arXiv:2302.02898v1 [cs.RO] 6 Feb 2023

演示：Arena-Web——用于自主导航方法的基于网页的开发与基准测试平台

Linh Kästner[1], *、Reyk Carstens[1], *、Lena Nahrwold[1],Christopher Liebig[1],Volodymyr Shcherbyna[1],Subhin Lee[1],和 Jens Lambrecht[1]

摘要—近年来，由于应用领域广泛，从医疗保健到仓储物流，移动机器人导航方法变得越来越重要。尤其是，深度强化学习方法在机器人导航领域日益受到欢迎，但对于非专业人士而言，这些方法并不易于使用，开发过程也较为复杂。近年来，人们不断努力使这些复杂的方法能够被更广泛的受众所使用。本文介绍了 Arena-Web，这是一个基于网页的开发与评估套件，用于开发、训练和测试适用于各种机器人平台与场景的基于深度强化学习的导航规划器。该界面旨在做到直观且富有吸引力，以吸引非专业人士，并使更广泛的受众能够使用这项技术。借助 Arena-Web 及其界面，无需编写一行代码，即可简化深度强化学习智能体的训练与开发，使其变得轻松易行。该网页应用可免费使用，并通过补充材料中所提供的链接公开访问。

一、引言

近年来，深度强化学习（DRL）在导航和运动规划领域不断取得进展，多项研究将 DRL 应用于其方法中 [1], [2]。然而，开发深度强化学习智能体面临诸多障碍和困难，例如训练复杂、配置繁琐或硬件成本高昂，因此相关工作往往局限于该领域的专家和从业者。这使得不同方法之间难以进行比较[3]。不过，对 DRL 方法进行基准测试和测试，对于将其与经典方法进行比较，并评估其性能和可行性至关重要。近年来，人们也为规范基准测试流程付出了努力 [4], [5], [6]。然而，这类应用由多个软件包和代码库组成，普通用户可能难以访问和理解。此外，这些应用还需要安装和配置，或依赖指定硬件，这进一步提高了普通用户和新入行从业者的使用门槛。另一方面，基于 Web 的应用程序已被证明能够带来更直观的体验，且更易于获取。已有多项工作致力于将机器人应用引入网络，并通过 Web 应用为用户提供帮助 [7], [8],[9]。

网页应用可从所有操作系统轻松访问，并且无需特定硬件即可启动复杂任务。此外，无需经历

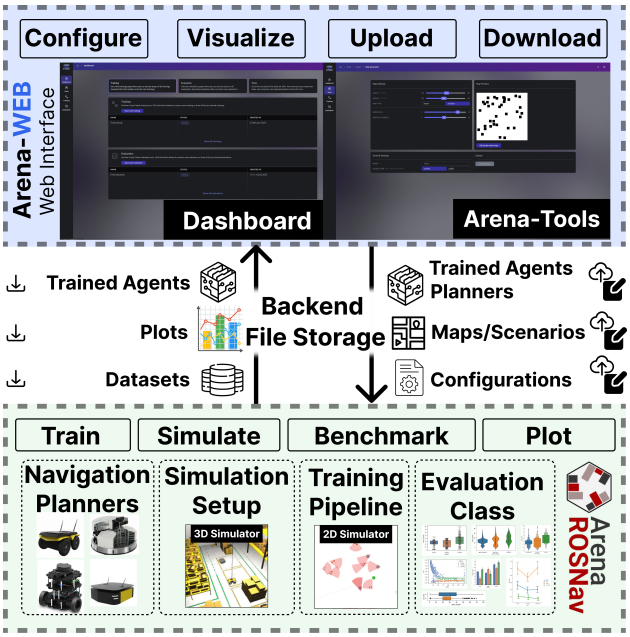

图 1：Arena-Web 提供了一个基于网页的界面，可在任意计算机上便捷地开发、训练和测试导航方法。它提供了用于生成多种实用内容的界面，例如地图、场景、训练参数和任务，以便定义训练和测试运行。用户界面采用直观且美观的设计，用户可以在一系列预定义步骤的引导下完成操作，同时不会牺牲高度的定制可能性。该网页应用构建于我们先前开发的 Arena-Rosnav 之上，Arena-Rosnav 在后端运行，并包含执行仿真和生成数据所需的全部功能。

复杂的安装流程可能导致错误和问题，从而削弱人们投身这一领域的动力。因此，能够促进轻松便捷开发的应用程序，不仅能帮助有志于从事深度强化学习和运动规划方法研究的初学者，也能帮助希望借助统一、简洁的平台对其方法进行基准测试，并节省时间和成本的专业科研人员。为此，我们提出了 Arena-Web，这是一款基于 Web 的应用程序，提供开发深度强化学习智能体的功能，并可利用多种工具，将这些导航方法与经典方法进行基准测试。用户可以通过多种方式设置并生成不同的地图、世界、场景或任务模式。流程中的各个环节均可单独自定义和编辑，用户还可以根据自身需求进行设置。用户可以自定义要测试智能体的任务模式，也可以创建新的世界和场景、神经网络

[1]Linh Kästner、Reyk Carstens、Christopher Liebig、Volodymyr Shcher- byna、Lena Nahrworld、Subhin Lee 和 Jens Lambrecht 均来自柏林工业大学电气工程与计算机科学学院工业级网络与云教席，德国柏林linhdoan@tu-berlin.de

*贡献相同

架构等。Arena-Web 是一个网页界面，整合了我们先前开发的 Arena-Rosnav [10]和 Arena-Bench [5]的功能，并增添了多项功能。具体而言，该网页应用能够：在自定义地图上使用自定义奖励值和自定义神经网络架构训练深度强化学习智能体；评估这些方法，并与其他基于学习的方法及经典导航规划器进行比较和基准测试；下载生成的模型和评估文件，以供后续使用，例如绘制结果图表或在真实机器人上运行模型。此外，平台还提供多种编辑器，可用于创建场景、神经网络架构，设置超参数、地图和自定义奖励。所有功能都集成在一个直观且富有吸引力的网络应用程序中，以提升用户体验、理解度和可访问性。本研究的主要贡献如下：

- 据我们所知，这是首个基于网络的深度强化学习开发与测试平台，允许用户使用多种经典和基于学习的导航方法创建、训练和评测深度强化学习智能体。该 Web 应用构建于我们先前的工作 Arena-Bench 之上，后者具有模块化架构，并可扩展支持其他规划器和方法。
- 可以自定义整个训练流程，包括自定义网络架构、超参数、奖励函数和场景。
- 可以自定义整个评估流程，包括自定义场景、任务模式和规划器。平台提供大量定性和定量评估指标以及绘图功能。

本文结构如下。第二节首先介绍相关工作。随后，第三节介绍所采用的方法。第四节介绍可使用该应用执行的示例用例。最后，第五节给出结论并展望未来工作。

II. 相关工作

事实证明，基于深度强化学习的导航方法是一种前景广阔的替代方案，已成功应用于各种机器人应用，并取得了显著成果。因此，许多研究工作将深度强化学习融入其导航系统中 [1], [2], [11], [12], [13], [14], [15], [10], [16]。

然而，大多数此类研究工作有时需要繁琐的安装和配置，才能测试和验证相关方法。过时的版本往往需要手动调整和修复错误，耗费大量成本和时间。此外，能够将基于深度强化学习的方法与经典方法并排进行基准测试的标准化平台十分少见且有限 [3], [5]。目前已有一些平台致力于为这些方法的基准测试和开发提供支持 [4], [5], [17], [6]。然而，这些平台也需要安装和配置，或依赖指定硬件，因此对普通用户或新手而言不够直观，也难以使用。另一方面，基于网络的应用程序已被证明能够带来更直观的体验，并且更广泛地向公众开放，让更多人能够使用。

一直以来，人们都在努力将机器人操作系统提供的服务作为网络应用程序使用。过去二十年间，人们已开始尝试将机器人技术与基于 Web 的应用程序相结合。2000 年，Schulz 等人[18] 提出了基于 Web 的界面，以便通过互联网远程控制机器人。研究人员提供了一个基于 Web 的用户界面，用户可以在其中选择地图，并引导机器人前往地图中的指定位置。类似地，Simmons 等人提出了功能更丰富、可视化效果更完善的扩展改进版本。其他用于机器人控制和交互的基于 Web 的平台还包括 [7], [8], [19], [9]。使用网页界面也有助于改善人机协作与交互，因为与命令行工具或开发环境相比，Web 服务更易于访问，也更开放。因此，较新的机器人网络应用程序涵盖了医疗保健领域的机器人辅助 [20], [21], [22], 、制造业 [23], 、引导机器人 [24], [25] 或教育用途 [26], [27], [28],[29]

此外，随着机器人操作系统（ROS）的出现， [30],人们提出了将 ROS 与网络相结合的工具。Osentoski 等人提出的开源库（如 ros.js） [31] ，以及 Toris 等人提出的 Robot Web Tools [32],, 通过 rosbridge 以 Web 应用的形式提供 ROS 的主要功能，有助于促进基于网络的机器人开发与使用，从而扩大开发者群体。最近，Webviz [33] 和 Foxglove [34] 等工具致力于将基于 ROS 的可视化与监控完全转移到网络上。

鉴于统一评估与基准测试平台的局限性，本文提出 Arena-Web，以促进基于深度强化学习的方法的开发，并将其与其他基于学习的方法和经典导航规划器进行评估。该平台完全通过网络应用程序提供服务，以促进快速开发和直观使用，从而加快开发周期。该工具旨在让更广泛的受众参与机器人导航方法的深度强化学习开发与基准测试。

III. 功能

Arena-Web 是一个网页界面，整合了我们此前作品 Arena-Rosnav [10] 和 Arena-Bench [5], 的功能，并在此基础上增加了多项功能，同时作为一个开放访问的用户界面。该应用的主要任务分为三类：工具、训练和评估。本节将概述 Arena-Web 当前具备的功能。值得注意的是，该应用仍在不断添加新功能。

A. 工具

我们提供了一些工具，用于生成开展训练或评估任务所需的多种重要实用资源。图2展示了目前所有可用工具的概览。

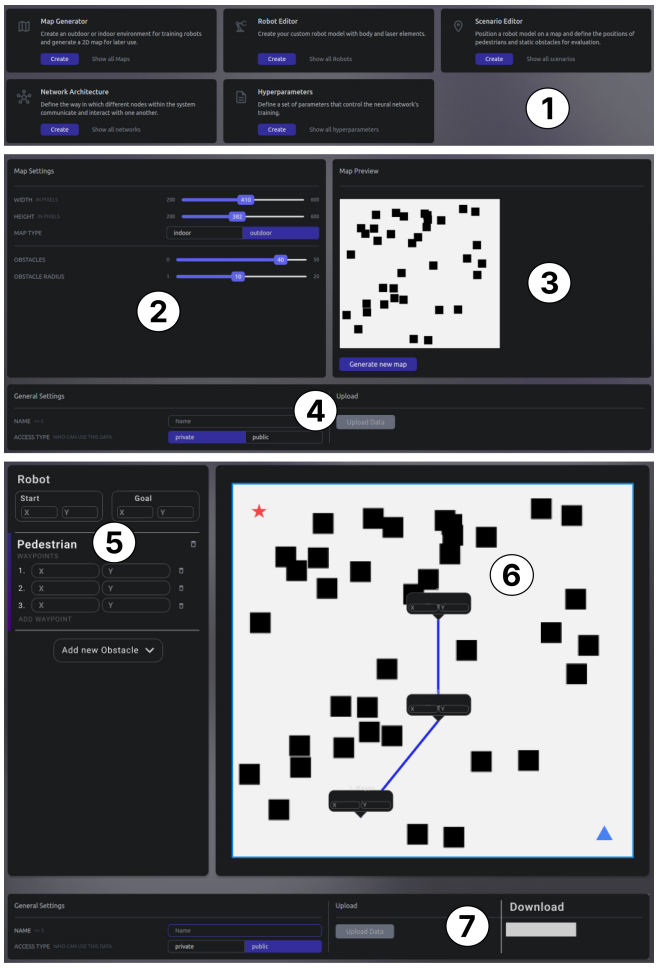

图2：我们提供了多种编辑器/生成器，用于生成各种工具，例如占据地图、具有预设障碍物数量和行为的场景，以及用于生成新神经网络架构、选择训练超参数、奖励系统等的训练流程编辑器。(1) 用于选择工具的仪表板；(2) 可通过滑块调整地图相关参数；(3) 实时预览地图，用户可以直接观察变化；(4) 也可以上传现有地图；(5) 在场景编辑器中，用户可以通过拖放输入机器人和障碍物的位置，数值会显示在侧边栏中；(6) 向用户预览场景；(7) 用户可以上传现有场景。

地图生成器： 地图生成器能够为室内和室外环境生成二维世界。用户可以指定多种参数，例如地图的高度和宽度、静态障碍物的数量和尺寸、房间之间通道的宽度，以及室内地图中房间的尺寸和数量。地图生成器将生成占据栅格地图。

场景生成器： 场景编辑器允许用户选择地图并在其中放置动态障碍物、定义其速度，并通过设置航点来定义其行为。还需设置机器人的起点和终点。用户可以通过拖放编辑所有位置，从而简化操作并提升用户体验。之后，可在评估阶段使用该场景，为不同规划器提供相同的环境，以便进行一致的比较。

网络生成器： 使用网络生成器，用户可以为深度强化学习智能体的训练创建个性化的神经网络架构。用户可以指定不同层的输入和输出大小，并可添加或删除这些层。目前支持全连接层、卷积层和 ReLU层，但我们的模块化配置便于在不久的将来轻松添加更多层类型。

奖励生成器： 由于使用深度强化学习来训练智能体，因此必须为特定动作定义奖励。调整奖励会对训练结果产生显著影响。因此，我们通过奖励生成器，让用户能够根据自身的具体需求调整所有奖励。

超参数生成器： 与奖励类似，超参数也是训练过程中不可或缺的重要组成部分。由于设置超参数往往需要反复试错，超参数生成器让用户能够创建自己的超参数组并尝试不同的数值。

B. 训练

我们的网页应用的一项主要功能是为机器人训练新的深度强化学习模型。因此，我们提供了 10 种预设机器人供用户选择。用户可以使用预定义或自定义的网络架构和地图进行训练。此外，用户还可以指定多种超参数，例如学习率、批次大小和评估频率。训练在服务器上运行期间，用户随时可以直接下载当前的最佳模型，并查看日志输出以监控训练过程。训练结束后，用户会收到通知。由于训练始终取决于机器人，我们提供了多种预先配置的机器人，涵盖所有主要的机器人运动学类型。其中包括 Turtlebot3、Robotino Festo、Carobot4、AGV Ota、Ridgeback、Kuka Youbot、Clearapath Jackal、Clearapath Husky、Dingo 和 TiaGo。

C. 评估

评估任务可用于评估训练后的智能体，或 Arena-Rosnav 套件中的其他现有规划器。表 I 列出了所有可用的规划器。我们提供了大量先进的学习型规划器以及经典规划器。用户可以为评估选择预定义场景，从而确保所有评估使用相同的环境，进行一致的比较。用户可随时下载当前评估数据，并通过日志文件查看评估运行的状态。评估运行结束后，用户可以下载生成的 .csv 文件，并使用我们的评估软件包在本地绘制结果。用户

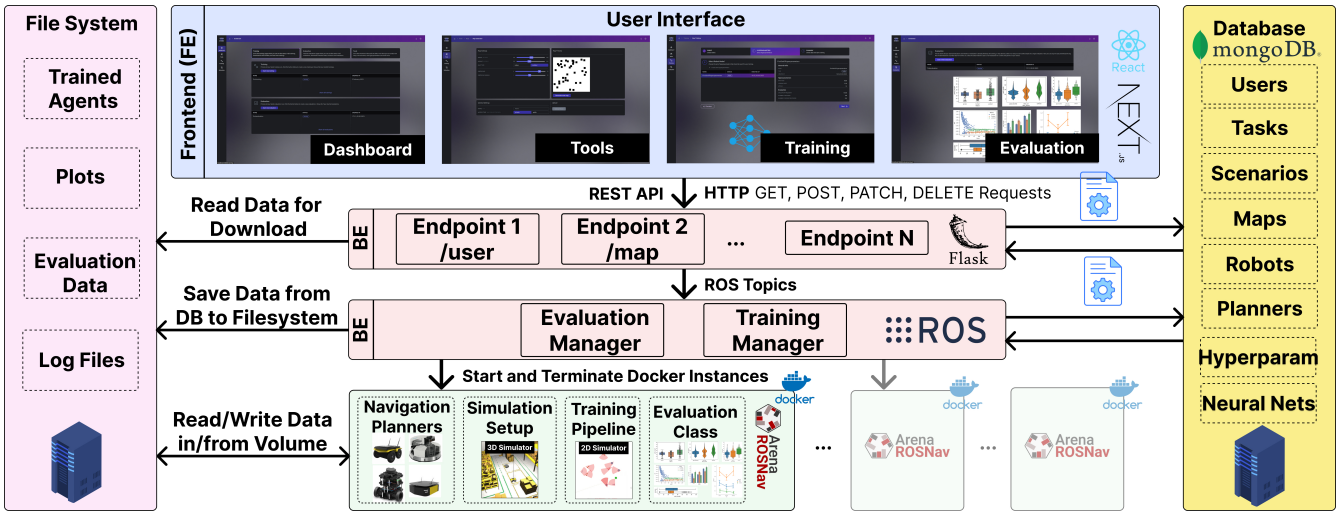

图 3：Arena-Web 的系统设计。前端（FE）提供用户界面，用于自定义和配置训练与评估运行，并创建地图和场景等必要工具。该网页应用基于 Arena-Rosnav 构建（Arena-Rosnav 的系统设计见图 4），后者提供工具，用于训练深度强化学习规划器，并将其与其他基于学习的方法或经典方法进行基准测试，同时绘制各项指标的结果图表。因此，前端通过 REST API 和端点与后端（BE）通信，并通过一个由两个 ROS 节点组成的附加 ROS 层，在后端创建和管理 Arena-Rosnav Docker 容器。生成的文件存储在 MongoDB 数据库中，用户生成的数据则存储在文件系统中。后端和数据库运行在我们的 GPU服务器上。

还可以创建自己的绘图声明文件，用于定义要生成哪些图表以及使用哪些数据。这样，即使是经验不足的用户也能快速绘制结果。目前，我们正致力于将绘图功能直接集成到 Web 应用中，以改善用户体验。

D. 可用机器人和规划器

目前已有多种机器人可供使用，涵盖了所有主要的机器人运动学类型。其中包括 Turtlebot3、Robotino Festo、Carobot4、AGV Ota、Ridgeback、Kuka Youbot、Clearapath Jackal、Clearapath Husky、Dingo 和 TiaGo。此外，表 I 列出了所有可用的规划器。

表 I：可用导航规划器

经典	混合	基于学习
TEB [35]	Applr [36]	ROSNavRL [10]
DWA [37]	LfLH [3]	RLCA [38]
模型预测控制 [39]	Dragon	Crowdnav [11]
Cohan [40]	TRAIL	SARL [41]
		Arena [42]
		CADRL [2]

IV. 方法

A. 系统设计概述

整个系统的示意图如图 3 所示。用户通过前端的用户界面与系统交互，该界面使用 React.js 实现。前端可分为三个主要部分：工具、训练和评估。前端通过 REST API 与后端通信。后端使用 Flask 编写，并为 Web 应用的所有功能提供多个不同的端点，以支持 Web 应用。为实现持久、高效的存储，系统使用 MongoDB 数据库；与关系型数据库相比，MongoDB 能够处理非结构化数据，因此被选用。开始新的训练和评估时，Arena-Rosnav 会作为 Docker 容器在服务器上运行，以处理相关任务。Arena-Rosnav 的系统设计如图 4 所示。由于用户创建的工具应能自定义处理流程，且用户也应能够访问输出和生成的文件，因此服务器文件系统中的部分内容会以卷的形式挂载到 Docker 容器中，以便从容器外部保存和加载文件。为进一步增强安全性并便于管理不同的 Docker 容器，系统在 Flask 后端与 Docker 容器之间实现了一层由 ROS 节点组成的轻量级中间层。

B. 前端

前端提供用户界面，可用于启动新的训练和评估运行，以及创建新的实用工具文件，以配合 III 中所述的工具用于训练或评估。出于安全和管理方面的考虑，整个前端仅向已登录的注册用户开放。这样一来，用户创建的数据和任务始终可以关联到特定用户，从而确保整个交互过程的私密性。我们决定取消实用工具文件与特定任务之间的直接关联，以便创建的文件能够与其他工具结合使用，并用于不同的多个任务。这提高了模块化程度，也让实验所需的工作更少。

在前端中，不同的工具用于为定制训练或评估创建必要条件。因此，我们实现了地图、场景、网络架构、超参数和

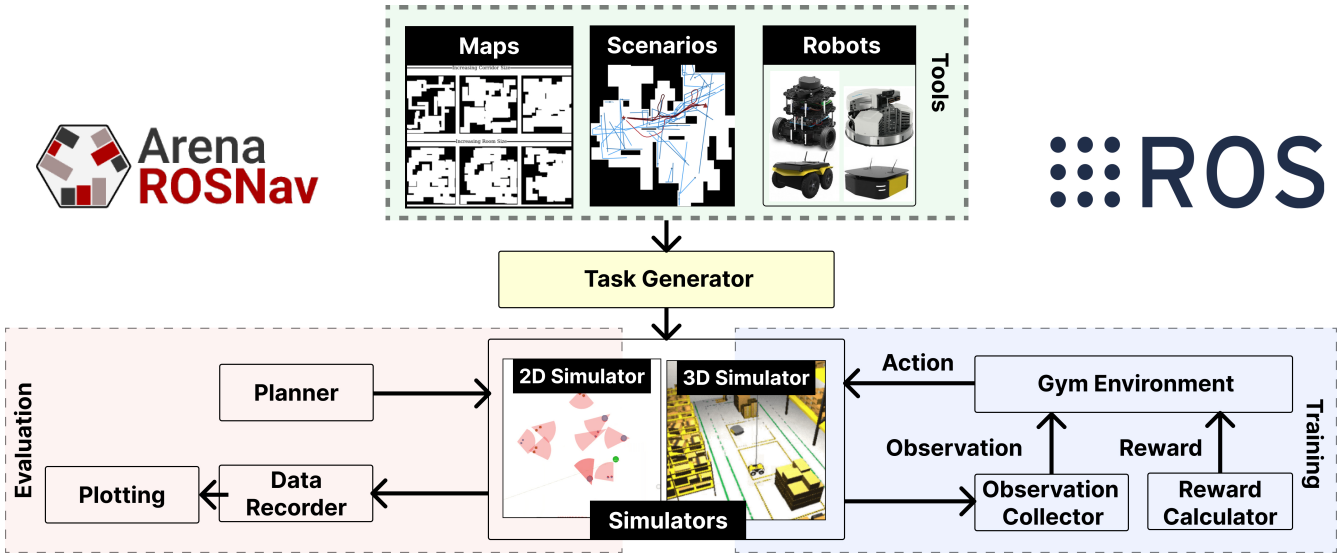

图 4：Arena-Rosnav 的系统设计。该系统提供工具，可在不同环境和场景中开发、训练和评估深度强化学习规划器，尤其适用于高度动态环境。因此，它包含多种仿真环境，例如 Flatland（二维）、Gazebo（三维）和 Unity（三维）。请注意，目前用户只能选择二维环境中的仿真。我们正在开发对三维环境的支持。训练流程会执行深度强化学习训练，用户可以对其进行自定义。随后，训练后的智能体可与其他导航方法进行基准测试。Arena-Rosnav 提供了一套表 I 所列的多种最先进的规划器，以及一个包含数据记录器和绘图类的评估流程。数据记录器会在评估运行期间记录所有必要的数据，绘图类则可利用这些文件，绘制导航效率、安全性和平滑度等多项导航指标的结果。

奖励生成器，这些内容已在第三节中介绍。我们决定将训练和评估中使用的所有实用工具文件与特定任务分开创建，以便在不同的多个运行中重复使用，而无需为不同的训练或评估运行反复创建这些文件。

C. 后端

后端提供了各种不同的 REST 端点，用于实现前端中所有可执行的功能。后端使用 MongoDB 数据库持久、高效地存储通过工具创建的所有数据。使用工具创建的所有实用工具文件，以及训练和评估信息，都会存储在数据库中，以便更轻松地进行操作，并确保后端各个部分都能一致地访问这些内容。除 Flask 后端外，服务器还运行着一个小型 ROS 节点拓扑，用于管理 Docker 容器，训练和评估任务便在这些容器中运行。因此，在启动新任务时，Flask 后端会通知 ROS 拓扑，根据请求中包含的指定配置启动定义的任务。随后，ROS 节点会创建必要的实用工具文件，并将其存储在服务器的文件系统中。接着，系统会为该任务启动一个新的 Docker 容器，并通过 Docker 卷将实用工具文件加载到 Docker 容器中。为了允许用户下载创建的数据，Docker 容器中的文件输出目录也会作为卷加载，使其能够在服务器文件系统中显示，并供后端使用。

使用 Docker 卷不仅可以在之后继续使用创建的数据，还允许针对不同任务以及不同的实用工具文件使用同一个 Docker 容器。使用这一额外的 ROS 节点层具有多项优势。首先，它提供了额外的安全层，因为该层不与后端相连，因此 Flask 后端中的故障或遭受的攻击不会影响 Docker 容器。此外，它使系统能够独立于 Flask 后端运行，并让其进一步集成到其他应用程序中变得更加容易。最后，它能够为用户带来更快的响应，因为端点本身只需发送一条 ROS 消息，而不必自行存储所有文件并启动 Docker 容器。这样，计算密集型任务便可以异步地执行。

与 Arena-Bench 之前的训练和评估能力相比，我们进行了一系列改进。为便于灵活使用不同的神经网络架构、奖励、地图和场景，我们定义了全新的数据存储格式，并集成了高效的解析器，以便在 *Arena-Bench* 中读取和使用数据。此外，前端和后端都会验证所有数据，以便尽早发现潜在错误并减少错误。

V. 示例用例工作流程

本章将说明如何针对不同的用例和需求使用该应用。具体而言，

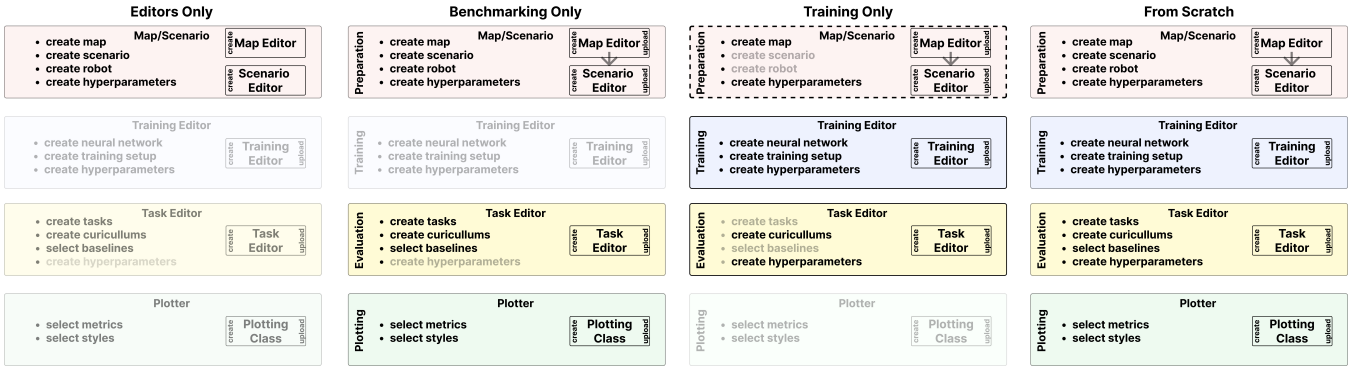

图 5：示例工作流程及其所需模块（针对相应使用场景突出显示）。Arena-Web 可用于满足不同的使用场景和需求。用户可以利用各种编辑器，通过直观的界面创建 ROS 占据栅格地图或不同场景，从而简化这些文件的生成过程，以便在 Arena-Rosnav 本地实例中使用。用户还可以使用网页应用从头开始进行深度强化学习训练，并将训练得到的智能体与多个规划器进行基准测试。其他用途包括利用该应用仅使用现有地图和智能体进行训练，或在不访问开发电脑的情况下，仅使用预先存在的智能体、规划器和场景执行评估任务。例如，这有助于团队在任何地点进行多次评估运行，而无需具备完全相同的电脑配置。我们目前正在开发移动版本，以便用户也能随时随地启动或监控这些任务，从而使开发流程和协作更加灵活高效。

主要有三种用例，可通过网页应用实现：创建新的实用工具文件，例如地图、场景、奖励和超参数；训练深度强化学习智能体；以及评估现有规划器或用户通过网页应用训练的深度强化学习智能体。对于每种用例，我们都会定义一个示例，并展示用户在网页应用中的操作流程，同时说明所需的工具和交互。图 5 展示了示例用例及其所需工具。示例用例如下：

- 创建新场景
- 开始新的训练并下载模型。
- 使用训练好的深度强化学习智能体启动新的评估，下载评估数据并绘制结果图表。

补充材料中提供了一段演示视频，展示了所有功能。

A. 创建新场景

场景在评估中作为固定环境，以便更好地比较不同的规划器方法。场景指定机器人的起点和目标位置，以及不同的动态障碍物（如行人）。此外，还可以通过航点指定障碍物的具体行为，并设置障碍物的速度。要创建新场景，用户必须选择一张地图，该地图可以使用地图生成工具创建。

登录后，用户会进入仪表板，其中简要展示了最新任务的概况。用户可以从这里直接跳转到工具选择页面并开始创建新地图。选择参数后，即可上传地图。随后，用户会返回工具选择页面，并选择创建新场景。此时，系统会要求用户从其有权访问的地图中选择一张。接下来，用户将进入实际的场景生成器页面，地图会显示在右侧。用户可以在左侧列表中设置不同选项，并可将机器人的目标位置和起始位置拖放到地图上，或直接输入 X 和 Y 坐标。此外，用户还会看到下方的下拉菜单，可从中选择要添加到场景中的新动态障碍物类型。对于每个障碍物，都可以通过输入字段更改其速度，并添加新的航点；用户可以在地图上拖动航点进行调整，也可以直接输入坐标值。完成场景设置后，系统会提示用户为场景命名，并选择是否允许其他用户查看和使用所创建的场景。点击上传后，整个流程即告完成，用户可以直接下载该场景，也可以将其用于新的评估。

B. 开始新的训练任务

训练可以进行多种自定义设置，包括选择和指定网络架构、超参数、地图和奖励。由于可选项和组合数量众多，可能会令用户感到不知所措，因此我们设计了一个由四个步骤组成的交互式直观流程，用于选择不同的工具并启动训练。在所有步骤中，我们都提供了多种可公开访问的文件供用户使用，并预设了一些默认值。

1) 选择地图和机器人：进入应用并来到仪表板后，用户可以直接跳转到创建新训练的页面。在那里，系统会提示用户选择用于训练的地图。用户可以从大量现有地图中选择，也可以选择自行生成或上传的地图。选好地图后，用户需要选择运行训练后模型的机器人。根据这一选择，神经网络架构步骤会考虑该机器人的观测和

动作空间，因为所有机器人都有不同的观测空间。

2）创建超参数集：如前所述，训练所需的超参数集应在用户界面中配置。用户界面如图6所示。



图6：训练任务的超参数设置。用户可以从多个参数中进行选择，例如任务模式、训练步长、评估步长等。

第一步是“设置”，用户可以将每个参数设为所需的数值（见图6）。为了让超参数生成过程比直接在输入字段中填写数字或字符串更加直观，界面中显示了按钮和滑块。每个超参数下方都有一段简短说明，介绍其用途。使用滑块和按钮也能避免将参数设为超出范围的数值。对于以按钮选项呈现的参数，只能设为界面显示的数值；滑块则可在一定范围内调整参数。用户进行配置之前，所有超参数均设为预定义数值，如图6所示。

配置超参数后，用户下一步需要在“常规信息”中选择其访问类型：选择公开以启用公开使用，或选择私有。根据这一决定，其他用户可能会看到所创建的超参数。在此步骤中，系统还会要求用户为所创建的超参数命名。这对于避免用户产生混淆非常重要。最后，系统会向用户概述所选的设置和访问类型，用户可以提交超参数。

3) 创建神经网络架构：为简单起见，创建自定义神经网络架构的过程与创建一组超参数类似。神经网络架构生成器同样分为三个步骤。第一步“架构”会引导用户创建网络架构。不过，这部分比超参数生成器更复杂。神经网络架构

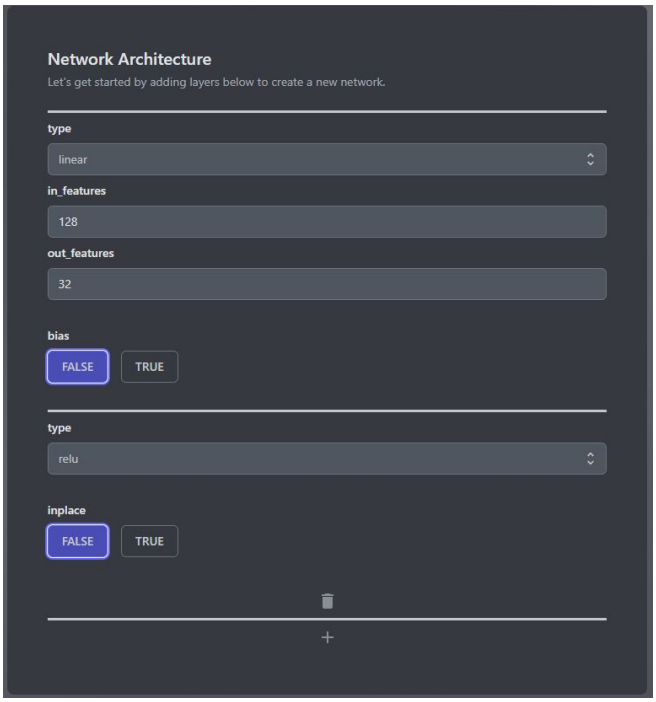

图7：用于定义深度强化学习智能体神经网络架构的用户界面。用户可以定义层的类型、输入和输出大小，以及激活函数。系统会将架构可视化，以便用户更好地理解。此外，根据机器人型号，如果输入和输出大小与机器人的规格不匹配，系统可能会显示错误消息。

在任何时候配置的架构都会保存到一个数组中。用户界面如图7所示。点击添加按钮，用户即可扩展该架构，添加一个新模块，从而在现有数组中增加另一个模块。用户也可以使用垃圾桶图标删除最近添加的模块。模块由类型及该类型所需的参数组成。在此示例中可以看到，线性层的参数包括输入特征、输出特征和偏置。添加新模块时，默认类型为线性层，但用户可以将其更改为所有可用模块中的任意一种。完成架构设置后，用户可以在“常规信息”步骤中选择公开或保持私有，并为其设置名称。在“摘要”步骤中，用户可以再次查看组合使用的模块类型和访问类型，然后将其提交至后端。

4) 指定奖励系统并开始训- 练：最后一步，用户应通过为若干预设惩罚项指定数值，并使用奖励生成器设置奖励系统

。奖励系统已设置默认值，因此用户也可以跳过此步骤并立即开始训练。提供所有必要输入后，用户可以提交训练任务，系统会将一个 JSON 对象发送到后端，其中包含训练名称以及所选机器人、网络架构和超参数的 ID。后端将根据所选配置启动 *Arena-bench* 和训练。训练开始后，用户将被重定向到一个页面，在该页面可以查看训练的日志输出，并下载当前的最佳模型。

C. 开始新的评估任务

网页应用的另一项重要功能是评估导航方法的性能。用户可以利用该应用对多个规划器进行纯粹的基准测试，以相互比较。用户可以选择要记录哪些指标，而已实现的数据记录器只会记录计算这些指标所必需的主题。与训练任务类似，开始新的评估需要选择大量不同的配置，这可能会让用户不知所措。因此，我们的任务编辑器提供了简明易用的操作流程，并明确列出需要执行的步骤。用户可以在仪表板上直接选择创建新的评估任务。第一步是选择场景文件。虽然我们强烈建议使用场景，因为这有助于在多次评估运行之间进行更稳定、一致的比较，但我们也提供了在随机任务上进行评估的选项；在此模式下，障碍物以及机器人在地图上的起点和目标位置均为随机放置。因此，用户可以在第一步中切换随机模式和场景模式。此外，用户还可以指定评估回合的数量。第二步，用户选择用于评估的机器人。根据这一选择，用户必须在第三步中选择一个规划器。用户可以从表 I 所列的多种规划器中进行选择，也可以选择一个使用 Web 应用创建的、经过训练的深度强化学习智能体。最后一步，系统会再次提示用户为评估命名，以便之后更好地识别。

启动评估后，用户会被重定向到一个新页面，在那里可以查看评估的日志输出。日志文件和最新的 .csv 评估文件均可下载。我们以 Jupyter 笔记本的形式提供评估类，其中预先编写的代码和说明可引导用户，并允许用户单独绘制结果。这样，用户便能根据自身需求高度自定义图表。图 9 展示了使用我们的评估类生成的示例图表。未来，我们计划在网页应用中增设图表实时显示功能。评估过程中，我们会呈现大量定量和定性指标

表 II 中提供了相关内容。网页应用所需的工作流程和界面如图 8 所示。

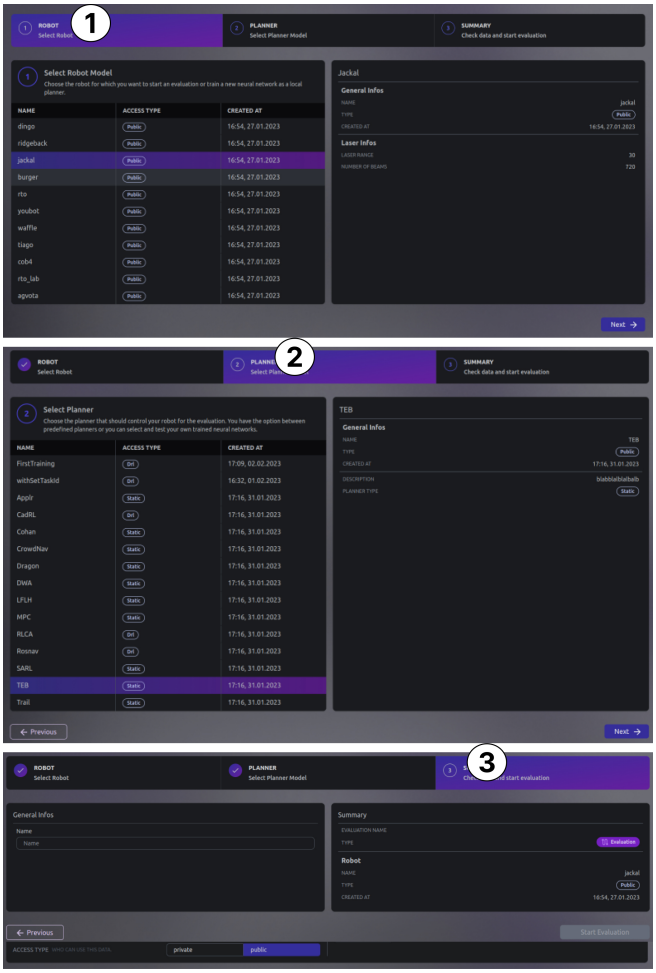

图 8：执行评估运行时的示例流程。该流程包含三个步骤：(1) 选择机器人模型，(2) 选择规划器，以及 (3) 选择评估任务。

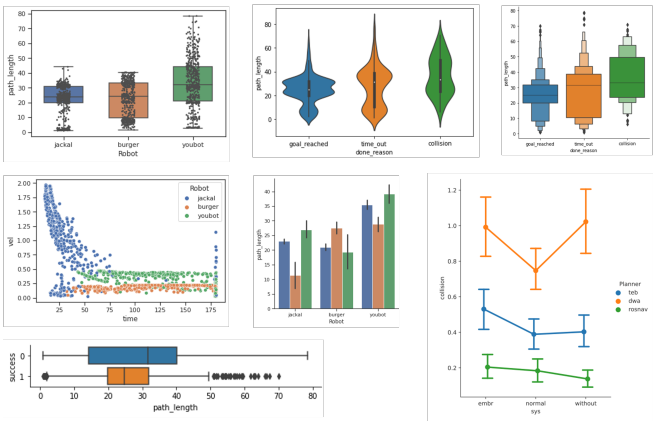

图 9：使用评估 类。

表 II：评估指标

超参数	数值	说明	计算
安全性			
碰撞[1,2]	-	碰撞总数	
障碍物间距 [2]	-	与障碍物的距离	
鲁棒性			
成功率[2]	%	发生碰撞的回合比例 <2 碰撞且未超时	
方差[2]	-	所有测量值之间的方差	$v = \sqrt{\sum_0^N (sN-}$ sN
效率			
路径长度[1,2]	$[m]$	路径长度（米）	
到达目标时间[2]	$[s]$	到达目标所需时间	
平滑度			
运动急动度[2]	$[\frac{m}{s^3}]$	平均加速度	$j = \lvert\frac{\Delta a}{\Delta t}\rvert$，$a = \frac{\Delta v}{\Delta t}$
粗糙度[2]	-	轨迹平滑度倒数	$R(\mathbf{x}_i) = \frac{\text{三角形面积}(\mathbf{x}_i,\ \mathbf{x}_{i+1},\ \mathbf{x}_{i+2})}{\lvert\mathbf{x}_{i+2}-\mathbf{x}_i\rvert^2}$
角度[2]	$[\frac{1}{2}]$	角度与路径长度之比	$c_{norm}(\mathbf{x}_i) = \frac{c(\mathbf{x}_i)}{\lvert\mathbf{x}_{i+1}-\mathbf{x}_i\rvert+\lvert\mathbf{x}_{i+2}-\mathbf{x}_{i+1}\rvert}$

VI. 结论

本文提出了 Arena-Web，这是一个基于网络的应用程序，用于在不同的机器人系统上开发、训练和基准测试基于深度强化学习的导航规划器。因此，该 Web 应用为先进的移动机器人基准测试框架提供了高度开放且易于访问的用户界面。凭借简洁美观、设计精良的用户界面，即使是缺乏经验的用户，也能轻松开始训练和基准测试导航规划器，而不会被应用为资深用户提供的众多功能所淹没，从而让更多人接触移动机器人学，降低入门门槛。Arena-Web 提供直观的工具，可用于生成新地图、场景、神经网络架构、奖励值和超参数，从而支持高度定制化的训练和评估流程。该网页应用提供大量可用的机器人和规划器，以及预定义的地图和场景。评估类包含多种导航指标。采用现代软件框架和结构良好的系统设计，不仅使系统高度安全，也为扩展和添加新功能提供了充足空间。用户可以公开分享自己的数据，从而有望建立一个大型社区，让所有人都能从他人的创作中受益。用户还可以下载所有创建的数据，因此不会受限于 *Arena-Web* 的任何功能，也可以单独使用应用中的小部分功能。未来的工作计划在网络应用程序中加入更多可视化内容和实时视频（例如训练和评估运行的视频），以进一步提升用户的理解和体验。如前所述，目前正在进行的工作包括将绘图功能集成到 Web 应用中，这也将进一步提升易用性和便捷性。此外，我们目前还在开发更多功能，例如机器人建模和多智能体基准测试集成。该网页应用免费使用，并可通过补充材料中提供的链接公开访问。

参考文献

[1] D. Dugas、J. Nieto、R. Siegwart 和 J. J. Chung，“Navrep：用于动态人类环境中机器人导航强化学习的无监督表征”，*arXiv*预印本 *arXiv:2012 04406*，2020。.[2] M. Everett、Y. F. Chen 和 J. P. How，“利用深度强化学习在动态决策智能体之间进行运动规划”，载于*2018*年*IEEE/RSJ*智能机器人与系统国际会议*(IROS)*。IEEE，2018年，第3052–3059页。[3] X. Xiao、B. Liu、G. Warnell 和 P. Stone，“利用机器学习进行移动机器人导航的运动规划与控制：综述”，《*Autonomous Robots*》，第1–29页，2022年。[4] E. Heiden、L. Palmieri、L. Bruns、K. O. Arras、G. S. Sukhatme 和 S. Koenig，“Bench-mr：轮式移动机器人的运动规划基准”，*IEEE*机器人与自动化快报，第6卷，第3期，第4536–4543页，2021年。[5] L. Kästner、T. Bhuiyan、T. A. Le、E. Treis、J. Cox、B. Meinardus、J. Kmiecik、R. Carstens、D. Pichel、B. Fatloun 等，“Arena-bench：高度动态环境中避障方法的基准测试套件”，*IEEE*机器人与自动化快报，第7卷，第4期，第9477–9484页，2022年。[6] N. Tsoi、A. Xiang、P. Yu、S. S. Sohn、G. Schwartz、S. Ramesh、M. Hussein、A. W. Gupta、M. Kapadia 和 M. Vázquez，“Sean 2.0：为机器人导航形式化并生成社交情境”，*IEEE*机器人与自动化快报，第7卷，第4期，第11 047–11 054页，2022年。[7] R. Marin、P. J. Sanz 和 J. S. Sánchez，“用于通过网页远程操控机器人的超高层级界面，包含增强现实功能”，载于*2002*年*IEEE*机器人与自动化国际会议论文集（第*3*卷）。IEEE，2002年，第2725–2730页。[8] R. Siegwart 和 P. Saucy，“网页上的交互式移动机器人”，载于*IEEE*机器人与自动化国际会议*(ICRA)*，1999年。[9] S. Thrun、M. Bennewitz、W. Burgard、A. B. Cremers、F. Dellaert、D. Fox、D. Hahnel、C. Rosenberg、N. Roy、J. Schulte 等，“Minerva：第二代博物馆导览机器人”，载于*1999*年*IEEE*机器人与自动化国际会议论文集（第*3*卷）。IEEE，1999年。[10] L. Kästner、T. Buiyan、X. Zhao、L. Jiao、Z. Shen 和 J. Lambrecht，“将基于深度强化学习的避障部署到传统自主导航系统”，*arXiv*预印本 *arXiv:2104.03616*，2021年。[11] C. Chen、Y. Liu、S. Kreiss 和 A. Alahi，“人群与机器人交互：利用基于注意力机制的深度强化学习实现人群感知机器人导航”，载于*2019*年机器人与自动化国际会议*(ICRA)*。IEEE，2019年，第6015–6022页。[12] A. Faust、K. Oslund、O. Ramirez、A. Francis、L. Tapia、M. Fiser 和 J. Davidson，“Prm-rl：结合强化学习与基于采样的规划完成长距离机器人导航任务”，载于*2018*年*IEEE*机器人与自动化国际会议*(ICRA)*。IEEE，2018年，第5113–5120页。

[13] Y. F. Chen、M. Everett、M. Liu 和 J. P. How，“利用深度强化学习进行社会感知运动规划”，载于*2017*年*IEEE/RSJ*智能机器人与系统国际会议*(IROS)*。IEEE，2017 年，第 1343–1350 页。[14] H.-T. L. Chiang 、 A. Faust、M. Fiser 和 A. Francis，“利用autorl端到端学习导航行为”，*IEEE*机器人与自动化快报，第4卷，第2期，第2007–2014页，2019年。[15] R. Guldenring、M. Görner、N. Hendrich、N. J. Jacobsen 和 J. Zhang，“学习室内环境中用于人类感知导航的局部规划器”，载于*2020*年*IEEE/RSJ*智能机器人与系统国际会议*(IROS)*。IEEE，第6053–6060页。[16] L. Kästner、X. Zhao、T. Buiyan、J. Li、Z. Shen、J. Lambrecht 和 C. Marx，“利用航点生成器将基于深度强化学习的避障与传统全局规划器相结合”，载于*2021*年*IEEE/RSJ*智能机器人与系统国际会议*(IROS)*。IEEE，第1213–1220页。[17] N. Tsoi、M. Hussein、J. Espinoza、X. Ruiz 和 M. Vázquez，“Sean：自主导航的社会环境”，载于第八届人机交互国际会议论文集，2020年，第281–283页。[18] D. Schulz、W. Burgard、D. Fox、S. Thrun 和 A. B. Cremers，“公共场所移动机器人的网页界面”，*IEEE*机器人与自动化杂志，第7卷，第1期，第48–56页，2000年。[19] J. Koch、M. Reichardt 和 K. Berns，“机器人控制框架的通用网页界面”，载于*2008*年*IEEE/RSJ*智能机器人与系统国际会议。IEEE，2008年，第2336–2341页。[20] T. Bhattacharjee、E. K. Gordon、R. Scalise、M. E. Cabrera、A. Caspi、M. Cakmak 和 S. S. Srinivasa，“自主性越高就总是越好吗？探索行动障碍用户对机器人辅助喂食的偏好”，载于*2020*年*ACM/IEEE*人机交互国际会议论文集，2020年，第181–190页。[21] A. Di Nuovo、F. Broz、T. Belpaeme、A. Cangelosi、F. Cavallo、R. Esposito 和 P. Dario，“面向robot-era多机器人服务老年用户的基于网页的多模态界面”，载于*2014*年*IEEE*系统、人类与控制论国际会议*(SMC)*。IEEE，2014年，第2186–2191页。[22] S. Co̧sar、M. Fernandez-Carmona、R. Agrigoroaie、J. Pages、F. Ferland、F. Zhao、S. Yue、N. Bellotto 和 A. Tapus，“Enrichme：家用老年人辅助机器人的感知与交互”，国际社会机器人学期刊，第12卷，第779–805页，2020年。[23] A. Perzylo、M. Rickert、B. Kahl、N. Somani、C. Lehmann、A. Kuss、S. Profanter、A. B. Beck、M. Haage、M. R. Hansen 等，“Smerobotics：面向柔性制造的智能机器人”，*IEEE*机器人与自动化杂志，第26卷，第1期，第78–90页，2019年。[24] F. Del Duchetto、P. Baxter 和 M. Hanheide，“导览机器人Lindsey：博物馆长期部署中的使用模式”，载于*2019*年第*28*届*IEEE*机器人与人类交互通信国际会议*(RO-MAN)*。IEEE，2019年，第1–8页。[25] J. Baˇcík、P. Tkáˇc、L. Hric、S. Alexoviˇc、K. Kyslan、R. Olexa 和 D. Perduková，“Phollower——适用于工业和民用环境、配备符合安全要求的新冠肺炎杀菌附加装置的通用自主移动机器人”，应用科学，第10卷，第21期，第7682页，2020年。[26] J. M. Cañas、E. Perdices、L. García-Pérez 和 J. Fernández-Conde，“基于ROS的智能机器人教育开源工具”，应用科学，第10卷，第21期，第7419页，2020年。[27] L. A. Hormaza、W. M. Mohammed、B. R. Ferrer、R. Bejarano 和 J. L. M. Lastra，“通过虚拟现实对机器人任务进行在线训练与监控”，载于*2019*年第*17*届*IEEE*工业信息学国际会议*(INDIN)*，第1卷。IEEE，2019年，第841–846页。[28] A. Faina、B. Nejati 和 K. Stoy，“Evobot：用于科学实验的开源模块化液体处理机器人”，应用科学，第10卷，第3期，第814页，2020年。[29] R. Bejarano、B. R. Ferrer、W. M. Mohammed 和 J. L. M. Lastra，“实现人机协作装配工作站”，载于*2019*年第*17*届*IEEE*工业信息学国际会议*(INDIN)*，第1卷。IEEE，2019年，第557–564页。[30] M. Quigley、K. Conley、B. Gerkey、J. Faust、T. Foote、J. Leibs、R. Wheeler、A. Y. Ng 等，“ROS：一个开源机器人操作系统”，载于*ICRA*开源软件研讨会，第3卷，第3.2号。日本神户，2009年，第5页。[31] S. Osentoski、G. Jay、C. Crick、B. Pitzer、C. DuHadway和O. C. Jenkins，“机器人即网络服务：使用rosjs实现可复现实验和应用开发”，载于*2011*年*IEEE*机器人与自动化国际会议。IEEE，2011年，第6078–6083页。[32] R. Toris、J. Kammerl、D. V. Lu、J. Lee、O. C. Jenkins、S. Osentoski、M. Wills和S. Chernova，“机器人网络工具：面向云机器人的高效消息传递”，载于*2015*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*）。IEEE，2015年，第4530–4537页。[33] “Webviz”，https://webviz.io/，访问日期：2023-02-03。[34] “Foxglove”，https://foxglove.dev/，访问日期：2023-02-03。[35] C. Rösmann、F. Hoffmann和T. Bertram，“用于时间最优点到点非线性模型预测控制的定时弹性带”，载于*2015*年欧洲控制会议（*ECC*）。IEEE，2015年，第3352–3357页。[36] X. Xiao、B. Liu、G. Warnell、J. Fink和P. Stone，“Appld：基于示范的自适应规划器参数学习”，*IEEE*机器人与自动化快报，第5卷，第3号，第4541–4547页，2020年。[37] O. Khatib，“面向机械臂和移动机器人的实时避障”，载于自主机器人车辆。Springer，1986年，第396–404页。[38] P. Long、T. Fan、X. Liao、W. Liu、H. Zhang和J. Pan，“通过深度强化学习实现最优分布式多机器人避碰”，载于*2018*年*IEEE*机器人与自动化国际会议（*ICRA*）。IEEE，2018年，第6252–6259页。[39] C. Rösmann，“时间最优非线性模型预测控制”，博士学位论文，多特蒙德工业大学，2019年。[40] P. T. Singamaneni、A. Favier和R. Alami，“面向多样化人机交互情境的人类感知导航规划器”，载于*IEEE/RSJ*智能机器人与系统国际会议（*IROS*），2021年。[41] K. Li、Y. Xu、J. Wang和M. Q.-H. Meng，“SARL：基于深度强化学习的室内环境移动机器人人类感知导航”，载于*2019*年*IEEE*机器人与仿生国际会议（*ROBIO*）。IEEE，2019年，第688–694页。[42] L. Kästner、C. Marx和J. Lambrecht，“基于深度强化学习的动态环境移动机器人语义导航”，载于*2020*年*IEEE*第*16*届自动化科学与工程国际会议（*CASE*）。IEEE，2020年，第1110–1115页。