

# MEDUSA: Scalable Multi-View Biometric Sensing in the Wild with Distributed MIMO Radars

Yilong Li<sup>1</sup>  
University of Wisconsin-Madison<sup>1</sup>  
United States  
yilong@cs.wisc.edu

Ramanujan K Sheshadri<sup>2</sup>  
Nokia Bell Labs<sup>2</sup>  
United States  
ram.sheshadri@nokia-bell-labs.com

Karthik Sundaresan<sup>3</sup>  
Georgia Institute of Technology<sup>3</sup>  
United States  
karthik@ece.gatech.edu

Eugene Chai<sup>2</sup>  
Nokia Bell Labs<sup>2</sup>  
United States  
eugene.chai@nokia-bell-labs.com

Yijing Zeng<sup>1</sup>  
University of Wisconsin-Madison<sup>1</sup>  
United States  
yijingzeng@cs.wisc.edu

Jayaram Raghuram<sup>1</sup>  
University of Wisconsin-Madison<sup>1</sup>  
United States  
jraghuram@wisc.edu

Suman Banerjee<sup>1</sup>  
University of Wisconsin-Madison<sup>1</sup>  
United States  
suman@cs.wisc.edu

## ABSTRACT

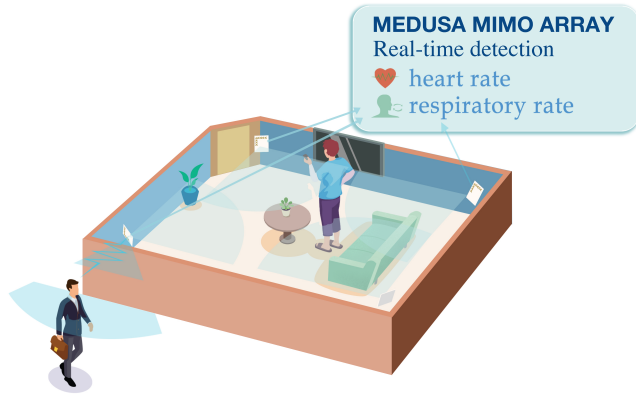
Radio frequency (RF) techniques have shown promise for continuous contactless healthcare applications. However, real-world indoor environments pose challenges for existing systems, which may struggle to detect subtle physiological signals. This paper proposes MEDUSA, a novel wireless vital-sign sensing system designed for multi-view setups. It enables users to deploy distributed Multiple Input Multiple Output (MIMO) arrays into their daily living environments, facilitating vital-sign sensing in real-world settings. Unlike most existing single Commercial Off-The-Shelf (COTS) radar-based systems that operate under controlled settings MEDUSA’s primary novelty lies in the design of a first-of-its-kind flexible *multi-view* vital sign sensing system that is view-agnostic, pose-agnostic, and contactless and can sense basic human vitals with good accuracy. Through our well-engineered hardware and software co-design, MEDUSA enables real-time processing of large distributed MIMO arrays, while balancing the tradeoff between Signal-to-Noise Ratio (SNR) and spatial diversity gain across each of its four distributed  $4 \times 4$  sub-arrays for increased robustness. This is achieved using our *novel unsupervised learning* model which effectively recovers vital sign waveforms by decomposing the received signals. Extensive evaluations with 21 participants demonstrate MEDUSA’s spatial diversity gain for real-world vital-sign monitoring, enabling free movement and orientation of subjects in both familiar and unfamiliar indoor environments.

## 1 INTRODUCTION

Over the last ten years, digital wellness has become increasingly popular, especially in the field of passive health monitoring (PHM) without the need for on-body devices. This trend supports a range of applications, such as remote physical rehabilitation, vital sign monitoring, and fall detection in indoor settings [6, 11, 13, 41, 44]. These systems offer numerous advantages to both users and physicians. Users are unburdened from the correct usage and maintenance of sensor devices, allowing them to participate freely in daily activities. On the other hand, physicians can implement continuous patient-monitoring protocols and facilitate proactive management of their health.

There have been a lot of recent advancements in wireless-driven passive health monitoring (PHM), ranging from simple respiration rate monitoring (e.g., [6]) to more advanced biometric sensing solutions using various technologies (e.g., UWB [11, 13, 41, 44], mmWave [13, 15], FMCW Radar [7], Wi-Fi [20, 24, 39]). While prevalent contact sensing technologies, including wristbands and chest vests, allow vital sign monitoring, such devices present limitations in prolonged use due to discomfort arising from contact with the skin and lack of robustness in detecting moving human subjects.

Despite their potential, existing radar-based vital-sign monitoring systems face challenges and decreased performance in real-world environments. (i) Lack of Robustness to Blockage and Motion: furniture and other environmental obstacles can block radar signals in real-world settings. Also, body parts may obstruct the radar path during subject movement. Commercial off-the-shelf (COTS) mmWave



**Figure 1: Vital-sign monitoring of human subjects in their daily habitats enables free movement and multiple subjects.**

radar devices face limitations in penetration, which hinders their ability to reliably distinguish the desired signals in the real world. **(ii) Limited Applicability:** Existing radar-based vital-sign sensing systems operate under controlled conditions due to limited antenna arrays, requiring subjects to face the radar directly. They allow only limited motion, either stationary or small movements, within a narrow field-of-view (FoV). **(iii) Limited Coverage from Single-View Sensing:** single-view vital-sign sensing systems, whether using a single Tx-Rx COTS radar or a single-perspective MIMO configuration, can only capture partial information from a human subject, especially in NLoS. Single-view systems are particularly vulnerable to occlusions caused by the subject’s body parts when the subject is not directly facing the radar. These occlusions can obstruct the radar signals, resulting in inaccurate detection.

The primary question posed in this research is: *Is it feasible to transition radar sensing of biometric signals from restrictive deployments to practical **in-the-wild** operation, allowing for the monitoring of subjects in their daily life environment without sacrificing the sensing performance?*

We propose a novel RF-based vital-sign monitoring system with distributed MIMO arrays, enabling the detection of human subjects in ambulatory settings with unconstrained movement and orientation. The core intuition behind MEDUSA is to leverage the diverse signal paths and perspectives enabled by spatially distributed MIMO radar arrays for “multi-view sensing”, as shown in Fig. 1. Unlike existing “single-view” systems that can only capture partial vital-sign information as people change orientations, our distributed MIMO approach allows the gathering of multi-aspect data.

While distributed MIMO arrays contribute to a throughput improvement in communication systems, it can potentially

deliver a more profound impact for wireless sensing – going from constrained operation to enabling practical “in-the-wild” sensing operation that is robust and immune to varying target locations, orientations, blockages, and mobility. Towards realizing this objective, we present MEDUSA – a large systems effort that addresses the following key challenges, leading to its contributions:

**1. Distributed Radar System:** How do we create a practical, distributed MIMO radar system that has sufficient gain to operate in real-time and reliably for practical room-scale deployments? While UWB and mmWave radar systems offer high resolution (1-4 GHz bandwidth), radar sensing offers a very limited operational range (less than 3m) with today’s systems that have limited arrays (e.g.,  $4 \times 3$  elements virtual mmWave array [1]). This is further compounded by the potential interference that arises when distributed radars transmit simultaneously without tight synchronization – a feature that is hard to accomplish yet important for a coherent MIMO system, the lack of which leads to confounding and corruption of the received target signals with other reflections.

**Contribution 1.** MEDUSA builds a first-of-its-kind 16 UWB radar elements MIMO system that enables flexible, coherent, distributed vital sign sensing in a multi-view way (shown in Fig. 2). The wireless synchronization facilitates easy deployment and ensures robustness across diverse target and environment configurations. mmWave radar suffers from signal blockage in NLoS scenarios. In contrast, UWB’s better penetration capability (3-10 GHz) compared to mmWave, coupled with the large antenna gain in MEDUSA, allows us to detect targets as far as 6m, even in NLoS reliably.

**2. Balancing MIMO SNR Gain vs. Spatial Diversity Gain:** The theory of distributed MIMO radar well understood [8]. However, dispersing the MIMO array into various sub-arrays can amplify spatial diversity gain, even if the SNR of individual sub-arrays is reduced. We revisit the tradeoff between antenna gain (SNR) and diversity gain (outage probability) in the context of sensing: A single 256 virtual MIMO ( $16Tx \times 16Rx$ ) array maximizes gain/coverage from one perspective but lacks robustness to diverse target orientations, locations, NLoS conditions, blockages, and motion. Conversely, distributing sub-arrays across multiple locations enhances robustness to “in-the-wild” configurations at the cost of reduced antenna gain and coverage per sub-array. The challenge is efficiently managing this tradeoff to ensure accurate and robust sensing across varied target and environmental conditions.

**Contribution 2.** Through extensive experimental and motivational characterization, MEDUSA identifies an ideal operating point, wherein one sub-array per quadrant around the target area strikes the most effective balance – any further distribution only reduces the gain per sub-array without

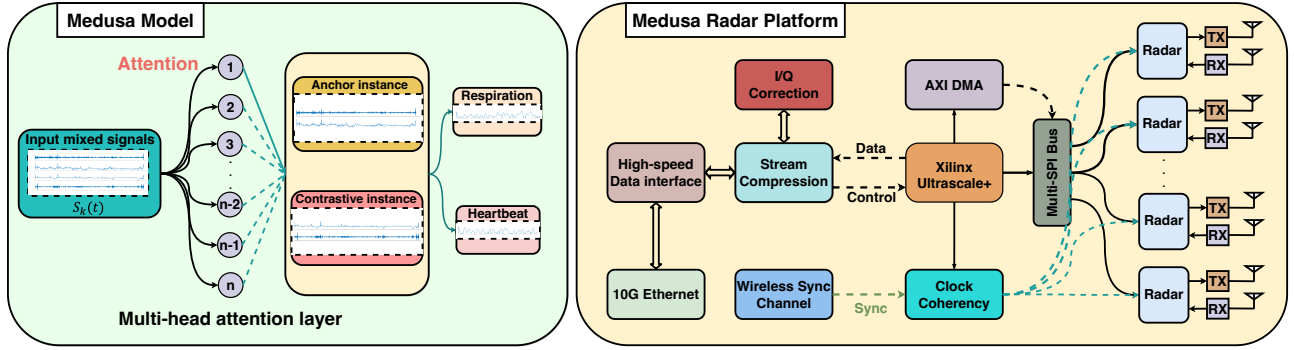


Figure 2: System architecture diagram of MEDUSA.

adding appreciably to the diversity gain. In contrast, less diversity diminishes the robustness in sensing performance. In particular, MEDUSA employs four  $4 \times 4$  sub-arrays, each with 16 virtual MIMO elements, to establish its distributed radar system for rooms. This configuration also offers flexibility to adapt to other room shapes as needed.

**3. Multi-view Information Fusion and Signal Extraction:** While MEDUSA enables optimal balanced SNR sensing signals and multi-view information in dynamic environments, it is still a hard challenge to isolate the vital sign signals mixed with numerous other reflections (distortions from human motion, multipath, etc.) in a non-linear manner. Additionally, extracting desired signals from different angles of the subject requires synchronizing and fusing the collected vital sign signals. Machine learning offers a scalable approach to extract our desired signal, given its recent success in several wireless sensing problems [11, 13, 15, 37, 44]. However, these learning models are limited to controlled settings and single-view COTS radar data, lacking the ability to process multi-view information. Respiration belts and other professional contact vital sign sensors [2, 4]) suffer from signal distortion caused by human movement, which leads to a lack of reliable ground truth data for supervised learning. Furthermore, every environment leaves its artifacts in a dataset, significantly affecting the model’s generality when tested on unseen environments with different configurations.

**Contribution 3.** MEDUSA leverages an unsupervised approach to extract the desired target breathing signals in various environment configurations while allowing for free motion and movement inherent to everyday tasks. To leverage the diversity offered by the distributed sub-arrays, MEDUSA employs a multi-head attention mechanism that allows the model to appropriately and implicitly attend to the different sub-arrays based on the estimated location and orientation of the target as well as the configuration of the environment (blockages).

Our comprehensive evaluation with 21 participant subjects reveals that MEDUSA improves median respiration measurement by over 20% accuracy compared to prior art and baselines in practical indoor environments, characterized by varying subject locations and orientations, and sustains errors under 5% even in NLoS (obstacles) conditions, where other approaches falter. Further, leveraging the diversity benefits of its distributed platform allows its model to (generalize) sustain performance accuracy even in unseen environments, subject mobility, and multiple subjects.

To the best of our knowledge, MEDUSA is the first-its-kind real-time vital sign monitoring system designed to function effectively in dynamic real-world environments, which is pioneering in incorporating multi-view information fusion into wireless vital sign sensing.

## 2 MOTIVATION

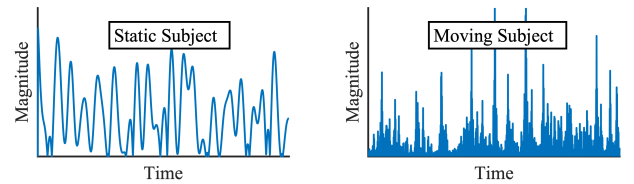


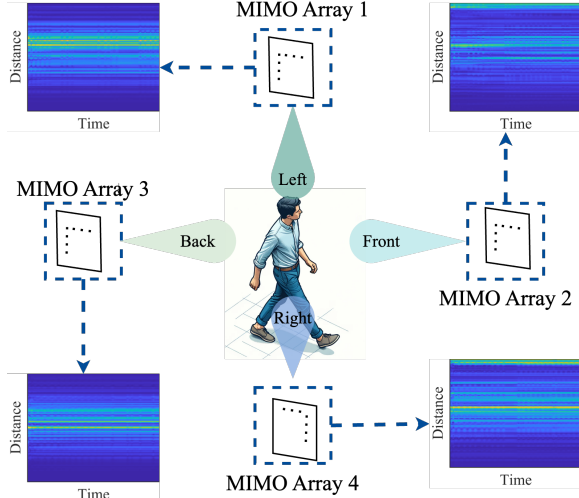
Figure 3: Respiration signals by vibration of COTS radar: Static subject vs. Moving subject.

In this section, we empirically study the impact of human movement and orientation on wireless vital sign monitoring and motivate the advantages of *distributed* MIMO array-based sensing (multi-view) over traditional co-located MIMO antenna systems (single-view).

### 2.1 Case for Distributed MIMO Sensing

The high-level idea of MEDUSA is to strike a tradeoff between MIMO antenna gain for SNR vs. the spatial diversity gain (multi-view) for a dynamic environment. As depicted in Fig. 4, the radar data from multiple angles provides diverse “viewpoints” of the target. In real environments, furniture and obstacles may obstruct some signal paths. However, as

subjects move, it's improbable that all directions are simultaneously blocked. Multi-view information from distributed MIMO radar arrays can prevent performance drops in traditional single-view sensing systems. Sensing information from angles beyond the front view of the human subject provides complementary data, thereby achieving diversity gain through these different "viewpoints." Our motivational experiments illustrate how spatial diversity solves the challenges discussed before.



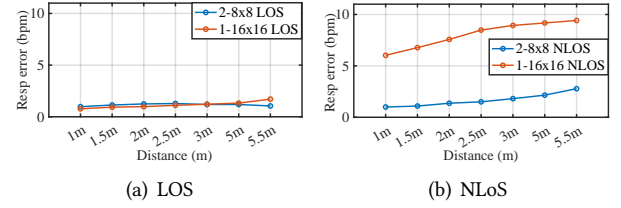
**Figure 4: Spatial diversity from distributed MIMO arrays (1, 2, 3, 4) in multi-view.**

**Lack of Robustness to Blockage and Movement:** Environmental structures can block radar signals in real-world settings, and as people move, their own body parts may also obstruct the signal paths. The respiration signals for static and moving subjects, shown in Fig. 3, demonstrate the distortion caused by movement, making it difficult to extract accurate breathing patterns.

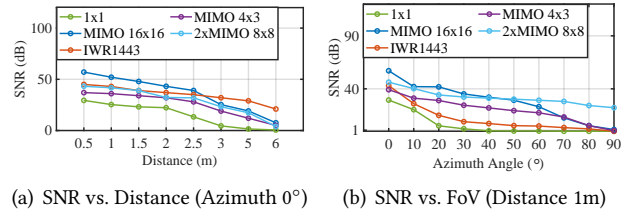
The result in Fig 5 highlights that distributing our  $16 \times 16$  modular array into two  $8 \times 8$  arrays (placed at 90 degrees to each other) into multi-view setup and illustrates a real-world deployment scenario for a static subject with furniture and obstructions (e.g., wooden or metal barrier) in the room space. For a fair comparison, we employ one  $16 \times 16$  MIMO array and two  $8 \times 8$  MIMO arrays, ensuring an equal number of antennas. The  $16 \times 16$  MIMO array is subjected to blockage by an obstacle barrier placed directly in front of it. It highlights the substantial performance degradation of co-located radar systems (Single-View) in such situations, with a notable error percentage of over 25%.

**Limited Applicability:** COTS UWB radar sensors used in previous studies [11, 41, 44] have a SISO configuration, limiting the operating range. MIMO radars offer superior spatial resolution and higher SNR gain compared to SISO

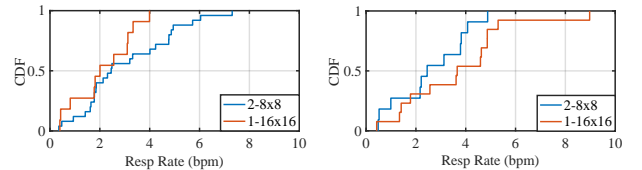
systems, offering increased operating range across both distance and angular dimensions. This is evident in the  $16 \times 16$  (256-element virtual) radar, as shown in Fig. 6. It achieves a substantial 10-30 dB SNR gain, particularly when the target is directly facing the antennas (Azimuth  $0^\circ$ ), enabling practical operational ranges exceeding 6m. However, the performance of co-located MIMO systems deteriorates when the target is outside the field-of-view (FoV) in Fig 6(b), resulting in a degraded SNR. While the mmWave MIMO radar delivers a good SNR when the target directly faces it (Azimuth  $0^\circ$ ), its performance degrades significantly out of FoV.



**Figure 5: Single-view  $16 \times 16$  MIMO vs. two  $8 \times 8$  MIMO radars in LOS and NLOS scenarios with subjects facing the radar.**



**Figure 6: Distance and FoV measurement ( $16 \times 16$  MIMO,  $4 \times 3$  MIMO,  $1 \times 1$  COTS UWB radar, TI IWR1443 mmWave radar and two  $8 \times 8$  MIMO).**



**Figure 7: Single-view (one  $16 \times 16$  radar) vs. multiple-view radar (two  $8 \times 8$ ).**

**Limited Coverage of Single-View Radars:** While conventional single-view (co-located) MIMO antennas improve SNR and coverage of a radar's aperture, their ability to illuminate the target and obtain sufficient reflections diminishes significantly when the target is not oriented front-facing the radar. On the other hand, a spatially distributed MIMO system will



provide the necessary spatial diversity for vital monitoring irrespective of the target's distance and orientation. The result in Fig 7 illustrates that two  $8 \times 8$  arrays (placed at 90 degrees to each other) of multi-view setup sacrifice some accuracy (median error still under 5%) compared to a single  $16 \times 16$  array when the static target is facing front. While it is impossible to eliminate the blockage of a co-located MIMO array (for instance, when an object in the living room, such as a refrigerator, obstructs the signal path, causing 6-7 dB attenuation), distributed MIMO arrays with multiple viewpoints can still operate effectively, by increasing the probability of an unobstructed path.

## 2.2 MEDUSA Design Challenges

While our experimental study motivates the need for a distributed MIMO radar for robust HVM across target and environment dynamics, realizing such a system is challenging.

**Synchronization and phase coherency:** In developing an active MIMO radar system, it is imperative to guarantee synchronization of each radar element that operates with its own transmit/receive (TX/RX) chain. This poses a challenge in hardware design, where the system must provide adequate gain across all the distributed subarrays, while concurrently synchronizing them, lest leading to interference and degradation of target SNR and consequently obscured by additional reflections. Although challenging, high-precision, over-the-air clock synchronization is crucial for scalable deployment and operation of distributed MIMO arrays.

**Balancing SNR vs. Diversity:** As seen in the study, while large MIMO arrays (e.g.,  $16 \times 16$ ) increase the SNR and hence the coverage of the radar, they fail to bring robustness to sensing in practical scenarios. In contrast, while spatially distributing a large MIMO array into several smaller subarrays brings much-needed diversity, over-distribution can significantly reduce the gain and coverage per aperture without adding diversity appreciably. A careful balance of the tradeoff between coverage and sensing robustness need to be addressed in determining an efficient distribution of MIMO sub-arrays for sensing.

**Vital Sign Signal Extraction:** To realize the true potential of distributed MIMO sensing for robustness towards target and environment dynamics, we need a scalable approach that can leverage and effectively fuse signals from the distributed sub-arrays to accurately extract the human vital signal from a complex non-linear mixture with other undesired signals (e.g., human motion, multipath reflections) in real-time.

MEDUSA addresses these challenges by building a MIMO UWB radar sensing with accurate time synchronization for distributed radar elements in real-world situations and an unsupervised learning model.

## 3 SYSTEM DESIGN

MEDUSA employs a modular design that enables flexible deployment for optimal performance in a wide range of environments. Multiple radar sensors, each of up to  $16 \times 16$  array size, can be distributed to cover an area in its entirety while guaranteeing fully coherent distributed MIMO sensing across all radar sensors. Motivated by the observations made in Sec. 2, we now describe MEDUSA's original hardware design, its efficient deployment and operation, and tightly integrated software models that leverage its distributed deployment.

### 3.1 MEDUSA Hardware Design

MEDUSA consists of a *baseboard* onto which multiple RF *daughterboards* (up to 16). Each daughterboard is built around a Novelda X4 UWB radar chip [3] which drives one Tx and one Rx antenna. These antennas are mounted on a separate *antenna board* that is connected to the daughterboard via RF cables. These different components of MEDUSA are shown in Fig. 8.

**3.1.1 MEDUSA Radar Baseboard.** The radar baseboard comprises a Xilinx Zynq UltraScale+ MPSoC FPGA [5] and socket interfaces for 16 radar daughterboards. Each socket is connected to the Zynq FPGA via 16 SPI buses and a pair of differential RF clock lines. This fan design enables significant operating flexibility across the daughterboards: all 16 daughterboards to be operated as a single coherent MIMO radar; alternatively, the daughterboards can be divided into subgroups of smaller MIMO radars, with each subgroup operating independently of the others. Timing skews within the clock signals and SPI to the daughterboards must be eliminated to ensure coherent MIMO radar operation. MEDUSA distributes a phase-coherent 243MHz Low-Voltage Differential Signal (LVDS) clock signal from the Zynq FPGA to all attached daughterboards over the differential RF lines to eliminate clock noise and interference. Impedance matching is also carefully calibrated to ensure equal clock and SPI line lengths, further reducing any timing skews in the hardware. For coherent MIMO operation, MEDUSA must also ensure coherence across the internal state of all X4 UWB radar chips. Finally, the baseboard streams I/Q data from all daughterboards, in real-time, to a host PC via a 10GbE Ethernet connection.

**3.1.2 MEDUSA Daughterboard.** Each daughterboard is built around a Novelda X4 chip [3]. The daughterboard routes differential clock signals and SPI commands from the baseboard to the X4 chip and forwards I/Q data from the X4 radar back to the baseboard, all with minimal time delay. Each daughterboard is physically pluggable into the baseboard via an 18-socket interface.

**3.1.3 MEDUSA Antenna Design.** Each daughterboard drives one Tx and one Rx antenna. To minimize errors in vital

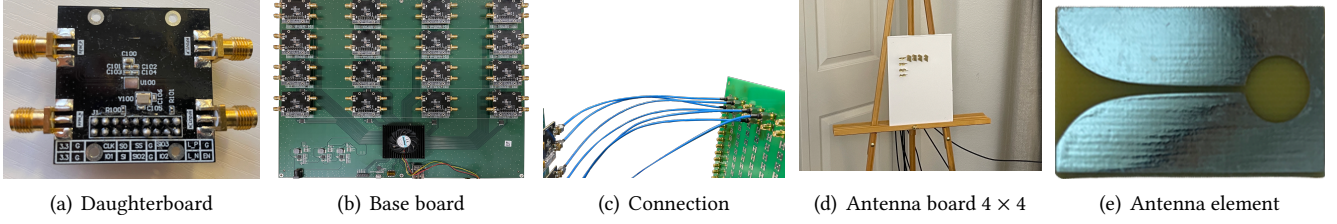


Figure 8: MEDUSA hardware platform.

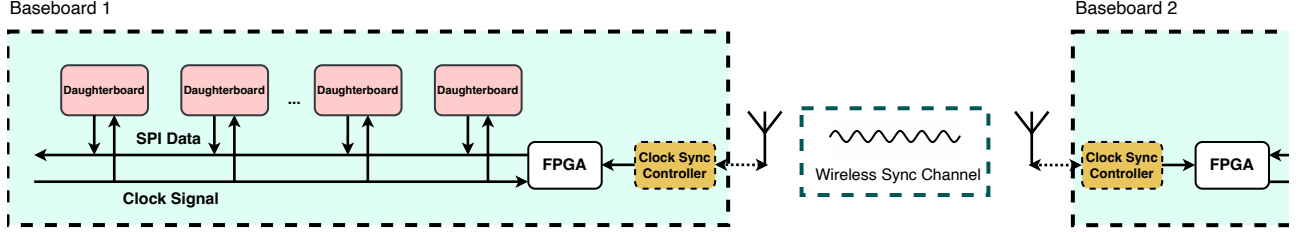


Figure 9: MEDUSA's scalable design with all elements clock synchronization.

sign monitoring, MEDUSA employs custom high-gain directional Vivaldi antenna elements that provide optimal SNR for radar returns, as shown in Fig. 8(e). The Novelda X4 chip employs differential RF lines for TX and RX, requiring 100Ω differential antennas. These antennas are connected to the daughterboard using SMA connectors.

### 3.2 Wireless Clock Synchronization

Efforts have been made to synchronize multiple radar systems over the years [9, 27]. However, these approaches are unsuited for wireless synchronization in indoor distributed MIMO array systems. The higher frequency of clock signals required for phase coherency makes conventional multiple radar synchronization systems ineffective in UWB radar systems. MEDUSA employs multiple baseboards with daughterboards and antennas for distributed MIMO sensing. To maintain coherence between physically separated daughterboards, MEDUSA uses wireless clock synchronization via COTS Software-Defined Radio (SDR) on each baseboard. One of the baseboards is designated as the *clock server*, while all others are the *clock clients*. Clock signals are transmitted from the server baseboard to all the client baseboards. Fig. 9 shows the clock generation design for each client. The wireless clock signals received by the client are sent into a phase-locked loop (PLL) which “cleans-up” the noisy wireless clock signal. Coherent MIMO requires that the phase of this signal be aligned at all the client baseboards. Nonetheless, the RF propagation distance between the server and the client influences the clock signal phase. A phase-offset correction is applied to obtain a clean reference clock signal, which is subsequently used to clock the X4 chip during standard operation.

**3.2.1 Clock Distribution.** The wireless clocks signals suffer from carrier frequency offset (CFO) due to the minute clock differences between the server and client baseboards. To correct this, the clock signals are transmitted as *differential* two tones of frequencies signals. Instead of transmitting a single clock signal, two different clock signals at frequencies  $f_1$  and  $f_2$  are sent. These are received at the client as  $\hat{f}_1 = f_1 + \delta f$  and  $\hat{f}_2 = f_2 + \delta f$  where  $\delta f$  is the CFO between the server and client transceivers. The client baseboard then derives the final reference clock from the difference of these two received signals,  $\hat{f}_1 - \hat{f}_2 = f_1 - f_2$ , that are no longer affected by the CFO.

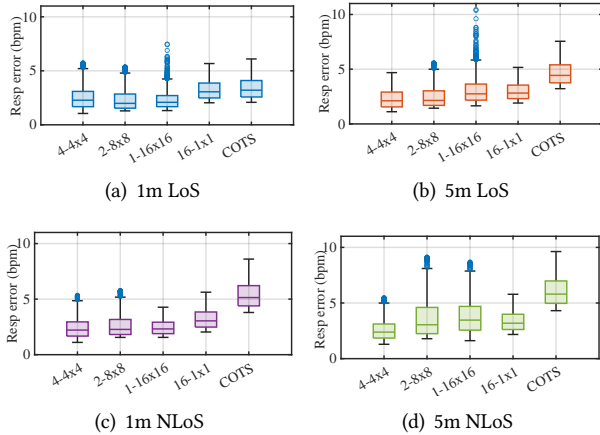
### 3.3 MEDUSA Deployment - Diversity vs. SNR

To achieve optimal performance, it is crucial to find a balance of two key factors: diversity gains from spatially distributed MIMO arrays and SNR gains from each co-located MIMO antenna array. We provide analytical discussion for deployment, supported by empirical experiments.

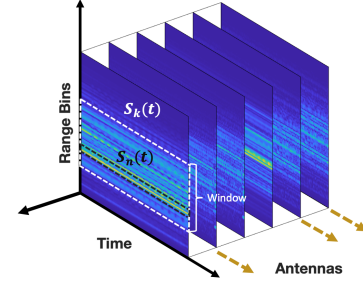
**Discussion:** MEDUSA's insight is that each single-view radar captures only partial vital sign information, affected by signal blockages and subject movement, especially in NLoS. The captured vital signs varies significantly with the subject's orientation. As Fig. 4 illustrates, radar data from four orthogonal directions (front, back, left, and right) demonstrates substantial diversity. Each MIMO radar potentially captures information from one or two orthogonal directions. SNR, proportional to antenna count, determines signal quality. While a  $16 \times 16$  array offers higher SNR than a  $4 \times 4$  array, it provides limited spatial diversity. Conversely, sixteen  $1 \times 1$  arrays maximize spatial diversity but sacrifice SNR. To balance deployment convenience, complete spatial diversity, and adequate SNR, we propose an optimal configuration of

four  $4 \times 4$  MIMO arrays in orthogonal directions. This setup ensures comprehensive vital sign information from moving subjects in NLoS conditions while maintaining signal quality in a low deployment workload.

**Experiments:** Then we empirically conduct thorough experiments leveraging the MEDUSA's modular design to validate various configurations (e.g., one  $16 \times 16$ , two  $8 \times 8$ , four  $4 \times 4$  and sixteen  $1 \times 1$  radars) and evaluating the measured breaths per minute (bpm) accuracy with a static target positioned at different indoor locations (see Fig 18) for **both LoS and NLoS scenarios**. (Details on ground truth and data collection can be found in Sec 4). Fig 10 shows a subset of our experimental results. In LoS conditions, when the target is close to a radar (e.g., 1m), the  $16 \times 16$  configuration delivers the best performance due to its high SNR gains with an average respiratory error of 2.01 bpm. However, when the target moves away to 5m, the  $16 \times 16$  configuration experiences a decline in accuracy as the SNR decreases, resulting in mean respiratory error of 2.98 bpm. Meanwhile, the spatially distributed four  $4 \times 4$  and two  $8 \times 8$  configurations together help compensate for the SNR loss at individual radars. In NLoS scenarios, the  $16 \times 16$  distributed single antenna arrays, while benefiting from spatial distribution and achieving more stable performance than single-view  $16 \times 16$  MIMO arrays, do not gain any SNR advantages, resulting in a median BPM error of 3.41 bpm at 1m and 4.12 bpm at 5m. The four  $4 \times 4$  configuration, on the other hand, offers the best trade-off (median BPM error 2.11 bpm) between SNR and diversity gains and perform the best. Consequently, MEDUSA adopts the four  $4 \times 4$  configuration.



**Figure 10: Respiration rate errors for various MIMO array configurations at 1m and 5m distances in both LoS and NLoS scenarios.**



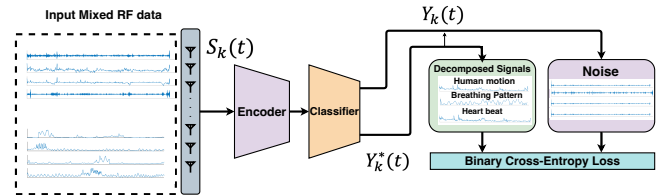
**Figure 11: Radar frame. The radar data structure of MEDUSA comprises CIR matrices associated with each radar element corresponding to each antenna.**

### 3.4 Human Vitals Sensing with MEDUSA

The X4 chip uses UWB time-of-flight pulses to measure ranges. With each UWB pulse, the X4 chip returns an 186-bin range measurement, with each bin containing the amplitude and phase of a reflected signal from the indicated distance. Fig 11 illustrates the returned range data from multiple daughterboard antennas over time. Changes in human activities and vitals (e.g. breathing, heartbeats, etc) will induce changes in the reflected signals captured in each radar frame. This is inter-mingled with signal reflections due to other movement in the environment, along with multipath distortions. These radar frames are encoded as I/Q data streams and transmitted over the 10 GbE link, where MEDUSA employs an unsupervised ML model to separate RF changes due to human vitals from those due to other unrelated activities.

**3.4.1 Extracting Human Vitals.** MEDUSA recovers the human vital signals mixed with other interfering signals using independent component analysis (ICA) [16]. Consider a *single* radar receiver,  $k \in \{1, \dots, M\}$ , where  $M$  is the total number of radar receivers. Let  $X_k(t) = [x_1(t), \dots, x_N(t)]^T$  be a vector of the  $N$  source signals induced by human respiration, heartbeats, and motion for one or more monitored individuals. These signals are combined in a non-linear fashion at the radar sensors as

$$S_k(t) = f([x_1(t), \dots, x_N(t)]^T) = [s_1(t), \dots, s_N(t)]^T \quad (1)$$



**Figure 12: Arhitecture of our unsupervised learning model to decompose and recover the waveforms.**

where  $S_k(t)$  is the vector of  $N$  received signals at the  $k^{th}$  radar receiver. In MEDUSA, these  $N$  signals  $s_n(t)$  arrive in

	Medusa Platform	IWR1443BOOST	AWR2243 Cascade
Frequency Band	6.5-9.5GHz	76-81 GHz	76-81 GHz
TX/RX	16TX/16RX	3TX/4RX	12TX/16RX
Azimuth Array	256 element virtual array	12 element virtual array	86 element virtual array
Max Angular Resolution	0.448 degree	9.53 degree	1.4 degree
Min Spacing Separation	0.039m at 5m	0.841m at 5m	0.122m at 5m
Frame Rate (FPS)	50 – 200	10	5

**Table 1: Comparison of MEDUSA platform, TI IWR1443BOOST, and TI AWR2243 Cascade.**

$N$  different range bins. The objective of non-linear ICA is to find the approximate inverse  $f^{-1}$  to recover the source signals  $X_k(t)$  from  $S_k(t)$ . To do this, we employ a contrastive learning model, as shown in Fig. 12.

In MEDUSA,  $S_k(t)$  represents a window of  $N$  range bins from receiver  $k$  at some time  $t$ , where  $1 \leq N \leq 186$ . We generate positive  $Y_k(t)$  and negative  $Y_k^*(t)$  augmented samples from  $S_k(t)$  for contrastive training, defined as

$$Y_k(t) = \begin{pmatrix} S_k(t) \\ S_k(t - T) \end{pmatrix}, Y_k^*(t) = \begin{pmatrix} S_k(t) \\ S_k(t - \delta) \end{pmatrix} \quad (2)$$

where  $T$  is a constant and  $\delta$  is a randomly selected time offset.

Let  $E(\cdot)$  define the encoder network used in the contrastive model. If we train the model to discriminate between  $E(Y_k(t))$  and  $E(Y_k^*(t))$ , we obtain the representation of  $h(S_k(t)) \approx f^{-1}(S_k(t))$  [17]. Note however, that this model does not yet account for coherent radar signals from multiple receivers. To this end, MEDUSA uses a multi-head attention step [30] in its unsupervised model as described next.

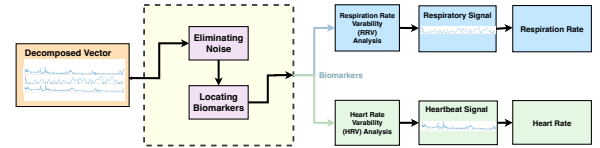
**3.4.2 Multi-Receiver Fusion.** Instead of utilizing the received signals from each antenna  $S_k(t)$  directly, MEDUSA uses a *multi-head attention layer* to fuse signal information from all the radar receivers, separately for the positive and negative augmented signals, prior to the contrastive training:

$$\begin{aligned} [Z_1(t), \dots, Z_W(t)] &= \mathbf{A}([Y_1(t), \dots, Y_M(t)]), \\ [Z_1^*(t), \dots, Z_W^*(t)] &= \mathbf{A}([Y_1^*(t), \dots, Y_M^*(t)]). \end{aligned} \quad (3)$$

Here  $W$  is the number of heads in the attention layer [30], and  $\mathbf{A}(\cdot)$  is the attention layer function that maps the  $M$  radar receiver signals into  $W$  head outputs. We then train the contrastive model to discriminate between the encoded pair  $E(Z_w(t))$  and  $E(Z_w^*(t))$  of each head output  $w \in \{1, \dots, W\}$ .

**3.4.3 Features and Vital Signs Identification.** MEDUSA identifies the breathing and heartbeat signals by analyzing the Respiratory Rate Variability (RRV) [21] and Heart Rate Variability (HRV) [26] of each of the  $W$  signals, as illustrated in Fig. 13. RRV and HRV are key indicators of general health and respiratory or cardiac complications. Normal breathing and heart-rate exhibit relatively constant rates and volumes, but variations within these rhythms are labeled as RRV and HRV, respectively. We use RRV and HRV analysis to identify the correct breathing or heart-beat signals from the output

features of the trained model. We use the extracted waveforms to identify if they are in respiratory rhythm or show normal variations in heart rate, and distinguish them into respiratory waveforms, heart-rate signals, or noise. This allows us to detect and track vital signs in the radar data accurately.



**Figure 13: Workflow of separating respiration, heart-rate, and motion patterns from the mixed radar signals. The resulting heartbeat and breath waveforms are identified.**

**3.4.4 Window Selection.** The UWB radar X4 chip extracts 186 range bins representing reflection amplitudes in distance. To reduce computing time, we use a window of size  $N$  to select a range bin subset before inputting the data into the model. We select the  $N$  consecutive bins with the highest reflected signal power, as these are most likely to contain human reflections. We choose a 30-bin section with higher amplitudes, reducing computational time for model training. This approach enables more effective detection and tracking of human reflections in radar data.

## 4 IMPLEMENTATION AND EVALUATION

We conduct MEDUSA's evaluation in two parts. First, we compare MEDUSA's custom-designed platform accuracy with that of the COTS mmWave MIMO radar (as indicated in Table 1) used previously in [11–13, 41–44]. Next, we evaluate MEDUSA's efficacy to measure the respiration and heart rate of multiple diverse targets in real-time, in real-world environments.

### 4.1 MEDUSA MIMO radar Micro-Benchmark

**AOA accuracy:** We compare the AoA accuracy of the MEDUSA's MIMO radar with the COTS UWB and mmWave radars in Table 1. We use a reflector box and position it at various distances and angles (but at the same height as the radar) measure the range and AoA of the strongest reflected signal, and compare it with the ground-truth. Fig. 14 shows



the AoA performance. For AoA in LoS, measured at 5m distance, MEDUSA MIMO radars perform better than the COTS mmWave radar. Median angular errors for MEDUSA is only 2 deg and a max AoA error of 8.2 deg, while the COTS mmWave radar's AoA median errors are 6.2 deg, while max angular error can go up to 12.5 deg.

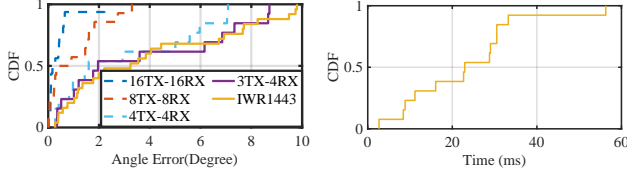


Figure 14: AoA accuracy. Figure 15: CDF of Latency.

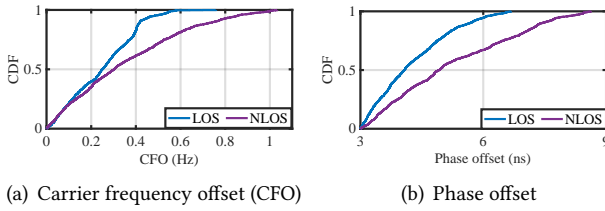


Figure 16: Wireless clock sync. in LoS and NLoS

**System Latency and Wireless Coherency:** Fig.15 shows a median latency of 22.94ms for respiration waveform reconstruction, from the time radar data was received, proving that MEDUSA can indeed run in real-time.

Next, we show the efficacy of MEDUSA's of wireless clock synchronization within reasonable ranges. As shown in Fig. 16(a), the median carrier frequency offset (CFO) is 0.25Hz in LoS and 0.3Hz in NLoS. Fig. 16(b) displays the phase offset during the wireless synchronization of clock signals.

## 4.2 Experiment Setup and Baseline

**Ground truth (GT):** We employed Vernier's breathing belt [4] and heart rate sensor [2] as Ground Truth (GT) for vital signs, which are also commonly utilized in prior research [7, 11, 13, 14, 44]. All the human participants in our evaluation wear these sensors during experiments. Necessary consent and IRB approvals were sought before all experiments. Fig 17 shows the overall experiment setup. The baseline TI radar sensor was aimed at the human chest. Operating at a lower frequency, UWB radars have wider beams that cover the entire body, allowing for a flexible height as depicted in Figure 17.

**Baselines:** In our evaluation, we employ two baselines. For respiration pattern detection, we utilize prior work [11, 13, 44] that uses Novelda UWB radar as a baseline. For heart rate monitoring, we compare MEDUSA with RFSCG [15], which is implemented on the TI mmWave radar IWR1443 [1].

**Data Collection:** We collected data from 27 volunteers (14 men, 13 women) aged 21-34 years (average age 25), weighing

52-102 kg (average 81.2 kg), and ranging in height from 164-187 cm (average 175 cm). Data collection took place in four indoor locations, primarily in a lab measuring  $18ft \times 30ft$  (540 sq ft), which exceed the size of average US bedrooms. Volunteers wore casual attire during 10-minute sessions, and none reported cardiovascular issues. We collect data of each person performing the following actions: (1) *static dataset*: standing, sitting with their body oriented in different directions. (2) *mobile dataset*: Arbitrary walking and jogging across a room in different directions, standing up and sitting down in continuous and staggered motion, jogging at the same spot with different body orientations, and performing various hand gestures. In total, we collected 3.75TB of data: 1.89TB from static activities and 1.86TB from mobile activities. We split the data into training and test sets with an 80:20 ratio, allocating approximately 3.04TB for training and 710GB for testing. The model underwent training with the designated training dataset and evaluation with the test dataset.

## 4.3 Unsupervised model

MEDUSA leverages an unsupervised learning model to recover human vital signals, and here we describe the network architecture and training process in detail.

**Network structure:** The unsupervised multi-head attention model includes a classifier and an encoder with a multi-head attention layer, implemented using TensorFlow by Python 3.8. Binary cross-entropy is utilized to classify loss. We added a multi-head attention layer before the Dense layer to assign weights to each MIMO antenna array, which helps the system focus on the most significant input data from influential antennas.

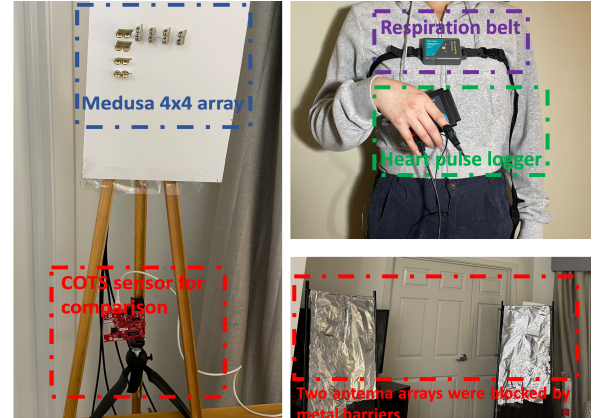
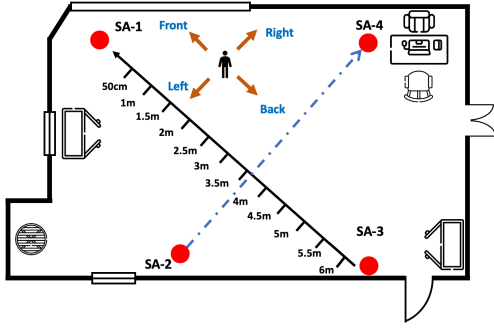


Figure 17: Experiment setup with MIMO radars, GT respiration belt, heart rate sensor and IWR1443BOOST mmWave sensor. And also the screen shield barrier for NLoS experiments.

**Training details:** The model is trained on a PC featuring an AMD Ryzen 6900 CPU and an RTX 3080 Ti graphics



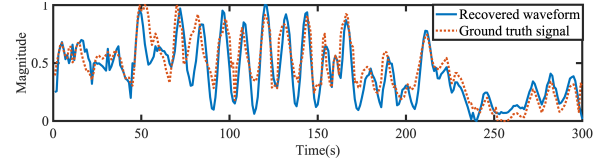
**Figure 18: Top-view of experiment-setup with four 4x4 radar placements (marked in red).**

card. Upon deploying the trained model, data is transmitted from the baseboard to the PC through a high-speed Ethernet interface. For a measurement duration of 2 minutes, the collected data has a size of 1.8GB and requires approximately 2 hours of model training using a single RTX 3080 Ti graphics card with 200 iterations. During training, datasets are split 80:20 for training and testing, using the leave-one-person-out cross-validation (LOPO CV) technique. The model is trained with the training dataset and evaluated with the test dataset, iteratively executed for each individual, ensuring each person's data is used as the test dataset once. This process is carried out iteratively for each participant in the dataset, guaranteeing that each individual's data is used as the test dataset precisely once.

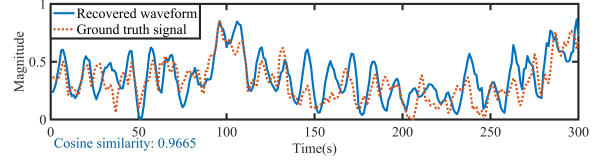
#### 4.4 Performance Evaluation:

**4.4.1 MEDUSA: Accuracy and Robustness.** We begin evaluating MEDUSA's robustness to monitor the Respiration-Per-Minute (BPM) of static and mobile users in LoS and NLoS when they are oriented in different directions. The experiment setup with the room layout, radar locations, items of furniture, as well as distances between radars are shown in Fig 18. The four  $4 \times 4$  subarray of MEDUSA is located at corners and we named them SA-1, SA-2, SA-3 and SA-4. NLoS experimental setup employs a common approach in that we use screen shields as obstacles between the target and radars rather than placing the target near the radars. This configuration more accurately simulates real-world NLoS conditions.

**Waveform Recovery:** Fig 19 displays the reconstructed waveform from MEDUSA for a randomly chosen participant positioned at 5 meters and facing left. The reconstructed waveform closely resembles the actual ground-truth waveform, exhibiting a cosine similarity of 0.987. Fig 19(b) demonstrates the increased complexity of respiration waveform recovery in Non-Line-of-Sight (NLoS) conditions, as evidenced by the reduced cosine similarity of 0.966.

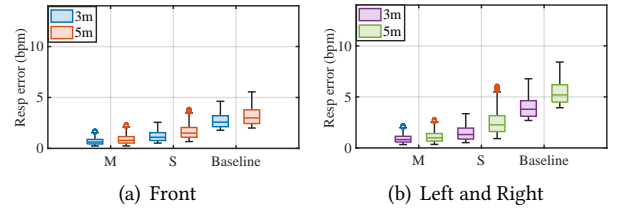


(a) Respiration waveform reconstructed for a moving target in LoS



(b) Respiration waveform reconstructed for a static target in NLoS

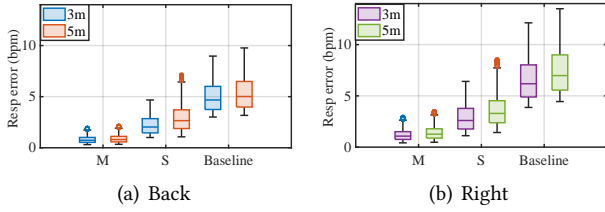
**Figure 19: Respiration waveforms reconstructed by MEDUSA.**



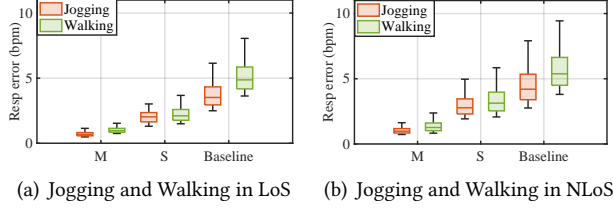
**Figure 20: LoS Static targets: BPM with different orientations. Comparison of MEDUSA ('M'), Co-located single-view MIMO radar ('S'), and Baseline.**

**Static Line-of-Sight (LOS):** Fig 20 presents the box-plot of respiration accuracy for static users (measured individually) when they are either standing or sitting and facing two different directions (body orientations) at distances 3m and 5m distance. The respiratory rate accuracy decreases at 5m for the single-view  $16 \times 16$  radar facing frontward, dropping by 12.1% when compared to the accuracy when facing left. The baseline's performance is worse than the single-view  $16 \times 16$  radar due to its poor SNR, especially when the user is 5m away, facing either front or left. MEDUSA, on the contrary, with  $4 \times 4$  MIMO distributed subarrays compensates for the drop in SNR (see Fig 17) when the target is 5m away by combining signals from the other three radars, providing an improvement of 5.8% and 18.5% over the baseline and single-view deployments, respectively. The benefits of MEDUSA are seen further when the target is oriented to the left and right at the 5m mark. When the baseline and the single-view solutions suffer, MEDUSA still estimates respiration with median errors  $< 2.8$  bpm.

**Static None-Line-of-Sight (NLOS):** To create NLoS conditions, we manually place the barrier screen (refer to Fig 17) in front of the back and right radars. Fig 21 shows the box-plot for respiration errors for all the individuals standing at the same 3m and 5m mark but oriented back and right (refer to Fig 18). When participants face the back array, MEDUSA's median errors are 2.15 bpm at 3 meters and 2.33 bpm at 5



**Figure 21: NLoS static users: Respiration errors (bpm) with different orientations.**



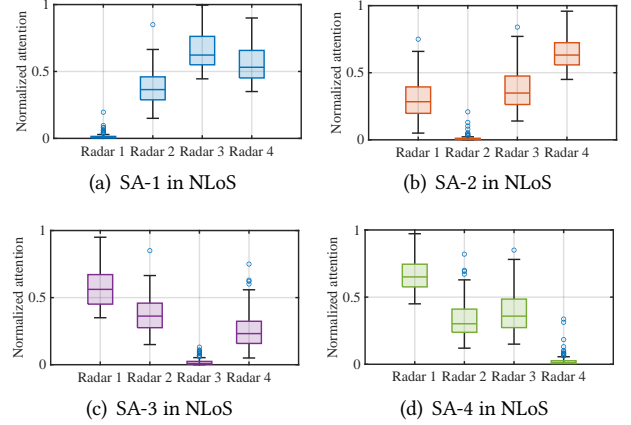
**Figure 22: Respiration errors (bpm) for targets with movements in LoS and NLoS.**

meters. Meanwhile, when facing the right, the median errors measured by MEDUSA are 2.21 bpm at 3 meters and 2.47 bpm at 5 meters, a gain of 17.5% and 21.9% over single-view and baseline solutions at 5 meters, respectively.

**Mobility in LoS and NLoS:** We next show the result when targets are mobile in LoS and NLoS (2 covered radars) conditions. Fig 22 showed the box plot of the average respiration errors per target when walking and jogging in random directions in LoS and in NLoS. In LoS, MEDUSA’s measured errors are 2.15 bpm and 2.36 bpm during walking and jogging, respectively, while in NLoS, MEDUSA’s measured median respiration errors are 2.67 bpm and 2.56 bpm, respectively. On the other hand, the single-view ( $16 \times 16$ ) single radar and the baseline solutions suffer, especially in NLoS conditions, with their respective median BPM errors increasing to 14.9% and 28%, respectively. MEDUSA delivers median gains of 10.3% and 21.7%, and max. gains of 21% and 38% over the single-view and baseline solutions, respectively.

**4.4.2 Impact on Attention-weights in NLoS.** We show how MEDUSA modifies attention weights for each radar depending on the received signal quality. In this experiment, we sequentially block one subarray (SA) at a time (SA-1 to SA-4) and measure the respiration of static targets at 3m. Fig 23 illustrates the normalized attention weight distribution when different SAs are obstructed. MEDUSA reduces attention weights for radars with poor SNR (NLoS) and increases weights for radars with higher SNRs (LoS). Low SNR SAs are given lower preference, focusing instead on radars with better-received signals.

**4.4.3 MEDUSA in unknown environments.** Next, we evaluate MEDUSA’s ability to generalize to different environments

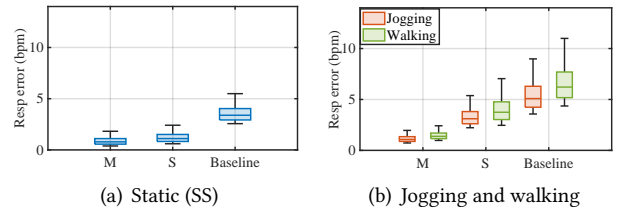


**Figure 23: Normalized attention weights when one of the  $4 \times 4$  subarray (SA) of MEDUSA in NLoS.**

than where it was trained to ensure a true “in-the-wild” deployment. We train MEDUSA using data from a residence and use the model in a university lab.

**Targets in LoS:** We perform static and mobility experiments with various targets in the radars’ LoS. In the static experiment, targets stand anywhere in the room, in any orientation. In the mobility experiment, targets continuously walk and jog within the room. Fig 24 displays BPM errors for static and mobile experiments. MEDUSA’s median BPM errors are 1.98 bpm for static targets and 2.12 bpm and 2.46 bpm for jogging and walking, respectively. MEDUSA outperforms the single-view and baseline solutions with accuracy gains of 11.2% and 26%, respectively.

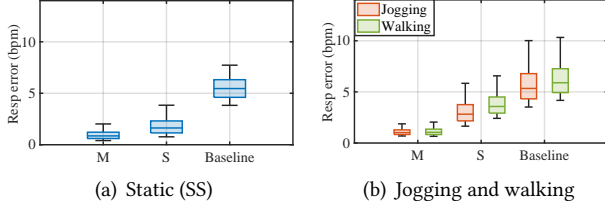
**Targets in NLoS:** We evaluate BPM errors for static and mobile targets in NLoS conditions, as shown in Fig 25. With two blocked radars, single-view and baseline solutions’ median errors shoot up to 2.67 bpm and 5.98 bpm for static targets, with max errors of 18.3% and 38%, respectively, while MEDUSA’s median error is 2.21 bpm, improving median accuracy by 2.4% and 20.8% over the other solutions. For mobile targets, MEDUSA accuracy gains are 2.32 bpm and 19.2% for jogging and 2.43 bpm and 24.4% for walking, over the single-view and baseline solutions, respectively.



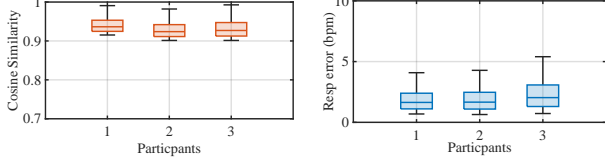
**Figure 24: Respiration errors (bpm) in LoS of Static and Moving targets in the untrained environment.**

**4.4.4 Case of Multiple Targets.** With multiple targets inside a room, MEDUSA can extract each individual’s respiration





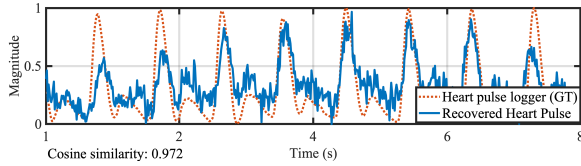
**Figure 25: Respiration errors (bpm) in NLoS of Static and Moving targets in the untrained environment.**



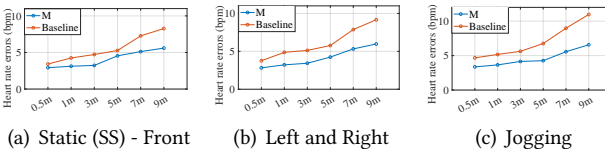
(a) Cosine similarity of the recovered respiration waveform (b) Respiration Error

**Figure 26: Cosine similarity and BPM of respiration rate detection for multiple targets.**

waveforms by decomposing the composite RF signals. However, it cannot map these extracted waveforms to individual targets. Mapping the recovered individual waveforms to specific targets requires continuous localization and tracking of each target, which we leave as future work. Nonetheless, we show MEDUSA's ability to recover the individual respiration waveforms and to accurately measure the individual BPM by manually associating the recovered waveform to the target. Fig 26 shows the cosine similarity of the recovered waveforms for three mobile targets in a room and their BPM accuracy. Median cosine similarity for the three waveforms are 0.923, 0.934, and 0.9216, respectively, and median errors are 2.14 bpm, 2.23 bpm, and 2.57 bpm.



**Figure 27: Recovered waveforms of heart pulses. Red dotted lines represent the ground-truth. Blue lines show the recovered waveforms of heart pulses by MEDUSA.**



**Figure 28: Heart rate error (bpm) across different distances (0.5m - 9m) of static and jogging targets.**

**4.4.5 Measuring Heart Rate.** Lastly, we demonstrate MEDUSA's heartbeat sensing capability. To comprehensively evaluate heartbeat detection, participants were asked to perform random motions at varying distances 28. We compare our results with the COTS mmWave TI radar [1] used by previous work [13, 15] as the baseline. Fig. 27 illustrates the recovered waveform, while Fig. 28 presents the heart rate error. At the closest 0.5m distance, MEDUSA clearly offers the accurate estimation for both static subjects (3.22 bpm - front, 3.47 bpm - left and right) and subjects with motion (3.77 bpm), surpassing the baseline solution (4.87 bpm). As distance increases, the performance of the baseline, which uses a single-view TI mmWave radar, significantly worsens compared to MEDUSA, with a widening discrepancy. The baseline achieved a heart rate error exceeding 10 bpm when tracking moving subjects at 9m, underscoring the limitations of single-view systems in practical scenarios.

## 5 RELATED WORK

Real-world applications demand robust human vital monitoring systems resilient to changes in target position, range, orientation, motion, as well as environmental factors like line-of-sight (LoS) and non-line-of-sight (NLoS) blockages.

**RF-based health sensing:** Research has made significant advances in using RF to passively (contact-free) sense people and their vital signs (i.e., human breathing and heart rates).

**Radar sensing:** Researchers in [11–13, 15, 41, 42, 44] have leveraged the large channel bandwidths available in UWB and mmWave bands to passively capture human breathing and heart rate. However, with most of these solutions employing either the COTS UWB radar [3] with a single Tx/Rx antenna that the operational range is limited to only 3-5m distance, and some of the work the mmWave radar array (4 Tx, 3 RX [1]) operating at a high (75GHz-77GHz) frequency, even in LoS and within a narrow FoV of  $\approx 50^\circ$ , making these solutions inadequate for practical room-scale deployments, especially with NLoS. Moreover, having a single radar additionally constrains the target to be oriented in a particular direction (e.g., facing the radar) to measure the physiological movements accurately.

On the other hand, solutions [19, 22, 32] that use two single-antenna radars, or a single radar with two antennas [29], require the two elements to be placed at specific locations with respect to a human body (e.g., front and back of a target) to eliminate signal-blockages caused by human action and body orientation, providing the simplest form of diversity, but are unable to offer practical ranges or NLoS operation with single antenna elements. Other solutions [6, 13, 18, 23, 35, 38, 43] that use COTS devices, while capable of measuring vitals of multiple targets within an enclosed space, can do so only when the targets are reasonably stationary and facing the radar in LoS. In [13], the



approach of using COTS mmWave sensor takes advantage of a well-established correlation between HR/RR and motion intensity, moving away from the traditional time-frequency analysis, which is often disrupted by motion. However, it faces challenges in NLoS conditions and multiple targets. Additionally, [38] introduced a method that employs linear ICA to decompose respiratory signals from multiple individuals who remain stationary, such as those asleep in set positions.

**Wi-Fi sensing:** WiFi-based sensing solutions [20, 24, 39, 40] rely on Channel State Information (CSI) to capture the minute displacements on a human body. However, with CSI being extremely sensitive to the environment, these solutions need extensive calibration and fingerprinting, making them impractical for *in-the-wild* deployments.

**Non-RF based health sensing:** In addition to RF-based sensing, researchers have explored acoustic [28, 31, 33, 34] and camera-based [10, 25, 36] solutions to monitor breathing and heart rates, respectively passively. However, they require targets to be relatively static in LoS, facing the radar and are impacted by background noise and human motion.

To the best of our knowledge, existing solutions cannot simultaneously cater to the different dimensions of target and environment dynamics, which are central to delivering a robust and practical HVM solution. MEDUSA takes an important step in this direction towards robust HVM sensing.

## 6 CONCLUSION

In summary, MEDUSA presents a novel multi-view 256-element virtual MIMO radar system that facilitates robust and precise distributed radar sensing across diverse target-environment configurations. Our comprehensive experiments demonstrate that MEDUSA significantly outperforms existing solutions in terms of vital-sign measurement accuracy for both stationary and moving targets. MEDUSA lays the foundation for a feasible, contact-free vital monitoring solution that can be effectively implemented at scale in real-world settings.

## REFERENCES

- [1] IWR1443. <https://www.ti.com/tool/IWR1443BOOST/>.
- [2] Neulog Heart Pulses Logger. <https://neulog.com/heart-rate-pulse/>.
- [3] Novelda XeThru X4. <https://novelda.com/>.
- [4] Respiration Belt. <https://www.vernier.com/product/go-direct-respiration-belt/>.
- [5] Xilinx Zynq UltraScale+ MPSoC XCZU4EV. <https://www.xilinx.com/products/silicon-devices/soc/zynq-ultrascale-mpsoc.html>.
- [6] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, page 837–846, New York, NY, USA, 2015. Association for Computing Machinery.
- [7] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, page 837–846, New York, NY, USA, 2015. Association for Computing Machinery.
- [8] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor. *MIMO wireless communications*. Cambridge university press, 2007.
- [9] F. A. Butt, M. A. Aslam, M. T. Zafar, I. H. Naqvi, and U. Riaz. Synchronization of Long-Range, Widely-Separated MIMO Radar Network using GSM Protocol. In *2018 19th International Radar Symposium (IRS)*, pages 1–10, 2018.
- [10] W. Chen and D. McDuff. Deepphys: Video-based physiological measurement using convolutional attention networks. *CoRR*, abs/1805.07888, 2018.
- [11] Z. Chen, T. Zheng, C. Cai, and J. Luo. Movi-fi: Motion-robust vital signs waveform recovery via deep interpreted rf sensing. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking, MobiCom '21*, page 392–405, New York, NY, USA, 2021. Association for Computing Machinery.
- [12] Z. Chen, T. Zheng, and J. Luo. Octopus: A practical and versatile wideband mimo sensing platform. *MobiCom '21*, page 601–614, New York, NY, USA, 2021. Association for Computing Machinery.
- [13] J. Gong, X. Zhang, K. Lin, J. Ren, Y. Zhang, and W. Qiu. RF Vital Sign Sensing under Free Body Movement. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 5(3), sep 2021.
- [14] J. Gong, X. Zhang, K. Lin, J. Ren, Y. Zhang, and W. Qiu. RF Vital Sign Sensing under Free Body Movement. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2021.
- [15] U. Ha, S. Assana, and F. Adib. Contactless seismocardiography via deep learning radars. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2020.
- [16] A. Hyvärinen. Independent component analysis: Recent advances. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, page 20110534, 2013.
- [17] A. Hyvärinen and H. Morioka. Unsupervised feature extraction by time-contrastive learning and nonlinear ica. In *Proceedings of the 30th International Conference on Neural Information Processing Systems, NIPS'16*, page 3772–3780, Red Hook, NY, USA, 2016. Curran Associates Inc.
- [18] O. J. Kaltiokallio, H. Yigitler, R. Jäntti, and N. Patwari. Non-invasive respiration rate monitoring using a single cots tx-rx pair. In *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, IPSN '14*, page 59–70. IEEE Press, 2014.
- [19] C. Li and J. Lin. Random body movement cancellation in doppler radar vital sign detection. *IEEE Transactions on Microwave Theory and Techniques*, (12):3143–3152, 2008.
- [20] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng. Tracking vital signs during sleep leveraging off-the-shelf wifi. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc '15*, page 267–276, New York, NY, USA, 2015. Association for Computing Machinery.
- [21] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. H. A. Chen. NeuroKit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4):1689–1696, feb 2021.
- [22] J.-M. Muñoz-Ferreras, Z. Peng, R. Gómez-García, and C. Li. Random body movement mitigation for fmcw-radar-based vital-sign monitoring. In *2016 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless)*, pages 22–24, 2016.
- [23] P. Nguyen, X. Zhang, A. Halbower, and T. Vu. Continuous and fine-grained breathing volume monitoring from afar using wireless signals. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9, 2016.
- [24] J. Ni, F. Zhang, J. Xiong, Q. Huang, Z. Chang, J. Ma, B. Xie, P. Wang, G. Bian, X. Li, and C. Liu. Experience: Pushing indoor localization

- from laboratory to the wild. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*, MobiCom '22, page 147–157, New York, NY, USA, 2022. Association for Computing Machinery.
- [25] A. Pai, A. Veeraraghavan, and A. Sabharwal. Hrvcam: robust camera-based measurement of heart rate variability. *Journal of Biomedical Optics*, 26, 02 2021.
- [26] T. Pham, Z. J. Lau, S. H. A. Chen, and D. Makowski. Heart rate variability in psychology: A review of hrv indices and an analysis tutorial. *Sensors*, 21(12), 2021.
- [27] S. Prager, M. S. Haynes, and M. Moghaddam. Wireless Subnanosecond RF Synchronization for Distributed Ultrawideband Software-Defined Radar Networks. *IEEE Transactions on Microwave Theory and Techniques*, 68(11):4787–4804, 2020.
- [28] X. Song, B. Yang, G. Yang, R. Chen, E. Forno, W. Chen, and W. Gao. Spirosonic: Monitoring human lung function via acoustic sensing on commodity smartphones. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, MobiCom '20, New York, NY, USA, 2020. Association for Computing Machinery.
- [29] M.-C. Tang, F.-K. Wang, and T.-S. Horng. Single self-injection-locked radar with two antennas for monitoring vital signs with large body movement cancellation. *IEEE Transactions on Microwave Theory and Techniques*, 65(12):5324–5333, 2017.
- [30] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [31] A. Wang, J. E. Sunshine, and S. Gollakota. Contactless infant monitoring using white noise. In *The 25th Annual International Conference on Mobile Computing and Networking*, MobiCom '19, New York, NY, USA, 2019. Association for Computing Machinery.
- [32] F.-K. Wang, T.-S. Horng, K.-C. Peng, J.-K. Jau, J.-Y. Li, and C.-C. Chen. Single-antenna doppler radars using self and mutual injection locking for vital sign detection with random body movement cancellation. *IEEE Transactions on Microwave Theory and Techniques*, 59(12):3577–3587, 2011.
- [33] T. Wang, D. Zhang, Y. Zheng, T. Gu, X. Zhou, and B. Dorizzi. C-fmcw based contactless respiration detection using acoustic signal. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(4), jan 2018.
- [34] X. Xu, J. Yu, Y. Chen, Y. Zhu, L. Kong, and M. Li. Breathlistener: Fine-grained breathing monitoring in driving environments utilizing acoustic signals. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '19, page 54–66, New York, NY, USA, 2019. Association for Computing Machinery.
- [35] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra. Vital sign and sleep monitoring using millimeter wave. *ACM Trans. Sen. Netw.*, 13(2), apr 2017.
- [36] Z. Yu, W. Peng, X. Li, X. Hong, and G. Zhao. Remote heart rate measurement from highly compressed facial videos: an end-to-end deep learning solution with video enhancement, 2019.
- [37] S. Yue, H. He, H. Wang, H. Rahul, and D. Katabi. Extracting Multi-Person Respiration from Entangled RF Signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, jul 2018.
- [38] S. Yue, H. He, H. Wang, H. Rahul, and D. Katabi. Extracting multi-person respiration from entangled rf signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(2), jul 2018.
- [39] Y. Zeng, D. Wu, R. Gao, T. Gu, and D. Zhang. Fullbreathe: Full human respiration detection exploiting complementarity of csi phase and amplitude of wifi signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(3), sep 2018.
- [40] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang. Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas. 3(3), sep 2019.
- [41] F. Zhang, J. Xiong, Z. Chang, J. Ma, and D. Zhang. Mobi2sense: Empowering wireless sensing with mobility. MobiCom '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [42] S. Zhang, T. Zheng, Z. Chen, and J. Luo. Can we obtain fine-grained heartbeat waveform via contact-free rf-sensing? In *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*, pages 1759–1768, 2022.
- [43] T. Zheng, Z. Chen, C. Cai, J. Luo, and X. Zhang. V2ifi: In-vehicle vital sign monitoring via compact rf sensing. 4(2), jun 2020.
- [44] T. Zheng, Z. Chen, S. Zhang, C. Cai, and J. Luo. More-fi: Motion-robust and fine-grained respiration monitoring via deep-learning uwb radar. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, SenSys '21, page 111–124, 2021.