IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 21 April 2026, accepted 5 May 2026, date of publication 7 May 2026, date of current version 12 May 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3691174

RESEARCH ARTICLE

mmHSense: Multi-Modal and Distributed mmWave ISAC Datasets for Human Sensing

NABEEL NISAR BHAT[1], MAKSIM KARNAUKH[1], STEIN VANDENBROEKE[1], WOUTER LEMOINE[1], JAKOB STRUYE[1], SIDDHARTHA KUMAR[2], MOHAMMAD HOSSEIN MOGHADDAM[2], JESUS OMAR LACRUZ[3], JOERG WIDMER[3], RAFAEL BERKVENS[1], AND JEROEN FAMAEY[1]

[1]IDLab–imec Research Group, University of Antwerp, Antwerp 2000, Belgium
[2]Qamcom Research and Technology AB, 41285 Gothenburg, Sweden
[3]IMDEA Networks Institute, 28918 Madrid, Spain

Corresponding author: Nabeel Nisar Bhat (nabeelnisar.bhat@uantwerpen.be)

This work was supported in part by the Research Foundation-Flanders (FWO) Project WaveVR under Grant G034322N; in part by the European Commission through the Horizon Europe Joint Undertaking (JU) Smart Networks and Services (SNS) Project Hexa-X-II under Grant 101095759; in part by the European Union's Horizon Europe Program through the SNS-JU through under Grant 101192521 (MultiX); in part by the Comunidad de Madrid through Projects DISCO6G-CM and TUCAN6-CM under Grant TEC-2024/COM-360 and Grant TEC-2024/COM-460 under ORDEN 5696/2024; and in part by Ministerio de Ciencia e Innovación (Ministry of Science and InnovaEon) (MCIN)/Agencia Estatal de InvesEgación (State Research Agency) (AEI)/10.13039/501100011033/Fondo Europeo de Desarrollo Regional (European Regional Development Fund) (FEDER), European Union (EU), under Grant Project PID2022-136769NB-I00 (ELSA). The work of Nabeel Nisar Bhat was supported by the FWO Strategic Basic (SB) Ph.D. Fellowship under Grant 1SH5X24N.

ABSTRACT This article presents *mmHSense*, a set of six labeled mmWave datasets to support human sensing research within Integrated Sensing and Communication (ISAC) systems. The datasets can be used to explore mmWave ISAC for various end applications such as gesture recognition, person identification, pose estimation, and localization. Moreover, the datasets can be used to develop and advance signal processing and deep learning research on mmWave ISAC. This article describes the testbed, experimental settings, and signal features for each dataset. The datasets cover both Wi-Fi and 5G signals. Furthermore, the utility of the datasets is demonstrated through validation on different downstream tasks such as pose estimation, gesture recognition, localization, and person identification. In addition, we demonstrate the use of parameter-efficient fine-tuning to adapt ISAC models to different tasks, significantly reducing computational complexity while maintaining performance on prior tasks.

INDEX TERMS mmWave, integrated sensing and communication, datasets, ISAC, deep learning, testbeds, extended reality.

I. INTRODUCTION

Integrated Sensing and Communication (ISAC) [1] enables communication networks to double as intelligent sensing systems, enabling advanced human sensing applications. For instance, ISAC can enable Wi-Fi routers to recognize human gestures in smart-home applications [2]. In Extended Reality (XR) setups, the same communication signals can track body pose without external cameras and hand-held controllers, creating more immersive and portable experiences. In autonomous vehicles, ISAC can assist in precise localization [2]. Since ISAC reuses the existing wireless infrastructure, these sensing capabilities can be deployed widely without the high costs of dedicated sensors.

ISAC at higher frequencies, such as millimeter waves (mmWave) [3] provides added benefits through its high bandwidth and large antenna arrays, enabling fine range and angular resolution. This leads to more precise sensing, which can greatly enhance the performance of applications such as gesture recognition and localization. Although sub-6 GHz ISAC has experienced rapid advancements in recent years, progress in mmWave ISAC research has been comparatively slower. This slower progress is mainly due to the limited availability of commercial-off-the-shelf (COTS) mmWave devices, the absence of open-source tools for mmWave systems that give access to signal features such as channel state information (CSI), and the high cost

The associate editor coordinating the review of this manuscript and approving it for publication was Julien Sarrazin.

associated with building custom experimental mmWave ISAC platforms [4].

This work introduces *mmHSense*,[1] a collection of six mmWave ISAC datasets designed to advance research on human sensing and bridge existing gaps in the field. The datasets incorporate various signal features, including mmWave CSI, beam SNR, and power per beam (PPBP), collected using both COTS devices and custom experimental platforms (software-defined radio). Our datasets cover sub-6 GHz, mmWave Wi-Fi signals and 5G mmWave orthogonal frequency division multiplexing (OFDM) signals, across bi-static and multi-static setups. Most existing mmWave datasets are radar-based. While DISC [4] uses communication hardware, it is limited by its monostatic Transmitter (Tx)–Receiver (Rx) configuration, which does not reflect real-world deployment scenarios. In practical deployments, Tx and Rx are spatially separated. Moreover, the dataset includes only five activities.

Further, unlike radar-based approaches that rely on dedicated waveforms for sensing [5], *mmHSense* relies on communication waveforms that are not specifically designed for sensing. As a result, sensing parameters are derived from signals affected with various *impairments* [6] such as synchronization errors, carrier frequency offset, phase noise, and other hardware impairments, reflecting conditions encountered in practical ISAC deployments. Therefore, our datasets benchmark ISAC-based sensing against radar-level performance across diverse downstream tasks, without dedicated hardware.

The key contributions of our work can be summarized as follows:

- **Comprehensive mmWave ISAC datasets:** We introduce *mmHSense*, a collection of mmWave ISAC datasets to support a range of end applications, including gesture recognition, pose estimation, localization, and person identification. This work addresses the existing shortage of publicly available mmWave ISAC datasets for human sensing.
- **Scalable COTS-based setups**: Most of our datasets are collected using COTS Wi-Fi devices rather than dedicated mmWave radars. This makes the setups scalable and easily reproducible, ideal for real-world deployment, and also representative of fully integrated ISAC systems, where communication signals are leveraged for sensing tasks.
- **Signal features beyond CSI**: Our datasets also include beam SNR and PPBP as signal features, in addition to CSI. Unlike CSI, beam SNR and PPBP are readily available as part of the standard sector sweep or beam sweeping process, requiring no additional overhead for extraction.
- **User, environment, and gesture diversity**: Unlike many existing mmWave ISAC datasets, which often focus on a limited number of users, environments, or gesture types, our datasets encompass a wide variety of users and environments (ranging from isolated rooms to open corridors), and include both predefined and natural gestures and poses. Furthermore, the data collection campaigns were conducted across multiple countries, including Spain, Belgium, and Sweden, ensuring greater geographic and environmental diversity.
- **Distributed ISAC:** Most ISAC experiments focus on a single Tx-Rx setup, which at mmWave frequencies requires line-of-sight between users and devices. To overcome this limitation, one of our datasets, *DISAC-mmVRPose*, explores distributed ISAC with four Rxs and a single Tx, enabling more robust coverage and sensing capabilities.
- **Natural poses**: Moreover, most ISAC datasets focus on predefined gestures and poses [7]. In contrast, *DISAC-mmVRPose* dataset captures realistic and natural gestures by allowing users to interact with a Virtual Reality (VR) game, resulting in more genuine poses and movements.
- **Multi-modal sensor fusion**: In most of our datasets, we incorporate multi-modal setups combining mmWave, sub-6 GHz Wi-Fi, and vision-based sensors. This enables pose estimation and supports advanced multi-sensor fusion techniques.
- **5G mmWave OFDM signals:** While most existing ISAC research relies on Wi-Fi signals [4], [8], our *5GmmGesture* dataset explores OFDM mmWave waveforms for gesture recognition. This serves as a primer for future 6G research, highlighting the potential of next-generation communication signals for ISAC.
- **Efficient fine-tuning with LoRA:** We explore the use of Low-Rank Adaptation to efficiently fine-tune ISAC models on new tasks without losing performance on the prior task. This significantly reduces training computational overhead, enabling scalable and resource-efficient training for 6G foundation models [9].
- **Data and code:** Datasets are available via IEEE Dataport[2] and the code is available via GitHub.[3]

II. STRUCTURE OF THE PAPER

The remainder of the paper is structured as follows: Section III details the related work concerning mmWave ISAC. Section IV describes the mmHSense dataset. Section V discusses the results on diverse tasks. Section VI describes the link between communication and sensing. Section VII outlines novel opportunities and challenges to advance ISAC sensing. Finally, Section VIII describes the conclusion.

[1] This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Antwerp Ethics Committee for the Social Sciences and Humanities (EA SHW) under Application No. SHW_2023_313_2.

[2] https://ieee-dataport.org/documents/mmwavexr-multi-modal-and-distributed-mmwave-isac-datasets-human-sensing

[3] https://github.com/nisarnabeel/Multi-Modal-and-Distributed-mmWave-ISAC-Datasets-for-Human-Sensing/tree/main

TABLE 1. Overview of the mmHSense. GR = gesture recognition, PE = pose estimation, Loc = localization, ID = identification.

Specs.	mmWGesture	mmWPose	mmW-GaitID	mmW-Loc	5GmmGesture	DISAC-mmVRPose
Hardware	Talon-AD7200	wAP60Gx3	wAP60Gx3	wAP60Gx3	Sivers EVK06002	Sivers EVK06002
Signal feature	Beam SNR	CSI amplitude	CSI amplitude	CSI amplitude	PPBP	CIR
#Hz mmWave	10	22	10	10	1540	2775
Task	GR	PE	Person ID	Loc	GR	PE
#Subjects	3	3	20	20	8	8
Labels	10 classes	Regression	8 classes	20 classes	8 classes	Regression
Geometry	1 Tx–Rx	1 Tx–Rx	2 Tx–Rx pairs (X)	2 Tx–Rx pairs (X) + Grid 4 × 5	1 Tx–1 Rx	1 Tx–4 Rx
Technology	IEEE 802.11ad	IEEE 802.11ad	IEEE 802.11ad	IEEE 802.11ad	5G NR	IEEE 802.11ay
Aux Modality	5 GHz CSI	Kinect skeletons	5 GHz CSI	5 GHz CSI	NA	Kinect skeletons
Dataset size (examples)	854	2904	1318	2177	34496	32480
Duration (minutes)	220	110	40	60	72	14

III. RELATED WORK

A. OVERVIEW OF EXISTING mmWave ISAC DATASETS

mmWave ISAC has gained significant traction in recent years; however, there is a scarcity of publicly available large-scale datasets supporting diverse downstream tasks.

Yu et al. [10], in their work, used COTS mmWave Wi-Fi for human pose and seat occupancy detection. The authors used Talon AD-7200 Wi-Fi devices in their experimental setup. Beam SNR was used as a signal feature. The experiments were limited to a single person and involved a set of 8 poses only. Pegoraro et al. [4] set up a software-defined radio testbed supporting the IEEE 802.11ay standard at 60 GHz frequency in an indoor environment. The downstream task was activity recognition through channel impulse responses. However, only 5 different activities were considered for a total of 7 subjects. Moreover, it employs a monostatic Tx-Rx configuration, which does not reflect practical real-world deployment scenarios. Ali et al. [11] investigate the use of mmWave radar for gesture recognition; however, such approaches rely on specialized and often costly dedicated sensing hardware.

Other datasets, such as MM-FI [7], also support multi-modal setups, consisting of LiDAR, mmWave radar, RGB images, and 5 GHz Wi-Fi. The dataset comprises 27 categories of human actions, all predefined and performed in place, without accounting for natural user movement. In contrast, DISAC-mmVRPose stimulates natural interactions based on an actual VR game, allowing subjects to act and move freely.

While there are well-established datasets for sub-6 GHz Wi-Fi sensing (e.g., SenseFi [8]), the *mmWave bands* (24.25 GHz to 71.0 GHz) lack similar publicly available benchmark datasets, hindering research advances, reproducibility, and fair comparison.

IV. mmWAVE ISAC DATASETS: mmHSense

Table 1 shows the specifications (specs.) of different datasets, such as hardware used, signal feature, sampling rate (#Hz) of signal feature, task, subjects, task type (labels), geometry, technology, auxiliary modality (Aux modality), dataset size, and duration. As illustrated in Figure 1, all datasets share a common experimental foundation based on mmWave wireless devices and bi-static or multi-static Tx–Rx configurations, which form the *sensing backbone* across deployments. From these configurations, different signal features, such as CSI, beam SNR, or PPBP, are extracted depending on the dataset. These features can be, for example, directly fed to a neural network in an *end-to-end* fashion, or undergo additional signal processing when required. Collectively, these datasets support various downstream tasks such as gesture recognition, localization, person identification, and pose estimation, and enable systematic benchmarking of mmWave sensing, in contrast to existing sub-6 GHz benchmarks such as SenseFi [8]. Three of the datasets are collected using identical hardware configurations, enabling systematic investigation of *transfer learning and foundation learning*-based approaches. Furthermore, while mmWPose considers simple and predefined gestures, DISAC-mmVRPose captures natural poses stimulated by an actual VR game. This enables evaluation of mmWave sensing performance as interaction complexity increases from *constrained gestures* to *immersive continuous user interactions* in a VR setting.

A. GESTURE RECOGNITION

1) mmWGesture

In the first dataset, we use a pair of TP-Link Talon AD 7200 mmWave Wi-Fi devices that support the IEEE 802.11ad standard in a bi-static configuration (as shown in Figure 1). These devices periodically perform a sector sweep to determine the optimal beam for communication based on signal strength. Using an open-source tool,[4] we get access to beam SNR for 36 such sectors. A user interrupts the line-of-sight between two devices functioning as Tx and Rx and performs a set of gestures and poses that trigger a change in beam SNR.

We consider static gestures (poses) as well as dynamic gestures, including head rotations. We collect data in multiple environments, the home and typical office space. In the home environment, we collected 125 minutes of activity data from a single user, resulting in 486 gesture examples.

[4] https://seemoo.de/talon-tools

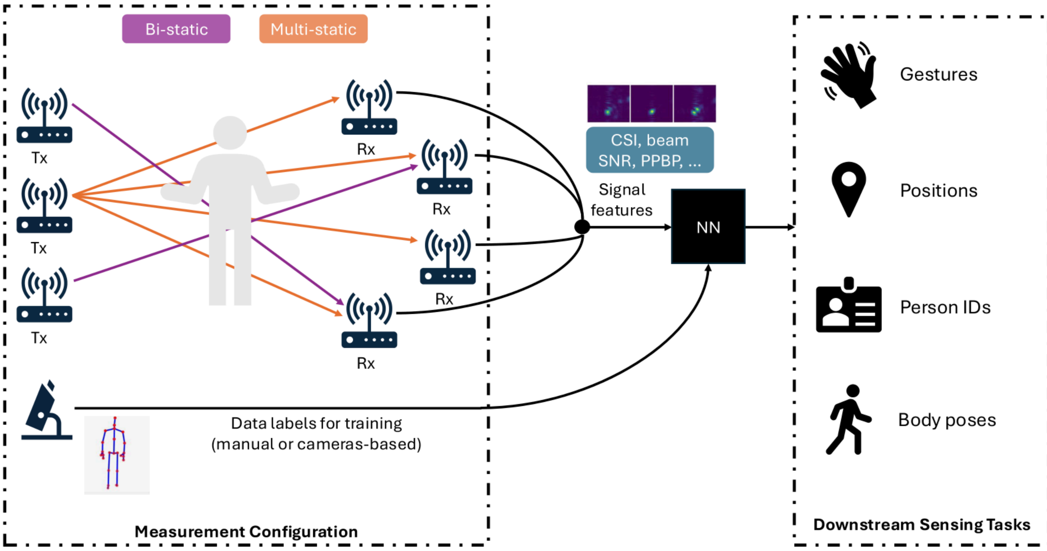


FIGURE 1. Illustration of the mmWave ISAC-based experimental setup and processing pipeline. Multiple transmitter–receiver configurations (bi-static, two-pairs, and multi-static 1 Tx - 4 Rxs) are deployed to capture wireless signal features (e.g., CSI, beam SNR, or PPBP). The extracted features are processed by a neural network (NN) to enable diverse downstream tasks, including gesture recognition, localization, person identification, and pose estimation. Skeleton data and manual labels are incorporated only during the training phase.

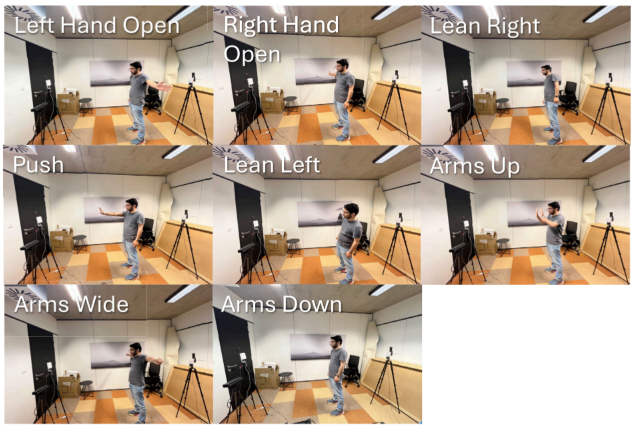


FIGURE 2. Gestures and poses considered for Pose and 5GmmGesture dataset.

In the office environment, we recorded 221 gesture examples from three different users. Additionally, we conducted a rotation-based experiment in the home environment, where a single user performed 147 gestures while rotating 90 degrees, capturing orientation variability. Across all experiments, the total recorded activity time amounts to approximately 220 minutes. Note that the sampling frequency of the beam SNR is tied to the beacon interval of 102.4 ms, resulting in approximately 10 samples per second. More details on the performed gestures are provided in our previous work [12].

The overall size of the dataset is 427000 × 36, where 36 represents the number of sectors, and 427000 represents the number of samples. This can be, for example, reshaped into 854 × 36 × 500, where 500 is the time index for each sample. Thus, a single gesture instance corresponds to a matrix capturing the temporal evolution of 36 features across 500 time steps.

2) 5GmmGesture

This ISAC system consists of bi-static 5G mmWave nodes configured as a Tx and Rx, implemented using two Sivers EVK06002 evaluation kits. These EVKs include front-end modules supporting frequencies in the 57–71 GHz range and are controlled via a programmable RF System-on-Chip (RFSoC).

The system transmits a 5G New Radio OFDM waveform consisting of a Synchronization Signal Block (SSB) and randomized data over a single carrier comprising 792 subcarriers, spaced 960 kHz apart—resulting in a total bandwidth of approximately 760 MHz centered around 60 GHz. The transmitter continuously sends a 10 ms frame in compliance with 5G standards, composed of 10 subframes and 112 OFDM symbols. Each OFDM symbol includes a cyclic prefix (CP) of variable length and is processed using a 1024-point FFT. Beam sweeping is conducted over 50 transmit and 56 receive beams, with two OFDM symbols transmitted per beam pair. The RFSoC is configured to sample the received signal such that exactly two OFDM symbols—excluding their CPs—are captured for each Tx-Rx beam pair. The sampled received signal is then convolved with the corresponding transmit signal (excluding CPs) in the time domain. To ensure efficient on-chip processing, the convolution is performed using an FFT-based method, computed in blocks of up to 1024 samples. The resulting convolved signals are used to calculate the power per beam pair. This is done by computing the signal power and averaging across the samples and OFDM symbols. Finally, the averaged PPBP values are reshaped into a 50 × 56 matrix to form the power grid, which serves as the input feature for gesture classification. This grid captures spatial energy distributions across the entire beam space, enabling fine-grained sensing and classification. We collect power per beam pair data from 8 diverse users performing 8 distinct gestures (as shown in Figure 2) for a total of 74.6 minutes of activity, with each gesture performed for 7 seconds and repeated 10 times. Unlike existing works that focus on CSI or micro-Doppler, PPBP is more robust (e.g., to phase noise), and practical (e.g., not requiring advanced synchronization). Moreover, the PPBP measurements can be easily extracted from the communication signals themselves, rather than relying on dedicated sensing packets. This approach is especially suitable for ISAC, entailing no overhead or additional resource utilization.

The overall size of the dataset is 34496 × 20 × 50 × 56, where 34496 is the number of examples, 20 represents the time dimension, 50 and 56 represent the number of transmitter and receiver beams, respectively.

B. POSE ESTIMATION

1) mmWPose

This dataset is designed for skeletal body pose estimation. Data is collected using a single Tx–Rx pair, while a synchronized Kinect system provides the ground truth. The synchronization was done using Network Time Protocol (NTP). The Kinect captures the 3D coordinates (x, y, z) of 25 human body joints. We employ MikroTik wAP 60Gx3 routers operating under the IEEE 802.11ad standard, with CSI amplitude extracted from 30 antenna elements using an open-source tool.[5] The CSI is sampled at 22 Hz. Three participants perform eight distinct poses, shown in Figure 2. Each pose was held for a duration of 15 seconds, and the entire session involved around 20 rounds of performing the set of eight poses, yielding a total of 110 minutes of activity. The multimodal pairing of mmWave and vision data facilitates cross-modal learning and radio-to-vision translation.

The overall size of the dataset is 145200 × 30, where 145200 represents the number of samples, and 30 represents the number of antenna elements. This is reshaped into 2904 × 30 × 50, where 50 represents the time index.

2) DISAC-mmVRPose

We deploy a distributed ISAC system based on the Mimorph testbed [13]. For this data collection, the setup consists of a single mmWave Tx and four mmWave Rxs, all based on the Sivers Semiconductors EVK06002 phased-array antenna kits[6] operating at a 60.48 GHz carrier frequency, and all connected to the same RFSoC board. Full synchronization across the distributed receivers is achieved by leveraging the RFSoC multi-tile synchronization (MTS) capabilities, which provide deterministic alignment of sampling clocks and phase coherence across ADC tiles, enabling phase-level synchronization among all Rxs [14]. The four receivers additionally share a common 45 MHz local oscillator, ensuring tight frequency alignment and limiting residual carrier frequency offsets. Any remaining frequency mismatch is compensated in post-processing using standard CFO estimation techniques based on the IEEE 802.11ay preamble structure [15]. Each receiver is connected to the RFSoC through long cables, enabling their spatial distribution in the environment, as shown in Figure 3; these cables may introduce fixed timing and phase offsets, which manifest as small, static biases in range estimation and can be readily calibrated out during system setup.

The Tx sends training packets following the IEEE 802.11ay standard [16], each including training fields (TRN) transmitted with different beam patterns that uniformly sweep the azimuth range from −45° to +45°. Each packet also includes a preamble composed of a short training field (STF) and channel estimation fields (CEF), which are used for packet detection, synchronization, and channel estimation. From the received TRN fields, channel impulse responses (CIRs) are obtained via correlation with the known Golay sequences and recorded for all receivers and beam patterns. The resulting CIR time series can be used to extract range profiles and micro-Doppler signatures by (i) aligning CIRs across consecutive packets to remove residual timing offsets, (ii) compensating carrier frequency offset using a static reference path, and (iii) applying time–frequency analysis to the phase evolution of selected CIR taps, following the processing pipeline described in [17]. Packets are transmitted with an inter-frame spacing of 0.3 ms, enabling Doppler and micro-Doppler analysis of human motion.

Ground-truth human poses are captured using a Microsoft Kinect that is time-synchronized with the Mimorph testbed. Both systems are controlled from the same server and synchronized using NTP, enabling timestamp alignment with millisecond-level accuracy. Data collection involves participants performing a series of gestures while interacting with a VR game, enabling the capture of realistic and naturalistic pose and gesture sequences. 8 volunteers participated to ensure diversity in the dataset. For each volunteer, we record approximately 105 seconds of VR interaction.

Unlike most existing works, which are often limited to predefined gestures or a single Tx–Rx pair, our dataset captures unconstrained, natural movements from multiple *synchronized* receivers, thereby providing multiple viewpoints of each action and enabling research into advanced multi-view data fusion techniques.

The dataset has dimensions 32480 × 4 × 67 × 1536, where the first dimension represents the number of examples, the second dimension corresponds to the four Rxs, the third dimension indexes time (67 samples), and the last dimension encodes the beam–range grid (12 × 128 = 1536; 12 beams and 128 range bins).

[5] https://github.com/IMDEANetworksWNG/Mikrotik-researchertools/tree/main

[6] Sivers Semiconductors, "Evaluation kit evk06002 for 57–71 GHz, unlicensed 5G mmWave (IEEE 802.11ad)," https://www.sivers-semiconductors.com/5g-millimeter-wave-mmwave-and-satcom/wireless-products/evaluation-kits/evaluation-kit-evk06002/,2025

C. PERSON IDENTIFICATION AND LOCALIZATION FINGERPRINTING

In these datasets, we employ MikroTik wAP 60Gx3 routers operating under the IEEE 802.11ad standard (same as mmWPose), with CSI amplitude extracted from 30 antenna elements using an open-source tool.[7]

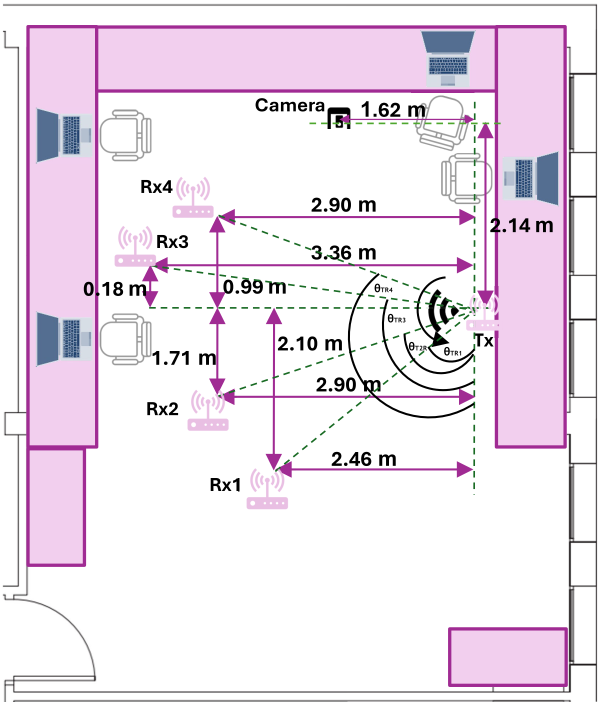


FIGURE 3. Distributed ISAC setup VR Pose estimation: DISAC-mmVRPose dataset.

The mmWave CSI is sampled at 10 Hz. These datasets also have an additional auxiliary modality, 5 GHz Wi-Fi, along with mmWave Wi-Fi. Since only CSI amplitudes are used, these datasets do not require additional preprocessing, such as *phase sanitization*, making them particularly suitable for *end-to-end deep learning approaches*.

1) mmW-GaitID

This dataset focuses on gait-based person identification. The mmWave setup consists of four wAP 60Gx3 units arranged as two independent Tx-Rx pairs in an X-shaped configuration. Figure 4 shows the experimental configuration of the setup. Arrows indicate the walking path. For cross-frequency comparison, a matching 5 GHz system (IEEE 802.11ac) is deployed using two ASUS RT-AC86U routers (AP + passive monitor) with CSI extraction obtained via Nexmon.[8] An Intel laptop sends ICMP echo requests to the AP, with the monitor capturing the CSI responses. The 5 GHz CSI is downsampled to 10 Hz to match the 60 GHz data.

Data collection spans three days with 20 participants (7/7/6), each walking for ∼2 minutes in a straight-line path with turns at each end. Background samples without any person present are also recorded. The 60 GHz CSI has shape ≈ 26,360 × 30 per device pair (30 representing the number of antenna elements), while the 5 GHz CSI has shape ≈ 520,000 × 52, where 52 represents the number of subcarriers. The 60 GHz CSI can be reshaped, for example, 1318 × 60 × 20 across two devices.

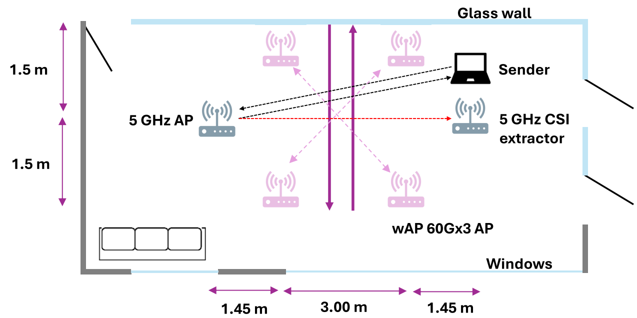


FIGURE 4. Environment and experimental setup for mmW-GaitID and mmW-Loc.

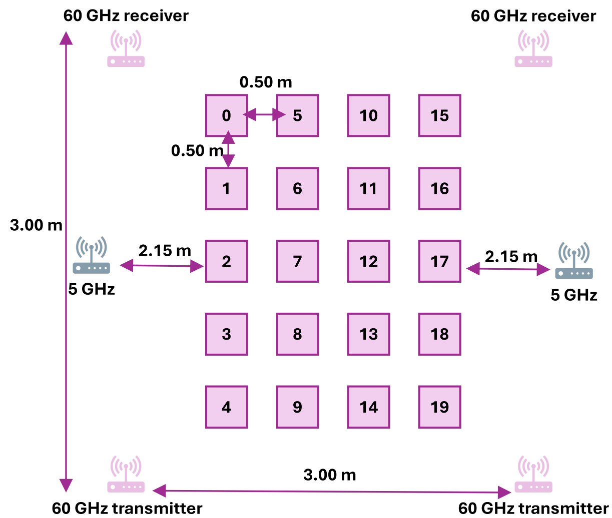


FIGURE 5. Localization fingerprinting grid for mmW-Loc dataset.

2) mmW-Loc

This dataset addresses person localization fingerprinting over a discrete 4 × 5 spatial grid. The hardware configuration and the environment are identical to GaitID, with simultaneous recording of 60 GHz and 5 GHz CSI. Data is collected from 20 participants, each holding a fixed pose (standing still) for 10 s at each of the 20 grid locations, shown in Figure 5. This setup enables evaluation of cross-frequency localization performance and robustness across both spatial variation and participant diversity.

The overall size of the mmWave CSI is 43,540 × 30 per device pair, which can be reshaped into 2177 × 60 × 20across two devices. For the 5 GHz CSI, the size is 1,404,060 × 52.

V. RESULTS

This section describes the validation of the datasets on specific end tasks. For the mmWGesture dataset and the

[7]https://github.com/IMDEANetworksWNG/Mikrotik-researcher-tools/tree/main

[8]https://github.com/seemoo-lab/nexmon

mmWPose datasets, the results are described in our previous works [12], [18]. To briefly summarize, we achieved 96.7% accuracy in gesture recognition using mmWave Wi-Fi in a single-person and single environment setting with the mmWGesture dataset [12]. With the mmWPose dataset [18], we showed that mmWave Wi-Fi can be used to achieve Mean Per Joint Position Error (MPJPE) in the range of 6.9 to 7.2 cm, for human pose estimation, comparable to vision-based methods [19].

A. 5GmmGesture

We split the PPBP dataset into training, validation, and test sets with proportions of 70%, 10%, and 20%, respectively. The model is trained for 150 epochs using a learning rate of 3×10^{-4} and a fixed random seed of 42. The same experimental setup is applied consistently across all other datasets. Here, the three sets contain PPBPs from all 8 users (in-domain settings). We leverage a ResNet-styled network [20] to automatically extract features from PPBPs and map changes in PPBPs to the corresponding gestures, without any additional processing. We use ResNet because of skip connections and the ability to reduce overfitting when training a deeper model. Table 2 shows the classification accuracy of the neural network for gesture recognition along with size and GPU VRAM. Specifically, the model size represents the amount of RAM required to load the model parameters into memory. GPU VRAM indicates the approximate memory needed during training, which includes not only the model weights but also additional storage for gradients, optimizer states, and intermediate activations. These metrics provide a clear indication of the feasibility of performing such tasks on edge or resource-constrained devices, where memory and computational resources are limited. We can see that ResNet18 achieves 97.75% accuracy on the test set while ResNet34 lags, due to complexity, resulting in slight overfitting.

However, when we evaluate the model in a cross-domain setting, we observe a notable drop in performance. In particular, the model is trained on data from seven users and tested on the held-out user. Table 3 shows the results for the unseen user. The accuracy varies considerably, ranging from 54% to 85%, with an overall average of 71%. This result underscores the challenges introduced by user-specific differences, commonly referred to as the *domain gap*.

B. DISAC-mmVRPose

This dataset enables radio-to-vision translation, specifically mapping radio signals to 3D skeletons. It can also be leveraged for evaluating distributed sensing frameworks. This is a regression task; therefore, *Mean Squared Error* is used as a loss function instead of *Cross Entropy Loss* for classification tasks. Skeletons obtained from a Kinect camera serve as ground truth.

From the raw data obtained from four Rxs, it is possible to obtain channel impulse response (CIR), and micro-Doppler. In this work, our focus is on CIR. We show pose estimation performance evaluated on 8 individuals. Table 2 reports the *MPJPE*, which is a commonly used metric for pose estimation. MPJPE measures the average Euclidean distance between the predicted and ground-truth 3D joint locations across all joints and frames. MPJPE is defined as:

$$\text{MPJPE} = \frac{1}{N \cdot J} \sum_{n=1}^{N} \sum_{j=1}^{J} \left\| \hat{\mathbf{p}}_{n,j} - \mathbf{p}_{n,j} \right\|_2 , \tag{1}$$

where N is the number of frames, J is the number of joints, $\hat{\mathbf{p}}_{n,j} \in \mathbb{R}^3$ denotes the predicted 3D position of joint j in frame n, and $\mathbf{p}_{n,j} \in \mathbb{R}^3$ is the corresponding ground-truth position.

We achieve an MPJPE of approximately 6.2 cm using ResNet18 model, shown in Table 2, which falls within the same magnitude as that achieved by camera-based computer vision methods. For example, Zhang et al. [21] in their recent work attain around 4.9 cm MPJPE using a spatio-temporal encoder, while in complex environments Gerats et al. [19] achieve MPJPE in the range of 8–12.6 cm. Therefore, our mmWave-only system produces competitive accuracy, on par with camera-based computer vision systems, while relying entirely on mmWave CIR at test time.

C. mmW-Loc AND mmW-GaitID

In this experiment, we present results on mmWave CSI, where the GaitID dataset is reshaped to $1318 \times 60 \times 20$ and the mmW-Loc dataset to $2177 \times 60 \times 20$, with 60 representing the concatenation of 30 antenna elements from the two Rxs, 20 representing the time index, while 1318 and 2177 represent the number of examples for the two datasets, respectively. Beyond reporting performance on identification and localization fingerprinting, we emphasize a critical challenge in ISAC-based sensing systems: catastrophic forgetting [22], [23]. Specifically, when a neural network trained on one task is subsequently retrained on another, its performance on the original task deteriorates significantly. This problem is not limited to ISAC-based systems but is widely observed in neural networks trained sequentially on multiple tasks. This poses a major obstacle for 6G foundation models, which are expected to generalize across diverse tasks while maintaining consistent performance on the previous tasks.

Classical Fine-tuning: Fine-tuning refers to training a pre-trained model on a new task to leverage the knowledge it has learned from the original task. Usually, some parameters or the entire model is retrained. However, this can lead to catastrophic forgetting, where the model loses knowledge of the original task while adapting to the new one.

To validate this effect, we first train a Temporal Convolutional Network (TCN) on an initial localization task using the mmW-Loc dataset. We use the TCN model from our previous work [24]. This pretrained model then serves as a baseline. We then fine-tune the same TCN model (all parameters) on a new person identification task using the GaitID dataset. Then, we re-evaluate performance on Task 1 (inference) after

TABLE 2. Results across datasets: 5GmmGesture, DISAC-mmVRPose, mmW-Loc and mmW-GaitID.

Architecture	5GmmGesture (Gesture Recognition) Accuracy (%)	Trainable Params (M)	Size (MB)	GPU VRAM (MB)
ResNet18	97.75	11.2	42.7	213.6
ResNet34	96.15	21.3	81.2	487
Architecture	**DISAC-mmVRPose (Pose Estimation) MPJPE (cm)**	**Trainable Params (M)**	**Size (MB)**	**GPU VRAM (MB)**
ResNet18	6.2	11.2	42.7	213.6
ResNet34	6.0	21.3	81.2	487
Architecture	**mmW-Loc (Task 1) and mmW-GaitID (Task 2) Accuracy (%)**	**Trainable Params**	**Size (MB)**	**GPU VRAM (MB)**
Standard Fine-Tuning (TCN Model)				
TCN (Task 1)	80.50	140K	0.53	3.2
TCN (Task 2)	79.17	140K	0.53	3.2
TCN (Task 1 after Task 2)	9.15	–	–	–
Standard Fine-Tuning (Transformer Model)				
Poolformer-s24 (Task 1)	64	21M	80.1	480
Poolformer-s24 (Task 2)	66.29	21M	80.1	480
Poolformer-s24 (Task 1 after Task 2)	10.09	–	–	–
LoRA Fine-Tuning (Classification Head Only: TCN Model)				
TCN (Task 1)	81.88	140K	0.53	3.2
TCN (Task 2)	73.80	2.2K	0.01	0.0050
TCN (Task 1 after Task 2)	81.88	–	–	–
LoRA Fine-Tuning (Adapters Across Network, Transformer Model)				
Poolformer-s24 (Task 1)	63.07	21M	80.1	480
Poolformer-s24 (Task 2)	67	92K	0.35	2.1
Poolformer-s24 (Task 1 after Task 2)	63.07	–	–	–

TABLE 3. Leave-one-user-out evaluation on the 5GmmGesture dataset. Each row shows accuracy when training on 7 users and testing on the held-out user.

Test User	Accuracy (%)
1	57.44
2	66.00
3	85.70
4	74.70
5	70.00
6	77.08
7	68.30
8	75.00
Average	**71.78**

training on Task 2 to measure the performance retention of the model. Table 2 shows the results of the model on the two tasks. While the model performs decently on the two tasks, fine-tuning the model makes it lose performance on the original task. We see that after fine-tuning, the performance of the model on the original task drops to 9.15% for TCN and 10.09% for the transformer-based Poolformer-s24 model.

LoRA-based fine-tuning: To tackle this, we adopted Low-Rank Adaptation (LoRA) [25]. LoRA is a parameter-efficient fine-tuning method that adapts pre-trained models by introducing small low-rank matrices to the frozen weights. By decomposing weight updates into two compact matrices, it greatly reduces the number of trainable parameters, enabling faster, more memory-efficient, and computationally cheaper fine-tuning compared to full model adaptation.

We tackle the challenge of maintaining performance on Task 1 while adapting to Task 2 using two variants of LoRA-based fine-tuning. In the first approach, we freeze the entire pre-trained TCN and attach a lightweight low-rank adapter only to the classification head. During training, only the adapter parameters are updated, while the original network weights remain fixed. This strategy greatly reduces computational overhead and fully preserves Task 1 performance. However, since the majority of the network remains frozen and the adapter has limited capacity, the model struggles to fully adapt to Task 2, leading to a drop in accuracy on the new task (from 79.17% to 73.80%).

In the second approach, we adopt the PoolFormer transformer architecture for this setup, as the PEFT[9] library used for LoRA fine-tuning requires standard architectures and cannot be directly applied to custom models. In this approach, the pre-trained PoolFormer backbone is kept frozen, and adapter layers are inserted into every convolutional block throughout the network. By distributing trainable parameters across multiple layers, the model gains greater representational capacity, enabling it to better capture features relevant to Task 2. This method improves Task 2 accuracy compared to adding adapters only to the classification head, while still preserving Task 1 performance and avoiding catastrophic forgetting.

From Table 2, we observe that LoRA significantly reduces both the number of trainable parameters and GPU memory requirements. For example, in the TCN model, the trainable parameters drop from 140K to only 2.2K, while the GPU VRAM usage decreases from 3.2 MB to just 0.005 MB, representing a 640× reduction. Similarly, for the Poolformer-s24, LoRA decreases trainable parameters from 21M to 92K

[9]https://huggingface.co/docs/peft/index

and GPU VRAM from 480 MB to 2.1 MB, achieving a 228× reduction. These drastic reductions are particularly important for training on mobile and edge devices, which have limited memory and computational resources. As models continue to scale toward 6G ISAC foundation models, potentially containing billions of parameters, full fine-tuning becomes computationally infeasible. In contrast, low-rank adapters make it possible to adapt large pre-trained models to new tasks efficiently, enabling real-time on-device training and rapid deployment in resource-constrained environments.

VI. LINK BETWEEN COMMUNICATION AND SENSING

In this work, we focused on the *sensing* aspect of the ISAC. However, our setups do not modify the existing protocol stack. Instead, sensing information is extracted without impacting the communication performance. For example, in *mmWGesture*, beam SNR is extracted from sector sweep, a part of the IEEE 802.11ad standard. Sector sweep is part of the periodic beacon interval. As a result, sensing information is extracted without any signaling overhead or impacting communication performance. Similarly, for *mmW-Pose, mmW-GaitID, mmW-Loc*, periodic pilot signals are sent at a lower rate to extract mmWave CSI. In *5GmmGesture* and *DISAC-mmVRPose*, the sampling frequency of signal features (PPBP and CIR) can be varied to analyze the trade-off between communication and sensing performance. In our recent work [26], we showed that a sensing airtime of 25% only reduces classification accuracy by 0.15 percentage points from full-time sensing on *5GmmGesture* dataset.

VII. NOVEL OPPORTUNITIES AND CHALLENGES

The main goal of this work is to show the validity of the datasets for mmWave ISAC-based human sensing. Future work may focus on developing new architectures, improving signal processing techniques, and exploring advanced learning methods, which remain open challenges for the research community. These datasets enable researchers to explore and address important challenges such as:

A. DOMAIN ADAPTATION

The results in Section V showed a significant reduction in accuracy on unseen users. Researchers can advance domain adaptation techniques using our datasets. Conventional domain adaptation methods, such as ADDA [27], often fail in this scenario, as they rely on the assumptions that (i) the source and target domains share a common feature space and (ii) the conditional distribution of classes remains invariant across domains. ADDA and other domain adaptation methods mostly align feature distributions across domains irrespective of class (i.e., matching marginal distributions). These methods do not guarantee that features of class i in the source domain map close to features of the same class i in the target domain. Without some mechanism to preserve which class a feature belongs to (or enforce class-conditional alignment), target features may drift toward the wrong source classes. Formally, given a source domain $\mathcal{D}_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$ and a target domain $\mathcal{D}_t = \{x_j^t\}_{j=1}^{N_t}$, where $x \in \mathcal{X}$ represents the input features and $y \in \mathcal{Y}$ denotes the corresponding gesture labels, these methods assume:

$$P_s(Y \mid X) \approx P_t(Y \mid X), \quad \forall X \in \mathcal{X}, \tag{2}$$

where P_s and P_t denote the conditional class distributions in the source and target domains, respectively. However, for ISAC-based gesture recognition, this assumption often does not hold due to significant variations across users and environments, leading to a large domain gap. Therefore, such adaptation techniques may result in mismatches; for example, source features of one gesture may be mapped to target features of another gesture. This dataset provides a valuable benchmark for developing more robust domain adaptation techniques adapted to gesture recognition in ISAC systems.

B. SEMI-SUPERVISED LEARNING (SSL) AND FOUNDATION LEARNING

Collecting labeled data at mmWave frequencies is challenging due to limited research hardware and complex data collection processes. Moreover, existing unsupervised domain adaptation techniques often fail to transfer directly to ISAC systems. Our datasets provide a foundation for advancing SSL research in mmWave ISAC. In our recent work on the 5GmmGesture dataset [28], we show that SSL-based methods can improve performance up to 25 percentage points compared to a supervised baseline when using only 0.6% of labeled data. Furthermore, these datasets open opportunities for unsupervised and self-supervised approaches, ultimately contributing to the development of mmWave foundation models [9]. For example, mmW-Loc, mmWPose, and mmW-GaitID have been collected using the same COTS hardware.

C. SENSING AND COMMUNICATION TRADE-OFF

Our 5GmmGesture dataset could be used to analyze the trade-off between gesture recognition (sensing) and communication performance, specifically examining how the number of Tx and Rx beams, both in space and time, impacts the sensing performance and communication overhead in ISAC [26].

D. DISTRIBUTED INFERENCE

Our dataset can also support research on split inference, where neural networks are distributed across multiple devices to reduce sensing data transfer bottlenecks and enable efficient, distributed multi-static ISAC [29]. Early results from our previous work [30] on the DISAC-mmVRPose dataset indicate that split inference can reduce communication overhead by three orders of magnitude at the cost of only 10% loss of accuracy.

E. IMPACT OF SENSING GEOMETRY

DISAC-mmVRPose dataset can also be used to investigate the impact of sensing geometry for pose estimation. The dataset allows evaluation of how pose estimation accuracy evolves as the number of Rxs is reduced or certain viewpoints are removed. This can be useful for practical mmWave ISAC deployments in determining the minimum number and placement of viewpoints (Txs/Rxs) required for reliable pose estimation in VR environments.

F. COMPARISON OF mmWave AND SUB-6 GHz ISAC

mmW-GaitID and mmW-Loc allow systematic comparison between 5 GHz and 60 GHz CSI for person identification and localization. Fusion techniques combining the mmWave and 5 GHz CSI for improved ISAC can also be explored using the same datasets. Initial results from our previous work [24] suggest mmWave Wi-Fi outperforms sub-6 GHz Wi-Fi for person identification.

VIII. CONCLUSION

In this paper, we introduced *mmHSense*, a set of mmWave ISAC datasets designed to support a wide range of downstream tasks. These datasets capture diverse signal characteristics and include cross-modal settings. We described the testbed, experimental setup, and implementation details to ensure that our experiments are easily reproducible.

Furthermore, we demonstrated the utility of these datasets by evaluating their performance on several representative downstream tasks and highlighted key challenges encountered in real-world deployment. We believe this work serves as a valuable resource for the research community and invite researchers to leverage these datasets to advance deep learning techniques and address the open challenges in mmWave ISAC.

REFERENCES

[1] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[2] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Netw.*, vol. 35, no. 5, pp. 158–167, Sep. 2021.

[3] Z. Gao, Z. Wan, D. Zheng, S. Tan, C. Masouros, D. W. K. Ng, and S. Chen, "Integrated sensing and communication with mmWave massive MIMO: A compressed sampling perspective," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1745–1762, Mar. 2023.

[4] J. Pegoraro, P. Saucedo, J. O. Lacruz, M. Rossi, and J. Widmer, "DISC: A dataset for integrated sensing and communications in mmWave systems," *IEEE Commun. Mag.*, vol. 63, no. 10, pp. 94–100, Oct. 2025.

[5] A. D. Singh, S. S. Sandha, L. Garcia, and M. Srivastava, "RadHAR: Human activity recognition from point clouds generated through a millimeter-wave radar," in *Proc. 3rd ACM Workshop Millimeter-wave Netw. Sens. Syst.*, Oct. 2019, pp. 51–56.

[6] D. Zhang, D. Wu, K. Niu, X. Wang, F. Zhang, J. Yao, D. Jiang, and F. Qin, "Practical issues and challenges in CSI-based integrated sensing and communication," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jul. 2022, pp. 836–841.

[7] J. Yang, H. Huang, Y. Zhou, X. Chen, Y. Xu, S. Yuan, H. Zou, C. X. Lu, and L. Xie, "MM-fi: Multi-modal non-intrusive 4D human dataset for versatile wireless sensing," in *Proc. Adv. Neural Inf. Process. Syst.*, 2023, pp. 18756–18768.

[8] J. Yang, X. Chen, H. Zou, C. X. Lu, D. Wang, S. Sun, and L. Xie, "SenseFi: A library and benchmark on deep-learning-empowered WiFi human sensing," *Patterns*, vol. 4, no. 3, Mar. 2023, Art. no. 100703.

[9] J. Du, T. Lin, C. Jiang, Q. Yang, C. F. Bader, and Z. Han, "Distributed foundation models for multi-modal learning in 6G wireless networks," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 20–30, Jun. 2024.

[10] J. Yu, P. Wang, T. Koike-Akino, Y. Wang, P. V. Orlik, and H. Sun, "Human pose and seat occupancy classification with commercial mmWave Wi-Fi," in *Proc. IEEE Globecom Workshops (GC Wkshps.)*, Mar. 2020, pp. 1–6.

[11] A. Ali, P. Parida, V. Va, S. Ni, K. N. Nguyen, B. L. Ng, and J. C. Zhang, "End-to-end dynamic gesture recognition using mmWave radar," *IEEE Access*, vol. 10, pp. 88692–88706, 2022.

[12] N. N. Bhat, R. Berkvens, and J. Famaey, "Gesture recognition with mmWave Wi-Fi access points: Lessons learned," in *Proc. IEEE 24th Int. Symp. World Wireless*, Aug. 2023, pp. 127–136.

[13] J. O. Lacruz, R. R. Ortiz, and J. Widmer, "A real-time experimentation platform for sub-6 GHz and millimeter-wave MIMO systems," in *Proc. 19th Annu. Int. Conf. Mobile Syst.*, 2021, pp. 427–439, doi: 10.1145/3458864.3466868.

[14] L. Crockett, D. Northcote, and R. W. Stewart, *Software Defined Radio With Zynq UltraScale+ RFSoC*. Glasgow, U.K.: Strathclyde Academic, 2023. [Online]. Available: https://www.RFSoCbook.com

[15] Y.-S. Huang, W.-C. Liu, and S. Jou, "Design and implementation of synchronization detection for IEEE 802.15.3c," in *Proc. Int. Symp. VLSI Design*, Jun. 2011, pp. 1–4.

[16] Y. Ghasempour, C. R. C. M. da Silva, C. Cordeiro, and E. W. Knightly, "IEEE 802.11ay: Next-generation 60 GHz communication for 100 Gb/s Wi-Fi," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 186–192, Dec. 2017.

[17] J. Pegoraro, J. O. Lacruz, T. Azzino, M. Mezzavilla, M. Rossi, J. Widmer, and S. Rangan, "JUMP: Joint communication and sensing with unsynchronized transceivers made practical," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9759–9775, Aug. 2024.

[18] N. N. Bhat, J. Sameri, J. Struye, M. T. Vega, R. Berkvens, and J. Famaey, "Multi-modal pose estimation in XR applications leveraging integrated sensing and communication," in *Proc. 1st ACM Workshop Mobile Immersive Comput., Netw., Syst.*, 2023, pp. 261–267.

[19] B. G. A. Gerats, J. M. Wolterink, and I. A. M. J. Broeders, "3D human pose estimation in multi-view operating room videos using differentiable camera projections," *Comput. Methods Biomechanics Biomed. Eng., Imag. Visualizat.*, vol. 11, no. 4, pp. 1197–1205, 2022.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[21] J. Zhang, Z. Tu, J. Yang, Y. Chen, and J. Yuan, "MixSTE: Seq2seq mixed spatio-temporal encoder for 3D human pose estimation in video," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2022, pp. 13222–13232.

[22] J. Kirkpatrick, R. Pascanu, N. C. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwińska, D. Hassabis, C. Clopath, D. Kumaran, and R. Hadsell, "Overcoming catastrophic forgetting in neural networks," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 13, pp. 3521–3526, 2017.

[23] R. Akbarian Bafghi, C. Bagwell, A. Ravichandran, A. Shrivastava, and M. Raissi, "Fine tuning without catastrophic forgetting via selective low rank adaptation," 2025, *arXiv:2501.15377*.

[24] N. N. Bhat, M. Karnaukh, J. Struye, R. Berkvens, and J. Famaey, "Beyond sub-6 GHz: Leveraging mmWave Wi-Fi for gait-based person identification," in *Proc. IEEE 23rd Consum. Commun. Netw. Conf. (CCNC)*, Feb. 2025, pp. 1–6.

[25] J. E. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, W. Lu, and C. Weizhu, "LoRA: Low-rank adaptation of large language models," in *Proc. Int. Conf. Learn. Represent.*, 2022. [Online]. Available: https://openreview.net/forum?id=nZeVKeeFYf9

[26] J. Struye, N. N. Bhat, S. Kumar, M. H. Moghaddam, and J. Famaey, "Millimeter-wave gesture recognition in ISAC: Does reducing sensing airtime hamper accuracy?" in *Proc. IEEE 23rd Consum. Commun. Netw. Conf. (CCNC)*, Feb. 2026, pp. 1–6.

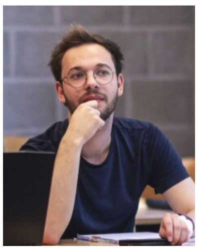
[27] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2017, pp. 2962–2971.

[28] N. N. Bhat, S. Kumar, M. H. Moghaddam, J. Struye, J. O. Lacruz, J. Pegoraro, J. Widmer, R. Berkvens, and J. Famaey, "MmGAN: Semi-supervised GAN for improved gesture recognition in mmWave ISAC systems," *IEEE Open J. Commun. Soc.*, vol. 7, pp. 95–117, 2025.

[29] Z. Zhuang, D. Wen, Y. Shi, G. Zhu, S. Wu, and D. Niyato, "Integrated sensing-communication-computation for over-the-air edge AI inference," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 3205–3220, Apr. 2024.
[30] W. Lemoine, N. N. Bhat, J. Struye, A. Belogaev, J. O. Lacruz, J. Widmer, and J. Famaey, "WIP: Distributed inference for human pose estimation using mmWave Wi-Fi," in *Proc. IEEE 26th Int. Symp. World Wireless*, Jun. 2025, pp. 141–144.

NABEEL NISAR BHAT received the M.Sc. degree in communications and computer networks engineering from Politecnico di Torino, in 2021. He is currently a Ph.D. Researcher in joint communication and sensing with the IDLab Research Group, University of Antwerp, and imec Research Institute, Belgium. His current research focuses on leveraging mmWave communication signals for pose estimation in extended reality applications. He has experience in signal processing, wireless communications, and deep learning.

MAKSIM KARNAUKH received the bachelor's degree in computer science and the master's degree in data science and artificial intelligence, in 2023 and 2025, respectively. He is currently working in the industry as a Data Scientist, where his work focuses primarily on natural language processing. His research interests include deep learning and sequence modeling.

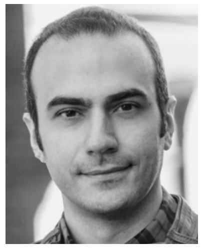


STEIN VANDENBROEKE is currently pursuing the master's degree in data science and artificial intelligence with the University of Antwerp, Belgium. He is a Student Assistant with the Bachelor of Computer Science program, where he teaches courses in computer architecture. His current research interests include computer vision, specifically exploring advancements in weakly supervised image object segmentation.

WOUTER LEMOINE received the M.Sc. degree in computer science from Hasselt University, in 2022. He is currently a Ph.D. Researcher in joint communication and sensing with the IDLab Research Group, University of Antwerp, and imec Research Institute, Belgium. His current research focuses on arbitrary transmitter localization using reconfigurable intelligent surfaces.

JAKOB STRUYE received the Ph.D. degree in computer science from the University of Antwerp, in 2024. He is currently a Postdoctoral Researcher in wireless networking with the IDLab Research Group, University of Antwerp, and imec Research Institute, Belgium. His current research focuses on mmWave ISAC for XR applications. He has experience in several deep learning subdomains, including reinforcement learning and time series prediction.

SIDDHARTHA KUMAR received the M.Sc. degree in communications engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 2015, and the Ph.D. degree in informatics from the University of Bergen, Bergen, Norway, in 2018. He is currently a Senior Systems Engineer with Qamcom Research and Technology AB, Sweden. His research interests are geared towards wireless communications, particularly on the subject of joint communications and sensing, information-theoretic privacy and security in distributed storage and computing.



MOHAMMAD HOSSEIN MOGHADDAM received the Ph.D. degree in electrical engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 2023. He is currently a Senior Systems Engineer with Qamcom Research and Technology AB, Gothenburg, where he has been contributing to joint communication and sensing research, since 2022. He has been actively involved in experimentation and prototyping of 5G/6G integrated sensing and communication testbeds. His research interests include statistical signal processing and its applications in wireless communication, autonomous driving (AD), and advanced driver assistance systems (ADAS).

JESUS OMAR LACRUZ received the bachelor's degree in electrical engineering from Universidad de Los Andes, Venezuela, in 2009, and the Ph.D. degree in electronic engineering from Universidad Politécnica de Valencia, Spain, in 2016. He is currently a Senior Researcher with IMDEA Networks Institute, Spain, where he has been contributing to cutting-edge wireless systems research, since 2017. His research interests include the design and implementation of signal processing algorithms and high-speed digital communication systems on FPGA/SoC platforms, encompassing digital communications, millimeter-wave systems, and joint communication and sensing technologies. He has been actively involved in experimentation and prototyping for 6G wireless networks, integrated sensing and communication systems, and advanced FPGA-based testbeds.

JOERG WIDMER is currently a Research Professor and the Research Director of IMDEA Networks, Madrid, Spain. Before, he held positions at DOCOMO Euro-Labs, Munich, Germany, and EPFL, Switzerland. He authored more than 200 conference and journal papers and three IETF RFCs, and holds 14 patents. His research focuses on wireless networks, ranging from extremely high frequency millimeter-wave communication and MAC layer design to mobile network architectures. He was awarded an ERC Consolidator Grant, the Friedrich Wilhelm Bessel Research Award of the Alexander von Humboldt Foundation, a Mercator Fellowship of the German Research Foundation, a Spanish Ramon y Cajal Grant, and 16 best paper awards. He is a Distinguished Member of ACM.

RAFAEL BERKVENS received the master's degree in applied engineering: electronics-ICT and the Ph.D. degree in indoor location information quantification from the University of Antwerp, Belgium, in 2012 and 2017, respectively. Currently, he is an Assistant Professor with IDLab–imec, Department of Electronic, Information, and Communication Technology, Faculty of Applied Engineering, University of Antwerp. His research interests include perception through ambient wireless communication, integrated sensing and communication, and intelligent transportation systems. He chairs the subject group of communication, teaches several networking courses, and coordinates part of the master's project. He is also a member of the IEEE Communications Society.

JEROEN FAMAEY is currently an Associate Research Professor with the Department of Computer Science, University of Antwerp, and imec, Belgium. He has published more than 170 peer-reviewed journal articles and conference papers and has eight granted patents. He was listed as one of the top 2% most cited scientists worldwide by Stanford University both, in 2022 and 2023. His research interests include performance modeling and optimization of wireless network protocols.