numerical methods, such as the finite-difference time-domain (FDTD) method (e.g., [169]), but calculation cost becomes then huge and inversion becomes extremely complex or even ill-posed for this application.

- (4) **Removal of the background scene.** Using the superposition principle, the first frame, for which the user hand is not in the field of interest, is subtracted from the radar signal to remove any remaining reflections of static reflectors, such as walls, furniture, or other objects (Fig. 13c). This step ensures accurate feature extraction during the inversion step, as reflections from other (static) objects could be confused with the user's hand.
- (5) **Inverse Fast Fourier Transform (IFFT).** The filtered radar signal is transformed to the time domain to simplify the computations during the next steps.
- (6) **Time gating.** The time-domain data are truncated to keep only the portion of the signal that is relevant for gesture recognition (Fig. 13d). The signal received within a given time window, and therefore range, is kept. This approach removes useless information, such as dynamic objects that are further away from the user, and further improves accuracy and reduces the processing time of the next steps. The user body, which may move during gesture recording, is also removed by time gating. Static objects were removed in the previous step.
- (7) **Inversion.** Two physically meaningful features are extracted from the filtered time-domain radar data by performing full-wave inversion [72] (Fig. 13e). The first feature, represented by the hand-radar distance, provides information about the hand trajectory. The second feature, represented by the effective permittivity of the medium defined by the hand, gives information on the configuration of the hand (e.g., whether the palm is facing the radar). The permittivity is effective in our case as it physically represents the permittivity of an equivalent infinite plane and not the physical permittivity of the hand, given the Green function defined above. An equivalent model is used because closed-form solutions of Maxwell's equations do not exist for complex targets. This drastic data reduction, while keeping just the essential information, enables the use of simple template matching algorithms for gesture recognition. Reduced data combined with multiple antennas can also improve accuracy with a limited impact on recognition time. It is worth noting that while radar range resolution is typically considered as the capacity to visually distinguish between two targets at different ranges, which is calculated from the bandwidth and wave velocity, full-wave inversion enables much finer resolution (of about one order of magnitude). When two targets are so close that only one reflection is visible, the distortion of that reflection due to the two reflectors can, to some extent, still be detected by inverse modeling.
- (8) **Filtering.** When the effective permittivity is under a given threshold (i.e., $\epsilon \leq 1.10$), the estimated distance and effective permittivity values are discarded and set to the default values of 75 cm (29.5 in) and 1.0, which removes most of the incorrect measurements when the hand is not clearly detected (e.g., too far from the radar or in a position that does not reflect well the waves towards the radar sensor). In addition, a moving-average filter with a window of length 8 is applied to smooth variations caused by errors from the inversion process (Fig. 13f). In the absence of filtering, abrupt changes in the estimated distance and/or permittivity values could negatively impact gesture recognition accuracy. The need for filtering may be reduced with radars with a larger dynamic range.

Fig. 13 shows the radar signal of the “push with palm” gesture recorded by one of the antenna pairs of the Walabot throughout the main steps of the RadarSense pipeline. Before any filtering is applied (Fig. 13a), the signal of the hand is lost in the imperfections of the radar. We can only identify a strong signal very close to the radar, mostly caused by

internal reflections and transmissions. After removing the radar source and antenna effects (Fig. 13b), the reflection of the user's hand becomes visible as an arc at the top of the image. Some parasitic reflections, probably caused by the box supporting the radar, can be identified at the top of the image. After removing the background scene (Fig. 13c), the static reflections are removed and therefore the signal is slightly cleaner. After time gating (Fig. 13d), all the reflected signals received after 4.5ns are discarded. The reflection of the user's hand thus occupies most of the image. Some visualisation

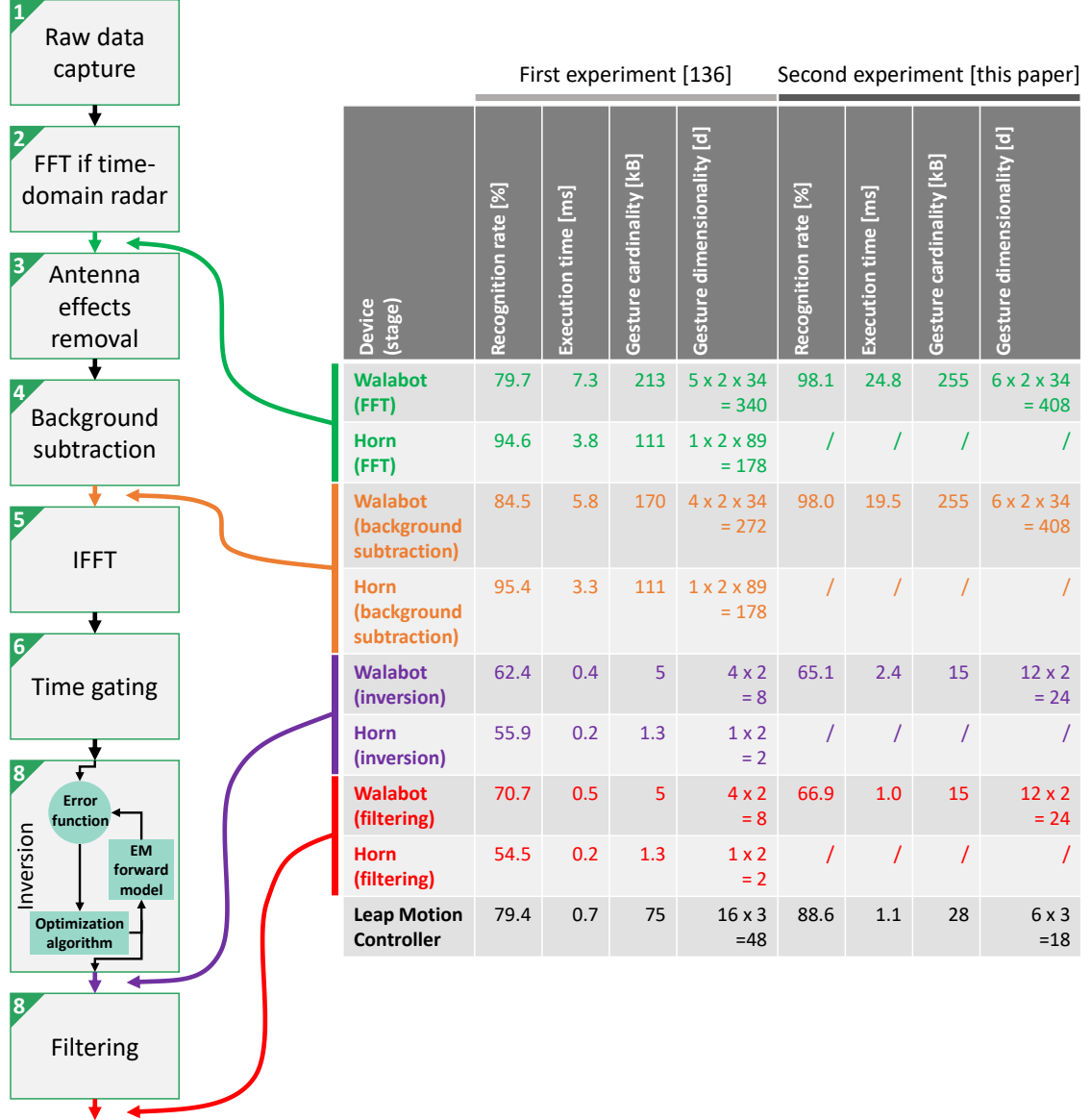


Fig. 12. The radar data processing pipeline. The recognition results of the best performing configurations are displayed for four significant steps of the pipeline in the user-dependent scenario: FFT (Fig. 21a, 23a), background subtraction (Fig. 21b, 23b), inversion (Fig. 21c, 23c), and filtering (Fig. 21d, 23d). Gesture size is estimated for a gesture duration of 2s (*i.e.*, 20 Horn frames, 80 Walabot frames, and 200 LMC frames, respectively).

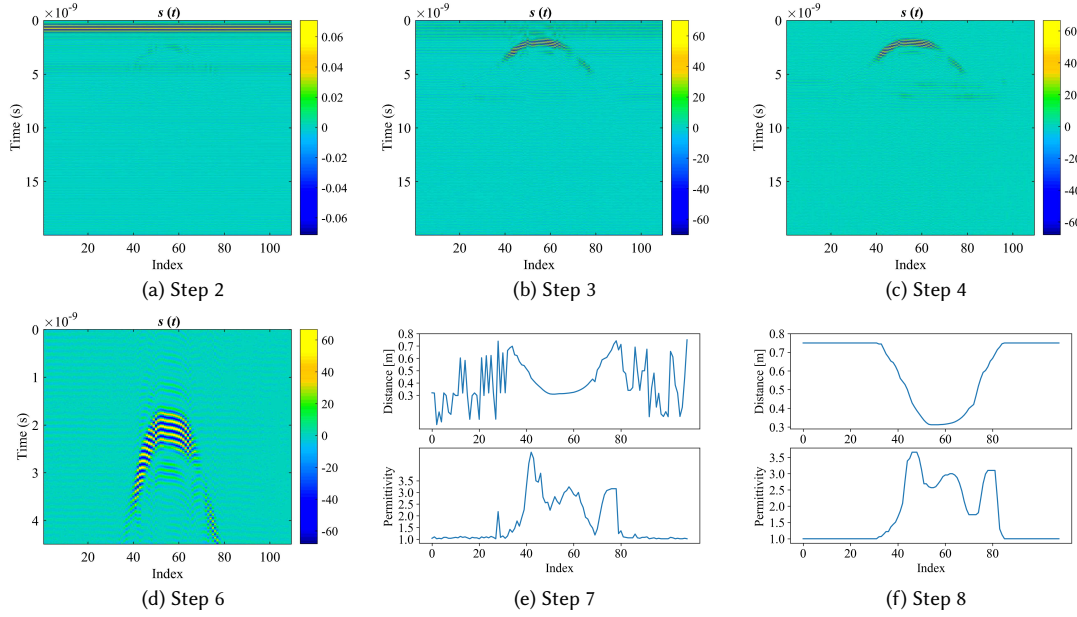


Fig. 13. Intermediate results of the processing of the “push with palm” gesture recorded with one antenna pair of the Walabot.

techniques facilitate the analysis of radar images by humans [118]. After inversion (Fig. 13e), two parameters are extracted. The upper graph shows the hand-radar distance. The distance oscillates between 0 and 70 cm (27.6 inches) when the hand is away from the radar, as the inversion process mostly fails to retrieve a correct distance value when the hand is not visible by the radar. However, the distance is accurately extracted when the hand moves in front of the radar (from index 35 to 80). The lower graph shows the relative permittivity, which stays around 1.0 (free-space relative permittivity) when the hand is away from the radar, and hovers around 2.5 when the hand is in front of the radar. After filtering (Fig. 13f), the incorrect estimated distances are set to the default value of 75 cm (29.5 in) and small imperfections are smoothed out, resulting in a cleaner signal.

Practitioners can employ only a part of the pipeline, depending on specific constraints imposed by the usage context. Fig. 12 also shows for steps no. 2, 4, 7, and 8 the best recognition accuracy rate obtained in our evaluation (Section 6) as well as the corresponding execution time and average gesture size in these processing steps. The latter corresponds to the estimated space occupied by a typical gesture in system memory and depends on the duration of the gesture, *e.g.*, 2s, the number of antenna pairs, and the size of the data at each processing step (*e.g.*, steps 7 and 8 return two 64-bit floating point numbers for each antenna pair).

4.2 Rationale Behind Modeling by Inversion

Due to the dependence of recognition algorithms to the radar type (mostly, CNNs - see O_6), we wanted to pursue some independence of the gesture recognition with respect to the radar type, which required some model-based approach.

By adopting a model-based approach by inversion as proposed in this paper, the physical features are extracted from the raw radar data not through the CNN but with full-wave modeling and inversion, namely, a physical distance and an effective permittivity related to the reflection strength, both dynamic in time. These features are meant to be independent of the radar for a given bandwidth. Nevertheless, different radar antenna radiation patterns may still affect the results as the radiation pattern is not accounted for in the full-wave electromagnetic model. Gestures that



Fig. 14. The main steps of a gesture recording.

deviate too much from the antenna axis may be affected, to some extent, by that simplification [173]. Accounting for the antenna radiation pattern is possible, but either (1) the radar-antenna calibration procedure then becomes a heavy task as the radar-antenna characteristic functions have to be determined for all angles in both the E- and H-planes, or (2) the more complex, generalized radar equation has to be used [71]. Nevertheless, if the gestures do not span a too large angle with respect to the main radiation lobe of the radar, the effects on the gesture recognition can be made negligible.

The main benefits of the model-based approach by inversion are:

- The radar equation permits to filter out the radar-antenna effects for a given frequency range, including the source, the antenna internal reflections, and the radar-target interactions. The resulting signal is a Green function [72] that represents a normalized quantity, which still preserves the independence from the radar type, *e.g.*, frequency versus time domain radar. Thus, a training dataset can be applied to different radar types and especially to different radars of the same model, knowing that two identical radar models generally produce (slightly) different signals.
- The inversion process produces two physical quantities, *i.e.*, the target distance and the effective permittivity, from the radar signal. This significantly reduces the quantity of data (2-3 orders of magnitude) while preserving the useful information contained in the radar signal. This represents a major asset for designing and feeding any subsequent recognizer, whether template-based or deep learning. In particular, the architecture of the recognition algorithm can be simplified and memory usage can be significantly reduced. This is especially important for real-time processing.
- As the quantities provided by the full-wave inversion benefit from a physical meaning, an interpretation of the gesture can be considered, thereby providing valuable insights into the gesture types that can be recognized as well as for designing and optimizing the recognition approach.

5 DATASET

We collected an extension of the gesture dataset from [136]. Our dataset contains the same 16 gesture types (Fig. 16), that we collected from three participants and two sensors: the European version of the Walabot and LMC as the baseline. The participants were all men aged 56, 45 and 23 years old. Their height were 202, 183, and 176 cm (79.5, 72.0, and 69.3 in), and their weight 101, 76, and 78 kg, were respectively. Lee et al. [80] provided a formula to estimate the hand surface area (HSA) from its length and circumference: Estimated HSA (cm^2) = $1.219 \times \text{Hand length (cm)} \times \text{Hand circumference (cm)}$. Based on this formula, we computed the following HSA values for our participants, respectively: 848, 548, and 525 cm^2 (132, 85, and 81 in^2).

Each participant produced 10 repetitions of each gesture type, resulting in a total of $16 \text{ (gestures)} \times 3 \text{ (participants)} \times 10 \text{ (repetitions)} = 480$ recordings per sensor. The Walabot was placed on a table at a distance of about 70 cm (27.6 in) in front of the subject (Fig. 15). An empty cardboard box separated the Walabot from the table by a distance of about

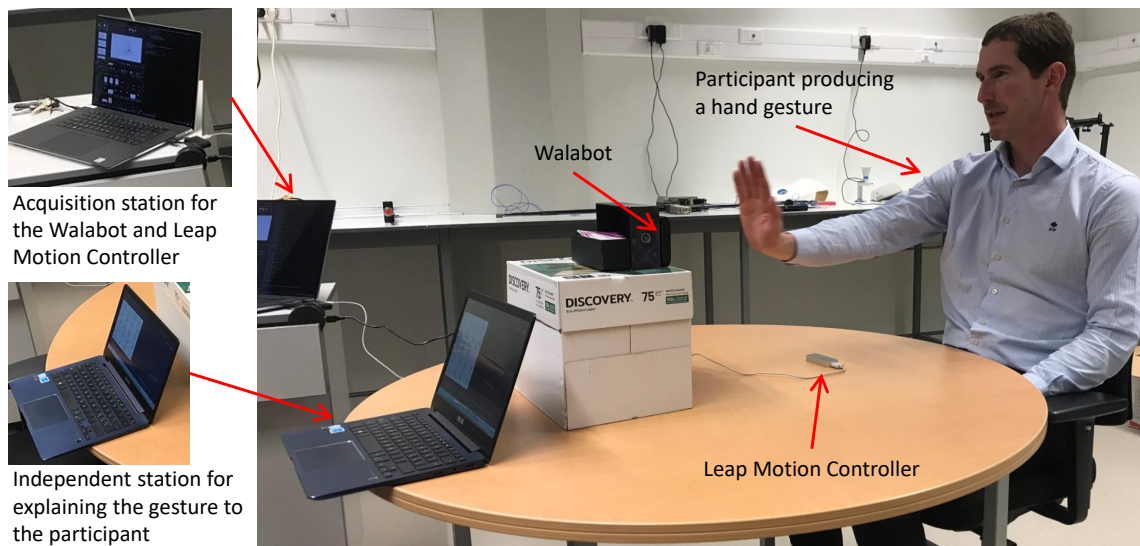


Fig. 15. Setup used for acquiring the hand gesture datasets.

26 cm (10.2 in). The LMC was placed on the desk. The following steps were repeated for each gesture sample: (1) the subject places their hands on their lap (Fig. 14a), (2) the experimenter starts the recording and prompts the participant to perform the gesture, (3) the participant performs the gesture (Figs. 14b and 14c), (4) the participant puts his hands back on their lap (Fig. 14d), and (5) the experimenter stops the recording. Each gesture was recorded simultaneously with 12 antenna pairs of the Walabot (PROF_SENSOR_NARROW) at around 40 frames per seconds and with the LMC at around 100 frames per second using a custom software.

6 EVALUATION

6.1 Procedure and Apparatus

All the studies from this section follow the same testing procedure. *Recognition rate* (i.e., the ratio of correctly recognized gestures divided by the total number of trials) and *execution time* (i.e., the time to process a candidate gesture) were calculated for each combination of variables (e.g., the number of training templates) in two different scenarios [152]: (1) *user-dependent*, in which the training and testing samples are collected from the same user, and (2) *user-independent*, in which the training samples are collected from different users than the gestures used for testing. For each combination of variables and for each type of gesture, we repeated the following steps 100 times: (1) select a random testing sample, (2) train the gesture recognizer using a set of randomly selected samples, and (3) evaluate the recognition accuracy. We use the Jackknife recognizer [146] that was introduced as a modality-agnostic approach to gesture classification. The evaluation was carried out on a laptop with an Intel i7-10875H CPU and 32GB of DDR4 RAM running Windows 10.

For each testing scenario and combination of variables, a confusion matrix was generated. The confusion matrix visualization enables one to quickly identify the reasons for a specific value of recognition rate, as it highlights which gestures were confused together by the recognizer. The diagonal elements of a confusion matrix indicate correctly predicted gestures, while off-diagonal elements indicate mispredicted gestures. Ideally, we would like the diagonal to stand out as much as possible, as it would indicate low confusion between gestures. In this section, we show only the

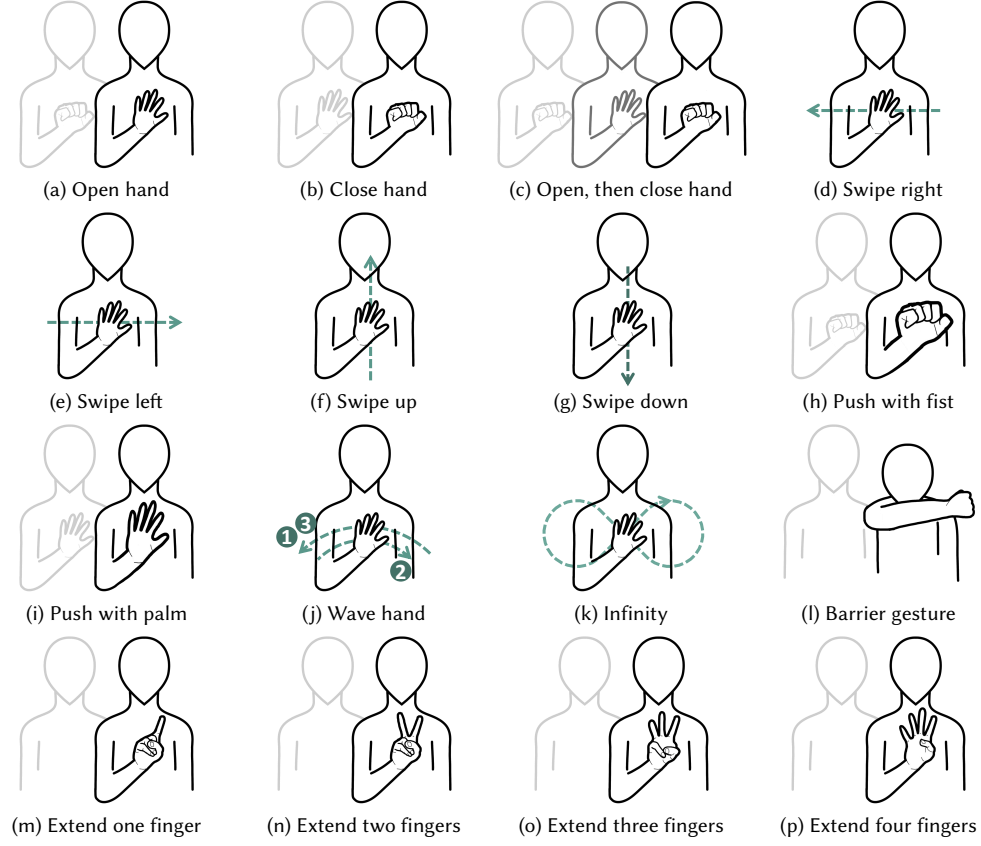


Fig. 16. The sixteen gesture types used in our dataset [136].

confusion matrices of the most relevant configurations. We also generated graphs that display the evolution of the recognition rate with respect to the number of sampling points and training samples. They enable us to identify how stable the accuracy of a recognizer is as the number of sampling points changes. In general, we expect the recognition rate to increase with the number of sampling points and training samples, up to a certain point [152]. Finally, we built two tables (Tables 3 and 4) summarizing the recognition rate and execution time of the recognizer measured in the best-performing configuration for all the sets of antenna pairs tested in four steps of the RadarSense pipeline. They enable an easy comparison of accuracy for the different pipeline steps and sets of antenna pairs.

6.2 Leap Motion Controller

In this first study, we evaluate the accuracy of the Jackknife recognizer on the raw LMC data, following the procedure described in Section 6.1. The study has a within-subject design with three independent variables:

- (1) NUMBER-OF-TEMPLATES: numerical variable with 4 conditions in the user-dependent scenario and 5 conditions in the user-independent scenario, representing the number of samples per gesture type used to train the recognizer: $T=\{1, 2, 4, 8\}$ (user-dependent) and $T=\{1, 2, 4, 8, 16\}$ (user-independent), respectively.
- (2) NUMBER-OF-POINTS: numerical variable with 37 conditions representing the number of points used to represent a gesture: $N=\{x \in \mathbb{N} \mid 4 \leq x \leq 40\}$.

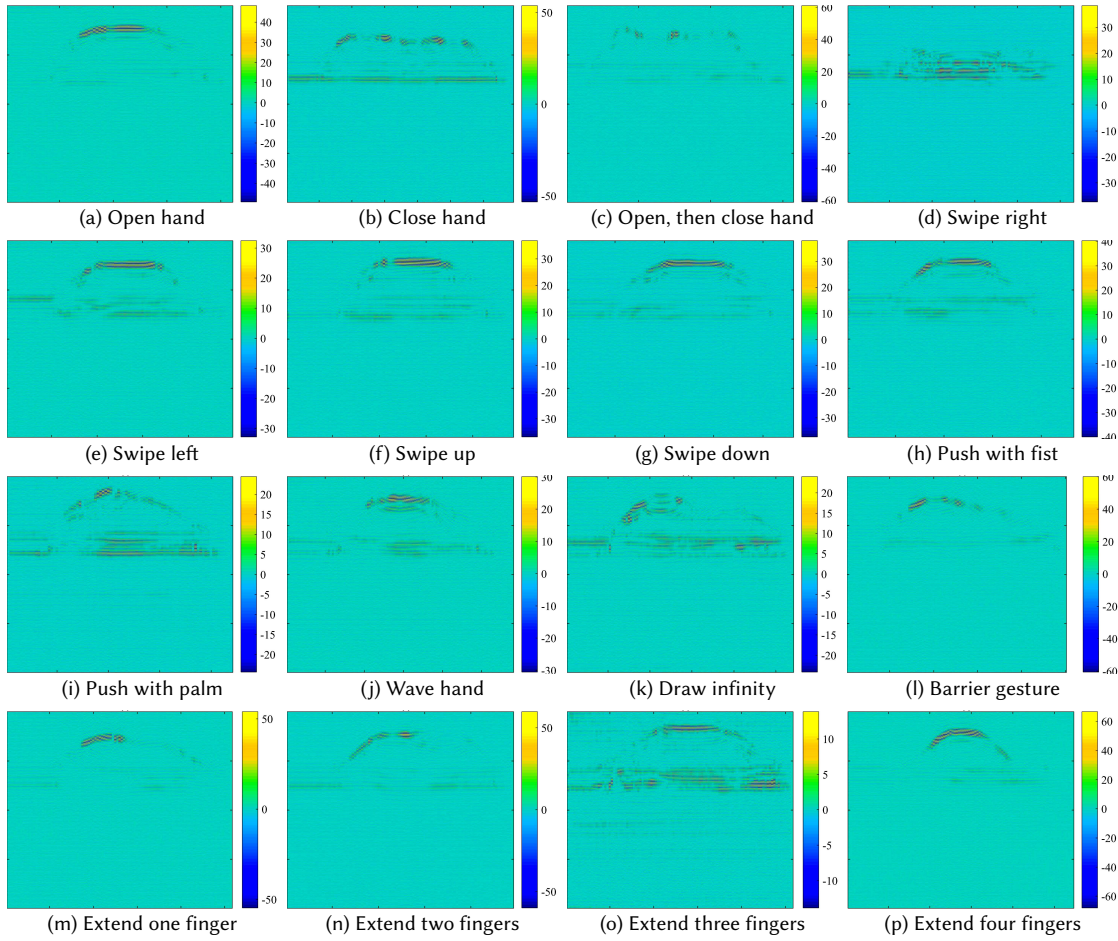


Fig. 17. The sixteen gestures used in our dataset displayed in the time-domain (after step 6).

- (3) **NUMBER-OF-JOINTS**: numerical variable with three conditions (Fig. 20), representing the number of joints employed by the recognizer: $J \in \{6, 11, 16\}$.

Each gesture sample consists of a sequence of frames, where each frame has two elements: a timestamp and a vector of length $3 \times J$ that result from the concatenation of the 3D coordinates (x, y, z) of the selected joints. Fig. 18 shows the evolution of the recognition accuracy rate with respect to the number of sampling points per gesture in user-dependent and user-independent scenarios. The accuracy rate is mostly stable for higher values of N , but decreases for $N < 5$ (user-dependent) and $N < 16$ (user-independent). In addition, the impact of the set of joints seems to be very limited, although accuracy rate is slightly higher for $J=6$. A smaller set of joints could accommodate better small variations in gesture execution, especially across gestures produced by different users. In both scenarios, the accuracy rates increase with the number of training samples per gesture type. However, the gain provided by more training samples becomes smaller as the number of samples increases.

The average recognition accuracy rate is high in the user-dependent scenario, reaching 88.6% in one configuration ($T=8, J=6, N=37$), and the execution time is short (1.1 ms). In this configuration, 10 of the 16 gesture types are recognized

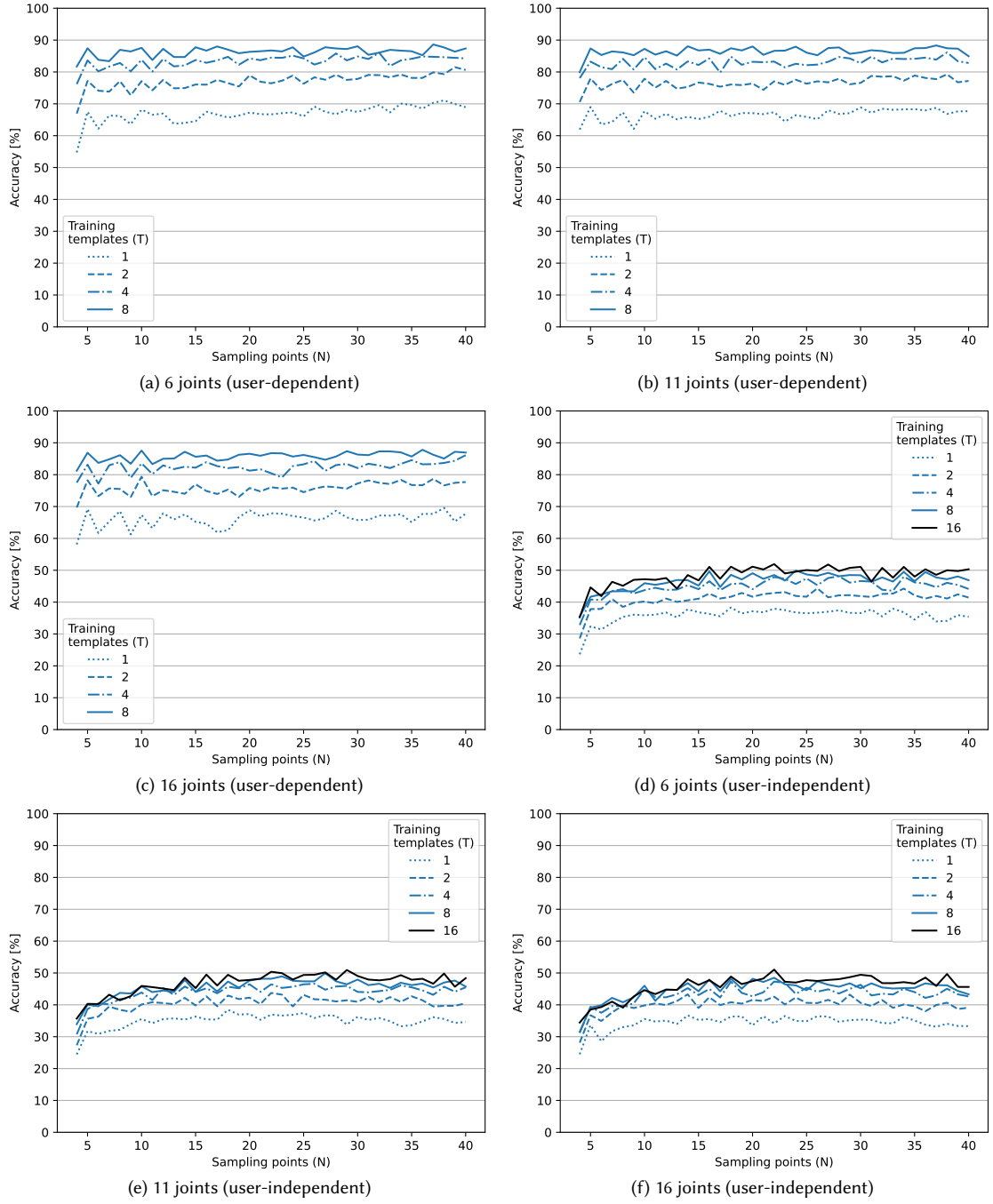


Fig. 18. Gesture recognition accuracy with respect to the number of sampling points and training samples on the LMC dataset in the user-dependent (top) and user-dependent scenario (bottom).

with over 90% accuracy (Fig. 19a). For some gesture classes, such as the “barrier gesture”, performed close to the body and at the limit of the LMC sensing range, the accuracy rates are lower. Also the “extend one finger” gesture was often confused with “push with fist” because they both present the same motion, although different hand poses.

The average recognition rate drops sharply in the user-independent evaluation scenario, reaching a maximum of 51.9% ($T=16$, $J=6$, $N=22$). The execution time is still fast, about 1.2 ms, in this configuration. Increasing the number of training samples from $T=8$ to $T=16$ has a limited impact on the recognition accuracy rates. Fig. 19b shows the confusion matrix for the best configuration with two notes: (1) there is a lot of confusion between similar gestures and (2) some of the gestures are not correctly recognized. Both issues were also visible in the user-dependent evaluation, but are amplified by the user-independent scenario. Three groups of gestures are concerned by the first issue: (1) gestures no. 1, 2, and 3 (“open hand”, “close hand”, and “open then close hand”), (2) gestures 1, 8, 9, and 13 (“open hand”, “push with fist”, “push with palm”, and “show one finger”), and (3) gestures 13 to 16 (“show one/two/three/four finger(s)”). LMC does not identify correctly the exact hand pose of these gestures, resulting in high similarity. Gesture 12 (“barrier gesture”) is subject to the second problem, as it is executed just outside the field of view of the LMC. Thus, its motion is not always sensed by LMC.

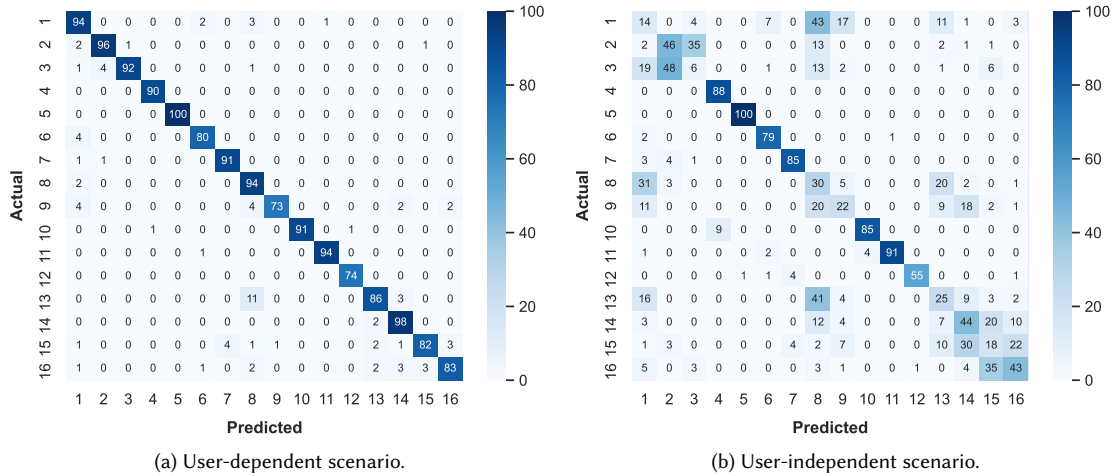


Fig. 19. Normalized confusion matrices for the best configuration with data acquired from the LMC in the user-dependent and user-independent testing scenarios. The values are represented as percentages.

6.3 Walabot

In this second study, we evaluated the impact of various parameters, including the processing step, the number of training samples and the selected pairs of antennas, on the accuracy of the Jackknife gesture recognizer in two testing scenarios. This study follows the procedure described in Section 6.1 and has four independent variables:

- (1) PROCESSING-STEP: nominal variable with 4 conditions, representing the processing step of the training data: FFT, Background subtraction, Inversion, Filtering.
- (2) NUMBER-OF-TEMPLATES: numerical variable with 4 conditions in the user-dependent scenario and 5 conditions in the user-independent scenario, respectively: $T=\{1, 2, 4, 8\}$ (user-dependent) and $T=\{1, 2, 4, 8, 16\}$ (user-independent).

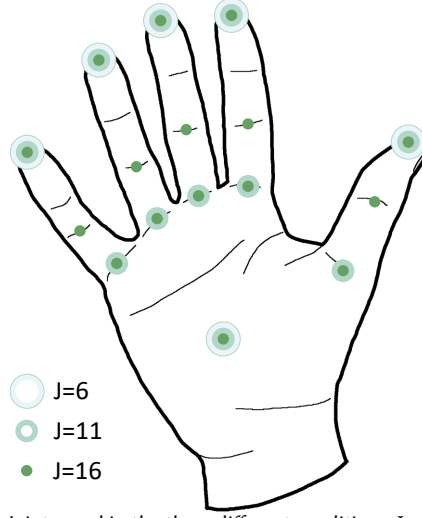


Fig. 20. Joints used in the three different conditions $J \in \{6, 11, 16\}$.

- (3) **NUMBER-OF-POINTS**: numerical variable with 37 conditions representing the number of points per gesture: $N = \{x \in \mathbb{N} \mid 4 \leq x \leq 40\}$.
- (4) **ANTENNA-PAIRS**: nominal variable with 15 conditions, representing unique antenna pairs, the two sets of six non-redundant antenna pairs, and the combination of the 12 antenna pairs: $AP = \{(1), (2), (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (1, 2, 3, 6, 8, 9), (4, 5, 7, 10, 11, 12), (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)\}$.

Regarding the independent variable **ANTENNA-PAIRS**, only a small and relevant subset of all the 4095 possible combinations was tested. In the inversion and filtering steps, the vector (length $2 \times AP$) is the concatenation of the distance and permittivity values retrieved from all the antenna pairs. In the other steps, the vector (length $2 \times 34 \times AP$) results from the concatenation of the real and imaginary parts of the frequency-domain radar signal (34 frequencies) from the pairs of antennas.

6.3.1 FFT (Step 2). We obtained a high accuracy rate in the user-dependent scenario (Table 3) with 98.1% in a configuration ($T=8$, $N=39$, $AP=(4, 5, 7, 10, 11, 12)$), which is 10% higher than the accuracy rate achieved with the LMC data. This result is due to the fact that the Walabot can pick up small differences in hand pose better than LMC and the different positions of the two sensors. However, the execution time is longer than 24.8 ms. The accuracy rate increases with the number of training samples T and the number of sampling points N , but growth slows as N increases. Fig. 23a shows the confusion matrix for the best performing configuration. The recognition accuracy rate is greater than 90% for all gesture types and reaches 100% for 7 out of the 16 gestures.

The recognition accuracy rate decreases in the user-independent scenario (Table 4), with a maximum of 13.2% ($T=16$, $N=39$, $AP=(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)$). In this configuration, recognition takes 86ms on average. Increasing the number of training samples does not significantly improve the recognition rate (Fig. 22a). The impact of the set of antenna pairs on the recognition accuracy is small, with the worst performing set achieving only 10.1% accuracy. Fig. 24a shows that only gesture no. 12 (“barrier gesture”) was correctly identified. This gesture differs from all of the others in terms of scale and articulation, which makes it easier to differentiate. On the other hand, the accuracy rate stays below 33% for the other gestures. This result is due to a combination of factors: the motion of the hand is similar for gestures

AP	FFT				Background subtraction				Inversion				Filtering			
	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>
1	89.3 (9.7)	4.8	8	38	88.4 (11.4)	3.6	8	32	21.8 (11.6)	1.3	8	38	32.4 (14.0)	0.3	8	22
2	89.9 (9.5)	5.5	8	40	90.6 (7.2)	3.9	8	33	19.1 (9.5)	1.5	8	39	27.2 (9.2)	0.3	8	21
3	87.7 (8.1)	4.1	8	35	89.5 (9.4)	4.1	8	35	22.1 (13.7)	1.4	8	39	27.3 (12.2)	0.2	8	12
4	90.9 (6.7)	4.4	8	39	89.8 (7.9)	4.8	8	40	17.2 (11.6)	1.4	8	37	27.9 (11.1)	0.7	8	36
5	89.2 (7.9)	3.9	8	36	90.1 (6.9)	4.6	8	39	17.9 (10.9)	1.1	8	32	24.7 (11.6)	0.2	8	14
6	87.2 (9.4)	4.1	8	38	89.1 (8.7)	4.4	8	39	28.5 (15.0)	1.5	8	39	35.6 (15.5)	0.7	8	39
7	89.2 (7.5)	4.5	8	40	89.4 (10.2)	4.8	8	40	23.3 (9.1)	1.2	8	34	29.4 (15.3)	0.3	8	18
8	89.3 (9.1)	4.4	8	40	90.4 (6.4)	3.9	8	36	21.6 (11.3)	1.4	8	38	29.2 (9.0)	0.3	8	16
9	89.6 (7.2)	4.0	8	37	91.2 (5.9)	4.2	8	38	20.8 (8.3)	1.2	8	34	31.5 (12.4)	0.6	8	32
10	87.9 (9.2)	4.1	8	38	90.4 (7.2)	4.3	8	37	21.4 (9.3)	1.5	8	39	33.2 (9.8)	0.4	8	26
11	89.1 (6.6)	4.4	8	39	90.1 (7.9)	4.5	8	38	22.4 (11.1)	1.3	8	37	30.9 (12.5)	0.6	8	36
12	87.2 (8.9)	4.4	8	40	90.6 (6.7)	4.7	8	40	25.9 (9.5)	1.2	8	37	35.0 (14.5)	0.3	8	24
1, 2, 3, 6, 8, 9	97.7 (2.2)	21.2	8	36	98.0 (2.6)	19.5	8	35	61.3 (14.4)	1.7	8	35	65.7 (13.3)	1.7	8	38
4, 5, 7, 10, 11, 12	98.1 (2.6)	24.8	8	39	97.7 (2.2)	21.2	8	38	58.7 (12.1)	1.3	8	31	63.1 (10.0)	1.0	8	26
All 12 pairs	97.9 (2.2)	42.1	8	38	97.8 (2.6)	37.9	8	35	65.1 (15.6)	2.4	8	36	66.9 (8.3)	1.0	8	21

Table 3. Recognition rates and execution times obtained for the Walabot data at four steps of the RadarSense pipeline for all the tested antenna pairs (user-dependent scenario). $T=8$, best value of N . AP=antenna pairs, RR=recognition rate [%], ET=execution time [ms].

AP	FFT				Background subtraction				Inversion				Filtering			
	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>	RR (<i>M, SD</i>) [%]	ET [ms]	<i>T</i>	<i>N</i>
1	10.5 (19.1)	4.3	16	22	10.6 (20.7)	5.6	16	26	13.4 (13.4)	0.4	16	11	14.1 (10.8)	0.3	16	10
2	10.4 (20.7)	8.4	16	38	9.9 (17.7)	5.6	16	25	11.5 (8.4)	2.9	16	40	13.6 (9.4)	1.0	16	27
3	10.8 (21.3)	7.5	16	35	10.8 (22.0)	8.4	16	36	12.8 (9.9)	0.2	16	7	16.2 (10.4)	0.4	16	12
4	10.1 (18.9)	7.6	16	35	10.0 (18.8)	9.2	16	40	10.7 (6.8)	0.1	16	5	13.4 (8.7)	0.3	16	10
5	12.4 (21.1)	7.0	16	33	10.8 (20.3)	4.6	16	23	12.4 (10.8)	0.2	16	8	16.5 (8.0)	1.0	16	25
6	10.2 (19.1)	4.0	16	21	10.4 (16.1)	0.8	4	16	15.2 (11.0)	1.2	16	21	18.6 (16.2)	0.5	16	14
7	10.3 (20.2)	2.1	8	22	10.3 (21.7)	5.9	16	26	11.7 (8.7)	0.5	4	36	17.0 (13.2)	0.4	8	18
8	11.6 (24.3)	5.6	16	27	10.9 (21.6)	5.0	16	23	13.3 (9.1)	1.2	16	20	17.4 (12.2)	0.5	16	12
9	11.1 (21.7)	3.4	8	32	11.7 (23.2)	8.9	16	40	13.4 (8.7)	0.1	16	5	17.8 (8.8)	0.4	16	10
10	11.9 (22.1)	6.3	16	30	11.8 (22.3)	7.3	16	34	13.1 (10.1)	0.2	16	7	16.0 (13.8)	0.5	16	13
11	11.6 (20.5)	4.2	8	38	11.6 (22.2)	8.1	16	37	12.9 (8.5)	0.3	16	9	15.8 (9.1)	1.1	16	30
12	10.4 (19.6)	1.4	8	16	9.9 (19.9)	4.3	8	37	16.0 (9.9)	0.5	16	10	18.8 (12.4)	1.3	16	39
1, 2, 3, 6, 8, 9	12.9 (20.9)	35.6	16	32	11.5 (17.7)	18.8	8	34	18.1 (18.9)	2.5	16	29	18.8 (16.9)	2.2	16	32
4, 5, 7, 10, 11, 12	12.2 (21.0)	45.0	16	40	11.4 (19.6)	22.2	8	40	16.2 (15.2)	3.4	16	34	18.6 (14.6)	1.6	16	23
All 12 pairs	13.2 (21.2)	86.4	16	39	11.4 (20.7)	73.8	16	31	17.3 (16.1)	3.2	16	32	18.8 (19.8)	2.3	16	24

Table 4. Recognition rates and execution times for the Walabot dataset at four steps of the RadarSense pipeline for all the tested antenna pairs (user-independent scenario). $T=16$, best value of N . AP=antenna pairs, RR=recognition rate [%], ET=execution time [ms].

with low recognition accuracy and the size of the hand varies between participants. Because hand poses are mostly identified using the intensity of the reflected radar signal, a difference in hand size may be interpreted as two different hand poses, resulting in a different gesture type. This aspect suggests that a normalization of the reflection amplitude may be required and/or that recognition should give priority to the relative dynamics of the frames. The lower accuracy in the user-independent scenario may also be explained in part by slight differences in gesture articulation between the different users, as mentioned in Section 6.2.

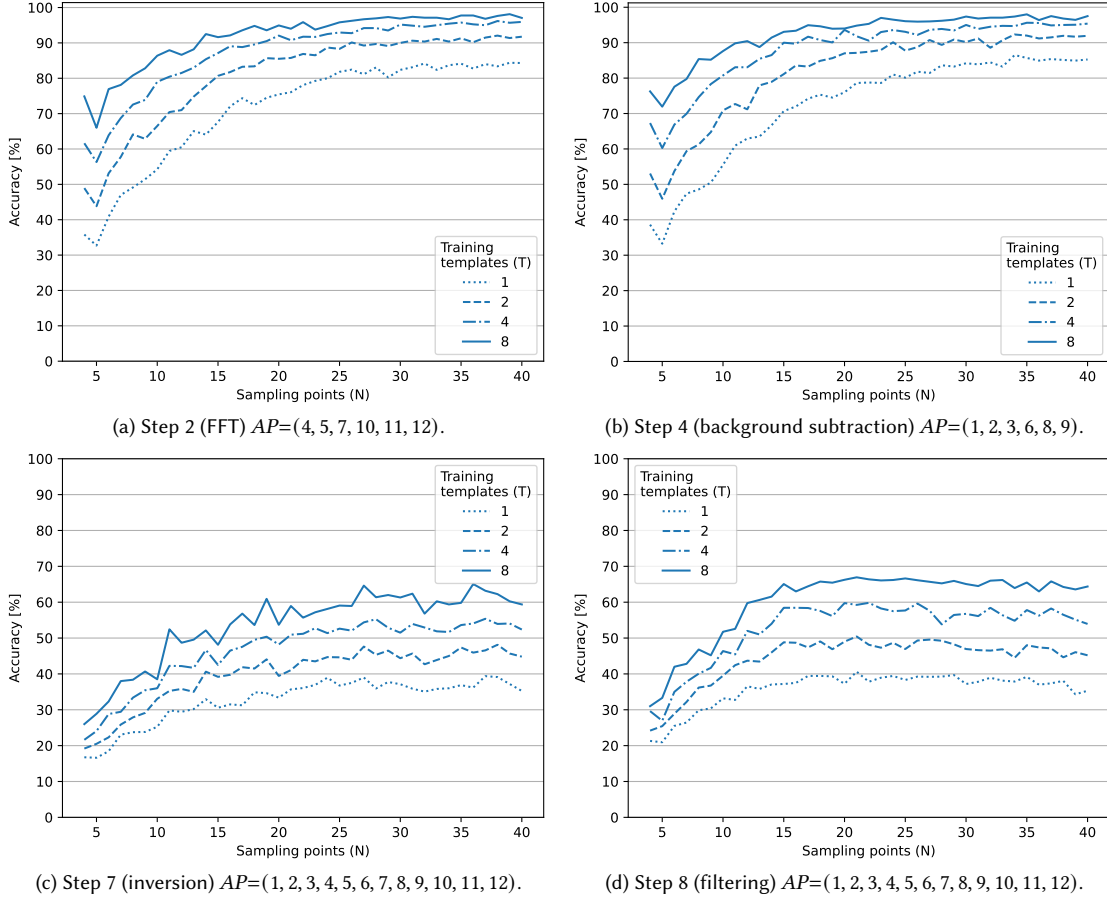


Fig. 21. Gesture recognition accuracy rates function of the number of sampling points and training samples for the best performing antenna pair with data acquired from the Walabot at four different steps of the pipeline (user-dependent scenario).

6.3.2 Background Subtraction (Step 4). The results for step 4 are similar to those obtained after step 2 for both the user-dependent and user-independent scenarios. In the user-dependent scenario, recognition accuracy reaches 98.0% ($T=8$, $N=32$, $AP=(4, 5, 7, 10, 11, 12)$) with an execution time of 19.5 ms. Fig. 21b shows the confusion matrix for this configuration. The recognition accuracy rate is above 90% for all but one gesture type and reaches 100% for 6 of the 16 gestures. The “push with palm” gesture was recognized less accurately (88%) as it was mostly confused with the “extend one/three/four finger(s)” and “push with fist” gestures, which have different hand poses but similar motion. The accuracy rate of individual antenna pairs is still high, between 88.4% and 91.2%.

In the user-independent scenario, accuracy rates reached a maximum of 11.8% ($T=16$, $N=34$, $AP=(10)$), lower than those obtained in step 2. In this configuration, execution time is 7.3ms on average. The confusion matrix (Fig. 24b) looks similar to that of the previous step with gesture no. 12 (“barrier gesture”) being correctly identified.

6.3.3 Inversion (Step 7). In the user-dependent scenario, the recognition rate for all sets of antenna pairs after the inversion step is lower than after the first two steps (Table 3). Such a drop is expected because the size of each frame

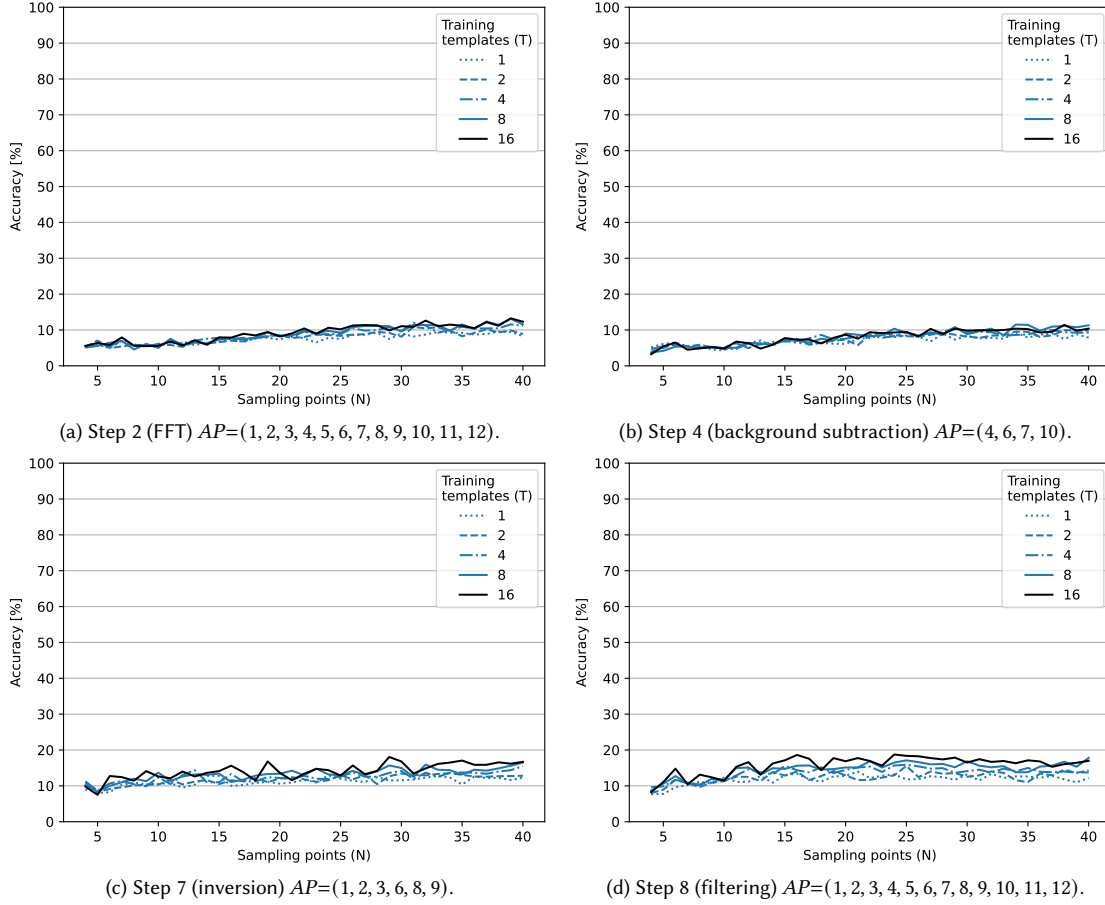


Fig. 22. Gesture recognition accuracy rates function of the number of sampling points and training samples for the best performing antenna pair with data acquired from the Walabot at four different steps of the pipeline (user-independent scenario).

is smaller by a factor of 34 compared to the processing steps. However, we attribute the lower performance to the limited signal-to-noise ratio in the radar waveforms, to which full-wave inversion is particularly sensitive. Echo from the hand may become relatively small as a function of its distance and orientation with respect to the radar. Moreover, a slight change in orientation affects the reflection amplitude, and hence, the retrieved effective permittivity. The best performing combination of antenna pairs delivered a 65.1% accuracy ($T=8$, $N=36$, $AP=(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)$) with an average execution time of 2.4 ms. Fig. 23c shows the confusion matrix for this configuration. Seven of the 16 gestures were recognized with accuracy rates greater than 70%: “open hand”, “swipe right”, “swipe left”, “swipe up”, “swipe down”, “push with palm”, and “wave hand.” Three gestures were recognized with less than 50% accuracy: “close hand” and “extend one/two finger(s).” The high accuracy for some of these gestures can be explained by the nature of effective permittivity and distance metrics. The estimated permittivity increases when the hand opens (larger reflecting surface) and decreases when it closes (smaller reflecting surface), explains the difference between the results of the “open hand”, “push with palm”, and “extend one/two finger(s)” results. Similarly, the “push with palm” gesture is easy to

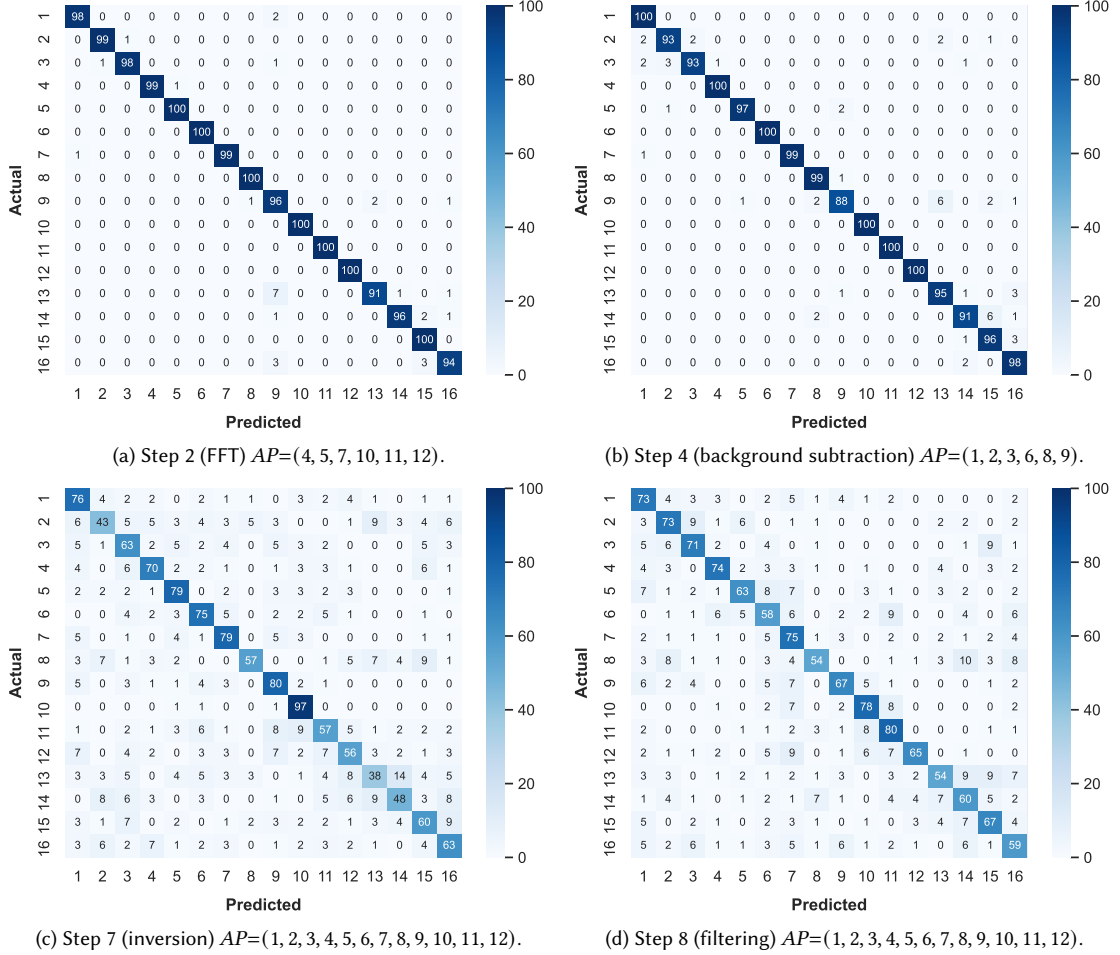


Fig. 23. Normalized confusion matrices for the best configuration with data acquired from the Walabot at four different steps of the pipeline (user-dependent scenario). The values in each cell are percentages.

differentiate from the other gestures. In general, the recognition accuracy rate increases with the number of training samples T and the number of sampling points N .

In the user-independent scenario, the recognition accuracy rates were low regardless of the set of antenna pairs (Table 4), the number of sampling points and training templates (Fig. 22c), but was larger than the ones from the first two steps. The best performing set of antenna pairs resulted in an average accuracy of 18.1% and average execution time of 2.5ms ($T=16$, $N=29$, $AP=(1, 2, 3, 6, 8, 9)$). The low accuracy is reflected by the confusion matrix (Fig. 24c). Only four gesture types with more than 20% accuracy were recognized, namely, “swipe right”, “swipe left”, “swipe down” and “wave hand.”

6.3.4 Filtering (Step 8). The recognition accuracy rate increases slightly in both scenarios after the filtering step and reaches 66.9% ($T=8$, $N=21$, $AP=(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)$) in the user-dependent scenario and 18.8% ($T=16$, $N=24$,

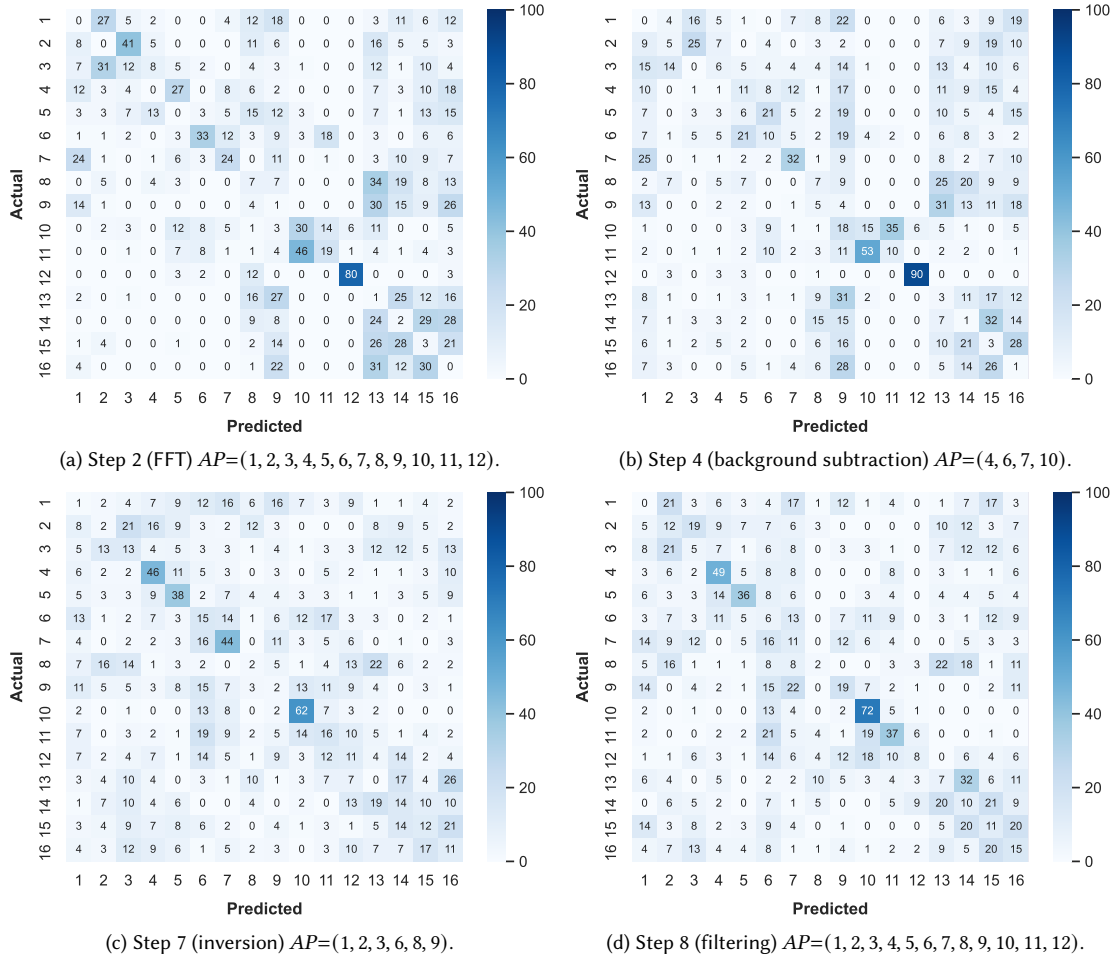


Fig. 24. Normalized confusion matrices for the best configuration with data acquired from the Walabot at four different steps of the pipeline (user-independent scenario). The values in each cell are percentages.

$AP=(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)$) in the user-independent scenario (Tables 3 and 4). Execution time is still fast at 1ms and 2.3ms, respectively. Furthermore, the recognition rate increases faster and is more stable than when using unfiltered inversion data (Fig. 21d). Fig. 23d shows the confusion matrix for the best performing configuration, which is similar to that from step 7 (Section 6.3.3). Accuracy rates increased for some gesture types, such as “close hand”, “draw infinity”, “barrier gesture”, and “extend one/two finger(s)”, resulting in all gesture types being recognized with over 50% accuracy rate.

Despite the filtering, the recognition rate in the user-independent scenario is still low. The confusion matrix (Fig. 24d) for the best performing configuration is similar to the one from the previous step, with the gesture of the “wave hand” being the most accurately recognized at 72%. A more in-depth analysis is provided in Section 6.3.3.

7 DISCUSSION

We now summarize the results of our evaluation. We then draw implications for designing hand gesture interaction with radar sensors and discuss the limitations of this work.

7.1 Summary of the Evaluation

Despite a large number of gesture classes and few training samples, the Jackknife recognizer achieved very high accuracy in the user-dependent scenario, reaching around 98% with data from steps 2 and 4, albeit with relatively slow execution. This is better than what we achieved using LMC data, which sometimes failed to correctly identify the user's hand pose, and is very close to the results found in the literature. For example, Lee et al. [79] obtained 99.1% accuracy on a set of seven hand gestures using a 3D-CNN. Wang et al. [164] reached 96.2% accuracy with their TS-I3D network on a set of 10 hand gestures. We did not observe a significant improvement between step 2 (FFT) and step 4 (background subtraction), probably because the Jackknife recognizer could isolate user hand reflections from the background and antenna effects. Still, we expect data from step 2 to be very beneficial for other algorithms, such as CNNs. Accuracy dropped after the full-wave inversion, reaching 65.1% at step 7 and 66.9% at step 8, which may be explained by the limited signal-to-noise ratio in the radar waveforms, to which the full-wave inversion is particularly sensitive. We expect better results using higher-quality radar systems.

However, our system needs further improvements before it is suitable for user-independent scenarios, as it reached only 13.2% before inversion (steps 2 and 4) and 18.8% after inversion (steps 7 and 8) in the best configurations. The slight improvement observed when using the signal from steps 7 and 8 likely stems from the fact that the signal after inversion is reduced to only two parameters, the distance and relative permittivity, which are less influenced by anatomical differences between users than the raw radar signal. Some approaches proposed in the literature perform significantly better in a user-dependent scenario, such as Choi et al. [29]'s LSTM encoder, which reached 98.5% accuracy when tested with gestures from a new participant after being trained on a dataset of 10 hand gestures performed by 10 participants. Another example is GestureVLAD [22], which reached 91.4% accuracy on a dataset of 11 gesture classes when trained on gestures from nine participants. The low accuracy reached by our system highlights the need for normalization with respect to the user (Section 7.3).

Overall, we observed that accuracy increased generally with the number of training templates T and the number of sampling points N in most of the configurations tested, but the growth in accuracy decreased as N increased. This effect was even more noticeable with the LMC, as the growth seemed to plateau from smaller values of N . Combinations of pairs of antennas generally performed better than isolating one pair of antenna. The spatial separation of the different pairs enables them to collect slightly different information, which helps distinguish between gestures. Higher accuracy could be reached by using fewer, more spaced out (pairs of) antennas, such as in Leem et al. [82].

7.2 Implications for Designing Radar-based Mid-Air Hand Gesture Interaction

11. Use just the steps of the recognition pipeline that are necessary for a given application. Not all the steps of our pipeline are necessary for all the applications. For instance, steps 7 and 8 may not be useful if the pipeline is combined with image classification algorithms, such as a Convolutional Neural Network (CNN). In that case, the data from step 4 may be appropriate, as it provides a signal without antenna effects, noise, and clutter. Other steps may hinder performance without a significant gain in recognition accuracy. In contrast, the data from Step 8 is better suited for simple template matching algorithms that enable end users to enter their own gestures.

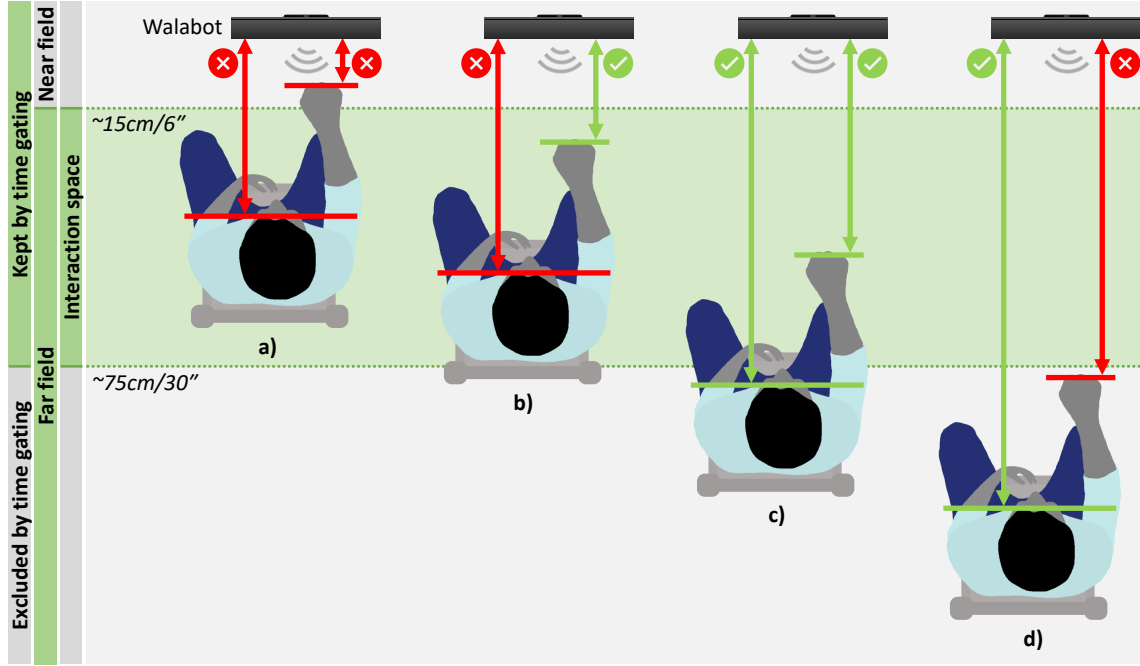


Fig. 25. Hand and body distance constraints for appropriate radar range.

12. Ensure that hand gestures are performed within the appropriate range of the radar sensor. Hand gestures should be performed neither too close to the radar (otherwise a more complicated near-field algorithm should be used to model the radar data and remove radar-antenna effects [71]) nor too far away (otherwise, the reflection may be too weak, reaching the noise level of the radar, or below the threshold used for the time window). Fig. 25 shows both correct and wrong examples of distances: both the hand and body are too close to the radar (a), the hand is far enough from the radar (within 15-75 cm or 6-30 in), but the body is not excluded by the time gating (b), both the hand and the body are within an appropriate range (c), and both the hand and the body are out of range (d). For example, the “barrier gesture” is recognized with 100% accuracy in Steps 2 and 4 (Fig. 23a and 23b), but with just 56% and 65% accuracy in Steps 7 and 8 (Figs. 23c and 23d), as most of the gesture motion is performed outside the time window, making it difficult to identify after step 6 (time gating).

13. Collect gestures from multiple users. Hands can vary widely in size and surface area between users. The length of the hand can range from 15 to 30 cm (5.9 to 11.8 in) and the circumference of the hand from 15 to 28 cm (5.9 to 11.02 in), giving an idea of the potential surface range. Hand and palm surface areas vary with gender, age, and morphology. For example, Göker and Gülhal Bozkir [48] reported that the average hand surface area for Turkish women and men is $M=158.34 \text{ cm}^2$ ($SD=14.76 \text{ cm}^2$) and $M=127.87 \text{ cm}^2$ ($SD=14.75 \text{ cm}^2$), respectively, and palm surface area $M=63.91 \text{ cm}^2$ ($M=11.84 \text{ cm}^2$) and $M=82.98 \text{ cm}^2$ ($SD=10.74 \text{ cm}^2$), respectively. Variations in the hand surface area result in variations in the radar signal that should be taken into account to ensure high recognition accuracy, especially for the user-independent scenario. Other parameters can equally vary between different users, such as arm length and position relative to the radar (e.g., elevation, distance), which may also result in signal variations. Gesture training sets for user-independent recognition should include samples produced by a variety of users.

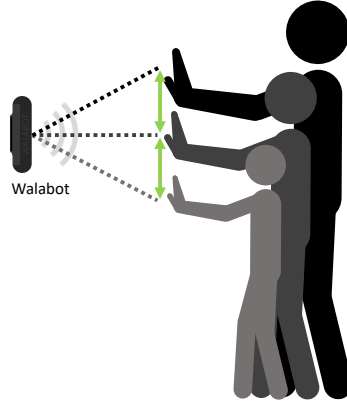


Fig. 26. Examples of variations in hand surface.

14. Favor gestures with motion parallel to the radar beam. As the angular resolution of a radar is significantly lower than its range resolution, it is easier to recognize gestures that are performed in a parallel plane with respect to the radar than in other configurations, such as orthogonal plane. As a negative example, the “Swipe left/right/up/down” gestures are more difficult to differentiate, especially with only one pair of antenna. As a positive example, “Push with palm/fist” will be easier to differentiate from other gestures (e.g., “Pull”), even with one antenna pair due to the relatively good range resolution of the radar.

15. Favor gestures with a highly differentiable surface of exposure. For example, in the user-dependent scenario, there was little or no confusion between the “push with palm” and “push with fist” gestures (Fig. 21) because the palm has a larger exposure surface than the fist, resulting in a larger amplitude of the received radar signal. This difference in amplitude can be used by a recognizer to differentiate between the two gestures. On the other hand, gestures such as “extend one/two/three/four fingers” were less accurately recognized, especially in the user-independent scenario (Fig. 22) because their surface of exposure was similar.

16. Favor the most accurately recognized gestures. Priority should be given to gestures that are the most accurately recognized. Positive examples for the Walabot are gestures no. 1 to 4, 7, 10, and 11 with more than 70% recognition accuracy rates.

17. Acquire a minimum of four templates per gesture type. Our evaluation in the user-dependent scenario showed that recognition accuracy rates improved when the number of training samples increased from 1 to 8, with $N=4$ achieving more than 90% accuracy rate in some configurations (Fig. 21a and 21b). The number of training samples should be chosen to balance gains in recognition accuracy and the cost of collecting more gesture samples. Keeping the number of training samples small is beneficial for end users that can define their own gestures with minimal effort. This aspect also represents a motivation in our approach to combine physical electromagnetic modeling with pattern matching algorithms that enables users to customize the gesture set without extensive training.

18. Use multiple antennas. Combining multiple pairs of antennas greatly improves the accuracy of recognition with the inversion data from Steps 7 and 8 (Sections 6.3.3 and 6.3.4). In a previous paper [136], inversion data resulted in poor recognition rates with a single high-end radar antenna (Fig. 12). However, the number of antennas seems to be a more relevant parameter than the quality of the antennas when using inversion.

I9. Identify the best combinations of antenna pairs. Our evaluation revealed that gesture recognition accuracy varies with the antenna pairs of the Walabot (Table 3 and 4). For instance, in the user-dependent scenario, the best performing antenna pairs were $AP = (4, 5, 7, 10, 11, 12)$ for step 2 (Section 6.3.1), $AP = (1, 2, 3, 6, 8, 9)$ for step 4 (Section 6.3.2), and $AP = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)$ for steps 7 (Section 6.3.3) and 8 (Section 6.3.4). It may not be necessary to use all transmitters/receivers of a radar system if the same or better accuracy can be achieved with only a subset.

I10. Select gestures based on the level of criticality of your application. For non-critical application (e.g., interacting with a TV at home) gestures with lower accuracy, such as “extend one/two/three/four fingers”, may be selected if they are well suited to perform a specific action. In such applications, an incorrect result from the recognizer would only impact the user experience of the application. In any case, using small gesture sets composed of training samples from the same person that uses the system are the key to provide high gesture recognition accuracy. The ability of our system to support very few training samples facilitates this, as it takes only a few minutes for a user to generate their own gesture set. For safety-critical applications, such as performing high-precision medical procedures, developers should rely on small gesture sets composed exclusively of gestures that can be recognized with very high accuracy (> 99%). It is strongly advised to train the system with data from the same person(s) that would use the system in order to guarantee high accuracy. In fact, an incorrectly recognized gesture in these applications could have life-threatening consequences.

7.3 Limitations and Future Work

Execution time. All our data pre-processing and testing was performed offline, as pipeline execution is currently too slow for applications that require response in interactive time. Future work will focus on optimizing the various steps of the pipeline (inversion in particular) and investigating techniques for gesture segmentation to identify candidate gestures from a continuous stream of radar data.

Data normalization. The low recognition accuracy obtained for user-independent scenarios highlights the need for more work on data normalization in our pipeline with respect to the user performing the gesture. Anatomical differences between users (e.g., hand size, arm length) and other differences, such as the position with respect to the radar, introduce variations that, when combined, reduce the effectiveness of the recognition algorithm. For example, a larger hand will reflect more signal back to the radar, resulting in a greater relative permittivity measured for the same gesture. As a consequence, the following changes are recommended for the pipeline to take into account such differences: (1) normalize the amplitude of user reflections with respect to their hand and body size, (2) extract the hand-body distance instead of the hand-radar distance in the inversion stage, and (3) normalize the hand-body distance with respect to the user’s arm length. In addition to improving the recognition accuracy for the user-independent scenario, the above improvements would enable the RadarSense pipeline to accommodate better small variations between repetitions of the same gesture performed by the same user.

Participants in our Study. Our study used a small number of users to train the recognizer and hand gestures were performed in a controlled environment: one user in front of the radar and very little clutter. Therefore, it must be confirmed whether our results can be transposed to a wider user population, including left- and right-handed people or people of different heights. Although we did not evaluate the impact of such factors on the recognition performance of the RadarSense pipeline, we expect that changes to the pipeline, e.g., new normalization steps, will be necessary.

Context of Use. The *context of use* [26, 36] is defined as the set of users engaged in tasks together with the available devices (the radar in our case) and the computing platforms (the laptop in our study) as well as the physical and social environments. Contextual differences can determine other types of variation that we did not consider in our evaluation. For example, when only one user performs the gestures, only their body can cause interference. If another user enters the active sensing area of the radar, *e.g.*, for the collaboration scenario, for example, an additional source of interference must be taken into account in the modeling and inversion steps.

Radar Types. Since the bandwidth of the US version of the Walabot device is larger than that we could use with the European model, we recommend replication of our experiment. We hypothesize that the higher bandwidth will result in raw signals of better quality and thus increased recognition accuracy rates. Furthermore, future work should investigate whether relying on a larger set of antenna pairs (*e.g.*, the 40 antenna pairs available in the PROF_SENSOR profile of the Walabot) or using multiple radar sensors could enable the accurate recognition of a wider range of gestures. The placement of radars at various locations [131] and the effect of the recognition accuracy of gestures sensed from those locations are another interesting direction for future work. The choice of the Walabot for our evaluation was mainly dictated by its availability on the market, its small weight, low price, and available API. When other radar sensors offer similar conditions, such as Vayyar’s single chip platform, they should be considered for a comparison evaluation.

Environments. Radars are prized for being operational under various environmental conditions with respect to lighting conditions and occlusion. However, it remains to be determined the influence of experimental conditions on the recognition accuracy of gestures sensed with radars. Examples include gestures performed in the dark (*e.g.*, under critical conditions), gestures performed behind a surface (*e.g.*, in an inaccessible room), or gestures resulting from a reflection (*e.g.*, in a corridor perpendicular to the one covered by the radar). For these conditions, gesture elicitation studies [155] would be useful to understand users’ preferences.

Gesture Recognition Algorithms. Our paper focused on the template matching approach for gesture recognition, which we showed to be accurate with just a few training samples. In particular, we used the Jackknife recognizer [146], that supports multiple data types, such as frequency-domain radar signal, sound, 3D position. For future work, it would be interesting to examine whether some steps of our pipeline, such as step 3 (removal of radar source and antenna effects) and 4 (background subtraction), could also improve the recognition performance of other techniques (*e.g.*, CNNs or LSTMs).

Gesture Set. The gestures in our dataset were selected based on their diverse shapes, sizes, and potential mapping to system functions. It would be interesting to conduct a gesture elicitation study [151, 155] to determine which gestures are preferred for specific tasks in specific applications. In addition, our gesture set consists of mid-air hand gestures exclusively, but other gesture types, including gestures performed with the feet and head, will be interesting to explore as well.

8 CONCLUSION

The use of microwave radar sensors for gesture recognition has several advantages compared to vision-based sensors, including less sensitivity to environmental conditions. However, because of the complexity of radar signals, most of the existing work on this topic has relied on complex machine learning approaches. As an alternative, we proposed a new processing pipeline for radar gesture recognition that filters the raw radar signals and reduces them to two physically meaningful features, the hand-radar distance and the effective permittivity of the hand. We observed that recognition

accuracy rates of up to 98% can be achieved using both raw and filtered frequency domain signals with user-dependent training. The inversion step of our pipeline considerably reduced the size of the gestures and the average execution time at the expense of accuracy. However, using more antenna pairs resulted in an increase of accuracy. We discussed limitations and proposed implications of our findings for future development of radar-based gesture recognition.

OPEN SCIENCE

Our website <https://sites.uclouvain.be/ingenious/2022/11/15/radarsense-accurate-recognition-of-mid-air-hand-gestures-with-radarsensing-and-few-training-examples/> provides the reader with useful resources, including SLR data (Section 3) the LMC and Walabot datasets (Section 5), and evaluation data (Section 6).

ACKNOWLEDGMENTS

The authors are grateful to anonymous IUI '22 reviewers for their constructive comments on a first version [136] of this work containing preliminary results. The authors also thank the study participants for their participation in the collection of the gesture dataset employed in this paper. This work was supported by Wallonie-Bruxelles International (WBI), Belgium under grant no. SUB/2021/519018 and UEFISCDI, Romania under grant no. PN-III-CEI-BIM-PBE-2020-0001/1BM/2021 (project “RadarSense”). Arthur Sluÿters is funded by the Fonds de la Recherche Scientifique (FNRS) under Grants n°40001931 and n°40011629.

REFERENCES

- [1] Gianluca Agresti and Simone Milani. 2019. Material Identification Using RF Sensors and Convolutional Neural Networks. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing* (Brighton, UK) (ICASSP '19). IEEE, Piscataway, NJ, USA, 3662–3666. <https://doi.org/10.1109/ICASSP.2019.8682296>
- [2] Misbah Ahmad, Milind Ghawale, Sakshi Dubey, Ayushi Gupta, and Poonam Sonar. 2021. GigaHertz: Gesture Sensing Using Microwave Radar and IR Sensor with Machine Learning Algorithms. In *Image Processing and Capsule Networks*, Joy Long-Zong Chen, João Manuel R. S. Tavares, Subarna Shakya, and Abdullah M. Ilyasu (Eds.). Vol. 1200. Springer International Publishing, Cham, 422–434. https://doi.org/10.1007/978-3-030-51859-2_39 Series Title: Advances in Intelligent Systems and Computing.
- [3] Shahzad Ahmed and Sung Ho Cho. 2020. Hand Gesture Recognition Using an IR-UWB Radar with an Inception Module-Based Classifier. *Sensors* 20, 2 (2020), 1–18. <https://doi.org/10.3390/s20020564>
- [4] Shahzad Ahmed, Faheem Khan, Asim Ghaffar, Farhan Hussain, and Sung Ho Cho. 2019. Finger-Counting-Based Gesture Recognition within Cars Using Impulse Radar with Convolutional Neural Network. *Sensors* 19, 6 (2019), 1–14. <https://doi.org/10.3390/s19061429>
- [5] Shahzad Ahmed, Dingyang Wang, Junyoung Park, and Sung Ho Cho. 2021. UWB-gestures, a public dataset of dynamic hand gestures acquired using impulse radar sensors. *Scientific Data* 8, 102 (April 2021), 1–9. <https://doi.org/10.1038/s41597-021-00876-0>
- [6] Ahmad Akl and Shahrokh Valaee. 2010. Accelerometer-based gesture recognition via dynamic-time warping, affinity propagation, & compressive sensing. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. 2270 – 2273. <https://doi.org/10.1109/ICASSP.2010.5495895>
- [7] Akram Al-Hourani, Robin J. Evans, Peter M. Farrell, Bill Moran, Marco Martorella, Sithamparanathan Kandeepan, Stan Skafidas, and Udaya Parampalli. 2018. Chapter 7 - Millimeter-wave integrated radar systems and techniques. In *Academic Press Library in Signal Processing, Volume 7*, Rama Chellappa and Sergios Theodoridis (Eds.). Academic Press, 317–363. <https://doi.org/10.1016/B978-0-12-811887-0.00007-9>
- [8] Mohammed Alloulah, Anton Isopoussu, and Fahim Kawsar. 2018. On Indoor Human Sensing Using Commodity Radar. In *Proceedings of the ACM International Joint Conference and International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore) (UbiComp '18). Association for Computing Machinery, New York, NY, USA, 1331–1336. <https://doi.org/10.1145/3267305.3274180>
- [9] Moeness G. Amin, Zhengxin Zeng, and Tao Shan. 2019. Hand Gesture Recognition based on Radar Micro-Doppler Signature Envelopes. In *Proceedings of the IEEE Int. Radar Conference* (Boston, MA, USA) (RADAR '19). IEEE, Piscataway, NJ, USA, 1–6. <https://doi.org/10.1109/RADAR.2019.8835661> ISSN: 2375-5318.
- [10] Moeness G. Amin, Zhengxin Zeng, and Tao Shan. 2020. Arm Motion Classification Using Curve Matching of Maximum Instantaneous Doppler Frequency Signatures. In *Proceedings of the IEEE International Radar Conference* (RADAR '20). 303–308. <https://doi.org/10.1109/RADAR42522.2020.9114779> ISSN: 2640-7736.
- [11] Moeness G. Amin, Zhengxin Zeng, Tao Shan, and R.G. Guendel. 2019. Automatic Arm Motion Recognition Using Radar for Smart Home Technologies. In *Proceedings of the IEEE International Radar Conference* (Boston, MA, USA) (RADAR '19). IEEE, Piscataway, NJ, USA, 1–4. <https://doi.org/10.1109/RADAR41533.2019.171318> ISSN: 2640-7736.

- [12] Dimitra Anastasiou and Eric Ras. 2020. Gestures in Tangible User Interfaces. *ERCIM News* 2020, 120 (2020), 52 pages. <https://ercim-news.ercim.eu/en120/special/gestures-in-tangible-user-interfaces>
- [13] Leonardo Angelini, Denis Lalanne, Elise Van den Hoven, Omar Abou Khaled, and Elena Mugellini. 2015. Move, Hold and Touch: A Framework for Tangible Gesture Interactive Systems. *Machines* 3, 3 (2015), 173–207. <https://doi.org/10.3390/machines3030173>
- [14] Caroline Appert and Shumin Zhai. 2009. Using Strokes as Command Shortcuts: Cognitive Benefits and Toolkit Support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) (*CHI '09*). Association for Computing Machinery, New York, NY, USA, 2289–2298. <https://doi.org/10.1145/1518701.1519052>
- [15] Ferran Argelaguet, Mélanie Ducoffe, Anatole Lécuyer, and Remi Gribonval. 2017. Spatial and rotation invariant 3D gesture recognition based on sparse representation. In *Proc. of IEEE Symposium on 3D User Interfaces* (Los Angeles, CA, USA) (*3DUI '17*). IEEE, Piscataway, NJ, USA, 158–167. <https://doi.org/10.1109/3DUI.2017.7893333>
- [16] M. Arsalan and A. Santra. 2019. Character Recognition in Air-Writing Based on Network of Radars for Human-Machine Interface. *IEEE Sensors Journal* 19, 19 (Oct. 2019), 8855–8864. <https://doi.org/10.1109/JSEN.2019.2922395>
- [17] Kongphum Arthamanolap, Somprasong Gabbualoy, and Pattarapong Phasukkit. 2019. Doppler Radar for Dynamic Hand Gesture Recognition based on Signal Image Processing. In *Proceedings of 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (Pattaya, Thailand) (*ECTI-CON'19*). IEEE, Piscataway, NJ, USA, 931–934. <https://doi.org/10.1109/ECTI-CON.2019.8955217>
- [18] Nuwan T. Attygalle, Luis A. Leiva, Matjaz Kljun, Christian Sandor, Alexander Plopski, Hirokazu Kato, and Klen Copic Pucihar. 2021. No Interface, No Problem: Gesture Recognition on Physical Objects Using Radar Sensing. *Sensors* 21, 17 (2021), 5771. <https://doi.org/10.3390/s21175771>
- [19] Daniel Avrahami, Mitesh Patel, Yusuke Yamaura, and Sven Kratz. 2018. Below the Surface: Unobtrusive Activity Recognition for Work Surfaces Using RF-Radar Sensing. In *23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) (*IUI '18*). Association for Computing Machinery, New York, NY, USA, 439–451. <https://doi.org/10.1145/3172944.3172962>
- [20] Daniel Avrahami, Mitesh Patel, Yusuke Yamaura, Sven Kratz, and Matthew Cooper. 2019. Unobtrusive Activity Recognition and Position Estimation for Work Surfaces Using RF-Radar Sensing. *ACM Trans. Interact. Intell. Syst.* 10, 1, Article 11 (Aug. 2019), 28 pages. <https://doi.org/10.1145/3241383>
- [21] Alan Bannon, Richard Capraru, and Matthew Ritchie. 2020. Exploring gesture recognition with low-cost CW radar modules in comparison to FMCW architectures. In *Proceedings of the IEEE International Radar Conference* (Washington, DC, USA) (*RADAR'20*). IEEE, Piscataway, NJ, USA, 744–748. <https://doi.org/10.1109/RADAR42522.2020.9114650> ISSN: 2640-7736.
- [22] Abel Díaz Berenguer, Meshia Cédric Oveneke, Habib-Ur-Rehman Khalid, Mitchel Alioscha-Perez, André Bourdoux, and Hichem Sahli. 2019. GestureVLAD: Combining Unsupervised Features Representation and Spatio-Temporal Aggregation for Doppler-Radar Gesture Recognition. *IEEE Access* 7 (2019), 137122–137135. <https://doi.org/10.1109/ACCESS.2019.2942305>
- [23] Alan Bole, Alan Wall, and Andy Norris. 2014. Chapter 2 - The Radar System – Technical Principles. In *Radar and ARPA Manual (Third Edition)* (third edition ed.), Alan Bole, Alan Wall, and Andy Norris (Eds.). Butterworth-Heinemann, Oxford, 29–137. <https://doi.org/10.1016/B978-0-08-097752-2.00002-7>
- [24] Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. 2011. Code Space: Touch + Air Gesture Hybrid Interactions for Supporting Developer Meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Kobe, Japan) (*ITS '11*). Association for Computing Machinery, New York, NY, USA, 212–221. <https://doi.org/10.1145/2076354.2076393>
- [25] Sanders Brandon. 2014. *Mastering Leap Motion*. Packt Publishing, Birmingham.
- [26] Gaëlle Calvary, Joëlle Coutaz, David Thevenin, Quentin Limbourg, Laurent Bouillon, and Jean Vanderdonck. 2003. A Unifying Reference Framework for multi-target user interfaces. *Interact. Comput.* 15, 3 (2003), 289–308. [https://doi.org/10.1016/S0953-5438\(03\)00010-9](https://doi.org/10.1016/S0953-5438(03)00010-9)
- [27] Zhaoxi Chen, Gang Li, Francesco Fioranelli, and Hugh Griffiths. 2019. Dynamic Hand Gesture Classification Based on Multistatic Radar Micro-Doppler Signatures Using Convolutional Neural Network. In *Proceedings of the IEEE International Radar Conference* (Boston, MA, USA) (*RADAR '19*). IEEE, Piscataway, NJ, USA, 1–5. <https://doi.org/10.1109/RADAR.2019.8835796> ISSN: 2375-5318.
- [28] Hong Cheng, Lu Yang, and Zicheng Liu. 2016. Survey on 3D Hand Gesture Recognition. *IEEE Transactions on Circuits and Systems for Video Technology* 26, 9 (2016), 1659–1673. <https://doi.org/10.1109/TCSVT.2015.2469551>
- [29] J. Choi, S. Ryu, and J. Kim. 2019. Short-Range Radar Based Real-Time Hand Gesture Recognition Using LSTM Encoder. *IEEE Access* 7 (2019), 33610–33618. <https://doi.org/10.1109/ACCESS.2019.2903586>
- [30] Klen Copić Pucihar, Nuwan T. Attygalle, Matjaz Kljun, Christian Sandor, and Luis A. Leiva. 2022. Solids on Soli: Millimetre-Wave Radar Sensing through Materials. *Proc. ACM Human-Computer Interaction* 6, EICS, Article 156 (jun 2022), 21 pages. <https://doi.org/10.1145/3532212>
- [31] Klen Copić Pucihar, Christian Sandor, Matjaz Kljun, Wolfgang Huerst, Alexander Plopski, Takafumi Taketomi, Hirokazu Kato, and Luis A. Leiva. 2019. The Missing Interface: Micro-Gestures on Augmented Objects. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (*CHI EA '19*). Association for Computing Machinery, New York, NY, USA, 1–6. <https://doi.org/10.1145/3290607.3312986>
- [32] Adrien Coyette, Sascha Schimke, Jean Vanderdonck, and Claus Vielhauer. 2007. Trainable Sketch Recognizer for Graphical User Interface Design. In *Proc. of 11th IFIP TC 13 International Conference on Human-Computer Interaction, INTERACT '07* (Rio de Janeiro, Brazil) (*Lecture Notes in Computer Science, Vol. 4662*), Maria Cecília Calani Baranauskas, Philippe A. Palanque, Julio Abascal, and Simone Diniz Junqueira Barbosa (Eds.). Springer, 124–135. https://doi.org/10.1007/978-3-540-74796-3_14
- [33] Albéric De Coster, Anh Phuong Tran, and Sébastien Lambot. 2016. Fundamental Analyses on Layered Media Reconstruction Using GPR and Full-Wave Inversion in Near-Field Conditions. *IEEE Transactions on Geoscience and Remote Sensing* 54, 9 (2016), 5143–5158. <https://doi.org/10.1109/>

TGRS.2016.2556862

- [34] Quentin De Smedt, Hazem Wannous, Jean-Philippe Vandeborre, J. Guerry, B. Le Saux, and D. Filliat. 2017. 3D Hand Gesture Recognition Using a Depth and Skeletal Dataset: SHREC'17 Track. In *Proceedings of the Workshop on 3D Object Retrieval* (Lyon, France) (3DOR '17). Eurographics Association, Goslar, DEU, 33–38. <https://doi.org/10.2312/3dor.20171049>
- [35] Bastiaan Dekker, Sebastiaan Jacobs, A. S. Kossen, Maarten Kruithof C., Albert G. Huizing, and M. Geurts. 2017. Gesture recognition with a low power FMCW radar and a deep convolutional neural network. In *Proceedings of the European Radar Conference (EURAD '17)*. 163–166. <https://doi.org/10.23919/EURAD.2017.8249172>
- [36] Anind K. Dey. 2001. Understanding and Using Context. *Pers. Ubiquitous Comput.* 5, 1 (2001), 4–7. <https://doi.org/10.1007/s007790170019>
- [37] C. Du, L. Zhang, X. Sun, J. Wang, and J. Sheng. 2020. Enhanced Multi-Channel Feature Synthesis for Hand Gesture Recognition Based on CNN With a Channel and Spatial Attention Mechanism. *IEEE Access* 8 (2020), 144610–144620. <https://doi.org/10.1109/ACCESS.2020.3010063>
- [38] C. Y. Du, X. H. Wang, Z. X. Yuan, and Y. Xu. 2019. Design of Gesture Recognition System Based on 77GHz Millimeter Wave Radar. In *2019 International Conference on Microwave and Millimeter Wave Technology (ICMMT)*. 1–3. <https://doi.org/10.1109/ICMMT45702.2019.8992849>
- [39] M. Eggimann, J. Erb, P. Mayer, M. Magno, and L. Benini. 2019. Low Power Embedded Gesture Recognition Using Novel Short-Range Radar Sensors. In *2019 IEEE SENSORS*. 1–4. <https://doi.org/10.1109/SENSORS43011.2019.8956617> ISSN: 2168-9229.
- [40] M. G. Ehrnsperger, T. Brenner, U. Siart, and T. F. Eibert. 2020. Real-Time Gesture Recognition with Shallow Convolutional Neural Networks Employing an Ultra Low Cost Radar System. In *2020 German Microwave Conference (GeMiC)*. 88–91. ISSN: 2167-8022.
- [41] M. G. Ehrnsperger, H. L. Hoese, U. Siart, and T. F. Eibert. 2019. Performance Investigation of Machine Learning Algorithms for Simple Human Gesture Recognition Employing an Ultra Low Cost Radar System. In *2019 Kleinheubach Conference*. 1–4.
- [42] X. Feng, Q. Song, Q. Guo, D. Liu, Z. Zhao, and Y. Zhao. 2019. Hand Gesture Recognition with Ensemble Time-Frequency Signatures Using Enhanced Deep Convolutional Neural Network. In *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 1602–1605. <https://doi.org/10.1109/APSIPAASC47483.2019.9023254> ISSN: 2640-0103.
- [43] L. O. Fhager, S. Heunisch, H. Dahlberg, A. Evertsson, and L. Wernersson. 2019. Pulsed Millimeter Wave Radar for Hand Gesture Sensing and Classification. *IEEE Sensors Letters* 3, 12 (Dec. 2019), 1–4. <https://doi.org/10.1109/LENS.2019.2953022>
- [44] Zak Flintoff, Bruno Johnston, and Minas Liarokapis. 2018. Single-Grasp, Model-Free Object Classification using a Hyper-Adaptive Hand, Google Soli, and Tactile Sensors. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 1943–1950. <https://doi.org/10.1109/IROS.2018.8594166>
- [45] A. Ghaffar, F. Khan, and S. H. Cho. 2019. Hand Pointing Gestures Based Digital Menu Board Implementation Using IR-UWB Transceivers. *IEEE Access* 7 (2019), 58148–58157. <https://doi.org/10.1109/ACCESS.2019.2914410>
- [46] Bogdan-Florin Gheran, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2018. Gestures for Smart Rings: Empirical Results, Insights, and Design Implications. In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018, DIS 2018, Hong Kong, China, June 09-13, 2018*, Ilpo Koskinen, Youn-Kyung Lim, Teresa Cerratto Pargman, Kenny K. N. Chow, and William Odom (Eds.). ACM, 623–635. <https://doi.org/10.1145/3196709.3196741>
- [47] Andrew Gigie, Smriti Rani, Arijit Chowdhury, Tapas Chakravarty, and Arpan Pal. 2019. An agile approach for human gesture detection using synthetic radar data. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 558–564. <https://doi.org/10.1145/3341162.3349332>
- [48] Pinar Göker and Memduha Gülhal Bozkir. 2017. Determination of hand and palm surface areas as a percentage of body surface area in Turkish young adults. *Trauma and Emergency Care* 2, 4 (May 2017), 1–4. <https://doi.org/10.15761/TEC.1000135>
- [49] P. Goswami, S. Rao, S. Bharadwaj, and A. Nguyen. 2019. Real-Time Multi-Gesture Recognition using 77 GHz FMCW MIMO Single Chip Radar. In *2019 IEEE International Conference on Consumer Electronics (ICCE)*. 1–4. <https://doi.org/10.1109/ICCE.2019.8662006> ISSN: 2158-4001.
- [50] Changzhan Gu, Jian Wang, and Jaime Lien. 2019. Motion Sensing Using Radar: Gesture Interaction and Beyond. *IEEE Microwave Magazine* 20, 8 (aug 2019), 44–57. <https://doi.org/10.1109/MMM.2019.2915490>
- [51] S.Z. Gurbuz, A. C. Gurbuz, E. A. Malaia, D. J. Griffin, C. Crawford, E. Kurtoglu, M. M. Rahman, R. Aksu, and R. Mdrafi. 2020. ASL Recognition Based on Kinematics Derived from a Multi-Frequency RF Sensor Network. In *2020 IEEE Sensors*. 1–4. <https://doi.org/10.1109/SENSORS47125.2020.9278864> ISSN: 2168-9229.
- [52] Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 5, 14 pages. <https://doi.org/10.1145/3411764.3445367>
- [53] S. Hazra and A. Santra. 2018. Robust Gesture Recognition Using Millimetric-Wave Radar System. *IEEE Sensors Letters* 2, 4 (Dec. 2018), 1–4. <https://doi.org/10.1109/LENS.2018.2882642>
- [54] S. Hazra and A. Santra. 2019. Radar Gesture Recognition System in Presence of Interference using Self-Attention Neural Network. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*. 1409–1414. <https://doi.org/10.1109/ICMLA.2019.00230>
- [55] S. Hazra and A. Santra. 2019. Short-Range Radar-Based Gesture Recognition System Using 3D CNN With Triplet Loss. *IEEE Access* 7 (2019), 125623–125633. <https://doi.org/10.1109/ACCESS.2019.2938725>
- [56] M. Hoffman, P. Varcholik, and J. J. LaViola. 2010. Breaking the status quo: Improving 3D gesture recognition with spatially convenient input devices. In *2010 IEEE Virtual Reality Conference (VR)*. 59–66.

- [57] Jinmiao Huang, Prakhar Jaiswal, and Rahul Rai. 2019. Gesture-based system for next generation natural and intuitive interfaces. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 33, 1 (2019), 54–68. <https://doi.org/10.1017/S0890060418000045>
- [58] Cloe Huesser, Simon Schubiger, and Arzu Çöltekin. 2021. Gesture Interaction in Virtual Reality. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer International Publishing, Cham, 151–160.
- [59] K. Ishak, N. Appenrodt, J. Dickmann, and C. Waldschmidt. 2018. Human Motion Training Data Generation for Radar Based Deep Learning Applications. In *2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*. 1–4. <https://doi.org/10.1109/ICMIM.2018.8443559>
- [60] Z. Kang, L. Shengchang, and Z. Guiyuan. 2019. Mining Spatio-Temporal Features from mmW Radar echoes for Hand Gesture Recognition. In *2019 IEEE Asia-Pacific Microwave Conference (APMC)*. 93–95. <https://doi.org/10.1109/APMC46564.2019.9038713>
- [61] N. Kern, M. Steiner, R. Lorenzin, and C. Waldschmidt. 2020. Robust Doppler-Based Gesture Recognition With Incoherent Automotive Radar Sensor Networks. *IEEE Sensors Letters* 4, 11 (Nov. 2020), 1–4. <https://doi.org/10.1109/LENS.2020.3033586>
- [62] Faheem Khan, Seong Kyu Leem, and Sung Ho Cho. 2017. Hand-Based Gesture Recognition for Vehicular Applications Using IR-UWB Radar. *Sensors* 17, 4 (2017), 1–18. <https://doi.org/10.3390/s17040833>
- [63] F. Khan, S. K. Leem, and S. H. Cho. 2020. In-Air Continuous Writing Using UWB Impulse Radar Sensors. *IEEE Access* 8 (2020), 99302–99311. <https://doi.org/10.1109/ACCESS.2020.2994281>
- [64] Rami N. Khushaba and Andrew John Hill. 2022. Radar-Based Materials Classification Using Deep Wavelet Scattering Transform: A Comparison of Centimeter vs. Millimeter Wave Units. *IEEE Robotics Autom. Lett.* 7, 2 (2022), 2016–2022. <https://doi.org/10.1109/LRA.2022.3143200>
- [65] S. Y. Kim, H. G. Han, J. W. Kim, S. Lee, and T. W. Kim. 2017. A Hand Gesture Recognition Sensor Using Reflected Impulses. *IEEE Sensors Journal* 17, 10 (May 2017), 2975–2976. <https://doi.org/10.1109/JSEN.2017.2679220>
- [66] Y. Kim and B. Toomajian. 2016. Hand Gesture Recognition Using Micro-Doppler Signatures With Convolutional Neural Network. *IEEE Access* 4 (2016), 7125–7130. <https://doi.org/10.1109/ACCESS.2016.2617282>
- [67] Y. Kim and B. Toomajian. 2017. Application of Doppler radar for the recognition of hand gestures using optimized deep convolutional neural networks. In *2017 11th European Conference on Antennas and Propagation (EuCAP)*. 1258–1260. <https://doi.org/10.23919/EuCAP.2017.7928465>
- [68] Barbara Kitchenham, Riallette Pretorius, David Budgen, O. Pearl Brereton, Mark Turner, Mahmood Niazi, and Stephen Linkman. 2010. Systematic literature reviews in software engineering – A tertiary study. *Information and Software Technology* 52, 8 (2010), 792–805. <https://doi.org/10.1016/j.infsof.2010.03.006>
- [69] D. Kohlsdorf, T. Starner, and D. Ashbrook. 2011. MAGIC 2.0: A web tool for false positive prediction and prevention for gesture recognition systems. In *Face and Gesture 2011*. 1–6.
- [70] H. Kulhandjian, P. Sharma, M. Kulhandjian, and C. D’Amours. 2019. Sign Language Gesture Recognition Using Doppler Radar and Deep Learning. In *2019 IEEE Globecom Workshops (GC Wkshps)*. 1–6. <https://doi.org/10.1109/GCWkshps45667.2019.9024607>
- [71] Sébastien Lambot and Frédéric André. 2014. Full-Wave Modeling of Near-Field Radar Data for Planar Layered Media Reconstruction. *IEEE Transactions on Geoscience and Remote Sensing* 52, 5 (2014), 2295–2303. <https://doi.org/10.1109/TGRS.2013.2259243>
- [72] Sébastien Lambot, E.C. Slob, Idesbald van den Bosch, Benoit Stockbroeckx, and Marnik Vanclooster. 2004. Modeling of ground-penetrating Radar for accurate characterization of subsurface electric properties. *IEEE Transactions on Geoscience and Remote Sensing* 42, 11 (2004), 2555–2568. <https://doi.org/10.1109/TGRS.2004.834800>
- [73] S. Lan, Z. He, W. Chen, and L. Chen. 2018. Hand Gesture Recognition Using Convolutional Neural Networks. In *2018 USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)*. 147–148. <https://doi.org/10.1109/USNC-URSI.2018.8602809>
- [74] S. Lan, Z. He, H. Tang, K. Yao, and W. Yuan. 2017. A hand gesture recognition system based on 24GHz radars. In *2017 International Symposium on Antennas and Propagation (ISAP)*. 1–2. <https://doi.org/10.1109/ISANP.2017.8228827>
- [75] S. Lan, Z. He, K. Yao, and W. Chen. 2018. Hand Gesture Recognition using a Three-dimensional 24 GHz Radar Array. In *2018 IEEE/MTT-S International Microwave Symposium - IMS*. 138–140. <https://doi.org/10.1109/MWSYM.2018.8439658> ISSN: 2576-7216.
- [76] Gierad Laput and Chris Harrison. 2019. *Sensing Fine-Grained Hand Activity with Smartwatches*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300568>
- [77] Gierad Laput, Robert Xiao, Xiang ‘Anthony’ Chen, Scott E. Hudson, and Chris Harrison. 2014. Skin buttons: cheap, small, low-powered and clickable fixed-icon laser projectors. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, HI, USA) (*UIST ’14*), Hrvoje Benko, Mira Dontcheva, and Daniel Wigdor (Eds.). ACM, 389–394. <https://doi.org/10.1145/2642918.2647356>
- [78] Joseph J. LaViola. 2013. 3D Gestural Interaction: The State of the Field. *International Scholarly Research Notices* 2013, Article 514641 (2013), 18 pages. <https://doi.org/10.1155/2013/514641>
- [79] Hyo Ryun Lee, Jihun Park, and Young-Joo Suh. 2020. Improving Classification Accuracy of Hand Gesture Recognition Based on 60 GHz FMCW Radar with Deep Learning Domain Adaptation. *Electronics* 9, 12 (2020), 1–24. <https://doi.org/10.3390/electronics9122140>
- [80] Joo-Young Lee, Jeong-Wha Choi, and Ho Kim. 2007. Determination of hand surface area by sex and body shape using alginate. *Journal of physiological anthropology* 26, 4 (2007), 475–483. <https://doi.org/10.2114/jpa2.26.475>
- [81] S. K. Leem, F. Khan, and S. H. Cho. 2020. Detecting Mid-Air Gestures for Digit Writing With Radio Sensors and a CNN. *IEEE Transactions on Instrumentation and Measurement* 69, 4 (April 2020), 1066–1081. <https://doi.org/10.1109/TIM.2019.2909249>

- [82] S. K. Leem, F. Khan, and S. H. Cho. 2020. Remote Authentication Using an Ultra-Wideband Radio Frequency Transceiver. In *2020 IEEE 17th Annual Consumer Communications Networking Conference (CCNC)*. 1–8. <https://doi.org/10.1109/CCNC46108.2020.9045438> ISSN: 2331-9860.
- [83] Feifei Li, Yujun Li, Baozhen Du, Hongji Xu, Hailiang Xiong, and Min Chen. 2019. A Gesture Interaction System Based on Improved Finger Feature and WE-KNN. In *Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence (Chegndu, China) (ICMAI 2019)*. Association for Computing Machinery, New York, NY, USA, 39–43. <https://doi.org/10.1145/3325730.3325759>
- [84] G. Li, R. Zhang, M. Ritchie, and H. Griffiths. 2017. Sparsity-based dynamic hand gesture recognition using micro-Doppler signatures. In *2017 IEEE Radar Conference (RadarConf)*. 0928–0931. <https://doi.org/10.1109/RADAR.2017.7944336> ISSN: 2375-5318.
- [85] G. Li, R. Zhang, M. Ritchie, and H. Griffiths. 2018. Sparsity-Driven Micro-Doppler Feature Extraction for Dynamic Hand Gesture Recognition. *IEEE Trans. Aerospace Electron. Systems* 54, 2 (April 2018), 655–665. <https://doi.org/10.1109/TAES.2017.2761229>
- [86] G. Li, S. Zhang, F. Fioranelli, and H. Griffiths. 2018. Effect of sparsity-aware time–frequency analysis on dynamic hand gesture classification with radar micro-Doppler signatures. *IET Radar, Sonar Navigation* 12, 8 (2018), 815–820. <https://doi.org/10.1049/iet-rsn.2017.0570>
- [87] H. Li, X. Liang, A. Shrestha, Y. Liu, H. Heidari, J. Le Kerneec, and F. Fioranelli. 2020. Hierarchical Sensor Fusion for Micro-Gesture Recognition With Pressure Sensor Array and Radar. *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology* 4, 3 (Sept. 2020), 225–232. <https://doi.org/10.1109/JERM.2019.2949456>
- [88] Hao-Yang Li, Han-Ting Zhao, Meng-Lin Wei, Heng-Xin Ruan, Ya Shuang, Tie Jun Cui, Philipp del Hougne, and Lianlin Li. 2020. Intelligent Electromagnetic Sensing with Learnable Data Acquisition and Processing. *Patterns* 1, 1 (2020), 100006. <https://doi.org/10.1016/j.patter.2020.100006>
- [89] Tuanjie Li and Mengmeng Ge. 2009. Human motion recognition using ultra-wideband radar and cameras on mobile robot. *Transactions of Tianjin University* 15, 5 (Oct. 2009), 381–387. <https://doi.org/10.1007/s12209-009-0067-5>
- [90] Tianxing Li, Xi Xiong, Yifei Xie, George Hito, Xing-Dong Yang, and Xia Zhou. 2017. Reconstructing Hand Poses Using Visible Light. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 71 (sep 2017), 20 pages. <https://doi.org/10.1145/3130937>
- [91] H. Liang, X. Wang, M. S. Greco, and F. Gini. 2020. Enhanced Hand Gesture Recognition using Continuous Wave Interferometric Radar. In *2020 IEEE International Radar Conference (RADAR)*. 226–231. <https://doi.org/10.1109/RADAR42522.2020.9114807> ISSN: 2640-7736.
- [92] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihhood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics* 35, 4 (July 2016), 142:1–142:19. <https://doi.org/10.1145/2897824.2925953>
- [93] C. Liu, Y. Li, D. Ao, and H. Tian. 2019. Spectrum-Based Hand Gesture Recognition Using Millimeter-Wave Radar Parameter Measurements. *IEEE Access* 7 (2019), 79147–79158. <https://doi.org/10.1109/ACCESS.2019.2923122>
- [94] Haipeng Liu, Yuheng Wang, Anfu Zhou, Hanyue He, Wei Wang, Kunpeng Wang, Peilin Pan, Yixuan Lu, Liang Liu, and Huadong Ma. 2020. Real-time Arm Gesture Recognition in Smart Home Scenarios via Millimeter Wave Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (Dec. 2020), 140:1–140:28. <https://doi.org/10.1145/3432235>
- [95] Jiayang Liu, Lin Zhong, Jehan Wickramasuriya, and Venu Vasudevan. 2009. UWave: Accelerometer-Based Personalized Gesture Recognition and Its Applications. *Pervasive Mob. Comput.* 5, 6 (Dec. 2009), 657–675. <https://doi.org/10.1016/j.pmcj.2009.07.007>
- [96] Yu Liu, Yuheng Wang, Haipeng Liu, Anfu Zhou, Jianhua Liu, and Ning Yang. 2020. Long-Range Gesture Recognition Using Millimeter Wave Radar. In *Green, Pervasive, and Cloud Computing*, Zhiwen Yu, Christian Becker, and Guoliang Xing (Eds.). Vol. 12398. Springer International Publishing, Cham, 30–44. https://doi.org/10.1007/978-3-030-64243-3_3 Series Title: Lecture Notes in Computer Science.
- [97] Olga Lopera, Evert C. Slob, Nada Milisavljevic, and Sébastien Lambot. 2007. Filtering Soil Surface and Antenna Effects From GPR Data to Enhance Landmine Detection. *IEEE Transactions on Geoscience and Remote Sensing* 45, 3 (mar 2007), 707–717. <https://doi.org/10.1109/TGRS.2006.888136>
- [98] X. Lou, Z. Yu, Z. Wang, K. Zhang, and B. Guo. 2018. Gesture-Radar: Enabling Natural Human-Computer Interactions with Radar-Based Adaptive and Robust Arm Gesture Recognition. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 4291–4297. <https://doi.org/10.1109/SMC.2018.00726> ISSN: 2577-1655.
- [99] Mehran Maghoumi and Joseph J. LaViola. 2019. DeepGRU: Deep Gesture Recognition Utility. In *Advances in Visual Computing*, George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Daniela Ushizima, Sek Chai, Shinjiro Sueda, Xin Lin, Aidong Lu, Daniel Thalmann, Chaoli Wang, and Panpan Xu (Eds.). Springer International Publishing, Cham, 16–31.
- [100] Nathan Magrofuoco, Paolo Roselli, and Jean Vanderdonckt. 2022. Two-dimensional Stroke Gesture Recognition: A Survey. *Comput. Surveys* 54, 7 (2022), 155:1–155:36. <https://doi.org/10.1145/3465400>
- [101] G. Malysa, D. Wang, L. Netsch, and M. Ali. 2016. Hidden Markov model-based gesture recognition with FMCW radar. In *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. 1017–1021. <https://doi.org/10.1109/GlobalSIP.2016.7905995>
- [102] Jess McIntosh, Mike Fraser, Paul Worgan, and Asier Marzo. 2017. DeskWave: Desktop Interactions using Low-cost Microwave Doppler Arrays. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 1885–1892. <https://doi.org/10.1145/3027063.3053152>
- [103] Antigoni Mezari and Ilias Maglogiannis. 2017. Gesture Recognition Using Symbolic Aggregate Approximation and Dynamic Time Warping on Motion Data. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare (Barcelona, Spain) (PervasiveHealth '17)*. Association for Computing Machinery, New York, NY, USA, 342–347. <https://doi.org/10.1145/3154862.3154927>
- [104] Elishiah Miller, Zheng Li, Helena Mentis, Adrian Park, Ting Zhu, and Nilanjan Banerjee. 2020. RadSense: Enabling one hand and no hands interaction for sterile manipulation of medical images using Doppler radar. *Smart Health* 15 (2020), 100089. <https://doi.org/10.1016/j.smhl.2019.100089>

- [105] Julien Minet, Sébastien Lambot, Evert C. Slob, and Marnik Vanclooster. 2010. Soil Surface Water Content Estimation by Full-Waveform GPR Signal Inversion in the Presence of Thin Layers. *IEEE Transactions on Geoscience and Remote Sensing* 48, 3 (2010), 1138–1150. <https://doi.org/10.1109/TGRS.2009.2031907>
- [106] P. Molchanov, S. Gupta, K. Kim, and K. Pulli. 2015. Multi-sensor system for driver’s hand-gesture recognition. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Vol. 1. 1–8. <https://doi.org/10.1109/FG.2015.7163132>
- [107] M. Q. Nguyen, A. Flores-Nigaglioni, and C. Li. 2018. Range-gating technology for millimeter-wave radar remote gesture control in IoT applications. In *2018 IEEE MTT-S International Wireless Symposium (IWS)*. 1–4. <https://doi.org/10.1109/IEEE-IWS.2018.8400811>
- [108] Prathyusha Nimmagadda and Hemalatha Inteti. 2019. Signal Analysis based Radar Imaging Using Radar Sensor. *Pramana Research Journal* 9, 6 (July 2019), 1919–1924. <https://doi.org/10.10089.PRJ.2019.V9I6.19.3724>
- [109] J. Oberhammer, N. Somjit, U. Shah, and Z. Baghchehsaraei. 2013. 16 - RF MEMS for automotive radar. In *Handbook of Mems for Wireless and Mobile Applications*, Deepak Uttamchandani (Ed.). Woodhead Publishing, 518–549. <https://doi.org/10.1533/9780857098610.2.518>
- [110] Institute of Electrical and Electronics Engineers. 2020. IEEE Standard Letter Designations for Radar-Frequency Bands. *IEEE Std 521-2019 (Revision of IEEE Std 521-2002)* (2020), 1–15. <https://doi.org/10.1109/IEEESTD.2020.8999849>
- [111] Matthew J Page, David Moher, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, Roger Chou, Julie Glanville, Jeremy M Grimshaw, Asbjørn Hróbjartsson, Manoj M Lalu, Tianjing Li, Elizabeth W Loder, Evan Mayo-Wilson, Steve McDonald, Luke A McGuinness, Lesley A Stewart, James Thomas, Andrea C Tricco, Vivian A Welch, Penny Whiting, and Joanne E McKenzie. 2021. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 372 (2021), 1–36. <https://doi.org/10.1136/bmj.n160> arXiv:<https://www.bmj.com/content/372/bmj.n160.full.pdf>
- [112] Sameera Palipana, Dariush Salami, Luis A. Leiva, and Stephan Sigg. 2021. Pantomime: Mid-Air Gesture Recognition with Sparse Millimeter-Wave Radar Point Clouds. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (March 2021), 27:1–27:27. <https://doi.org/10.1145/3448110>
- [113] Joseph Paradiso, Craig Abler, Kai-yuh Hsiao, and Matthew Reynolds. 1997. The Magic Carpet: Physical Sensing for Immersive Environments. In *CHI ’97 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia) (CHI EA ’97). Association for Computing Machinery, New York, NY, USA, 277–278. <https://doi.org/10.1145/1120212.1120391>
- [114] J. Park and S. H. Cho. 2016. IR-UWB Radar Sensor for Human Gesture Recognition by Using Machine Learning. In *2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*. 1246–1249. <https://doi.org/10.1109/HPCC-SmartCity-DSS.2016.0176>
- [115] J. Park, J. Jang, G. Lee, H. Koh, C. Kim, and T. W. Kim. 2020. A Time Domain Artificial Intelligence Radar System Using 33-GHz Direct Sampling for Hand Gesture Recognition. *IEEE Journal of Solid-State Circuits* 55, 4 (April 2020), 879–888. <https://doi.org/10.1109/JSSC.2020.2967547>
- [116] Avishek Patra, Philipp Geuer, Andrea Munari, and Petri Mähönen. 2018. mm-Wave Radar Based Gesture Recognition: Development and Evaluation of a Low-Power, Low-Complexity System. In *Proceedings of the 2nd ACM Workshop on Millimeter Wave Networks and Sensing Systems (mmNets ’18)*. Association for Computing Machinery, New York, NY, USA, 51–56. <https://doi.org/10.1145/3264492.3264501>
- [117] Claudio Patriarca, Sébastien Lambot, M.R. Mahmoudzadeh, Julien Minet, and Evert Slob. 2011. Reconstruction of sub-wavelength fractures and physical properties of masonry media using full-waveform inversion of proximal penetrating radar. *Journal of Applied Geophysics* 74, 1 (2011), 26–37. <https://doi.org/10.1016/j.jappgeo.2011.03.001>
- [118] Jorge Luis Pérez Medina, Jean Vanderdonckt, Albéric De Coster, and Sébastien Lambot. 2019. Mobile Direct Visualization of Pipes, Apexes, and Water Leaks in Water Distribution Networks. In *2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG)*. 172–179. <https://doi.org/10.1109/ICEDEG.2019.8734376>
- [119] A. Ren, Y. Wang, X. Yang, and M. Zhou. 2020. A Dynamic Continuous Hand Gesture Detection and Recognition Method with FMCW Radar. In *2020 IEEE/CIC International Conference on Communications in China (ICCC)*. 1208–1213. <https://doi.org/10.1109/ICCC49849.2020.9238935> ISSN: 2377-8644.
- [120] Yuwei Ren, Jiuyuan Lu, Andrian Beletchi, Yin Huang, Ilia Karmanov, Daniel Fontijne, Chirag Patel, and Hao Xu. 2021. Hand gesture recognition using 802.11ad mmWave sensor in the mobile device. In *Proc. of IEEE Wireless Communications and Networking Conference Workshops (WCNCW ’21)*. 1–6. <https://doi.org/10.1109/WCNCW49093.2021.9419978>
- [121] M. Ritchie, R. Capraru, and F. Fioranelli. 2020. Dop-NET: a micro-Doppler radar data challenge. *Electronics Letters* 56, 11 (2020), 568–570. <https://doi.org/10.1049/el.2019.4153>
- [122] M. Ritchie, A. Jones, J. Brown, and H. D. Griffiths. 2017. Hand gesture classification using 24 GHz FMCW dual polarised radar. In *International Conference on Radar Systems (Radar 2017)*. 1–6. <https://doi.org/10.1049/cp.2017.0482>
- [123] J. Rożman, H. Hagrass, J. A. Perez, D. Clarke, B. Müller, and S. F. Data. 2020. Privacy-Preserving Gesture Recognition with Explainable Type-2 Fuzzy Logic Based Systems. In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. 1–8. <https://doi.org/10.1109/FUZZ48607.2020.9177768> ISSN: 1558-4739.
- [124] S. Ryu, J. Suh, S. Baek, S. Hong, and J. Kim. 2018. Feature-Based Hand Gesture Recognition Using an FMCW Radar and its Temporal Feature Analysis. *IEEE Sensors Journal* 18, 18 (Sept. 2018), 7593–7602. <https://doi.org/10.1109/JSEN.2018.2859815>
- [125] T. Sakamoto, X. Gao, E. Yavari, A. Rahman, O. Boric-Lubecke, and V. M. Lubecke. 2017. Radar-based hand gesture recognition using I-Q echo plot and convolutional neural network. In *2017 IEEE Conference on Antenna Measurements Applications (CAMA)*. 393–395. <https://doi.org/10.1109/CAMA.2017.8273461>

- [126] T. Sakamoto, X. Gao, E. Yavari, A. Rahman, O. Boric-Lubecke, and V. M. Lubecke. 2018. Hand Gesture Recognition Using a Radar Echo I-Q Plot and a Convolutional Neural Network. *IEEE Sensors Letters* 2, 3 (Sept. 2018), 1–4. <https://doi.org/10.1109/LSSENS.2018.2866371>
- [127] Ugo Braga Sangiorgi, François Beuven, and Jean Vanderdonck. 2012. User interface design by collaborative sketching. In *Proc. of ACM Int. Conf. on Designing Interactive Systems* (Newcastle Upon Tyne, United Kingdom) (*DIS '12*). ACM, 378–387. <https://doi.org/10.1145/2317956.2318013>
- [128] P. S. Santhalingam, Y. Du, R. Wilkerson, A. A. Hosain, D. Zhang, P. Pathak, H. Rangwala, and R. Kushalnagar. 2020. Expressive ASL Recognition using Millimeter-wave Wireless Signals. In *2020 17th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. 1–9. <https://doi.org/10.1109/SECON48991.2020.9158441> ISSN: 2155-5494.
- [129] Panneer Selvam Santhalingam, Al Amin Hosain, Ding Zhang, Parth Pathak, Huzefa Rangwala, and Raja Kushalnagar. 2020. mmASL: Environment-Independent ASL Gesture Recognition Using 60 GHz Millimeter-wave Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (March 2020), 26:1–26:30. <https://doi.org/10.1145/3381010>
- [130] Ben Shneiderman. 1992. Tree Visualization with Tree-Maps: 2-d Space-Filling Approach. *ACM Trans. Graph.* 11, 1 (jan 1992), 92–99. <https://doi.org/10.1145/102377.115768>
- [131] Alexandru-Ionut Sian, Cristian Pamparau, and Radu-Daniel Vatavu. 2022. Scenario-based Exploration of Integrating Radar Sensing into Everyday Objects for Free-Hand Television Control. In *Proceedings of the ACM International Conference on Interactive Media Experiences (IMX '22)*. ACM, New York, NY, USA, 357–362. <https://doi.org/10.1145/3505284.3532982>
- [132] S. Skaria, A. Al-Hourani, and R. J. Evans. 2020. Deep-Learning Methods for Hand-Gesture Recognition Using Ultra-Wideband Radar. *IEEE Access* 8 (2020), 203580–203590. <https://doi.org/10.1109/ACCESS.2020.3037062>
- [133] S. Skaria, A. Al-Hourani, M. Lech, and R. J. Evans. 2019. Hand-Gesture Recognition Using Two-Antenna Doppler Radar With Deep Convolutional Neural Networks. *IEEE Sensors Journal* 19, 8 (April 2019), 3041–3048. <https://doi.org/10.1109/JSEN.2019.2892073>
- [134] S. Skaria, D. Huang, A. Al-Hourani, R. J. Evans, and M. Lech. 2020. Deep-Learning for Hand-Gesture Recognition with Simultaneous Thermal and Radar Sensors. In *2020 IEEE Sensors*. 1–4. <https://doi.org/10.1109/SENSOR47125.2020.9278683> ISSN: 2168-9229.
- [135] Evert Slob, Motoyuki Sato, and Gary Olhoeft. 2010. Surface and borehole ground-penetrating-radar developments. *Geophysics* 75, 5 (2010), 75A103–75A120. <https://doi.org/10.1190/1.3480619> arXiv:https://doi.org/10.1190/1.3480619
- [136] Arthur Sluÿters, Sébastien Lambot, and Jean Vanderdonck. 2022. Hand Gesture Recognition for an Off-the-Shelf Radar by Electromagnetic Modeling and Inversion. In *Proc. of 27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) (*IUI '22*), Giulio Jacucci, Samuel Kaski, Cristina Conati, Simone Stumpf, Tuukka Ruotsalo, and Krzysztof Gajos (Eds.). ACM, 506–522. <https://doi.org/10.1145/3490099.3511107>
- [137] Arthur Sluÿters, Mehdi Ousmer, Paolo Roselli, and Jean Vanderdonck. 2022. QuantumLeap, a Framework for Engineering Gestural User Interfaces based on the Leap Motion Controller. *Proc. ACM Human-Computer Interaction* 6, EICS, Article 161 (jun 2022), 47 pages. <https://doi.org/10.1145/3532211>
- [138] K. A. Smith, C. Csech, D. Murdoch, and G. Shaker. 2018. Gesture Recognition Using mm-Wave Sensor for Human-Car Interface. *IEEE Sensors Letters* 2, 2 (June 2018), 1–4. <https://doi.org/10.1109/LSSENS.2018.2810093>
- [139] W. Stern. 1929. Versuch einer elektrodynamischen Dickenmessung von Gletschereis. *Gerlands Beitrage zur Geophysik* 27 (1929), 292–333. <https://cir.nii.ac.jp/crid/1570572700956541952>
- [140] J. S. Suh, S. Ryu, B. Han, J. Choi, J. Kim, and S. Hong. 2018. 24 GHz FMCW Radar System for Real-Time Hand Gesture Recognition Using LSTM. In *2018 Asia-Pacific Microwave Conference (APMC)*. 860–862. <https://doi.org/10.23919/APMC.2018.8617375>
- [141] Y. Sun, T. Fei, S. Gao, and N. Pohl. 2019. Automatic Radar-based Gesture Detection and Classification via a Region-based Deep Convolutional Neural Network. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 4300–4304. <https://doi.org/10.1109/ICASSP.2019.8682277> ISSN: 2379-190X.
- [142] Y. Sun, T. Fei, X. Li, A. Warnecke, E. Warsitz, and N. Pohl. 2020. Multi-Feature Encoder for Radar-Based Gesture Recognition. In *2020 IEEE International Radar Conference (RADAR)*. 351–356. <https://doi.org/10.1109/RADAR42522.2020.9114664> ISSN: 2640-7736.
- [143] Y. Sun, T. Fei, X. Li, A. Warnecke, E. Warsitz, and N. Pohl. 2020. Real-Time Radar-Based Gesture Detection and Recognition Built in an Edge-Computing Platform. *IEEE Sensors Journal* 20, 18 (Sept. 2020), 10706–10716. <https://doi.org/10.1109/JSEN.2020.2994292>
- [144] Y. Sun, T. Fei, F. Schliep, and N. Pohl. 2018. Gesture Classification with Handcrafted Micro-Doppler Features using a FMCW Radar. In *2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*. 1–4. <https://doi.org/10.1109/ICMIM.2018.8443507>
- [145] Adrian Tang, Robert Carey, Gabriel Virbila, Yan Zhang, Rulin Huang, and Mau-Chung Frank Chang. 2020. A Delay-Correlating Direct-Sequence Spread-Spectrum (DS/SS) Radar System-on-Chip Operating at 183–205 GHz in 28 nm CMOS. *IEEE Transactions on Terahertz Science and Technology* 10, 2 (2020), 212–220. <https://doi.org/10.1109/TTHZ.2020.2969105>
- [146] Eugene M. Taranta II, Amirreza Samiei, Mehran Maghoumi, Pooya Khaloo, Corey R. Pittman, and Joseph J. LaViola Jr. 2017. Jackknife: A Reliable Recognizer with Few Samples and Many Modalities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). ACM, New York, NY, USA, 5850–5861. <https://doi.org/10.1145/3025453.3026002>
- [147] A. Tzadok, A. Valdes-Garcia, P. Pepeljugoski, J.-O. Plouchart, M. Yeck, and H. Liu. 2020. AI-driven Event Recognition with a Real-Time 3D 60-GHz Radar System. In *2020 IEEE/MTT-S International Microwave Symposium (IMS)*. 795–798. <https://doi.org/10.1109/IMS30576.2020.9224112> ISSN: 2576-7216.
- [148] Laurens van der Maaten, Eric Postma, and H. Herik. 2009. Dimensionality Reduction: A Comparative Review. *Journal of Machine Learning Research* 10 (1 2009), 66–71. https://members.loria.fr/moberger/Enseignement/AVR/Exposes/TR_Dimensiereductie.pdf

- [149] Baptist Vandersmissen, Nicolas Knudde, Azarakhsh Jalalvand, Ivo Couckuyt, Tom Dhaene, and Wesley De Neve. 2020. Indoor human activity recognition using high-dimensional sensors and deep neural networks. *Neural Computing and Applications* 32, 16 (Aug. 2020), 12295–12309. <https://doi.org/10.1007/s00521-019-04408-1>
- [150] Radu-Daniel Vatavu, Laurent Grisoni, and Stefan Gheorghe Pentiu. 2007. Gesture Recognition Based on Elastic Deformation Energies. In *Gesture-Based Human-Computer Interaction and Simulation, 7th International Gesture Workshop, GW 2007, Lisbon, Portugal, May 23-25, 2007, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 5085)*, Miguel Sales Dias, Sylvie Gibet, Marcelo M. Wanderley, and Rafael Bastos (Eds.). Springer, 1–12. https://doi.org/10.1007/978-3-540-92865-2_1
- [151] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (Seoul, Republic of Korea) (CHI 2015)*, Bo Begole, Jinwoo Kim, Kori Inkpen, and Woontack Woo (Eds.). ACM, 1325–1334. <https://doi.org/10.1145/2702123.2702223>
- [152] Radu-Daniel Vatavu. 2013. The impact of motion dimensionality and bit cardinality on the design of 3D gesture recognizers. *International Journal of Human-Computer Studies* 71, 4 (2013), 387 – 409. <https://doi.org/10.1016/j.ijhcs.2012.11.005>
- [153] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2012. Gestures as Point Clouds: A \$P Recognizer for User Interface Prototypes. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction (Santa Monica, California, USA) (ICMI '12)*. Association for Computing Machinery, New York, NY, USA, 273–280. <https://doi.org/10.1145/2388676.2388732>
- [154] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2018. \$Q: A Super-Quick, Articulation-Invariant Stroke-Gesture Recognizer for Low-Resource Devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (Barcelona, Spain) (MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, Article 23, 12 pages. <https://doi.org/10.1145/3229434.3229465>
- [155] Santiago Villarreal-Narvaez, Alexandru-Ionuț Șean, Arthur Sluÿters, Radu-Daniel Vatavu, and Jean Vanderdonckt. 2022. Informing Future Gesture Elicitation Studies for Interactive Applications That Use Radar Sensing. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces (AVI 2022)*. Association for Computing Machinery, New York, NY, USA, Article 50, 3 pages. <https://doi.org/10.1145/3531073.3534475>
- [156] O. Viunytskyi and A. Totsky. 2017. Novel bispectrum-based wireless vision technique using disturbance of electromagnetic field by human gestures. In *2017 Signal Processing Symposium (SPSymposium)*. 1–4. <https://doi.org/10.1109/SPS.2017.8053684>
- [157] Q. Wan, Y. Li, C. Li, and R. Pal. 2014. Gesture recognition for smart home applications using portable radar sensors. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 6414–6417. <https://doi.org/10.1109/EMBC.2014.6945096> ISSN: 1558-4615.
- [158] L. Wang, Z. Cao, Z. Cui, C. Cao, and Y. Pi. 2020. Negative Latency Recognition Method for Fine-Grained Gestures Based on Terahertz Radar. *IEEE Transactions on Geoscience and Remote Sensing* 58, 11 (Nov. 2020), 7955–7968. <https://doi.org/10.1109/TGRS.2020.2985421>
- [159] L. Wang, Z. Cui, Z. Cao, S. Xu, and R. Min. 2019. Fine-Grained Gesture Recognition Based on High Resolution Range Profiles of Terahertz Radar. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. 1470–1473. <https://doi.org/10.1109/IGARSS.2019.8900601> ISSN: 2153-7003.
- [160] P. Wang, J. Lin, F. Wang, J. Xiu, Y. Lin, N. Yan, and H. Xu. 2020. A Gesture Air-Writing Tracking Method that Uses 24 GHz SIMO Radar SoC. *IEEE Access* 8 (2020), 152728–152741. <https://doi.org/10.1109/ACCESS.2020.3017869>
- [161] Saiwen Wang, Jie Song, Jaime Lien, Ivan Poupyrev, and Otmar Hilliges. 2016. Interacting with Soli: Exploring Fine-Grained Dynamic Gesture Recognition in the Radio-Frequency Spectrum. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. Association for Computing Machinery, New York, NY, USA, 851–860. <https://doi.org/10.1145/2984511.2984565>
- [162] Yong Wang, Xiuqian Jia, Mu Zhou, Liangbo Xie, and Zengshan Tian. 2019. A novel F-RCNN based hand gesture detection approach for FMCW systems. *Wireless Networks* (July 2019), 1–14. <https://doi.org/10.1007/s11276-019-02096-2>
- [163] Y. Wang, A. Ren, M. Zhou, W. Wang, and X. Yang. 2020. A Novel Detection and Recognition Method for Continuous Hand Gesture Using FMCW Radar. *IEEE Access* 8 (2020), 167264–167275. <https://doi.org/10.1109/ACCESS.2020.3023187>
- [164] Y. Wang, S. Wang, M. Zhou, Q. Jiang, and Z. Tian. 2019. TS-I3D Based Hand Gesture Recognition Method With Radar Sensor. *IEEE Access* 7 (2019), 22902–22913. <https://doi.org/10.1109/ACCESS.2019.2897060>
- [165] Y. Wang, S. Wang, M. Zhou, W. Nie, X. Yang, and Z. Tian. 2019. Two-Stream Time Sequential Network Based Hand Gesture Recognition Method Using Radar Sensor. In *2019 IEEE Globecom Workshops (GC Wkshps)*. 1–6. <https://doi.org/10.1109/GCWkshps45667.2019.9024691>
- [166] Yong Wang, Zedong Zhao, Mu Zhou, and Jinjun Wu. 2019. Two Dimensional Parameters Based Hand Gesture Recognition Algorithm for FMCW Radar Systems. In *Wireless and Satellite Systems*, Min Jia, Qing Guo, and Weixiao Meng (Eds.). Vol. 280. Springer International Publishing, Cham, 226–234. https://doi.org/10.1007/978-3-030-19153-5_23 Series Title: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering.
- [167] Z. Wang, G. Li, and L. Yang. 2020. Dynamic Hand Gesture Recognition Based on Micro-Doppler Radar Signatures Using Hidden Gauss-Markov Models. *IEEE Geoscience and Remote Sensing Letters* 18, 2 (2020), 291–295. <https://doi.org/10.1109/LGRS.2020.2974821>
- [168] Zhu Wang, Xinye Lou, Zhiwen Yu, Bin Guo, and Xingshe Zhou. 2019. Enabling non-invasive and real-time human-machine interactions based on wireless sensing and fog computing. *Personal and Ubiquitous Computing* 23, 1 (Feb. 2019), 29–41. <https://doi.org/10.1007/s00779-018-1185-7>
- [169] Craig Warren, Antonios Giannopoulos, and Iraklis Giannakis. 2016. gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar. *Computer Physics Communications* 209 (2016), 163–170. <https://doi.org/10.1016/j.cpc.2016.08.020>
- [170] Frank Weichert, Daniel Bachmann, Bartholomäus Rudak, and Denis Fisseler. 2013. Analysis of the Accuracy and Robustness of the Leap Motion Controller. *Sensors* 13 (05 2013), 6380–6393. <https://doi.org/10.3390/s130506380>

- [171] Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI world: designing natural user interfaces for touch and gesture* (1 ed.). Morgan Kaufmann Publishers, Burlington, Massachusetts, United States. <https://doi.org/10.1016/C2009-0-64091-5>
- [172] Christian Wolff. 2022. Waves and Frequency Ranges. <https://www.radartutorial.eu/07.waves/WavesandFrequencyRanges.en.html>
- [173] Kaijun Wu and Sébastien Lambot. 2022. Effect of Radar Incident Angle on Full-Wave Inversion for the Retrieval of Medium Surface Permittivity for Drone-Borne Applications. *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), 1–10. <https://doi.org/10.1109/TGRS.2022.3157370>
- [174] Kaijun Wu, Gabriela Arambulo Rodriguez, Marjana Zajc, Elodie Jacquemin, Michiels Clément, Albéric De Coster, and Sébastien Lambot. 2019. A new drone-borne GPR for soil moisture mapping. *Remote Sensing of Environment* 235 (2019), 111456. <https://doi.org/10.1016/j.rse.2019.111456>
- [175] Te-Yen Wu, Shutong Qi, Junchi Chen, Mujie Shang, Jun Gong, Teddy Seyed, and Xing-Dong Yang. 2020. Fabriccio: Touchless Gestural Input on Interactive Fabrics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376681>
- [176] Z. Xia, Y. Luomei, C. Zhou, and F. Xu. 2020. Multidimensional Feature Representation and Learning for Robust Hand-Gesture Recognition on Commercial Millimeter-Wave Radar. *IEEE Transactions on Geoscience and Remote Sensing* 59, 6 (2020), 4749–4764. <https://doi.org/10.1109/TGRS.2020.3010880>
- [177] L. Yang and G. Li. 2018. Sparsity aware dynamic gesture recognition using radar sensors with angular diversity. *IET Radar, Sonar Navigation* 12, 10 (2018), 1114–1120. <https://doi.org/10.1049/iet-rsn.2018.5206>
- [178] Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. 2016. RadarCat: Radar Categorization for Input & Interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (Tokyo, Japan) (UIST '16)*. Association for Computing Machinery, New York, NY, USA, 833–841. <https://doi.org/10.1145/2984511.2984515>
- [179] Hui-Shyong Yeo, Ryosuke Minami, Kirill Rodriguez, George Shaker, and Aaron Quigley. 2018. Exploring Tangible Interactions with Radar Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4, Article 200 (Dec. 2018), 25 pages. <https://doi.org/10.1145/3287078>
- [180] Hui-Shyong Yeo and Aaron Quigley. 2017. Radar Sensing in Human-Computer Interaction. *Interactions* 25, 1 (Dec. 2017), 70–73. <https://doi.org/10.1145/3159651>
- [181] J. Yu, L. Yen, and P. Tseng. 2020. mmWave Radar-based Hand Gesture Recognition using Range-Angle Image. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*. 1–5. <https://doi.org/10.1109/VTC2020-Spring48590.2020.9128573> ISSN: 2577-2465.
- [182] Myoungseok Yu, Narae Kim, Yunho Jung, and Seongjoo Lee. 2020. A Frame Detection Method for Real-Time Hand Gesture Recognition Systems Using CW-Radar. *Sensors* 20, 8 (2020), 1–17. <https://doi.org/10.3390/s20082321>
- [183] Wei Zeng, Cong Wang, and Qinghui Wang. 2018. Hand Gesture Recognition Using Leap Motion via Deterministic Learning. *Multimedia Tools Appl.* 77, 21 (Nov. 2018), 28185–28206.
- [184] Zhengxin Zeng, Moeness G. Amin, and Tao Shan. 2020. Arm Motion Classification Using Time-Series Analysis of the Spectrogram Frequency Envelopes. *Remote Sensing* 12, 3 (2020), 1–20. <https://doi.org/10.3390/rs12030454>
- [185] Z. Zeng, M. G. Amin, and T. Shan. 2020. Automatic Arm Motion Recognition Based on Radar Micro-Doppler Signature Envelopes. *IEEE Sensors Journal* 20, 22 (Nov. 2020), 13523–13532. <https://doi.org/10.1109/JSEN.2020.3004581>
- [186] Shumin Zhai, Per Ola Kristensson, Caroline Appert, Tue Haste Andersen, and Xiang Cao. 2012. Foundational Issues in Touch-Surface Stroke Gesture Design - An Integrative Review. *Found. Trends Hum. Comput. Interact.* 5, 2 (2012), 97–205. <https://doi.org/10.1561/1100000012>
- [187] Bo Zhang, Lei Zhang, Mojun Wu, and Yan Wang. 2021. Dynamic Gesture Recognition Based on RF Sensor and AE-LSTM Neural Network. In *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*. 1–5. <https://doi.org/10.1109/ISCAS51556.2021.9401065>
- [188] G. Zhang, S. Lan, K. Zhang, and L. Ye. 2020. Temporal-Range-Doppler Features Interpretation and Recognition of Hand Gestures Using mmW FMCW Radar Sensors. In *2020 14th European Conference on Antennas and Propagation (EuCAP)*. 1–4. <https://doi.org/10.23919/EuCAP48036.2020.9135694>
- [189] G. Zhang, K. Zhang, Y. Yun, G. Lu, and S. Lan. 2019. Implementation of C4.5 decision tree in Human Gesture Recognition based on Doppler radars. In *2019 International Symposium on Antennas and Propagation (ISAP)*. 1–3.
- [190] J. Zhang and Z. Shi. 2017. Deformable deep convolutional generative adversarial network in microwave based hand gesture recognition system. In *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*. 1–6. <https://doi.org/10.1109/WCSP.2017.8170976> ISSN: 2472-7628.
- [191] Jiajun Zhang, Jinkun Tao, and Zhiguo Shi. 2019. Doppler-Radar Based Hand Gesture Recognition System Using Convolutional Neural Networks. In *Communications, Signal Processing, and Systems*, Qilian Liang, Jiasong Mu, Min Jia, Wei Wang, Xuhong Feng, and Baoju Zhang (Eds.). Vol. 463. Springer Singapore, Singapore, 1096–1113. https://doi.org/10.1007/978-981-10-6571-2_132 Series Title: Lecture Notes in Electrical Engineering.
- [192] K. Zhang, Z. Yu, D. Zhang, Z. Wang, and B. Guo. 2020. RaCon: A gesture recognition approach via Doppler radar for intelligent human-robot interaction. In *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. 1–6. <https://doi.org/10.1109/PerComWorkshops48775.2020.9156109>
- [193] S. Zhang, G. Li, M. Ritchie, F. Fioranelli, and H. Griffiths. 2016. Dynamic hand gesture classification based on radar micro-Doppler signatures. In *2016 CIE International Conference on Radar (RADAR)*. 1–4. <https://doi.org/10.1109/RADAR.2016.8059518>
- [194] X. zhang, Q. Wu, and D. Zhao. 2018. Dynamic Hand Gesture Recognition Using FMCW Radar Sensor for Driving Assistance. In *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*. 1–6. <https://doi.org/10.1109/WCSP.2018.8555642> ISSN: 2472-7628.
- [195] Zhenyuan Zhang, Zengshan Tian, Zhou Mu, and Yi Liu. 2018. Application of FMCW Radar for Dynamic Continuous Hand Gesture Recognition. In *Proceedings of the 11th EAI International Conference on Mobile Multimedia Communications (MOBIMEDIA '18)*. ICST (Institute for Computer

- Sciences, Social-Informatics and Telecommunications Engineering), Brussels, BEL, 298–303.
- [196] Z. Zhang, Z. Tian, Y. Zhang, M. Zhou, and B. Wang. 2019. u-DeepHand: FMCW Radar-Based Unsupervised Hand Gesture Feature Learning Using Deep Convolutional Auto-Encoder Network. *IEEE Sensors Journal* 19, 16 (Aug. 2019), 6811–6821. <https://doi.org/10.1109/JSEN.2019.2910810>
 - [197] Z. Zhang, Z. Tian, and M. Zhou. 2018. Latern: Dynamic Continuous Hand Gesture Recognition Using FMCW Radar Sensor. *IEEE Sensors Journal* 18, 8 (April 2018), 3278–3289. <https://doi.org/10.1109/JSEN.2018.2808688>
 - [198] Z. Zhang, Z. Tian, and M. Zhou. 2019. SmartFinger: A Finger-Sensing System for Mobile Interaction via MIMO FMCW Radar. In *2019 IEEE Globecom Workshops (GC Wkshps)*. 1–5. <https://doi.org/10.1109/GCWkshps45667.2019.9024578>
 - [199] Z. Zhang, Z. Tian, M. Zhou, W. Nie, and Z. Li. 2018. Riddle: Real-Time Interacting with Hand Description via Millimeter-Wave Sensor. In *2018 IEEE International Conference on Communications (ICC)*. 1–6. <https://doi.org/10.1109/ICC.2018.8422765> ISSN: 1938-1883.
 - [200] Z. Zhao, Y. Wang, M. Zhou, X. Yang, and L. Xie. 2019. Interference Suppression Based Gesture Recognition Method with FMCW Radar. In *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*. 1–6. <https://doi.org/10.1109/WCSP.2019.8928108> ISSN: 2472-7628.
 - [201] Zhi Zhou, Zongjie Cao, and Yiming Pi. 2018. Dynamic Gesture Recognition with a Terahertz Radar Based on Range Profile Sequences and Doppler Signatures. *Sensors* 18, 1 (2018), 1–15. <https://doi.org/10.3390/s18010010>
 - [202] Shangyue Zhu, Junhong Xu, Hanqing Guo, Qiwei Liu, Shaoen Wu, and Honggang Wang. 2018. Indoor Human Activity Recognition Based on Ambient Radar with Signal Processing and Machine Learning. In *2018 IEEE International Conference on Communications (ICC)*. 1–6. <https://doi.org/10.1109/ICC.2018.8422107>

A TYPES OF RADAR

Type	Frequency band	Reference(s)
CW	L	[156]
	S	[125, 126, 157]
	C	[190, 191]
	X	[40, 41, 70, 102, 175, 177, 193]
	K	[9–11, 17, 21, 47, 74, 75, 84–86, 91, 98, 104, 133, 167, 168, 177, 182, 184, 185, 189]
	V	[129]
FMCW	K	[2, 21, 27, 35, 51, 73, 106, 121, 122, 124, 140, 160, 195–197, 199]
	V	[16, 22, 29, 31, 53–55, 79, 92, 116, 123, 138, 142, 143, 147, 161]
	W	[37, 38, 42, 49, 51, 60, 61, 93, 94, 96, 101, 107, 107, 119, 128, 141, 144, 149, 162–166, 176, 181, 188, 194, 198, 200]
	THz	[158, 159, 201]
	Unknown	[59]
Pulse	S	[65, 115]
	C	[62, 66, 67, 81, 82, 87, 132]
	X	[3, 4, 45, 51, 63]
	K	[192]
	V	[39, 43]
	Unknown	[89, 114, 134]
DSSS	V	[92]
Unknown	Unknown	[88]

Table 5. Distribution of papers from our systematic literature review according to the type and frequency band of the radar(s).

Received June 2022; revised November 2022