

# DT-RaDaR: Digital Twin Assisted Robot Navigation using Differential Ray-Tracing

Sunday Amatare\*, Gaurav Singh\*, Raul Shakya\*, Aavash Kharel\*, Ahmed Alkhateeb† and Debashri Roy\*

The University of Texas at Arlington\* and Arizona State University†

Emails: {saa3326, gaurav.singh4, rxs6339, axk8168}@mavs.uta.edu\*, aalkhateeb@asu.edu†, debashri.roy@uta.edu\*

**Abstract**—Autonomous system navigation is a well-researched and evolving field. Recent advancements in improving robot navigation have sparked increased interest among researchers and practitioners, especially in the use of sensing data. However, this heightened focus has also raised significant privacy concerns, particularly for robots that rely on cameras and LiDAR for navigation. Our innovative concept of Radio Frequency (RF) map generation through ray-tracing (RT) within digital twin environments effectively addresses these concerns. In this paper, we propose DT-RaDaR, a robust privacy-preserving, deep reinforcement learning-based framework for robot navigation that leverages RF ray-tracing in both static and dynamic indoor scenarios as well as in smart cities. We introduce a streamlined framework for generating RF digital twins using open-source tools like Blender and NVIDIA’s Sionna RT. This approach allows for high-fidelity replication of real-world environments and RF propagation models, optimized for service robot navigation. Several experimental validations and results demonstrate the feasibility of the proposed framework in indoor environments and smart cities, positioning our work as a significant advancement toward the practical implementation of robot navigation using ray-tracing-generated data.

**Index Terms**—robot navigation, privacy-preservation, RF digital twins, ray-tracing, reinforcement learning.

## I. INTRODUCTION

The domain of autonomous robots is expanding swiftly on a global scale, with a broadening range of new uses and increasing interests from established sectors such as automotive, freight, public transit, industrial operations, and defense. One area where autonomous robots have seen increased interest in recent years is e-commerce. This is largely attributed to traffic congestion, especially in urban areas, and the difficulties companies encounter with tight schedules for picking up and delivering items [1]. Additionally, the recent increase in package deliveries, as more people shop online instead of visiting physical stores, has further intensified the demands on human workers. Particularly, autonomous delivery robots (ADR) have become a practical solution for these tasks. These robots can assist human workers in areas where their capabilities are constrained, particularly by helping to reduce their workload. According to Facts and Factors, delivery robot market is projected to reach \$55 billion by 2026, with an annual growth rate of 20.4% [1]. Additionally, certain companies assert that these robots can lower delivery expenses by as much as 90%. Companies like Panasonic, Starship, and Robomart are leading the development of ADRs [1]. In addition, major delivery services such as Amazon, FedEx and Uber are utilizing ADRs to improve their delivery operations [2].

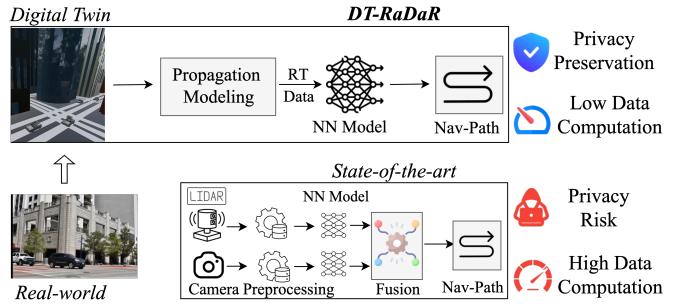


Fig. 1. Comparison of sensor-based navigation and DT-RaDaR. We show that the DT-RaDaR is privacy preserving and data efficient.

**Sensor-enabled Robot Navigation.** Sensing data plays a crucial role in autonomous systems, especially for service robots and ADRs engaged in tasks such as mapping, localization, and navigation within unstructured environments. Among the laser-based solutions for mobile robot navigation and localization, 2D light detection and ranging (LiDAR) sensors stand out as a popular option due to their accessibility and trusted performance. Although, a significant limitation of the 2D LiDAR sensor is its dependence on a single horizontal scan line for sensing, which restricts its capacity to gather detailed spatial data and hinders effective navigation in complex terrains. On the other hand, 3D LiDAR enables a more detailed understanding of the environment by collecting spatial data in multiple dimensions. Yet, its increased expense can be a barrier, discouraging its use in certain applications. In recent years, depth cameras have gained popularity primarily due to their affordability, accessibility, and ability to capture detailed depth information at a lower cost [3]. These features make depth cameras a competitive option for various applications, particularly in situations where more expensive solutions, like 3D LiDAR, are impractical. However, depth cameras have limitations in range and accuracy.

**Concerns with Sensor-enabled Robot Navigation.** The integration of sensory information into service robots and ADRs presents significant privacy challenges that require careful consideration. Key concerns include the risk of unauthorized monitoring, where individuals may be surveilled without their consent, leading to potential violations of personal privacy. Moreover, the possibility of data breaches raises alarms, as sensitive information collected by these robots could be compromised, exposing users to various security threats. These concerns underscore the necessity of implementing a privacy-preserving system to protect sensitive information collected by

service robots. This system should prioritize ethical standards, such as obtaining explicit user consent and upholding transparent data usage policies throughout both the development and deployment stages of these technologies. By doing so, it fosters trust among users and stakeholders, as highlighted in [4]. Furthermore, it is essential to recognize that these sensors require significant computational resources due to their need to provide high precision to capture the environmental characteristics critical to effective robot navigation [4].

**Privacy Preservation through Ray-tracing (RT).** The concept of rays is easy to grasp based on our everyday encounters with sunlight. When sunlight passes through a relatively large opening in a wall and enters a room, we can observe the ‘ray’ traveling in a straight path [5]. RF ray-tracing is an established technology that can effectively model the propagation of electromagnetic waves across various relevant scenarios. It has become an effective approach for selectively capturing and analyzing visual and spatial data from the environment while ensuring privacy protection. Ray-tracing models electromagnetic waves as distinct rays that travel in straight lines. When these rays interact with objects in their environment, they can generate various phenomena, such as reflection, refraction, diffraction, and scattering. Differential ray-tracing (DRT) is a specialized extension of conventional ray-tracing that detects slight alterations in the environment, allowing for precise sensitivity analysis [6], [7]. It also identifies essential environmental characteristics required to produce accurate directives for actuation in complex settings, which can be utilized to inform robot navigation decisions. Furthermore, the data generated through DRT includes numerical values of different ray properties, making it a potentially efficient means of representing the environment.

**Motivation and Contributions of this Paper.** Most state-of-the-art service robots use cameras or LiDAR systems to perceive and interpret their surroundings, enabling efficient navigation and task performance across diverse applications, including autonomous delivery robots. While these technologies enhance the robots’ ability to operate in complex environments, they often collect large amounts of visual and spatial data, potentially capturing sensitive information such as individuals’ physical characteristics, locations, or behaviors. This poses significant privacy risks, raising concerns about data security, consent, and potential misuse [8], [9]. With this motivation in mind, we propose an innovative RF-based framework, DT-RaDaR, aimed at ensuring privacy in robot navigation while advancing technology. The RF-based framework requires robots to be equipped with RF devices, which are often integrated into modern service robots but typically used solely for communication purposes. These RF devices are used for ray-tracing to generate maps, aiding the robots in navigating through unstructured environments. As illustrated in Fig. 1, in DT-RaDaR, the novel application of RT for robot navigation requires: (a) Creation of static blueprint digital twins and dynamic digital twins that accurately simulate real-world scenarios, (b) a DRT-based propagation modeling and data generation process within the digital twin, and (c) the

implementation of a reinforcement learning (RL) approach that leverages the propagation characteristics of the digital twin to generate optimal navigation paths for the robot. The proposed work requires validation through a specific type of simulation, which includes creating a digital twin from real-world, integrating the digital twin with ray-tracing software, performing propagation characteristics within the imported digital twin, and generating data to train the proposed RL-based robot navigation model. Formally, this paper’s contributions are:

**C1.** We propose a high-fidelity digital twin creation method that accurately models both indoor and outdoor environments by extracting real-world features through the open-source Blender tool and its OpenStreetMap (OSM) add-on.

**C2.** We propose a methodology for precisely configuring scenes and generating the propagation characteristics of various environments. This approach employs DRT on the digital twin of each scene using NVIDIA’s Sionna RT-based software to generate propagation data across all grids within the scene.

**C3.** We propose a RL-based algorithm by designing: (a) a customized environment using the propagation data generated from the ray-tracing tool and (b) a customized state extraction process from the data generated using the ray-tracing, four actuation actions for the robot, and a reward function that provides feedback to the (DQN) agent received from the environment. Following the realistic scenario, we train the DQN agent on a blueprint digital twin and perform inference on dynamic digital twins.

**C4.** We present a first-of-its-kind (to the best of our knowledge) ray-tracing-based digital twin dataset that replicates both indoor laboratory environments and outdoor smart city scenarios, specifically modeled on downtown Dallas and downtown Houston. These highly realistic digital twins are curated using publicly available software tools, ensuring accuracy and relevance to real-world settings.

**C5.** We validate our overall framework, DT-RaDaR, through rigorous experimental validations on the collected datasets. We also conduct experiments where the DT-RaDaR agent is trained on a twin and its inference performance is analyzed on another. We observe that RT-based training requires  $\sim < 74.4\%$  of training time of similar methods which uses LiDAR/Camera sensors. Furthermore, upon acceptance of this article, we pledge to release our codebase, digital twin models, ray tracing datasets in [10].

## II. RELATED WORKS AND MOTIVATION

**Indoor Robot Navigation:** In this type of navigation, robots employ various techniques and systems to understand their environment, plan paths, and avoid obstacles while moving through spaces such as homes, hospitals, offices, and warehouses. Almeida *et al.* [11] propose a precise magnetic mapping method that is reliable and accurate, effectively addressing many indoor localization challenges using magnetic data. Noh *et al.* [12] present an autonomous mobile robot system that uses open-source ROS packages to address the challenge of avoiding collisions with both static and dynamic objects. Bavle *et al.* [13] introduce using odometry readings

and planar surfaces extracted from 3D LiDAR scans to construct a Situational Graph for robot pose estimation. Amatare *et al.* [14] introduce an RF-based robot navigation system leveraging Q-Learning in a static environment. Kulhanek *et al.* [15] propose a deep reinforcement learning method for visual navigation in real-world settings, based on the parallel advantage actor-critic algorithm, enhanced with auxiliary tasks and curriculum learning.

**Outdoor Robot Navigation:** For outdoor navigation the robots utilize advanced technologies and methods to navigate and operate effectively due to the challenges such as varying terrain, dynamic obstacles, and changing weather conditions. Palazzo *et al.* [16] introduce a deep learning-based approach to assess and predict the traversability of various routes within the field of view of an onboard RGB camera. Karnan *et al.* [17] propose a method combining imitation learning and inverse reinforcement learning for social navigation in mobile robots. They utilize a socially compliant navigation dataset to address the scarcity of large-scale datasets that capture socially appropriate robot behavior in real-world environments. Elnoor *et al.* [18] present a technique that leverages proprioceptive data to assess terrain traversability in real time for legged robots. Their approach also incorporates sensors to accurately evaluate terrain stability, the risk of entrapment, and potential crashes. Sorokin *et al.* [19] introduce a quadrupedal robot designed to navigate sidewalks by following a route generated by a public map service. The robot uses onboard sensors to steer clear of obstacles and pedestrians, ensuring safe and collision-free movement.

**Motivation for Designing DT-RaDaR:** All the highlighted approaches to robot navigation face challenges related to data privacy and the computational demands associated with the use of high-definition sensors, such as RGB-D cameras and LiDAR systems. Additionally, previous RF-based approach focused solely on indoor, static environment, whereas in practice, scene environments frequently change, both indoors and outdoors, across nearly all applications [20]. In this paper, we propose DT-RaDaR, a robust privacy-preserving, deep reinforcement learning-based framework for robot navigation that leverages RF ray-tracing in both static and dynamic indoor scenarios as well as in smart cities.

### III. SYSTEM MODEL AND PROBLEM FORMULATION

#### A. Problem Formulation

We consider robot  $R$  in the environment with  $O$  obstacles. In our setting, the robot is equipped with RF receiver and edge devices. We have a transmitter in the environment which is transmitting in the 2.4GHz band. The RF receivers listen to the 2.4GHz band and perform ray-tracing. Each edge device of the robot runs a decision algorithm from the ray-tracing data and generates the next coordinate for navigating through the environment (see Fig. 1). The path planning policy targets to minimize the number of steps reaching to the target, while avoiding the obstacles. We use deep reinforcement learning algorithm to find the optimum policy that meets the above

requirements. Thus, given a path planning policy  $\theta$ , we formulate our objective as:

$$\text{Maximize}_{\theta}: \mathbb{E}_{\theta} \left[ \sum_{t=0}^{T-1} f_R^{\theta}(s_t, a_t) \right], \quad (1a)$$

$$\text{s.t } a_t = \theta(s_t), \quad (1b)$$

$$\sum_{t=1}^T \varphi_t = 0. \quad (1c)$$

Here,  $s_t$  and  $a_t$ , denote the state and action for the robot  $R$  at step  $t$  ( $T$  total steps). The action space includes movement either in the X or in the Y directions. Moreover,  $f_R^{\theta}$  denotes a reward function, which is a function of state and action. Finally,  $\varphi_t$  is a Boolean predicate, with  $\varphi_t$  to be 1 if collision happens, and 0 otherwise, that ensures avoiding collisions.

#### B. System Architecture in DT-RaDaR

An overall view of our framework is shown in Fig. 2 and consists of three main modules as follows:

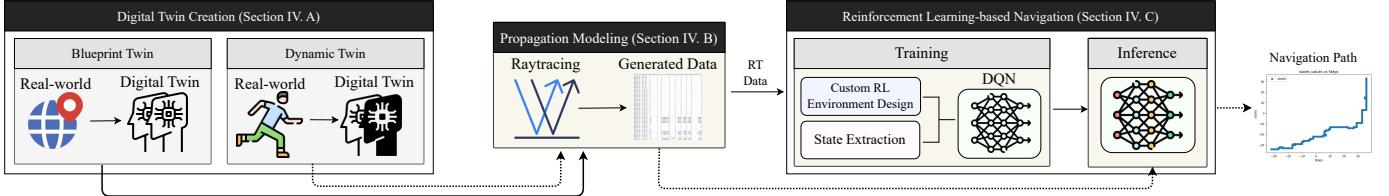
- **Digital Twin Creation (Module 1):** We create a digital twin by directly extracting the real world features to digital world via sensors. The training of RL algorithm is done over a *blueprint digital twin*, and the inference and deployment is conducted over the real-time twins, named as *dynamic digital twin*. While the transmitter is placed at the roof area, the receiver (robot) is placed on the floor of the room alongside obstacles (details in Sec. IV-A).
- **Propagation Modeling (Module 2):** The autonomous robot receives the instantaneous local ray-tracing data from the transmitting device in the digital twin and acts appropriately to detect the obstacles in close and distant ranges (details in Sec. IV-B).
- **RL-based Navigation (Module 3):** We design a RL-based algorithm that leverages the ray-tracing data and the propagation characteristics from the digital twin for the robot navigation (details in Sec. IV-C).

**Remark 1.** Our digital twin generation method is designed to generalize to realistic scenarios by utilizing training data collected from an initial blueprint digital twin. As the blueprint digital twin is continuously updated over time, re-training is conducted to ensure model accuracy and adaptability.

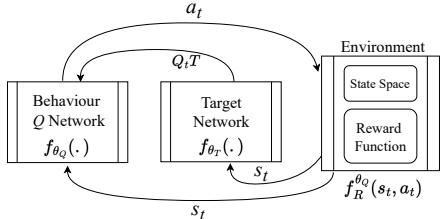
## IV. DT-RADAR FRAMEWORK

#### A. Module 1: Digital Twin Creation

In the DT-RaDaR, we consider properties with respect to map precision and radio propagation properties. While we integrate several specific key metrics in the DT-RaDaR, our baseline can be extended in the future to different environmental setup. We initialize the created digital twin as  $\mathcal{E} = f(\text{map}, O, \rho)$ . Here,  $\text{map}$  and  $O$  denote the imported Blender [21] map and present structures or obstacles for the twin  $\mathcal{E}$ , and  $\rho$  are number of allowed reflections for the created twin. We characterize objects present in the twin with  $O_u = \{(x_k, y_k, d_{0,k}, d_{1,k}, d_{2,k})\}_{k=1}^{N^{\text{obs}}}$ , where  $N^{\text{obs}}$  is the number of structures or obstacles that are present in the twin  $\mathcal{E}$ . Moreover,  $(x_k, y_k)$  are the obstacle coordinates on the map,



**Fig. 2.** Overview of the proposed system architecture. DT-RaDaR continuously monitors the environment. In *Module 1*, we generate the digital twins. The re-training is performed over an interval (*Module 2*), we update the *Module 3* which eventually updates path planning policy by optimizing the navigation (*Module 3*). The solid line shows the training path, and dashed line represent inference path.



**Fig. 3.** Overview of the proposed DQN agent. We formulate the state and reward space to optimize navigation.

whereas  $d_{0,k}$ ,  $d_{1,k}$ ,  $d_{2,k}$  represent the height, width, and length of the obstacle- $k$ .

As mentioned in Sec. III-B, the digital twins are categorized as (a) blueprint digital twin which provides the blueprint of the scenario with static objects for the RL agent to train on and (b) dynamic digital twin which captures the real-time changes within the scenario during inference.

### B. Module 2: Propagation Modeling

We use the open source Sionna RT [22] tool to generate the propagation characteristics of the created digital twin  $\mathcal{E}$  by employing RF ray-tracing. For a given transmitter  $TX$ , a propagation map is a rectangular surface with arbitrary orientation subdivided into rectangular cells of size  $|\mathcal{C}|$ . A parameter  $\eta$  controls the granularity of the map. The propagation map associates with every cell  $(\mathcal{C}_i, \mathcal{C}_j)$ . In DT-RaDaR, for every ray  $n$  intersecting the propagation map cell  $(\mathcal{C}_i, \mathcal{C}_j)$ , the channel coefficients  $\Theta$ , phase shifts  $\phi$ , the azimuth angles of departure  $\alpha_d$  and arrival  $\alpha_a$ , zenith angles of departure  $\zeta_d$  and arrival  $\zeta_a$  are computed.

### C. Module 3: RL-based Navigation (Training)

Once the propagation modeling is performed for the digital twin  $\mathcal{E}$ , we use reinforcement learning algorithm to solve Eq. 1 and obtain the optimum policy for navigating robot  $R$ . We use the Deep Q-Network (DQN) based training for performing the reinforcement learning to identify the policy for optimum path planning as shown in Fig. 3. The agent takes as input the state array  $s_t$  and outputs the action  $a_t$  following the policy  $\theta_Q$ .

1) *Custom RL Environment Design:* An RL environment represents the context in which an RL agent functions, serving as a simulated model with which the agent interacts by executing actions, receiving rewards or penalties, and transitioning to new states accordingly. This environment furnishes the agent with its initial state and updates it to the next state following each action taken by the agent, while also providing the reward back to the agent received at each new

state. However, given the novelty of our proposed RT-based robot navigation approach, custom RL environment must be generated to accurately capture the propagation characteristics of the digital twin  $\mathcal{E}$ . As depicted in Fig. 3, the custom RL environment encapsulates both state space and reward function which are further elaborated in the subsequent sections.

2) *State Extraction:* The state space is extracted from the Environment. We define the state space  $s_t$  of the robot  $R$  at step  $t$  as:  $s_t = \{R_x^t, R_y^t, \zeta_d^t, \alpha_d^t, \zeta_a^t, \alpha_a^t, |\Theta|^t, \phi^t\}$ , where  $R_x^t$  is the X coordinate of the robot location on the map,  $R_y^t$  is the Y coordinate of the robot location on the map,  $\zeta_d^t$  is zenith angle of departure,  $\alpha_d^t$  is azimuth angle of departure,  $\zeta_a^t$  is zenith angle of arrival,  $\alpha_a^t$  is azimuth angle of arrival,  $|\Theta|^t$  is channel coefficient magnitude, and  $\phi^t$  is phase shift at step  $t$ . The overall state array has 8 elements, encapsulating both RF propagation characteristics and navigation elements.

3) *DQN Agent Design:* Given the state array at the timestep  $t$ , the DQN agent is trained by interleaving optimization of two neural networks: (a) one neural network which estimates the actions of the agent using the current state based on the Q values and trains them is named as *behavior Q network*, denoted as  $f_{\theta_Q}(\cdot)$  with  $\theta_Q$  as *behavior policy*; and (b) another neural network which sets the target outputs for the behavior Q network to be trained on is named as *target Q network*, denoted as  $f_{\theta_T}(\cdot)$  with  $\theta_T$  as *target policy*.

**Behavior Q Network.** The behavior Q network policy ( $\theta_Q$ ) is generated by training the behavior Q network which estimates the Q values as:  $Q_t^Q = f_{\theta_Q}(s_t)$ .

**Action Space.** The action is taken by following the  $\epsilon$ -greedy policy,  $a_t = \epsilon - \text{greedy}(Q_t^Q)$ . The action  $a_t = \{X-, X+, Y-, Y+\}$ , representing traversing either in X direction ( $X-$  and  $X+$ ) or Y direction ( $Y-$  and  $Y+$ ) to reach the next state's X and Y coordinate  $(R_x^{t+1}, R_y^{t+1})$  by the robot  $R$  at timestep  $(t+1)$ .

**Reward Function.** The reward is calculated based on the next state reached by the robot by taking the action generated by the behavior deep Q network (DQN) in the current state. We define the overall reward as (details are in Table I):  $f_R^{\theta_Q}(s_t, a_t) = r_{d,d'}^t + r_{arr}^t + r_{obs}^t$ . Here,  $s_t$  and  $a_t$  denote the state array and action of the robot  $R$  respectively at timestep  $t$ . The reward is computed following the policy  $\theta_Q$ , encapsulating the DQN learning agent. Intuitively, the relative distance reward ( $r_{d,d'}^t$ ) is increased when the robot gets closer to the target (target coordinates:  $(Tg_x^t, Tg_y^t)$  at timestep  $t$ ) and vice versa. On reaching the goal state the robot achieves the

Table I: The reward function of the proposed DQN agent.

Symbol	Reward Description	Reward Value
$r_{d,d'}^t$	Relative distance to target	$-( Tg_x^t - R_x^t ^2 +  Tg_y^t - R_y^t ^2)$
$r_{arr}^t$	Reaching target	5000
$r_{obs}^t$	Distance to obstacle	-5000 (if $( Tg_x^t - R_x^t ^2 +  Tg_y^t - R_y^t ^2) = 0$ ) 0 (otherwise)

reward of  $5K$ . On the other hand ( $r_{obs}^t$ ) punishes the robot when colliding with any obstacle with a reward of  $-5K$ .

**Target Network.** The target network policy ( $\theta_T$ ) is defined using a feed forward network called as target network which takes the next state  $s_{t+1}$  as input and estimates the target Q values for training the behavior Q network. The target Q values are denoted as:  $Q_t^T = f_{\theta_T}(s_{t+1})$ . The rewards received by traversing to the next state is also added to the target Q values discounted by a factor of gamma representing the future sum of rewards, eventually:  $Q_t^T = Q_t^T + f_R^{\theta_Q}(s_t, a_t)$ .

**Model Training on Blueprint Digital Twins.** At the training phase, the DQN agent interacts with the environment by passing the current state  $s_t$  to the behavior Q network  $f_{\theta_Q}(\cdot)$  which generates the action traversing the agent to the next state  $s_{t+1}$  using the behavior Q values  $Q_t^Q$  following the  $\epsilon$ -greedy policy. This next state  $s_{t+1}$  is passed to the target network  $f_{\theta_T}(\cdot)$  which generates the target Q values  $Q_{t+1}^T$ , which works as target values after discounting them and adding reward to them, for the behavior Q network. Hence we design a custom loss function  $\mathcal{L}(\cdot)$  using the Bellman equations [23] and a discount factor of  $\gamma$ . The loss function  $\mathcal{L}(\cdot)$  considers  $Q_t^Q$  as current Q value and  $Q_{t+1}^T$  as target Q value. The  $\mathcal{L}(\cdot)$  uses Huber loss which acts like the mean squared error when the error is small, but like the mean absolute error when the error is large. Formally,

$$\mathcal{L}(Q_t^Q, Q_{t+1}^T) = \begin{cases} \frac{1}{2}(Q_{t+1}^T - Q_t^Q)^2 & \text{for } |Q_{t+1}^T - Q_t^Q| \leq 1 \\ |Q_{t+1}^T - Q_t^Q|^2 - \frac{1}{2} & \text{Otherwise} \end{cases} \quad (2)$$

Depending on the loss, the behavior model weights are updated through back propagation. After certain number of training iterations, the model weights of the behavior Q network is copied to the target Q network so that the target Q values are generated using the updated trained network.

#### D. Module 3: RL-based Navigation (Inference)

During the inference, the states are extracted by following *Module 2* (Sec. IV-B), denoted as:  $s_t^d$  (details in Sec. IV-C2).

**Model Inference on both Blueprint and Dynamic Digital Twins.** The generated action during inference:  $a_t^d = \theta(s_t^d)$ , where  $a_t^d \in \{X-, X+, Y-, Y+\}$  and  $\theta$  is the trained model from *Module 3*, Sec. IV-C. The inference can be applied to both the blueprint and dynamic digital twins.

## V. DT-RADAR DATASET

In this section we introduce, to the best of our knowledge, a pioneering ray-tracing-based digital twin dataset that accurately replicates both indoor laboratory environments and outdoor smart city scenarios, specifically modeled after downtown Dallas and downtown Houston. These meticulously crafted

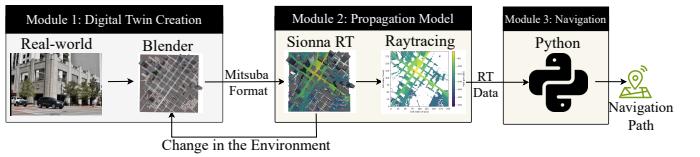


Fig. 4. Overview of various used tools for implementing DT-RaDaR.

digital twins, developed using publicly available software tools, offer high realism and relevance to real-world conditions. Next, we provide step-by-step details of the dataset generation process for enabling the research community to replicate and further explore its applications.

**Experimental Platform.** We collect the DT-RaDaR dataset using: (a) an Intel i7 system with Ubuntu, Blender LTS v3.6.12 [21], and NVIDIA's Sionna RT [6] for *Module 1* and 2. The dataset generation includes digital twin creation, subsequent ray-tracing and propagation characterization using Sionna RT, as shown in the first two blocks of Fig. 4. Details about the robot navigation on the collected data is presented in subsequent section Sec. VI.

#### A. Experimental Scenarios

We consider four distinct scenarios, encompassing two indoor and two outdoor environments. The indoor environments include a laboratory cubicle area and our lab's meeting space, while the outdoor environments are represented by smart cities: downtown Dallas and downtown Houston. The indoor scenes include office equipments like tables, chairs, TVs, computers, walls, and floors. The material properties of these objects are represented as wood, metal, glass, marble, or concrete. We present the features of the physical world for each of the experimental scenarios.

- **Indoor. Lab Cubicle:** The lab cubicle area features 7 metal chairs that serve as student workstations, along with a table surrounded by 4 additional chairs for dining and relaxation. The representative physical world is shown in Fig. 5 (a).

- **Indoor. Lab Meeting:** The lab meeting area includes an open space with various objects towards the wall of the meeting room. We have a wooden table and circulating 8 metal chairs in the middle of the room, shown in Fig. 5 (d).

- **Outdoor. Downtown Dallas:** Downtown Dallas measuring 0.4 by 0.5 km, with coordinates [32.78243, -96.80136] for the northwest and [32.77759, -96.79671] for the southeast. The scene includes 37 buildings, along with roads and 10 parking lots, with object materials represented as wood, metal, brick, concrete, or marble. The considered physical area is shown in Fig. 5 (g) in map form.

- **Outdoor. Downtown Houston:** For Downtown Houston, we choose an area measuring 0.8 by 0.8 km, with coordinates [29.75959, -95.37055] for the northwest and [29.75249, -95.36208] for the southeast. This scene includes 58 buildings, 29 parking lots, and numerous roads, with materials represented as wood, metal, glass, marble, brick, or concrete. The considered map is shown in Fig. 5 (j).

#### B. Digital Twin Creation

We create the digital twin from our physical indoor laboratory environment, as shown in the first block of Fig. 4.

By harnessing Blender and its OpenStreetMap (OSM) add-on, we digitally recreate an accurate indoor environment from the real world, including scene objects with various radio material properties. The inclusion of obstacles in the digital twin is important for providing information about the environment and guiding path planning and decision-making. This is because effective obstacle detection, avoidance, and understanding are essential capabilities for autonomous robots operating in diverse and dynamic environments. This digital twin is exported from Blender in *Mitsuba 3 .xml* file format which is responsible for rendering the scene and importation to Sionna RT [6] for propagation modeling. In Sionna, a set of international telecommunication union (ITU) materials with corresponding radio properties [6] is available for each object, ensuring realism and compatibility. Note that the utilization of Blender is solely for the digital twin generation and is not involved in making decisions for robot navigation. More details about the digital twin creation are presented in our previous work [20], [14], [24].

1) *Blueprint Digital Twins*: We generate a blueprint digital twin that incorporates only *static* objects. The intuition behind generating this set of twins is to provide the RL agent the blueprint data of an environment to get trained on. Various properties of the such designed digital twins are:

- **Lab Cubicle Blueprint Digital Twin.** In this digital twin, we model 11 chairs, 7 cubicles, 1 table and wall replicating the physical world. The carpet floor is represented with *ITU-concrete*, the wall with *ITU-brick*, the chairs with *ITU-wood*, and the cubicles with *ITU-metal* in Blender. The generated twin is shown in Fig. 5 (b).

- **Lab Meeting Blueprint Digital Twin.** For our lab meeting space, we model 10 chairs, 4 tables, 1 TV, 2 monitors, 1 board, 2 racks, walls, and carpet to simulate the real world. In Blender, chairs, tables, boards, and racks are modeled with *ITU-wood*, the TV and monitors with *ITU-metal*, the walls with *ITU-brick*, and the carpet with *ITU-concrete*. The twin representation is in Fig. 5 (e).

- **Downtown Dallas Blueprint Digital Twin.** For the digital twin creation of downtown Dallas, we model the measured 0.4 by 0.5 km area consisting of 37 buildings, 10 parking lots, and numerous roads to replicate real-world characteristics. In Blender, the buildings are modeled with their physical properties using *ITU-brick*, *ITU-marble*, *ITU-concrete*, or *ITU-glass*. The parking lots and roads are modeled with *ITU-concrete*. The twin is shown in Fig. 5 (h).

- **Downtown Houston Blueprint Digital Twin.** We model the measured 0.8 by 0.8 km area of downtown Dallas, encompassing 58 buildings, 29 parking lots, and various roads to replicate the physical properties, shown in Fig. 5 (k). In Blender, the buildings are modeled with their physical properties using *ITU-brick*, *ITU-marble*, *ITU-concrete*, or *ITU-glass*. The parking lots and roads are modeled with *ITU-concrete*.

2) *Dynamic Digital Twins*: While the blueprint digital twins provide the training data, however it does not incorporate the mobile objects of a dynamic environment. Hence, we

generate another set of twins of the same scenarios by adding various mobile objects within the scene. These dynamic digital twins replicate the real-time scenarios. The details of various added dynamic objects are listed below:

- **Lab Cubicle Dynamic Digital Twins.** We add an additional chair, adjust the orientation of three chairs, and randomly position two people in the scene, while modeling all added objects with their respective ITU-defined materials. The representation of the dynamic and corresponding blueprint twins are presented in Fig. 6 (b) and (a), respectively.

- **Lab Meeting Dynamic Digital Twin.** We place two laptops on the table, add two chairs and a stack of boxes, and modify the orientation of two chairs within the scene, ensuring all objects are modeled with their ITU-defined materials. The generated twins are shown in Fig. 6 (c) and (d).

- **Downtown Dallas Dynamic Digital Twin.** We randomly position five cars and five buses on two roads to simulate mobility within the scene, representing all objects with their ITU-defined materials. The representations are shown in Fig. 6 (e) and (f).

- **Downtown Houston Dynamic Digital Twin.** We place four cars and two buses at random locations on two roads to show mobility in the scene, using ITU-specified materials to represent all objects. The generated twin and corresponding blueprint is shown in Fig. 6 (h) and (g), respectively.

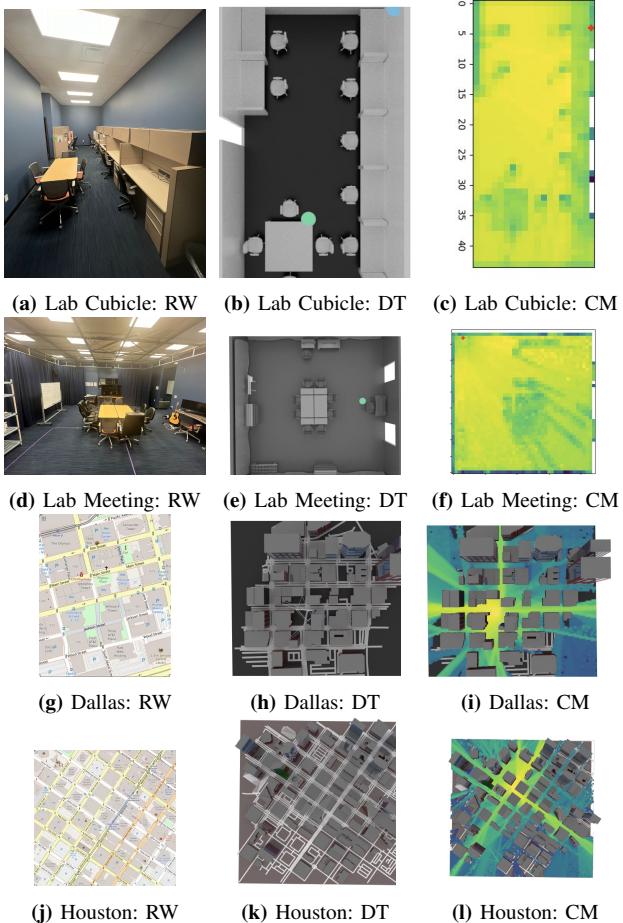
### C. Propagation Modeling

The propagation modeling process in Sionna RT commences with the integration of Sionna with the *Low Level Virtual Machine (LLVM) toolchain* to ensure smooth compatibility. This integration guarantees seamless scene loading, transmitters and receivers creation and configuration, replicating real-world characteristics within the scene. To accurately model RF propagation, signal behavior, and communication performance within a scene, Sionna classifies all scene components based on their *radio material* types and associated *material properties*. The transmitter and receiver antennas are configured as  $8 \times 2$  *tr38901* [22] with dual polarization and  $1 \times 1$  *dipole* planner array, respectively. We calculate propagation paths using the *compute\_paths* function in Sionna to generate ray-traced data and produce propagation map by setting *max-depth = 5* and *num-samples = 200*. Overall, our Sionna based implementation ensures improved computational efficiency by incrementally updating the ray paths based on changes in the receiver (or robot) location [22].

Overall, the coverage maps corresponding to the propagation models of the blueprint digital twins are presented in Fig. 5 (c, f, i, l) and Fig. 7 for quadrant-wise experiments. In these plots, the obstacles in the physical world yields to lower signal strength regions. The white regions of Fig. 7 represents the various buildings, hence absence of any signals.

### D. Generated Datasets

Following propagation modeling, a ray-tracing dataset is extracted comprising channel coefficients  $\Theta$ , phase shifts  $\phi$ , azimuth angles of departure  $\alpha_d$  and arrival  $\alpha_a$ , as well as



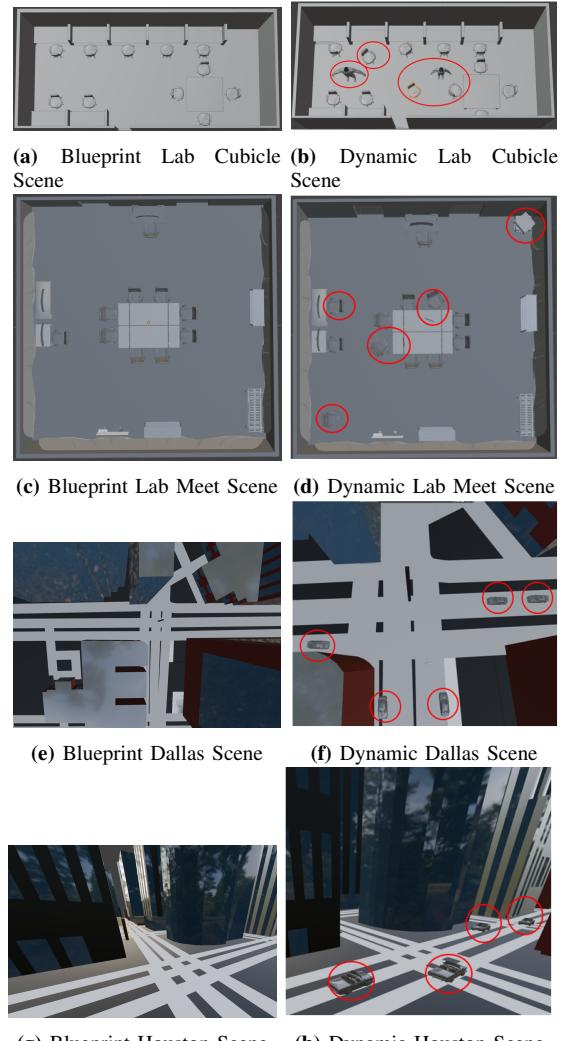
**Fig. 5.** The scene maps and coverage maps for both indoor and outdoor environments (details in Sec. IV-B). RW, CM and DT represents real-world, coverage map and digital twin, respectively.

zenith angles of departure  $\zeta_d$  and arrival  $\zeta_a$  for each cell of the generated propagation map. Overall we have generated the ray-tracing datasets corresponding to: (a) all four blueprint digital twins and (b) all four dynamic digital twins.

**(a) Blueprint Datasets.** At first, we generate the ray-tracing data for each of the four static blueprint digital twins. The dataset encompasses 4692 rows, equating to 4692 states, with a total size of 632KB. Various statistics of the dataset are presented in Table. II.

For collecting the ray-tracing data for Dallas and Houston downtown, we divide the total downtown region in four quadrants, following the realistic scenario of having multiple base-stations or access points for a city downtown area. The details about the transmitter and receiver positions specific to those are presented in Tables III and IV. The third and fourth rows of Table. II, shows the statistics of the merged data for all quadrants. These datasets will be released for public usage upon acceptance of this article.

**(b) Dynamic Datasets.** The dynamic datasets are collected by running the ray-tracing algorithm on the dynamic twins presented in Sec. V-B2. The Dallas dynamic twin is created on by updating the Dallas quadrant 3 (Q3) blueprint twin with the



**Fig. 6.** A snapshot of the dynamic digital twins (shown in second column) and corresponding blueprint twins (in the first column), details in Sec. V-B2. We have circled the added objects in the figure to highlight dynamicity of the scenario.

Table II: The description of the DT-RaDaR dataset encompassing the blueprint digital twins.

Scene Name	Grid Size	Data Size	Data Row
Cubicle	[42, 18]	95 KB	757
Meeting Room	[38, 40]	170 KB	1521
Downtown Dallas	[137, 130]	39 MB	456483
Downtown Houston	[507, 473]	20 MB	232363

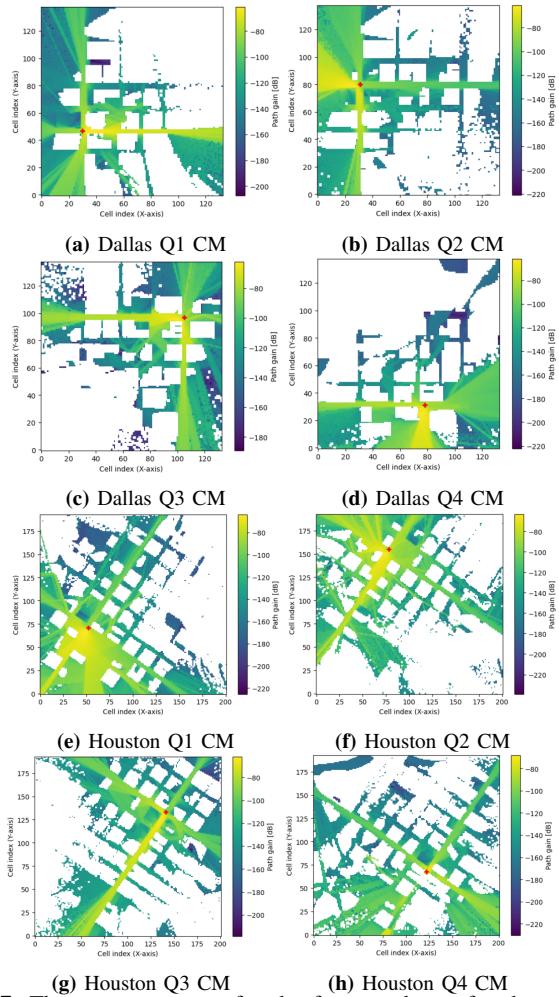
Table III: Overview of the scene configuration in downtown Dallas.

Downtown Dallas	Transmitter Position	Receiver Position
1st Quadrant (Q1)	[-170, -92, 18]	[-80, -92, 1.5]
2nd Quadrant (Q2)	[-165, 74, 18]	[-319, 0, 1.5]
3rd Quadrant (Q3)	[205, 159, 18]	[0, 0, 1.5]
4th Quadrant (Q4)	[70, -170, 18]	[0, -328, 1.5]

Table IV: Overview of the scene configuration in downtown Houston.

Downtown Houston	Transmitter Position	Receiver Position
1st Quadrant (Q1)	[-258, -118, 20]	[-507, -473, 1.5]
2nd Quadrant (Q2)	[-127, 304, 20]	[-507, 0, 1.5]
3rd Quadrant (Q3)	[189, 195, 20]	[0, 0, 1.5]
4th Quadrant (Q4)	[96, -134, 20]	[0, -473, 1.5]

added objects mentioned in Sec. V-B2. Similarly, the Houston



**Fig. 7.** The coverage maps for the four quadrants for the outdoor Dallas and Houston blueprint digital twin scenarios (Details in Sec. V-B1).

Table V: The description of the DT-RaDaR dataset encompassing the dynamic digital twins.

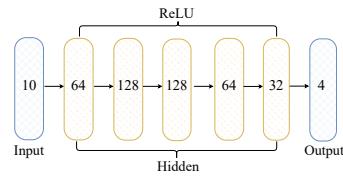
Scene Name	Grid Size	Data Size	Data Row
Cubicle	[42, 18]	90 KB	757
Meeting Room	[38, 40]	142 KB	1521
Downtown Dallas	[137, 130]	10 MB	123825
Downtown Houston	[180, 200]	5 MB	61216

dynamic twin is created on top of the Houston quadrant 2 (Q2) blueprint twin. The statistics of the generated ray-tracing data for the dynamic twins are shown in Table V.

**Remark 2.** The overall generated ray-traced data, representing the blueprint and dynamic digital twins, are  $< 100MB$  and  $< 10MB$ , which enables the DQN agent to have a faster training and inference, respectively (validates Contribution 4).

## VI. PERFORMANCE EVALUATION

**Experimental Platform.** We implement and validate the DT-RaDaR DQN agent (*Module 3*) on the DT-RaDaR dataset. The experiments are performed on an Intel (R) Xeon w7-2495x processor using Python with Pytorch, numpy, and matplotlib libraries.



**Fig. 8.** Neural network architecture used in DT-RaDaR framework.

**Evaluation Metrics:** To measure the performance of the training of the robot navigation, we evaluate the total rewards at each episode. We also consider the minimum number of steps the robot takes to reach the final destination. Finally, we also present the training time and end-to-end decision delay for the DT-RaDaR DQN agent.

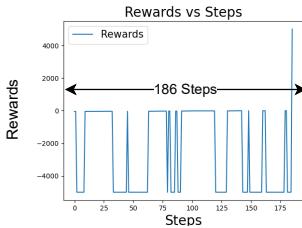
### A. Experimental Dataset

We validate our proposed DT-RaDaR framework using the blueprint and dynamic digital twins for two indoor and two outdoor scenarios, details in Sec. V. The experiments have been configured based on various training and inference scenarios involving the blueprint and dynamic digital twins.

### B. DQN Training

The deep neural network model we use to train our DQN agent within the DT-RaDaR framework consists of an input layer with 10 neurons representing the 10 features that the agent received in each state which are X coordinate ( $R_x^t$ ), Y coordinate ( $R_y^t$ ), azimuth angle of arrival ( $\alpha_a^t$ ), zenith angle of arrival ( $\zeta_a^t$ ), azimuth angle of departure ( $\alpha_d^t$ ), zenith angle of departure ( $\zeta_d^t$ ), channel coefficient's real and imaginary values ( $|\Theta|^t$ ), and propagation delay ( $\phi^t$ ) within the digital twin. The details about these inputs are given in Sec. IV-C2. It is followed by five dense layers each having the rectified linear unit (ReLU) as the activation function, as shown in Fig. 8. The output layer had 4 Dense units with linear activation function representing the Q-values of the four actions that the agent chooses from. We trained the agent using the DQN algorithm for 80 iterations in the outdoor environment and 500 iterations in the indoor environment. For each iteration, we take the number of steps as long as the agent reaches the target location from the starting location. We randomly set the starting location and the target location on the map. The behavior policy  $\theta$  is trained by updating the behavior DQN architecture using the Q values from the target DQN architecture. We follow an  $\epsilon$ -greedy policy to learn the Q values with an exploration rate  $\epsilon$  of 0.1. We use Bellman's optimality equation [23] to update the Q-values. Q-values of the behaviour network is trained with reference to the Q-values of the next state from the target network discounted with a factor of 0.9, of the action followed to reach that next state and added to the reward received in that next state.

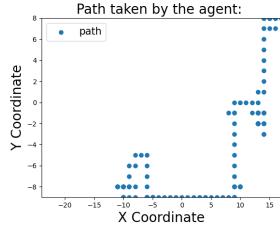
The source and target coordinates during training and inference of all experiments are taken randomly, considering that the coordinate is not an obstacle or building in the corresponding scenario.



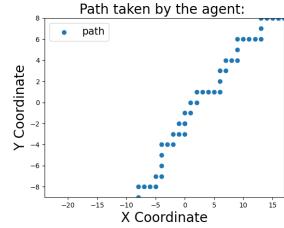
(a) Rewards vs Steps @Ep0



(b) Rewards vs Steps @Ep13



(c) Path Taken @Ep0



(d) Path Taken @Ep13

**Fig. 9.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of lab cubicle. We see 75.8% of reduction is required number of steps to 45th episode.

### C. Performance of DT-RaDaR DQN Agent (Trained/Inference on Blueprint Twin)

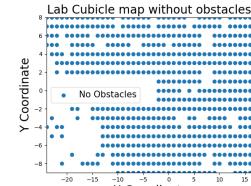
In this first set of experiments, we validate the DT-RaDaR DQN agent by training and inferring on the same scenarios.

- **Lab Cubicle (Training).** For the lab cubicle, we observe that the DQN agent took more than 185 steps to reach a specific target, leading to more fluctuating reward plots in the first episode, which is episode 0 in our experiments, as shown in Fig. 9 (a). The larger number of dips in the reward plot signify a huge number of collisions in this plot, which later improves over training in Fig. 9 (b). It is also evident from Fig. 9 (c) that the path taken by the agent at the first episode is not optimal or shortest, rather it is zigzagged at various locations, which is due to the lack of training of the agent and the random Q-values at the beginning of the training. After training for 45 number of episodes, we observe that the DQN agent learned to reach the target while avoiding the obstacles and yielding to lesser collisions, plotted in Fig. 10 (b). The path taken by the agent at this stage is also shown in Fig. 9 (d). The start and end coordinates are picked randomly which does not overlap with any obstacle region.

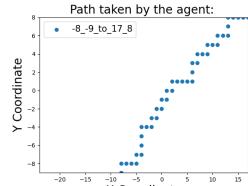
- **Lab Cubicle (Inference).** The DQN agent navigates the scene following the optimal path from the starting point at the coordinates [-8, -9] to the target point at [17, 8] in the XY coordinates system, as shown in Fig. 10 (b). These specific coordinates are picked randomly which does not coincide with the obstacles. Throughout navigation, the agent successfully avoids all obstacles, resulting in no collisions.

- **Indoor Meeting Space (Training).** We observe similar trend of the reward values and navigation path for the meeting space scenario, the plots are shown in Fig. 11.

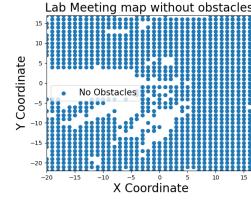
- **Indoor Meeting Space (Inference).** Similar to the cubicle scene, the agent navigates the meeting scene by following the optimal path from the source at coordinates [-20, 22] to the destination at coordinates [13, 17], as illustrated in Fig. 10 (d). Similar to previous experiments, these specific coordinates are



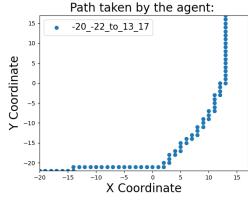
(a) Cubicle



(b) Nav Path

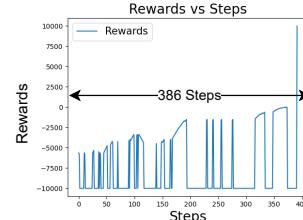


(c) Meeting

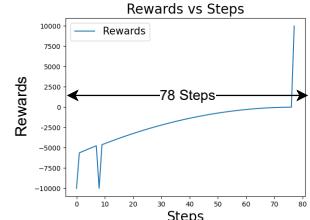


(d) Nav Path

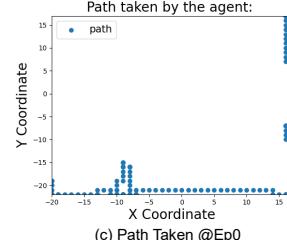
**Fig. 10.** Inference on blueprint digital twins (lab cubicle and meeting space). In (a) and (c), the areas with and without collisions are illustrated for the cubicle and meeting scenes, respectively. In contrast, (b) and (d) depict the robot's successful navigation, where it follows an optimal path while avoiding obstacles, from the starting point to the target in both the cubicle and meeting scenes, respectively.



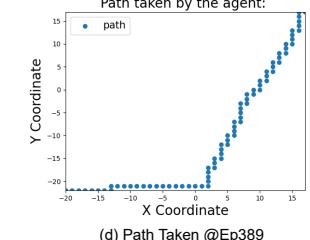
(a) Rewards vs Steps @Ep0



(b) Rewards vs Steps @Ep38



(c) Path Taken @Ep0



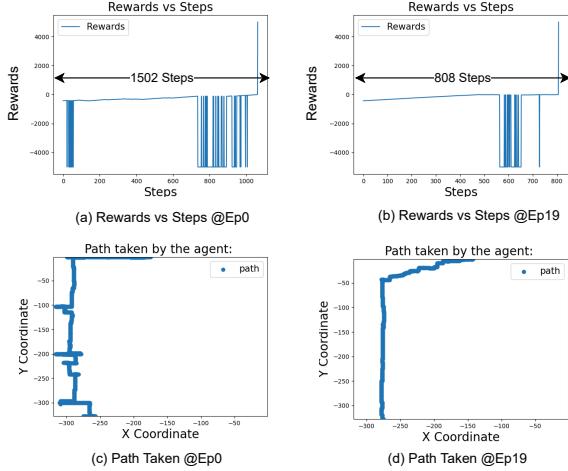
(d) Path Taken @Ep38

**Fig. 11.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of lab meeting space. We see 79.9% of reduction in required number of steps at 78th episode.

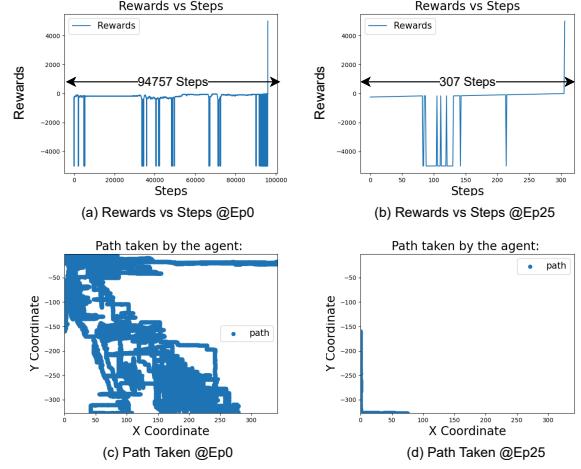
picked randomly which does not coincide with any obstacle. Observing the navigation area in Fig. 10 (c), it is clear that the agent successfully avoids all obstacles.

- **Downtown Dallas (Training).** In the downtown Dallas scenario, we perform quadrant-wise training over the blueprint twins, details of quadrants are in Sec. V. The training performance of all four different quadrants for this case is presented in Figs. 12, 13, 14, and 15. We observe, the DQN agent learned the optimal path around 19th, 32nd, 37th, and 25th episodes for Dallas first (Q1), second (Q2), third (Q3), and fourth (Q4) quadrants, respectively.

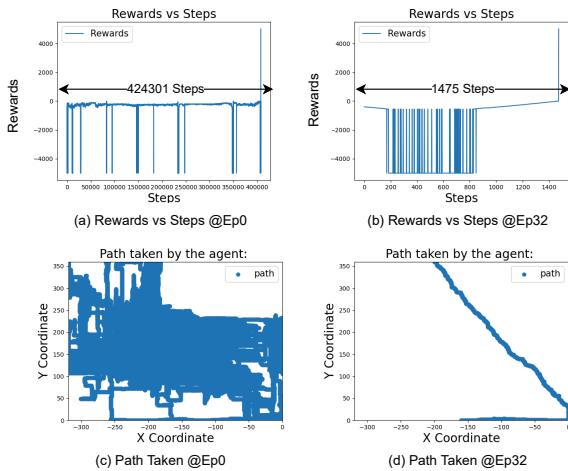
- **Downtown Dallas (Inference).** For the inference, we perform various quadrant-wise experiments: (i) inference on same quadrant, the model is trained on, and (ii) inference on different quadrant than the model is trained on. For the case (i), we observe that the trained DQN agent was able to



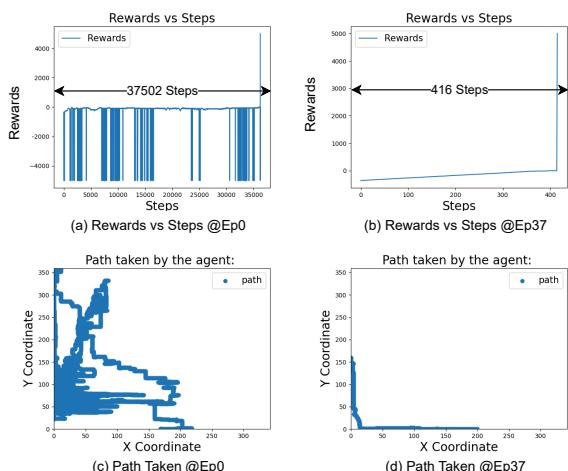
**Fig. 12.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Dallas Q1. We see a 46.2% of reduction in the required number of steps by the 19th episode.



**Fig. 15.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Dallas Q4. We see a 99.6% of reduction in the required number of steps by the 25th episode.



**Fig. 13.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Dallas Q2. We see a 99.6% of reduction in the required number of steps by the 32nd episode.



**Fig. 14.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Dallas Q3. We see a 98.8% of reduction in the required number of steps by the 37th episode.

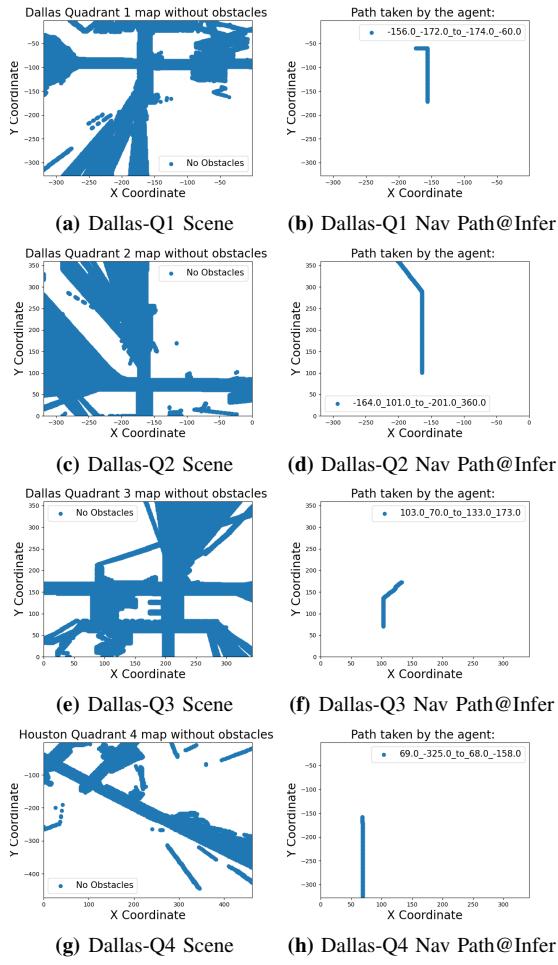
reach to target for most of the time, the detailed plots are in Fig. 16. For the case (ii), we conduct experiments using the DQN agent trained on all quadrants (Q1, Q2, Q3, and Q4) individually, and infer on the Q2 digital twin, the results are shown in Fig. 17 (a). From the plot, it is evident that, during inference, the trained agent on Q2 performed best on Q2, while competitive performance was given by the agent trained on Q1. Intuitively, this is due to the fact of similar landscape of the areas. We observe similar characteristics for other quadrants of the Dallas digital twin.

• **Downtown Houston (Training).** For Houston scenario, we observe, the minimum required episodes for the agent to learn is 14th, 67th, 55th, and 72nd for Houston Q1, Q2, Q3, and Q4 scenarios, respectively. The detailed plots showing the training performance for these experiments are in Figs. 18, 19, 20, 21 for Houston Q1, Q2, Q3, and Q4 scenarios, respectively.

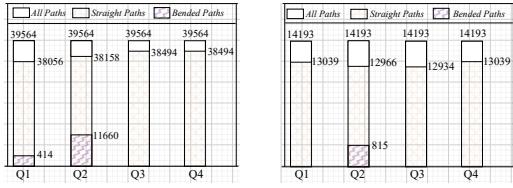
• **Downtown Houston (Inference).** Similar to the experiments conducted for Dallas, we also conduct quadrant-based testing for Houston. This involves: (i) running inference on the same quadrant the model was trained on, and (ii) running inference on a different quadrant from the one the model was trained on. In case (i), we observe that the trained DQN agent reaches the target almost every time successfully, with detailed plots presented in Fig. 22. In case (ii), we use the DQN agent trained in Q2 and evaluate its performance across all quadrants, with the results illustrated in Fig. 17 (b). The plot shows interesting behavior that the trained agent on Q2 performs best in Q3 digital twin of Houston.

**Observation 1. Training:** *We observe that the trained DT-RaDaR DQN agent learns the optimal path by avoiding obstacles and minimizes the collisions for all four blueprint digital twins (see Figs. 9, 11, 12, 13, 14, 15, 18, 19, 20, 21).*

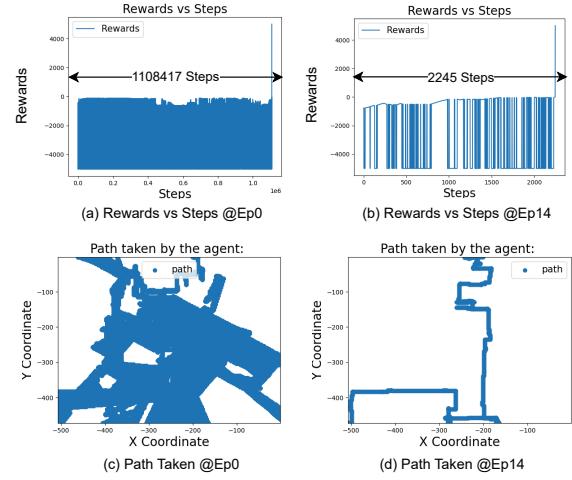
**Observation 2. Inference:** *We observe that the generated navigation paths by the DT-RaDaR DQN agent reaches the target location by avoiding the obstacle regions of various indoor and outdoor scenarios (see Figs. 10, 16, 22).*



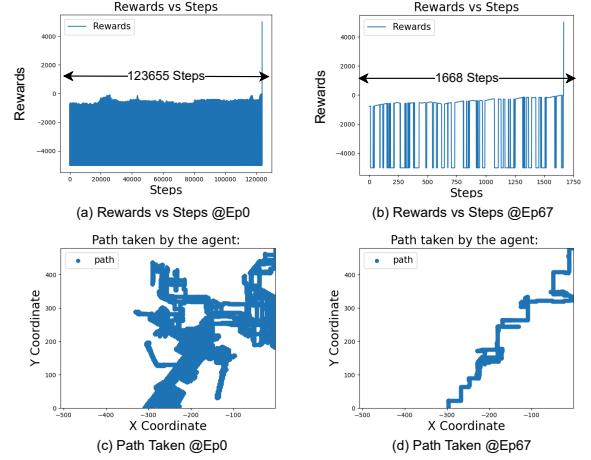
**Fig. 16.** Inference on blueprint digital twins (downtown Dallas). The first column (a), (c), (e), and (g), illustrates the non-obstacle regions within in each quadrant. Plots in (b), (d), (f), and (h) show the robot's successful movement from one point on the street to the target point (changing the X and Y coordinates) while avoiding obstacles in quadrants Q1, Q2, Q3, and Q4, respectively.



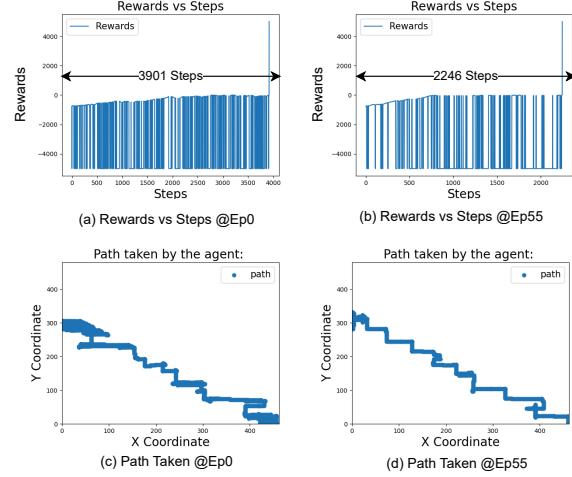
**Fig. 17.** Performance of quadrant-wise inference for Dallas and Houston when the model is trained on one part of the city (i.e., one quadrant) and analyzed on another part. The performance is intuitive to the phenomenon that the best performance is received when the model is trained and inferences on the twin from same quadrant. Each bar represents when the model is trained on that quadrant.



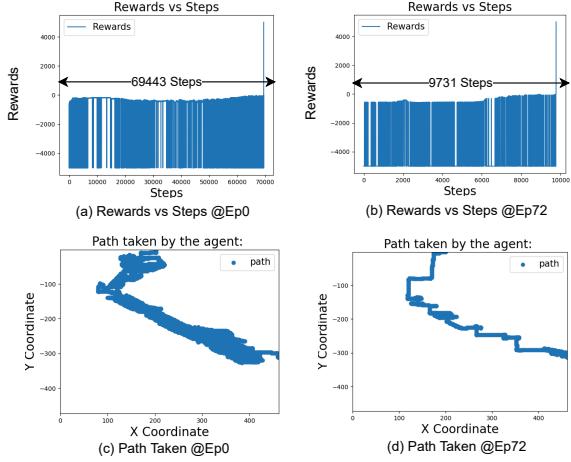
**Fig. 18.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Houston Q1. We see 99.7% of reduction is required number of steps to 14th episode.



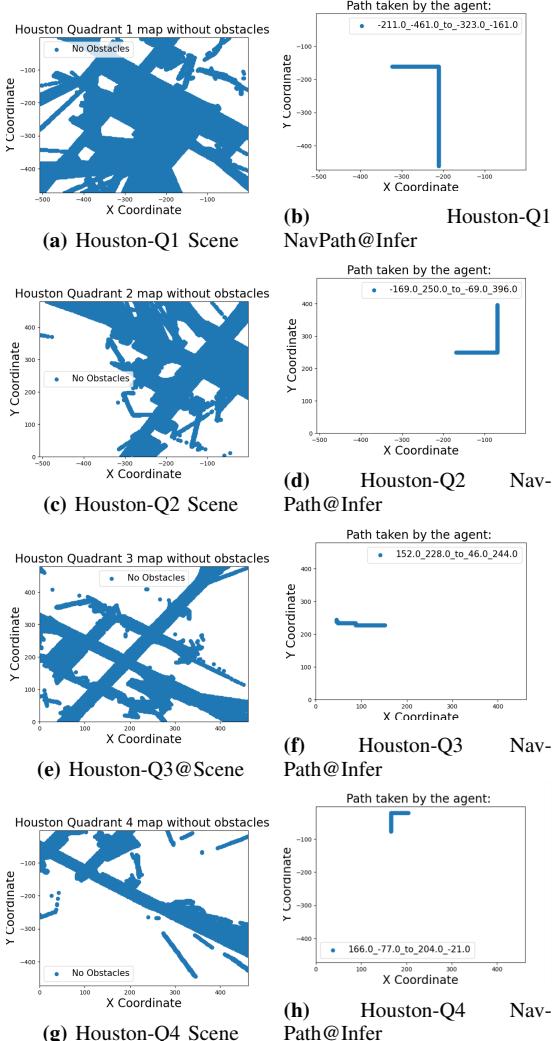
**Fig. 19.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Houston Q2. We see 98.6% of reduction is required number of steps to 67th episode.



**Fig. 20.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Houston Q3. We see 42.4% of reduction is required number of steps to 55th episode.



**Fig. 21.** The training performance of the DT-RaDaR DQN agent on the blueprint twin of Houston Q4. We see 85.98% of reduction is required number of steps to 72nd episode.



**Fig. 22.** Inference on blueprint digital twins (downtown Houston). Plots (a), (c), (e), and (g) illustrate the non-obstacle regions within each quadrant. Plots (b), (d), (f), and (h) highlight the robot's successful traversal between points, adjusting its X and Y coordinates and avoiding obstacles in quadrants Q1, Q2, Q3, and Q4.

**Table VI:** Glimpse of the training performance for all the blueprint twins w.r.t. minimum steps taken. The X and Y coordinates are relative locations received from the robot odometer sensor.

Scene Description	X and Y Coordinates	Minimum Steps Taken
Lab Cubicle	[-8, -9] [17, 8]	45 in 13 Episode
Lab Meeting	[-20, -22] [17, 17]	78 in 389 Episode
Dallas Q1	[-275, -328] [-175, -1]	808 in 19 Episode
Dallas Q2	[-160, 0] [-200, 360]	1475 in 32 Episode
Dallas Q3	[200, 0] [0, 160]	416 in 37 Episode
Dallas Q4	[75, -328] [0, -160]	307 in 25 Episode
Houston Q1	[-570, -473] [-1, -161]	2245 in 14 Episode
Houston Q2	[-297, 0] [-1, 140]	1668 in 67 Episode
Houston Q3	[460, 0] [0, 300]	2246 in 55 Episode
Houston Q4	[-462, -301] [200, -1]	6431 in 72 Episode

**1) Minimizing the Collisions in DT-RaDaR:** During training, the agent explores the environment and initially experiences a large number of negative rewards due to frequent collisions. However, over time, it learns to minimize these collisions by avoiding obstacles and improving its overall training performance for all blueprint twins (Figs. 9, 11, 12, 13, 14, 15, 18, 19, 20, and 21) illustrates the agent's collision frequency during training and how it gradually improved, as reflected in the increase in rewards and the reduction in steps taken. After training, the agent successfully learned the optimal path to reach its target from the starting point while avoiding obstacles as much as possible, as demonstrated in Figs. 9, 11, 12, 13, 14, 15, 18, 19, 20, and 21. Table VI presents the scene description, the number of steps taken in the episode, and the X and Y coordinates from the source to the destination for training in all the blueprint twins.

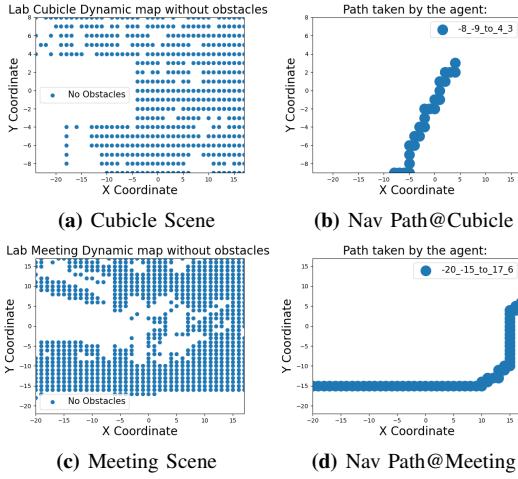
#### D. Performance of DT-RaDaR DQN Agent (Trained on Blueprint Twin and Inference on Dynamic Twin)

In the second set of experiments, we explore the performance of DT-RaDaR DQN agent when it is trained on a specific digital twin and later the digital twin is updated with more (or less) objects within the scene. The details about the dynamic digital twin generation is discussed in Sec. V-B2.

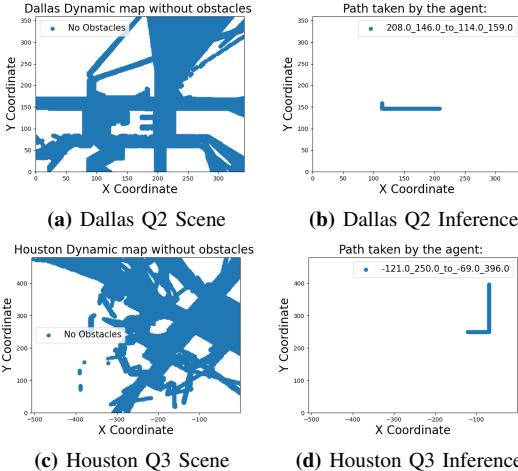
• **Dynamic Lab Cubicle (Inference).** The DQN agent which is trained on the blueprint lab cubicle digital twin is analyzed on the dynamic lab cubicle digital twin, the plots are shown in Fig. 23 (a) and (b). The blue dots in Fig. 23 (a) shows no obstacle regions or open areas within the dynamic digital twin of cubicle, which is different from the map shown in blueprint twin of the same in Fig. 10 (a). The navigation path shown in Fig. 23 (b) shows that the trained DQN agent is able to successfully navigate even with the dynamic changes.

• **Dynamic Meeting Space (Inference).** Similarly, the trained DQN agent on the blueprint lab meeting digital twin is analyzed on the dynamic lab meeting digital twin, the plots are in Fig. 23 (b) and (d). Fig. 23 (a) shows no obstacle regions or open areas within the dynamic digital twin of meeting space (the blueprint twin of the same is in Fig. 10 (c)). The navigation path is shown in Fig. 23 (d).

• **Dynamic Downtown Dallas (Inference).** For the outdoor scenarios, the trained model on the blueprint twins are analyzed over the dynamic twin for Dallas digital twin. The plots to signify the non-blocked areas and a sample navigation path are given in Fig. 24 (a) and (b), respectively. Note that



**Fig. 23.** Inference on dynamic digital twins (lab cubicle and meeting space). In (a) and (c), the areas with and without obstacles are shown for the cubicle and meeting space scenes, respectively. While (b) and (d) depict the robot's successful navigation, avoiding obstacles, for cubicle and meeting spaces' dynamic digital twins, respectively.



**Fig. 24.** Inference on dynamic digital twins (Dallas and Houston). In (a) and (c), the areas with and without blockages and obstacles are shown for the Dallas and Houston downtown, respectively. The paths in (b) and (d) depict the robot's successful navigation, avoiding blockages and obstacles.

generated map signifying the non-blocked areas (i.e., where there are no buildings) in Fig. 24 (a) is different than the blueprint twin of the same shown in Fig. 16 (e).

• **Dynamic Downtown Houston (Inference).** The plots for non-blocked areas and a sample navigation path for the Houston are given in Fig. 24 (c) and (d), respectively. Similar to the Dallas case, for Houston also the generated maps signifying the non-blocked areas in Fig. 24 (c) is different than the blueprint twin of the same shown in Fig. 22 (c).

**Observation 3.** *The trained DQN agent is able to successfully navigate within the dynamic changes (see Figs. 23 and 24).*

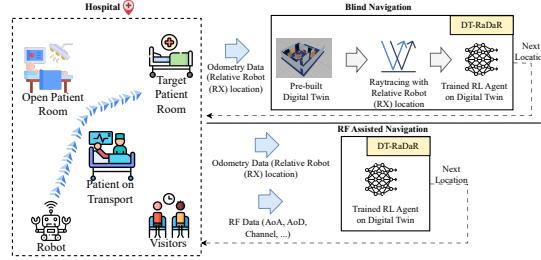
**Remark 3.** *The in-depth analysis of the percentage of changes a trained model can adapt, is open for further investigation.*

#### E. Comparison with State-of-the-art

We compare the overall training time with the state-of-the-art [25], [26], [27], it is evident in Table. VII, that DT-RaDaR

**Table VII:** Comparison with state-of-the-art w.r.t. training time.

Paper	Steps	Sensors	Computing Platform	Training Time
Niu <i>et al.</i> [25]	25K	LiDAR	NVIDIA TITAN RTX GPU	133 minutes
Beomsoo <i>et al.</i> [26]	600K	LiDAR	Not mentioned	180 minutes
Hu <i>et al.</i> [27]	10K	LiDAR	NVIDIA GTX 1080 GPU and Intel Core i9 CPU	40 minutes
DT-RaDaR	600K	Ray-tracing Data	32 GB Ryzen CPU	46 minutes



**Fig. 25.** Applicability of DT-RaDaR in a hospital scenario. We show two approaches on how DT-RaDaR can be used for robot navigation in the hospital scenario: (a) blind and (b) RF-assisted navigation.

outperforms the other methods in terms of training time even when trained on a computing platform with only CPU.

**Observation 4.** *We observe that the proposed DT-RaDaR framework is able to yield competitive training time with the state-of-the-art approaches. DT-RaDaR requires lesser training time ( $\sim < 74.4\%$ ) even when run on a CPU computing platform (see Table VII, validates Contribution 5).*

#### F. End-to-end Decision Delay

The end-to-end delay is calculated based on the two measurements: (a) the ray-tracing time per step, which is 16ms and (b) the inference time to generate the next action for the robot, which is 17ms. Overall, the end-to-end decision delay is summed up to be: 33ms when ran on an edge device.

#### G. Applicability of DT-RaDaR

The DT-RaDaR framework is particularly applicable in privacy-sensitive environments, such as hospitals. In these dynamic settings, where doctors, patients, and visitors are constantly moving, traditional SLAM algorithms struggle with navigation challenges [28]. DT-RaDaR enables robots to navigate effectively in crowded spaces while protecting patient privacy. As shown in Fig. 25, a robot can autonomously reach a patient's room to initiate a teleconference for doctor-patient communication, utilizing ray-tracing data to avoid potential privacy risks associated with conventional sensors. This approach ensures efficient navigation while addressing the privacy concerns of all individuals present. Next we discuss, how the DT-RaDaR can be used in such scenario:

1) **DT-RaDaR in Blind Navigation:** One way of using DT-RaDaR is blind navigation strategy where only the robot's odometry data is fed to our framework as shown in Fig. 25. The robot will be using off-the-shelf DT-RaDaR framework which includes a pre-build digital twin, ray-tracing suite, and trained RL model. However, this approach may result in sub-optimal performance if the real-world environment undergoes significant changes that differ from the digital twin model.

2) *DT-RaDaR in RF-assisted Navigation:* Various RF data such as angle-of-arrival, angle-of-departure, and channel coefficients [29], [30] are fed to the trained RL agent of DT-RaDaR, as shown in Fig. 25. The RL agent is trained on a digital twin of the hospital scenario, enabling real-time navigation with RF data for accuracy and privacy preservation.

#### H. Limitations and Future Possibilities of DT-RaDaR

The DT-RaDaR framework establishes a foundational approach for privacy-preserving robot navigation using ray-tracing data, but it has limitations that highlight future research opportunities. These include potential discrepancies between real-world environments and digital twin representations, the impact of external factors like weather on signal quality, and the current inability of digital twins to capture fine-grained, rapid changes in the physical world. Additionally, while the framework's privacy benefits are a key contribution, further emphasis on existing work addressing privacy concerns in sensor-based navigation is needed to clearly demonstrate DT-RaDaR's superior protection. Addressing these challenges could involve implementing dynamic update loops between digital twins and real environments, developing robust interference mitigation strategies, identifying optimal application scenarios, and conducting comparative privacy analyses against traditional sensor-based methods.

## VII. CONCLUSION

We propose a DQN based robot navigation approach which is applied on the digital twins of a real-world scenario for the navigation of delivery or service robots. The digital twins are designed using open source Blender software and the ray-tracing data is generated through the Sionna RT software. The thorough experimentation shows that the idea of using ray tracing data for robot navigation is feasible. It also provides some resiliency when trained over the digital twins of one scenario and analyzed over the digital twins of another type of scenario. The experimentation on quantifying such resiliency and extending to more generalized twins would be our next step for future work.

## ACKNOWLEDGMENT

We acknowledge usage of AI tools for solely language polishing purpose only.

## REFERENCES

- [1] M. Hossain, "Autonomous delivery robots: A literature review," *IEEE Engineering Management Review*, vol. 51, no. 4, pp. 77–89, 2023.
- [2] S. Srinivas, S. Ramachandiran, and S. Rajendran, "Autonomous robot-driven deliveries: A review of recent developments and future directions," *Transportation research part E: logistics and transportation review*, vol. 165, p. 102834, 2022.
- [3] M. Camplani, A. Paiment, M. Mirmehdi, D. Damen, S. Hannuna, T. Burghardt, and L. Tao, "Multiple human tracking in rgb-depth data: a survey," *IET computer vision*, vol. 11, no. 4, pp. 265–285, 2017.
- [4] S. Gatesichapakorn, J. Takamatsu, and M. Ruchanurucks, "Ros based autonomous mobile robot navigation using 2d lidar and rgb-d camera," in *First international symposium on instrumentation, control, artificial intelligence, and robotics (ICA-SYMP)*. IEEE, 2019, pp. 151–154.
- [5] Z. Yun and M. F. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE access*, vol. 3, pp. 1089–1100, 2015.
- [6] J. Hoydis, F. A. Aoudia, S. Cammerer, M. Nimier-David, N. Binder, G. Marcus, and A. Keller, "Sionna rt: Differentiable ray tracing for radio propagation modeling," *arXiv preprint arXiv:2303.11103*, 2023.
- [7] J. Hoydis, F. A. Aoudia, S. Cammerer, F. Euchner, M. Nimier-David, S. ten Brink, and A. Keller, "Learning radio environments by differentiable ray tracing," 2023.
- [8] B. Song, H. Xu, W. Hu, Y. Li, and Y. Guo, "How to calculate privacy: privacy concerns and service robots' use intention in hospitality," *Current Issues in Tourism*, pp. 1–17, 2023.
- [9] A. Arabo, I. Brown, and F. El-Moussa, "Privacy in the age of mobility and smart devices in smart homes," in *International Conference on Privacy, Security, Risk and Trust*. IEEE, 2012, pp. 819–826.
- [10] <https://github.com/TWIST-Lab/SionnaNetworkedRobotics>.
- [11] D. Almeida, E. Pedrosa, and F. Curado, "Magnetic mapping for robot navigation in indoor environments," in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2021, pp. 1–8.
- [12] S. Noh, J. Park, and J. Park, "Autonomous mobile robot navigation in indoor environments: Mapping, localization, and planning," in *International Conference on Information and Communication Technology Convergence (ICTC)*, 2020, pp. 908–913.
- [13] H. Bavle, J. L. Sanchez-Lopez, M. Shaheer, J. Civera, and H. Voos, "Situational graphs for robot navigation in structured indoor environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9107–9114, 2022.
- [14] S. Amatare, M. Samson, and D. Roy, "Testbed design for robot navigation through differential ray tracing," in *2024 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2024, pp. 173–174.
- [15] J. Kulhánek, E. Derner, and R. Babuška, "Visual navigation in real-world indoor environments using end-to-end deep reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4345–4352, 2021.
- [16] S. Palazzo, D. C. Guastella, L. Cantelli, P. Spadaro, F. Rundo, G. Muscato, D. Giordano, and C. Spampinato, "Domain adaptation for outdoor robot traversability estimation from rgb data with safety-preserving loss," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 10014–10021.
- [17] H. Karnan, A. Nair, X. Xiao, G. Warnell, S. Pirk, A. Toshev, J. Hart, J. Biswas, and P. Stone, "Socially compliant navigation dataset (scand): A large-scale dataset of demonstrations for social navigation," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11807–11814, 2022.
- [18] M. Elnoor, A. J. Sathyamoorthy, K. Weerakoon, and D. Manocha, "Pronav: Proprioceptive traversability estimation for legged robot navigation in outdoor environments," *IEEE Robotics and Automation Letters*, vol. 9, no. 8, pp. 7190–7197, 2024.
- [19] M. Sorokin, J. Tan, C. K. Liu, and S. Ha, "Learning to navigate sidewalks in outdoor environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3906–3913, 2022.
- [20] S. Amatare, G. Singh, M. Samson, and D. Roy, "RagNAR: Ray-tracing based Navigation for Autonomous Robot in Unstructured Environment," in *IEEE Global Communications Conference*, December 2024.
- [21] B. O. Community, *Blender - a 3D modelling and rendering package*, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. [Online]. Available: <http://www.blender.org>
- [22] N. Inc., "Ray tracing," 2024, last accessed 6 April 2024. [Online]. Available: <https://nvlabs.github.io/sionna/api/rt.html>
- [23] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018.
- [24] S. Amatare, G. Singh, A. Kharel, and D. Roy, "Real-time localization of objects using radio frequency propagation in digital twin," *Available at SSRN 4937841*, 2024.
- [25] H. Niu, Z. Ji, F. Arvin, B. Lennox, H. Yin, and J. Carrasco, "Accelerated sim-to-real deep reinforcement learning: Learning collision avoidance from human player," in *IEEE/SICE International Symposium on System Integration (SII)*, 2021, pp. 144–149.
- [26] H. Beomsoo, A. A. Ravankar, and T. Emaru, "Mobile robot navigation based on deep reinforcement learning with 2d-lidar sensor using stochastic approach," in *IEEE International Conference on Intelligence and Safety for Robotics (ISR)*. IEEE, 2021, pp. 417–422.
- [27] J. Hu, H. Niu, J. Carrasco, B. Lennox, and F. Arvin, "Voronoi-based multi-robot autonomous exploration in unknown environments via deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14413–14423, 2020.
- [28] C. Yu, Z. Liu, X.-J. Liu, F. Xie, Y. Yang, Q. Wei, and Q. Fei, "Ds-slam: A semantic visual slam towards dynamic environments," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1168–1174.

- [29] M. Yang, B. Ai, R. He, C. Huang, Z. Ma, Z. Zhong, J. Wang, L. Pei, Y. Li, and J. Li, "Machine-learning-based fast angle-of-arrival recognition for vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1592–1605, 2021.
- [30] S.-L. Shih, C.-K. Wen, S. Jin, and K.-K. Wong, "Angle-of-arrival estimation with practical phone antenna configurations," *IEEE Internet of Things Journal*, vol. 10, no. 22, pp. 20 006–20 020, 2023.