Interactive RF Game Design for Deciphering Real-World Human Motion: Activities, Gestures, and Sign Language

Sevgi Z. Gurbuz[1], Chris Crawford[2], Darrin J. Griffin[3], Emre Kurtoglu[1], Oladipupo Adeoluwa[1], Josh Haeker[1]
[1]*Dept. of Electrical and Computer Engineering,* [2]*Dept. of Computer Science,* [3]*Dept. of Communication Studies*
The University of Alabama, Tuscaloosa, AL / USA
szgurbuz@ua.edu; crawford@cs.ua.edu; djgriffin1@ua.edu; {ekurtoglu, oadeoluwa, jdhaeker}@crimson.ua.edu

***Abstract*—Current methods for acquiring datasets for human motion classification are limited to controlled settings where participants are directed by a human experiment organizer. Datasets acquired in controlled settings often cannot capture natural human behaviors and are inadequate for obtaining large amounts of real-world data in a sustainable fashion. This paper proposes a new paradigm for automated acquisition of natural human movements based on Interactive RF Gaming. The training of AI/ML models is considered over an evolution of time: pre-deployment via physics-aware batch training and post-deployment via continual learning from interactions. Algorithms needed to address real-world considerations such as the parsing of continuous streams of data, computational constraints, real-time processing, and game design considerations based on cyber-physical human system requirements are also discussed.**

***Index Terms*—micro-Doppler, RF sensing, radar, human-computer interaction, human sensing, human activity recognition, gesture recognition, sign language**

I. Introduction

The advent of low-cost, low-power radio frequency (RF) sensors has paved the way for many new commercial applications involving indoor monitoring, telemedicine, automotive perception, and human computer interaction. As technology advances and an increasing number of devices enter our homes and workplace, however, an important aspect of autonomous system design involves human-system collaboration that is optimized through improved understanding of human behavior [1]. RF sensing represents an exciting new frontier in this context, as radar provides an "invisible" mode of perception that can operate in an ambient fashion and is effective when alternative sensors may not. Among advantages of RF sensors includes their ability to operate at night, in the dark, and without acquiring of privacy invasive information, such as facial or background imagery. Unlike wearables, RF sensors cannot be "forgotten" and are a non-intrusive modality. Of course, RF sensors can be utilized in a stand-alone or in a collaborative fashion with other Internet-of-Things (IoT) sensors.

However, accurate and robust machine understanding of human motion with RF sensors has been precluded by several fundamental challenges [2]:

1) ***Human motion has nearly unlimited diversity.*** Yet, current techniques are designed to recognize just a small number (10-20) of pre-defined classes. Thus, it is almost certain that the sensor will encounter motions other than the pre-defined classes. Lumping all open-set data into a single "unknown" class may prevent some confusions but does not address the fact that huge numbers of motions will remain unidentified.
2) ***Physical factors can distort, block, or mask patterns seen in the RF data.*** RF sensors measure radial, not absolute velocity, so the path of motion and angle of observation modulate the amplitude of micro-Doppler frequencies measured, resulting in mismatch with training data that is acquired by participants moving along a straight line and reducing classification accuracy to just 5% in some cases [3]. Also, obstructions and the presence of other motion sources, such as other people, pets, or ceiling fans, can result in missing components or intertwined signatures. But current techniques cannot effectively decouple such returns, preventing correct classification. Most studies consider only straight, linear paths, without any obstruction or masking.
3) ***Human data is a time-varying, continuous stream of sequential motion.*** Most current machine learning (ML) approaches convert the complex time-stream of RF data into 2D real images representing the micro-Doppler signature (mDS) of the data over a finite duration window. But, this approach forces the RF data to conform to structure of deep neural networks (DNNs) originally designed for computer vision applications, and is not well suited to the processing of sequential data. Although there have been some works [4]–[8] investigating recurrent neural networks (RNNs) and their variants, these do not account for the transition period between motion classes. The patterns and duration of the transitions depend on the classes involved and constraints imposed by biomechanics. Thus, sequential classification remains an open problem.
4) ***Environmental factors can also vary with time.*** Life is dynamic: not just the person of interest, but other vehicles, people, animals, and motion sources also move. Obstructions and masking are typically transient, not

This work was funded in part by NSF CNS/CPS Award #1932547, AFOSR Award #FA9550-22-1-0384 and the American Association of University Women (AAUW).

persistent, causes for degradation. Sources of RF interference can block the receive frequency band and corrupt the received human return. Most current techniques in the literature are effective only under ideal conditions, and do not account for dynamically changing real-world environmental factors.

5) ***Insufficient Training Data for Conventional DNNs***: The training of deep models requires a lot of data that spans all probable subject profiles; e.g., people of all heights and walking styles – a typically unfeasible task, resulting in most RF studies using only several thousand samples. This in turn severely limits DNN depth, and, hence, the accuracy and generalization of resulting models to real-world conditions.

Among all these challenges, the issue of how to most effectively train AI/ML models with limited data remains the most critical. The problem is not just the *quantity* of data, but the ***quality*** of data. Ideally, we would want to obtain data representative of every single variant of human movement, acquired under every possible antenna-target geometry and environment, for as many different people as possible. This is quite a daunting and impractical task, especially considering the current way in which RF human subject data is acquired: namely, measurements conducted in controlled settings under direction of a study organizer, under highly restrictive conditions, and with just a few participants.

This approach to acquiring human subjects data is neither realistic nor sustainable. It is not sustainable because it requires continually time and effort on behalf of the study organizer to direct the experiments. And it is not realistic because controlled experiments do not capture the "natural flow of life." If you tell someone to do a certain activity, it will be cognizantly articulated. The same person doing the same activity at home during the day without thinking about it will move differently. Moreover, *when people know that they are being observed, they tend to behave differently.* This is a common challenge for in-hospital quantitative gait analysis methods, which utilize force plates and cameras, to monitor and assess gait: doctor assessments can be skewed by the patient trying harder to walk well, when in fact they may experience more instabilities walking at home not thinking about what they are doing.

This is effect is even more pronounced when we consider gesture and sign language recognition, and the initial and subsequent positions of the hands effect the way the gesture/sign actually looks in the time-frequency domain. This phenomenon is referred to as *co-articulation* and is a significant factor that degrades recognition accuracy when training data is recorded for specific starting and ending positions.

The current state-of-the-art in radar micro-Doppler based human movement recognition has shown that AI/ML models can recognize about a dozen distinct activities with high accuracy [2]. Future applications rely on the recognition of more subtle changes and sophisticated movements. In applications of remote health monitoring, gait analysis, and fall risk assessment, as well as gesture and sign language

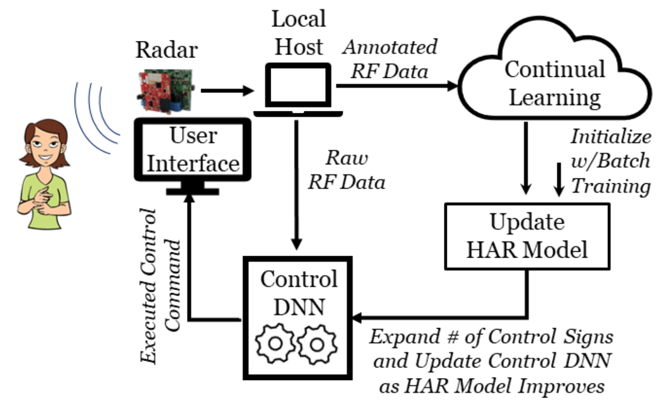

Fig. 1. Flow diagram of the proposed Interactive RF Gaming paradigm.

recognition, advancing AI/ML models requires the ability to observe natural human movements without the intervention of a human experiment organizer. In prior work [9], an Open Radar Dataset together with a complete, integrated and replicable hardware and software pipeline, which allowed users to not only download existing data, but also to acquire additional data with a compatible format was proposed.

While common datasets are important for comparing and benchmarking algorithms, this paper instead proposes a new paradigm for *Interactive* RF data acquisition. Rather than directing human participants to articulate a particular motion class, human behavior is guided via a software environment and interactions of the participants with this software. The resulting RF data acquired is thus more representative of real-world human motion. This in turn requires and end-to-end processing chain that can temporally segment or parse continuous time streams of data, recognize intervals of movement versus activity, understand which intervals contain the relevant information (movement facilitating interaction), recognize/classify relevant motions, and select what parts of the recordings will be used to update models. Moreover, the InteractiveRF software will need to manage and update the training of the AI/ML model not only starting from pre-deployment but also during the duration of operation via continual learning. This paper discusses methods for addressing each one of these processing and model training stages. Figure 1 illustrates the overall end-to-end implementation of the proposed pipeline.

In Section II, first the problem of parsing continuous RF data streams is considered. An approach for segmenting and recognizing intervals where there is movement is presented. Next, in Section III, methods for overcoming the data scarcity problem are presented. Generative methods and implications for classification and continual learning of real-world human motion is discussed. In Section IV, command line control of an RF sensing system is presented to enable data acquisition without a human operator. This ability is essential to closing-the-loop and integrating the RF sensor with the Interactive RF gaming software, discussed in Section V, that implicitly guides interactions. Section VI discusses conclusions and future work.

II. Parsing Continuous RF Data Streams

A fundamental requirement of any Interactive RF system is the ability to parse continuous RF data streams. Natural human movements are inherently continuous and sequential. Although there has been some work on sequential human activity recognition [6], [10], [11], these works process the entire recording in aggregate after acquisition. To enable real-time processing, data streams must be temporally parsed and only the intervals where there is motion extracted and supplied to a classifier. One approach could be to set a fixed processing interval, and apply sequential recognition on this interval. However, in previous work [12], we found that the sequential recognition accuracy over fixed intervals varied depending on the duration of the interval, and that this duration also drove the latency for actually computing the spectrogram. Reducing the duration by half reduces the spectrogram computation time by half. As the duration shortens, however, the likelihood that a movement will be interrupted increases - such cropping can result in clipping movements so that they are not recognizable. Alternatively, first segmenting the continuous data stream according to intervals of movement versus no-movement can both save in computation time (on average) while preserving the micro-Doppler signature of connected motion sequences.

A variety of methods have been proposed for identification of motion intervals. While thresholding of the power burst curve (PBC) has been proposed [13] for arm movement classification, this method is prone to a high rate of false triggering, especially in the presence of noise, because the threshold is not adaptive. A dual windowing approach, using short-time and long-time duration windows, can be used to mitigate false triggering somewhat, but the fixed duration of the windows still prevents sufficient adaptation to the dynamics of the motion. While long windows may be suitable to detect slow movements, performance can degrade significantly when attempting to segment rapidly changing activities, such as gestural communications.

Instead, we found that utilizing a variable window short-time average over long-time average (VW-STA/LTA) technique on the absolute distance between the upper and lower envelopes of the micro-Doppler signature achieved much improved results in determining the starting and ending time of motion. The STA(t) and LTA(t) are the leading and lagging windows defined at time t for each recording i, and may be expressed as

$$STA(t) = \frac{1}{T_1} \sum_{k=t+1}^{t+T_1} v_i(k), \; LTA(t) = \frac{1}{T_2} \sum_{k=t-T_2+1}^{t} v_i(k) \tag{1}$$

where T_1 and T_2 are the lengths of short and long windows respectively. The starting point of a motion is detected when the following conditions are satisfied:

$$STA(t) > \sigma_1 \quad \text{and} \quad \frac{STA(t)}{LTA(t)} > \sigma_2 \tag{2}$$

where σ_1 and σ_2 are predefined detection thresholds. Similarly, the ending point is detected if

$$STA(t) < \sigma_3 \quad \text{and} \quad \frac{STA(t)}{LTA(t)} < \sigma_2 \tag{3}$$

where σ_3 is the detection threshold for the stopping point.

The segmentation accuracy for the identification of motion detected intervals (MDI) was evaluated by comparing VW-STA/LTA based segmentation with ground truth generated by a human analyst. We found that the proposed approach maintained a consistent segmentation accuracy of about 85%, whereas STA/LTA applied with fixed window length has a maximum 76% accuracy with the window durations of 2.3 seconds, dropping to below 40% as shorter durations were utilized. Thus, the proposed approach offered a more consistent and robust segmentation method to extract MDI.

Once the MDI are extracted, any classifier of choice may be utilized to recognize the sequence of activities contained within. Approaches typically involve utilization of a recurrent neural network (RNN), such as a Long Short-Term Memory (LSTM) network, either alone or in combination with a Convolutional Neural Network (CNN). Although micro-Doppler signatures offer a richer source of information relative to range-Doppler or range-Angle maps, and are hence a better basis for classification, our recent work [12] has shown that the performance of a joint-domain multi-input multi-task learning (JD-MIMTL) approach surpasses that of using micro-Doppler only.

III. Learning Effective Models

A. Generative Batch Pre-Training

A critical limitation to any supervised classification approach, however, is the availability of a sufficient amount of diverse data to train deep neural networks (DNNs) for classification. Human subject experiments are time-consuming and expensive, while the diversity in the subjects and scenarios for which the data is collected is necessarily limited and cannot span all possible positions, movements, and environments.

Consequently, there has been much work on micro-Doppler signature synthesis [14]–[18]. These works may be grouped according to two principle approaches: (1) synthesis by taking the time-frequency transform of the received radar return computed from a skeletal model comprised of point targets animated using motion capture (MOCAP) data, and (2) direct synthesis of the micro-Doppler signature using generative adversarial network (GANs) trained from a small number of measured signatures. Model-based synthesis from MOCAP has the advantage of allowing estimation of the resulting target micro-Doppler for any desired antenna-target geometry. Although the MOCAP data itself is still subject-specific data, a diversification technique [19] that applies data augmentation to the underlying skeleton has allowed for relatively few MOCAP measurements to be used to generate thousands of statistically independent animations, which better span the range of expected human profiles. However, one major drawback is that model-based synthesis does not account for changes in signal-to-noise ratio, sensor-related artifacts, non-stationary clutter, interference or signal dispersion induced by frequency-dependent barriers such as walls. In contrast, GANs have been shown effective in modeling sensor imperfections, noise and clutter in radar human activity data synthesis. A wide range of GANs have been utilized in synthesizing radar micro-Doppler signatures for human motion recognition.

However, previous research [20] has revealed that GAN-generated RF mDS exhibit systemic flaws in generation of target kinematics, which correspond to physically impossible features in the synthetic data. Examples of some of these kinematic flaws include disjoint components, malformed shapes, inconsistencies in peak values, subdued regions, and additional non-zero micro-Doppler components that make the signature resemble a different activity class altogether. A hard, impulsive fall may instead resemble a slower progressive fall. A walking signature may include a period over which the person is actually stopped and not moving at all. These kinematic aberrations can significantly degrade classification accuracy when the synthetic data is used for training.

In contrast, physics-aware GANs [21] integrate the domain knowledge of human kinematics into the design of the GAN architecture and loss functions, thereby improving the accuracy with which human mDS are synthesized. Because the envelope constrains the maximum velocity incurred during motion and differences between human gaits is captured by the envelope, it is essential that the process for generating synthetic samples consistently and realistically replicates the envelopes characteristics of ambulatory classes. Gross kinematic errors can be precluded by supplying the micro-Doppler envelopes as inputs to additional branches in the discriminator and utilizing an additional physics-based loss term in the GAN loss function. If the envelope is considered as a time-series or a curve, the choice of the distance metric is tied to the ability of the metric to produce a significant quantitative difference between the two envelopes based on how dissimilar/similar they are. One way of quantifying the similarity between envelopes is the Dynamic Time Warping (DTW) distance. In this case, the new loss function is

$$Loss = \{D(x) - D(G(z))\} + GP + \lambda_P L_{physics}, \tag{4}$$

where GP is the gradient penalty, $L_{physics}$ is computed from the DTW, $D(x)$ is the discriminator's estimate of the probability that real data instance x is real; $G(z)$ is the generator's output when given noise z; $D(G(z))$ is the discriminator's estimate of the probability that a fake instance is real.

B. *Continual Learning via Interactive RF*

Studies conducted for both activity recognition [21] and sign language [22] have shown improvements in the kinematic fidelity of synthetic human signatures using PhGAN architectures, and, hence, improvements in human motion recognition accuracy. However, regardless of how well the signatures for a given motion class are synthesized, there remains the fundamental limitation that any synthetic training database will be comprised of a limited number of motion classes. Given the nearly boundless range in human movements, it is likely that in real-world scenarios an open-set problem will be encountered: in other words, our test data will not belong to the same class as our training data. Moreover, when RF systems are deployed in the real world, the test subjects will always be different from those utilized in training. While data synthesis significantly helps in improving the ability of a model to generalize, the differences in the received signal due to subject-specific and environment-specific shifts in the distribution will need to be surmounted for robust performance. Such shifts in distribution will need be bridged by continual learning on data that is interactively acquired over extended periods of time *in situ*.

IV. Interactive RF Game Design

The idea of framing data acquisition of human movements in the context of interactive games is appealing for several key reasons: First, the game itself serves as the interface automatically guiding the types of movements expected in response to the gaming environment. For example, conversations can be designed to facilitate acquisition of natural sign language; embodiment games can be designed requiring whole body movements, such as in sports games, or activity-based games, where the user does construction by virtually moving objects; gesture-based games can be designed to control virtual objects or vehicles. As a long-term vision for the possibilities engendered by this paradigm, consider that virtual reality could be paired so as to trigger responses a virtual environment to prompt desired responses. Second, the game provides an enjoyable experience that will stimulate more human participation. In controlled subjects testing, the experiences is repetitive and mundane. Oftentimes, the participant fees given ($25-$50) are not incentive enough for continued participation, inherently limiting the amount of acquired data. With an interactive game, participants who enjoy the experience will want to continue to interact with the system, providing a steady supply of data that can be used to continually learn more effective models initialized with synthetically generated batch pre-training. In the next sub-sections, details about the implementation of Interactive RF is given.

A. *Command Line Control of Radar*

Texas Instruments (TI) provides a device firmware package for users who want to control their radars using terminal-based applications. These applications utilize the command line interface (CLI) instead of TI's graphical user interface (GUI) to operate the radar. Although the provided GUI is intuitive and mostly self-explanatory, to promote modularity of data acquisition applications, using the CLI as the application backbone becomes essential as the manufacturer GUI is too rigid and introduces restrictions on the functionality of the radar. Using the CLI option, users can initialize the radar chip, begin data capture sequences, and load data into data capture devices without interrupting the higher-level application. Moreover, configuration of the radar parameters, chirp profiles, antenna configurations, and all other configurable parameters can be dynamically changed while the main program is running.

In order to use a TI radar board coupled with a data capture card (e.g., DCA1000 EVM) in the CLI mode, users set the xWRxxxx BOOST radar module to use Serial Peripheral Interface (SPI) Communication. The SPI configuration is achieved with SOP4 mode module(i.e., placing jumpers on the SOP0 pin, while SOP1 and SOP2 pins are left open).

TI provides user examples for different operating systems (OS) to be able to instantiate data collection procedures. These examples are contained in the software development kit (SDK) provided by the manufacturer for users who want to develop their own applications around their application programming interface. These applications can have extended functionality such as creating advanced frames, advanced chirp, continuous mode, dynamic chirp, and dynamic profile configurations. Users can enable or disable these advanced features in an inbuilt configuration file. They perform three key tasks:

- Initialize radar chip.
- Downloads the meta/OS image over serial peripheral interface (SPI).
- Reads the application programming interface (API) parameters from the configuration text file.

After the radar initialization setup is completed, the radar starts transmitting and receiving packets. A secondary application that serves as a packet listener stores the packets in a data capture card. After executing the application, the DCA1000EVM card connected through ethernet accesses the configuration parameters from a custom JSON file which contains information regarding the configuration and data capture parameters of xWRxxxx BOOST device. Four sequences are executed during the data capture application:

- Connection and configuration of DCA1000 to a PC.
- Setting up a user-set packet delay.
- Initiating packet recording.
- Terminating the data capture.

B. Closing-the-Loop: User Interface

Using the CLI tools, a user can design their own higher level application to interact with the radar system, such as a simple interactive game. However, our goal is to replace keyboard commands with gestural or sign language based commands. With simple games, such as that described in Section IV-C, the actual data we wish to record are not the command signs, but activities or signs prompted during the course of the game. Thus, there are actually two DNN models utilized in the Interactive RF paradigm: one model for the control DNN, which operates on raw RF data via a complex neural network (e.g. Complex SincNet [23]) to provide real-time predictions of a low set of command signs, and the actual model we wish to train long-term based on the data acquired during the game. The long term model may take as input various RF data representations, as desired, and over time will increasingly become more and more accurate via utilization of continual learning to update the model based on new annotated data acquired via the Interactive RF game. The initial model, batch trained for 140 ASL signs using fluent signing data and PhGAN [21] synthesized signatures (accessible from *http://github.com/ci4r/ASL_Dataset*), yields a CNN-based prediction of over 65%, 82% and 86% testing accuracies for Top-1, Top-2 and Top-3 accuracies, respectively.

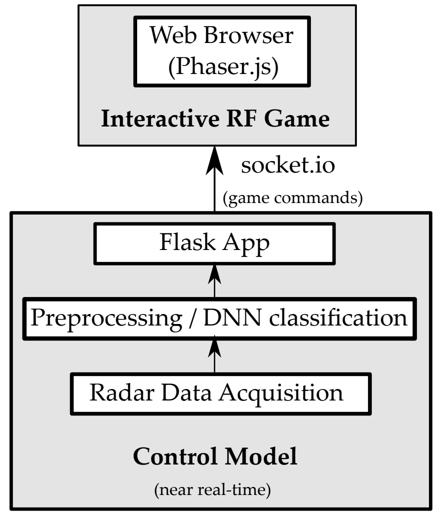

Fig. 2. Control model processing pipeline for end-to-end game design.

When utilizing mDS spectrograms, computationally the application pipeline can be divided into four phases: data acquisition, raw data reading, spectrogram generation and prediction. While data acquisition and saving is an instant step with no delay, reading of raw data and extracting the radar data cube (RDC) takes 0.2 sec for a 3 seconds recording. RDC extraction time is proportional to the data recording duration. Next, generation of a high resolution spectrogram with 4096 Fast Fourier Transform (FFT) points and windowing overlap factor of 90% takes around 0.7 sec which can also be reduced drastically by reducing the overlap factor or the number of FFT points. These are mostly application dependent hyperparameters that user can optimize according to their needs. Finally, the prediction for a sample spectrogram takes 1.4 sec and 0.03 sec without and with the graphical processing unit (GPU) support. It is evident that the GPU support is essential for low-latency required applications. The GPU used in this work is Nvidia's GTX 1660 Ti. Overall, the total time spent from data acquisition to making a prediction does not exceed 1 sec and can be further reduced if needed. Such radar-based recognition system can be used to enable users to control various applications using their hand and body motions.

C. Connecting Predictions to a Game Environment

Modern web-based technologies were used to translate predictions from the raw radar processing steps to game commands. In particular, Flask [24], socket.IO [25], and Phaser.js [26] were used to build an interactive RF game that operates in near real-time. As shown in Figure 2, processed data from the radar system is passed to a Flask App. Flask is a python web framework that enables a pythonic approach to web application development. This approach enables the configuration of a client/server architecture that allows the

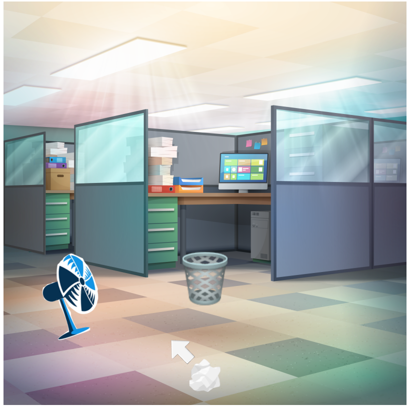

Fig. 3. Example of a game controllable with sign language.

system to take advantage of both python libraries for data processing/classification and JavaScript tools for rapid game prototype development. The game prototype shown in Figure 3 was implemented using Phaser.js. Phaser.js is a lightweight JavaScript game engine that supports the development of 2D games capable of running in a web browser. Socket.IO manages communications between the python server and the web client that displays the interactive RF game.

The game shown in Figure 3 is inspired by the mobile game Paper Toss. The objective of the game is to throw the ball of paper into the trash can. The game also features a fan that influences the flight path of the paper ball. The player is tasked with selecting a flight angle (indicated by the arrow) that will result in the paper ball colliding with the trash can launched. A hybrid approach is currently used to control the game. In particular, the arrow keys are used to update the paper ball flight angle. However, the player must perform a gesture to trigger the launch command. While this game contains limited commands, more complex interactive RF systems could be designed using these core components.

V. Conclusion

This paper proposes an Interactive RF gaming paradigm to address the challenge faced due to inadequate data - both in terms of quantity and quality. Cyber-physical human systems rely on the understanding of complex human behaviors, which can be quite nuanced. Advancement of radar-based AI/ML algorithms requires appropriate data that captures natural human movements. The novel interactive RF gaming paradigm proposed enables the continual acquisition of real-world RF data without the intervention of a human experiment organizer. An example game based on paper toss that is controlled using sign language is presented. In future work, we plan to further develop the gaming environment, continually acquire data, improve our ASL models, and expand interactive functionality as the game is played over an extended period of time by fluent ASL users.

References

[1] Y. Yildiz, "Cyberphysical human systems: An introduction to the special issue," *IEEE Control Systems Magazine*, vol. 40, no. 6, pp. 26–28, 2020.

[2] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 16–28, 2019.

[3] S. Z. Gurbuz, B. F. Erol, B. Cagliyan, and B. Tekeli, "Operational assessment and adaptive selection of micro-doppler features," *Iet Radar Sonar and Navigation*, vol. 9, pp. 1196–1204, 2015.

[4] A. Shrestha, H. Li, J. Le Kernec, and F. Fioranelli, "Continuous human activity classification from fmcw radar with bi-lstm networks," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13 607–13 619, 2020.

[5] 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. on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[6] H. Li, A. Mehul, J. Le Kernec, S. Z. Gurbuz, and F. Fioranelli, "Sequential human gait classification with distributed radar sensor fusion," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7590–7603, 2021.

[7] *Proc. 29th Annual Symp. on User Interface Software and Tech.* New York, NY, USA: Association for Computing Machinery, 2016.

[8] Z. Zhang, Z. Tian, and M. Zhou, "Latern: Dynamic continuous hand gesture recognition using fmcw radar sensor," *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3278–3289, 2018.

[9] D. Gusland, J. M. Christiansen, B. Torvik, F. Fioranelli, S. Z. Gurbuz, and M. Ritchie, "Open radar initiative: Large scale dataset for benchmarking of micro-doppler recognition algorithms," in *2021 IEEE Radar Conference (RadarConf21)*, 2021, pp. 1–6.

[10] H. Abedi, A. Ansariyan, P. P. Morita, J. Boger, A. Wong, and G. Shaker, "Sequential deep learning for in-home activity monitoring using mm-wave fmcw radar," in *IEEE Int. Symp. on Antennas and Prop. and USNC-URSI Radio Science Meeting (APS/URSI)*, 2021, pp. 1499–1500.

[11] E. Kurtoglu, A. C. Gurbuz, E. Malaia, D. Griffin, C. Crawford, and S. Z. Gurbuz, "Sequential classification of asl signs in the context of daily living using rf sensing," in *IEEE Radar Conf.*, 2021, pp. 1–6.

[12] E. Kurtoğlu, A. C. Gurbuz, E. A. Malaia, D. Griffin, C. Crawford, and S. Z. Gurbuz, "Asl trigger recognition in mixed activity/signing sequences for rf sensor-based user interfaces," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 4, pp. 699–712, 2022.

[13] M. G. Amin, Z. Zeng, and T. Shan, "Arm motion classification using curve matching of maximum instantaneous doppler frequency signatures," in *IEEE Int. Radar Conf.*, 2020, pp. 303–308.

[14] S. Z. Gurbuz, *Deep Neural Network Design for Radar Applications*. SciTech Publishing, 2020.

[15] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, "Simulation and analysis of human micro-dopplers in through-wall environments," *IEEE Trans. Geosci. Rem. Sens.*, vol. 48, no. 4, pp. 2015–2023, 2010.

[16] I. Alnujaim, S. S. Ram, D. Oh, and Y. Kim, "Synthesis of micro-doppler signatures of human activities from different aspect angles using generative adversarial networks," *IEEE Access*, vol. 9, pp. 46 422–46 429, 2021.

[17] J. A. Nanzer, "Simulations of the millimeter-wave interferometric signature of walking humans," in *Proc.IEEE Int. Symp. Ant. Prop.*, 2012.

[18] B. Erol and S. Z. Gurbuz, "A kinect-based human micro-doppler simulator," *IEEE Aero. Elec. Sys. Mag.*, vol. 30, no. 5, pp. 6–17, 2015.

[19] M. S. Seyfioglu, B. Erol, S. Z. Gurbuz, and M. G. Amin, "DNN transfer learning from diversified micro-doppler for motion classification," *IEEE Trans. Aerosp. Elec. Sys.*, vol. 55, no. 5, pp. 2164–2180, 2019.

[20] B. Erol, S. Z. Gurbuz, and M. G. Amin, "Motion classification using kinematically sifted acgan-synthesized radar micro-doppler signatures," *IEEE Trans. Aerosp. Elec. Sys.*, vol. 56, no. 4, pp. 3197–3213, 2020.

[21] M. Rahman, S. Gurbuz, and M. Amin, "Physics-aware generative adversarial networks for radar-based human activity recognition," *IEEE Trans. Aerosp. Elec. Sys.*, 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. Elec. Sys.*, vol. 58, no. 4, pp. 2732–2745, 2022.

[23] S. Biswas, C.O. Ayna, S.Z. Gurbuz, and A.C. Gurbuz, "Complex sincnet for more interpretable radar based activity recognition," in *Proc. IEEE Radar Conference*, 2023.

[24] M. Grinberg, *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.", 2018.

[25] (2023) Socket.io. [Online]. Available: http://Socket.IO/

[26] (2023) Phaser.js. [Online]. Available: https://phaser.io/