学习轻推：一种用于密集杂物环境中安全机器人交互的可扩展屏障函数框架

Haixin Jin*, 1, Nikhil Uday Shinde*, 1, Soofiyan Atar[1],Hongzhan Yu[1], Dylan Hirsch[1],Sicun Gao[1], Michael C.Yip[1],Sylvia Herbert[11]加利福尼亚大学圣迭戈分校{haj013, nshinde, satar, hoy021, dhirsch, sig049, yip, sherbert}@ucsd.edu

arXiv:2601.02686v1 [cs.RO] 6 Jan 2026

摘要——**在日常环境中运行的机器人必须在杂乱拥挤的空间内导航和操作，而与周围物体发生物理接触不可避免。传统安全框架将接触视为不安全，因此限制机器人只能进行碰撞规避，妨碍其在密集的日常环境中发挥作用。随着物体数量增加，基于模型的安全操作方法在计算上变得难以处理；与此同时，学习方法通常将安全性与当前任务绑定，因此难以迁移到新任务，除非重新训练。本文提出密集接触屏障函数（DCBF）。我们的方法不再显式建模多物体动力学，而是学习一个可组合、以物体为中心的函数，隐式捕捉物理交互产生的安全约束，从而绕过显式建模的计算复杂性。通过少量物体的交互离线训练后，学习得到的 DCBF 可在运行时针对任意物体集合进行组合，生成单一的全局安全过滤器，其计算规模随物体数量线性增长，并且无需重新训练即可迁移到不同任务。我们通过密集杂物环境中的仿真实验验证了该方法，展示了它在适宜环境中实现无碰撞导航以及安全、接触丰富的交互的能力。**

I. 引言

机器人能够在密集杂乱环境中安全地进行交互并穿行，是推动其广泛应用于日常非结构化空间的关键一步。此类能力可使家用机器人安全地穿行于拥挤的食品储藏室，也能让仓储机器人在狭窄、多人共用的商业场所补充货架。这些环境中物体密集且动态排列，机器人往往必须进行物理接触才能完成任务。例如，机器人从堆满物品的食品储藏室中取出某件物品时，可能需要轻轻推开相邻物品；而物流机器人则可能需要接触易碎物品，同时避免将其撞倒。

机器人安全性研究历来侧重于碰撞规避。以 Hamilton-Jacobi 可达性（HJR） [1]、控制屏障函数（CBFs） [2]和人工势场（APF） [3] 等经典方法为代表的大多数既有研究，都以与所有障碍物保持一定的最小距离为前提。即使是专注于复杂操作的研究，也往往通过将被操作的物体纳入必须避免碰撞的物体集合来处理安全问题 [4], [5]。然而，在杂乱环境中，这种碰撞规避准则存在根本性的局限。若家用机器人一味避免碰撞，那么当它需要触及食品储藏室中第一排物品之后的任何物体时，就会束手无策。相比之下，人类会经常接触、轻推和移动物体，同时本能地理解如何通过互动来防止物品洒落或损坏。本研究旨在弥合这一差距，使机器人能够推理并执行安全接触。

尽管已有一些研究探索了物理交互，但这些方法往往未能解决密集多物体场景所需的可扩展性和泛化能力问题。显式建模机器人与物体之间复杂耦合动态的方法 [6]会随着物体数量增加而变得计算上难以处理，因为潜在交互的组合数量会急剧膨胀。诸如 [6], [7]的研究成功展示了与单个复杂可变形物体进行安全接触，但未能明确指出如何扩展到多物体系统。强化学习 [8], [9],等学习方法能够处理复杂系统，但往往将安全性与性能纠缠在一起，因此每遇到新任务都需要从头训练。这凸显了对一种方法的需求：它能在这些交互过程中兼顾性能与安全性，适用于数量众多且可变的物体，并能轻松泛化到各种不同的排列方式。

本文提出密集接触屏障函数（DCBF），这是一种以物体为中心的控制屏障函数，可计算隐式交互，从而在密集物体环境中实现安全且可扩展的交互。我们框架的核心是一个可组合 DCBF，它通过少量物体参与的交互数据进行离线训练。在运行时，该函数会针对场景中的每个物体进行评估，并将结果汇总为单个安全过滤器。这种组合式方法使我们的方法能够扩展到包含任意数量物体的环境，并且无需重新训练即可泛化到多种不同配置。我们的方法通过以物体为中心、可扩展的方式学习隐式捕捉复杂交互产生的安全约束，避免了显式多物体交互建模与推理的棘手难题。我们的主要贡献如下：

- 一种隐式交互控制屏障函数（CBF）表述，无需显式建模即可捕捉多物体交互对安全性的影响。
- 一种可组合、以物体为中心的框架，可针对数量可变的物体实现可扩展的安全推理。

*这些作者对本研究的贡献相同。

- 一种高交互式细化流程，用于降低学习得到的控制屏障函数的保守性
- 在密集杂乱场景中进行安全交互的仿真实验验证。

II. 相关工作

安全自主系统的开发是一个至关重要的研究领域，尤其针对在家庭或商业设施等复杂动态环境中运行的机器人。尽管已有大量研究致力于确保机器人安全，但其中大部分都集中于碰撞规避。这一范式虽然在开放空间中有效，但在机器人必须进行物理接触并与物体交互以完成任务的密集环境中却过于严格。我们的工作通过提出一种基于可组合屏障函数和隐式交互建模的安全机器人—物体交互可扩展方法，弥补了这一不足。

碰撞规避的安全自主性： 传统上，机器人安全一直等同于避免碰撞。Hamilton-Jacobi 可达性（HJR）等经典方法计算安全价值函数，用于定义控制不变安全集 [1]。这些函数已被在线用于最优控制以及作为最低侵入性的安全滤波器 [10]。然而，HJR受到维数灾难的影响，并且需要解析模型，因此其应用受限于低维控制仿射系统，进而只能实现较为简单的基于碰撞规避的安全性。近期的研究（如 [11], [12]）通过基于学习的近似扩展了HJR，缓解了维度限制，但仍依赖已知的解析模型，因而难以扩展到具有未知动态物体交互的环境中。

即使在以物体交互为目标的操作任务中，对碰撞规避的关注仍然存在 [13]。人工势场等经典方法 [3], [14] 以及基于模型预测控制的规划器 [15] ，一直致力于与障碍物保持距离并尽量减少接触。近期在杂乱环境中的操作 [5] 和可变形物体操作 [4]方面的研究，通过将被操作物体视为系统的一部分，扩展了安全碰撞规避的概念，从而确保物体也能避免与环境发生碰撞。对于拥挤的环境，这些方法仍然过于保守，因为避免任何接触会导致许多任务无法完成。

可扩展性与泛化方面的挑战： 能够探索更复杂系统乃至物理交互的方法，往往难以扩展至密集物体环境。强化学习方法已被用于处理复杂动力学，例如将安全性作为软约束 [16], [17]，或采用受贝尔曼启发的更新 [8], [9]。然而，这些方法通常将安全性与特定任务策略纠缠在一起，限制了它们作为通用、任务无关安全过滤器的应用。其他方法则利用 XPBD [18] 和 IsaacLab [19] 等仿真技术的进步，为复杂系统实现规划。例如， [6] 将 XPBD 与基于模型的规划结合，在满足安全约束的同时操控可变形物体，以实现主动感知。其他方法，如 [7] ，则利用学习到的潜空间动力学，在可变形物体操作任务中推理更复杂的安全定义。像 [20]这样的方法甚至会结合额外的感知模态（例如力反馈），以安全地操控单个关节物体。虽然这些方法在单物体场景中效果显著，但计算开销仍然很大，或难以泛化至多物体的杂乱环境。关键需求是开发一种可扩展的方法，能够针对数量可变的可移动物体进行推理，同时不过度保守。

控制屏障函数： 为应对上述挑战，我们以控制屏障函数（CBF）框架为基础。控制屏障函数是在系统状态空间中定义控制不变安全集的函数，可保证存在一种控制动作，使智能体始终处于该集合中 [2], [21]。控制屏障函数已被广泛用于通过侵入性最小的安全过滤器在线保障安全；鲁棒控制屏障函数（RCBF） [22]等扩展方法则将其推广至不确定动力学情形。

经典控制屏障函数面临的一项关键挑战是其综合过程 [23]，因为这通常需要精确的系统知识。这一局限历来将其应用范围限制在碰撞规避等较简单的任务上。尽管可以采用基于HJR的综合方法 [10], ，但其可扩展性较差。为克服这一问题，神经控制屏障函数通过直接从数据中学习安全约束 [24], ，开创了一类新的方法：纳入感知 [25], 、处理不确定性 [26], ，并扩展至密集环境 [27]。其他研究则在此基础上继续发展，利用HJR改进候选神经控制屏障函数 [28]。例如， [27] 等研究利用可组合控制屏障函数提升了密集环境下安全保障的可扩展性，从而实现人群中的安全导航。尽管取得了这些进展，控制屏障函数的应用仍主要聚焦于碰撞规避，能够管理安全物理交互的方法仍存在关键缺口。

在这项工作中，我们研究密集的多物体环境中的安全交互问题，在此类环境中，物理接触不可避免。我们通过引入密集接触屏障函数，将控制屏障函数的应用从碰撞规避扩展至逐物体安全性，并以可组合的方式进行推理。这使得我们能够构建可扩展且具有泛化能力的安全过滤器，在不牺牲交互和完成任务能力的前提下，处理配置各异的多个物体。

III. 预备知识

A. 问题陈述

我们研究机器人如何在物体密集分布的环境中安全运行，例如家庭储藏室，或商店、仓库的货架。机器人可能会与物体发生物理交互，但必须确保其动作不会导致物体进入不安全状态。尽管我们的表述可推广至各种基于状态的安全约束，本文重点关注物体倾斜超过规定倾斜角这一失效条件。该问题描述涵盖了多种相关任务，例如机械臂伸向食品储藏室深处的玻璃杯、地面机器人推开家具，或仓储机器人重新整理货架。在每种情况下，安全性

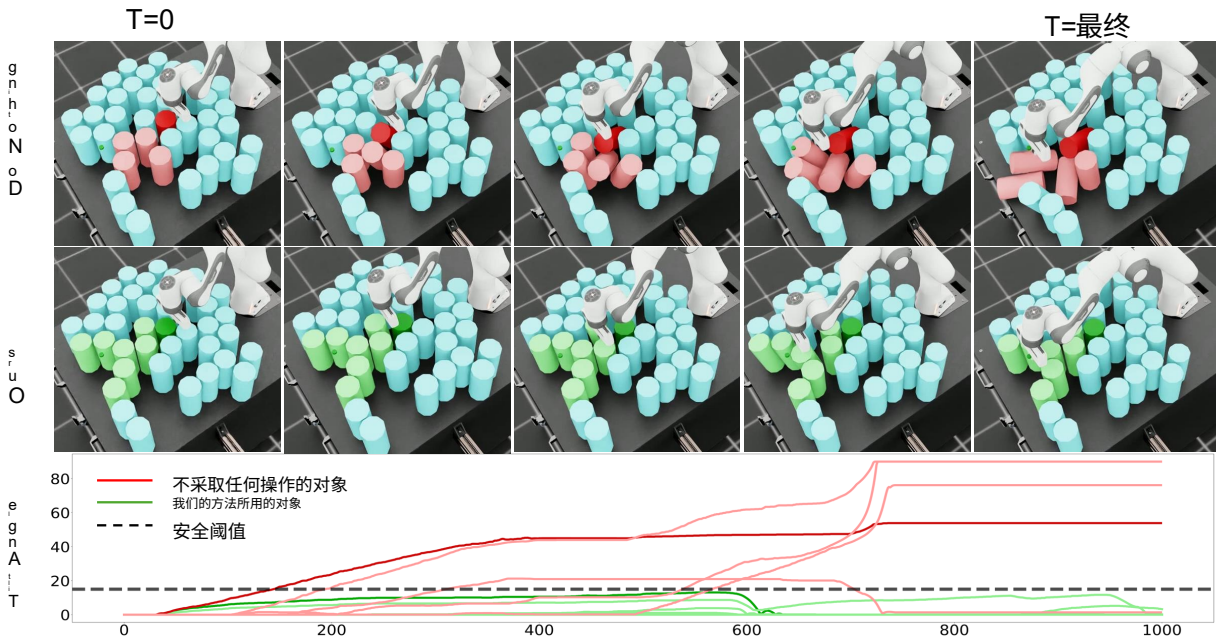

图1：不同方法下模拟机器人从起点到目标点的轨迹。第一行展示了什么都不做基线方法，机器人直接推过杂乱环境，导致多个物体被撞倒。发生交互的物体以红色突出显示，其中最靠近末端执行器的起始物体显示为深红色，其他物体显示为浅红色。第二行展示了我们提出的方法，该方法能够安全地与物体交互，在不违反安全性的情况下到达目标点。发生交互的物体以绿色突出显示，其中最靠近末端执行器的起始物体显示为深绿色，其他物体显示为浅绿色。第三行绘制了发生交互的物体随时间变化的倾斜角，曲线采用相同的颜色编码（深/浅红色和深/浅绿色），虚线表示15°的安全阈值。红色曲线超过倾斜角阈值，表明存在安全违规；而绿色曲线始终低于该阈值，证明了安全交互。

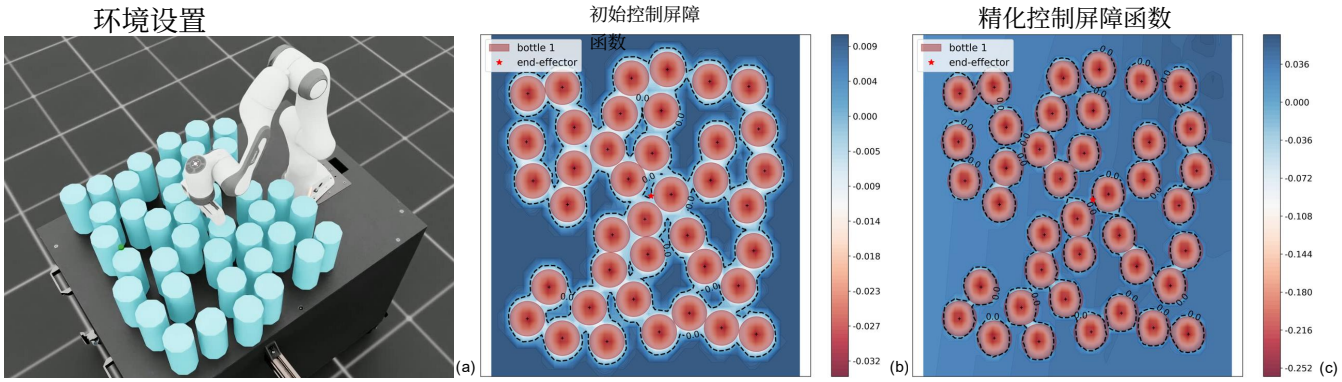

图2：初始模型与精化模型（$\sigma = 0.01$）的安全边界对比：(a) 发生接触时的环境快照。末端执行器与一个瓶子发生轻微接触，使其倾斜了一个小角度，同时保持在安全区域内。(b) 接触情况下初始模型的全局控制屏障函数值图。保守边界错误地将该交互视为不安全，常常导致机器人停滞。(c) 接触情况下精化模型的全局控制屏障函数值图。该边界能够适应安全交互，体现出针对杂乱环境中安全性的更少保守且更符合实际的推理

都以物体稳定性来定义，同时机器人必须能够有效地执行任务。

我们首先对这个问题进行形式化定义。将机器人的状态定义为 $\mathbf{r} \in \mathbb{R}^n$ ，其动力学为

$$\mathbf{r}^{t+1} = f(\mathbf{r}^t, u^t), \tag{1}$$

其中 $u^t \in \mathbb{R}^a$ 为控制输入。环境中包含N个物体，每个物体的状态为 $\mathbf{o}_i \in \mathbb{R}^m \ \forall i \in [1, N]$。每个物体的安全性由函数 $q_i(\mathbf{o}_i)$ 定义；当该安全代价超过已知阈值时，即发生违规，也就是 $q_i(\mathbf{o}_i) \geq q_T$。

对于仅包含一个物体 $\mathbf{o}$ 的场景，其动力学由以下公式给出

：$\mathbf{o}^{(t+1)} = p(\mathbf{r}^t, u^{(t)}, \mathbf{o}^{(t)})$。然而，对于包含多个物体的场景，动力学由

$$\{\mathbf{o}_1^{(t+1)}, \mathbf{o}_2^{(t+1)}, \ldots, \mathbf{o}_N^{(t+1)}\} = F(\mathbf{r}^{(t)}, u^{(t)}, \{\mathbf{o}_1^{(t)}, \mathbf{o}_2^{(t)}, \ldots, \mathbf{o}_N^{(t)}\}), \tag{2}$$

受到机器人和物体间交互的共同影响。在 F 中对交互进行推理的一项关键挑战是，物体间交互的数量会随物体数量呈指数级增长。通过显式建模来推理每个物体的安全性，很快就会变得难以实现。此外，密集环境中的接触物理模型通常计算效率过低，无法实际使用。本文假设可以访问机器人的前向动力学模型 (1)，但无法在运行时访问能够对物体进行前向仿真的物理模型 (2)。

B. 控制屏障函数

考虑一个具有动力学 $\mathbf{x}^{t+1} = f(\mathbf{x}^t, u^t)$ 和状态空间 $\mathbf{x} \in \mathbb{R}^n$ 的离散时间系统。我们称集合 $\mathcal{C} \subseteq \mathbb{R}^n$ 为控制不变 集，如果对于任意 $\mathbf{x} \in \mathbb{R}^n$，都存在一个动作 u，使后续状态保持在该集合中，即 $f(\mathbf{x}\ u) \in \mathcal{C}$。

此外，如果对于每个状态 $\mathbf{x} \in \mathbb{R}^n$，条件成立，我们称函数 h: $\mathbb{R}^n \to \mathbb{R}$ 为（离散时间）控制障碍函数 (CBF)：

$$h(f(\mathbf{x}, u)) - h(\mathbf{x}) \geq -\gamma(h(\mathbf{x}))$$

成立，其中 $\gamma : \mathbb{R} \to \mathbb{R}$ 是任意扩展类 $\mathcal{K}$ 函数。

我们注意到，对于控制屏障函数 h，其上水平集 $\mathcal{C}_h := \{\mathbf{x} \in \mathbb{R}^n \mid h(\mathbf{x}) \geq 0\}$ 是控制不变的。

IV. 密集接触屏障函数:

我们的目标是将物体间接触动力学对每个物体的影响隐式编码到学习得到的控制障碍函数 $B(\mathbf{r}_i^{t+1}, \mathbf{O}_i^t)$ 的结构中，其中 $\mathbf{r}_i^{t+1}$ 和 $\mathbf{O}_i^t$ 分别是以物体为中心的机器人状态，以及与物体 i相对应的机器人与物体交互的以物体为中心的历史记录。由此将得到一个任务无关的安全过滤器，使机器人能够在穿过环境时与物体安全接触。本节将介绍学习得到的控制屏障函数模型中的三个关键细节：1）隐式交互，即将物体间交互隐式编码到控制屏障函数中；2）状态历史，它有助于提升对交互动力学的隐式认知；3）以物体为中心的状态表示，它支持组合，并能推广到密度各异的环境。

A. 隐式交互

为了说明机器人-物体交互如何隐式编码在 B 中，我们首先考虑单物体环境的情况，其中机器人动力学为 (1)，物体动力学为 $\mathbf{o}^{t+1} = p(\mathbf{o}^t, \mathbf{r}^t, u^t)$。由于接触动力学难以进行解析建模，且模拟计算效率低下，我们需要学习一个数据驱动的标称接触模型 $\mathbf{o}^{t+1} \approx \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)$。换句话说，

$$p(\mathbf{o}^t, \mathbf{r}^t, u^t) = \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t) + e(\mathbf{o}^t, \mathbf{r}^t, u^t). \tag{3}$$

其中， e 表示真实模型与学习得到的模型之间的误差。

根据 [22], ，我们称函数 $h(\mathbf{r}^t, \mathbf{o}^t)$ 是关于学习到的动力学 $\hat{p}$ 的鲁棒控制屏障函数，若存在某个 $\sigma > 0$，使得不等式

$$\sup_{u^t \in \mathbb{R}^m} h\big(f(\mathbf{r}^t, u^t), \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) - h(\mathbf{r}^t, \mathbf{o}^t) - \sigma \geq -\gamma(h(\mathbf{r}^t, \mathbf{o}^t))$$

对每个 $\mathbf{r}^t$ 和 $\mathbf{o}^t$ 均成立。参数 σ 称为鲁棒性裕度。相应的鲁棒可行控制集 为:

$$K_{\text{RCBF}}(\mathbf{r}^t, \mathbf{o}^t) = \Big\{u \in \mathbb{R}^m \mid h\big(f(\mathbf{r}^t, u^t), \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) - h(\mathbf{r}^t, \mathbf{o}^t) - \sigma \geq -\gamma(h(\mathbf{r}^t, \mathbf{o}^t))\Big\}. \tag{4}$$

为理解鲁棒控制屏障函数的作用，请看以下内容。假设我们有一个鲁棒控制屏障函数 h ，且模型失配的上界不超过鲁棒性裕度，即

$$h\big(f(\mathbf{r}^t, u^t), \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) - h\big(f(\mathbf{r}^t, u^t), p(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) \leq \sigma$$

对于所有 $\mathbf{r}^t$、$\mathbf{o}^t$ 和 u^t 均成立。因此可得

$$\begin{aligned}
&\sup_{u^t \in \mathbb{R}^m} h\big(f(\mathbf{r}^t, u^t), p(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) - h(\mathbf{r}^t, \mathbf{o}^t) \\
&= \sup_{u^t \in \mathbb{R}^m} \Bigg[\underbrace{h\big(f(\mathbf{r}^t, u^t), \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) - h(\mathbf{r}^t, \mathbf{o}^t)}_{\text{nominal}} \\
&\quad - \underbrace{\Big(h\big(f(\mathbf{r}^t, u^t), \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) - h\big(f(\mathbf{r}^t, u^t), p(\mathbf{o}^t, \mathbf{r}^t, u^t)\big)\Big)}_{\text{mismatch}}\Bigg] \\
&\geq \sup_{u^t \in \mathbb{R}^m} h\big(f(\mathbf{r}^t, u^t), p(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) - h(\mathbf{r}^t, \mathbf{o}^t) - \sigma \\
&\geq -\gamma(h(\mathbf{r}^t, \mathbf{o}^t)).
\end{aligned}$$

因此，如果 h 是关于学习到的动力学的鲁棒控制屏障函数，且模型失配的上界不超过鲁棒性裕度，那么 h 也是关于真实动力学的控制屏障函数。特别是，任何控制器 $u = k(\mathbf{r}^t, \mathbf{o}^t) \in K_{\text{RCBF}}(\mathbf{r}^t, \mathbf{o}^t)$ 都能保证安全性，不仅是在学习到的动力学下，也同样是在真实动力学下。

正因如此，我们将寻求同时学习物体动力学和一个鲁棒控制屏障函数（不过正如稍后所述，我们将以隐式方式完成这一点）。此外，根据均值定理，对于某个 $\mathbf{o}^*$ ，其位于 $\hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)$ 与 $p(\mathbf{o}^t, \mathbf{r}^t, u^t)$ 之间，有 $h(f(\mathbf{r}^t, u^t), \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)) - h\big(f(\mathbf{r}^t, u^t)\ p(\mathbf{o}^t, \mathbf{r}^t, u^t)\big) = -\frac{\partial h}{\partial \mathbf{o}}(f(\mathbf{r}^t, u^t), \mathbf{o}^*) e(\mathbf{o}^t, \mathbf{r}^t, u^t)$。给定控制屏障函数值 h 随物体位置变化的利普希茨界 L_o ，以及误差界 $\varepsilon \geq \|e\|$，一个足够的鲁棒性裕度为

$$\begin{aligned}
\sigma^* := L_0 \varepsilon &\geq -\frac{\partial h}{\partial \mathbf{o}}(f(\mathbf{r}^t, u^t), \mathbf{o}^*) e(\mathbf{o}^t, \mathbf{r}^t, u^t) \\
&\geq h(f(\mathbf{r}^t, u^t), \hat{p}(\mathbf{o}^t, \mathbf{r}^t, u^t)) - h\big(f(\mathbf{r}^t, u^t), p(\mathbf{o}^t, \mathbf{r}^t, u^t)\big).
\end{aligned}$$

随着模型$\hat{p}$ 不断改进， $\|e\|$ 也随之按比例缩小。概念上，这表明随着训练数据变得更能代表可能发生的交互，使得

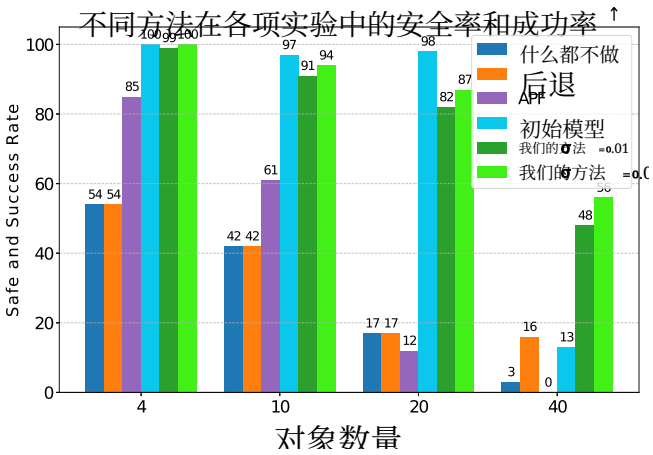

图3：在物体密度逐渐增加（4、10、20 和 40 个物体）的不同环境中评估不同方法的安全率和成功率。该指标展示了机器人既到达目标，又在整个轨迹过程中始终保持安全的轨迹所占的百分比。每种方法在每个环境中的结果均基于 100 条随机生成的轨迹计算。结果凸显了基线方法的局限性，尤其是在物体密度较高时；同时也体现了我们的方法在兼顾安全性与任务成功率方面的鲁棒性。

$\hat{}p$ 能够更好地捕捉动力学，这一充分的鲁棒性裕度 σ^* 会减小。

最后，我们不再显式地分别学习控制屏障函数 h 和动力学p，而是提出通过学习函数$B(\mathbf{r}^{t+1},\mathbf{o}^t) \approx h(\mathbf{r}^{t+1},\mathbf{o}^{t+1})$，以隐式且同步的方式完成学习。在实际运行中，函数$B$ 可按如下方式使用。假设我们有一个机器人-物体状态 $(\mathbf{r}^t,\mathbf{o}^t)$。在选择动作 u^t之前，我们利用机器人模型确定执行该动作后的机器人状态，即计算 $\mathbf{r}^{t+1} = f(\mathbf{r}^t, u^t)$。接下来，查询 $B(\mathbf{r}^{t+1},\mathbf{o}^t) \approx h(\mathbf{r}^{t+1}, \mathbf{o}^{t+1})$ 的值，以确定所选动作的安全值。如果对应的机器人-物体状态是安全的，即 $B(\mathbf{r}^{t+1}, \mathbf{o}^t) \geq 0$，我们就可以安全地执行所选动作。

B. 物体状态历史

受 [27]，启发，我们修改了神经控制屏障函数模型，使其不仅接收单步物体状态$\mathbf{o}^t$，还接收过去 T 个时间步的物体状态历史：$\mathbf{O}^t_{\mathbf{abs}} = [\mathbf{o}^t, \mathbf{o}^{t-1}, \ldots\ldots, \mathbf{o}^{t-(T-1)}]$。这样做是为了更好地编码潜在交互效应（机器人-物体和物体-物体）以及未观测属性（例如摩擦力、物体质量）。

C. 以物体为中心的表示

隐式交互建模也很容易推广到多个物体。然而，由于组合复杂性以及高得令人望而却步的数据需求，对多物体联合交互进行建模可能难以处理。相反，受 [29]，[27] 启发，我们采用以物体为中心的方法，通过改变输入表示来学习一个可扩展到任意数量物体的单一模型。

我们方法的核心是一个单一的隐式交互控制障碍函数（CBF），B，它从每个物体的独立视角推理安全性。我们通过将系统状态转换到每个物体对应的相对坐标系中来实现这一点。给定将状态 a 映射到 b 的相对坐标的变换 $R(a,b)$，我们将物体 i 的模型输入改为相对于物体 i $t-T$ $_i^t = R$ ^{t} $_i^{t-T}$，$_i^t =$ 在时 R 间 处的初始位置：$(\mathbf{r}(\mathbf{r},\mathbf{o})\mathbf{O}(\mathbf{O}^t_{\mathbf{abs}},\mathbf{o}_i^{t-T}))$。这种表示有助于隐式推断交互对每个物体的独立影响：如果一个物体在被向前推动时开始向左移动，这很可能意味着它正在与另一个影响其运动的物体发生交互。因此，这种以物体为中心的表述具有可扩展性和数据高效性，因为在具有代表性的交互上训练的单一模型，可以泛化到包含任意数量物体和新颖布局的复杂场景。

训练好的以物体为中心的模型 $B(\mathbf{r}_i^t, \mathbf{O}_i^{t-1}) \in \mathbb{R}$ 预测：给定物体历史$\mathbf{O}_i^{t-1}$，当 $B(\cdot) \geq 0$ 时，在状态$\mathbf{r}_i^t$下与机器人的交互是安全的。我们将 N 个独立的以物体为中心的屏障函数组合成一个全局场景级控制障碍函数，以验证所有物体的联合安全性：

$$B_{\text{global}}(\mathbf{r}^t_{1:N}, \mathbf{O}^{t-1}_{1:N}) = \min_{i=1,\ldots,N} B(\mathbf{r}^t_i, \mathbf{O}^{t-1}_i) \tag{5}$$

这种组合在计算上高效，其计算量随物体数量线性增长，并且各个屏障函数的计算可以轻松并行化。

V. 方法实现：

本节详细介绍我们如何训练和部署所提出的方法，以实现在密集物体环境中的安全、可扩展交互。我们的方法分为两个阶段：离线学习控制屏障函数，以及将其作为安全过滤器进行在线部署。离线流程包括：**1) 数据收集：**在仿真中收集交互丰富的数据**2) 初始训练：**基于收集的数据学习神经控制屏障函数**3) 细化：**迭代细化学习得到的控制障碍函数，以降低其保守性。在线部署阶段使用细化后的控制障碍函数作为安全过滤器进行控制合成。本节将介绍各个流程的具体细节。

A. 在线：安全集与安全滤波器

给定离散控制障碍函数 B，为保持物体 i 安全而诱导出的控制不变集为 $\mathcal{C} = \{(\mathbf{r}_i^t, \mathbf{O}_i^{t-1}) | B(\mathbf{r}_i^t, \mathbf{O}_i^{t-1}) \geq 0\}$。等价地，全局控制障碍函数诱导出一个控制不变集，以维护相对于所有物体的安全性，$\{$ $_{1\ N}^{t+1}$ $_{1:N}^{t}$ $|B_{\text{global}}$ $_{1\ N}^{t+1}$ $_{1:N}^{t}$ $\geq 0\}$ $\mathcal{C}_{\text{global}} =$ $(\mathbf{r}:$ ，$\mathbf{O})$ $(\mathbf{r}:$ ，$\mathbf{O})$。在运行时，全局控制障碍函数充当一种最小侵入式安全过滤器。给定B_{global}、名义控制 u_{nom}以及可能的机器人动作集合 $\mathbb{U}$，可以求解一个约束优化问题，以生成满足控制障碍函数不变性条件的最小侵入式安全控制 u_{safe}：

$$u_{\text{safe}} = \arg\min_{u \in \mathbb{U}} ||u - u_{\text{nom}}|| \ s.t. \tag{6}$$

$$B_{\text{global}}(f(\mathbf{r}^t_{1:N}, u), \mathbf{O}^t_{1:N}) - B_{\text{global}}(\mathbf{r}^t_{1:N}, \mathbf{O}^{t-1}_{1:N}) \gamma(B_{\text{global}}(\mathbf{r}^t_{1:N}, \mathbf{O}^{t-1}_{1:N})) \geq 0 \tag{7}$$

上述优化问题可能具有较高的计算难度，因为我们不假设动力学具有控制仿射形式。作为一种实用的替代方案，我们采用以下基于采样的方法。通过检查潜在下一状态的全局控制障碍函数值，可以验证机器人的安全性。如果名义动作 u_{nom} 满足$B_{\text{global}}(f(\mathbf{r}^t_{1:N},\ u_{\text{nom}}),\ \mathbf{O}^t_{1:N}) \geq 0$，则认为该动作是安全的。如果名义动作不安全，我们可以通过围绕名义控制 u_{nom} 进行采样来构造候选动作集 $\mathbb{U}$。在安全动作的子集 $\mathbb{U}_{\text{safe}} = \{u \in \mathbb{U} | B_{\text{global}}(f(\mathbf{r}^t_{1:N},\ u),\ \mathbf{O}^t_{1:N}) \geq 0\}$ 中，我们选择最接近名义控制的动作：$u_{\text{safe}} = \arg\min_{u \in \mathbb{U}_{\text{safe}}} ||u - u_{\text{nom}}||$。这样既能确保安全性，又能将偏离程度降至最低。

如果不并行执行，那么每次动作之前为每个物体计算控制屏障函数的代价可能很高。在实践中，通常只需检查与机器人距离在固定范围内，或近期状态发生过变化的物体。这些物体与安全性最为相关，因为它们与机器人的交互，无论是直接交互还是通过其他物体间接交互，都可能对机器人的安全造成重大影响。这种选择性评估可以降低计算开销，同时不损害安全性。

B. 离线：数据收集

为了通过隐式交互建模有效地学习控制屏障函数，我们首先在简单的仿真环境中收集丰富的交互数据，其中包含多达 N 个随机初始化的物体。只要能够涵盖机器人与物体、物体与物体之间交互的大部分变化，即使 N 远少于我们希望部署该模型时所面对的物体数量也无妨。数据集还必须具有代表性，并使用能够涵盖安全和不安全动作的控制策略，以便控制屏障函数正确学习安全边界。此过程为每个物体 i生成转移对：$\{(\mathbf{r}^t_i,\ \mathbf{O}^{t-1}_i),\ (\mathbf{r}^{t+1}_i,\ \mathbf{O}^t_i)\},\ \forall i \in [1, \ldots, N]$；根据后一个状态是否满足真实安全判据，为每个转移对标注安全或不安全。这就构成了我们的安全数据集 $\mathcal{D}_s$ 和不安全数据集 $\mathcal{D}_u$ 。除了数据点，我们还保存环境对应的完整状态，以供细化时使用。

C. 离线：初始训练

我们使用神经网络来表示 B。网络架构由一个用于编码物体历史状态的长短期记忆网络（LSTM）和一个用于编码机器人状态的多层感知器（MLP）组成。这些特征经过融合后，再输入最终的多层感知器，以输出标量屏障值。

初始训练通过最小化损失函数 $L(\theta)$来优化神经控制屏障函数的参数 θ:

$$L_s(\theta) = \frac{1}{|\mathcal{D}_s|} \sum_{(\mathbf{r}^t_i, \mathbf{O}^{t-1}_i) \in \mathcal{D}_s} \Big[- B(\mathbf{r}^t_i,\ \mathbf{O}^{t-1}_i) \Big]_+, \tag{8}$$

$$L_u(\theta) = \frac{1}{|\mathcal{D}_u|} \sum_{(\mathbf{r}^t_i, \mathbf{O}^{t-1}_i) \in \mathcal{D}_u} \Big[B(\mathbf{r}^t_i,\ \mathbf{O}^{t-1}_i) \Big]_+, \tag{9}$$

$$L_d(\theta) = \frac{1}{|\mathcal{D}|} \sum_{((\mathbf{r}^t_i, \mathbf{O}^{t-1}_i), (\mathbf{r}^{t+1}_i, \mathbf{O}^t_i)) \in \mathcal{D}} \Big[(1-\gamma)\, B(\mathbf{r}^t_i,\ \mathbf{O}^{t-1}_i) - B(\mathbf{r}^{t+1}_i,\ \mathbf{O}^t_i) + \sigma \Big]_+, \tag{10}$$

$$L(\theta) = \eta_s \cdot L_s(\theta) + \eta_u \cdot L_u(\theta) + \eta_d \cdot L_d(\theta). \tag{11}$$

其中，$[\cdot]_+$ 是ReLU函数。这里，$L_s(\theta)$确保控制屏障函数在安全样本（$\mathcal{D}_s$）上输出正值，$L_u(\theta)$确保其在不安全样本（$\mathcal{D}_u$）上输出负值，而 $L_d(\theta)$则对违反控制屏障函数条件的情况（$\mathcal{D} := \mathcal{D}_s \cup \mathcal{D}_u$）施加惩罚。

D. 离线：精炼

使用该初始策略训练的控制屏障函数通常过于保守，由于初始数据采集策略并非最优，它会将许多状态判定为不安全。这种保守性会妨碍机器人完成任务。同样，初始训练效果不佳可能导致将不安全状态误判为安全状态，从而损害安全集的控制不变性。为解决这一问题，我们对安全边界进行精炼，以降低保守性并提高不变性。具体而言，我们从数据集中识别边界附近的样本，其中 $|B(\mathbf{r}^t_i,\ \mathbf{O}^{t-1}_i)| \leq \delta$ ，其中δ 是一个可调的边界精炼阈值。随后，我们重新模拟这些关键场景，并使用完整的多物体场景来捕捉复杂交互。从这些状态出发，我们施加“最安全”的控制 $u_{\text{refinement}}$，即使全局屏障值最大化的动作：

$$u_{\text{refinement}} = \arg\max_{u \in \mathbb{R}^m} B_{\text{global}}(f(\mathbf{r}^t_{1:N}, u), \mathbf{O}^t_{1:N}) \tag{12}$$

我们最多执行 s 个细化步骤，采用这种最安全的控制策略并收集所有数据对，再根据真实值仿真中最终得到的安全性，将状态重新标记为安全或不安全。可以调整这一细化过程，使其仅关注满足 $B(\mathbf{r}^t_i, \mathbf{O}^{t-1}_i) \leq 0$ 或 $B(\mathbf{r}^t_i, \mathbf{O}^{t-1}_i) \geq 0$ 的样本，从而分别使控制屏障函数更保守或更不保守。随后，使用更新后的数据集重新训练控制屏障函数。这一迭代过程可持续进行，直至控制屏障函数收敛，且数据集中不再发现误分类，从而得到最终的、控制不变的学习得到的控制屏障函数。

VI. 实验与结果

我们在仿真中评估所提出的方法，以展示其在物理接触不可避免的杂乱环境中实现安全且可扩展交互的能力。我们设计了实验，以评估随着物体数量增加、场景复杂度提高，方法的性能和泛化能力会如何变化。

实验设置： 实验在自定义的 IsaacLab 环境中进行，该环境包含一条 Franka Emika Panda 机械臂和若干圆柱形瓶子，如图2(a)所示。机器人需要到达随机目标位置，同时避免任何瓶子的倾斜角超过 15° 的安全阈值。机器人的末端执行器被限制在二维（$x,\ y$）平面运动中，动作定义为二维航路点；训练期间，最大步长为 1cm。每个瓶子的状态由其笛卡尔坐标和倾斜角描述：($x,\ y,\ z,\ \theta$)，相对变换通过计算

在笛卡尔空间中减去（x，y）得到。我们通过在每次试验中随机改变每个瓶子不可观测的物理属性，测试该方法的鲁棒性，包括质量（$1.3-2.0$kg）和摩擦系数（静摩擦系数：$(0.5-0.7)$，动摩擦系数：$(0.3-0.49)$）。可观测参数包括瓶子半径（5cm）和高度（20cm），这些参数保持固定。不过，也可以通过输入条件化将可观测的变化纳入我们的框架。

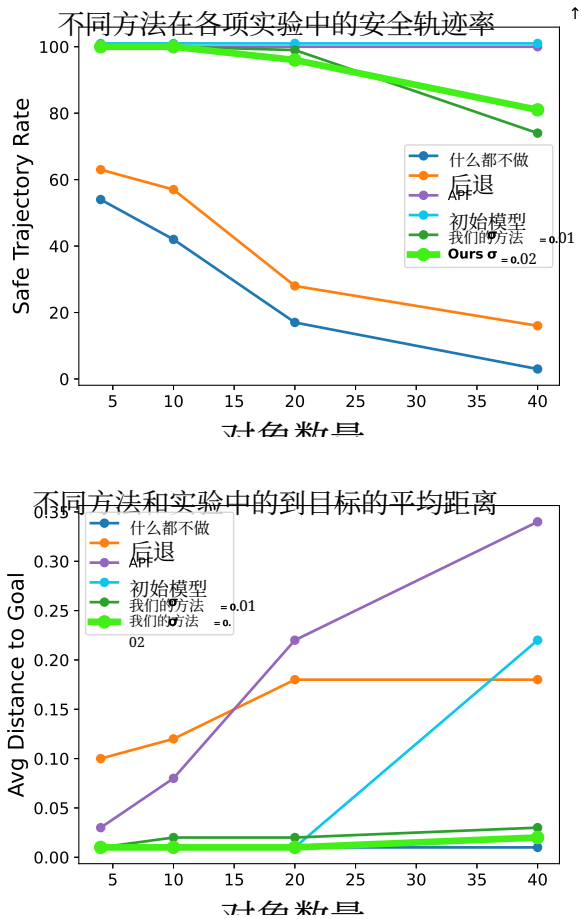

图4：不同方法在物体密度各异的环境（4、10、20 和 40 个物体）中的表现。安全率衡量安全轨迹所占的百分比（越高越好），目标平均最终距离衡量到目标的最终距离（越低越好）。这些指标基于图3所示的同一实验计算得出，进一步展示了各方法的性能。这些结果凸显了基线方法（什么都不做、反步法）在保持安全方面的局限性、人工势场的保守性、初始模型 在更高密度下的保守性，以及我们提出的方法在兼顾安全性和任务成功率方面的鲁棒性。

训练与精炼：我们的控制屏障函数是在仅包含 4 个瓶子的环境中训练的。我们使用$\sigma = 0.01$ 和 $\sigma = 0.02$进行训练。首先，我们使用朴素反步策略，从 1200 条轨迹中收集了一个包含 9600000 对数据的初始数据集。该策略会在接近安全阈值时退回到机器人的上一个航路点。我们通过丢弃自由空间中的样本来平衡该数据集，以确保其中既包含丰富的安全交互示例，也包含接近不安全的交互示例。为了解决仅从朴素控制器收集数据所带来的局限，尤其是难以获取安全边界附近数据的问题，我们对模型进行迭代细化。利用 IsaacLab 的并行仿真能力，我们初始化了物体交互程度较高的状态，收集新的数据批次，并将精炼步骤设为 $s = 4$，随后对控制屏障函数进行微调。该过程重复了 47 个批次。

可视化学习到的安全边界：为了可视化学习得到的控制障碍函数，我们为图2(b)(c)中的 40 个物体场景生成了全局值图。该图表明，初始控制屏障函数（细化前）过于保守，形成了无法通行的区域，导致机器人停滞。相比之下，细化后的控制障碍函数学习到了更紧凑的安全边界，识别出允许物理交互的安全路径，从而能够在密集环境中成功导航。

定量评估：我们在包含 4、10、20 和 40 个物体的环境中，将最终模型与四种基线方法进行了评估，其中成功的定义是在保持安全性的同时到达目标。基线方法包括：**什么都不做：**直线朝目标移动；**反向步进：** 用于初始数据收集的反应式策略；**人工势场（APF）**：一种使用排斥场的经典方法；以及**初始模型：** 细化过程之前的控制屏障函数模型。我们还比较了使用不同鲁棒性裕度 $\sigma = 0.01$ 和 $\sigma = 0.02$ 训练的方法。我们将不执行任何操作基线用作标称控制器。反向步进被调节为在任何接近 14° 的倾斜角下向后退步；人工势场参数被调节为 $K_P = 5.0$ （吸引势增益）、 $\eta = 50.0$ （排斥势增益）、长度为 1.2 米（桌面尺寸）的势场区域，以及振荡检测长度为 3，从而允许发生一定程度的接触，同时避免过于保守。图1展示了使用我们的方法（我们的方法）和不执行任何操作基线生成的轨迹样本，以及所关注高亮物体的倾斜角。基线方法会使多个物体违反安全要求，而我们的方法则将所有物体保持在安全阈值以下。

图3中的定量结果表明，我们提出的方法在安全性和任务成功率方面始终优于所有基线方法，尤其是在物体密度增加时。该指标衡量的是机器人既保持安全又到达目标的轨迹所占的百分比。图4进一步分解了每种方法的安全率和目标平均最终距离。朴素的什么都不做和反向步进方法未能充分考虑机器人—物体交互的影响，经常违反安全约束并导致物体倾倒。相比之下，人工势场和我们未经精炼的初始模型过于保守。尽管安全，但它们在密集场景中会在距离目标更远处停滞，从而导致较高的任务失败率以及较大的最终目标距离。我们的最终精炼方法在两者之间取得了必要的平衡，在穿越密集物体排列时实现了较高的到达成功率和安全率。比较 $\sigma = 0.01$ 和 $\sigma = 0.02$可见，对安全裕度进行低估可能会

导致安全违规。原则上，过近似可能导致过度保守。即使我们的方法出现违规，失败轨迹中的平均最大倾斜角在20个和40个物体的情况下分别为17.65°和18.88°，仅表现出轻微违规；而4个和10个物体的场景始终保持安全（报告于 $\sigma = 0.02$）。相比之下，基线方法什么都不做的违规幅度大得多，分别为66.81°、68.45°、68.16°和46.40°。值得注意的是，尽管我们的模型仅在4个物体场景上进行训练，但它仍能有效泛化到包含40个物体的复杂环境中，展现出良好的可扩展性和性能。

VII. 结论

在本研究中，我们提出了密集接触屏障函数，这是一种可扩展的框架，能够在密集杂乱环境中实现安全机器人交互。我们的方法成功地从仅使用少量物体进行训练，推广到包含许多物体的复杂场景，而无需显式建模难以处理的多物体动力学。尽管研究结果令人鼓舞，未来仍有多个值得探索的方向。我们计划验证该方法处理更多样化物体的能力，包括具有不同可观测参数的物体。此外，我们还计划将该方法扩展到更复杂的场景，涉及不受约束的全身机器人运动和可变形物体操控，并最终在实体硬件上展示其有效性。本研究是迈向部署能够在现实世界中无结构、高接触环境里安全且熟练地穿行的机器人的关键一步。

参考文献

[1] S. Bansal、M. Chen、S. Herbert 和 C. J. Tomlin，“Hamilton-Jacobi 可达性：简要概述与近期进展”，载于*2017 IEEE 第56届年度决策与控制会议（CDC）*。IEEE Press，2017年，第2242–2253页。 [在线]。可访问：https://doi.org/ 10 1109 2017 8263977 。/CDC。

[2] A. D. Ames、S. Coogan、M. Egerstedt、G. Notomista、K. Sreenath 和 P. Tabuada，“控制障碍函数：理论与应用”，载于*2019年第18届欧洲控制会议（ECC）*，2019年，第3420–3431页。

[3] M. C. Lee 和 M. G. Park，“基于人工势场的移动机器人路径规划：虚拟障碍物概念的应用”，载于*2003 IEEE/ASME 国际高级智能机电一体化会议论文集（AIM 2003）*，第2卷，2003年，第735–740页，第2卷。

[4] Y. Tang、X. Chu、J. Huang 和 K. W. Samuel Au，“用于受约束可变形线状物体操控的带安全滤波器的基于学习的模型预测控制”，*IEEE 机器人与自动化快报*，第9卷，第3期，第2877–2884页，2024年。

[5] X. Ding、H. Wang、Y. Ren、Y. Zheng、C. Chen 和 J. He，“杂乱环境中机器人操作臂的安全关键最优控制”，2022年11月。

[6] N. Shinde、X. Liang、F. Liu、Y. Zhang、F. Richter、S. Herbert 和 M. Yip，“Jiggle：用于可变形手术环境中边界参数估计的主动感知框架”，2024年7月。

[7] K. Nakamura、L. Peters 和 A. Bajcsy，“通过潜在空间可达性分析将安全性推广至避碰之外”，2025年。 [ ] 2502 00935 在线。可访问：https://arxiv.org/abs/。

[8] O. So、C. Ge 和 C. Fan，“使用强化学习求解最小成本到达-避障问题”，2024年。 [在线]。可访问：https://arxiv.org/abs/ 2410 22600 。

[9] W. Sharpless、D. Hirsch、S. Tonkens、N. Shinde 和 S. Herbert，“采用新型Hamilton-Jacobi-Bellman公式的双目标强化学习”，2025年。 [在线]。可访问：https://arxiv.org/abs/2506.16016

[10] A. Begzadic、N. Shinde、S. Tonkens、D. Hirsch、K. Ugalde、M. Yip、J. Cortes 和 S. Herbert，“回到起点：通过带到达-避障安全滤波器的安全重置实现免干预学习”，载于第七届年度学习动力学与控制会议论文集，丛书《机器学习研究论文集》，N. Ozay、L. Balzano、D. Panagou 和 A. Abate 编，第283卷。PMLR，2025年6月4日至6日，第1154–1166页。 [在线]。可访问：https://proceedings.mlr.press/v283/begzadic25a.html

[11] S. Bansal 和 C. Tomlin，“DeepReach：一种面向高维可达性的深度学习方法 ”，2020 年。 [在线]。可访问：https: //arxiv.org/abs/ 2011 02082 。

[12] Z. Feng、L. Qiu 和 S. Bansal，“连接模型预测控制与深度学习，实现可扩展可达性分析”，2025年。 [在线]。可访问：https://arxiv.org/abs/ 2505 03830 。

[13] D. Morton 和 M. Pavone，“采用操作空间控制障碍函数的安全、任务一致的操作”，2025年。 [在线]。可访问：https://arxiv.org/abs/ 2503 06736 。

[14] D. A. Dsouza、S. Shenoy、M. Wang 和 A. R. Chowdhury，“在受限工作空间中利用改进人工势场实现人机协作的综合安全架构”，Robotica，第43卷，第4期，第1373–1393页，2025年4月。

[15] J. Lee、M. Seo、A. Bylard、R. Sun 和 L. Sentis，“采用奇异点容忍分层任务控制的工业机械臂实时模型预测控制”，2022年。

[16] J. Achiam、D. Held、A. Tamar 和 P. Abbeel，“约束策略优化”，2017年。 [在线]。可访问：https://arxiv.org/abs/1705。10528

[17] S. Gu、B. Sel、Y. Ding、L. Wang、Q. Lin、M. Jin 和 A. Knoll，“安全强化学习中的奖励与安全优化平衡：梯度操控视角”，2025年。 [在线]。可访问：https://arxiv.org/abs/ 2405 01677 。

[18] M. Macklin、M. Muller 和 N. Chentanez，“XPBD：基于位置的柔顺约束动力学仿真”，载于第九届国际游戏运动会议论文集，2016年，第49–54页。

[19] M. Mittal、C. Yu、Q. Yu、J. Liu、N. Rudin、D. Hoeller、J. L. Yuan、R. Singh、Y. Guo、H. Mazhar、A. Mandlekar、B. Babich、G. State、M. Hutter 和 A. Garg，“Orbit：用于交互式机器人学习环境的统一仿真框架”，IEEE 机器人与自动化快报，第8卷，第6期，第3740–3747页，2023年。

[20] L. Wei、J. Ma、Y. Hu 和 R. Zhang，“通过隐式触觉校准确保视觉引导机器人操作中的力安全”，2024年。

[21] E. Shakhesi、A. Katriniok 和 W. P. M. H. M. Heemels，“反例引导的鲁棒离散时间控制障碍函数综合”，IEEE控制系统快报，第9卷，第1574–1579页，2025年。 [在线]。可访问：http://dx.doi.org/10.1109/ LCSYS.2025.3578971

[22] A. Alan、T. G. Molnar、A. D. Ames 和 G. Orosz，“保证不确定性下安全性的参数化障碍函数”，IEEE控制系统快报，第7卷，第2077–2082页，2023年。

[23] K. Garg、J. Usevitch、J. Breeden、M. Black、D. Agrawal、H. Parwana 和 D. Panagou，“控制障碍函数理论进展：应对自主系统与机器人系统安全控制综合中的实际挑战”，2023年。 [在线]。可访问：https://arxiv.org/abs/ 2312 16719 。

[24] S. Liu、C. Liu 和 J. Dolan，“输入受限条件下利用神经控制障碍函数实现安全控制”，2022年。 [在线]。可访问：https://arxiv.org/abs/ 2211 11056 。

[25] S. Raja、A. Kumar 和 P. Gupta，“OGM-CBF：基于占据栅格地图的控制障碍函数，用于机器人安全控制”，载于 IEEE 国际机器人与自动化会议（ICRA）论文集，2024年，第5678–5685页。

[26] M. Tayal、H. Zhang、P. Jagtap 和 A. Clark，“在随机环境中学习形式化验证的控制障碍函数”，2024年。

[27] H. Yu、C. Hirayama、C. Yu、S. Herbert 和 S. Gao，“用于可扩展动态避障的序贯神经障碍函数”，2023年。 [ ] 2307 03015 在线。可访问：https://arxiv.org/abs/。

[28] S. Tonkens 和 S. Herbert，“通过Hamilton-Jacobi 可达性优化控制障碍函数”，2022年。 [在线]。可访问：https://arxiv.org/abs/ 2204 12507 。

[29] N. U. Shinde、J. Johnson、S. Herbert 和 M. C. Yip，“用于采用非参数方法进行交互式在线学习的以物体为中心的表征”，2023年。 [在线]。可访问：https://arxiv.org/abs/2307.10063