

RNBF: Real-Time RGB-D Based Neural Barrier Functions for Safe Robotic Navigation

Satyajeet Das^{*1}, Yifan Xue², Haoming Li² and Nadia Figueroa²

Abstract—Autonomous safe navigation in unstructured and novel environments poses significant challenges, especially when environment information can only be provided through low-cost vision sensors. Although safe reactive approaches have been proposed to ensure robot safety in complex environments, many base their theory off the assumption that the robot has prior knowledge on obstacle locations and geometries. In this paper, we present a real-time, vision-based framework that constructs continuous, first-order differentiable Signed Distance Fields (SDFs) of unknown environments entirely online, without any pre-training, and is fully compatible with established SDF-based reactive controllers. To achieve robust performance under practical sensing conditions, our approach explicitly accounts for noise in affordable RGB-D cameras, refining the neural SDF representation online for smoother geometry and stable gradient estimates. We validate the proposed method in simulation and real-world experiments using a Fetch robot.

I. INTRODUCTION

Safety in robotics is a critical challenge, especially when operating in unstructured environments where obstacle states and geometries are unavailable [1]. In recent years, Control Barrier Functions (CBFs) were widely studied due to their ability to provide safety guarantees for any control affine systems and adopted in various applications including vehicle cruise control [2], indoor navigation [3], quadrotor control [4], legged robot locomotion [5], etc. All the aforementioned works assume that the robot has prior knowledge of obstacle states and geometries, or even a predefined map. However, for robot navigation in unknown environments, such assumptions are unrealistic and the robot can only access noisy environment information from its onboard sensors.

CBFs's performance in enforcing robot safety relies not only on the robot's ability to collect accurate nearby obstacle state information but also on how well the collected information is processed to meet the requirement of valid CBFs. Therefore, when environment descriptions are unavailable or noisy, the safety guarantee of robot navigation is jeopardized. Building valid CBFs adaptively using noisy data collected from robots' on-board sensors in real time remains a challenge. Many recent works use LiDAR sensors to construct CBFs and ensure robot safety in unknown environments [6]–[9]. In this work, we instead demonstrate that safe reactive

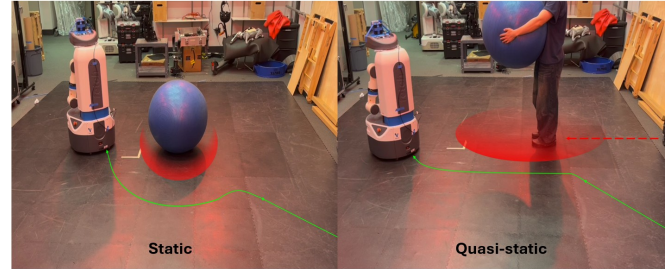


Fig. 1: Hardware setup for real-world experiments showing the RNBF-Control pipeline navigating around static and quasi-static obstacles without any prior knowledge about the obstacles.

control can be achieved with RGB-D camera alone, without relying on additional sensing modalities such as LiDAR.

Seldom works have investigated perception-based CBFs using RGB-D camera inputs. Existing vision-based CBF pipelines utilize machine-learning techniques to incorporate robot observations into CBF formulations. For example, [10] leverages a conditional Generative Adversarial Network to approximate CBFs given RGB-D camera inputs; [11] predicts future RGB-D observations with Neural Radiance Fields (NeRFs) and feeds them into discrete-time CBFs. These methods require significant offline training and their performance can be affected by the quality of their training dataset. Additionally, using CBFs in discrete-time settings may weaken their safety guarantees when the update frequency is not fast enough. Other works, like [12], approximate mathematically CBFs and their gradients from point cloud data collected through LiDARs or depth cameras.

Among the methods listed above, it is noteworthy that none has explicitly considered the effects of sensory noises on perception-based CBF. Further, most existing vision-based CBFs have made approximations to standard CBFs to accommodate robot observation formats. Assumptions that CBFs are founded on, such as the continuity and first-order differentiability of the barrier function are typically relaxed, weakening their safety guarantees. Additionally, altering standard CBFs in vision-based safe navigation hinders the adaptation of well-developed none-sensor-based safe reactive control algorithms into vision-based robot platforms. Many safe reactive controllers, including Control Barrier Function Quadratic Programs (CBF-QP), are developed assuming the system's access to SDFs and their gradients of the obstacles near the robot [13], [2], [14], [15], [16] and [3]. In robot safe navigation, SDFs refer to the orthogonal distance of a given point to the boundary of an unsafe set.

^{*}Corresponding author.

¹Satyajeet Das is with the University of Southern California, Los Angeles, CA, 90007 USA, satyajee@usc.edu. This research was primarily conducted while at the University of Pennsylvania.

²Yifan Xue, Haoming Li, and Nadia Figueroa are with the University of Pennsylvania, Philadelphia, PA, 19104 USA, yifanxue@seas.upenn.edu, lihaomingforreal@gmail.com, nadiafig@seas.upenn.edu.

Therefore, in this paper, we offer an online barrier function generation framework capable of smoothing sensor noises and constructing continuous and first-order differentiable neural SDFs in real-time given RGB-D camera inputs without pre-training for any unknown environments. With the advancements presented by NeRF [17], the quality and efficiency of environmental reconstruction have significantly improved. However, limited work explored generating SDFs using NeRF, most of which were designed for an offline setting [18], [19]. A notable exception is iSDF [20], which introduced a method for real-time SDF generation. Recently, [21] proposed a dual-level SDF approach combining pre-trained object SDFs with a scene-level iSDF for indoor navigation. Their method focused on trajectory optimization and relies on offline components, without incorporating any safety-filtering framework. In contrast, our vision pipeline generates real-time, continuous, and differentiable SDFs, specifically designed for constructing continuous CBFs for safe robot navigation. Instead of altering CBFs to accommodate vision observations, as in prior works, we establish continuous and first-order differentiable SDFs online, which enables our methods to be readily adaptable to SDF-based safe reactive controllers such as standard CBF-QPs [22], [23] and dynamical system based modulation [24], [25]. Our approach is validated in both gazebo simulations and real-life hardware experiments using the Fetch robot in static and quasi-static obstacle environments.

Our **key contributions** are as follows:

- We introduce an approach to generate real-time (5-15 Hz) SDFs from raw RGB-D data using an online neural representation compatible with established SDF-based reactive controllers.
- To account for depth noise inherent in affordable sensors like RGB-D cameras, we mitigate its impact on neural SDF representations adopting the huber loss [26].
- Leveraging the neural SDF representation, we propose a method to construct barrier functions in real-time without prior training, leveraging cost-effective RGB-D camera for safe navigation in novel environments.

II. PROBLEM FORMULATION

Consider a control-affine nonlinear system of the form

$$\dot{x} = f(x) + g(x)u, \quad (1)$$

where $x \in \mathbb{R}^n$ is a state vector, $u \in \mathbb{R}^m$ is a control input vector, $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, and $g: \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ define the control-affine nonlinear dynamics. We define the notion of robot safety and consider safe control algorithms based on barrier functions.

Given a continuously differentiable function $h: \mathbb{R}^n \rightarrow \mathbb{R}$, h is a barrier function if the safe set C (outside the obstacle), the boundary set ∂C (on the boundary of the obstacle), and the unsafe set $-C$ (inside the obstacles) of the system can be defined as in (2), (3), (4) [22].

$$C = \{x \in \mathbb{R}^n : h(x) > 0\} \quad (2)$$

$$\partial C = \{x \in \mathbb{R}^n : h(x) = 0\} \quad (3)$$

$$-C = \{x \in \mathbb{R}^n : h(x) < 0\} \quad (4)$$

In practice, $h(x)$ is often measured as the shortest distance from state x to the boundary set ∂C , i.e. a signed distance function (SDF).

Assumptions: We assume the robot has access to its poses (through odometry or external sensing) in real-time during the vision-based navigation. The entire vision-based navigation pipeline takes in 2 poses: the camera pose x_{cam} and the robot actuation center pose x_{rob} measured in the world frame. Given $h(x)$ that returns the Euclidean distance to the nearest obstacle, it can be deduced that $|h(x_{\text{rob}}) - h(x_{\text{cam}})| < \zeta$, where ζ is the Euclidean distance between the robot and camera centers. Since our framework is designed specifically for robots with onboard RGB-D cameras, where ζ is a small restricted value, we assume that $h(x_{\text{rob}}) \approx h(x_{\text{cam}}) = h(x)$ and ensure robot safety by defining a safety margin greater or equal to ζ .

Problem Statement: Given a robot with accurate knowledge of its onboard camera poses in an unexplored environment, the goal of our vision-based navigation pipeline is to (i) take in noisy RGB-D camera inputs, (ii) generate continuous and first-order differentiable SDFs as barrier functions and (iii) compute an admissible input u that will ensure the state of the robot x is always within the safe set C defined in (2) using standard CBF-QP.

III. PRELIMINARY

A. Control Barrier Function

To guarantee safety of the controlled agent, CBF-QP utilizes the Nagumo Set Invariance Theorem.

Definition 3.1 (Nagumo Set Invariance)

$$C \text{ is set invariant} \iff \dot{h}(x) \geq 0 \forall x \in \partial C \quad (5)$$

Control Barrier Functions are designed by extending Nagumo set invariance theorem to a ‘‘control’’ version [22], where the condition $\forall x \in \partial C$ is rewritten mathematically using an extended \mathcal{K}_∞ function α ,

$$C \text{ is set invariant} \iff \exists u \text{ s.t. } \dot{h}(x, u) \geq -\alpha(h(x)). \quad (6)$$

Definition 3.2 (Extended \mathcal{K}_∞ Functions) An extended \mathcal{K}_∞ function is a function $\alpha: \mathbb{R} \rightarrow \mathbb{R}$ that is strictly increasing and with $\alpha(0) = 0$; that is, extended \mathcal{K}_∞ functions are defined on the entire real line $(-\infty, \infty)$.

The CBF condition in (6) can be used to formulate a quadratic programming (QP) problem that guarantees safety by enforcing the set invariance of the safety set C defined in (2). For general control affine systems as in (1), CBF-QP is defined as

$$u_{\text{cbf}} = \arg \min_{u \in \mathbb{R}^m} \frac{1}{2} \|u - u_{\text{nom}}\|_2^2 \quad (7)$$

$$L_f h(x) + L_g h(x)u \geq -\alpha(h(x)) \quad (8)$$

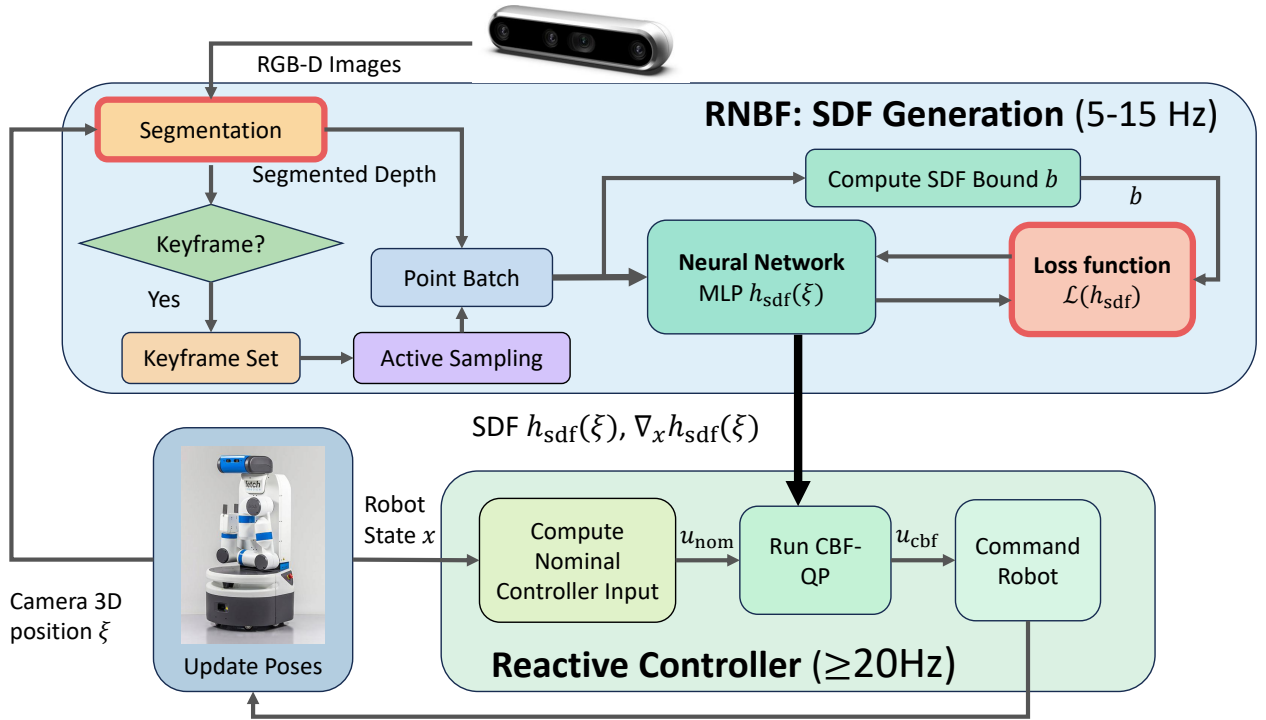


Fig. 2: Block diagram of the proposed RNBF-based safe reactive navigation pipeline. The SDF Generation module (5–15 Hz) constructs a continuous, differentiable signed distance field from real-time RGB-D input, while the Reactive Controller (≥ 20 Hz) ensures collision-free navigation using a CBF-QP. The black thick arrow highlights the main coupling step between the perception and control pipelines. Red outlines highlight components in the SDF generation stage that distinguish the RNBF pipeline from the original iSDF pipeline.

B. Signed Distance Function (SDF)

When representing three-dimensional geometry, a signed distance field $h_{\text{sdf}} : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a scalar field that maps a 3D coordinate $\xi = [p_x, p_y, p_z]^\top$ to the scalar signed distance value [20]. The surface S is the zero level-set of the field:

$$S = \{\xi \in \mathbb{R}^3 \mid h_{\text{sdf}}(\xi) = 0\} \quad (9)$$

By definition, signed distance fields (SDF) satisfy the requirement of barrier functions in Eq. (2),(3) and (4) when $n = 3$. Signed distance fields are good barrier function candidates because they retain differentiability almost everywhere (depending on the formulation), and their gradients are identical to their surface normals.

C. iSDF: Real-time SDF Reconstruction

iSDF [20] proposed a continual learning system for real-time estimates a SDF from a stream of depth images with known camera poses. The SDF is represented by a multilayer perceptron (MLP) mapping a 3D location to its signed distance. At each iteration, the system selects a small set of depth frames via active keyframe replay to mitigate catastrophic forgetting. For each selected frame, pixels and depths are sampled along back-projected rays to form a batch of 3D query points. The model is evaluated at these points and trained with a self-supervised objective; parameters are updated by backpropagation. The work reported real-time

performance and evaluates the method on real and synthetic indoor datasets [20].

IV. METHODOLOGY

A. SDF Generation Pipeline

1) *Core Methodology*: Our real-time SDF pipeline (Figure 2, “SDF Generation”) follows the iSDF [20] template but differs primarily in two major components that are critical for safe robot navigation: (i) an explicit *floor segmentation* stage (ii) a *robust near-surface loss* (Sec. IV-A.3) that replaces the L_1 term used in iSDF to attenuate RGB-D depth noise.

Given a sequence of posed RGB-D images, Our pipeline first segments the floor and masks the corresponding depth pixels to prevent the large planar ground from being treated as an obstacle, following standard RGB-D scene understanding practice [27]. The segmentation module is modular, and any modern method (e.g., SAM [28]) can be used. In our implementation, for simplicity we use a fast color-based segmenter on the RGB image and apply the resulting mask to the depth image. At the heart of our pipeline also lies a MLP, which transforms a three-dimensional point ξ into its corresponding signed distance. The network (Sec. IV-A.2) is randomly initialized and refined online from the incoming masked depth stream. We also maintain a small keyframe set, select a few frames per update via active replay, and sample query points along back-projected rays using the sampling

strategy as in iSDF [20]. The neural SDF $h_{\text{sdf}}(\xi)$ is queried at these sampled points and optimized with the objective in section IV-A.3, yielding a continuous SDF and gradients available online (5–15 Hz) to the reactive safety controller (running at ≥ 20 Hz) for safe robot navigation.

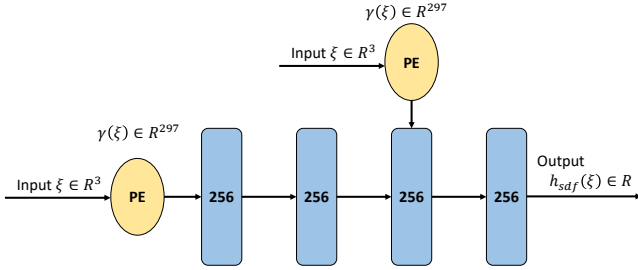


Fig. 3: Architecture of the neural SDF network $h_{\text{sdf}}(\xi)$. The model is a fully connected MLP composed of four hidden layers, each with 256 units (depicted as blue rectangles). Yellow elliptical blocks marked “PE” represent positional encodings, where input coordinates $\xi \in \mathbb{R}^3$ are transformed via a periodic embedding $\gamma(\xi) \in \mathbb{R}^{297}$ to enable high-frequency detail reconstruction. The final linear layer outputs the scalar signed distance value $h_{\text{sdf}}(\xi) \in \mathbb{R}$.

2) *Neural Architecture*: The neural SDF (Figure 3) is defined by a parameterized MLP, $h_{\text{sdf}}(\xi; \theta)$, where $\xi \in \mathbb{R}^3$ denotes the input coordinates, H denotes the hidden layers of MLP and $\theta = \{\mathbf{W}^{(i)}, \mathbf{c}^{(i)}\}_{i=1}^H \cup \{\mathbf{W}, \mathbf{c}\}$ represents the learnable weight matrices and biases. For brevity, we will write $h_{\text{sdf}}(\xi)$ instead of $h_{\text{sdf}}(\xi; \theta)$, recognizing that θ is implicitly updated during training. It comprises four hidden layers, each with 256 nodes, utilizing the Softplus activation with a high β in intermediate layers to ensure smooth gradient updates while preserving non-linearity. The final output layer remains linear to maintain the integrity of the SDF values. Finally, inspired by Mip-NeRF 360 [29] & iSDF [20], we also apply positional embeddings (PE) to the input coordinates ξ and concatenate the embedding at an intermediate layer (the third layer) to better capture high-frequency detail.

3) *Loss Function*: The loss function $\mathcal{L}(h_{\text{sdf}})$ of our framework consists of three key components: SDF loss, gradient loss, and Eikonal loss. The SDF loss supervises the network by enforcing consistency between the predicted SDF and the estimated ground truth. It consists of two terms: the free space loss and the near-surface loss. The free space loss, $L_{\text{free_space}}$, applies to points in free space and is defined as:

$$L_{\text{free_space}}(h_{\text{sdf}}(\xi), b) = \max(0, e^{-\beta h_{\text{sdf}}(\xi)} - 1, h_{\text{sdf}}(\xi) - b). \quad (10)$$

Here, b represents the estimated ground truth SDF value. Since computing exact signed distances for all surface points is computationally expensive, we employ the batch distance method [20] to approximate the ground-truth SDF efficiently while balancing performance and computational overhead.

Near the surface, the loss should be more stringent. The near-surface loss, $L_{\text{near_surf}}$, directly supervises the SDF prediction within these confines and is formulated using a huber loss [26]:

$$L_{\text{near_surf}}(h_{\text{sdf}}(\xi), b, \delta) = \begin{cases} \frac{1}{2}(h_{\text{sdf}}(\xi) - b)^2, & |h_{\text{sdf}}(\xi) - b| \leq \delta, \\ \delta(|h_{\text{sdf}}(\xi) - b| - \frac{1}{2}\delta), & |h_{\text{sdf}}(\xi) - b| > \delta. \end{cases} \quad (11)$$

Previous works like iSDF [20] primarily use L1 loss, but our experiments demonstrate that huber loss mitigates the impact of noisy depth values, improving the quality of neural SDF reconstruction. A qualitative analysis comparing huber and L1 loss under real-world noisy depth settings is presented in section V-C. The combined SDF loss is given by:

$$L_{\text{sdf}}(h_{\text{sdf}}(\xi), b) = \begin{cases} \lambda_{\text{surf}} L_{\text{near_surf}}(h_{\text{sdf}}(\xi), b, \delta), & |D[u, v] - s| < \epsilon, \\ L_{\text{free_space}}, & \text{otherwise.} \end{cases} \quad (12)$$

The second component, the gradient loss, refines the SDF prediction by supervising its gradient and ensure smooth, accurate gradient. To enforce gradient consistency, the gradient loss penalizes the cosine distance between the predicted gradient and the approximated gradient \mathbf{g} :

$$L_{\text{grad}}(h_{\text{sdf}}(\xi), \mathbf{g}) = 1 - \frac{\nabla_{\xi} h_{\text{sdf}}(\xi) \cdot \mathbf{g}}{\|\nabla_{\xi} h_{\text{sdf}}(\xi)\| \|\mathbf{g}\|}. \quad (13)$$

The third component, the Eikonal loss, ensures that the learned model satisfies the Eikonal equation, enforcing a valid distance field:

$$L_{\text{eik}}(h_{\text{sdf}}(\xi), \mathbf{g}) = \begin{cases} 0, & \|\nabla_{\xi} h_{\text{sdf}}(\xi)\| - 1 < \alpha', \\ \|\|\nabla_{\xi} h_{\text{sdf}}(\xi)\| - 1\|, & \text{otherwise.} \end{cases} \quad (14)$$

where α' is a threshold determining when the regularization is applied. The total loss function $\mathcal{L}(h_{\text{sdf}})$ is a weighted sum of all three terms:

$$\mathcal{L}(h_{\text{sdf}}) = L_{\text{sdf}} + \lambda_{\text{grad}} L_{\text{grad}} + \lambda_{\text{eik}} L_{\text{eik}}. \quad (15)$$

where λ_{grad} and λ_{eik} control the relative influence of gradient and Eikonal constraints.

4) *Continuity & Differentiability of Neural SDF*: As noted in section IV-A.2, our neural SDF $h_{\text{sdf}}(\xi)$ first embeds the input $\xi \in \mathbb{R}^3$ into a higher-dimensional space via PE, which employs smooth (infinitely differentiable) sin and cos functions [29]. We then feed $\gamma(\xi)$ into an MLP with H hidden layers. At each layer i , an affine map $\mathbf{z}_i = \mathbf{W}^{(i)} \mathbf{q}_{i-1} + \mathbf{c}^{(i)}$ is followed by a Softplus activation $\mathbf{q}_i = \text{Softplus}(\mathbf{z}_i; \beta)$, where $\text{Softplus}(z; \beta) = \frac{1}{\beta} \ln(1 + e^{\beta z})$ [30]. We set the initial layer to $\mathbf{q}_0 = \gamma(\xi)$. The final output layer remains linear to preserve the signed-distance property:

$$h_{\text{sdf}}(\xi) = \mathbf{w}^T \mathbf{q}_H + \mathbf{c}. \quad (16)$$

Because $h_{\text{sdf}}(\xi)$ is composed of continuous functions (positional embedding, affine transformations, Softplus), it

is continuous on \mathbb{R}^3 . Moreover, each component is differentiable, so by the chain rule, the gradient $\nabla h_{\text{sdf}}(\xi)$ exists and remains continuous everywhere [31]–[33]. Concretely,

$$\nabla h_{\text{sdf}}(\xi) = \frac{\partial h_{\text{sdf}}}{\partial \mathbf{q}_H} \frac{\partial \mathbf{q}_H}{\partial \mathbf{z}_H} \cdots \frac{\partial \mathbf{z}_1}{\partial \mathbf{q}_0} \frac{\partial \mathbf{q}_0}{\partial \xi}, \quad (17)$$

where, for instance, $\frac{d}{dz} \text{Softplus}(z; \beta) = \frac{\beta e^{\beta z}}{1 + e^{\beta z}}$. Hence, $h_{\text{sdf}}(\xi)$ and its gradient $\nabla h_{\text{sdf}}(\xi)$ are both continuous and differentiable, satisfy the requirements needed for standard CBF formulations.

B. RNBF Pipeline

Unlike existing work that twists standard CBF-QP formulations (Eq. (7), (8)) to accommodate noises and instability in the learned SDFs, we are able to combine SDF outputs from the vision pipeline smoothly with existing reactive safe controller, like CBF-QP, since our SDF learning process has explicitly considered sensor noises. Assuming robot state x contains robot position, i.e. $x = [\xi; \xi']$ and ξ' are robot state elements that exclude 3D position, the CBF-QP in (7), (8) can be rewritten as:

$$u_{\text{cbf}} = \arg \min_{u \in \mathbb{R}^m} \frac{1}{2} \|u - u_{\text{nom}}\|_2^2 \quad (18)$$

$$\nabla_{\xi} h_{\text{sdf}}(\xi) \dot{\xi} \geq -\alpha(h_{\text{sdf}}(\xi)). \quad (19)$$

As stated in section II, we reasonably assume that $h_{\text{sdf}}(\xi_{\text{rob}}) = h_{\text{sdf}}(\xi_{\text{cam}}) = h_{\text{sdf}}(\xi)$. It is noteworthy that while given any control affine robot dynamics, \dot{x} is guaranteed to be actuated, i.e. \dot{x} explicitly depends on control input u . However, such guarantee does not necessarily hold for $\dot{\xi}$, in which case an appropriate robot model must be chosen. Our resulted outputs from the vision pipeline can be adopted by any robot safe navigation algorithms that plan collision-free trajectories given real-time distance and the surface normal (or gradient) to the nearest obstacle.

The overall RNBF-Control pipeline can be found in Figure 2, where the SDF generation algorithms are running at a slower rate of 5-15 Hz and the reactive safe controller concurrently at a higher frequency to ensure robot safety. The vision pipeline processes RGB-D camera inputs as discussed in section IV-A. As the robot navigates, the vision pipeline continuously updates the environment's SDF, while the control pipeline receives the current SDF $h_{\text{sdf}}(\xi)$, and its gradient $\nabla h_{\text{sdf}}(\xi)$ by querying neural SDF at the current camera position ξ . These inputs are fed into the CBF-QP algorithm, which calculates the necessary control commands. These commands adjust the robot's position within the environment, facilitating safe navigation towards the target location.

V. EXPERIMENTS

A. Experiment Setup

Our proposed approach is evaluated on a Fetch robot running at 20 Hz in simulation and robot hardware. For simulations, we employ a realistic hospital environment in Gazebo, as shown in Figure 4a, and define four scenarios

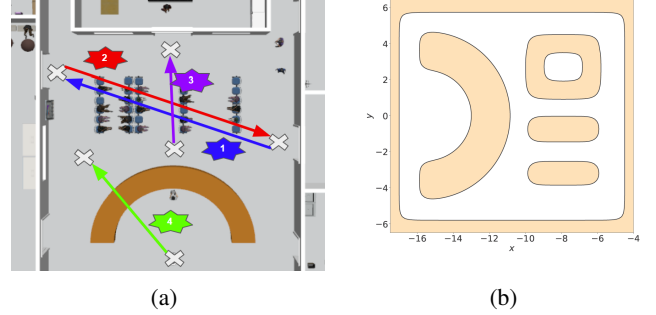


Fig. 4: (a) Simulation environment setup with four different scenarios. (b) Ground-truth SDF used by Parametric CBF-QP. Note, RNBF-CBF-QP builds its SDF online and does not use the ground-truth SDF.

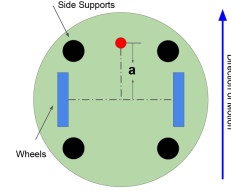


Fig. 5: Modified point of interest on Fetch robot.

with different initial and target combinations. For real-world validation, we equip the Fetch robot with an Intel RealSense D435i RGB-D camera (Figure 1) and utilize a motion capture system to acquire real-time camera poses. All experiments are run on an Intel Core i7-12700K CPU with a GeForce RTX 3090 Ti GPU.

B. Robot Dynamics

Differential drive robots like Fetch is often modeled as a standard unicycle (20).

$$\dot{x} = \begin{bmatrix} \dot{p}_x \\ \dot{p}_y \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \quad (20)$$

However, in our experiment, we choose to instead model Fetch using the shifted unicycle model by selecting a point of interest $a > 0$ m ahead of the wheel axis (see Figure 5). The unicycle model with the shifted coordinate system becomes the following.

$$\dot{x} = \begin{bmatrix} \dot{p}_x \\ \dot{p}_y \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & -a \sin \theta \\ \sin \theta & a \cos \theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \quad (21)$$

Note that for differential drive robot, $x = [p_x, p_y, \theta]^\top$, $\xi = [p_x, p_y]^\top$, and $u = [v, \omega]^\top$. Position in the z-axis is neglected for 2D navigation tasks. Our robot model decision is made due to the fact that $\dot{\xi}$ using the standard unicycle model depends on control input v while using the shifted model depends $\dot{\xi}$ on both v and ω . The CBF constraints for the shifted unicycle system can be written as

$$\frac{\partial h(p_x, p_y)}{\partial p_x} \dot{p}_x + \frac{\partial h(p_x, p_y)}{\partial p_y} \dot{p}_y \geq -\alpha(h(p_x, p_y)). \quad (22)$$






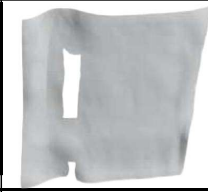

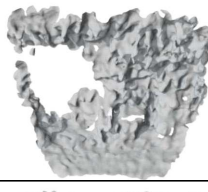
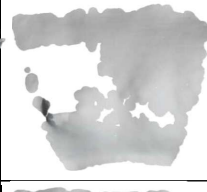
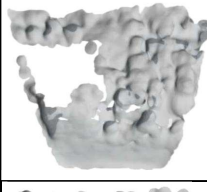
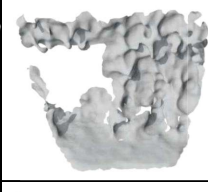
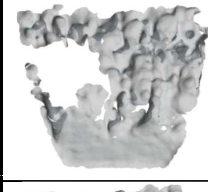
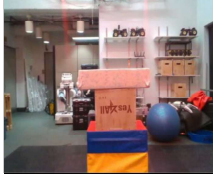
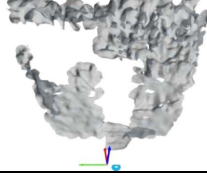
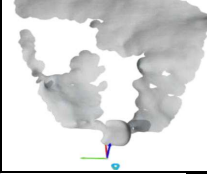
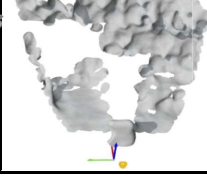
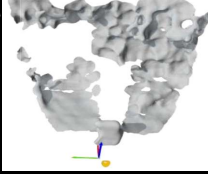
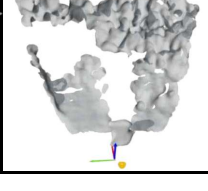
RGB IMAGE	L1	HUBER LOSS			
		$\delta = 0.01$	$\delta = 0.05$	$\delta = 0.1$	$\delta = 0.5$
					
					
					

Fig. 6: Comparison of SDF reconstructions using L1 loss & huber loss functions on depth data from an Intel RealSense D435i sensor. The leftmost column shows the RGB images of the scene. The second column presents reconstructions using the L1 loss. The remaining columns show reconstructions using the huber loss with varying δ values. Each row corresponds to a scene with a different camera-to-surface distance: in the first row, the wall is approximately 2 meters from the camera; in the second row, the racks and back wall are approximately 7 meters away; and in the third row, the closest obstacle is approximately 3 meters away.

Nominal Controller: Our nominal controllers are designed to approximately follow linear trajectories u_{nom}^l pointing from the robot’s position ξ to the target ξ^* with a speed of 0.5 m/s , where ψ is the angle difference between the current robot pose θ and the desired robot pose estimated using the orientation of u_{nom}^l , and Δt is the update timestep of the nominal controller. The approximation is not exact due to the non-holonomic property of differential drive systems.

$$u_{\text{nom}}^l = 0.5 \frac{\xi - \xi^*}{\|\xi - \xi^*\|_2} \quad v_{\text{nom}} = \|u_{\text{nom}}^l\|_2 \quad \omega_{\text{nom}} = \frac{\psi}{\Delta t} \quad (23)$$

C. Effect of Depth Noise on SDF

When deploying our pipeline on a real robot, we use a low-cost Intel RealSense D435i sensor, whose depth estimates are notably noisy. Such noises introduce spurious artifacts that would degrade the SDFs and their gradients from the neural network if left attended. Such artifacts will disqualify the SDFs from being approximate barrier functions. To mitigate this issue, we adopt a huber loss for $L_{\text{near_surf}}$ (11), which reduces the impact of outliers while still enforcing tight supervision near the surface. Figure 6 compares SDF reconstructions using RealSense depth data. In this work, sensor noise refers specifically to the depth noise inherent in affordable RGB-D sensors. The L1 loss based model amplifies sensor noise, leading to rough surfaces and unstable gradients, whereas the huber-based model produces smoother

geometry and more stable gradients. The choice of δ significantly influences reconstruction quality: smaller values of δ suppress noise more effectively but risk oversmoothing fine details, while larger values retain geometric features but may allow noise to propagate. In our real-world experiments, we find $\delta = [0.01, 0.05]$ to offer a favorable balance, though it remains a tunable parameter depending on the quality of depth sensor. Although we do not claim perfect fidelity, the result is sufficiently accurate for our pipeline. Overall, the huber loss preserves the key property of a smoothly varying SDF, essential for real-time CBF-based control under imperfect depth measurements.

D. Experiment Results

In this section, we present the experimental results that highlight the effectiveness of our *RNBF-based safety filter*. While in this work we demonstrate its use with a CBF-QP controller (*RNBF-CBF-QP*), the RNBF-based safety filter pipeline can be readily adapted to any SDF-based controller or safety filter that consumes an SDF and its gradients. RNBF-CBF-QP is compared against two baselines: (i) *Parametric CBF-QP*, which assumes full knowledge of all obstacle locations and geometries using the ground-truth SDF of the environment (Figure 4b), and (ii) *iSDF-CBF-QP*, which integrates the original iSDF reconstruction into a CBF-QP without floor segmentation or depth-denoising. We display in Figure 7 the trajectories generated by RNBF-CBF-QP, the parametric CBF-QP, iSDF-CBF-QP and the nominal

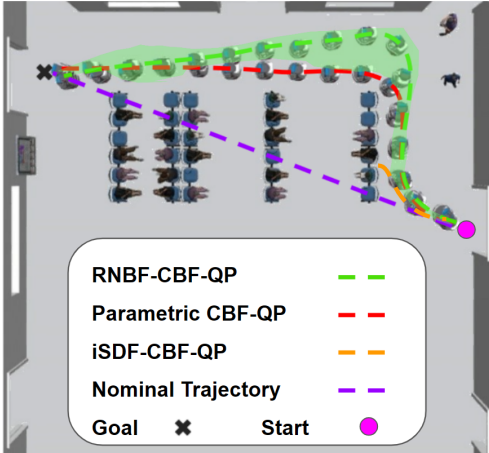


Fig. 7: **Scenario 1 (simulation): Trajectory comparison.** Dashed curves represent trajectories, with mean values shown where applicable. For RNBF-CBF-QP, a shaded tube depicts the variance of the mean trajectory over 10 rollouts. Parametric CBF-QP, which assumes perfect obstacle knowledge, produces nearly identical paths across runs; only one representative trajectory is plotted for clarity. iSDF-CBF-QP is likewise shown with a single representative run since it collided in all trials. The nominal controller is deterministic and follows the shortest path from start to goal.

controller for scenario 1.

Methods	Safe %	Reached %	Duration (s)
RNBF-CBF-QP	100	100	36.7
Parametric CBF-QP	100	100	29.0
iSDF-CBF-QP	0	0	-

TABLE I: Performance comparison of RNBF-CBF-QP with Parametric CBF-QP & iSDF-CBF-QP in Scenario 1 over 10 independent trials.

Table I reports quantitative results for ten independent runs of Scenario 1 with RNBF-CBF-QP, Parametric CBF-QP, and iSDF-CBF-QP. Performance is measured by the average traversal time from start to goal (*duration*, s), the proportion of collision-free trajectories (*safe %*), and the proportion of successful goal reaches (*reached %*). Both RNBF-CBF-QP and the Parametric CBF-QP achieved a 100% collision-free rate and a 100% success rate in reaching the destination. However, RNBF-CBF-QP requires slightly more time to reach the goal (36.7s vs. 29.0s) owing to its exclusive reliance on online RGB-D sensing to construct barrier functions, as opposed to the Parametric CBF-QP’s use of full prior knowledge of obstacle locations. In contrast, the iSDF-CBF-QP fails in all trials, with 0% success rate in reaching the destination and repeated collisions with the obstacles. Two factors primarily drive this failure: (1) *Floor misclassification*: without a segmentation module, large planar floor regions are incorrectly treated as obstacles, preventing the controller from finding a feasible path; this issue arises in both simulation and real-world deployments. (2) *Depth-noise artifacts*: sensor noise introduces spurious structures into the



Fig. 8: **Real-world experiments.** Shown are collision-free RNBF-CBF-QP trajectories overlaid on video frames for static (single obstacle - shape 1 & 2, and multi obstacle) and quasi-static obstacle scenarios.

neural SDF (section V-C), yielding unstable gradients and erratic control commands; while less relevant in simulation due to noise-free gazebo depth sensing, it significantly impacts physical robot navigation (Figure 9). RNBF-CBF-QP mitigates both issues through explicit floor segmentation and huber-loss based near-surface supervision, enabling stable gradients and reliable, collision-free navigation.

Beyond gazebo simulations, the proposed pipeline is further validated in real-world experiments involving both static and quasi-static obstacles (Figure 1 & Figure 8). For both types of obstacles, RNBF-CBF-QP reached the goal collision-free under partial observability and depth noise. In contrast, iSDF-CBF-QP exhibited the same failure modes as in simulation leading to collisions and failure to reach the goal as shown in Figure 9. We provide supplementary videos, which reveal that our method is capable of accurately *building and updating the barrier function on-the-fly*, thereby ensuring safety in novel environments.

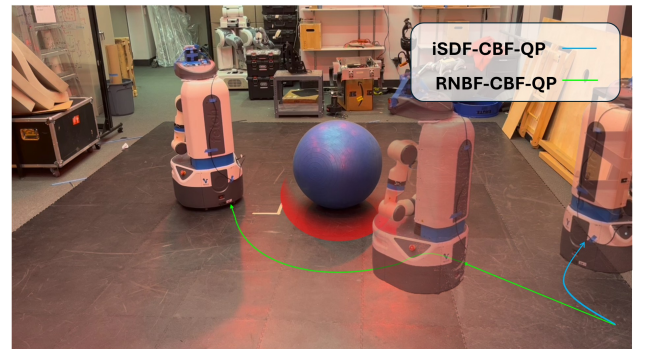


Fig. 9: Trajectory comparison of RNBF-CBF-QP vs iSDF-CBF-QP for real-world deployment with static obstacle.

VI. CONCLUSIONS

The paper introduces a novel pipeline for real-time CBF construction in unknown environments using affordable sensors like RGB-D cameras. Our framework is robust to sensor noise while generating a smooth and differentiable neural SDF representation, thereby preserving the core assumptions required by standard CBFs. Experiments in both simulation

and on a physical Fetch robot confirm that our approach enables safe, reactive navigation and can readily integrate with established SDF-based controllers. Moreover, the system's modular design allows it to serve as a dedicated safety layer in broader robotic applications, given only onboard RGB-D sensing.

Although our method demonstrates robust performance in static and quasi-static environments, it faces limitations in fully dynamic settings. Currently, our RNBF pipeline lacks a "forgetting" mechanism, causing outdated obstacle data to persist in the reconstructed map and reducing adaptability to rapidly changing scenarios. As part of future work, we plan to incorporate real-time memory management to discard stale obstacle information, as well as integrate temporal reasoning and motion prediction to further enhance performance in dynamic environments.

REFERENCES

- [1] G.-Z. Yang, J. Bellingham, P. E. Dupont, P. Fischer, L. Floridi, R. Full, N. Jacobstein, V. Kumar, M. McNutt, R. Merrifield *et al.*, "The grand challenges of science robotics," *Science robotics*, vol. 3, no. 14, p. eaar7650, 2018.
- [2] A. D. Ames, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs with application to adaptive cruise control," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 6271–6278.
- [3] Y. Xue and N. Figueroa, "No minima, no collisions: Combining modulation and control barrier function strategies for feasible dynamical collision avoidance," 2025. [Online]. Available: <https://arxiv.org/abs/2502.14238>
- [4] M. Khan, M. Zafar, and A. Chatterjee, "Barrier functions in cascaded controller: Safe quadrotor control," in *2020 American Control Conference (ACC)*, 2020, pp. 1737–1742.
- [5] R. Grandia, A. J. Taylor, A. D. Ames, and M. Hutter, "Multi-layered safety for legged robots via control barrier functions and model predictive control," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 8352–8358.
- [6] Y. Zhang, G. Tian, L. Wen, X. Yao, L. Zhang, Z. Bing, W. He, and A. Knoll, "Online efficient safety-critical control for mobile robots in unknown dynamic multi-obstacle environments," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 12370–12377.
- [7] C. Dawson, B. Lowenkamp, D. Goff, and C. Fan, "Learning safe, generalizable perception-based hybrid control with certificates," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1904–1911, 2022.
- [8] K. Long, C. Qian, J. Cortés, and N. Atanasov, "Learning barrier functions with memory for robust safe navigation," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4931–4938, 2021.
- [9] M. Harms, M. Kulkarni, N. Khedekar, M. Jacquet, and K. Alexis, "Neural control barrier functions for safe navigation," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 10415–10422.
- [10] H. Abdi, G. Raja, and R. Ghabcheloo, "Safe control using vision-based control barrier function (v-cbf)," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 782–788.
- [11] M. Tong, C. Dawson, and C. Fan, "Enforcing safety for vision-based controllers via control barrier functions and neural radiance fields," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 10511–10517.
- [12] M. De Sa, P. Kotaru, and K. Sreenath, "Point cloud-based control barrier function regression for safe and efficient vision-based control," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 366–372.
- [13] M. Zucker, N. Ratliff, A. D. Dragan, M. Pivtoraiko, M. Klingensmith, C. M. Dellin, J. A. Bagnell, and S. S. Srinivasa, "Chomp: Covariant hamiltonian optimization for motion planning," *The International journal of robotics research*, vol. 32, no. 9-10, pp. 1164–1193, 2013.
- [14] G. Notomista and M. Saveriano, "Safety of dynamical systems with multiple non-convex unsafe sets using control barrier functions," *IEEE Control Systems Letters*, vol. 6, pp. 1136–1141, 2022.
- [15] L. Huber, A. Billard, and J.-J. Slotine, "Avoidance of convex and concave obstacles with convergence ensured through contraction," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1462–1469, 2019.
- [16] C. K. Fourie, N. Figueroa, and J. A. Shah, "On-manifold strategies for reactive dynamical system modulation with nonconvex obstacles," *IEEE Transactions on Robotics*, vol. 40, pp. 2390–2409, 2024.
- [17] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.
- [18] P. Liu, K. Zhang, D. Tateo, S. Jauhri, J. Peters, and G. Chalkatzaki, "Regularized deep signed distance fields for reactive motion generation," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 6673–6680.
- [19] Z. Tang, B. Sundaralingam, J. Tremblay, B. Wen, Y. Yuan, S. Tyree, C. Loop, A. Schwing, and S. Birchfield, "Rgb-only reconstruction of tabletop scenes for collision-free manipulator control," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 1778–1785.
- [20] J. Ortiz, A. Clegg, J. Dong, E. Sucar, D. Novotny, M. Zollhoefer, and M. Mukadam, "isdf: Real-time neural signed distance fields for robot perception," *arXiv preprint arXiv:2204.02296*, 2022.
- [21] S. T. Bukhari, D. Lawson, and A. H. Qureshi, "Differentiable composite neural signed distance fields for robot navigation in dynamic indoor environments," *arXiv preprint arXiv:2502.02664*, 2025.
- [22] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *2019 18th European control conference (ECC)*. Ieee, 2019, pp. 3420–3431.
- [23] Y. Xue and N. Figueroa, "No minima, no collisions: Combining modulation and control barrier function strategies for feasible dynamical collision avoidance," *arXiv preprint arXiv:2502.14238*, 2025.
- [24] C. K. Fourie, N. Figueroa, and J. A. Shah, "On-manifold strategies for reactive dynamical system modulation with nonconvex obstacles," *IEEE Transactions on Robotics*, vol. 40, pp. 2390–2409, 2024.
- [25] M. Koptev, N. Figueroa, and A. Billard, "Reactive collision-free motion generation in joint space via dynamical systems and sampling-based mpc," *The International Journal of Robotics Research*, vol. 43, no. 13, pp. 2049–2069, 2024.
- [26] P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in statistics: Methodology and distribution*. Springer, 1992, pp. 492–518.
- [27] S. Gupta, P. Arbelaez, and J. Malik, "Perceptual organization of points in rgb-d images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 2–16, 2013.
- [28] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo *et al.*, "Segment anything," *arXiv preprint arXiv:2304.02643*, 2023.
- [29] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, "Mip-nerf 360: Unbounded anti-aliased neural radiance fields," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5470–5479.
- [30] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2011, pp. 315–323.
- [31] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [32] J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, "Deepsdf: Learning continuous signed distance functions for shape representation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 165–174.
- [33] M. Tancik, P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. Barron, and R. Ng, "Fourier features let networks learn high frequency functions in low dimensional domains," *Advances in neural information processing systems*, vol. 33, pp. 7537–7547, 2020.