

# Human Activity Recognition with a 6.5 GHz Reconfigurable Intelligent Surface for Wi-Fi 6E

Nuno Paulino, *Member, IEEE*, Mariana Oliveira, Francisco Ribeiro, Luís Outeiro, Pedro A. Lopes, Francisco Vilarinho, Sofia Inácio, Luís M. Pessoa, *Member IEEE*.

**Abstract**—Human Activity Recognition (HAR) is the identification and classification of static and dynamic human activities, which find applicability in domains like healthcare, entertainment, security, and cyber-physical systems. Traditional HAR approaches rely on wearable sensors, vision-based systems, or ambient sensing, each with inherent limitations such as privacy concerns or restricted sensing conditions. Recently, Radio Frequency (RF)-based HAR has emerged, relying on the interaction of RF signals with people to infer activities. Reconfigurable Intelligent Surfaces (RISs) offers significant potential in this domain by enabling dynamic control over the wireless environment, thus enhancing the information extracted from RF signals.

We present an Hand Gesture Recognition (HGR) approach that employs our own 6.5 GHz RIS design to manipulate the RF medium in an area of interest. We validate the capability of our RIS to control the medium by characterizing its steering response, and further we gather and publish a dataset for HGR classification for three different hand gestures. By employing two Convolutional Neural Networks (CNNs) models trained on data gathered under random and optimized RIS configuration sequences, we achieved classification accuracies exceeding 90%.

**Index Terms**—reconfigurable intelligent surface, reflect-array, WiFi-6E, antennas, beamsteering, RF sensing, human activity classification

## I. INTRODUCTION

HAR is a growing field dedicated to identifying and classifying all forms of human activity, encompassing both static and dynamic motion [1]. Figure 1 illustrates different types of human activity, which include static activities, dynamic activities, and activities involving postural transitions [2]. Static activities include lying, sitting, and standing, which are generally easier to identify but can be confused with similar postures. Dynamic activities involve continuous movement, such as walking or running, which require distinguishing between different motion patterns. Activities with postural transitions, such as sitting to standing or walking to jogging, involve shifts between static and dynamic states and can be particularly complex due to overlapping motion patterns.

The capability of performing accurate HAR has the potential to significantly impact various domains [3]. In healthcare, HAR can assist in monitoring patient conditions and detecting early signs of medical issues [4]. In smart environments, HAR can enhance home assistance services to improve quality of life and autonomy [5]. In sports, it can contribute to

All authors are with INESC TEC (Institute for Systems and Computer Engineering, Technology and Science) and with the Faculty of Engineering, University of Porto, (Portugal). e-mail: {nuno.m.paulino, mariana.s.fonseca, francisco.m.ribeiro, luis.outeiro, pedro.a.lopes, francisco.g.vilarinho, sofia.i.inacio, luis.m.pessoa}@inesctec.pt

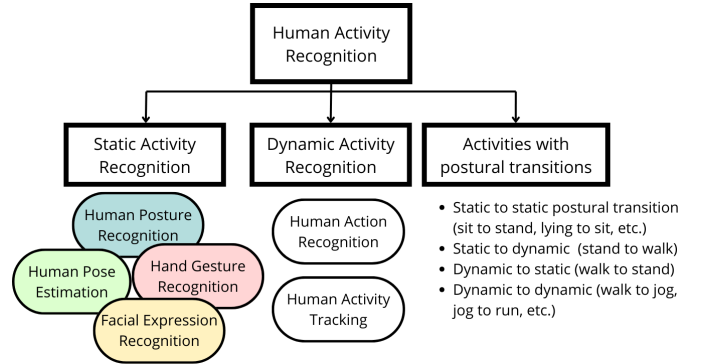


Fig. 1. Types of Human Activity Recognition and respective examples.

improving personal health and well-being [6]. In security, HAR can be used for surveillance to identify unusual events [7]. Additionally, HAR improves human-computer interaction (HCI) by enabling more intuitive and responsive interfaces [8].

However, HAR is complex, requiring selecting and deploying suitable sensors to collect information-rich data, pre-processing the data, and extracting relevant features. Afterward, a range of approaches can be considered to process the data, including machine learning and deep learning algorithms to recognize activities [9]. Regarding specific sensor types, HAR approaches may be vision-based, through monocular or stereo cameras [10], based on sensors such as accelerometers, which can be worn on the body or integrated into smartphones or smartwatches to capture direct movement data [11]. Additional sensors can provide secondary parameters like temperature and humidity, and provide further contextual information [12].

Another emergent method for HAR is the use of RF characteristics of the environment, where data relative to the interaction of stray or controllable RF signals with objects or people within a given Space-of-Interest (SoI) is used as the sensor component of the system. By relying on RF signals, activities can be detected or classified through the effects of body movements on the propagation of these signals, allowing for HAR techniques that address concerns about monitoring and privacy [13]. It does not require wearable or environmental sensors, providing a noninvasive and lower cost approach, instead exploiting the already existing plethora of RF signals in the environment that originate from ubiquitous commodity networks like 5G or WiFi. Since reflected RF signals are also captured by receiving antennas, information

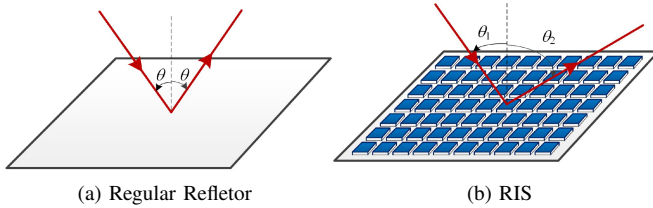


Fig. 2. Regular reflector vs. RIS, which achieves non-specular reflections through programmable phase shifts (adapted from [19]).

can be gathered even in Non-Line-of-Sight (NLoS) conditions [14], and depending on the signal frequency, even through-wall detection can be considered [15], [16]. Finally, the use of higher frequency signals can even provide a high spatial resolution, which has demonstrated advantages over vision-based systems for multi-person recognition [17].

Despite these advantages, wireless environments are complex and difficult to model. Not only is there no direct control over the stray RF signals in a region of interest, but there is also the complexity of multi-path interference, caused when signals reflect off different surfaces and reach the receiver through various paths [18]. Thus the difficulty of clearly relating and modeling the parameters of received RF signals to the status of the environment derives from a largely uncontrolled propagation medium.

Reconfigurable Intelligent Surfaces (RISs) are a promising solution to achieve this medium control. They are composed by a 2D matrix of antenna elements, also called unit cells. These cells contain no active RF chains and act only as reflectors. Simple electronics are used to change the reflection coefficient of each individual unit cell, and consequently a RIS can adjust the surface impedance to produce specific phase shifts for each of the reflected replicas of the incident signal, thus achieving control of the wireless environment [20], [21].

For instance, whereas a specular reflector like a metal surface shown in Figure 2a deflects a signal at an angle equal to the angle of incidence, specific control patterns applied to the unit cells of a RIS can be used to generate constructive interference to focus the total reflection into a particular direction, achieving non-specular reflections, as shown in Figure 2b.

The advantages for communication systems are increased spectral efficiency and power consumption, by reducing radiation towards unintended directions, and support for NLoS communications, since links can be redirected around obstacles. But applications go beyond communication, as this medium control places RISs as a candidate technology to also enable applications like localization, imaging, or sensing [22], [23], [24], [20], [25], [26], in emergent networks like 6G [27].

Concurrently, the WiFi-6E standard has recently been defined, promising to deliver a throughput of up to 9.6 Gbps by extending WiFi capability into the 5.925 GHz to 7.125 GHz range. This expansion allows for up to seven 160 MHz channels, compared to the two channels available in WiFi 5 and WiFi 6, or alternatively up to 59 channels with a bandwidth of 20 MHz. Despite this, this frequency range suffers from greater path loss relative to previous WiFi standards, and also greater attenuation through walls or objects,

compromising NLoS operations. This introduces the need for signal repeaters or greater transmission power.

Considering the signal manipulation capabilities explained before, namely signal steering, RISs are once again placed as a potential low-cost solution to fulfill the expected benefits of WiFi-6E. Moreover, the use of RISs operating in WiFi ranges has already demonstrated the viability of localization [28] and imaging [29] systems. However, to the best of our knowledge, RIS designs for Wi-Fi 6E are under-explored. Given the anticipated ubiquity of Wi-Fi networks and the advantages offered by Wi-Fi 6E, we address this gap by investigating the use of a Wi-Fi 6E RIS for sensing applications.

In this paper, we implement an HAR recognition approach by using a RIS to implement sensing of the medium, allowing for activity classification through RF channel data. Specifically, we will address the classification of static activities, namely HGR, which focuses on detecting hand movements and static hand positions. This technique is relevant for applications like human-computer interaction, virtual reality, and sign language recognition [30].

The resulting contributions are as follows:

- The design and fabrication of a PIN diode based unit cell for a RIS and with a 1-bit phase control, targeting the range of WiFi6-E (for  $f_c = 6.5$  GHz).
- Characterization of the magnitude and phase response of the unit cell in a waveguide environment, and of the radiation pattern of an 8x8 RIS tile in an anechoic chamber for different frequencies and steering angles.
- Modeled the environment and channel for an HGR system, and implemented RIS configuration sequences for sensing the environment through  $S_{21}$  parameters.
- Collected and made publicly available a dataset of  $S_{21}$  data designed for RF-based classification of three human hand gestures.
- Demonstrated the successful classification of gestures using this dataset, for two different CNN architectures.

The remainder of this paper is organized as follows: in Section II we review general aspects of HAR and RIS design and relevant works in both fields; in Section III we explain our approach for HAR – specifically HGR – using our own RIS design, including an explanation of our channel model, RIS configuration scheme, and gesture classification method; in Section IV we present the results of a series of experiments, including characterization of the fabricated RIS, and different sensing setups and dataset gathering, culminating in a CNN based classification of hand gestures through RF data; finally, Section V concludes the paper.

## II. RELATED WORK

Achieving high performance RF sensing is challenging in general, particularly since the areas of interest for sensing, and the RF medium itself, are subject to physical obstacles compromising Line-of-Sight (LoS), and the effects of multipath interference. Functional radio limitations such as the viable amount of information to be carried by transmission channels also compromise the viability of sensing.

To tackle this, some approaches rely on deep learning techniques to extract features directly from RF data without

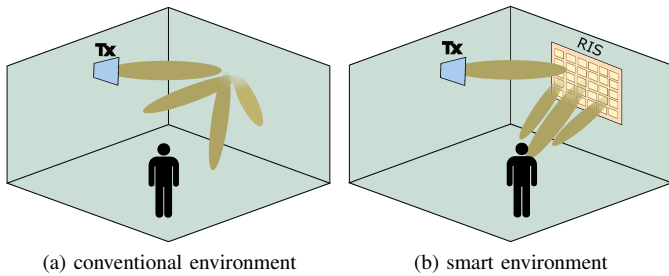


Fig. 3. a) Conventional environment with diffuse signal scattering; b) Reconfigurable environment with RIS enabling targeted signal beamforming.

explicit modeling of complex physical processes, enhancing data processing capabilities [31]. However, the major obstacle in RF sensing is the uncontrollability of the medium. The use of RISs allows for this control to be achieved, as the response of the medium is a function of known applied control.

A RIS can manipulate the propagation of electromagnetic waves, enabling dynamic adjustments of phase shifts and amplitude ratios along propagation paths. By reconfiguring these parameters, RIS creates more favorable channel conditions for RF sensing, optimizing signal transmission and reception [32]. Thus, unlike conventional surfaces, which scatter radio waves based on their material properties and boundary conditions, a RIS can direct reflected signals towards specific locations, enhancing the accuracy of RF sensing systems for human activity detection, as shown in Figure 3.

Given this context, we now review the state-of-the-art from two perspectives. Firstly we present relevant works in the field of RF sensing for Human Activity Recognition (HAR), focusing specifically on RIS based approaches. We also provide a respective summary in Table I at the end of the section, including our own approach. Secondly, we briefly review hardware designs for RIS designs comparable to the design we have employed in this paper.

#### A. RF Sensing through Reconfigurable Intelligent Surfaces

In [26] the authors propose a novel RF sensing system that actively customizes the propagation environment by optimizing the configurations of the RIS. This approach addresses the limitations of conventional RF sensing techniques, which often struggle with multi-path fading and restricted transmission channels. By leveraging the capabilities of the RIS, the system can create multiple independent paths for signal transmission, thereby enriching the information available for accurate posture recognition. The authors employ the Frame Configuration Alternating Optimization (FCAO) algorithm, which minimizes the mutual coherence of the measurement matrix. This reduction in mutual coherence enhances the system's ability to distinguish between different postures and facilitates this classification task. The experimental setup features a horn antenna that emits signals towards the RIS, while an omnidirectional antenna captures the reflections from the human body, utilizing  $S_{21}$  parameters as raw data to evaluate the system's performance. The effectiveness of their approach is demonstrated through practical experiments, resulting in a

remarkable improvement in recognition accuracy for 4 postures, achieving up to 96.7% with optimized configurations, compared to 82.1% in random configurations and 73.2% in non-configurable environments.

Li *et al.* [33] present a significant advancement in radio-frequency (RF) sensing systems with their intelligent metasurface imager and recognizer, which operates at approximately 2.4 GHz, one of the standard Wi-Fi frequency bands. This system enhances HAR by adaptively manipulating ambient RF signals to capture detailed images of human bodies and recognize gestures and vital signs in real-time. The approach relies on an Artificial Neural Networks (ANNs) to process RF signals, effectively transforming them into silhouette images of individuals by integrating a vision dataset. The authors adapt the configuration of the RIS in real-time, according to the processed RF data and the specific areas of interest identified by the ANNs. This capability allows the system to focus electromagnetic fields on selected body parts while minimizing interference from the surrounding environment and enables the system to distinguish between various hand signs and monitor vital signs, such as respiration, even when individuals are uncooperative or positioned behind obstacles. The experimental setup includes a large-aperture programmable metasurface, transmitting and receiving antennas, and a Vector Network Analyser (VNA) for data acquisition, along with a network of ANNs. This configuration enables the system to distinguish between various hand signs and monitor vital signs, such as respiration, even when individuals are uncooperative or positioned behind obstacles.

In [34] the authors present an approach of electromagnetic sensing utilizing a programmable metasurface for imaging and body gesture recognition tasks at 2.4 GHz. This innovative system shapes waves illuminating the scene and enhances sensing quality by employing multiple coding patterns of the metasurface, resulting in improved resolution in the image reconstruction of body postures. A key contribution of this work is the integration of two deep ANNs: the measurement ANN (m-ANN) for adaptive data acquisition and the reconstruction ANN (r-ANN) for instant data processing. This integrated framework allows for the simultaneous optimization of measurement strategies and data processing schemes, significantly enhancing the efficiency and accuracy of human activity recognition tasks. The experimental setup employs a dual-antenna configuration using horn antennas, where one antenna serves as the transmitter and the other as the receiver. These antennas are connected to a VNA, which captures the  $S_{21}$  parameters as the data.

Wang *et al.* [35] introduce an innovative camera system that integrates programmable electromagnetic metasurfaces with deep learning techniques, establishing a Bayesian inference framework to enhance data acquisition and processing. A key contribution of this work is the development of an intelligent EM metasurface camera capable of visualizing human behaviors behind a 60 cm-thick reinforced concrete wall. This capability is achieved through the use of a large-aperture programmable metasurface, which allows for adaptive data acquisition, and artificial neural networks that facilitate instant data processing. The system operates at a frequency of approx-

TABLE I  
SUMMARY OF RIS-BASED HUMAN ACTIVITY RECOGNITION  
APPROACHES

Approach	Anechoic Env.	Freq. (GHz)	Type of HAR	RIS Config.
Hu <i>et al.</i> [26]	Yes	3.2	Posture Recognition	FCAO algorithm
Li <i>et al.</i> [33]	No (indoor)	2.4	Hand Gesture Recognition, Vital Signs	Gerchberg-Saxton algorithm
Li <i>et al.</i> [34]	No (indoor)	2.4	Posture Recognition	m-ANN
Wang <i>et al.</i> [35]	No (indoor)	2.4	Behavior Monitoring	Random patterns
Li <i>et al.</i> [36]	Yes	2.4	Posture Recognition	Gerchberg-Saxton algorithm
This work	Yes	5.0 to 6.5	Hand Gesture Recognition	FCAO Algorithm

imately 2.4 GHz, demonstrating robust performance in real-world settings, even in visually obstructed environments. The experimental setup includes a low-cost commercial software-defined radio device (Ettus USRP X310), a transmitting antenna, and a three-antenna receiver, and a single controlling personal computer. The camera employs a compressive microwave measurement strategy, utilizing 18 control patterns per image to optimize the trade-off between imaging quality and efficiency. This innovative approach not only enhances the frame rate but also significantly improves the Signal-to-Noise Ratio (SNR) of the acquired data, making it a promising tool for smart community applications and beyond.

Li *et al.* [36] present a digital imaging system based on a reprogrammable metasurface that utilizes machine-learning techniques to optimize real-time image capture. The system employs a two-step training strategy where desirable radiation patterns are first learned using machine learning algorithms, and then the corresponding coding patterns of the metasurface are designed to achieve the desired radiation configuration. This approach enables the generation of optimized measurement modes that facilitate high-quality object recognition and imaging while significantly reducing the number of required measurements. By leveraging a modified Gerchberg-Saxton algorithm [37] for discrete-valued optimization, the system effectively tailors the electromagnetic wave manipulation in real-time, enhancing the accuracy and efficiency of the imaging process. A key contribution of this research is the implementation of an image that can be electronically reprogrammed in real-time, allowing for the generation of measurement modes required for machine learning techniques, such as Principal Component Analysis (PCA). The authors demonstrate that by using a prototype of the imager, it is possible to achieve high-accuracy object classification and image reconstruction, even in dynamic scenarios. The experimental results validate the effectiveness of the system, showing that it can operate with a significantly reduced data acquisition time, representing an important advancement in the field of imaging and recognition. The experimental setup consists of a  $2\text{ m} \times 2\text{ m} \times 2\text{ m}$  anechoic

chamber, equipped with a transmitting horn antenna and a waveguide receiver, with a VNA for  $S_{21}$  data acquisition.

Finally, the last row of Table I lists the approach presented in this paper. Relative to the state-of-the-art works we have reviewed, it is, to the best of our knowledge, the only RF sensing approach based considering the WiFi-6E frequency range, with or without the use of a RIS.

### B. Unit Cell and Reconfigurable Intelligent Surface Designs

In [38] a RIS design for 3.5 GHz is evaluated in a real-world environment. The design is fabricated on 1.52 mm thick F4BT450 substrate and contains 2430 unit cells in a near square layout. A varactor is employed per unit cell to achieve a phase control range of close to  $360^\circ$ . The experimental setup consisted of an RX in non-line-of-sight from the TX and used an SDR implementation to stream data using QPSK modulation. The RIS and RX were placed at three different relative orientations to demonstrate the RIS steering capabilities.

A RIS tuned for 5.8 GHz with  $16 \times 10$  elements is presented in [39]. Using one PIN diode per element a phase difference of  $171^\circ$  between states is achieved. The design uses a single layer Rogers RT 6002 substrate, and the diodes are controlled by shift registers and a micro-controller on an external PCB. The RIS is evaluated outdoors, with a feed horn placed 5 m away on the same horizontal plane as the RIS. An RX horn placed 10 m away is moved in a quarter circle range, verifying that the RIS improves the SNR at the RX by up to 15 dB.

In [40] the authors present an open-source design for a RIS and unit cell tuned for 5.5 GHz, fabricated on FR4 substrate with three layers. Each unit cell is controlled by a low-power RF switch, which sets the unit cell to either a short or open state. An ideal difference of  $180^\circ$  is achieved at the central frequency. The RIS contains 256 elements in a  $16 \times 16$  disposition and is controlled by an additional board assembled to its bottom layer which interfaces via USB or Bluetooth.

In [41] focus is on datasets of sub-6 GHz RIS characterization, highlighting that the lack of available datasets hinders research on codebook generation or optimization. They also make available two of their own datasets collected with a  $10 \times 10$  5.3 GHz RIS on an FR4 substrate with 0.5 mm thickness. The unit cells rely on SP8T RF switch which enables configuring the phase response to eight different values. Using a setup similar to ours, with the RIS and feed horn mounted on a turntable, the steering performance is characterized for a range of  $\pm 90^\circ$  and sub-arrays of elements of several sizes.

## III. IMPLEMENTING RECONFIGURABLE INTELLIGENT SURFACE BASED HAND GESTURE RECOGNITION

Recent work has shown the viability of using  $S_{21}$  parameters as data for HAR-related classification tasks [42], [43], [44]. This includes our own previous work, where we implemented Human Posture Recognition (HPR) using RF sensing [45]. However, we relied only on specular reflections, i.e., without resorting to a RIS. In this paper, we enhance the approach by using our own RIS, which we have designed and fabricated for a central frequency of 6.5 GHz, thus keeping in mind its application for WiFi-6E capable scenarios. The



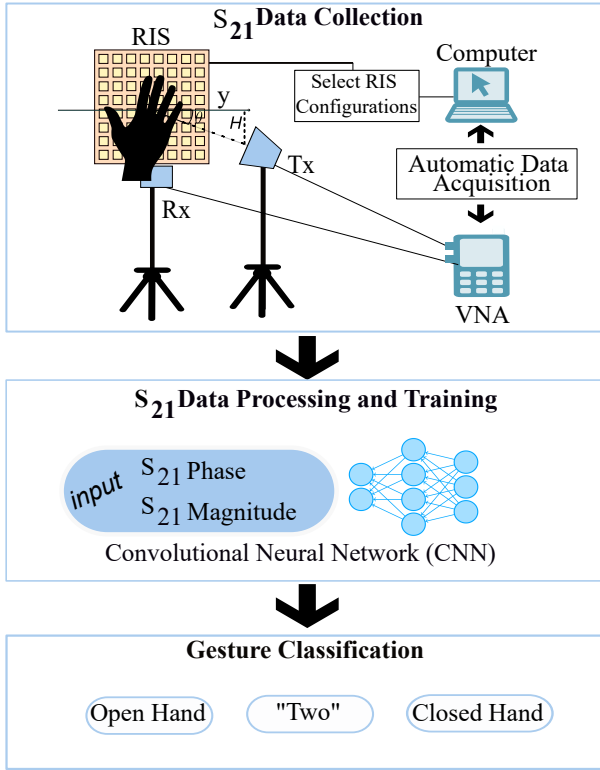


Fig. 4. System architecture diagram illustrating the three main components: Data Acquisition, Data Processing and Training, and Gesture Classification.

RIS enables dynamic and configurable signal paths, providing greater flexibility in RF signal manipulation [46], [47], [48].

Due to the dimensions of our RIS tile ( $18.46 \text{ cm}^2$ ), we focus specifically on Hand Gesture Recognition (HGR). That is, since the dimension of the RIS restricts it to a smaller area of interest, we consider hand gesture recognition a more viable application for this setup. Our approach and setup are illustrated in Figure 4, divided into three stages. Firstly, we design and implement the experimental setup for data collection. This includes the channel between a source antenna Tx, our RIS design, and the target receiver antenna Rx. The SoI, where the human hand whose gesture we wish to classify is present, is placed in front of the RIS. Data is gathered by configuring the RIS to a desired state, triggering the transmission of a signal from the Tx through a VNA, and recording the  $S_{21}$  parameters at the Rx. We gather data for three different hand gestures. To determine the RIS configurations to use, we adopted a strategy similar to the one presented by Hu *et al.* [26], as we explain in detail in Section III-B.

Secondly, we rely on CNNs to interpret RF signals as images, as presented by Li *et al.* [33]. Thus we process the  $S_{21}$  data in order to represent it as image data and train two different CNN architectures. Specifically, we train each architecture with two datasets, one gathered using random sequences of RIS configurations, and another using optimized sequences which increase the information extracted from the channel. Finally, we evaluate the learning rates and classification accuracies of both models for both datasets.

The following sections present the implementation aspects

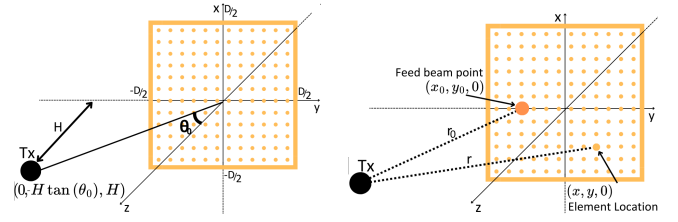


Fig. 5. Geometric configuration between Tx and RIS, considering distance ( $H$ ), offset angle  $\theta_0$ , and location of FBP.

TABLE II  
PARAMETERS FOR HGR EXPERIMENTAL SETUP

Parameter	Formula
Feed Location	$(x_c, y_c, z_c) = (0, -H \tan \theta_0, H)$
Feed Beam Point	$(x_0, y_0, 0)$
Element location	$(x, y, 0)$
$d_{f-FBP}$	$r_0 = \sqrt{(x_0^2 + y_0^2 + H^2 \sec^2(\theta_0) + y_0(2H \tan \theta_0))}$
$d_{f-e}$	$r = \sqrt{(x^2 + y^2 + H^2 \sec^2(\theta_0) + y(2H \tan \theta_0))}$
$d_{FBP-e}$	$s = \sqrt{(x - x_0)^2 + (y - y_0)^2}$

of each component. Firstly, in Section III-A, we explain the model of the physical setup of the approach, as well as the channel model including the RIS. We then explain in Section III-B how sensing can be achieved by then applying a sequence of RIS configuration states over this model, and also explain how we generate an optimized sequence of these states. Finally, we explain the hardware implementation aspects, namely of the unit cell and of the RIS tile Section III-C.

#### A. Modeling and Configuring the HGR Experimental Setup

Our approach requires determining a physical configuration between the RIS, TX, and RX. Proper antenna positioning is crucial to ensure the quality of the transmitted and received signals, and to maximize the efficiency of signal emission and reception. To achieve this, we require a model of the propagation environment between the RIS and SoI. We explain both these aspects in the following sections.

**Antenna Setup and Physical Layout** Regarding antenna selection, we considered canonical or pyramidal horns. Pyramidal horn antennas suit our setup best, as they combine E-plane and H-plane sectoral horns for equal radiation patterns [49]. In contrast, the beam widths in the principal planes E and H of the conical horn are generally unequal [50, p. 15-2], which does not meet the symmetrical characteristics required for our setup. For our analysis, we employed a broadband pyramidal horn antenna, specifically an LB-20180-SF, with a frequency range of 2 GHz to 18 GHz and 12 dBi gain. We chose the same horn for the receiver Rx, to ensure that we capture primarily signals from the SoI, while minimizing any LoS component of the Tx-Rx path.

We then modeled the physical layout between the RIS and Tx antenna, as shown in Figure 5. The center of the RIS is the origin of the coordinate system. The feed antenna Tx is characterized by an offset angle  $\theta_0$  measured between itself and the normal vector to the plane of RIS, and is situated at

a distance  $H$  from the  $xy$ -plane. Finally, the Tx is directed at a given FBP. With this model, we employ the principles regarding aperture efficiency, outlined in [51], in order to establish the ideal distances between components. Specifically, we aimed to achieve an edge taper of  $-10$  dB to ensure that strong diffractions do not occur at the aperture edge, thus minimizing unwanted radiation and diffraction effects that can degrade the quality of the transmitted signal. To facilitate this analysis and given the configuration shown in Figure 5, all the important geometric quantities involved are listed in Table II, based on the equations outlined in [52].

Firstly, we model the radiation of this antenna as a  $\cos^q$  radiation pattern. To determine the  $q$  value, we compared the best fit between datapoints of the radiation pattern of the antenna's datasheet and the  $\cos^q$  model, at our target central frequency of 6.5 GHz, leading to a choice of  $q = 5$ .

Secondly, we considered the aperture's spillover efficiency ( $\eta_s$ ) and its illumination efficiency ( $\eta_i$ ). The spillover efficiency quantifies the fraction of power radiated by the feed that is captured by the reflecting aperture, compared to the total radiated power, and is defined as shown in Equation (1). The numerator calculates the power intercepted by the aperture, while the denominator represents the total radiated power integrated over a spherical surface centered at the feed.

$$\eta_s = \frac{\iint_A P(r) dA}{\iint_{\text{sphere}} P(r) dA} \quad (1)$$

$$\text{where } P(r)dA = \frac{H}{r^3} \left( \frac{r_0^2 + r^2 - s^2}{2r_0r} \right)^{2q} dx dy \quad (2)$$

The illumination efficiency is related to how much of the target surface area is not illuminated, when considering the aperture of the feed and its distance, and is defined by Equation (3).

$$\eta_i = \frac{1}{A} \frac{|\iint_A I(x, y) dA|^2}{\iint_A |I(x, y)|^2 dA} \quad (3)$$

$$\text{where } I(x, y) = \frac{H^{q_e}}{r^{1+q_e}} \left( \frac{r_0^2 + r^2 - s^2}{2r_0r} \right)^q \quad (4)$$

The total aperture efficiency is defined by the product  $\eta_a = \eta_s \cdot \eta_i$ . Given this we optimized the aperture efficiency by performing sweeps over the physical configuration parameters of our experimental setup, considering  $q = 5$  and a conventional value of  $q_e = 1$  for the element pattern power factor. Specifically, we evaluate the efficiencies for a range of  $0^\circ$  to  $50^\circ$  for  $\theta_0$ , 0.2 m to 1.8 m for  $H$ , and  $\pm 15$  cm for  $y_0$ , while keeping  $x_0 = 0$ , and considering the RIS area of  $18.46 \text{ cm}^2$ .

Following this, the optimal values were determined to be  $H = 33 \text{ cm}$ ,  $\theta_0 = 35^\circ$ , and  $FBP = (0, -0.02 \text{ cm}, 0)$ . This setup resulted in about 35% efficiency due to the minimum possible  $H$  in the anechoic chamber. Higher efficiency would have required  $H \approx 16 \text{ cm}$ , but this was not possible due to physical limitations within the anechoic chamber and potential signal interference with the SoI. Finally, the SoI was shifted by  $15^\circ$  to prevent interference from the Tx antenna, and the Rx was placed directly below the RIS, oriented towards the center of the SoI.

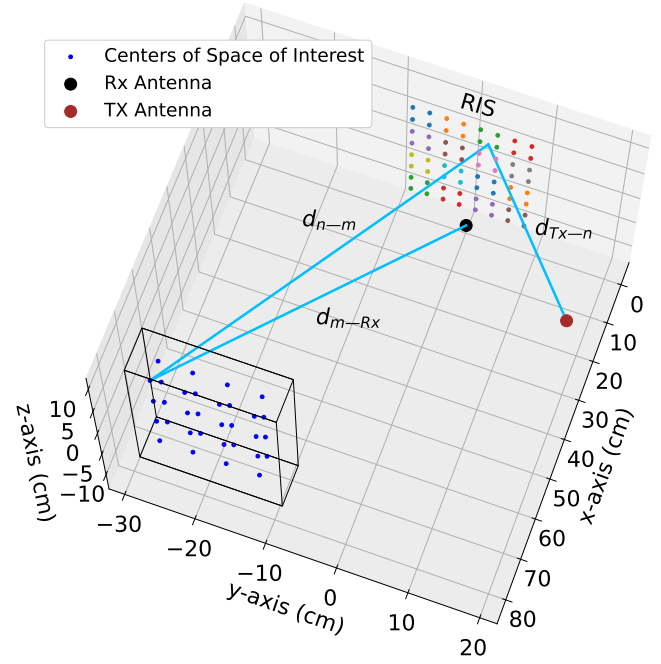


Fig. 6. Top-view of modeled environment for our HGR experiments.

**Tx-to-Rx Propagation Model through RIS and SoI** The setup shown in Figure 6 models the RIS, Space-of-Interest (SoI), and Tx and Rx antennas, at positions given by the analysis explained in the previous section. The RIS consists of  $N \times N$  elements (where  $N = 8$ ), divided into  $L = 16$  square groups, each containing  $N_l = 4$  elements. Each element of the RIS is controlled by a single PIN diode, which can be in two states, ON or OFF. Therefore, each element can switch between two states,  $s_i$ , with  $i \in \{1, 2\}$  and  $N_a = 2$ . For simplicity of control, it is assumed that all elements within the same group  $l$  are in the same state  $s_l$ . The SoI, located in front of the RIS, is divided into  $M = 32$  equal cuboids. Each cuboid is characterized by a reflection coefficient  $\eta_m$ , which is zero when no object is present and non-zero otherwise.

The most significant components are the multiple reflections from the transmitter to RIS and the subsequent propagation of the signal to the SoI, where further reflections occur off the object within the SoI, ultimately captured by the Rx antenna. Thus, considering the states of all groups as the vector  $\mathbf{s}_L$ , the signal at the receiver can be expressed as:

$$y_{rx} = \sum_{m \in [1, M]} \sum_{l \in [1, L]} \sum_{n \in \mathcal{N}_l} h_{n,m}(s_l, \eta_m) \cdot P_t \cdot x \quad (5)$$

where  $P_t$  represents the Tx Power,  $x$  is a unit baseband with frequency  $f_c$  and  $h$  represents the channel gain. Based on [53],  $h$  is represented by:

$$\phi_{n,m} = e^{-j2\pi(d_{nTx} + d_{nm} + d_{mRx})/\lambda}$$

$$h_{n,m}(s_l, \eta_m) = \frac{\lambda \cdot r_{n,m}(s_l) \cdot \eta_m \cdot \sqrt{gT} \cdot gR}{4\pi \cdot d_{nTx} \cdot d_{nm} \cdot d_{mRx}} \cdot \phi_{n,m} \quad (6)$$

Here,  $\lambda$  is the wavelength of the signal, and  $gT$  and  $gR$  denote the gains of the transmitter and the receiver, respectively. The distance from the Tx antenna to the  $n$ -th RIS element is given by  $d_{nTx}$ ,  $d_{nm}$  represents the distance from the  $n$ -th RIS

element in the  $l$ -th group to the  $m$ -th spatial block, and  $d_{mRx}$  is the distance from the  $m$ -th spatial block to the Rx antenna. Finally,  $r_{n,m}$  are the reflection coefficients of the RIS elements in a given state  $s_i$ , and  $\eta_m$  are the reflection coefficients of each cuboid  $m$  in the SoI.

This model demonstrates that the signal at the receiver is a function of the known state of the RIS elements, and unknown conditions of the SoI. The following section explains how we now manipulate these states in order to relate the received signal to these unknown conditions to implement sensing.

### B. RIS Configuration Sequences for Sensing

Given our propagation model, we now present the method to determine a RIS configuration sequence that optimizes the information extracted for the medium. The method protocol operates in a periodic time *frame*, where each *frame* has a duration  $\delta$ . Instead of changing the configurations of all RIS elements simultaneously, each RIS element sequentially transitions between states  $s_1$  and  $s_{N_a}$  within each *frame*.

For the  $l$ -th group of elements, the duration in each state is represented by the vector  $\tilde{t}_l = (\tilde{t}_{l,1}, \dots, \tilde{t}_{l,N_a})^T$ , with  $\sum_{i=1}^{N_a} \tilde{t}_{l,i} = \delta$ . The complete matrix describing the time configuration  $T$ , with dimensions  $K \times (LN_a)$  where  $K = 10$  frames, is defined as:

$$T = (\mathbf{t}_1 \quad \mathbf{t}_2 \quad \dots \quad \mathbf{t}_K)^T \quad (7)$$

where  $t_k = [t_{1,1} \ t_{1,N_a} \ \dots \ t_{L,1} \ t_{L,N_a}]$ . In conclusion, the received signal can be computed as:

$$y_{rx} = P_t \cdot x \cdot T \cdot A \cdot \eta, \quad \text{with } \Gamma = T \cdot A \quad (8)$$

where matrix  $A$  represents the channel gain, with elements denoted by  $\alpha$ . Each element represents the channel gain of the RF paths from Tx to Rx through the  $L$  RIS groups in the different states and the  $M$  spatial blocks, resulting in a dimension of  $LN_a \times M$ :

$$A = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_M]$$

$$\alpha_m = [\alpha_{m,1,1} \ \dots \ \alpha_{m,1,N_a} \ \dots \ \alpha_{m,L,1} \ \dots \ \alpha_{m,L,N_a}]^T$$

$$\alpha_{m,l,i} = \sum_{n \in N_l} \frac{\lambda \cdot r_{n,m}(s_l) \cdot \eta_m \cdot \sqrt{gT} \cdot gR}{4\pi \cdot d_{nTx} \cdot d_{nm} \cdot d_{mRx}} \cdot \phi_{n,m} \quad (9)$$

Thus, the characterization of the SoI, i.e., our RIS-base RF sensing, is based on the received signal  $y_{rx}$ , which can be reconstructed as:

$$y_{rx} = \Gamma \cdot \eta + z \quad (10)$$

where  $\Gamma \in \mathbf{C}^{K \times M}$  represents the measurement matrix,  $z \in \mathbf{C}^K$  denotes additive noise, and  $\eta \in \mathbf{C}^M$ .

In other words, for a measured signal  $y_{rx}$ , and a known sequence of RIS configuration states  $T$ , together with the channel gain of all unit cells, jointly represented by  $A$ , then we can derive the reflection coefficients  $\eta_m$  of each cuboid in the SoI. Thus, sensing of the space is implemented as each reflection coefficient  $\eta_m$  will vary with the conditions (e.g., human posture or gesture) of the space.

However, the sequence of RIS configuration states  $\Gamma$  should be designed such that the most information can be derived

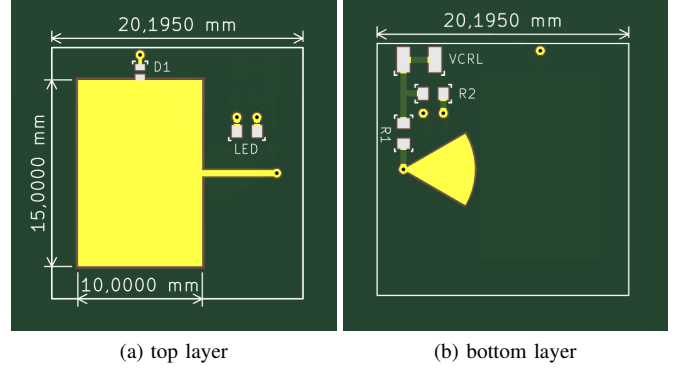


Fig. 7. 6.5 GHz Unit cell design with PIN diode control

from the space. In simple terms, the signal paths between each unit cell, the SoI, and the Tx antenna, should carry the least correlated amount of information possible.

We achieve this optimization of  $\Gamma$  by adapting the method presented by Hu *et al.* [26] to our two-state RIS. Specifically, we adapt their FCAO algorithm, which is proposed for dynamic control of RIS elements, to minimize the mutual coherence of the  $\Gamma$  matrix elements. We then employ the resulting RIS control patterns in our HGR experiments, demonstrating how the FCAO generated patterns extract richer information from the RF medium relative to randomly applied patterns.

### C. Unit Cell and RIS Tile Implementation

In this section we present our design for a 1-bit PIN diode based unit cell, and the 64-element RIS based on this element. All electromagnetic design and simulation was performed through CST Studio. Our objective was to design and test a RIS tile operating the WiFi-6E range and validate its capacity to manipulate the electromagnetic medium for a given range of frequencies centered on 6.5 GHz, providing the hardware necessary for our RF based HGR approach.

**6.5 GHz PIN Diode Based Unit Cell Design** The unit cells were designed to be tuned to for a center frequency of 6.5 GHz, corresponding to the central frequency of the WiFi-6E band. Its' top and bottom layer designs are shown in Figure 7. The top layer contains a rectangular patch of 10 mm by 15 mm. The total enclosing perimeter of the design is a square with a side of 23 mm. Placing two cells side by side provides a separation of half a wavelength. Also shown are the contacts for the D1 PIN diode used to control the phase response. Due to its placement, the unit cell has a vertical polarization. The control voltage for the PIN diode is applied on the bottom layer (Figure 7b), at VCRL. A radial stub is used to decouple the RF path from the DC path.

The PIN diode is an SMP1331-079LF, and the control voltage applied at the *ON* state is of 0.8 V. Given the 220  $\Omega$  value for  $R1$  and the value of 3.3 V applied at VCRL, a unit cell in an *ON* state draws 15 mA. Thus, if all PIN diodes are conducting, the RIS consumes 3.16 W, but the majority of control patterns only activate approximately half the unit cells. Resistor  $R2$  regulates the current for the status LED,



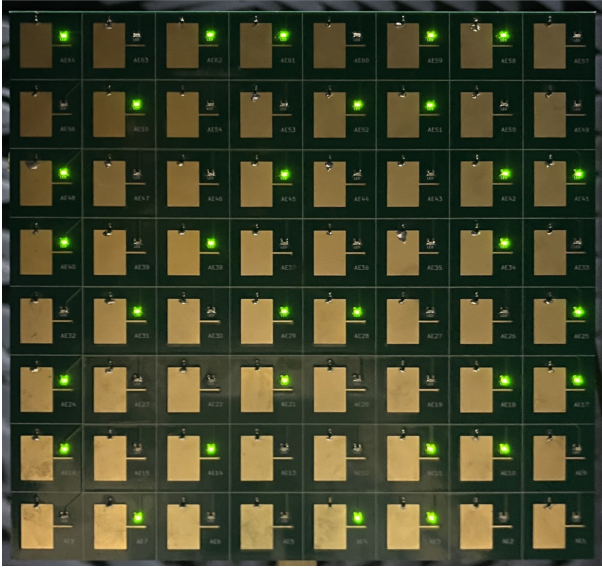


Fig. 8. 64 element RIS tuned for 6.5 GHz, with 1-bit phase control (18.46 cm by 18.46 cm). The status LEDs show some unit cells at the *ON* state.

which exists to allow for easy interpretation and validation of the applied control pattern throughout the RIS. The LEDs can be disabled globally, but each draws 1 mA if active.

The PCB has a composite substrate, where the top layer is a 4 mm F4B substrate, and the bottom layer is a 0.5 mm FR4 substrate. The cost of a tile totaled approximately 85€ for fabrication and approximately 60€ for components. The PIN diodes alone are around 50% of this cost.

**8x8 RIS Tile with 6.5 GHz Unit Cell** The fabricated RIS tile is shown in Figure 8. It is composed of 64 unit cells designed as shown in Figure 7, resulting in a square design with a side of 18.46 cm. In Figure 8 some of the LEDs, one per unit cell, are active, indicating that the respective unit cell is at its *ON* state, which should ideally produce a  $180^\circ$  of phase shift for the signal reflection. A total of 64 signals are needed to control the entire tile.

The tile allows for easy scalability without compromising ideal beamforming conditions. It was purposefully designed such that when physically placing two tiles side by side, with direct contact between the edges of the PCBs, the spacing between elements at the periphery of both tiles equals then spacing within a single tile. A multi-tile RIS can be assembled by resorting to a mounting structure that secures each tile via its bottom layer, as opposed to any frame based mounting which would physically displace the relative tile positions. We achieve this by assembling two SMD 40-pin female pin header connectors on the bottom layer of the tile. These connectors provide the physical mounting support, specifically by interfacing with a smaller control board, shown in Figure 9.

The control board is 11 cm by 7 cm, and hosts eight shift-registers. Namely, eight 74HC595PW ICs, which are 8-bit serial-in parallel-out shift registers. We have connected these in a daisy chain, and to set these shift registers, the control board contains an SPI interface to an ESP32-S3-DevKitC-1. The ESP32 sends 64 bits to the shift-registers, and issues the a

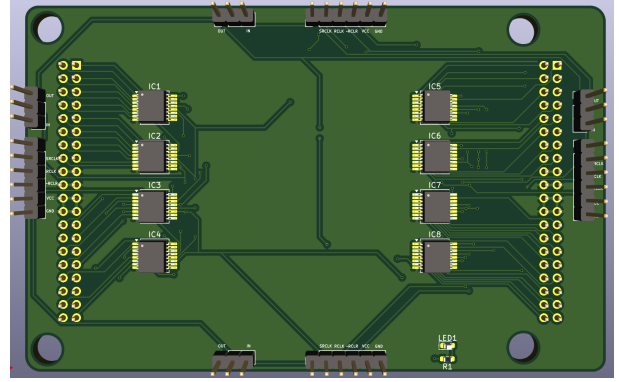


Fig. 9. Top layer of control board hosting shift registers to hold the configuration pattern. The bottom layer contains two 40 male pin-header connectors to interface with the tile.

*latch* signal which changes all unit cell configurations simultaneously. This synchronization prevents undesired intermediate states which would occur if the values were sent directly to the unit cells as their were being shifted into the tile.

We opt to implement this control board separately from the RIS, instead of assembling the shift-registers on its bottom layer, to mitigate any effects on beamforming performance and to simplify the RIS by homogenizing it, which also simplifies scalability. This separation also allows using the same control board for other RIS designs, e.g., for different unit cell designs. The total cost of the control board is approximately 40€.

Finally, besides allowing for the rear mounting which facilitates the physical scaling of the number of tiles, the four interfaces at the control board's periphery can be interconnected to control additional unit cells. For example, a single ESP32 can command the resulting 256 elements of 4 tiles as a single larger RIS, or four parallel daisy chains can allow for a more explicit tile-level control.

#### IV. EXPERIMENTAL EVALUATION

We conducted several experimental campaigns. Firstly, we validate the RIS design, which includes characterizing the magnitude and phase response of the fabricated unit cell versus simulations (Section IV-A), and measuring the beamforming capability of the RIS as a whole (Section IV-B). Once the correct functioning of the RIS was established, we conducted HGR related experiments, namely a preliminary trial to determine the viability of hand gesture recognition with our setup (Section IV-D), followed by a more extensive data gathering to perform the CNN based classification of gestures (Section IV-C).

**Automated Data Acquisition** To ensure fast and consistent execution of all necessary experimental run configuration, data acquisition, and post-processing tasks across experiments, we developed software to interface with our laboratory instruments. Specifically, through a laptop controlling the process, we synchronized processes such as RIS configuration, adjusting relative orientations of Tx and Rx, and measuring  $S_{21}$  data. After setting a particular physical and digital configuration of the system, we resort to a Python-based Virtual



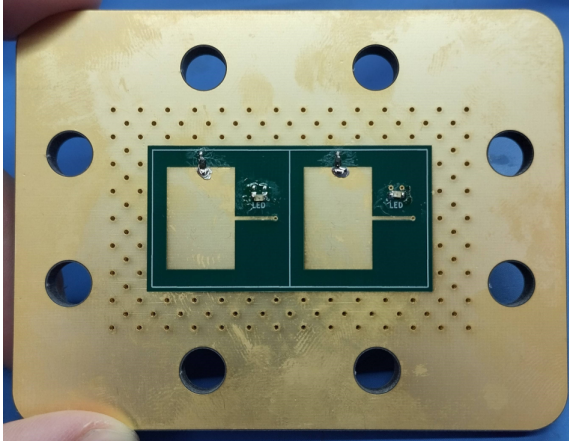


Fig. 10. PCB with two unit cells for validation of frequency and magnitude/phase response in a WR159 waveguide

Instrument Software Architecture (VISA) library to record the corresponding  $S_{21}$  measurements from the VNA. Specifically, we resort to a portable Keysight Fieldfox N9914A VNA, to facilitate the assembly of this data-gathering setup and self-contained experimental setups within our anechoic chamber. The design files and control software for these components are publicly accessible<sup>1</sup>. The following experiments (except for the unit cell characterization via a waveguide) rely on minor variants of this automated flow to gather all measurements.

#### A. Unit Cell Characterization in Waveguide

In this first experiment, we measured the true response of the fabricated unit cell, versus the expected response as given by CST simulation. This verification then determined the experimental parameters for the RIS characterization experiments (Section IV-B), namely the frequency with the best expected steering response, and also influenced the analysis of the HGR results (Sections IV-C and IV-D).

**Experimental Setup** Figure 10 shows the fabricated PCB for use with a WR159 waveguide, with two unit cells spaced identically to the RIS. To measure the response, we applied the same control to both unit cells via an external voltage source. We measured the  $S_{11}$  parameters of the unit cell using a Keysight N5224B, a VNA which operates up to 43 GHz. The simulations considered a Floquet port setup and periodic boundary condition in the 3D electromagnetic simulation software CST, and range from 5.8 GHz to 7.2 GHz, given the intended center frequency of 6.5 GHz. The measurements cover a lower range between 5 GHz to 7 GHz, given the response we observed.

**Experimental Results** Figure 11 shows the resulting  $S_{11}$  measurements and respective simulations, for the *ON* and *OFF* states. Figure 11a shows the magnitude response. At 6.5 GHz the response for both states is approximately equal, which is a desired behavior since the reflected power should not depend

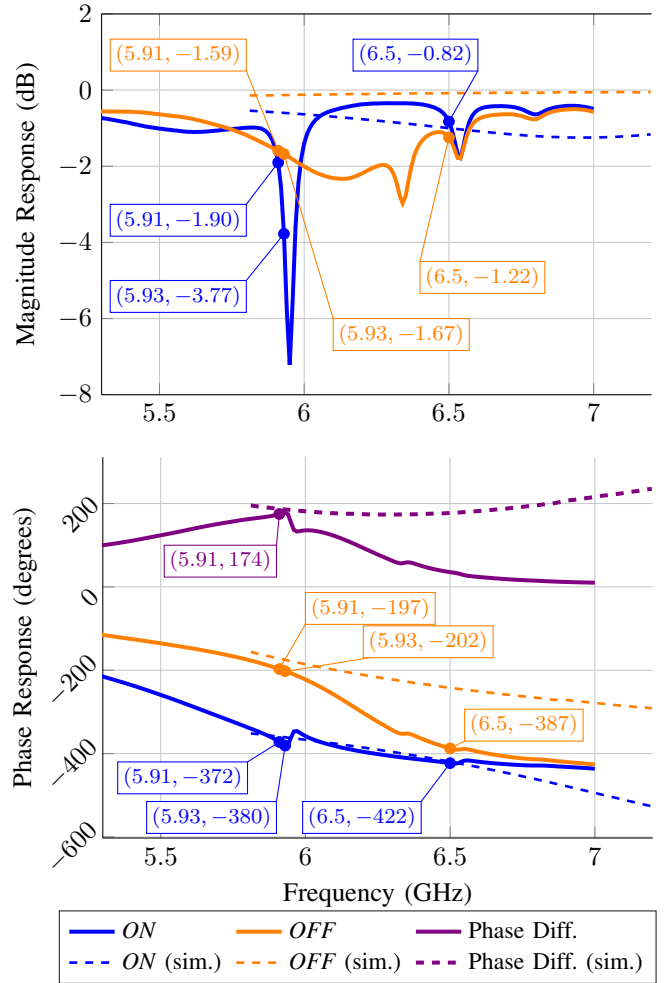


Fig. 11. Magnitude and phase response of unit cell measured in VNA, for both *ON* and *OFF* states. A phase difference closest to  $180^\circ$  is achieved for 5.93 GHz, but a better magnitude response is present at 5.91 GHz.

on the applied control, to avoid any distortions to the desired beamforming. However, the phase difference between states is only  $35^\circ$ , whereas the simulated response was  $180^\circ$ .

Instead, the greatest phase difference of  $178^\circ$  is observed for 5.93 GHz. However, at this frequency the *ON* state experiences a  $-3.77$  dB attenuation. Since the magnitude response of the *OFF* state at this frequency is  $-1.67$  dB, the disparity in magnitude response between states could affect beamforming performance. A trade-off between a balanced magnitude response and phase difference occurs for 5.91 GHz, where a  $174^\circ$  phase difference is attained for a more approximate magnitude response between states, and where the response for either state is not lesser than  $-3$  dB.

This deviation may be due to the LED which was not simulated, or to the tolerances of the PIN diode, whose datasheet reports a typical  $0.18$  pF capacitance up to a maximum of  $0.35$  pF. However, we note that the magnitude response was extremely sensitive to the alignment between the PCB and the waveguide, with the phase response less so. Additionally, it is known that the waveguide environment imposes an oblique wave incidence condition (of approximately  $38^\circ$  in this case), unlike the normal incidence obtained in free space conditions.

<sup>1</sup>F. Ribeiro and M. Oliveira, "SpecRF-Posture". GitHub, Apr. 03, 2024. <https://github.com/franciscombr/SpecRF-Posture>

**Conclusion** It is difficult to determine precisely the source of the deviations from the design, or the true response that is obtained in free space. In the absence of a better reference, for the following tests of the RIS, we consider 5.91 GHz as the central frequency for computing beamforming control patterns.

### B. RIS Characterization in Anechoic Chamber

To validate and characterize the RIS, we measure its beamforming capabilities for several combinations of digital beamsteering and relative orientation between the RIS and a receiver. Following we explain the physical setup in an anechoic chamber of approximately  $7\text{ m} \times 7\text{ m} \times 3\text{ m}$ , and the theory behind the generated RIS control.

**Experimental Setup** Figure 12 shows our setup in an anechoic chamber. We employed the LB-20180-SF horn antenna model identified in Section III-A for both the Rx and Tx, and the mentioned portable VNA to measure the  $S_{21}$  parameters. We resort to our automated data gathering setup, using it to control the RIS digital beamforming, the position of a rotor assembly, and to gather data from the VNA.

We consider that the center of the coordinate system is at the center of the RIS, that the XY plane is the broadside of the RIS, and that the direction of the X axis is downwards. The RIS is mounted on a rotating platform, and its X axis is aligned with the rotational axis of this platform. The feed horn was placed on a platform below the RIS, displaced by 25.5 cm on its X axis, and 36.5 cm on its Z axis, resulting in a distance of 40 cm to the feed point phase center. This distance was chosen as a compromise between the ideal distance of 20 cm, which would provide an edge taper of  $-9\text{ dB}$  (as per Section III-A), and a setup that would reduce the occlusion of the RIS by the feed horn itself. The Rx antenna was placed 170 cm away from the RIS, approximating far field conditions for the measured frequencies.

The RIS control board is attached behind the PCB, containing also the ESP32 micro-controller powered by a power bank, which receives the pattern to apply to the RIS from a laptop via WiFi (on a 2.4 GHz link). This laptop computes the configuration based on the desired steering, according to the following equations.

$$R = \sqrt{(x_c - x_{ij})^2 + (y_c - y_{ij})^2 + z_c^2} \quad (11)$$

$$\phi_{ij} = k(R - \sin(\theta)x_{ij} \cos(\phi) + y_{ij} \sin(\phi)) \quad (12)$$

The coordinates of the feed antenna are given by  $(x_c, y_c, z_c)$ , which for this RIS characterization setup are  $(d \sin 35^\circ, 0, d \cos 35^\circ)$ , given the distance  $d = 400\text{ mm}$  and  $35^\circ$  angle of the feed horn towards the RIS. The coordinates of the center of each unit cell,  $i, j$ , considering that the center of the coordinate system is at the center of the RIS, are given by  $(x_{ij}, y_{ij}) = (iC_d - \frac{H}{2}, jC_d - \frac{W}{2})$ , where  $C_d$  is the dimension of the unit cell, 23 mm, and  $H$  and  $W$  are the height and width of the RIS, both 18.4 cm.

The beam direction is given by  $\theta$  and  $\phi$ ,  $R$  is thus the Euclidean distance between each unit cell and the feed horn,  $k = \frac{2\pi}{\lambda}$  is the wave vector. For these calculations, we considered the wavelength for 5.91 GHz. Finally, given the two

control levels available through the PIN diodes, we quantize the resulting ideal phase responses,  $\phi_{ij}$ , to either zero (for  $0^\circ < \phi_{ij} < 180^\circ$ ) or one (for  $180^\circ \leq \phi_{ij} \leq 360^\circ$ ).

With this setup, and considering a frequency range from 5.3 GHz to 6.5 GHz (the limit of the VNA), we measured the  $S_{21}$  parameters for RIS steering configurations of  $0^\circ$ ,  $\pm 20^\circ$  and  $\pm 40^\circ$ , for a range of rotor orientations of  $\pm 60^\circ$  with a step of  $2^\circ$ .

**Experimental Results** We first retrieved the radiation pattern for RIS configuration of  $\theta = 0^\circ$ , with a rotor orientation of  $0^\circ$  (i.e., forward-facing). Our VNA resolution produced measurements for 200 frequencies, and Figure 13 shows the resulting radiation patterns, normalized to 0 dB.

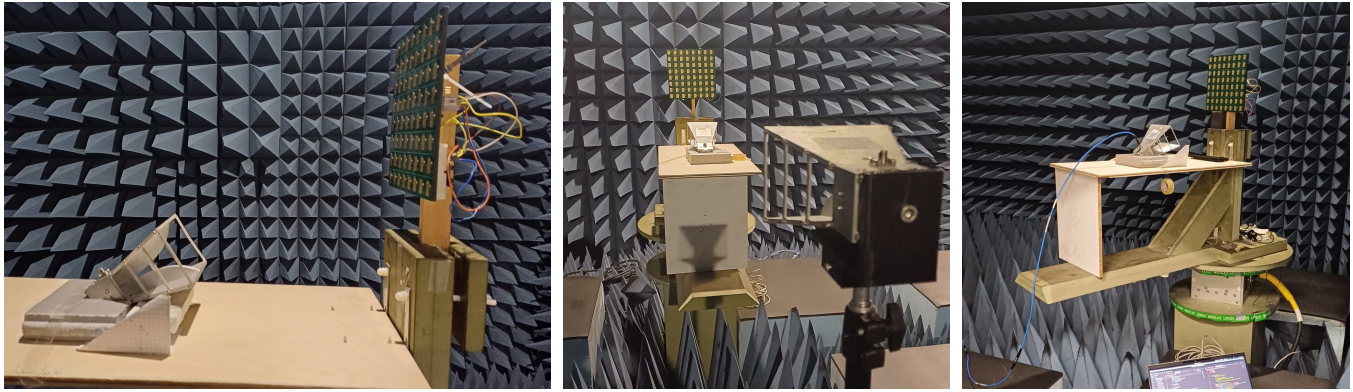
Plotted in light gray are the responses for a subset of all frequencies covering the measured range (specifically, from 5 GHz to 7 GHz with a step of 5.9 MHz). We highlighted the previously chosen frequency of 5.91 GHz, and four additional responses in a bandwidth of  $\pm 59\text{ MHz}$ . Despite the expected best phase response at this frequency, there is no noticeable steering towards the desired direction of  $0^\circ$ . Instead, the shown  $\pm 50\text{ MHz}$  bandwidth centered on 6.16 GHz shows the most response to this steering configuration.

We can attribute this to our measurement of the location of the feed point relative to the RIS center, which is difficult to determine precisely in real-world measurements. The calculation of the phase patterns, shown in Equations (11) and (12), depends on these coordinates as well as the presumed central frequency. So, a likely deviation to the true value of  $(x_c, y_c, z_c)$  resulted in phase profiles more effective for 6.16 GHz despite the value of 5.91 GHz considered for the calculations.

We then measured the radiation patterns for the RIS steering configurations of  $\pm 20^\circ$  and  $\pm 40^\circ$ . Given the previous conclusion, Figure 14 shows these profiles only for 6.16 GHz where a good steering response is confirmed. The RIS displays a symmetric behavior, without angle skew in regards to the steering direction, and with similar magnitude responses for opposing steering angles (e.g.  $\pm 20^\circ$ ).

Despite the good steering response for this frequency, it may not be the best performance achievable by the design. According to the unit cell response for 6.16 GHz seen in Figure 7, the phase difference between states is of approximately  $110^\circ$  and there is a difference of 1.88 dB between states. That is, although the computed phase profiles generated a good steering response at this frequency, more accurate measurements of the feed point or adjustments to its position could result in a better response for 5.91 GHz, exploiting a greater phase difference between states and therefore resulting in a more focused and power efficient beam. Finally, we also plot the radiation pattern measured with the RIS disabled, i.e., acting as a purely specular reflector. This shows that there is little passive component received at the RX, for this frequency, and that the previous profiles are a the result of the beamsteering.

**Conclusion** In summary, the RIS correctly performs beamforming, achieving the best steering response for phase control patterns computed assuming a central frequency of 5.91 GHz.



(a) relative TX to RIS positioning, with TX feed point positioned 400 mm away from center of RIS and a  $35^\circ$  angle (b) RX position relative to RIS mounted on rotor (c) rotor positioned at  $-60^\circ$  relative to its rotation axis, with relative position between TX and RIS maintained

Fig. 12. Setup for RIS data gathering with RIS and TX mounted on a rotational platform, RX antenna on a fixed support. The VNA (not shown) and rotor control are both connected to a laptop to drive automated data gathering.

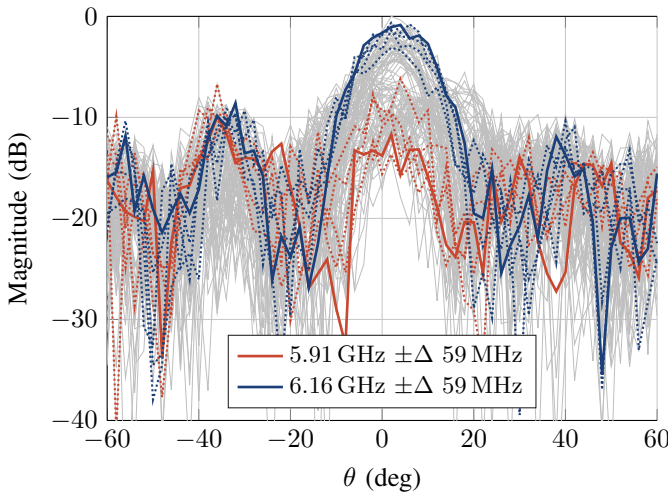


Fig. 13. RIS radiation patterns for a beamforming configuration of  $\theta = 0^\circ$  and  $\phi = 0^\circ$  (i.e., steering towards the normal vector of the RIS plane) and an angle range of  $\pm 60^\circ$ , with 5.91 GHz and 6.16 GHz highlighted.

Although a more pronounced steering response is observed for 6.16 GHz, the RIS does demonstrate steering behavior for a range of 5.0 GHz to 6.5 GHz. This means that, to a greater or lesser extent depending on the frequency, the electromagnetic medium is indeed influenced by the RIS in this broader band. In other words, we can apply phase control patterns generated by the FCAO based optimization approach, and measure the resulting  $S_{21}$  for this range to drive our HGR approach.

### C. RIS-based Hand Gesture Recognition Feasibility Study

The objective of this first experiment was to validate not just the overall experimental setup, but mainly our methodology Hand Gesture Recognition (HGR). Specifically, we aimed to prove that measurements of the  $S_{21}$  parameters, considering the channel through the SoI, reacted to the presence of a human hand to such a degree that different hand gestures could be differentiated.

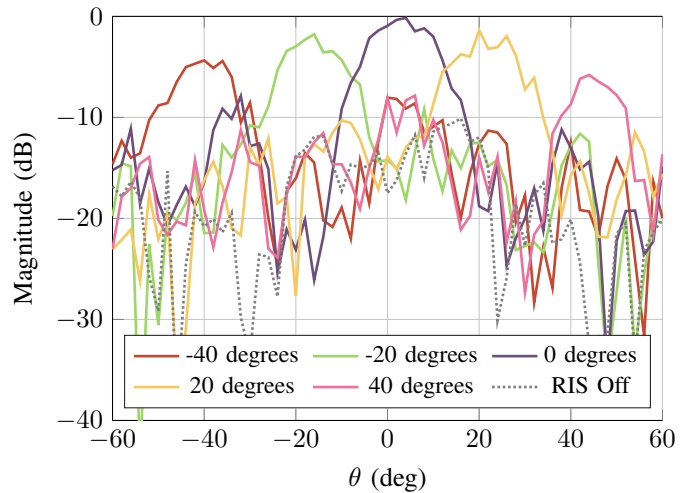
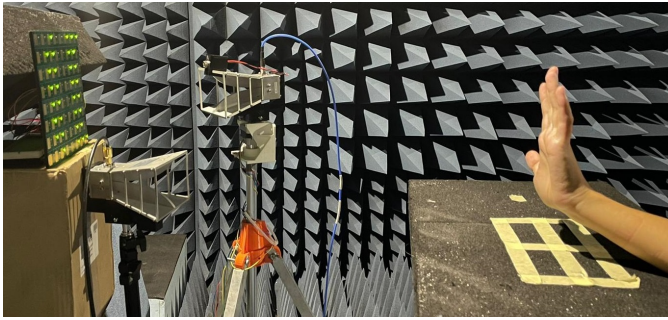


Fig. 14. RIS radiation patterns for six beamforming configurations, for 6.16 GHz, showing an effective steering capability which compensates the rotor positioning. Also shown is the specular response (i.e., RIS off).

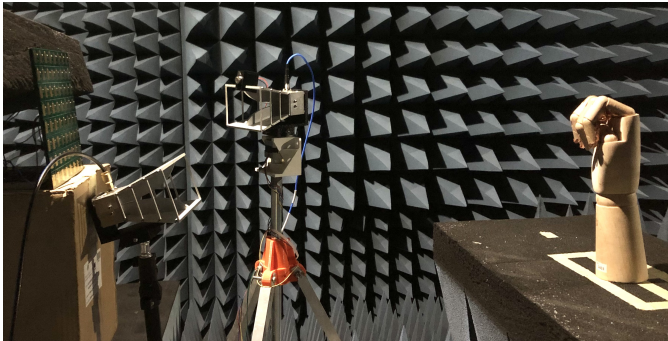
**Experimental Setup** The experimental setup was assembled as we presented in Section III-A, within the same previously mentioned anechoic chamber, resorting to the same pair of horn antennas, i.e., two A-INFO LB-20180 pyramidal horns. Additionally, a ZX60-153LN-S+ low-noise amplifier was added to the setup, to amplify the signal of the Tx antenna.

Following the model developed in Section III-A (shown in Figure 6), Figure 15a illustrates the experimental setup for this experiment. Data was collected for three different hand gestures while the RIS was disabled (*off*), and for four distinct RIS steering configurations, namely  $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ,  $20^\circ$ . Unlike other experiments, the presence of a human volunteer compromises the full automation of data gathering. Instead, for each of the three gestures, data acquisition involved manually saving the  $S_{21}$  parameters from the VNA after directly sending the RIS configuration and assuring the correct positioning of the gesture within the SoI.





(a) side view of experimental setup using real hand (employed in Section IV-C)



(b) side view of experimental setup using model wooden hand (employed in Section IV-D)

Fig. 15. Experimental setup for HGR with (a) real hand (b) wooden hand, inside the anechoic chamber. Data was captured for three different hand gestures, plus an empty SoI.

**Data Gathering** The collected dataset for the three distinct gestures (*open hand*, *two fingers*, and *closed hand*) was conducted through a single volunteer. For each gesture, the hand remained static while the five configurations of the RIS changed. This process was repeated 60 times (i.e., *runs*), resulting in 900 samples (1 subject  $\times$  5 RIS configurations  $\times$  60 *runs*). Additionally, noise measurements were taken without any hand in the SoI. Specifically, ten *runs* were performed for each of the five RIS configurations, resulting in 50 noise measurements. Each sample is characterized by the number of frequency points  $n_{res} = 201$ , between 5.0 GHz and 6.5 GHz.

During data capture, the signal varied slightly, which is expected given the difficulty in keeping the hand perfectly still. Considering that each *run* included changing the configuration and saving the data, each *run* took approximately 15 seconds, resulting in a total acquisition time of about 1 hour and 15 minutes (15 seconds  $\times$  5 RIS configurations  $\times$  60 *runs*), excluding re-setup and rest breaks. This data collection process was quite exhausting, making it impractical to use a real hand for collecting a large dataset required for HGR approach.

**Experimental Results** The obtained dataset was analyzed to observe if there were distinguishable differences between the hand gestures with this setup. Since the signal varied slightly due to the difficulty in maintaining a static hand position, averaging the signal would result in a loss of information. Therefore, the raw signals were analyzed. First, we examined the magnitude of the  $S_{21}$  parameters for each gesture at a specific RIS configuration that steered the signal by 20

degrees. This approach allowed us to observe the variability between different *runs* for the same gesture.

Figure 16 shows the phase response, derived from the respective  $S_{21}$  parameters, for one measurement of each of the four hand gestures. The responses shown were randomly selected from a set of 60 measurements of the respective hand gesture (shown on the right-hand side). We show the phase response for five configurations of the RIS for each gesture. Specifically, *off*, i.e., purely specular, and for steering angles  $0^\circ$ ,  $10^\circ$ ,  $15^\circ$ , and  $20^\circ$ . For brevity, we omit the magnitude response but note that the same distinction in responses between gestures. We illustrate only one measurement per gesture since we observed that there was very little variation between measurements of the same gesture. From the data in Figure 16 it is evident that each gesture produces a distinct phase response, leading to the conclusion that there are clear differences in the received signal at the Rx for each gesture.

**Conclusion** In summary, the  $S_{21}$  parameters measured within the SoI display distinct characteristics per hand gesture, which are consistent between measurements of the same gesture, and therefore the viability of the classification of gestures is demonstrated. Given the high effort of data gathering, we conduct the following experiment using a proxy hand model, instead of a human subject.

#### D. RIS-based Hand Gesture Recognition Classification

Having demonstrated that the  $S_{21}$  parameters within the space of interest display variations with hand gestures, we gathered a larger dataset to allow for CNN training and inference. The following sections present the data gathering setup, illustrate that the FCAO based RIS configuration provides more distinct  $S_{21}$  responses relative to random phase control patterns, and show training and classification results for two different CNN models.

**Experimental Setup** Analogous to the previous experiment, Figure 15b depicts the setup with a wooden hand, where the Rx antenna is slightly tilted upwards (relative to Figure 15a) to focus on the center of the wooden hand. This adjustment ensures more similarity between both setups by isolating the region of interest, i.e., the hand and fingers, which was shown in the previous experiment to produce  $S_{21}$  parameters capable of distinguishing gestures. That is, we wish to exclude potential response components from the wrist that compromise the differentiation in responses we observed in Section IV-C.

Following our RIS reconfiguration method (Section III-B) where configurations are grouped into *frames*, we now consider measurement runs with 10 frames, where each frame contains 39 distinct RIS configurations. These configurations were applied sequentially, resulting in a total of 390 measurements of the transmission coefficients  $S_{21}$  per run. The experiment was conducted twice: once with random configurations and once with the optimized configurations determined by the FCAO algorithm. The data acquisition system was adjusted to synchronize the configuration changes in the RIS with the data collection. For each of the 39



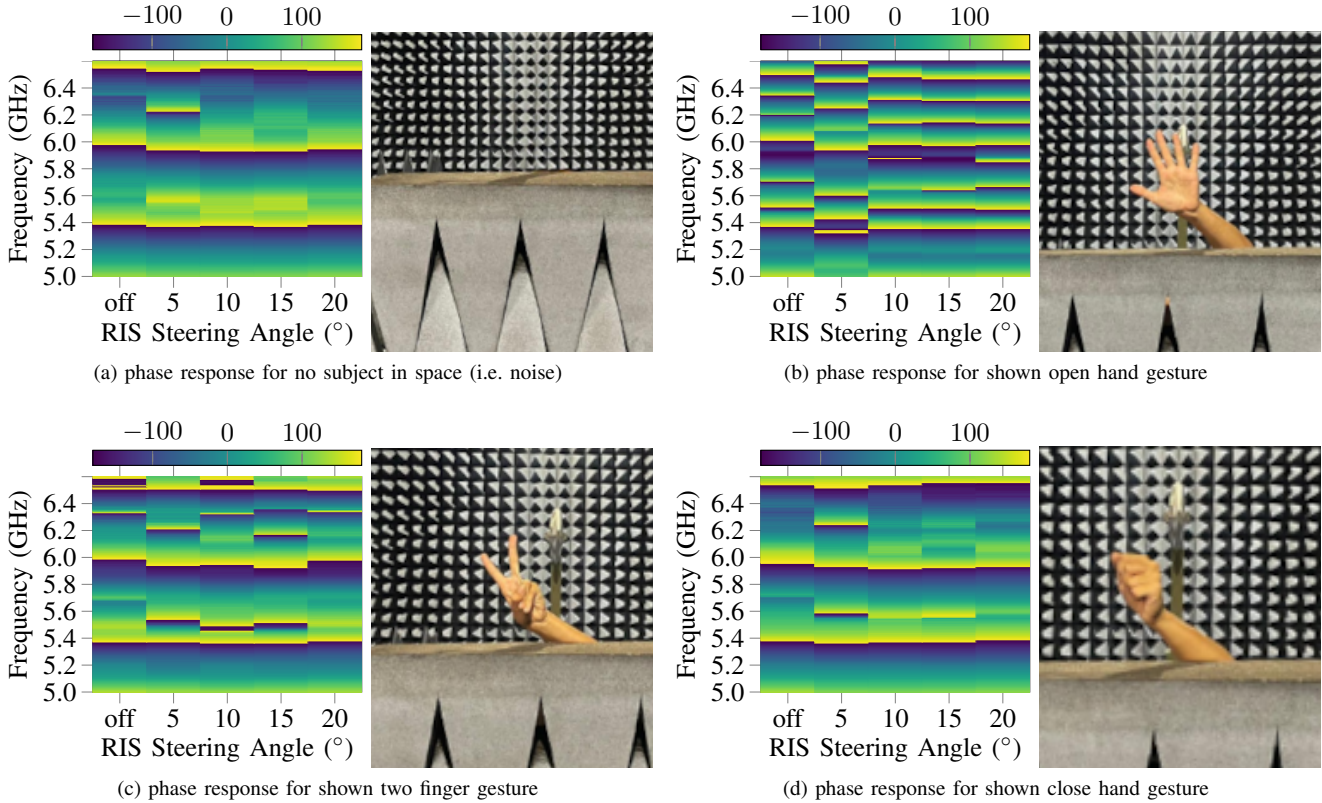


Fig. 16. Phase responses for a frequency range of 5.0 GHz to 6.5 GHz for several RIS configurations and hand gestures.

configurations per frame, the VNA was triggered to record the  $S_{21}$  parameters, resulting in a consistent data capture of 390 total measurements per run.

**Data Gathering** In this second experiment, a wooden hand was used for data collection. The collection process now applies the FCAO algorithm, and therefore, the concept of frames will be discussed again. The copper tape was applied to the fingers to enhance reflection, as wood is a material with low reflectivity [54]. The same three hand gestures from the first experiment – Open Hand, Two Fingers, and Close Hand – were studied. For each gesture, 115 runs were conducted using random RIS configurations and 115 runs with optimized configurations, both across 10 frames. Each sample contained transmission coefficients  $S_{21}$  across the frequency range of 5 GHz to 6.5 GHz, with 201 points per sample, as in the first experiment. The time required to complete one run, which involved capturing data for all 390 configurations (39 configurations per frame  $\times$  10 frames), was about 3.5 minutes, for both the random and optimized configurations. The data acquisition time for the random configurations was 3.5 minutes  $\times$  115 runs = 6.7 hours per gesture, with the same duration for the optimized configurations. Considering the three different gestures, this resulted in a total acquisition time of around 39 hours. This extended acquisition time demonstrated the impracticality of using a real hand for large datasets, as it demands prolonged data collection, which is physically exhausting for participants. The dataset created and used in

this experimental campaign has been made publicly available<sup>2</sup>.

**Random vs. FCAO Optimized RIS Configurations** We again plotted the collected  $S_{21}$  data as images, mapping the responses in magnitude and phase for our frequency range, as a function of the sequence of RIS configurations. Figure 17 illustrates this for each gesture, for the random (left-hand side) and optimized RIS (right-hand side) configurations. Axes ranges are identical for each subplot, and the horizontal axis represents the entire sequence of 390 configurations. Value ranges are shown by the color bars on top, with magnitude ranging between  $-35$  dB to  $-15$  dB, and phase between  $\pm 180^\circ$ . The respective hand gestures for each row of plots are shown on the right.

For the random configuration patterns (Figures 17c, 17f and 17i), we observe little signal variability in magnitude across configurations, except for the last *frame* of the configuration sequence, i.e., the very end of the sequence (configurations 351 to 390). The phase response also varies slightly throughout the configuration sequence. This suggests that having multiple frames with random configurations may not add much value, as the information across them is redundant. However, there is variability between gestures, particularly in the 5.2 GHz to 5.6 GHz range, where *closed hand* results shows less variability relative to the *open hand* gesture.

<sup>2</sup>M. S. F. de B. Oliveira, F. M. Ribeiro, N. Paulino and L. M. Pessoa, “RIS Based Hand Gesture Recognition Dataset”. Zenodo, Sep. 12, 2024. doi: 10.5281/zenodo.13754235.

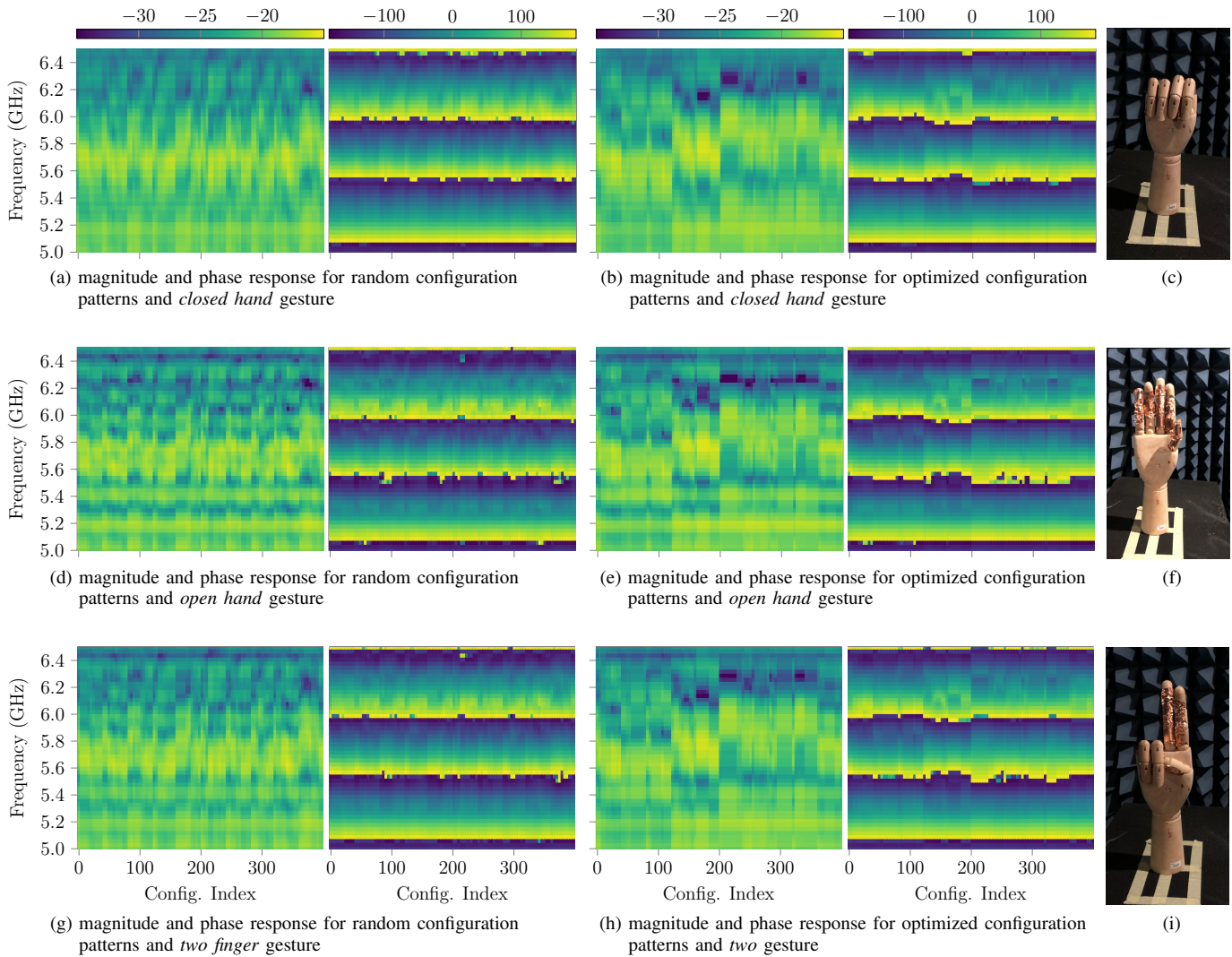


Fig. 17. Phase responses for a frequency range of 5.0 GHz to 6.5 GHz for random RIS configuration sequences (Figures 17a, 17d and 17g), and for the FCAO optimized sequences (Figures 17b, 17e and 17h), for each of the respective hand gestures (Figures 17c, 17f and 17i)

In contrast, the optimized patterns (Figures 17b, 17e and 17h) reveal more variability in both magnitude and phase across frames, indicating that more information is captured from the space of interest when using the configuration pattern computed by our FCAO implementation. Additionally, differences between gestures are more pronounced. For example, between 6.0 GHz and 6.5 GHz, the magnitude response is more attenuated relative to the random configuration sequence, and between the optimized cases, the attenuation is greater for *closed hand* relative to *open hand* in the same range.

For clarity, note that these plots show the average responses for each gesture, considering all 115 measurement runs per gesture. We present the averages directly, as opposed to analyzing specific chosen measurements runs per gesture since we observed very little variability for different measurements of the same gesture. For instance, considering the magnitude, and each set of 115 responses of the same set of measurements (i.e., same gesture and RIS configuration sequence), we verify that the mean of the absolute differences of a gesture relative to the average response of that gesture ranges between 0.2 dB

and 1.2 dB. Given the absolute range of  $-35$  dB to  $-15$  dB.

In other words, there is little variability in data between different runs for the same gesture, likely due to the anechoic environment and use of an immobile hand. Therefore, data augmentation techniques were employed to ensure a more robust dataset for training a neural network for hand gesture recognition.

**Data Augmentation for Machine Learning** We augmented our data in two ways. Firstly, we had observed during the experiments on the human hand subject that slight adjustments or minor motion of the hand position was enough to cause response changes. Therefore, additional measurements were taken with the wooden hand placed in 8 different minor variations for each gesture. Given the conclusion regarding variability between measurement of the same gesture we have just presented above, we determined that repeating measurements per gesture did not significantly contribute to new information. Therefore, only a single measurement was taken for each new orientation.

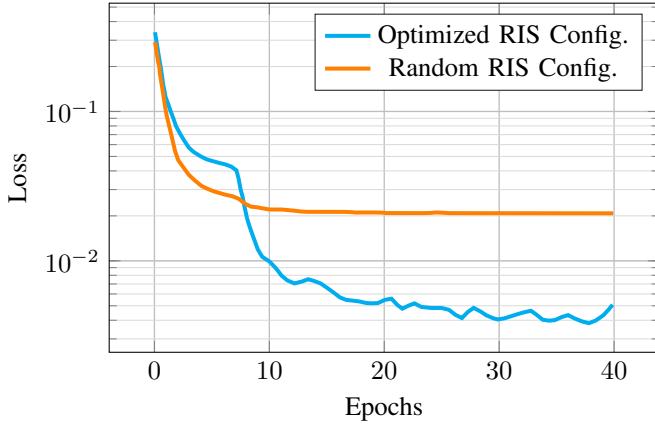


Fig. 18. Training epochs for CNN model #1, for both types of collected data

Instead, as our second technique for data augmentation, for each new variant position, we replicated the respective measurement 115 times, adding adding random white noise to each replica, with a mean of 0 and a standard deviation randomly chosen from [0.08, 0.1, 0.15, 0.2] as described in [55]. This approach ensured a more robust dataset, allowing the neural network to learn patterns beyond the limited original data and avoid overfitting. This process resulted in a total of 1035 samples per gesture for both random and optimized configurations (115 runs  $\times$  9 orientations). With three gestures, the final dataset comprised 3105 samples for random configurations and 3105 samples for optimized configurations, where each sample corresponds to an image used for CNN learning and classification.

**CNN Classification with Model #1** The first CNN model was adapted from the one presented by Hu *et al.* [26], from where we also based our FCAO implementation. We employ the same model here to arrive at a comparison of performance given the different input data collected. Specifically, in this reference paper, only the information at the operating frequency of the RIS was used for classification. In our case, due to the characterization results of the unit cell and RIS, this frequency  $f_c = 5.91$  GHz as we explained above (Section IV-A. Thus this model relies only on a single frequency and furthermore takes only the average of the response per each configuration frame (i.e., it considers the average of the responses of all configurations within a frame). In our case, and considering the inputs are the only real components of the complex values of the  $S_{21}$  parameters, resulting in a feature vector with dimension (1, 10).

Regarding architecture, the neural network consists of three fully connected layers: an initial layer with 64 neurons, a hidden layer with 32 neurons, both using the ReLU activation function and an output layer with *softmax* activation. The dataset was split into 80% for training, 10% for validation, and 10% for testing. We implemented the same model and trained it with both of our data subsets (i.e.,  $S_{21}$  data for random and optimized configuration sequences).

During training, shown in Figure 18, it is evident that in the initial epochs, the model converges faster with the

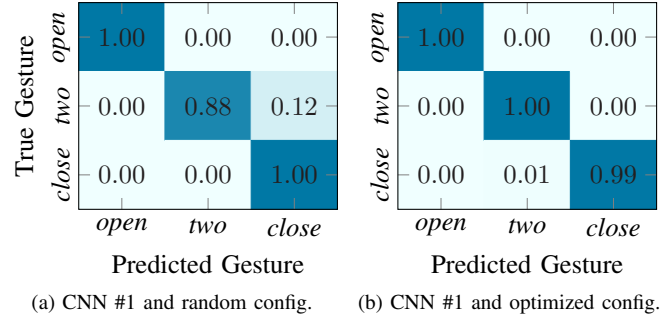


Fig. 19. Gesture recognition accuracy for the CNN architecture utilizing only  $S_{21}$  information for a frequency of 5.91 GHz using both random configurations or the FCAO optimized configurations

random configurations than with the optimized ones. This suggests that the optimized configuration contains more complex information, making it harder to learn at first. However, as training progressed the optimized configuration showed improved accuracy, achieving around  $5\times$  smaller loss. This pattern was also observed in the study by Hu *et al.* [26].

Figure 19 presents the confusion matrices for both configurations, demonstrating an advantage for the optimized setup. This confirms that the optimized configuration matrix, designed with low average mutual coherence, can achieve superior accuracy in gesture recognition. Specifically, the optimized configuration resulted in a 99.67% recognition accuracy, providing a 3.53% improvement compared to the random configuration, which achieved 96.14%.

**CNN Classification with Model #2** The architecture of our second CNN is presented in Figure 20. We now consider two channels, the magnitude and phase responses, represented as the image plots we have previously shown. That is, this model considers broadband information from 5.0 GHz to 6.5 GHz, and also the response of each individual configuration of the sequence, rather than an average response per frame. Thus the resulting dimension of the input layer is two feature vectors with dimensions (201, 390), and the output layer consists of three nodes for the hand gestures to be classified. This architecture allows the network to learn features from both the magnitude and phase components of the  $S_{21}$  parameters, potentially capturing distinct and complementary information.

The training for this network is shown in Figure 21, once again for both types of configuration strategies. In this case, the optimized case produces a lower loss per epoch even from the start of the training process. However, both configuration methods converge to a similar loss. That is, even though the gains from FCAO are less evident for this model and respective input data, its use still permits a faster learning convergence.

The resulting classification accuracies for both of our data subsets (i.e., random and optimized sequences) are shown in Figure 22. Both the random and optimized configurations achieved high recognition accuracies. The average accuracy for the dataset retrieved using random configurations is 99.047%, which increases minimally to 99.057% for the optimized configuration dataset. This smaller increase relative



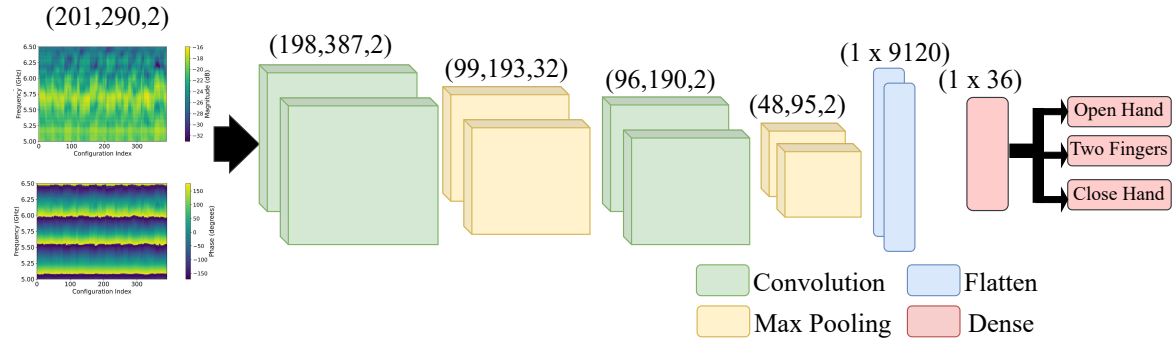


Fig. 20. Architecture of the CNN developed for gesture classification (Open Hand, Two Fingers, Close Hand). The input to the network consists of the magnitude and phase of the  $S_{21}$  parameters.

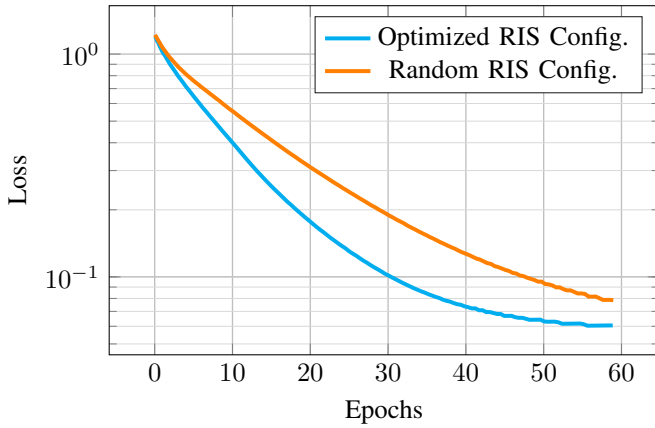


Fig. 21. Training epochs for CNN model #2, for both types of collected data

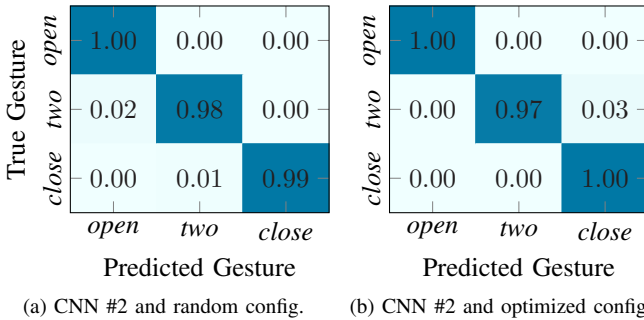


Fig. 22. Gesture recognition accuracy for the CNN architecture utilizing broadband information from 5.0 GHz to 6.5 GHz using both random configurations or the FCAO optimized configurations

to the previous CNN model again suggests that optimizing the RIS configuration yields little benefit, if also combined with data retrieved over a broader range of frequencies.

We can briefly justify this based on how the formulation of the channel and the FCAO algorithm operate. Underlying the approach are compressive sensing principles, where the FCAO algorithm can provide benefits if the number of measurements  $K$  is ideally much smaller than the number of distinct spatial regions of interest  $M$ . In other words, for the first CNN model with an input size of  $(1, K = 10)$  for our case where  $M = 32$ ,

these conditions are minimally satisfied.

However, in the second approach, by not averaging the signals across frames, we leverage all 390 distinct signals, which results in  $K = 390$ , greatly exceeding  $M$ . This substantial increase in measurements makes the application of the FCAO algorithm less impactful regarding the increase of information extracted from the channel. Both configurations achieve similarly high performance, demonstrating that with such a large number of measurements, the choice of configuration has minimal impact on the classification results.

In other words, while the use of an optimized configuration using FCAO increases the information captured from a narrowband signal, at the cost of the computational effort of determining this configuration, we achieve a similar level of information by capturing magnitude and phase data for a broader frequency band.

**Conclusion** We have successfully implemented an RF-based HGR approach, relying on our own RIS design. We have shown that  $S_{21}$  data for a bandwidth of 1.5 GHz within the WiFi-6E range provides sufficient information for gesture classification, even for a RIS design with only two control states per unit cell.

Regarding maximizing the information that can be extracted from the channel, Hu *et al.* [26] demonstrated this through the FCAO algorithm for a 2-bit RIS, achieving a 14.6% accuracy improvement in their Human Posture Recognition (HPR) setup, relative to random configurations sequences. Although not directly comparable, we achieve an analogous 3.53% increase in classification accuracy for our 3 gestures in our HGR setup using the same approach. The difference in accuracy increase can likely be attributed not just to different setups, but especially to the increased number of possible configuration states in a 2-bit RIS, and respective control of the RF medium. However, we have demonstrated that optimized configurations can still improve performance even for a 1-bit RIS, a smaller SoI, and different classification task. Moreover, we have shown that equivalent classification performances can be achieved even with a random sequence of RIS configuration states if a larger amount of RF data is directly collected from the channel.



## V. CONCLUSION

In this work, we have demonstrated a Hand Gesture Recognition (HGR) approach that is capable of classifying three distinct human hand gestures, relying only on RF data retrieved from a channel including the SoI within which the human hand target is located. The channel is a so called *smart radio* environment, as it includes a Reconfigurable Intelligent Surface (RIS) with which the response of the channel can be controlled. In this way, the information extracted from the SoI enriched, since different RIS configurations result in distinct  $S_{21}$  parameters.

We collected and published a large data set for HGR in an anechoic chamber setup, for three possible human hand gestures, and for a frequency range of 5.0 GHz to 6.5 GHz. For two Convolutional Neural Network (CNN) models, and two strategies for RIS reconfiguration, we have demonstrated a classification accuracy above 90%. Furthermore, we have demonstrated this using our own hardware designs for a RIS tile, and supporting control electronics, targeting the WiFi-6E range. Thus, we have advanced the state of the art on RF based solutions for Human Activity Recognition (HAR), specifically within the context of emergent sixth generation communications.

Our RIS design is a 8x8 tile using a PIN diode based unit cell was validated in anechoic chamber environment, where we analyzed the performance for several frequencies and steering angles. Despite a deviation from the intended central frequency, we demonstrated successful steering capabilities at 6.16 GHz, for a signal emitted by a feed horn located 40 cm away from the RIS. Our design will be available as open-source, accelerating research on RIS implementations and applications for WiFi-6E networks.

Regarding future work from a sensing perspective, we are assembling a setup with four tiles in order to achieve a total of 256 units cells. With this larger RIS, we expect that the HGR approach can be improved by potentially quadrupling the information from the channel. Due to a larger RIS, we can also consider a larger Space-of-Interest (SoI), and thus perform full body HPR. From a hardware design and communication perspective, We will also evaluate the improved steering capabilities with this larger RIS, and will also explore the multi-beam capabilities afforded by a tile level control. Finally, we will iterate on the unit cell design to mitigate deviations from the intended central frequency, which may include focusing on a 2-bit control, and also consider RIS designs for FR2 frequencies.

## REFERENCES

- [1] M. H. Arshad, M. Bilal, and A. Gani, "Human activity recognition: Review, taxonomy and open challenges," *Sensors*, vol. 22, no. 17, 2022.
- [2] M. A. R. Ahad, A. D. Antar, and M. Ahmed, *Human Activity Recognition: Data Collection and Design Issues*. Cham: Springer International Publishing, 2021, pp. 63–75.
- [3] P. Patel, B. Bhatt, and B. Patel, "Human body posture recognition — a survey," in *2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, 2017, pp. 473–477.
- [4] N. Dua, S. N. Singh, S. K. Challa, V. B. Semwal, and M. L. S. Sai Kumar, "A survey on human activity recognition using deep learning techniques and wearable sensor data," in *Machine Learning, Image Processing, Network Security and Data Sciences*, N. Khare, D. S. Tomar, M. K. Ahirwal, V. B. Semwal, and V. Soni, Eds. Springer Nature Switzerland, 2022, pp. 52–71.
- [5] D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, "A survey of human activity recognition in smart homes based on iot sensors algorithms: Taxonomies, challenges, and opportunities with deep learning," *MDPI Sensors*, vol. 21, no. 18, 2021.
- [6] Z. Zhuang and Y. Xue, "Sport-related human activity detection and recognition using a smartwatch," *MDPI Sensors*, vol. 19, no. 22, 2019.
- [7] A. Sunil, M. H. Sheth, S. E, and Mohana, "Usual and unusual human activity recognition in video using deep learning and artificial intelligence for security applications," in *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2021, pp. 1–6.
- [8] P. Xu, "A real-time hand gesture recognition and human-computer interaction system," 2017. [Online]. Available: <https://arxiv.org/abs/1704.07296>
- [9] Z. Hussain, Q. Z. Sheng, and W. E. Zhang, "A review and categorization of techniques on device-free human activity recognition," *Journal of Network and Computer Applications*, vol. 167, p. 102738, 2020.
- [10] A. Jalal, Y.-H. Kim, Y.-J. Kim, S. Kamal, and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused features," *Pattern Recognition*, vol. 61, pp. 295–308, 2017.
- [11] S. K. Bashar, A. Al Fahim, and K. H. Chon, "Smartphone based human activity recognition with feature selection and dense neural network," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020, pp. 5888–5891.
- [12] N. Roy, A. Misra, and D. Cook, "Ambient and smartphone sensor assisted adl recognition in multi-inhabitant smart environments," *Journal of ambient intelligence and humanized computing*, vol. 7, pp. 1–19, 2016.
- [13] Y. Yang, P. Hu, J. Shen, H. Cheng, Z. An, and X. Liu, "Privacy-preserving human activity sensing: A survey," *High-Confidence Computing*, vol. 4, no. 1, p. 100204, 2024.
- [14] S. Khunteta, P. Saikrishna, A. Agrawal, A. Kumar, and A. K. R. Chavva, "Rf-sensing: A new way to observe surroundings," *IEEE Access*, vol. 10, pp. 129 653–129 665, 2022.
- [15] M. Zhao, T. Li, M. A. Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human pose estimation using radio signals," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7356–7365.
- [16] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Trans. Graph.*, vol. 34, no. 6, Nov. 2015.
- [17] C. Xie, D. Zhang, Z. Wu, C. Yu, Y. Hu, and Y. Chen, "Rpm 2.0: Rf-based pose machines for multi-person 3d pose estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 1, pp. 490–503, 2024.
- [18] L. Wang, C. Zhang, Q. Zhao, H. Zou, S. Lasaulce, G. Valenzise, Z. He, and M. Debbah, "Generative ai for rf sensing in iot systems," 2024. [Online]. Available: <https://arxiv.org/abs/2407.07506>
- [19] M. Dajer, Z. Ma, L. Piazzini, N. Prasad, X.-F. Qi, B. Sheen, J. Yang, and G. Yue, "Reconfigurable intelligent surface: design the channel – a new opportunity for future wireless networks," *Digital Communications and Networks*, vol. 8, no. 2, pp. 87–104, 2022.
- [20] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Comm. Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.
- [21] H. Zhang, B. Di, L. Song, and Z. Han, *Introductions and Basics*. Cham: Springer International Publishing, 2021, pp. 1–17.
- [22] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using wifi," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*. New York, NY, USA: Association for Computing Machinery, 2015, p. 269–282. [Online]. Available: <https://doi.org/10.1145/2785956.2787487>
- [23] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5386–5402, 2021.
- [24] R. Chen, M. Liu, Y. Hui, N. Cheng, and J. Li, "Reconfigurable intelligent surfaces for 6g iot wireless positioning: A contemporary survey," *IEEE Internet of Things Journal*, vol. 9, no. 23, pp. 23 570–23 582, 2022.
- [25] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding, and M. Gidlund, "Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6g wireless networks," *Wireless Communications*, vol. 28, no. 6, p. 184–191, Dec. 2021.
- [26] J. Hu, H. Zhang, B. Di, L. Li, K. Bian, L. Song, Y. Li, Z. Han, and H. V. Poor, "Reconfigurable intelligent surface based rf sensing: Design, optimization, and implementation," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2700–2716, 2020.

- [27] G. C. Alexandropoulos, N. Shlezinger, I. Alamzadeh, M. F. Imani, H. Zhang, and Y. C. Eldar, "Hybrid reconfigurable intelligent metasurfaces: Enabling simultaneous tunable reflections and sensing for 6g wireless communications," *IEEE Vehicular Technology Magazine*, vol. 19, no. 1, pp. 75–84, 2024.
- [28] C. Li, Q. Huang, Y. Zhou, Y. Huang, Q. Hu, H. Chen, and Q. Zhang, "Riscan: Ris-aided multi-user indoor localization using cots wi-fi," in *Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems*, 2024, p. 445–458.
- [29] Y. He, D. Zhang, and Y. Chen, "High-resolution wifi imaging with reconfigurable intelligent surfaces," *IEEE Internet of Things Journal*, vol. 10, no. 2, pp. 1775–1786, 2023.
- [30] N. Mohamed, M. B. Mustafa, and N. Jomhari, "A review of the hand gesture recognition system: Current progress and future directions," *IEEE Access*, vol. 9, pp. 157422–157436, 2021.
- [31] T. Zheng, Z. Chen, S. Ding, and J. Luo, "Enhancing rf sensing with deep learning: A layered approach," *IEEE Communications Magazine*, vol. 59, no. 2, pp. 70–76, 2021.
- [32] Z. Li, Y. Xie, L. Shanguan, R. I. Zelaya, J. Gummeson, W. Hu, and K. Jamieson, "Programmable radio environments with large arrays of inexpensive antennas," *GetMobile: Mobile Computing and Communications*, vol. 23, no. 3, p. 23–27, 2020.
- [33] L. Li, Y. Shuang, Q. Ma, H. Li, H. Zhao, M. Wei, C. Liu, C. Hao, C. W. Qiu, and T. J. Cui, "Intelligent metasurface imager and recognizer," *Light: Science and Applications*, vol. 8, 12 2019.
- [34] H.-Y. Li, H.-T. Zhao, M.-L. Wei, H.-X. Ruan, Y. Shuang, T. J. Cui, P. del Hougne, and L. Li, "Intelligent electromagnetic sensing with learnable data acquisition and processing," *Patterns*, vol. 1, no. 1, p. 100006, 2020.
- [35] Z. Wang, H. Zhang, H. Zhao, T. J. Cui, and L. Li, "Intelligent electromagnetic metasurface camera: System design and experimental results," *Nanophotonics*, vol. 11, pp. 2011–2024, 4 2022.
- [36] L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C.-W. Qiu, and T. J. Cui, "Machine-learning reprogrammable metasurface imager," *Nature communications*, vol. 10, no. 1, p. 1082, 2019.
- [37] R. W. Gerchberg, "A practical algorithm for the determination of plane from image and diffraction pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972.
- [38] A. Araghi, M. Khalily, M. Safaei, A. Bagheri, V. Singh, F. Wang, and R. Tafazolli, "Reconfigurable intelligent surface (ris) in the sub-6 ghz band: Design, implementation, and real-world demonstration," *IEEE Access*, vol. 10, pp. 2646–2655, 2022.
- [39] B. G. Kashyap, P. C. Theofanopoulos, A. S. Shekhawat, A. Y. Modi, A. P. Sengar, S. K. Kumar, A. Chang, T. Osman, A. Alkhateeb, and G. C. Trichopoulos, "A reconfigurable intelligent surface for 5g wireless communication applications," in *2021 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/URSI)*, 2021, pp. 111–112.
- [40] M. Heinrichs, A. Sezgin, and R. Kronberger, "Open source reconfigurable intelligent surface for the frequency range of 5 ghz wifi," in *IEEE Intl. Symposium On Antennas And Propagation (ISAP)*, 2023, pp. 1–2.
- [41] M. Rossanese, P. Mursia, A. Garcia-Saavedra, V. Sciancalepore, A. Asadi, and X. Costa-Perez, "Open experimental measurements of sub-6ghz reconfigurable intelligent surfaces," *IEEE Internet Computing*, vol. 28, no. 2, pp. 19–28, 2024.
- [42] Y. Zhang, Y. Shao, L. Xiong, and J. Zhang, "Human activity recognition based on on-body antennas and semi-supervised learning," in *2023 17th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2023, pp. 1–5.
- [43] Y. Kim and Y. Li, "Human activity classification with transmission and reflection coefficients of on-body antennas through deep convolutional neural networks," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 5, pp. 2764–2768, 2017.
- [44] S. H. Gupta, A. Sharma, M. Mohta, and A. Rajawat, "Hand movement classification from measured scattering parameters using deep convolutional neural network," *Measurement*, vol. 151, p. 107258, 2020.
- [45] M. Oliveira, F. M. Ribeiro, N. Paulino, O. Yurduseven, and L. M. Pessoa, "Specrf-posture: Exploring specular reflections for human posture recognition," in *2024 IEEE International Mediterranean Conference on Communications and Networking (MediCom)*, 2024, pp. 227–232.
- [46] N. Rubab, S. Zeb, A. Mahmood, S. A. Hassan, M. I. Ashraf, and M. Gidlund, "Interference mitigation in ris-assisted 6g systems for indoor industrial iot networks," in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2022, pp. 211–215.
- [47] V. K. Chapala and S. M. Zafaruddin, "Reconfigurable intelligent surface empowered multi-hop transmission over generalized fading," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.
- [48] R. Zhou, X. Chen, W. Tang, X. Li, S. Jin, E. Basar, Q. Cheng, and T. J. Cui, "Modeling and measurements for multi-path mitigation with reconfigurable intelligent surfaces," in *2022 16th European Conference on Antennas and Propagation (EuCAP)*, 2022, pp. 1–5.
- [49] S. Ikram and G. Ahmad, "Design & implementation of a standard gain horn antenna for optimized gain and radiation pattern using mathcad & hfss," in *2008 Second International Conference on Electrical Engineering*, 2008, pp. 1–5.
- [50] R. Johnson and H. Jasik, *Antenna Engineering Handbook*, ser. Electronics Electrical Engineering. McGraw-Hill, 1993. [Online]. Available: <https://books.google.pt/books?id=xTSNJhVIHGgC>
- [51] P. Nayeri, F. Yang, and A. Z. Elsherbeni, "Reflectarray antennas: theory, designs, and applications," 2018.
- [52] A. Yu, F. Yang, A. Z. Elsherbeni, J. Huang, and Y. Rahmat-Samii, "Aperture efficiency analysis of reflectarray antennas," *Microwave and Optical Technology Letters*, vol. 52, no. 2, pp. 364–372, 2010.
- [53] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.
- [54] J. Gervacio and J. Cruz-Mandujano, "Measurement of light reflectance in different types of wood," vol. 6046, 02 2006, pp. 604617–604617.
- [55] M. Sheeny, A. Wallace, and S. Wang, "Radio: Parameterized generative radar data augmentation for small datasets," *Applied Sciences*, vol. 10, no. 11, 2020.