# A Paradigm Shift From an Experimental-Based to a Simulation-Based Framework Using Motion-Capture Driven MIMO Radar Data Synthesis

Sahil Waqar , Muhammad Muaaz , Stephan Sigg , and Matthias Pätzold

Abstract—The development of radar-based classifiers driven by empirical data can be highly demanding and expensive due to the unavailability of radar data. In this article, we introduce an innovative simulation-based approach that addresses the data scarcity problem, particularly for our multiple-input multiple-output (MIMO) radar-based direction-independent human activity recognition (HAR) system. To simulate realistic MIMO radar signatures, we first synthesize human motion and generate corresponding spatial trajectories. From these trajectories, a received radio frequency (RF) signal is synthesized using our MIMO channel model, which considers the non-stationary behavior of human motion and the multipath components originating from the scatterers on human body segments. Subsequently, the synthesized RF signals are processed to simulate MIMO radar signatures for various human activities. The proposed simulation-based direction-independent HAR system achieves a classification accuracy of 97.83% when tested with real MIMO radar data. A significant advantage of our simulation-based framework lies in its ability to facilitate multistage data augmentation techniques at the motion-layer, physical-layer, and signal-layer syntheses. This capability significantly reduces the training workload for radar-based classifiers. Importantly, our simulation-based proof-of-concept is applicable to single-input single-output (SISO) and MIMO radars in monostatic, bistatic, and multistatic configurations, making it a versatile solution for realizing other radar-based classifiers, such as gesture classifiers.

Index Terms—Aspect angle, data augmentation, data synthesis, deep learning, distributed multiple-input multiple-output (MIMO) radar simulation, human activity recognition (HAR), micro-Doppler analysis, motion capture (MoCap), motion synthesis, multiclass classification, virtual reality.
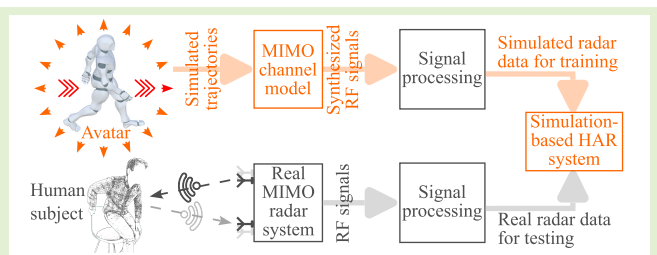
I. INTRODUCTION

A. Background

THE generation of area-specific synthetic data has been an important topic of interest among researchers [1], [2]. Device-specific or sensor-tailored simulation models help generate realistic sensory data and have been used to realize real-world solutions [3], [4]. Given the increasing prevalence of machine learning and artificial intelligence methodologies and applications today, the importance of the concept of device-specific synthetic data generation, as well as the significance of sensor modeling, cannot be overstated. For many sensing modalities such as magnetometer, infrared, light detection and ranging (LiDAR), sonar, and radar, data scarcity often hinders the realization of machine learning-based solutions [5], [6]. Sensor-tailored simulation models mitigate the data scarcity problem by providing clean and labeled synthetic datasets for various real-world conditions. Such synthetic datasets are important to develop machine learning-based applications, e.g., medical imaging [7].

Human activity recognition (HAR) [8], [9], [10], [11], [12], [13] remains an important and active research area facing the challenge of data scarcity, especially when using radio frequency (RF) sensors such as Wi-Fi [14] and radar [15], [16], [17]. Furthermore, for multiple-input multiple-output (MIMO) radar systems with user-defined (required) operating parameters and antenna configurations, readily available HAR

datasets are almost non-existent. Optimal radar operating conditions and antenna configurations are often not known in advance for different environmental conditions and applications. Synthetic data generation is therefore a pragmatic and promising approach to realizing radar-based classifiers, offering tremendous design control and system flexibility in a cost-effective manner. Realizing HAR systems through a simulation-based approach poses two main challenges: 1) how to synthesize human activities and 2) how to simulate single-input single-output (SISO) and MIMO radar signatures for the synthesized human activities. Before going into further details of synthetic data generation and our proposed simulation-based approach, we first provide an overview of the relevant research in Section I-B.

B. Related Work

The ongoing miniaturization and commercialization of radar sensors, as well as many Internet of Things (IoT) sensors, have encouraged the development of human-centric applications, including HAR. Small-scale radar systems are increasingly preferred by researchers for the development of HAR systems [18], [19], gesture [20], [21], and sign language [22] recognition systems. Realizing empirical-data-driven (experimental-based) HAR systems is often very challenging due to the low availability of recorded radar datasets. Among other challenging and monotonous tasks, the development of experimental-based HAR systems requires the involvement of human subjects, an actual SISO or MIMO radar system, and the manual labeling of the recorded data. Yu et al. [23] used manually labeled point cloud data to train the HAR system, which was built upon a long short-term memory (LSTM) network. By utilizing the measured features of a millimeter wave (mm-wave) radar, Zhao et al. [24] tackled the issue of HAR in multiview settings.

Recent studies have shown that, to some extent, data augmentation techniques can reduce the scarcity of empirical data for HAR systems. For instance, a rotation-shift technique was utilized in [25] to expand the 3-D point cloud dataset. A generative adversarial network (GAN)-based data augmentation technique was adopted in [26] to create varied radar signatures of human activities. The use of the few-shot learning method was suggested by Liu et al. [27], which offers a unique way of augmenting the capabilities of pre-trained and pre-existing HAR systems. According to a recent study [28], a two-stage domain adaptation approach can be used to alleviate the data scarcity issue as well. With this approach, the simulated micro-Doppler signatures can be translated into measurement-like micro-Doppler signatures by using small real datasets. Note that even with such data augmentation methods, time-consuming and tedious data collection cannot be avoided.

Radar-based classifiers may face unique challenges in different situations and application areas, which may necessitate the adaptation of radar antenna configurations and operating conditions. This exacerbates the problem of data scarcity in radar systems because the training dataset recorded from a radar system in one scenario may not be applicable and useful in another. Therefore, the synthetic data generation is the way to realize radar-based HAR systems. To date, only a few studies have been conducted in the context of RF sensing that deal with synthetic data generation for HAR. In this regard, the utilization of motion capture (MoCap) systems [29] is an effective means of modeling and reanimating complex human motion for further motion synthesis. For passive Wi-Fi radar (PWR), Vishwakarma et al. [30] devised a system, namely SimHumalator, to generate target returns. In [31], a simulation tool was created to evaluate the radar cross section of a walking individual in close proximity. However, this technique is inadequate for reproducing detailed and complex human movements.

C. Our Approach

In this article, we present a proof-of-concept that overcomes the problems related to radar data scarcity, offers significant design control and flexibility of the radar system, and allows the simulation of unbounded, clean, and labeled radar datasets. We emulate a 2×2 MIMO radar system with the help of our proposed simulation-based framework to realize a simulation-based direction-independent HAR system. First, we devise an activity simulation module that synthesizes multiple types of human activities in a virtual environment by using the 3-D animation tools from the Unity [32] and MotionBuilder [33] software. An appropriate avatar or a humanoid character, equipped with multiple simulated point scatterers on its body segments, is used to reanimate MoCap data in these programs (see Section III). Subsequently, we generate spatial trajectories corresponding to all simulated point scatterers or body segments of the avatar, which effectively characterize the overall humanoid motion.

The spatial trajectories of the body segments are processed by our channel model, which simulates the received RF signal from a frequency-modulated continuous wave (FMCW) radar system for software-defined antenna positions. While simulating the raw in-phase and quadrature (IQ) components of a received baseband signal, our channel model takes into account the multipath components originating from the non-stationary simulated (real) point scatterers with distinct time-variant (TV) propagation delays (see Section IV). In the proposed channel model, the long- and short-time stationarity characteristics of the scatterers are considered in an indoor wireless propagation environment. Additionally, to train the 2×2 MIMO radar-based direction-independent HAR system (see Section VII), we simulated five types of multidirectional human activities by rotating the transmitter and receiver antennas of the emulated MIMO radar system (see Section V).

Unlike conventional or experimental-based designs of HAR systems, the proposed simulation-based approach is highly versatile and offers numerous advantages. Our simulation-based approach is capable of simulating diverse training datasets to meet various radar-based applications and a wide range of operational requirements. For monostatic/bistatic/multistatic SISO/MIMO radar systems, the scatterer-level modeling of moving objects in our simulation-based framework opens up new research opportunities to further fine-tune the simulated radar signatures, such as TV micro-Doppler signatures (TV radial velocity distributions) and TV range distributions (see Sections IV and VI).

For example, the TV path gains of the scatterers (simulated point scatterers) can be adjusted or optimized to improve and augment the simulated radar signatures. Moreover, the simulation-based framework provides multistage data augmentation techniques (see Section V), which allow us to generate diverse and high-quality SISO/MIMO radar datasets in a flexible and cost-effective manner. For instance, at the motion-layer synthesis data augmentation stage, various animation parameters and avatar characteristics, e.g., speed and height, can be arbitrarily varied to simulate a range of human motions. Most importantly, the proposed simulation-based framework radically reduces the workload and resources for classifier training. As our simulation-based approach is versatile, it can be easily extended to implement many other SISO/MIMO radar-based classifiers, such as air-writing gesture classification [34].

D. Contributions

The key findings and contributions of this study can be delineated as follows.

1) This research proposes a simulation-based framework to significantly minimize the data collection workload required for devising real-world radar-based HAR systems. The simulation-based framework is capable of synthesizing realistic, diverse, and clean datasets for MIMO radar systems, regardless of their configuration: monostatic, bistatic, or multistatic. Although this study focuses on a 2×2 MIMO radar-based direction-independent HAR system, the utility of the simulation-based framework extends beyond the HAR application, making it also valuable for other radar-based applications, e.g., sign language detection.

2) We have developed a MoCap-data-driven activity simulation module that enables the synthesis of multiple types of human activities in a virtual environment. For a total of 21 simulated point scatterers placed on body segments of an avatar, the activity simulation module generates 3-D trajectories that essentially characterize the overall human motion. Our activity simulation module can integrate motion data from diverse sources, including biomechanical, wearable, and optical MoCap systems (see Section III). Additionally, the activity simulation module can generate arbitrary software-defined motion data.

3) We formulate a MIMO channel model that simulates realistic RF data or raw IQ data by using the spatial trajectory data from non-stationary simulated point scatterers. In this channel model, we study and simulate the TV propagation delays corresponding to the multipath components emanating from the non-stationary simulated point scatterers on the avatar's body segments. The proposed MIMO channel model helps generate unlimited radar datasets and provides extensive design control and versatility.

4) We present multistage data augmentation techniques for motion-layer synthesis, physical-layer synthesis, and signal-layer synthesis. For example, in the proposed simulation-based framework, we first diversified the target motion data in the motion-layer synthesis using the activity simulation module. And subsequently in the physical-layer synthesis, we augmented the radar data by varying physical-layer parameters such as radar orientation. Lastly, we further augmented the radar data by using multiple sets of simulated point scatterers' weights (TV path gains) at the signal-layer synthesis of the proposed simulation-based framework. The multistage, simulation-based data augmentation techniques allowed us to vary target motion characteristics and antenna configurations, simulate multiple radar sensors, and transform unidirectional motion data to multidirectional motion data.

5) For the simulated human activities, we generated high-quality MIMO radar signatures, such as TV radial velocity distribution and mean radial velocity, which closely resemble the actual MIMO radar signatures of actual human activities. This demonstrates the efficacy of the proposed simulation-based framework, which is highly versatile as it can be effortlessly extended to simulate radar signatures for various other moving objects such as cars, drones, and aircraft.

6) By employing our simulation-based framework, we generated a unique simulated dataset to train/realize classifiers based on (deep) machine learning. The training dataset for HAR incorporates simulated radar patterns, derived from software-defined avatar movements. This approach proves highly advantageous and practical as the training data is developed entirely from scratch, eliminating the need for real individuals and an actual MIMO radar system.

7) For the 2×2 MIMO radar framework, we realized a simulation-based HAR system by employing a deep convolutional neural network (DCNN). The system employed multiperspective simulated radar signatures as input features. To showcase the practical applicability of our simulation-driven HAR system, we evaluated its performance using actual mm-wave radar data collected from actual individuals. Our simulation-based multiperspective HAR system achieved an impressive classification accuracy of 97.83%, providing compelling evidence for its effectiveness.

E. Article Organization

The article is divided into eight sections. Section II deals with the system design and the general structures of the conventional and the proposed approaches. Human MoCap and synthesis techniques are presented in Section III. Section IV details channel modeling and simulation. Multistage data augmentation approaches are elucidated in Section V. Section VI discusses the generation of MIMO radar signatures. Section VII presents the design, training, and testing phases of our simulation-based direction-independent HAR system. Finally, we conclude our research in Section VIII.

II. SYSTEM DESIGN

In this section, we discuss a conventional experimental-based design of a HAR system and the proposed simulation-based realization of a HAR system. We also discuss

problems of conventional HAR systems and how the proposed end-to-end simulation framework resolves them. Note that SISO radar-based HAR systems struggle to classify multidirectional human activities [35], [36]. To classify different types of multidirectional human activities, we need multiple radar subsystems illuminating the environment from different perspectives. Therefore, in this article, we consider multidirectional human activities recorded by a multiperspective distributed MIMO radar system.

A. Conventional Experimental-Based Designs of HAR Systems

In radar sensing, state-of-the-art experimental-based HAR systems [9], [13], [14], [15], [17], [18], [19], [23], [24], [36] generally face challenges, such as data scarcity and their adaptability to environmental conditions. As an example of state-of-the-art experimental-based designs, we considered a direction-independent HAR system implemented with a mm-wave 2×2 MIMO radar system, as shown in Fig. 1(a). In Fig. 1(a), Radar$_i$ represents the ith radar subsystem of the distributed MIMO radar system, $A_i^{T_x}$ is the ith transmitter antenna, and $A_i^{R_x}$ is the ith receiver antenna for $i = 1, 2$. Note that the two horn antennas, namely $A_i^{T_x}$ and $A_i^{R_x}$, are arranged in a monostatic configuration for Radar$_i$. In the conventional experimental-based HAR system of Fig. 1(a), six human subjects performed the following types of multidirectional activities: falling on a mattress, walking, standing up from a chair, sitting down on a chair, and picking up an object from the floor.

The distributed MIMO radar system simultaneously illuminates the human subject from two aspect angles and generates the corresponding raw IQ data, as shown in Fig. 1(a). Then, the radar signal processing block (see Section VI) generates the TV micro-Doppler signatures or, equivalently, the TV radial velocity distributions for Radar$_1$ and Radar$_2$. These recorded radar signatures (TV radial velocity distributions) are accumulated to create a real radar dataset. In conventional experimental HAR systems, the real radar dataset is usually divided into a training subset and a testing subset to train and test these HAR systems, respectively. However, for this research, we only use the experimentally obtained radar dataset to test our proposed simulation-based HAR system (see Fig. 2).

Similar to any multiclass classifier, radar-based HAR systems require extensive amounts of recorded data for their training. However, unlike other sensing modalities such as cameras, radar systems often suffer from data scarcity. To experimentally design a HAR system, real human subjects must perform various types of activities in front of the MIMO radar system in multiple directions. These requirements make data collection time-consuming and costly. Additionally, the recorded radar training dataset usually cannot be reused for different antenna configurations and operating conditions. For instance, changing the position of a transmitter or a receiver antenna of the MIMO radar system can invalidate the entire recorded training dataset.

B. Simulation-Based Design of HAR Systems

In this article, we propose a feasible alternative to overcome the aforementioned limitations of radar-based classifiers, particularly with regard to the scarcity of radar data. To develop real-world HAR systems, we propose a comprehensive simulation-based framework that utilizes MoCap systems to synthesize realistic MIMO radar data, as depicted in Fig. 1(b). The objective is to generate a simulated MIMO radar-based training dataset by seamlessly simulating a large number of realistic MIMO radar signatures without real human subjects and a physical radar system.

The block diagram in Fig. 1(b) provides a general overview of the proposed end-to-end simulation framework for HAR systems. In Fig. 1(b), the activity simulation module synthesizes the five types of human activities in the 3-D space from motion data collected by the MoCap systems (see Section III). The activity simulation module simulates 3-D trajectories corresponding to different body segments of an avatar, e.g., head, neck, torso, and upper and lower extremities. To simulate the human activities in multiple directions as shown in Fig. 1(b), we rotate the positions of the transmitter antenna $A_i^{T_x}$ and receiver antenna $A_i^{R_x}$ in our simulation-based framework for $i = 1, 2$ (see Section V-B). For a desired antenna configuration of the MIMO radar system, our channel simulation module first transforms the 3-D trajectories into TV propagation delays. Then, the channel simulation module generates realistic RF or raw IQ data for the simulated TV propagation delays and a set of scatterer weights. Eventually, the radar signal processor arranges the simulated raw IQ data in the fast- and slow-time domain and processes it to simulate realistic radar signatures, i.e., range distribution, radial velocity distribution (micro-Doppler signature), and mean velocity (mean Doppler shift).

We synthesize numerous examples of the five types of human activities, simulate the corresponding radial velocity distributions (micro-Doppler signatures), and store them in our simulated radar dataset, as shown in Fig. 1(b). The proposed simulation-based framework has no limits on the generation of simulation data. The simulated radar dataset is used to train the simulation-based HAR system, which is based on a DCNN architecture. To demonstrate the practical importance and the generalizability of this simulation-based framework, we need to evaluate its performance in a real scenario. Therefore, the proposed simulation-based HAR system is evaluated on a previously unseen real radar dataset acquired with a mm-wave distributed MIMO radar system and real human subjects, as shown in the testing phase of Fig. 2. Note that we used an identical radar signal processing block in Fig. 1 because the simulated and real RF signals are structurally indistinguishable. More details on each block of the simulation-based HAR system are provided in the following sections.

III. HUMAN MoCap AND SYNTHESIS

This section explores several ways of capturing and synthesizing human motion. First, biomechanical modeling and its limitations will be briefly discussed. Second, wearable
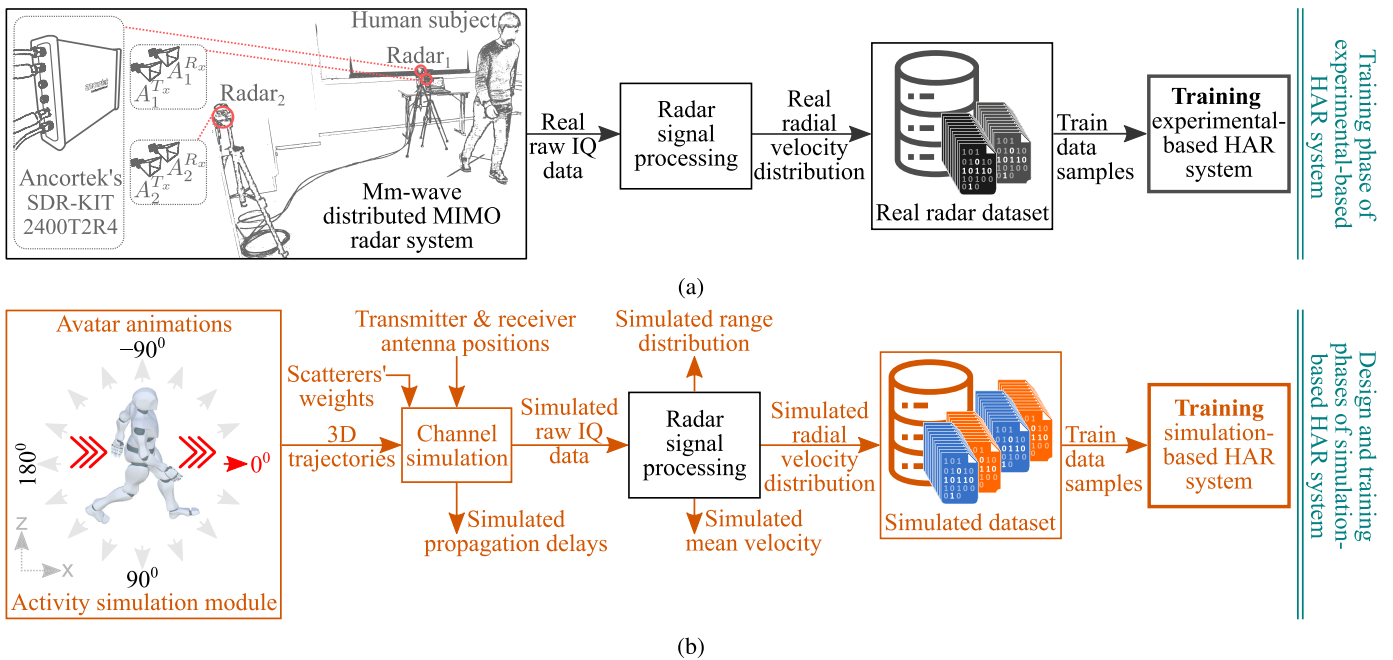
(a)

(b)

Fig. 1. (a) Design of conventional (experimental-based) direction-independent HAR systems that require human subjects and a MIMO radar system for their training. (b) Design of the proposed simulation-based HAR system that requires the simulated radar signatures for its training.
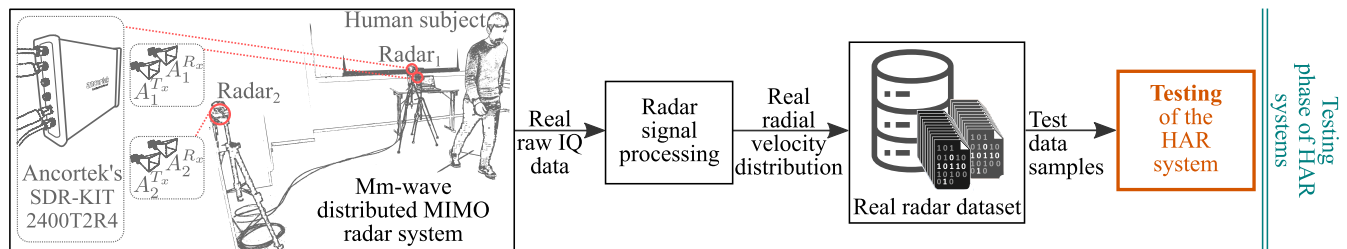

Fig. 2. Testing phase of both experimental and simulation-based direction-independent HAR systems. In the testing phase, the performance of the simulation-based HAR system is evaluated against unseen real radar signatures.

sensors as a means of MoCap systems are briefly mentioned. Third, we discuss optical MoCap systems such as Mixamo [37] and Qualisys [38]. It is important to highlight that the proposed simulation-based framework allows incorporating synthesized or recorded motion data from diverse sources such as biomechanical, wearable, and optical MoCap systems. Lastly, we explain the process of generating 3-D trajectories of human body segments using software such as Unity [32] and Autodesk's MotionBuilder [33]. These software programs (3-D animation tools) help us augment the motion data at the motion-layer synthesis.

A. Biomechanical Modeling of Human Body Segments

The utility of biomechanical modeling [39] for human body segments is undeniable, yet its complexity is inherently high, primarily due to the intricate nature of the human body. Also, it is difficult to develop generalizable biomechanical models because individuals differ in physiology, anatomy, and motor function. Moreover, the interaction between the human body and the environment can further increase the complexity of a biomechanical model.

Obtaining high-fidelity motion data of human body segments can be more feasible and accessible through MoCap repositories and systems such as Mixamo and Qualisys. In addition, the Unity and MotionBuilder software provide

a cost-effective and pragmatic alternative to biomechanical modeling, enabling the seamless and dynamic simulation of new motion data in a virtual environment. Therefore, we use MoCap systems to capture the human motion and employ 3-D animation tools from MotionBuilder and Unity to synthesize and subsequently augment human motion.

B. Wearable MoCap Systems

Wearable MoCap systems offer a versatile and cost-effective solution for capturing human movement data. The sensors, typically accelerometers and gyroscopes, are often integrated into garments to capture data on the orientation and acceleration of body segments. In this area, Rokoko Smartsuit Pro [40] is a viable choice with multiple inertial sensors for real-time tracking of an individual's skeletal movements. It facilitates seamless transfer of motion data to various applications such as sports, biomechanical analysis, and virtual reality. Compared to optical MoCap systems, wearable MoCap systems have limitations in terms of accuracy. Additionally, wearable MoCap systems can suffer from magnetic interference, which can affect the precision of the MoCap data.

C. Optical MoCap Systems

We used Mixamo and Qualisys optical MoCap systems to capture motion data for human activities. Mixamo is an

online platform that offers an extensive selection of readily available MoCap data captured from real performers [41]. Our Qualisys MoCap system was based on six Miqus M3 cameras connected in a daisy chain, capable of tracking passive reflective markers placed on a subject at 340 frames/s. The Qualisys MoCap system includes proprietary Qualisys track manager (QTM) software that provides an interface for tasks such as camera configuration and calibration, session setup and organization, marker-set definition, and MoCap measurements. Furthermore, QTM offers a suite of tools for marker labeling, data processing, analysis, and the export of MoCap data, thereby enabling seamless integration with third-party software. The camera system was calibrated according to the QTM guidelines to ensure accurate tracking of the markers and capturing their position and orientation in 3-D space. Next, 41 passive reflective markers were attached to a full body suit. The participant wore the suit, and we recorded a MoCap trial to generate an automatic identification of markers (AIMs) model. This model applies computer vision, localization, and motion estimation techniques to detect and track markers, facilitating an automated workflow for identifying and labeling markers. Once the AIM model was created, the skeleton solver function of QTM was used to calibrate the skeleton based on the marker positions. Next, a person's motion data was recorded for four activities: normal walking, standing up from a chair, sitting down onto a chair from a standing position, and picking up a small object from the floor. The recorded skeleton data was then exported in the Filmbox (FBX) file format and further processed in the MotionBuilder software. Note that for the falling activity, the MoCap data was relatively difficult to collect due to markers attached to the body. Therefore, we obtained MoCap data of the falling activity from Mixamo [37], a freely accessible online platform. In the next step, we import the acquired MoCap data into specialized software such as Unity or MotionBuilder, which are equipped with powerful tools that allow for the creation of comprehensive, meticulous, and lifelike 3-D animations.

D. 3-D Trajectories of Human Body Segments

By using the basic MoCap data and the 3-D animation tools, we synthesized, augmented, and visualized five human activities: falling on the floor, walking in an indoor environment, standing up from and sitting down on a chair, and picking up an object from the floor. Initially, the human activities were simulated and varied in a single direction or at an aspect angle of $0°$ with the help of 3-D animation tools, as shown in Fig. 1(b). However, we also needed to synthesize multidirectional human activities to realize a simulated MIMO radar-based direction-independent HAR system. Instead of using 3-D animation tools, we simulated multidirectional human activities more conveniently and efficiently by spatially rotating the transmitter and receiver antennas of the radar subsystem, Radar$_i$ (see Section V-B).

Following the synthesis of the human movement, we extract the spatial trajectories corresponding to each body segment of the avatar. To track the different body segments, 21 simulated point scatterers were placed on the avatar (see Fig. 3),
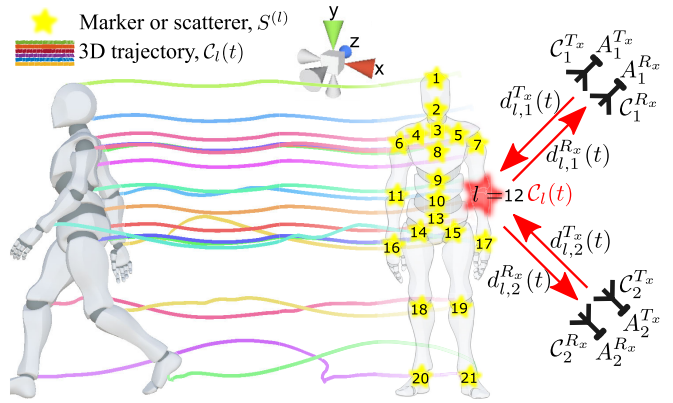

Fig. 3. Virtual 3-D propagation environment comprising a non-stationary avatar with 21 simulated point scatterers on its body segments and a simulated 2×2 multiperspective MIMO radar system.

these model the actual body scatterers that backscatter the transmitted RF signal to the receiver antennas of the 2×2 distributed MIMO radar system. We recorded the TV positions (trajectories) of the simulated point scatterers in the 3-D space for the simulated human activities. For example, the 3-D trajectories of the simulated point scatterers for a simulated walking activity are shown in Fig. 3.

At the outset, only 34 MoCap files were recorded, each representing one of the five distinct types of human activities. We visualized these activities using the Unity and Motion-Builder 3-D animation tools, and computed the corresponding 3-D trajectories. To expand the total number of synthesized human activities to 84, we applied data augmentation at the motion-layer synthesis using the Unity and MotionBuilder software (see Section V-A). Subsequently, we processed the 3-D trajectories in MATLAB for further data augmentation at the physical- and signal-layer syntheses. Although data augmentation at the motion-layer synthesis may require some attention to motion details, the physical-layer synthesis and signal-layer synthesis data augmentation stages in the proposed simulation-based framework are fairly automated. With the help of such multistage data augmentation techniques, we generated 2826 micro-Doppler signatures (TV radial velocity distributions) for each radar subsystem of the multiple-input multiple-output (MIMO) radar system. Section V provides more details on the multistage data augmentation techniques furnished by the proposed simulation-based framework.

IV. CHANNEL MODELING AND SIMULATION

In this section, we first present a geometrical 3-D indoor channel that models an indoor propagation scenario using the proposed simulation-based framework (see Fig. 3). Second, we investigate the multipath components caused by non-stationary simulated (real) point scatterers on avatar (human) body segments and simulate the corresponding TV propagation delays for a human activity. Lastly, we explain how the simulated propagation delays can be used to synthesize a received RF signal, specifically for an FMCW 2×2 MIMO radar system.

A. Geometrical Channel Model

We model and simulate a 3-D channel for an indoor environment, which consists of a 2×2 distributed MIMO radar system, a moving person, and stationary miscellaneous items such as furniture and electronics, as illustrated in Fig. 1(a). Recall that Radar$_i$ represents the ith radar subsystem of the distributed MIMO radar system, $A_i^{T_x}$ is the ith transmitter antenna, and $A_i^{R_x}$ is the ith receiver antenna for $i = 1, 2$. Let $[\cdot]^\top$ denote the vector transpose operation. Then, the position of the ith transmit (receive) antenna $A_i^{T_x}$ $(A_i^{R_x})$ of the 2×2 MIMO radar system is represented by $\mathcal{C}_i^{T_x} = [x_i^{T_x}, y_i^{T_x}, z_i^{T_x}]^\top$ $(\mathcal{C}_i^{R_x} = [x_i^{R_x}, y_i^{R_x}, z_i^{R_x}]^\top)$, as illustrated in Fig. 3.

A virtual propagation environment that resembles a real geometrical 3-D indoor channel is depicted in Fig. 3. In a real propagation environment, a moving human subject has countless non-stationary scatterers. For this research, we model these non-stationary bodily scatterers with $\mathcal{L} = 21$ non-stationary simulated point scatterers on a moving avatar, as shown in Fig. 3. Moreover, in Fig. 3, $\mathcal{C}_l(t) = [x_l(t), y_l(t), z_l(t)]^\top$ denotes the TV spatial trajectory of the lth marker $S^{(l)}$, $d_{l,i}^{T_x}(t)$ $(d_{l,i}^{R_x}(t))$ represents TV Euclidean distance between the lth marker $S^{(l)}$ and the ith transmit antenna $A_i^{T_x}$ (receive antenna $A_i^{R_x}$), where $i = 1, 2$ and $l = 1, 2, \ldots, \mathcal{L}$.

For the lth marker $S^{(l)}$ and the ith radar subsystem Radar$_i$, the TV radial distance $d_{l,i}(t)$ is equal to one-half of the overall propagation distance, i.e., $d_{l,i}(t) = (d_{l,i}^{T_x}(t) + d_{l,i}^{R_x}(t))/2$. Fig. 3 shows that the antenna configuration $\{\mathcal{C}_i^{T_x}, \mathcal{C}_i^{R_x}\}$ of the ith radar subsystem, Radar$_i$, follows a monostatic configuration, where $\mathcal{C}_i^{T_x} = \mathcal{C}_i^{R_x}$ for $i = 1, 2$. This leads to the following simplification: $d_{l,i}(t) = d_{l,i}^{T_x}(t) = d_{l,i}^{R_x}(t)$. The obtained TV radial distances $d_{l,i}(t)$ of the $\mathcal{L}$ non-stationary simulated point scatterers play an important role in simulating the TV propagation delays $\tau_i^{(l)}(t)$, as explained in Section IV-B.

B. Modeling of Multipath Components Caused by Human Body Segments

RF signals generally experience multipath propagation, particularly in indoor environments with numerous stationary and non-stationary reflective objects. In Fig. 3, the transmitted RF signal takes on multiple propagation paths, traveling from the transmitter antenna to the receiver antenna via multiple real (simulated) point scatterers on the human (avatar) body segments. Recall that in our simulation-based framework, the 21 simulated point scatterers on the avatar's body segments basically model the actual bodily scatterers that scatter the transmitted RF signal back to the receiver antennas of the 2×2 distributed MIMO radar system. For this study, by virtue of the cross-channel interference mitigation technique [16], we assume that the two radar subsystems, Radar$_1$ and Radar$_2$, of Ancortek's mm-wave radar system do not interfere with each other.

In the proposed simulation-based framework, we only consider multipath components originating from the $\mathcal{L} = 21$ non-stationary dominant and non-dominant scatterers located on various body segments of the avatar, as shown in Fig. 3. The multipath components originating from stationary dominant scatterers, such as walls, furniture, and floor, are excluded from the analysis because they are easily filtered out through signal preprocessing. Moreover, the bistatic components of the 2×2 distributed MIMO radar systems are not considered for this study. However, if required, the bistatic components of the 2×2 distributed MIMO radar system can be easily simulated in the proposed simulation-based framework.

The receiver antennas receive the multipath components, or multiple copies of the transmitted RF signal, with distinct TV propagation delays $\tau_i^{(l)}(t)$. For Radar$_i$, the lth TV propagation delay $\tau_i^{(l)}(t)$ is related to the lth TV radial distances $d_{l,i}(t)$ according to the relation $\tau_i^{(l)}(t) = 2d_{l,i}(t)/c_0$, where c_0 is the speed of light. Within the framework of radar sensing, the synthesized motion is completely characterized by the simulated TV propagation delays $\tau_i^{(l)}(t)$, as explained in Section IV-C.

For the five distinct types of simulated human activities and Radar$_1$, Fig. 4 shows the simulated TV propagation delays $\tau_i^{(l)}(t)$ of the $\mathcal{L} = 21$ simulated point scatterers. The lth TV propagation delay $\tau_i^{(l)}(t)$ depends solely on the spatial trajectory of the lth marker. Therefore, when a person suddenly falls, the abrupt change in the spatial positions of the upper-body segment is reflected in the corresponding TV propagation delays $\tau_1^{(l)}(t)$, as illustrated in Fig. 4. In Fig. 4, the TV propagation delays $\tau_1^{(l)}(t)$ demonstrate the repetitive nature of the walking activity. By analyzing the TV propagation delays $\tau_1^{(l)}(t)$, it is evident that the simulated walking activity comprised four steps toward Radar$_1$. In contrast, the TV propagation delays $\tau_1^{(l)}(t)$ in Fig. 4 for the other three types of simulated human activities in place, namely sitting, standing up and picking up an object, show smaller variations corresponding to the mobility of the simulated point scatterers.

C. Channel Modeling for RF Sensing

This section elucidates the simulation of a composite RF signal or equivalently, raw IQ data in fast time t' and slow time t, corresponding to a specific motion. To simulate the composite RF signal of Radar$_i$, we need user-defined scatterer weights, a user-defined antenna configuration $\{\mathcal{C}_i^{T_x}, \mathcal{C}_i^{R_x}\}$, and the simulated TV propagation delays $\tau_i^{(l)}(t)$ corresponding to the spatial trajectories of the simulated point scatterers for a specific motion or a human activity (see Figs. 1(b) and 4). For this study, we consider the $\mathcal{L}$ bodily scatterers to be long-time non-stationary over the slow time t, and short-time stationary over a limited chirp duration T_{sw} [42]. In the following, for the FMCW 2×2 distributed MIMO radar system placed in the indoor wireless channel, we synthesize the complex baseband signal called the composite beat signal $s_{b,i}(t', t)$ [43], where $i = 1, 2$. Additionally, we discuss an interpolation procedure that is integral to our channel-simulation module of Fig. 1(b), as it mitigates the issues of aliasing in the Doppler domain.

FMCW radar systems operate by repetitively emitting a chirp waveform $c(t')$ [44], which is scattered back to the receiver antenna by multiple stationary and non-stationary scatterers present on the human body segments and other objects in the environment. A quadrature mixture element integrated into the receiver chain of the FMCW 2×2 dis-
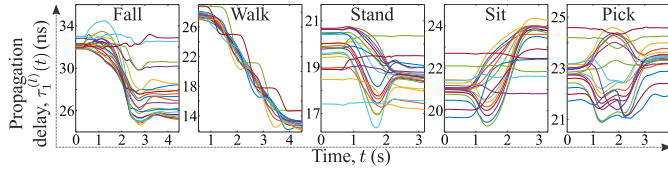
Fig. 4. Simulated TV propagation delays $\tau_1^{(l)}(t)$ of the L simulated point scatterers for the five distinct human activities and Radar$_1$.

tributed MIMO radar system is responsible for transforming the incoming passband RF signal into complex baseband (composite beat) signal $s_{b,i}(t', t)$. The received complex baseband signal $s_{b,i}(t', t)$ is sampled in the fast-time domain by the analog-to-digital converter (ADC) module of the receiver with the discrete sampling interval T_s in the fast-time domain. Subsequently, for the coherent processing interval (CPI) of the ith radar subsystem, Radar$_i$, the discrete samples of the received complex baseband signal $s_{b,i}(t', t)$ are organized in fast- and slow-time domains. During the CPI, the phase of Radar$_i$ is preserved. This organization or rearrangement of the discrete fast- and slow-time samples results in the radar's raw IQ data matrix $\mathcal{D}_i$ [42], which can be expressed as

$$
\mathcal{D}_i = \begin{bmatrix}
s_{b,i}(0, 0) & s_{b,i}(T_s, 0) & \dots \\
s_{b,i}(0, T_{sw}) & s_{b,i}(T_s, T_{sw}) & \dots \\
\vdots & \vdots & \vdots \\
s_{b,i}(0, (N_c - 1)T_{sw}) & s_{b,i}(T_s, (N_c - 1)T_{sw}) & \dots
\end{bmatrix}
$$

$$
\begin{matrix}
s_{b,i}(T_{sw} - T_s, 0) \\
s_{b,i}(T_{sw} - T_s, T_{sw}) \\
\vdots \\
s_{b,i}(T_{sw} - T_s, (N_c - 1)T_{sw})
\end{matrix} \quad (1)
$$

where N_c represents the number of chirps present within the CPI of the FMCW radar system.

We want to synthesize the actual received complex baseband signal $s_{b,i}(t', t)$ of the FMCW 2×2 distributed MIMO radar system, so that we can simulate the radar's raw IQ data matrices $\mathcal{D}_i$ for $i = 1, 2$. The received complex baseband signal $s_{b,i}(t', t)$ of Radar$_i$ can be synthesized by adding up the $\mathcal{L}$ distinct beat signals $s_{b,i}^{(l)}(t', t)$ [42], [43], each corresponding to the lth multipath component originating from the lth simulated point scatterer, i.e.,

$$
s_{b,i}(t', t) = \sum_{l=1}^{\mathcal{L}} s_{b,i}^{(l)}(t', t). \quad (2)
$$

For Radar$_i$, the lth beat signal $s_{b,i}^{(l)}(t', t)$ or the lth multipath component can be simulated by using the expression [42]

$$
s_{b,i}^{(l)}(t', t) = \sum_{n=0}^{\infty} a_i^{(l)}(t) \exp\left[j\left(2\pi f_{b,i}^{(l)}(t)t' + \phi_i^{(l)}(t) \right) \right]
$$
$$
\times \, \delta(t - \tau_i^{(l)}(t) - T_n) \quad (3)
$$

where $a_i^{(l)}(t)$, $f_{b,i}^{(l)}(t)$, and $\phi_i^{(l)}(t)$ denote the TV path gain, beat frequency, and phase of the lth beat signal $s_{b,i}^{(l)}(t', t)$, respectively, and $\delta(\cdot)$ denotes the Dirac delta function. The symbol T_n in (3) represents the nth discrete slow-time instance, which is determined by the chirp duration T_{sw}, such that $T_n = nT_{sw}$,

where n is a non-negative integer. Let γ represent the slope of the chirp signal. Then, the lth TV beat frequency $f_{b,i}^{(l)}(t)$ of Radar$_i$ in (3) is given by $f_{b,i}^{(l)}(t) = \tau_i^{(l)}(t)\gamma$. The ith and lth TV phase $\phi_i^{(l)}(t)$ component is determined by the TV propagation delay $\tau_i^{(l)}(t)$ according to $\phi_i^{(l)}(t) = 2\pi f_0 \tau_i^{(l)}(t)$, where f_0 is the carrier frequency.

The TV path gain $a_i^{(l)}(t)$ in (3) models the strength of the lth marker in the received signal. For Radar$_i$ and $\mathcal{L}$ simulated point scatterers, we use time-invariant path gains $a_i^{(l)}$ in (3) to avoid unnecessary complexity. Therefore, we have $a_i^{(l)}(t) = a_i^{(l)}$. In this study, for the five types of synthesized human activities, the values of the time-invariant path gains $a_i^{(l)}$ are adjusted by investigating the actual TV radial velocity distributions $p_i(v, t)$ (see Section VI) and the body surface area [45]. It is worth noting that by using different sets of time-invariant path gains, we can augment the radar data at the signal-layer synthesis for a synthesized human activity (see Section V-C).

We consider the $\mathcal{L}$ bodily scatterers to be long-time non-stationary over the slow time t, and short-time stationary over the fast time t' for a limited chirp duration T_{sw} [42]. Thus, the TV propagation delays $\tau_i^{(l)}(t)$, beat frequencies $f_{b,i}^{(l)}(t)$, and phases $\phi_i^{(l)}(t)$ of the $\mathcal{L}$ simulated point scatterers are only a function of the slow time t. For Radar$_i$ and the kth slow-time instant t_k [kth row of the raw IQ data matrix $\mathcal{D}_i$ in (1)], the short-time stationarity assumption simplifies the synthesis of the discrete complex baseband signal $s_{b,i}(t', t_k)$ for a synthesized human activity. At the slow-time instant t_k, the IQ components of the complex baseband signal $s_{b,i}(t', t_k)$ can be digitally simulated as a sum of tone signals, i.e., $s_{b,i}(t', t_k) = \sum_{l=1}^{\mathcal{L}} s_{b,i}^{(l)}(t', t_k)$, where the lth tone signal $s_{b,i}^{(l)}(t', t_k)$ has the constant (time-invariant) beat frequency $f_{b,i}^{(l)}(t_k)$ and phase $\phi_i^{(l)}(t_k)$.

Within the framework of radar sensing, the synthesized motion can be completely characterized by the simulated TV propagation delays $\tau_i^{(l)}(t)$ of the $\mathcal{L}$ simulated (real) point scatterers. The $\mathcal{L}$ TV propagation delays $\tau_i^{(l)}(t)$ are computed from the TV spatial trajectories $\mathcal{C}_l(t)$ of the $\mathcal{L}$ simulated point scatterers, which are animated with a fixed frame interval denoted by T_f. Therefore, the frame interval T_f is the slow-time sampling interval of the simulated TV spatial trajectories $\mathcal{C}_l(t)$ and the propagation delays $\tau_i^{(l)}(t)$. In actual radar systems, the slow-time sampling interval is equal to the radar's pulse repetition interval (PRI), which is smaller (better) than the frame interval T_f. Concretely, for the actual (simulated) raw IQ data matrix $\mathcal{D}_i$ in (1), the slow-time sampling interval T_{sw} is equal to the radar's PRI (frame interval T_f). Thus, to ensure that the simulated frame interval T_f is equal to the radar's PRI, we interpolate the spatial trajectories or the simulated TV propagation delays $\tau_i^{(l)}(t)$ in our simulation framework. This is necessary because the upper limit of the actual (synthesizable) radial velocity, denoted by v_{max} (v'_{max}), is determined by the radar's PRI (animation's frame interval T_f). Let λ denote the wavelength, then we have $v_{max} = \lambda/(4 \cdot \text{PRI})$, and $v'_{max} = \lambda/(4T_f)$.

V. Multistage Data Augmentation

In this section, we explore multistage data augmentation techniques (see Fig. 5) provided by the proposed simulation-based framework that allow us to simulate large quantities of quality radar signatures. First, we discuss a motion-layer data augmentation technique, where various animation parameters and avatar characteristics, e.g., size and speed, can be randomly varied to synthesize a variety of human motions. We then explain data augmentation of the physical layer that allows us to vary numerous physical-layer configurations and the radar's operating parameters, e.g., number of antennas and their setup and PRI. Lastly, we delve into a data augmentation technique at the signal-layer synthesis.

A. Motion-Layer Synthesis

For the five types of distinct human activities, we acquired a small and basic MoCap dataset from the Mixamo platform and the Qualisys MoCap system. A person with a height of about 1.74 m performed the activities several times in a room equipped with the Qualisys MoCap system. The MoCap dataset we acquired comprised only 34 MoCap files, each representing one of the five types of activities. The 3-D animation tools from both Unity and MotionBuilder software were used to visualize the basic MoCap data for the human activities. We complemented the basic MoCap data with the 3-D animation tools to render realistic and diverse motion data.

In this study, one of our objectives is to synthesize a large amount of data representing real human motions at the motion-layer synthesis of our simulation-based framework. To this end, we first adjusted the height of the avatar in the MotionBuilder software by reducing it to 1.52 m (5 ft) and increasing it to 1.83 m (6 ft). We then aligned the MoCap data to the avatars with different sizes to account for the effects of avatar dimensions and extended the data on the motion-layer synthesis. Therefore, in the Unity and MotionBuilder software, the total number of synthesized human activities were increased to 84 by applying data augmentation at the motion-layer synthesis stage, as indicated by Fig. 5. Note that we can synthesize complex, varied, and entirely new sequences of human movements by using the blend tree animation tool in the Unity software that facilitates seamless transitions between multiple humanoid animations. For the augmented human-motion data (synthesized human activities), we computed TV spatial trajectories (see Section III-D) and imported them into MATLAB for further data augmentation at the physical- and signal-layer syntheses (see Fig. 5).

B. Physical-Layer Synthesis

The simulation-based framework allows the adjustment of the radar operating parameters and physical-layer configurations, e.g., PRI, carrier frequency f_c, bandwidth B_w, and antenna configuration $\{\mathcal{C}_i^{T_x}, \mathcal{C}_i^{R_x}\}$. Through these adjustments, it is possible to both extend the simulated radar data and simulate specific scenarios. At the physical-layer synthesis data augmentation stage, appropriate antenna configurations $\{\mathcal{C}_i^{T_x}, \mathcal{C}_i^{R_x}\}$ were chosen to simulate the two radar subsystems, Radar$_1$ and Radar$_2$, as shown in Fig. 6. To maintain consistency with the actual 2×2 distributed MIMO

radar system depicted in Fig. 1(a), the emulated radar system's operating parameters, such as PRI, carrier frequency f_c, and bandwidth B_w, were kept the same.

We first simulated different positions of the radar subsystems, Radar$_1$ and Radar$_2$, by using the rotation matrix $R_y(\theta_{Ri})$, which can be expressed as [46]

$$R_y(\theta_{Ri}) = \begin{bmatrix} \cos\theta_{Ri} & 0 & \sin\theta_{Ri} \\ 0 & 1 & 0 \\ -\sin\theta_{Ri} & 0 & \cos\theta_{Ri} \end{bmatrix} \quad (4)$$

where θ_{Ri} denotes the clockwise angular rotation along the y-axis for Radar$_i$ and $i = 1, 2$. Initially, the simulated radar subsystems, Radar$_1$ and Radar$_2$, were placed at $\theta_{R1} = 0°$ ($\mathcal{C}_1^{T_x} \approx \mathcal{C}_1^{R_x} \approx [3, 1.1, 0]^\top$) and $\theta_{R2} = -90°$ ($\mathcal{C}_2^{T_x} \approx \mathcal{C}_2^{R_x} \approx [0, 1.1, 3]^\top$), respectively, i.e., $\mathcal{C}_2^{T_x} = R_y(-90°)\mathcal{C}_1^{T_x}$ and $\mathcal{C}_2^{R_x} = R_y(-90°)\mathcal{C}_1^{R_x}$. In other words, Radar$_2$ can be simulated by simply rotating Radar$_1$ counterclockwise by 90° along the y-axis, as illustrated in Fig. 6. Using this method, we emulated a 2×2 distributed MIMO radar system, similar to the actual radar system in Fig. 1(a), to simulate the MIMO radar signatures. Note that, with the use of the rotation matrix $R_y(\theta_{Ri})$, any number of radar subsystems, sensors, or nodes can be simulated at the physical-layer synthesis data augmentation stage.

Recall that the human activities were initially simulated with 3-D animation tools in a single direction or at an aspect angle of 0°. However, to develop a simulated MIMO radar-based direction-independent HAR system, we required multidirectional human activities. Compared to the motion-layer synthesis, the required multidirectional human activities can be simulated more easily and efficiently at the physical-layer synthesis data augmentation stage. The multidirectional human activities are simulated by spatially rotating the transmitter and receiver antennas of the radar subsystem Radar$_i$, for $i = 1, 2$. The angular difference between the two radar subsystems is always kept at 90°, i.e., $\theta_{R1} - \theta_{R2} = 90°$, as depicted in Fig. 6. The different rotations of Radar$_1$ and Radar$_2$ (θ_{R1}, θ_{R2}) correspond to the different directions of the human activities, where $(\theta_{R1}, \theta_{R2}) \in [-180°, 180°)$. We simulated 18 different directions of the human activities at the physical-layer synthesis data augmentation stage, namely Direction 1 to Direction 18, as illustrated in Fig. 6. For instance, for a human activity, Direction 11 in Fig. 6 corresponds to the scenario, where $(\theta_{R1}, \theta_{R2}) = (-160°, 110°)$.

To summarize, at the physical-layer synthesis, we first simulated the two radar subsystems, Radar$_1$ and Radar$_2$, to emulate the 2×2 distributed MIMO radar system. Second, by using the rotation method, we simulated the multidirectional human activities by simultaneously rotating the two radar subsystems, as illustrated in Fig. 6. Thus, our proposed simulation-based framework includes a physical-layer synthesis data augmentation stage, which efficiently and conveniently transforms and augments unidirectional motion data into multidirectional motion data and single radar data into multiple radar data.

C. Signal-Layer Synthesis

The signal-layer synthesis data augmentation stage of the proposed simulation-based framework allows to simulate real-
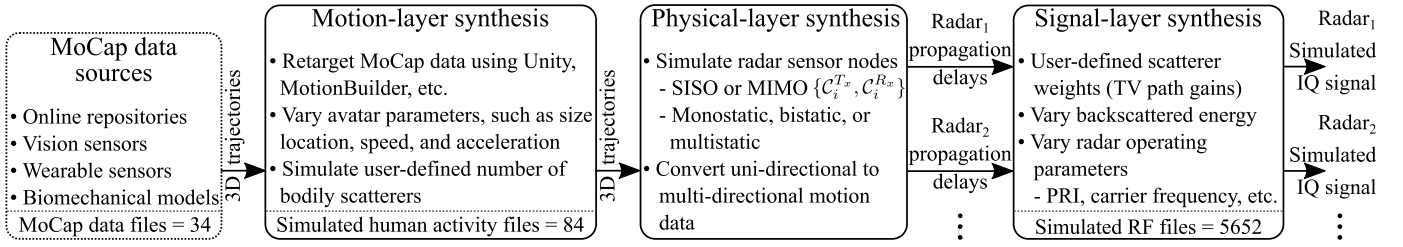
Fig. 5. Proposed multistage data augmentation techniques offered by our simulation-based framework.
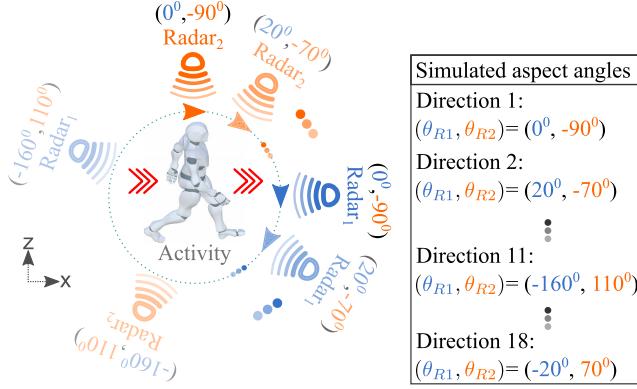

Fig. 6. Simulation of a multiperspective 2 × 2 MIMO radar system and multidirectional human activities.

istic and diverse TV radial velocity distributions $p_i(v, t)$ (micro-Doppler signatures) for a single human activity. Using (3), we can simulate numerous multipath components corresponding to the stationary and non-stationary scatterers in the received complex baseband signal $s_{b,i}(t', t)$ [see (2)]. In this research, multipath components originating from stationary scatterers, such as walls and furniture, are not considered, as they can be effectively filtered out during the signal preprocessing stage. However, if necessary, the signal-layer synthesis can simulate numerous complex propagation scenarios, e.g., those with or without radar clutter, by adjusting the path gains $a_i^{(l)}(t)$, beat frequencies $f_{b,i}^{(l)}(t)$, and phases $\phi_i^{(l)}(t)$ of the lth beat signal $s_{b,i}^{(l)}(t', t)$ for Radar$_i$.

For the five types of synthesized human activities, we first adjusted the values of the time-invariant path gains $a_i^{(l)}$ by looking into the actual radar signatures (TV radial velocity distributions $p_i(v, t)$ [see Section VI]) and the body surface area [45]. Subsequently, we augmented the simulated radar signatures by varying the power levels (time-invariant path gains $a_i^{(l)}$) of the individual multipath component. Therefore, at the signal-layer synthesis of the proposed simulation-based framework, we augmented the radar data by using different sets of time-invariant path gains $a_i^{(l)}$ for the five types of synthesized human activities.

In this section, we discussed three data augmentation techniques implemented at multiple layers of the proposed simulation-based framework: the motion-layer synthesis, physical-layer synthesis, and signal-layer synthesis. By applying these multistage data augmentation techniques, we simulated 2826 TV radial velocity distributions $p_i(v, t)$ (micro-Doppler signatures) for each radar subsystem of the 2 × 2 MIMO radar system. In other words, for the two radar

subsystems, Radar$_1$ and Radar$_2$, a total of 5652 TV radial velocity distributions $p_i(v, t)$ were simulated. To conclude, the multistage data augmentation methods in the proposed simulation-based framework are quite useful and they allowed for increased variability, realism, and diversity in the simulated radar dataset. With these methods, we were able to transform and augment the basic motion data (34 MoCap files) into 5652 radar signatures, which indicates the utility of the proposed simulation-based approach for realizing radar-based classifiers.

VI. MIMO RADAR SIGNATURES

In this section, we delineate the radar signal processing module of Fig. 1 that generates the MIMO radar signatures: range distribution, TV radial velocity distribution $p_i(v, t)$ (micro-Doppler signature), and mean velocity (mean Doppler shift). For $i = 1, 2$, the radar signal processing module transforms the actual and the simulated complex baseband signals $s_{b,i}(t', t)$ into the TV radial velocity distributions $p_i(v, t)$. The first step is to compute the beat frequency function $S_{b,i}(f_b, t)$ as [47]

$$S_{b,i}(f_b, t) = \int_0^{T_{sw}} s_{b,i}(t', t) e^{-j2\pi f_b t'} dt' \tag{5}$$

where f_b refers to the beat frequency.

Let f and $f_{b,\max}$ denote the Doppler frequency and maximum beat frequency, respectively. Then, the micro-Doppler signatures $S_i(f, t)$ are obtained from the beat frequency function $S_{b,i}(f_b, t)$ according to the relation [35]

$$S_i(f, t) = \left| \int_0^{f_{b,\max}} \int_{-\infty}^{\infty} S_{b,i}(f_b, t'') W_r(t'' - t) e^{-j2\pi f t''} dt'' df_b \right|^2 \tag{6}$$

where t'' denotes the running time, and $W_r(\cdot)$ denotes a rectangular window function that spans over 64 chirp intervals. According to [43], the TV radial velocity distribution $p_i(v, t)$ can be obtained as

$$p_i(v, t) = \frac{S_i\left(\frac{2f_0}{c_0} v, t\right)}{\int_{-\infty}^{\infty} S_i\left(\frac{2f_0}{c_0} v, t\right) dv} \tag{7}$$

where v denotes the radial velocity. From the TV radial velocity distribution $p_i(v, t)$ in (7), we can compute the TV mean radial velocity $\bar{v}_i(t)$ as [43]

$$\bar{v}_i(t) = \int_{-\infty}^{\infty} v p_i(v, t) dv. \tag{8}$$

For Radar$_i$, the TV beat-frequency signatures $S'_i(f_b, t)$ can be computed as

$$S'_i(f_b, t) = \left| \int_0^{\text{PRF}} \int_{-\infty}^{\infty} S_{b,i}(f_b, t'') W_r(t'' - t) e^{-j2\pi f t''} dt'' df \right|^2 \tag{9}$$

where PRF is the pulse repetition frequency of the radar system, and $i = 1, 2$. Finally, for the 2×2 MIMO radar system, the TV range distribution $p'_i(r, t)$ can be obtained as [42]

$$p'_i(r, t) = \frac{S'_i\left(\frac{2\gamma}{c_0} r, t\right)}{\int_{-\infty}^{\infty} S'_i\left(\frac{2\gamma}{c_0} r, t\right) dr}. \tag{10}$$

Recall that the real (simulated) point scatterers on the human (avatar) body segments, each with unique TV radial velocity components, scatter the transmitted RF signal back to the receiver antennas of the 2×2 distributed MIMO radar system. For Radar$_i$ and the $\mathcal{L}$ distinct non-stationary real (simulated) point scatterers, the TV radial velocity distribution $p_i(v, t)$ in (7) indicates the strengths of the radial velocity components over the slow time t (see Fig. 7). The TV mean radial velocity $\bar{v}_i(t)$ in (8), obtained from the TV radial velocity distribution $p_i(v, t)$, shows the weighted average of the velocity components of all $\mathcal{L}$ real (simulated) bodily scatterers over the slow time t (see Fig. 8). Moreover, the strengths of the TV radial distances of all $\mathcal{L}$ non-stationary real (simulated) point scatterers over the slow time t are provided by the TV range distributions $p'_i(r, t)$. Due to the current practical limitations of radar systems, the TV range distributions $p'_i(r, t)$ are not usually used to realize HAR systems, so their simulation results are omitted for brevity. However, for completeness and possible future applications, we have included the expression in (10) to simulate the TV range distribution $p'_i(r, t)$.

In Section V, we saw that multidirectional human activities can be simulated by simultaneously rotating the two radar subsystems, Radar$_1$ and Radar$_2$, as shown in Fig. 6. For some of the 18 directions and all five types of simulated (actual) human activities, the simulated (actual) TV radial velocity distributions, $p_1(v, t)$ and $p_2(v, t)$, are shown in Fig. 7(a) [Fig. 7(b)]. The images of the simulated (actual) TV radial velocity distributions, $p_1(v, t)$ and $p_2(v, t)$, were used to train (test) the proposed 2×2 MIMO radar-based direction-independent HAR system. In Section VII, the two colored images of the TV radial velocity distributions, $p_1(v, t)$ and $p_2(v, t)$, will serve as input feature maps to the HAR system. Moreover, for the five types of human activities and the two radar subsystems, Radar$_1$ and Radar$_2$, the simulated and actual TV mean radial velocities $\bar{v}_i(t)$ are depicted in Fig. 8. The utility and effectiveness of the proposed simulation-based framework is evident from the high-fidelity simulated radar signatures, which are quite similar to the actual radar signatures, as exemplified by Figs. 7 and 8.

To quantitatively assess the similarity between simulated and real radar signatures, we employ the dynamic time warping (DTW) algorithm [48]. Table I presents the normalized DTW distances between the real and simulated TV mean
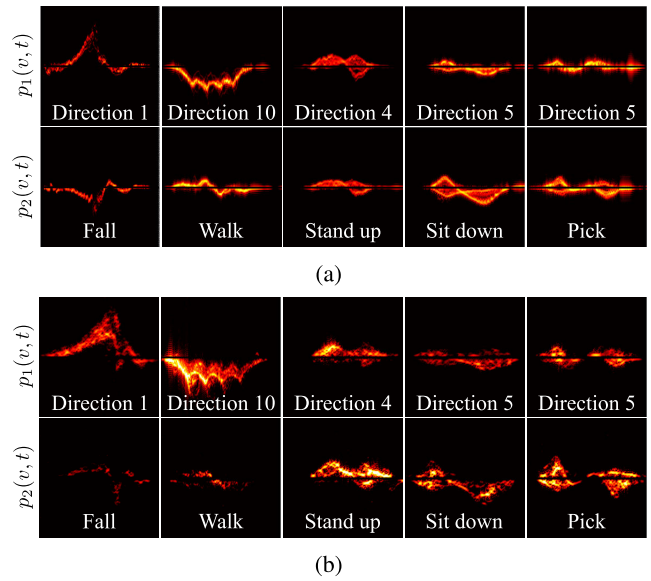

(a)

(b)

Fig. 7. (a) Simulated TV radial velocity distributions, $p_1(v, t)$ and $p_2(v, t)$, for the emulated multidirectional human activities. (b) Real TV radial velocity distributions, $p_1(v, t)$ and $p_2(v, t)$, for the real multidirectional human activities.
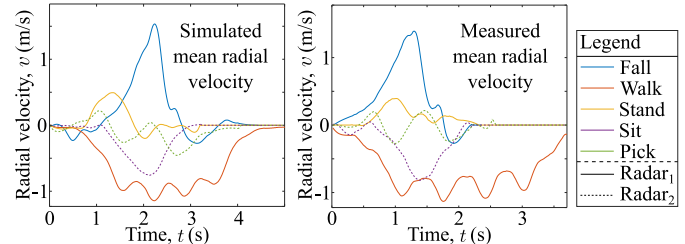

Fig. 8. Simulated and measured TV mean radial velocities $\bar{v}_1(t)$ and $\bar{v}_2(t)$ of Radar$_1$ (—) and Radar$_2$ (···), respectively.

radial velocities $\bar{v}_i(t)$ from Fig. 8 across five human activities. Remarkably, the DTW distance metric is minimized for each activity, indicating close resemblance between the simulated and real radar signatures. For example, for the sitting activity, a DTW distance of 0.01 between the simulated and real TV mean radial velocities $\bar{v}_i(t)$ demonstrates precise replication of this pattern. This consistent trend across all activities confirms the accuracy of our approach in simulating realistic radar data.

VII. SIMULATION-BASED HAR SYSTEM

This section elucidates the training and testing phases of our simulation-based direction-independent HAR system that was realized by using a DCNN-based multiclass classifier. First, we look into the design of the HAR classifier and its training with the simulated radar dataset. Then, to demonstrate the practical importance and the generalizability of our proposed simulation framework in real-world scenarios, we used a real 2×2 MIMO radar dataset to evaluate the classification performance of the trained simulation-based direction-independent HAR system.

A. Design of the Simulation-Based HAR System

To develop a simulation-based HAR system, we first created a large labeled dataset of simulated radar signatures. For Radar$_i$ ($i = 1, 2$) of the 2×2 MIMO radar system and the five types of humanoid activities, we simulated 2826 TV radial velocity distributions $p_i(v, t)$ by using the

TABLE I
DTW DISTANCE METRIC IS CALCULATED FOR THE SIMULATED AND
REAL (ACTUAL) TV MEAN RADIAL VELOCITIES $\bar{v}_i(t)$ OF FIG. 8

	Activity type	Real				
		fall	walk	stand	sit	pick
Simulated	fall	**0.03**	0.30	0.14	0.21	0.14
	walk	0.28	**0.01**	0.36	0.03	0.22
	stand	0.06	0.24	**0.03**	0.11	0.05
	sit	0.17	0.05	0.17	**0.01**	0.08
	pick	0.09	0.13	0.08	0.04	**0.02**

proposed multistage data augmentation techniques of our simulation-based framework (see Section V). Thus, the simulated radar dataset consisted of a total of 5652 simulated TV radial velocity distributions $p_i(v, t)$, which were used to train the proposed simulation-based direction-independent (multiperspective) HAR classifier.

The simulation-based direction-independent HAR system comprises two feature extraction networks (FENs) and a multi-layer perceptron (MLP) network. Fig. 9(a) illustrates the FEN that computes relevant features from the simulated (actual) TV radial velocity distributions $p_i(v, t)$ during the training (testing) phase for the ith radar subsystem, Radar$_i$. It consists of four convolutional layers, containing 64, 72, 80, and 96 2-D trainable kernels with dimension k_d either equal to 4×4 pixels or 3×3 pixels. Each 2-D kernel uses the rectified linear unit (ReLU) activation function to avoid the problem of vanishing gradients [49]. The max-pool layers were employed to reduce redundancies in the feature maps. To avoid overfitting the training data, we used dropout layers with the dropout rates of 10% and 15% for the FEN and MLP, respectively. The flatten layer of our FEN rearranges the extracted features into a vector of order 18816×1, as shown in Fig. 9(a).

The two FENs in the DCNN-based multiperspective HAR system are identical, as shown in Fig. 9(b). As Radar$_1$ and Radar$_2$ illuminate the indoor environment from multiple perspectives, the extracted features from the two TV radial velocity distributions, $p_1(v, t)$ and $p_2(v, t)$, are merged by the multiperspective feature fusion block, as shown in Fig. 9(b). Subsequently, based on the received multiperspective features, the MLP network is trained to detect the type of the human activity. The multiperspective feature fusion block enables the HAR classifier to recognize the human activities regardless of their directions. Note that the design of this multiperspective deep neural network closely resembles the architecture reported in [35]. To train the parameters of our DCNN-based multiperspective HAR classifier, we used the adaptive moment estimation (Adam) optimizer [50] and the simulated radar signatures of multidirectional human activities. The training dataset, comprising 2826 pairs of simulated TV radial velocity distributions $p_i(v, t)$, was further divided into training and validation subsets in an $80 : 20$ ratio. During the training phase, our DCNN-based multiperspective HAR classifier showed no signs of overfitting, as demonstrated by the training and validation curves in Fig. 10.

B. Testing of the Simulation-Based HAR System

To evaluate the performance of the trained 2×2 MIMO radar-based multiperspective HAR classifier in a real-world setting, we used a real radar dataset recorded by Ancortek
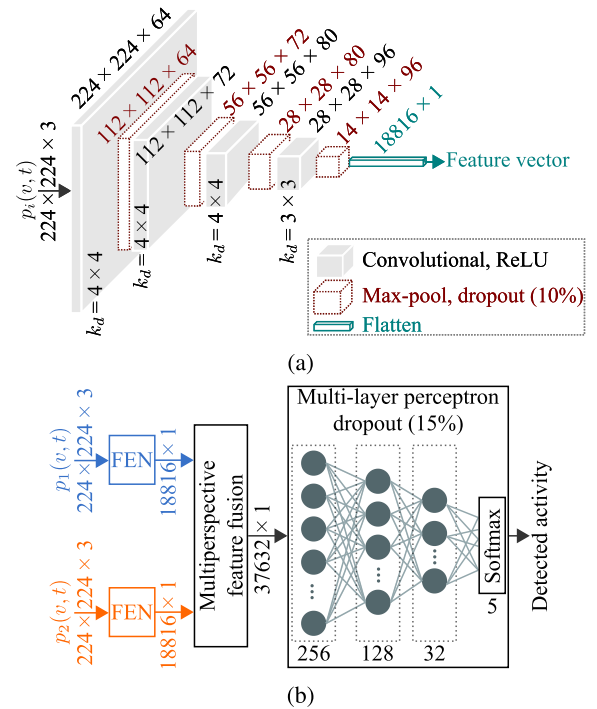

Fig. 9. (a) Feature extraction network, FEN, based on convolutional filters. (b) Proposed DCNN-based multiperspective HAR classifier that is trained (tested) using the simulated (real) HAR dataset.
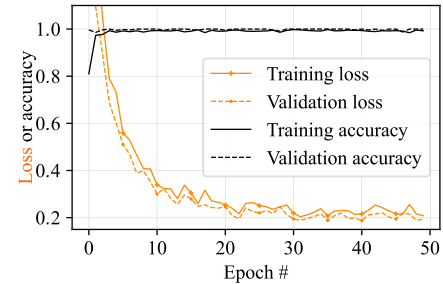

Fig. 10. Training history of our simulation-based direction-independent HAR system.

SDR-KIT 2400T2R4, as shown in Fig. 2. The operating parameters and antenna configurations of the real and the simulated 2×2 MIMO radar systems were kept similar for consistency. Specifically, we set the PRI, carrier frequency f_c, and bandwidth B_w of the real and simulated MIMO radar systems to 0.5 ms, 24.125 GHz, and 250 MHz, respectively. For Radar$_1$ and Radar$_2$, the antennas were placed at $\mathcal{C}_1^{T_x} \approx \mathcal{C}_1^{R_x} \approx [3, 1.1, 0]^\top$ and $\mathcal{C}_2^{T_x} \approx \mathcal{C}_2^{R_x} \approx [0, 1.1, 3]^\top$, respectively.

A total of 875 multidirectional human activities were recorded with the 2×2 MIMO radar system from six human subjects, including a female participant. Thus, the real radar dataset consisted of 1750 TV radial velocity distributions $p_i(v, t)$ (micro-Doppler signatures) for the two radar subsystems, Radar$_1$ and Radar$_2$. As a direct result of this extensive measurement campaign, the simulation–real (training–testing) data ratio approximately came out to be $76 : 24$. Our simulation-based framework enabled the realization of the simulation-based direction-independent HAR system, which exhibited remarkable performance and efficacy in the real world, as demonstrated by the confusion matrix in Fig. 11. For each of the five types of multidirectional human activities, the number of correct classifications is represented

True Class	Fall	Walk	Stand	Sit	Pick	
Fall	94	2	0	1	0	96.91% / 3.09%
Walk	2	151	0	0	0	98.69% / 1.31%
Stand	2	0	163	0	0	98.79% / 1.21%
Sit	0	1	3	223	4	96.54% / 3.46%
Pick	4	0	0	0	225	98.25% / 1.75%
	92.16% / 7.84%	98.05% / 1.95%	98.19% / 1.81%	99.55% / 0.45%	98.25% / 1.75%	97.83% / 2.17%
	Fall	Walk	Stand	Sit	Pick	

Predicted Class

Fig. 11. Confusion matrix of our simulation-based multiperspective HAR classifier with a classification accuracy of 97.83%.

TABLE II
COMPARING THE CLASSIFICATION PERFORMANCE OF STATE-OF-THE-ART RF-BASED HAR APPROACHES

Approach	Training data type & brief description	Accuracy
Wi-Sense [14]	Involves measurement data; Obtaining fingerprints from the Wi-Fi channel state information (CSI)	97.8%
Convolutional neural network (CNN)-LSTM [54]	Involves measurement data; Unsupervised learning and data fusion using LSTM	92%
Few-shot adversarial domain adaptation (FS-ADA) [55]	Involves measurement data; Discovering shared feature spaces in training datasets	91.6%
CNN-recurrent neural network (RNN) [56]	Involves measurement data; Extracting patterns over space and time	90.8%
Joint domain and semantic transfer learning (JDS-TL) [12]	Involves measurement data; Utilizing 10% labeled radar data for domain adaptation	87.6%
Wasserstein refined generative adversarial network with gradient penalty (WRGAN-GP) [57]	Synthesizing data with GANs utilizing limited measurements; Employing a refinement model to synthesize high-fidelity spectrograms	94.9%
Multibranch generative adversarial network (MBGAN) [58]	GAN-centered data generation using limited real data; Physics-informed method improving GANs for accurate micro-Doppler synthesis	89.2%
Vid2Doppler [53]	Simulation-based training; Converting video data into Doppler data	81.4%
Our simulation-based approach	Simulation-based training; Converting motion data into radar signatures	97.8%

by the first five diagonal entries of the confusion matrix. The green colored entries in the last row and column exhibit the precision and recall [51] in Fig. 11. Finally, the white colored entry of the confusion matrix shows the overall classification accuracy of our simulation-based direction-independent HAR system, which is 97.83%. As our test dataset was sufficiently balanced, the macro-average F1-score [52] came out to be approximately 97.6%, which is close to the overall classification accuracy.

For RF-based HAR systems, asserting the superiority of one method proves challenging, given their tailored design to address diverse research challenges. Nonetheless, Table II presents the performance of various contemporary HAR sys-

tems, utilizing classification accuracy for comparison. Notably, the measurement-based HAR methods and those partially utilizing measurement data demonstrate strong classification accuracies. The Vid2Doppler [53] method, which translates video to radar data, achieves an accuracy of 81.4%, while our simulation-based approach, converting MoCap data to radar data, achieves a higher accuracy of 97.8%, both utilizing entirely simulated training data.

This section demonstrated the utility and efficacy of the simulation-based framework in the real world. The classification accuracy of the simulation-based direction-independent HAR system is comparable to the current HAR systems [18], [59], with the additional consideration of the multidirectional HAR problem. Moreover, our simulation-based framework is unique in its ability to generate realistic, diverse, and unlimited labeled MIMO radar datasets with software-defined operating parameters and configurations. Therefore, the proposed simulation-based framework in Fig. 1(b) can be readily used to develop other SISO and MIMO radar-based classifiers, e.g., for sign language detection.

VIII. CONCLUSION

The progression of SISO and MIMO radar-based classifiers is primarily impeded by the unavailability of large labeled training datasets. Therefore, as a proof-of-concept, we have presented in this work a simulation-based approach to address the concern of data scarcity for monostatic, bistatic, and multistatic SISO and MIMO radar systems. Although our focus was on realizing a 2×2 MIMO radar-based direction-independent HAR system, the utility of our simulation-based framework extends beyond HAR applications.

The proposed simulation-based framework provides the flexibility to synthesize software-defined human movements using MoCap data-driven activity simulation. We proposed a MIMO channel model to convert simulated 3-D trajectories into received RF signals, while considering a user-defined antenna configuration of a distributed MIMO radar system and the multipath components emanating from the non-stationary simulated point scatterers. The synthesized RF signals were further processed to simulate the multiperspective MIMO radar signatures used to implement our simulation-based direction-independent HAR system.

To generate a diverse training dataset for radar-based HAR systems, we introduced multistage data augmentation techniques at the motion-layer synthesis, physical-layer synthesis, and signal-layer synthesis within our simulation-based framework. The multistage data augmentation techniques helped to gain absolute control over various factors, such as avatar size, location, velocity, acceleration, PRI, and radar antenna configuration. By using these techniques, we augmented the basic MoCap data to 5652 micro-Doppler signatures, drastically minimizing the overall training workload and demonstrating the effectiveness of our simulation-based approach for realizing radar-based classifiers. Our MIMO radar-based HAR system trained on the simulated micro-Doppler signatures achieved classification accuracy of 97.83% when tested with

actual radar data. As our study eliminates the need for direct involvement of human participants and an actual radar system, we believe that the proposed proof-of-concept will be of great importance for training future SISO/MIMO radar-based classifiers.

Our MIMO channel model opens up new research perspectives for modeling received RF signals at the scatterer level. For example, future studies can explore the optimization of scatterer-level parameters, such as the simulated TV path gains. A limitation of this research is that the methods discussed are not directly applicable to the moving clutter scenario where the radar antennas are non-stationary. This research gap is beyond the scope of this work and can be addressed in future studies.

ACKNOWLEDGMENT

REFERENCES

[1] C. Hu, S. Hudson, M. Ethier, M. Al-Sharman, D. Rayside, and W. Melek, "Sim-to-real domain adaptation for lane detection and classification in autonomous driving," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2022, pp. 457–463.

[2] M. Hernandez, G. Epelde, A. Alberdi, R. Cilla, and D. Rankin, "Synthetic data generation for tabular health records: A systematic review," *Neurocomputing*, vol. 493, pp. 28–45, Jul. 2022.

[3] B. Kiefer, D. Ott, and A. Zell, "Leveraging synthetic data in object detection on unmanned aerial vehicles," in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2022, pp. 3564–3571.

[4] K. Hagn and O. Grau, "Improved sensor model for realistic synthetic data generation," in *Proc. Comput. Sci. Cars Symp.*, Nov. 2021, pp. 1–9, doi: 10.1145/3488904.3493383.

[5] M. A. Bansal, D. R. Sharma, and D. M. Kathuria, "A systematic review on data scarcity problem in deep learning: Solution and applications," *ACM Comput. Surveys*, vol. 54, no. 10s, pp. 1–29, Jan. 2022.

[6] H. Tang and K. Jia, "A new benchmark: On the utility of synthetic data with blender for bare supervised learning and downstream domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 15954–15964.

[7] V. Thambawita et al., "SinGAN-seg: Synthetic training data generation for medical image segmentation," *PLoS ONE*, vol. 17, no. 5, May 2022, Art. no. e0267976, doi: 10.1371/journal.pone.0267976.

[8] M. A. Al-Qaness, A. Dahou, M. Abd Elaziz, and A. Helmi, "Multi-ResAtt: Multilevel residual network with attention for human activity recognition using wearable sensors," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 144–152, Jan. 2023.

[9] Z. Chen, Z. Cai, T. Zheng, J. Luo, J. Xiong, and X. Wang, "RF-based human activity recognition using signal adapted convolutional neural network," *IEEE Trans. Mobile Comput.*, vol. 22, no. 1, pp. 487–499, Jan. 2023.

[10] Y. Tang, L. Zhang, F. Min, and J. He, "Multiscale deep feature learning for human activity recognition using wearable sensors," *IEEE Trans. Ind. Electron.*, vol. 70, no. 2, pp. 2106–2116, Feb. 2023.

[11] Y. Tang, L. Zhang, Q. Teng, F. Min, and A. Song, "Triple cross-domain attention on human activity recognition using wearable sensors," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 6, no. 5, pp. 1167–1176, Oct. 2022.

[12] X. Li, Y. He, F. Fioranelli, and X. Jing, "Semisupervised human activity recognition with radar micro-Doppler signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5103112.

[13] L. Zhang, W. Cui, B. Li, Z. Chen, M. Wu, and T. S. Gee, "Privacy-preserving cross-environment human activity recognition," *IEEE Trans. Cybern.*, vol. 53, no. 3, pp. 1765–1775, Mar. 2023.

[14] M. Muaaz, A. Chelli, M. W. Gerdes, and M. Pätzold, "Wi-sense: A passive human activity recognition system using Wi-Fi and convolutional neural network and its integration in health information systems," *Ann. Telecommun.*, vol. 77, nos. 3–4, pp. 163–175, Apr. 2022.

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

[16] S. Waqar and M. Pätzold, "Interchannel interference and mitigation in distributed MIMO RF sensing," *Sensors*, vol. 21, no. 22, p. 7496, Nov. 2021, doi: 10.3390/s21227496.

[17] B. Erol and M. G. Amin, "Radar data cube processing for human activity recognition using multisubspace learning," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 6, pp. 3617–3628, Dec. 2019.

[18] W.-Y. Kim and D.-H. Seo, "Radar-based human activity recognition combining range–time–Doppler maps and range-distributed-convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 1002311.

[19] G. Lee and J. Kim, "Improving human activity recognition for sparse radar point clouds: A graph neural network model with pre-trained 3D human-joint coordinates," *Appl. Sci.*, vol. 12, no. 4, p. 2168, Feb. 2022, doi: 10.3390/app12042168.

[20] Z. Xia and F. Xu, "Time-space dimension reduction of millimeter-wave radar point-clouds for smart-home hand-gesture recognition," *IEEE Sensors J.*, vol. 22, no. 5, pp. 4425–4437, Mar. 2022.

[21] H. Liu et al., "M-gesture: Person-independent real-time in-air gesture recognition using commodity millimeter wave radar," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3397–3415, Mar. 2022.

[22] M. M. Rahman, E. A. Malaia, A. C. Gurbuz, D. J. Griffin, C. Crawford, and S. Z. Gurbuz, "Effect of kinematics and fluency in adversarial synthetic data generation for ASL recognition with RF sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 4, pp. 2732–2745, Aug. 2022.

[23] Z. Yu et al., "A radar-based human activity recognition using a novel 3-D point cloud classifier," *IEEE Sensors J.*, vol. 22, no. 19, pp. 18218–18227, Oct. 2022.

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

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

[26] Y. Yang, Y. Zhang, Y. Lang, B. Li, S. Guo, and Q. Tan, "GAN-based radar spectrogram augmentation via diversity injection strategy," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023.

[27] Z. Liu, C. Wu, and W. Ye, "Category-extensible human activity recognition based on Doppler radar by few-shot learning," *IEEE Sensors J.*, vol. 22, no. 22, pp. 21952–21960, Nov. 2022.

[28] Y. Yang, Y. Zhang, H. Ji, B. Li, and C. Song, "Radar-based human activity recognition under the limited measurement data support using domain translation," *IEEE Signal Process. Lett.*, vol. 29, pp. 1993–1997, 2022.

[29] M. J. Bocus et al., "*OPERAnet*, a multimodal activity recognition dataset acquired from radio frequency and vision-based sensors," *Scientific Data*, vol. 9, no. 1, p. 474, Aug. 2022, doi: 10.1038/s41597-022-01573-2.

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

[31] G. Manfredi, P. Russo, A. De Leo, and G. Cerri, "Efficient simulation tool to characterize the radar cross section of a pedestrian in near field," *Progr. Electromagn. Res. C*, vol. 100, pp. 145–159, 2020.

[32] *Unity Real-Time Development Platform | 3D, 2D VR & AR Engine*. Accessed: Feb. 20, 2023. [Online]. Available: https://unity.com/

[33] Autodesk. *Autodesk MotionBuilder Software*. Accessed: Jun. 21, 2023. [Online]. Available: https://www.autodesk.com/products/motionbuilder/overview?term=1-YEAR&tab=subscription

[34] S. Ahmed, W. Kim, J. Park, and S. H. Cho, "Radar-based air-writing gesture recognition using a novel multistream CNN approach," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23869–23880, Dec. 2022.

[35] S. Waqar, M. Muaaz, and M. Pätzold, "Direction-independent human activity recognition using a distributed MIMO radar system and deep learning," *IEEE Sensors J.*, vol. 23, no. 20, pp. 24916–24929, 2023, doi: 10.1109/JSEN.2023.3310620.

[36] M. Muaaz, S. Waqar, and M. Pätzold, "Orientation-independent human activity recognition using complementary radio frequency sensing," *Sensors*, vol. 23, no. 13, p. 5810, Jun. 2023, doi: 10.3390/s23135810.

[37] Mixamo. *Get Animated*. Accessed: Feb. 20, 2023. [Online]. Available: https://www.mixamo.com/#/

[38] Qualisys. *Qualisys | Motion Capture Systems*. Accessed: Jun. 21, 2023. [Online]. Available: https://www.qualisys.com/

[39] I. Roupa, M. R. da Silva, F. Marques, S. B. Gonçalves, P. Flores, and M. T. da Silva, "On the modeling of biomechanical systems for human movement analysis: A narrative review," *Arch. Comput. Methods Eng.*, vol. 29, no. 7, pp. 4915–4958, Nov. 2022.

[40] Rokoko. *Smartsuit Pro II—Quality Body Motion Capture in One Simple Mobile Mocap Suit*. Accessed: Jul. 10, 2023. [Online]. Available: https://www.rokoko.com/products/smartsuit-pro

[41] Mixamo. *Mixamo Motion Capture (behind the Scenes)*. Youtube. Accessed: Aug. 24, 2012. [Online]. Available: https://www.youtube.com/watch?v=q7jWMd7ix98

[42] S. Waqar and M. Pätzold, "A simulation-based framework for the design of human activity recognition systems using radar sensors," *IEEE Internet Things J.*, early access, Dec. 19, 2023, doi: 10.1109/JIOT.2023.3344179.

[43] S. Waqar, M. Muaaz, and M. Pätzold, "Human activity signatures captured under different directions using SISO and MIMO radar systems," *Appl. Sci.*, vol. 12, no. 4, p. 1825, Feb. 2022.

[44] S. Sana, S. Waqar, H. Yusaf, M. Waqas, and F. A. Siddiqui, "Software defined digital beam forming processor," in *Proc. 13th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2016, pp. 671–676.

[45] M. Mance, M. Prutki, A. Dujmovic, M. Milošević, V. Vrbanovic-Mijatovic, and D. Mijatovic, "Changes in total body surface area and the distribution of skin surfaces in relation to body mass index," *Burns*, vol. 46, no. 4, pp. 868–875, 2020.

[46] A. Vilfan, "Myosin v passing over Arp2/3 junctions: Branching ratio calculated from the elastic lever arm model," *Biophysical J.*, vol. 94, no. 9, pp. 3405–3412, May 2008.

[47] M. Muaaz, S. Waqar, and M. Pätzold, "Radar-based passive step counter and its comparison with a wrist-worn physical activity tracker," in *Proc. 4th Int. Conf. Intell. Technol. Appl.*, 2022, pp. 259–272.

[48] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series.," in *Proc. KDD Work.*, 1994, vol. 10, no. 16, pp. 359–370.

[49] H. Ide and T. Kurita, "Improvement of learning for CNN with ReLU activation by sparse regularization," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 2684–2691.

[50] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[51] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proc. 23rd Int. Conf. Mach. Learn. (ICML)*, 2006, pp. 233–240.

[52] H. Schutze, C. D. Manning, and P. Raghavan, *Introduction To Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[53] K. Ahuja, Y. Jiang, M. Goel, and C. Harrison, "Vid2Doppler: Synthesizing Doppler radar data from videos for training privacy-preserving activity recognition," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, 2021, pp. 1–10.

[54] H. Khalid, A. Gorji, A. Bourdoux, S. Pollin, and H. Sahli, "Multi-view CNN-LSTM architecture for radar-based human activity recognition," *IEEE Access*, vol. 10, pp. 24509–24519, 2022.

[55] X. Li, Y. He, J. A. Zhang, and X. Jing, "Supervised domain adaptation for few-shot radar-based human activity recognition," *IEEE Sensors J.*, vol. 21, no. 22, pp. 25880–25890, Nov. 2021.

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

[57] L. Qu, Y. Wang, T. Yang, and Y. Sun, "Human activity recognition based on WRGAN-GP-synthesized micro-Doppler spectrograms," *IEEE Sensors J.*, vol. 22, no. 9, pp. 8960–8973, May 2022.

[58] M. M. Rahman, S. Z. Gurbuz, and M. G. Amin, "Physics-aware generative adversarial networks for radar-based human activity recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 3, pp. 2994–3008, Jun. 2023.

[59] Y. Zhao, H. Zhou, S. Lu, Y. Liu, X. An, and Q. Liu, "Human activity recognition based on non-contact radar data and improved PCA method," *Appl. Sci.*, vol. 12, no. 14, p. 7124, 2022.

Sahil Waqar received the bachelor's (Hons.) degree in electrical engineering from the University of Engineering and Technology (UET), Taxila, Pakistan, in 2010, and the master's (Hons.) degree in electrical engineering from the National University of Science and Technology (NUST), Islamabad, Pakistan, in 2013. His master's research was related to probabilistic data association filters.

From 2012 to 2020, he worked on the design and development of digital communication systems, signal processors, radar systems, and embedded systems. Since October 2020, he has been working on the CareWell Project as a Ph.D. Research Fellow at the University of Agder, Grimstad, Norway. His current research interests include machine/deep learning, sensors, multiple-input multiple-output (MIMO) radars, communications, signal processing, and embedded systems.

Muhammad Muaaz received the M.Sc. degree in information and communication systems security from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2012, and the Ph.D. degree in computer science from the Johannes Kepler University Linz, Linz, Austria, in 2017.

From 2018 to 2020, he held a postdoctoral fellowships at the University of Agder, Grimstad, Norway; and the Johannes Kepler University Linz. He is currently working as a Researcher at the University of Agder. His research interests include design and development of secure intelligent systems, information and communication systems security, data science, machine (deep) learning, mobile computing, and physiological and behavioral biometrics.

Stephan Sigg is an Associate Professor at the Department of Communications and Networking, Aalto University, Espoo, Finland. With a background in the design, analysis and optimization of algorithms for distributed and ubiquitous systems, he focuses on sensing systems for environmental perception and usable (perception-based) security. Especially, his work covers proactive computing, distributed adaptive beamforming, context-based secure key generation, and device-free passive activity recognition.

Prof. Stephan is an Editorial Board Member of the *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, the *Elsevier Journal on Computer Communications*, and has been a Guest Editor of the Springer *Personal and Ubiquitous Computing Systems Journal*. He has served on the organizing and technical committees numerous prestigious conferences including IEEE PerCom, ACM Ubicomp, and IEEE ICDCS.

Matthias Pätzold received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from Ruhr University Bochum, Bochum, Germany, in 1985 and 1989, respectively, and the Habilitation degree in communications engineering from Hamburg University of Technology, Hamburg, Germany, in 1998.

From 1990 to 1992, he worked at ANT Nachrichtentechnik GmbH, Backnang, Germany, where he was involved in digital satellite communications. From 1992 to 2001, he was with the Department of Digital Networks, Hamburg University of Technology. Since 2001, he has been a Full Professor of Mobile Communications with the University of Agder, Grimstad, Norway. He has authored several books and numerous technical articles.

Dr. Pätzold has been actively participating in numerous conferences as a TPC Member and a TPC Chair. His publications received 14 best paper awards.