

TomoID: A Scalable Approach to Device Free Indoor Localization via RFID Tomography

Yang-Hsi Su, Jingliang Ren, Zi Qian, David Fouhey, and Alanson Sample
University of Michigan, Ann Arbor, MI, USA



Fig. 1: TomoID is a real-time indoor localization system based on battery-free RFID that can locate multiple users in real-world environments. Panel ‘A’ shows two users being simultaneously located in a conference room with the corresponding heatmap. Using the same machine learning model, Panel ‘B’ shows a user being located in a digital fabrication laboratory without the need for additional training data. Panel ‘C’ shows TomoID implemented in a hallway, using the same pre-trained model with the tags and readers concealed behind the wall.

Abstract—Device-free localization methods allow users to benefit from location-aware services without the need to carry a transponder. However, conventional radio sensing approaches using active wireless devices require wired power or continual battery maintenance, limiting deployability. We present TomoID, a real-time multi-user UHF RFID tomographic localization system that uses low-level communication channel parameters such as RSSI, RF Phase, and Read Rate, to create probability heatmaps of users’ locations. The heatmaps are passed to our custom-designed signal processing and machine learning pipeline to robustly predict users’ locations. Results show that TomoID is highly accurate, with an average mean error of 17.1 cm for a stationary user and 18.9 cm when users are walking. With multi-user tracking, results showing an average mean error of <72 cm for five individuals in constant motion. Importantly, TomoID is specifically designed to work in real-world multipath-rich indoor environments. Our signal processing and machine learning pipeline allows a pre-trained localization model to be applied to new environments of different shapes and sizes, while maintaining good accuracy sufficient for indoor user localization and tracking. Ultimately, TomoID enables a scalable, easily deployable, and minimally intrusive method for locating uninstrumented users in indoor environments.

Index Terms—Indoor Localization, Radio Tomography, RFID, Multi-User, Real-Time

I. INTRODUCTION

Effective means of determining one’s indoor location can empower users to explore their indoor environments, provide them with pertinent and contextually aware information about their surroundings, and enable their living spaces to respond and adapt to their needs. Research efforts in this area have focused on using localization methods to solve important problems such as indoor navigation [1], [2], assisted living [3], improving healthcare [4], and in-home activity detection [5]. Both academic and industry researchers have investigated a wide variety of underlying physical phenomena for localization [6] as well as localization algorithms and deployment strategies [7].

While these systems demonstrate impressive results, they rely on the fundamental assumption that users are *actively*

participating, by continuously wearing a localization hardware or carrying a smartphone. This assumption makes sense in highly interactive and purposeful usage scenarios such as navigating an unknown airport or shopping plaza. However, there are a wide number of application scenarios where it may be impractical to instrument all possible users. Users may not be able to wear or carry a device, or users may not have hardware or software that is compatible with the positioning system installed in the environment. This raises the need for device-free indoor localization systems that do not burden users with the requirement of continuous instrumentation.

In this work, we propose TomoID, a real-time, RFID-based tomography approach that uses radio waves to visualize the interior of living spaces to determine the location of users, as depicted in Figure 1. In this approach, battery-free RFID sticker tags are placed around the perimeter of living spaces along with RFID readers. While modeling radio propagation in the real world, multi-path environments remains an open challenge. Our signal processing pipeline utilizes low-level RFID communication channel parameters such as RSSI, RF Phase, and Read Rate to create tomograms (i.e. probability heatmaps) which depict disturbances in the radio traffic. This allows us to transform a relatively hard RF localization problem into the computer vision domain, a mature field that excels at extracting information from images. Our enhanced tomograms are passed through a custom-designed pipeline that consists of a modified U-Net convolutional neural network which is designed for image segmentation tasks followed by a Kalman filter for temporal smoothing. Results show that our system has an average accuracy of 17.1 cm for stationary users and 18.9 cm for mobile users at an update rate of one frame per second. Furthermore, TomoID demonstrates multi-user, device-free tracking of two individuals with an average accuracy of 39.6 cm, and up to five people continually moving with an average accuracy of 70 cm, which is sufficient for most indoor localization applications given that the average shoulder width of males and females is 46.5 cm and 39.5 cm.

In order to ensure that TomoID is a generalizable solution that is not specific to a particular testing environment, we demonstrate that our pre-trained model can be successfully applied to new environments of different size and geometry, with different reader configurations and number of RFID tags, while retaining sufficient accuracy for indoor localization tasks. Additionally, we show two applications of TomoID, tracking usage of a fabrication laboratory and monitoring the traffic of a hallway. Ultimately, TomoID enables a scalable and easily deployable method for locating uninstrumented users in real-world, multi-path rich environments.

This work makes the following contributions:

- The use of RF Phase and Read Rate channel parameters (along with RSSI) to create tomograms
- Highly accurate real-time localization of individuals and multiple people when stationary and moving
- A machine learning approach that allows pre-trained models to be deployed to new real-world environments with sufficient accuracy for indoor localization applications

II. RELATED WORK

Indoor positioning systems have drawn significant attention in recent years as researchers have sought to create location-aware services and applications. A wide range of physical phenomena such as visible and infrared light [8], [9], ultrasound [10], as well as capacitive and magnetic beacons [11], [12] have been investigated, with the use of propagating Radio Frequency (RF) waves emerging as the predominate method of choice [7], [13]. The following subsections review related work in the area of Device-Free localization and draw a distinction between *active* systems that generate and transmit radio signals, and *passive* systems that backscatter (or reflect) signals in order to transmit data.

A. Device-Free, Active RF Localization

Device-free, active RF localization systems have received a notable amount of attention as they do not require the user to be tagged in any way. Typically, these systems use the same radio hardware designed for wireless communication, but instead of focusing on transferring data, these systems monitor metadata related to the state of the communication channel (RSSI [14], [15], CSI [16], [17], etc.) for disturbances caused by the presence of people. Other examples include radio imaging systems that perform frequency sweeps to locate users by finding the blobs in the reconstructed images, which is computationally heavy and hard to implement in real time. Examples include millimeter-wave radar [18], [19], WiFi based approaches [20], [21], as well as traditional radar approaches [22], [23]. These systems can achieve good accuracy but have several drawbacks, including the use of custom hardware instead of commercially available hardware, the potential for interfering the existing wireless network, and the need of a custom-trained model for each space.

Active Radio Tomographic Imaging (RTI) systems place evenly spaced radio nodes around the perimeter of the

environment to be monitored. Each battery powered radio transceiver sequentially transmits to the other receivers. By monitoring the changes in RSSI caused by the presence of a person, it is possible to create a top down heatmap showing the location of the person [24], [25] with results showing decimeter to centimeter location accuracy.

While this method is conceptually similar to TomoID, active RTI systems require a continuous power source for each node. Thus, the high installation cost of adding wired power to each node or the need to maintain the sensor node's battery level actively limits the use of this approach, lowering the possibility of large-scale deployment. In contrast, TomoID has the advantage of using UHF RFID tags which cost as little as 15¢ per node [26], do not require batteries that need to be replaced, and uses a different physical layer that provides richer channel state information to boost performance.

B. Device-Free, Passive RFID Localization

Device-free localization of people and objects using passive RFID has also received research attention as system designers take advantage of the low-cost and battery-free nature of RFID tags. One popular approach is to deploy a two-dimensional grid of tags on walls [27], [28], floors [29], [30], poster board [31], and on keypad-style input devices [32]. When a user either touches a tag or enters its near-field region, the RFID reader will report a change in RSSI and RF Phase allowing the system to make a localization estimation based on which tag was touched or interfered. For these systems, location resolution is typically based on tag grid spacing, with room-scale applications having a localization resolution of 30 cm and a keypad application having a resolution of 3 cm.

Tomography-based approaches using passive RFID tags have also been explored in literature [33]–[36], but often simplify the testing environment or constrain the usage scenario to make the problem tractable. Notable examples include Wagner et al. [33] which uses specialized bi-static RFID readers that can transmit and receive on multiple readers antenna simultaneously to compute a full tomographic matrix, resulting in 45 cm accuracy. Ma et al. [37] use commercially available RFID readers and have demonstrated 24 cm localization accuracy, but required a multi-path free environment and used stationary aluminum foil targets as a stand-in for people. TomoID makes significant advances in the use of RFID tomography for people tracking, by creating a real-time system that can track multiple users freely walking around an unconstrained and multi-path rich environment. We further differentiate our work by creating a computer vision-based machine learning framework that, once trained, can be applied to new indoor environments with significantly different geometries, such as non-square-shaped rooms and hallways. The system's flexibility significantly decreases deployment burden while transferring the system to new environments whereas maintaining high localization accuracy.

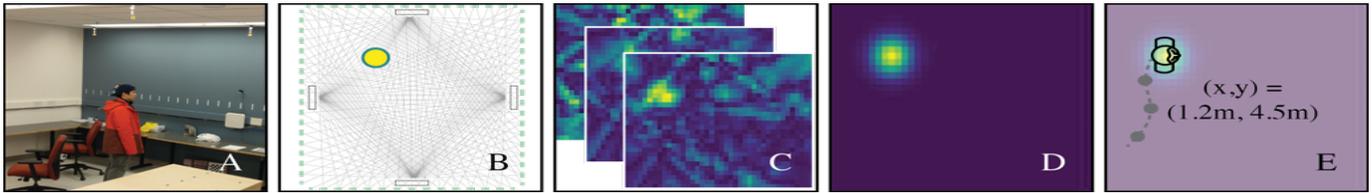


Fig. 2: An overview of the signal processing flow to transform RFID data into human location. Panel ‘A’ shows a tagless user in the office environment with RFID tags and antennas on the wall. Panel ‘B’ shows the invisible RFID signals in the air being blocked by the user. In panel ‘C’, 3 heatmaps are created using tomographic imaging methods generated from RSSI, RF Phase, and Read Rate channel parameters. Panel ‘D’ shows the denoised and segmented probability map created by a CNN from the 3 heatmaps in panel ‘C’. Panel ‘E’ shows the calculated current position of the user and the trajectory smoothed by a Kalman filter.

III. SYSTEM OVERVIEW

This section provides a holistic overview of TomoID’s components and their operation. Subsequent sections dive deeper into each of these components and provide greater details on methodology and implementation. At the highest level, TomoID uses an array of passive RFID tags along the perimeter of a living space to create a tomogram, also called a heatmap, of the interior of the room. Our signal processing and machine learning pipeline uses these heatmaps to make real-time predictions of users’ locations.

Unlike previous works, TomoID has been designed from the ground up to work in real-world living and office spaces as shown in Figure 2 panel ‘A’. RFID tags in the form of stickers are placed directly on the wall along the perimeter of the room using a basic cardboard template. In the long-term, tags can be manufactured into (or on the backside of) the drywall for easy deployment or in the form of wallpaper. A single Impinj Speedway R420 UHF RFID reader is used with one of the four antennas placed on each wall facing inward, and can be placed on, within, or even behind the wall. To further increase the deployability, equivalent UHF RFID readers with integrated RFID antennas and Wi-Fi capability such as ThingMagic A6-IN-WIFI or Alien ALR-9650 can be used to reduce cabling overhead as they only require wall power or Power over Ethernet (PoE). The RFID reader interrogates the tags in a probabilistic fashion as defined by the EPC Class 1 Gen 2 (ISO-18000-6C) protocol and reports back low-level channel parameters such as RSSI, RF Phase, and Read Rate of each tag. For simplicity, these read transactions can be thought of as a point-to-point connection between the tag and the reader, as shown in Figure 2 panel ‘B’. When a person enters the room, they block or disrupt some of the RF paths. Using a tomographic approach, we can create tomograms showing a virtual “hole” in the signal path as shown in figure 2 panel ‘C’, which visually depicts the RSSI, Phase, and Read Rate tomograms.

The challenge when using RF radio systems in real-world conditions is obstacles, especially metal objects and surfaces with complex geometries. These obstacles introduce nonidealities in signal propagation such as multi-path, scattering, and near-field effects. As a result, the tomograms often visually appear to be corrupted or incomplete, and it is not possible to use linear or rule-based methods to determine the person’s location with high accuracy. Furthermore, given the probabilistic nature of the RFID protocol used by the commercially

available off-the-shelf readers, coupled with the variations in per tag’s channel condition, the tomograms will be missing data and leads to limiting the maps’ update rate. To overcome these issues, a custom-designed convolutional neural network (CNN) based on U-Net [38] was optimized for model size and computation speed. Results of the trained network are shown as a probability map in Figure 2 panel ‘D’. Finally, the users’ (X, Y) positions are extracted from the probability map, and a Kalman filter is added at the end to smooth the trajectory as shown in Figure 2 panel ‘E’. The pre-trained model runs in real-time at a frame rate of 1Hz. Since the tomograph approach transforms the RF localization problem to a CV image processing problem, TomoID is able to apply the pre-trained model to other spaces and have good accuracy without collecting any new training data to update the model. Results will be described in Section VII and VIII, showing a real-time localization accuracy of 17.1 cm for a stationary user, 18.9 cm for a freely moving user, and near 70 cm for 3 to 5 users in motion. A demo video can be found here ¹.

IV. INSTRUMENTATION AND DATA ACQUISITION

This section outlines the initial deployment of readers and tags, which are placed uniformly around the perimeter of office space along with the data acquisition methodology. Section VII investigates which instrumentation and data acquisition parameters have the most significant impact on localization accuracy, and Section VIII analyzes which instrumentation parameters can be modified to improve system deployability.

To instrument the space, a line of RFID tags is placed approximately 1.4 m above the floor around the room’s perimeter, with a tag-to-tag spacing of 16.37 cm. This spacing is equal to the half-wavelength of the center frequency (915 MHz) of the UHF RFID band used in North and South America and has been selected to avoid coupling between adjacent tags. A simple cardboard template with spacing marks on it can speed up the installation process. Each reader antenna is mounted on the center of each wall below the line of tags. The position of each tag can be calculated once the location of the first tag is measured. The system takes the positions of the tags and the antennas to formulate the RF links shown in 2 panel ‘B’. Finally, the four antennas are connected to the reader, which sends data to the host PC via Ethernet. The whole instrumentation process can be done within 15 minutes

¹https://www.youtube.com/watch?v=XfPdPIWIY_I

or less, depending on the room’s size, by a person familiar with the system. Figure 1 A shows the conference room used for our initial testing and represents a typical usage scenario. For instance, the RFID tags are placed directly on the drywall mounted on steel studs, typical for commercial buildings in the United States. The suspended ceiling consists of metal tracks and gypsum ceiling tiles, and the floor is made of concrete strengthened with steel rebar. The conference room is filled with office desks, a large conference table, and chairs.

The reader operates in the 902 ~ 928MHz frequency band and pseudo-randomly frequency hops between 50 channels as required by the Federal Communications Commission (FCC). The output power of the RFID reader is 30dBm, and it is set to hybrid mode with ‘dual-target select’ ($B \rightarrow A$) search mode. This setting allows the reader to perform a deeper, cleaner scan of the whole tag population with the trade-off of decreasing absolute read rate of an individual tag.

Each tag exhibits different RF characteristics based on manufacturing variations and its surrounding environment. For instance, some tags may be placed near metal cabinets, electrical conduits, computer monitors, etc., while others will inevitably be placed near or directly in front of the steel support studs inside the wall, which are spaced 16 to 24 inches apart. All of which will cause an unknown amount of degradation in individual tag performance. To overcome variations in per-tag read performance, TomoID performs once daily background subtraction and adjusts the calculation matrix’s dimension dynamically, as described in the next section.

V. GENERATION OF RFID TOMOGRAMS

Radio Tomographic Imaging (RTI) methods transform a time series of RF measurements from multiple RF links into camera-like images called tomograms, which spatially convey information about interference in the communication channel between the reader and tags. RTI methods offer two advantages. First, they provide a formulation for inferring the obstacles in space, in this case humans in a living space. Second, they convert an RF signal interpretation problem into a computer vision problem, enabling the application of well-developed techniques for segmenting and localizing objects in images. Moreover, the tomogram itself converts information about link geometry (i.e., via the pixel-correspondence between locations in space in the tomograms and links) in a convenient format for subsequent learning systems.

A. Low-level UHF RFID signal characteristics of RF link

An RF link is formed when an RF signal is transmitted from the reader antenna, backscattered off an RFID tag, and then received by the same reader antenna. The RF measurements at the receiving end contain information of the physical space. While there will be minor variations over time, if a person blocks the *direct* line-of-sight of the RF link, there will be substantial variations: the received RSSI will be lower, the RF Phase will have a jump, and the Read Rate will drop.

As an illustrative example, Figure 3 shows a user walking perpendicularly through the LOS path of a single RF link

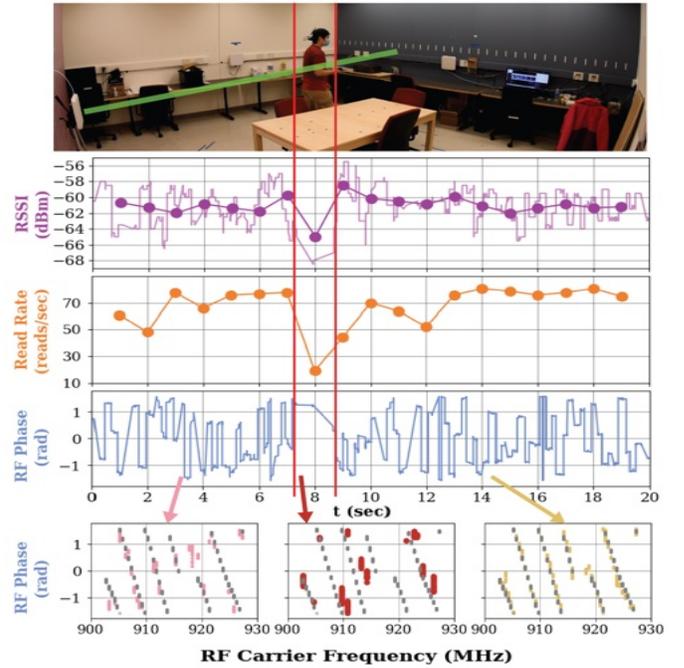


Fig. 3: Low-level RFID signals of an RF link (highlighted in green) over time as a user passes by with the dots showing averages over a one-second window. RSSI (purple) and Read Rate (orange) suffers a substantial drop at $t = 8$ seconds, while RF Phase has no obvious jump. The changes in RF Phase can be revealed by sorting the data with the carrier frequency used. The gray dots show the RF Phase when the room is empty without any users.

(highlighted in green). The purple and orange plots show RSSI and Read Rates respectively as a function of time, with dots showing averages over a one-second window. When the user walks into the LOS path at $t = 8$ seconds, there is a substantial dip in both RSSI and Read Rate. Due to the regulatory requirement that RFID readers pseudo randomly change their carrier frequency, the RF Phase (plotted in blue) appears to oscillate randomly. While a small change can be seen when the user walks into the path at $t = 8$, re-plotting the RF Phase data as a function of the carrier frequency (which is also reported by the RFID reader) reveals clear patterns when a person enters and leaves the LOS path. Using the background RF Phase measured (plotted in gray) when the room is empty with no users as a reference, an offset occurs at $t = 8$ (red). While the RF Phase before (pink) and after (yellow) $t = 8$ align with the reference. These variations in communication channel parameters caused by the presence of an individual or multiple people form the bases of the tomograms, which will be described next.

B. Mapping changes in an RF link onto pixels in tomograms

Our goal is to use the received RFID tag read data to infer the location of users and objects in the space. This formulation instantiates that there is a set of L RF links, and the space is discretized into a set of N grid cells. The changes in the physical space that happen in each cell contribute to the observed measurement results. Therefore, given the observation $\mathbf{y} \in \mathbb{R}^L$ (from the three RFID channel

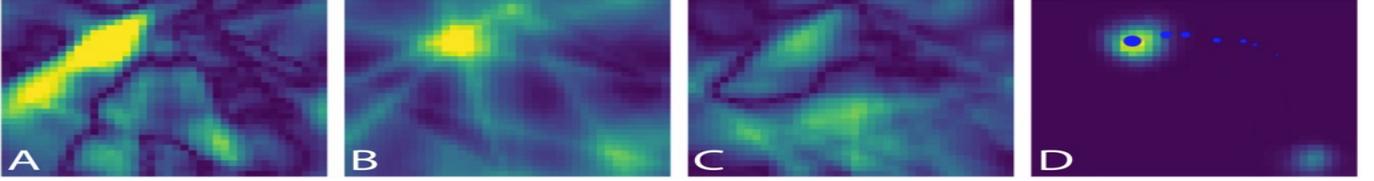


Fig. 4: A snapshot with one user in the space. Panel ‘A’ is the heatmap generated from RSSI, panel ‘B’ from RF Phase, and panel ‘C’ from Read Rate. Panel ‘D’ is the output of TomoID with the trace of the user in blue dots.

parameters, \mathbf{y}_{RSSI} , $\mathbf{y}_{RFPhase}$, and $\mathbf{y}_{ReadRate}$), we assume that it is generated by a linear function of information at each of the N grid cells; this information is represented by an $\mathbf{x} \in \mathbb{R}^N$, which is the vector form of a tomogram. Since we have the information about the RF link geometry, namely the approximate position of the antennas and the tags, we can construct a model mapping cell information to the RF link’s observations. The model is assumed to be linear, it takes the form $\mathbf{y} = \mathbf{W}\mathbf{x}$ where $\mathbf{W} \in \mathbb{R}^{L \times N}$ encodes link-cell information. From the low-level RFID signal characteristics discussed in the previous section, we have learned that the physical changes happening on the LOS path cause significant variations in the received RFID read data. We then define that cells in the direct LOS of the i th RF link contribute with weight w_i and zero otherwise, resulting in a \mathbf{W} of the form

$$\mathbf{W}_{ij} = \begin{cases} w_i, & \text{if link } i \text{ passes through cell } j \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The weight \mathbf{W} controls the relationship between the RF measurements \mathbf{y} and the tomogram \mathbf{x} . We define the weight w_i for an RF link i as the reciprocal of square root d_i , the distance from the antenna to the RFID tag for RSSI, RF Phase and Read Rate. When the RFID tag is far away from the antenna, the RF signal is weaker and is more likely to be degraded by noise or obstructions outside the LOS path. Therefore, the weight is designed to be smaller for longer RF links. Different weight models might construct tomograms that are visually brighter or darker. Despite the appearance of the tomograms, our custom neural net extracts information and generate a clean, less noisy tomogram as output which will be described in section VI.

With a set of observations \mathbf{y} and weight \mathbf{W} , one could solve for the tomogram \mathbf{x} that best explains the recorded observations in a least-square sense, or $\arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{W}\mathbf{x}\|_2^2$. The least-squares problem, however, is ill-posed and one has to impose a smoothing operation to the output. This objective results in the Tikhonov-regularized least-squares problem,

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{W}\mathbf{x}\|_2^2 + \alpha(\|\mathbf{D}_x\mathbf{x}\|_2^2 + \|\mathbf{D}_y\mathbf{x}\|_2^2), \quad (2)$$

where \mathbf{D}_x and \mathbf{D}_y are $N \times N$ horizontal and vertical difference operators in 2D, and α trades off between smoothness and reconstruction fidelity. Eqn. 2 can be minimized by directly solving the normal equations $(\mathbf{W}^T\mathbf{W} + \alpha(\mathbf{D}_x^T\mathbf{D}_x + \mathbf{D}_y^T\mathbf{D}_y))\mathbf{x} = \mathbf{W}^T\mathbf{y}$. Due to the randomness in the RFID communication protocol, unfortunately the number of responding links (and thus L) is different per-time step. TomoID varies the rows of \mathbf{W} to match the number of received RF links in the time window, enabling the RTI method to generate fixed-size \mathbf{x}

with varying sizes of \mathbf{y} at each time frame. To the best of our knowledge, this precludes the pre-computation of an inverse matrix for Eqn. 2. However, we find that solutions can be obtained sufficiently quickly in practice.

One remaining practical challenge is that one link may be weaker than another for reasons unrelated to a person walking through the link LOS path (e.g., due to nearby metal studs). To handle this, following [39], our observations \mathbf{y} represent the difference between a run-time observation \mathbf{y}' and calibration observation $\bar{\mathbf{y}}$ (or $\mathbf{y} = \mathbf{y}' - \bar{\mathbf{y}}$). Since the system is linear, the \mathbf{x}^* obtained from solving Eqn. 2 is the difference between run-time cell information \mathbf{x}' and calibration observation $\bar{\mathbf{x}}$. Thus, one can create a calibration set of observation data $\bar{\mathbf{y}}$ when the system $\bar{\mathbf{x}}$ is in a known state (e.g., nobody present). Experimental results show that calibration data only needs to be obtained once every one or two days as variations in tag performance changes slowly if at all. This data can easily be collected automatically by the system when users are likely to be asleep in the early mornings, and no movement in the room is detected. We obtain $\bar{\mathbf{y}}$ by taking the average of 5 minutes of data with no user in the space. Given this calibration information, the resulting reconstructions then show the difference from the calibrated state.

It should be noted that this construction of the tomograms helps resolve the variability of tag read rate and the probabilistic nature of the EPC Class 1 Gen 2 (ISO-18000-6C) protocol, which is particularly important for the neural network learning model discussed in the next section. The key insight is that tomograms \mathbf{x} are a uniform data representation that can be generated from sparse input data \mathbf{y} . For example, there is no guarantee that all the RFID tags will be read by the reader within a time window. Thus for a given time window the RSSI, RF Phase, and Read Rate data from the tags that are read will be used to update the weights of the pixels on the tomograms along their respective RF links, which is achieved by forming the weight matrix \mathbf{W} with only the RF links read. Missing tags will contribute no information to their respective RF links but importantly will not harm the construction of the tomograms. This provides system designers with flexibility in the duration of the acquisition time and the resulting system update rate. Details on the trade-off of window size and the use of history are described and explored in section VII.

Given three \mathbf{y} consisting of RSSI, RF Phase, and Read Rate, we produce three tomograms, shown in Figure 4. The tomograms have different blob sizes and shapes since the RF characteristics (including noise) of RSSI, RF Phase, and Read Rate are different. Sometimes the blob is not apparent in the RSSI tomogram but can be seen in the RF Phase or the

Read Rate tomogram. Thus the tomograms enrich each other compared to a traditional RSSI-based tomogram, which we demonstrate empirically. At the same time, given the different tomograms, it would be difficult to directly create an algorithm to extract the information. We thus turn to convolutional neural networks.

VI. CONVOLUTIONAL NEURAL NETWORK PROBABILITY MAP RECONSTRUCTION AND LOCALIZATION

After creating tomograms (converting the RF-based challenge into an image processing problem), we proceed to locate users in the tomograms. Having the tomograms lets us apply the mature field of keypoint localization [40], [41] to the problem of localizing people in RF data. Our architecture is built on the standard U-Net [38] architecture, which is common in image-to-image translation problems [38], [42]. Additionally, it is of particular interest to our case because U-Net is agnostic to the size of the input, which enables deployment in varying size rooms without retraining. Following keypoint detection approaches [40], [41], we train this U-Net to regress a small blob around the person’s location with a mean squared error. At training time, the location can be generated by a ground-truthing system; at test time, localization entails running the tomograms through the network and finding peaks.

A U-Net consists of an encoder that reduces spatial resolution while increasing the feature dimensions, and the encoder is followed by a decoder that increases spatial resolution while decreasing feature dimension. Throughout the network, there are skip connections that propagate high-resolution information from the encoder for use by the decoder. Our architecture takes input tomograms (typically with 3 channels, RSSI, RF Phase, and Read Rate in our case), converts them into features with two convolution layers with 64 channels, and then passes them through the encoder and decoder. These, and all convolution layers are size 3×3 and are followed by a ReLU [43]. The encoder consists of a set of *Downsampling* blocks. Each contains two convolution layers, the first of which increases feature dimension by $2\times$, followed by 2×2 max-pooling. The decoder mirrors this with *Upsampling* blocks that $2\times$ bilinearly upsampling their input, concatenating the corresponding pre-max-pool high-resolution encoder feature map, and then performing two convolution layers. This is finished with a single 1×1 convolution that maps to a single output channel. In total, the network can be seen as a function f that maps a tomogram of a given height and width to a one-channel output of the same height and width, which is parameterized by a set of weights and biases θ .

Given a set of ground-truth tomograms and corresponding locations, we can train our network to produce a heatmap indicating the location of the person. At training time, we need a set of N tomograms $\{\mathbf{T}_i\}_{i=1}^N$ (with $\mathbf{T}_i \in \mathbb{R}^{H \times W \times 3}$) and locations $\{(x_i, y_i)\}$ of a person. We then construct a heatmap $\mathbf{G}_i \in \mathbb{R}^{H \times W \times 1}$ by placing a Gaussian function centered at (x_i, y_i) with standard deviation 30cm. Supposing

that our network is f with parameters θ , the goal is to solve the problem

$$\arg \min_{\theta} \sum_{i=1}^N \sum_{j=1, k=1}^{H, W} (f(\mathbf{T}_i)_{j,k} - (\mathbf{G}_i)_{j,k})^2, \quad (3)$$

or to find the parameters that minimize the total per-pixel error in predicting the heatmaps. We train the deep network using the AdamW [44] optimizer (with learning rate = 10^{-3} and $\epsilon = 10^{-8}$).

Inference consists of running a new tomogram through the model, labeling the blobs with the *ndimage.label* function in *scipy* python package, and finding the blobs with peak value above the threshold. This, in conjunction with the skip-connections in the U-Net, enables TomoID to output a variable number of users rather than a fixed number of outputs (which we demonstrate empirically). The particular location is extracted from the blob by weighted average and a Kalman filter is applied to smooth outputs over time. The end result appears in Figure 4, panel ‘D’.

We currently obtain this using an ultra-wideband tracking system that acquires uninterrupted, continuous localization data as ground-truth to train and evaluate TomoID. With 9 Ciholas DWUSB ultra-wideband anchors on the ceiling, we can track an ultra-wideband node on a custom-made hat worn by a participant at a rate of 100Hz with ± 1 cm accuracy in our configuration. This high-precision, high-refresh-rate tracking setup provides TomoID with continuous and precise measurements that let TomoID see observations as the user moves around rather than standing still.

VII. EVALUATION

Different applications have widely varying localization requirements. For example, smart buildings that optimize energy consumption based on human occupancy require coarse accuracy and may only need to check if a person is in a room to decide when to control lighting or the heating and cooling system. On the other hand, in-home activity detection for elder care requires a higher update rate and higher accuracy to support applications such as activity inferencing and monitoring. A practical resolution for these applications should at least be lower than the average size of the human shoulder breadth, 47 cm for male and 40 cm for female in the U.S. [45], [46].

Scalably deploying learned device-free systems presents other challenges. In a single room, concerns include how much training data is required and whether the systems generalize to a variety of bodies. Extending device-free localization to multiple rooms raises questions of how much training is required to generalize different spaces. Finally, there are practical concerns about how cost-effective and easy-to-install the system is, which affects the deployability as well.

Our experiments aim to test both accuracy and deployment complexity. We analyze the relationship between system parameters, localization accuracy, and computation complexity in terms of parameter size of the NN model. We then analyze

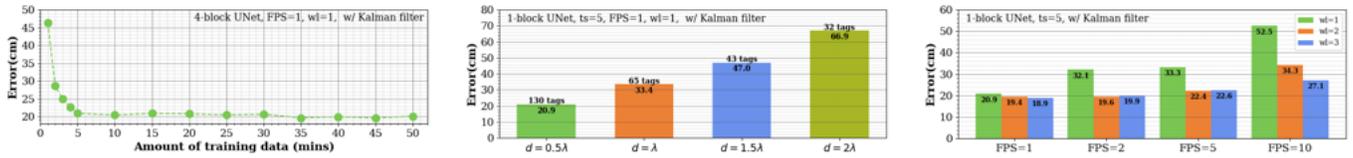


Fig. 5: (a)-(c) from left to right. (a) Performance of TomoID with different amounts of training data. Results show that performance saturates with just 5 minutes of training data. (b) Performance of TomoID with different RFID tag spacing of $0.5\lambda = 16.4$ cm, $\lambda = 32.8$ cm, $1.5\lambda = 49.2$ cm and $2\lambda = 65.6$ cm. (c) Performance of TomoID with different inference period FPS and number of past information wl .

the real-time performance of TomoID with both single and multiple users. Finally, we test whether the model trained on the original conference room described above can be transferred to other spaces (such as a large digital fabrication lab and a long narrow hallway), thereby eliminating the need to gather ground truth and training data for new deployments.

A. Data collection

A total of 130 UHF RFID tags with 16.37 cm spacing are placed on four walls in an office room of 31.2 square meters (336 sq. ft.) shown in Figure 1 panel ‘A’. We collected 90 minutes of data for a single person spread over three separate days (50 minutes, 25 minutes, and 15 minutes respectively), each with an additional 5 minutes of calibration data recorded. The 50 minutes of data from the first day is used as training data, while the 40 minutes from the other two days are used as test data. A 174 cm tall male participant moves and stops randomly out of his own will during the session. As for the ground truth, the Ultra-Wideband (UWB) localization system previously described was installed in the same office room. The participant wears a custom hat with a battery-powered UWB tag throughout the whole session. A total of nine UWB anchors are placed on the ceiling and connected to a host computer, locating the UWB tag on the user at a rate of 100 Hz. A second dataset lasting 30 minutes with two users (a 174 cm tall male and a 170 cm tall female) was collected using the same method, with both users wearing the UWB hats, moving throughout the space within the session. Three more datasets with 3, 4, and 5 users each under the same experimental setup across a 10 minutes session are collected. The participants are males in the range of 165 to 180 cm tall.

B. TomoID system parameter analysis

This section explores the relationship between accuracy and system parameters such as the amount of training data (ts), the number of U-Net blocks, spacing between RFID tags (d), update frames per second (FPS), and window length (wl) in TomoID through a series of experiments.

Collecting training data (for now) depends on using an alternate sensor like a UWB sensor, and practical deployment depends on this training step being not overly burdensome. Figure 5a shows the resulting accuracy while varying the amount of training data used to train the model. The average error saturates at ~ 20 cm with only 5 minutes of training data, showing the ability of the model to converge with a small amount of data provided.

U-Net has an encoder-decoder structure that shrinks and expands the data in a set of steps, or blocks. Increasing the

number of blocks increases the network’s learning capacity (and parameter count), as well as run-time. Table I shows the number of parameters and the localization accuracy for U-Nets with different numbers of blocks. The results suggest that all networks have similar accuracy. Therefore, we choose the 1-block U-Net as our final setting since it has 35x fewer parameters than the 4-block U-Net.

TABLE I: TomoID’s performance and neural network parameter count with different number of U-Net blocks under parameters setting of $ts = 5$, $FPS = 1$, $wl = 1$, with Kalman filter.

UNet Blocks	4	3	2	1
Parameter Size	17.27M	9.34M	2.25M	0.48M
Average Error (cm)	21.3	21.8	22.0	20.8

Increasing the spacing between tags means lower tag population, so the reader can complete an inventory scan cycle faster, and the system can be deployed at a lower cost. Figure 5b shows the accuracies under various tag spacing by removing the recorded tags in post-processing. As would be expected, with larger tag-to-tag spacing, accuracy decreases monotonically. Interestingly, the measured error is very close to the distance between tag spacing, suggesting that moving to smaller tag-to-tag spacing would increase system performance. However, near-field cross-coupling is expected between tags at distances under 0.5λ , which will degrade system performance.

Refresh rate is a critical element that decides the system’s responsiveness in real-time applications. When varying FPS , two crucial parameters are involved: computation time and tag reads per frame. First, the computation bottleneck of TomoID is solving Equation 2. Pre-computing the matrix is impossible since each cycle has different responding RF links, so \mathbf{W} is different each cycle. Second, higher reads per frame mean more RF links are measured, forming a more intact RF grid. The average read rate for our setup is 642 reads/sec for the whole office space. Assuming under the best case scenario, all tags have the same RF condition, and there is no collision, we can loop through all 310 RF links forming the whole RF grid in the room every 0.5 seconds. Therefore, with $FPS = 5$ and $FPS = 10$, the average reads in the time window are 130 and 64 reads, meaning at most 40% and 20% of the RF links are scanned even under the ideal condition, forming a very loose RF grid. However, we observed 62.9% of RF links read per second on average in the office space environment under real-world conditions. Given the tomograms formed by partial information of the RF grid, our custom 1-block U-Net can still extract the user’s position successfully. The corresponding accuracies are shown in Figure 5c.

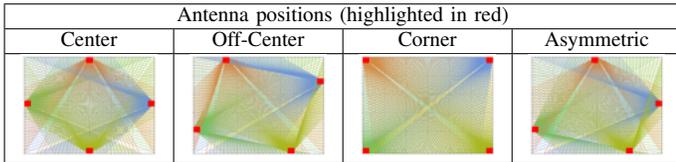


Fig. 6: TomoID with four antenna placement configurations. The plot shows the RF link distribution under specific configuration with the antennas highlighted in red. The average errors are calculated using the model trained with the antennas at the center. The number of RF links within the setup from left to right are 310, 310, 260, and 310. The average errors from left to right are 18.9 cm, 19.6 cm, 24.3 cm, and 19.5 cm.

While higher FPS results in lower accuracy, it is possible to improve the performance by adding RSSI, RF Phase, and Read Rate tomograms from the past time slices as input to the neural network. Figure 5c shows that by having a total of 9 tomograms across 3 consecutive time slices ($wl = 3$) as input at $FPS = 10$, TomoID can lower the average error by half to 27.1 cm compared to $FPS = 10$ with only the current time slice as input ($wl = 1$).

C. Performance

After the analysis above, we finalized the parameters to a 1-block U-Net, $ts = 5$, $d = 0.5\lambda$, $FPS = 1$, and $wl = 3$ to perform real-time localization. The results in this section were measured within the same office space where the 5 minutes of training data with only one person was collected.

Single-user performance: TomoID achieves **17.1 cm** average error with a single user standing still at 25 uniformly distributed points in the office space for 1 minute each. Although this method is widely used, it is not practical since the user’s motion is eliminated, forming more stable RF measurements. A testing situation that closely approximates real life with the user moving randomly around the space was measured with the device-based UWB localization system’s help as ground truth. The average error is **18.9 cm**, indicating the ability of TomoID to capture the user’s position with non-static RF measurements accurately.

Multi-user performance: Device-free localization systems have difficulty locating multiple users because of the shadowing effect created when more than one user is aligned on the same RF link from the antenna to the tag. Prior works [39], [47] tried to tackle this problem but were tested in well-controlled environments which do not represent real-world scenarios. Even being trained with only 1 user in the training data, our custom U-Net is still able to output multiple blobs because the unique NN structure extracts features from small local areas other than viewing the input as a whole. TomoID achieves an average error of **39.6 cm** when locating two users, with 90% of the error within 70 cm. As the number of users increases, the system might only locate some users in a frame. With the help of the Kalman filter, TomoID can interpolate the locations with past information to achieve 72.2cm, 72.1 cm, and 70.1 cm average error with 3, 4, and 5 users respectively.

Hardware placement flexibility: Sometimes it is not ideal to place the antenna in the center of the wall because of blocking furniture. TomoID can still operate when the

positions of the antennas are changed since it learns the heatmaps instead of the physical layout of the space. We moved the antennas to different locations without changing the position of the tags. Figure 6 shows the RF links under different antenna placements with the model trained with the antennas at the center. The average error stays at near 20 cm, proving the robustness of the neural network and the deployment flexibility of TomoID.

VIII. TOMOID IN NEW ENVIRONMENTS AND ITS APPLICATIONS

Neural networks excel at learning and extracting useful information from the RF parameters. However, the complex non-linear multipath relationship of RF parameters is highly dependent on the physical space. Therefore, one of the hardest challenges for ML-based, device-free indoor localization systems is maintaining high accuracy when applying a pre-trained model to a new area. TomoID overcomes this issue by featurizing RF measurements into tomograms to eliminate the neural network from learning room-specific information such as multipath from raw RF measurements. Moreover, the unique neural network structure can process input tomograms with flexible sizes, allowing the opportunity to apply a well-trained model to new spaces without retraining. This is important because the task of collecting high-fidelity ground truth data is one of the greatest barriers to using localization systems.

When deploying TomoID to new spaces, we install reader antennas and use a cardboard template to space out the tags onto the walls evenly. TomoID only needs the weight matrix \mathbf{W} described in section V-B, 5 minutes of background calibration data, and the pre-trained model to start inferring users’ locations.

We deploy TomoID in a digital fabrication laboratory and a hallway. Given that the two new spaces have fundamentally different sizes, geometries, reader locations and orientations, along with different RF multipath environments, we wanted to investigate the quality of the pre-trained model compared to a custom train model for that space. This way any change in performance could be compared to the ideal case of a custom-trained model. The result shows that the pre-trained model works effectively for the target applications described in the paper. Finally, we explore and evaluate several possible applications of TomoID in the digital fabrication laboratory and a hallway.

A. TomoID in Digital Fabrication laboratory

Figure 1 panel ‘B’ shows the digital fabrication laboratory, which is 60% larger than the office space used in the previous sections. There are workbenches, a laser cutter, a metal-backed whiteboard, a 3D printer, and tall metal cabinets that block and reflect the RF signals in the room. The digital fabrication laboratory’s shape is also different from the office space used to create the pre-trained machine learning model. The blacked-out area in the top-down floor plan shown in Figure 7a indicates a walled-off section of the space. A total of 139 RFID tags, creating 318 RF links, are in the space with four

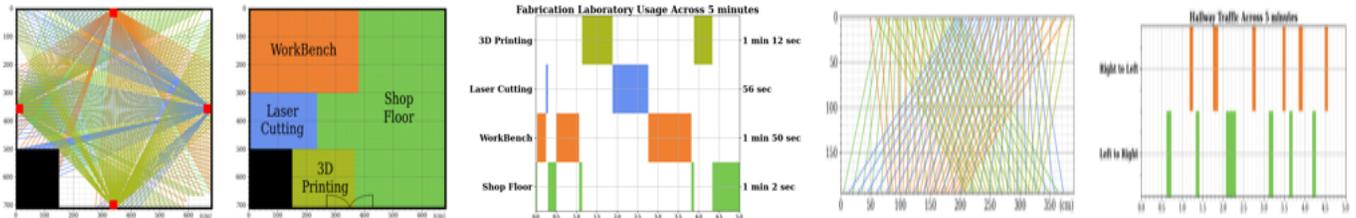


Fig. 7: (a)-(e) from left to right. (a) RF Link distribution of the fabrication laboratory with a cut out in black and RFID antennas in red. (b) Four sections of the fabrication laboratory. (c) The fabrication laboratory usage of a single user across 5 minutes. The user first worked at the workbench, issued a 3D printing job, and then did laser cutting, went back to the workbench, finally retrieved the printed 3D model and headed out of the space. (d) The RF link distribution plot of hallway scenario. Since RFID tags are only placed on two sides of the hallway, there are no horizontal RF links compared to the office space or fabrication laboratory. (e) Hallway traffic showing a person crossing the hallway back and forth across 5 minutes.

RFID reader antennas mounted on the walls denoted as red squares. To quantify the performance in this new space, we deployed the UWB localization system to obtain ground truth.

Within the 25 minutes test session with 5 minutes of data to train, the average error is 28.3 cm for the custom-trained model and 38.3 cm for the pre-trained model. The pre-trained model is only 10 cm less accurate than the custom-trained model, which is an impressive result given that the room is larger, has a different shape, and has significantly different RF characteristics than the room the model was trained on. Furthermore, lower accuracy does not necessarily mean poor system performance. A 38.3 cm average error is comparable to the average human shoulder breadth of 47 cm and 40 cm for males and females in the U.S. [45], [46], which is considered sufficient for many location-based applications.

As proof of viability, we show that the pre-trained model has sufficient accuracy in tracking space usage. In this test, the space was divided into 4 stations: 3D printing, a laser cutting area, a hand tool workbench, and an open shop floor, as shown in Figure 7b. A time-series usage plot, shown in figure 7c, visualizes the user’s activity within the time frame. Results show that TomoID could successfully track an individual’s space usage across time, which can be used to enhance shop maintenance and scheduled cleaning routines.

B. TomoID in Hallway - Seeing Through Walls

As a second example of using a pre-trained model on a new space, TomoID is deployed to monitor the number and direction of people walking down the hallway shown in Figure 1 panel ‘C’. A total of 46 RFID tags and 4 antennas were placed inside the rooms on the two sides of the hallway, resulting in an instrumented area of 4m x 2.25m filled with 92 RF links. The RFID reader antennas are arranged in a new ordination with two antennas on each side of the hallway—the open-ended hallway results in missing horizontal RF links as shown in Figure 7d. Therefore, it is harder to detect vertical movements and localize users at the ends of the hallway. Nonetheless, the pre-trained model from the office space performs well, resulting in a localization error of 39.2 cm. The custom-trained model trained with 5 minutes of data recorded in the hallway has a localization error of 28.6 cm, which is only a 10 cm improvement compared to the pre-trained model. While the performance of the pre-trained model is indeed

lower than the custom-trained model, it has sufficient accuracy for many indoor localization applications. Thus the pre-trained model has the significant advantage that a user does not have to collect ground truth training data, significantly improving the deployability of TomoID compared to previous work.

An example of monitoring foot traffic flow through the hallway is shown in Figure 7e while two participants were asked to walk up and down the hallway for 5 minutes. Results show that TomoID can indeed track the direction of users without the need for collecting a site-specific training set. Importantly, the fabrication lab and hallway experiments show the power of combining tomography and flexible machine learning models for device free RF base localization, as it allows the RF physical layer to be separated from the machine learning layer enabling pre-trained models to be used, thus lowering the burden for deploying the localization system.

IX. CONCLUSION

This paper proposes TomoID, a device-free real-time localization system that uses battery-free RFID technology to image the interior of living spaces to predict a user’s location. This radio tomography approach employs a novel signal processing pipeline that uses communication channel parameters such as RSSI, RF Phase, and Read Rate to create tomograms which are processed by our custom design convolutional neural network. Not only does this system provide high localization accuracy for a single moving user with an average mean error of 17.1 cm standing and 18.9 cm moving. It also enables multi-user tracking with an average mean error of 39.6 cm for two users (moving) and 72 cm for 3 to 5 moving users. Furthermore, this work demonstrates that a pre-trained model can be deployed to a new area (without the need for re-training) while still being able to achieve good localization of accuracy sufficient for many indoor localization applications.

Finally, our results reveal new system capabilities in terms of localization accuracy, real-time performance, frame rate, and deployability that have not been demonstrated in prior work. These advancements fundamentally lower the barrier to deploying device-free tomographic localization systems in the real-world. Ultimately, enabling users to benefit from location-aware services and smart environments without the need to continuously wear a transponder or carry a mobile device.

REFERENCES

- [1] N. Fallah, I. Apostolopoulos, K. Bekris, and E. Folmer, "Indoor human navigation systems: A survey," *Interacting with Computers*, vol. 25, no. 1, pp. 21–33, 2013.
- [2] U. Rehman and S. Cao, "Augmented-reality-based indoor navigation: A comparative analysis of handheld devices versus google glass," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 1, pp. 140–151, 2017.
- [3] T. H. Riehle, P. Lichter, and N. A. Giudice, "An indoor navigation system to support the visually impaired," in *2008 30th EMBS*, 2008, pp. 4435–4438.
- [4] X. Lin, T. Ho, C. Fang, Z. Yen, B. Yang, and F. Lai, "A mobile indoor positioning system based on ibeacon technology," in *2015 37th EMBC*, 2015, pp. 4970–4973.
- [5] D. Fortin-Simard, K. Bouchard, S. Gaboury, B. Bouchard, and A. Bouzouane, "Accurate passive rfid localization system for smart homes," in *2012 IEEE 3rd NESEA*, 2012.
- [6] D. Zhang, F. Xia, Z. Yang, L. Yao, and W. Zhao, "Localization technologies for indoor human tracking," in *2010 5th International Conference on Future Information Technology*, 2010, pp. 1–6.
- [7] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.
- [8] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta, "Luxapose: Indoor positioning with mobile phones and visible light," in *MobiCom '14*. New York, NY, USA: ACM, 2014, p. 447–458.
- [9] S. Hijikata, K. Terabayashi, and K. Umeda, "A simple indoor self-localization system using infrared leds," in *2009 Sixth International Conference on Networked Sensing Systems (INSS)*, 2009, pp. 1–7.
- [10] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, ser. *MobiCom '00*. New York, NY, USA: ACM, 2000, p. 32–43.
- [11] A. Ramezani Akhmareh, M. Lazarescu, O. Bin Tariq, and L. Lavagno, "A tagless indoor localization system based on capacitive sensing technology," *Sensors*, vol. 16, p. 1448, Sep 2016.
- [12] T. H. Riehle, S. M. Anderson, P. A. Lichter, N. A. Giudice, S. I. Sheikh, R. J. Knuesel, D. T. Kollmann, and D. S. Hedin, "Indoor magnetic navigation for the blind," in *EMBS*, 2012, pp. 1972–1975.
- [13] F. Liu, J. Liu, Y. Yin, W. Wang, D. Hu, P. Chen, and Q. Niu, "Survey on wifi-based indoor positioning techniques," *IET Communications*, vol. 14, no. 9, pp. 1372–1383, 2020.
- [14] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang, "Radio-frequency tomography for passive indoor multitarget tracking," *IEEE Transactions on Mobile Computing*, vol. 12, no. 12, pp. 2322–2333, 2013.
- [15] K. Ohara, T. Maekawa, Y. Kishino, Y. Shirai, and F. Naya, "Transferring positioning model for device-free passive indoor localization," in *UbiComp '15*. New York, NY, USA: ACM, 2015, p. 885–896.
- [16] J. Wang, J. Xiong, H. Jiang, K. Jamieson, X. Chen, D. Fang, and C. Wang, "Low human-effort, device-free localization with fine-grained subcarrier information," *IEEE Transactions on Mobile Computing*, vol. 17, no. 11, pp. 2550–2563, 2018.
- [17] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, "Indotrack: Device-free indoor human tracking with commodity wi-fi," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, sep 2017.
- [18] C. M. Watts, P. Lancaster, A. Pedross-Engel, J. R. Smith, and M. S. Reynolds, "2d and 3d millimeter-wave synthetic aperture radar imaging on a pr2 platform," in *2016 IEEE/RSJ IROS*, 2016, pp. 4304–4310.
- [19] A. Patra, L. Simic, and M. Petrova, "mmrti: Radio tomographic imaging using highly-directional millimeter-wave devices for accurate and robust indoor localization," in *2017 IEEE 28th PIMRC*. IEEE, 2017, pp. 1–7.
- [20] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via rf body reflections," in *NSDI '15*. USA: USENIX Association, 2015, p. 279–292.
- [21] 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 *CVPR*, 2018, pp. 7356–7365.
- [22] M. Bocca, O. Kaltiokallio, and N. Patwari, "Radio tomographic imaging for ambient assisted living," in *Communications in Computer and Information Science*, vol. 362, 01 2013, pp. 108–130.
- [23] S. Wang, S. Kueppers, H. Cetinkaya, and R. Herschel, "3d localization and vital sign detection of human subjects with a 120 ghz mimo radar," in *2019 20th International Radar Symposium (IRS)*, 2019, pp. 1–6.
- [24] O. Kaltiokallio, M. Bocca, and N. Patwari, "A fade level-based spatial model for radio tomographic imaging," *IEEE Transactions on Mobile Computing*, vol. 13, no. 6, pp. 1159–1172, 2013.
- [25] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang, "Radio-frequency tomography for passive indoor multitarget tracking," *IEEE Transactions on Mobile Computing*, vol. 12, no. 12, pp. 2322–2333, 2012.
- [26] R. J. LLC. (2015, sep) How much does an rfid tag cost today? [Online]. Available: <http://www.rfidjournal.com/faq/show?85>
- [27] X. Liang, Z. Huang, S. Yang, and L. Qiu, "Device-free motion & trajectory detection via rfid," *ACM Trans. Embed. Comput. Syst.*, vol. 17, no. 4, Aug. 2018.
- [28] W. Ruan, Q. Z. Sheng, L. Yao, T. Gu, M. Ruta, and L. Shanguan, "Device-free indoor localization and tracking through human-object interactions," in *2016 IEEE 17th WoWMoM*, 2016, pp. 1–9.
- [29] W. Ruan, L. Yao, Q. Z. Sheng, N. J. G. Falkner, and X. Li, "Tagtrack: Device-free localization and tracking using passive rfid tags," in *MobiQ-uitous*. ICST, 2014, p. 80–89.
- [30] D. Zhang, J. Zhou, M. Guo, J. Cao, and T. Li, "Tasa: Tag-free activity sensing using rfid tag arrays," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 4, pp. 558–570, 2010.
- [31] Z. Wang, F. Xiao, N. Ye, R. Wang, and P. Yang, "A see-through-wall system for device-free human motion sensing based on battery-free rfid," *ACM Trans. Embed. Comput. Syst.*, vol. 17, no. 1, Sep. 2017.
- [32] M.-J. Hsieh, J.-L. Guo, C.-Y. Lu, H. wei Hsieh, R.-H. Liang, and B.-Y. Chen, "Rftouchpads: batteryless and wireless modular touch sensor pads based on rfid," in *UIST 2019*. United States: ACM, Inc, Oct. 2019, pp. 999–1011.
- [33] B. Wagner, N. Patwari, and D. Timmermann, "Passive rfid tomographic imaging for device-free user localization," in *2012 9th Workshop on Positioning, Navigation and Communication*, 2012, pp. 120–125.
- [34] J. Han, C. Qian, X. Wang, D. Ma, J. Zhao, W. Xi, Z. Jiang, and Z. Wang, "Twins: Device-free object tracking using passive tags," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1605–1617, 2015.
- [35] W. Ruan, Q. Z. Sheng, L. Yao, X. Li, N. J. Falkner, and L. Yang, "Device-free human localization and tracking with uhf passive rfid tags: A data-driven approach," *Journal of Network and Computer Applications*, vol. 104, pp. 78–96, 2018.
- [36] Y. Ma, B. Wang, X. Gao, and W. Ning, "The gray analysis and machine learning for device-free multitarget localization in passive uhf rfid environments," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 802–813, 2019.
- [37] L. Ma, M. Liu, H. Wang, Y. Yang, N. Wang, and Y. Zhang, "Wallsense: Device-free indoor localization using wall-mounted uhf rfid tags," *Sensors*, vol. 19, no. 1, p. 68, Dec 2018.
- [38] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, 2015.
- [39] J. Wilson, N. Patwari, and F. G. Vasquez, "Regularization methods for radio tomographic imaging," in *2009 Virginia Tech Symposium on Wireless Personal Communications*. Citeseer, 2009.
- [40] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *ECCV*, 2016.
- [41] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *International Conference on Computer Vision*, 2017.
- [42] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *CVPR*, 2017.
- [43] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *ICML*, 2010.
- [44] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.
- [45] H. W. Stoudt, "Skinfolds, body girths, biacromial diameter, and selected anthropometric indices of adults, united states, 1960-1962," *Vital and health statistics. Series 11, Data from the national health survey ; no. 35;Public Health Service publication ; no. 1000, ser. 11, no. 35.*, 1970. [Online]. Available: <https://stacks.cdc.gov/view/cdc/12624>
- [46] Anthropometry - Measurements of People. (2022). [Online]. Available: "<http://personal.cityu.edu.hk/meachan/Online/%20Anthropometry/Sitemap.htm>"
- [47] B. Wagner and D. Timmermann, "Approaches for device-free multi-user localization with passive rfid," in *International Conference on Indoor Positioning and Indoor Navigation*, 2013, pp. 1–6.