

项目申请书

项目名称: 面向工业制造的具身智能基准测试套件

项目主导师: 郑子木

申请人: 程国亮

日期: 2025.06.03

邮箱: 15179143561@163.com

一、项目背景.....	1
二、项目目标.....	1
三、项目详细方案.....	1
1、具身智能工业数据集	1
2、实施细节	8
2.1 PyBullet 仿真平台	8
2.2 场景构建	9
2.3 数据集	10
2.4 KubeEdge Ianvs 评估算法	11
四、项目研发计划规划	12

一、项目背景

在工业制造智能化加速演进的时代背景下，工业机器人、柔性生产线与智能检测装备持续迭代革新。云边协同技术以其高效的数据处理与实时响应能力，成为驱动具身智能系统在复杂工业场景中落地应用的核心技术引擎。

本项目拟基于 KubeEdge-Ianvs 协同人工智能基准测试框架，构建包含工业场景测试数据集、仿真测试环境与多维度性能指标的完整体系，旨在打造适配工业制造领域的专业化、行业级具身智能测试能力。

值得关注的是，当前工业领域对具身智能服务的需求已向高精度感知决策、动态实时响应、跨设备协同控制等方向纵深发展。但现有的通用具身智能基准测试体系，因缺乏对工业场景独特需求与设备特性的针对性考量，难以满足精准评估需求，已然成为制约工业具身智能技术发展的关键瓶颈。

二、项目目标

项目 issue 地址：<https://github.com/kubeedge/ianvs/issues/197>。

项目成果仓库：<https://github.com/kubeedge/ianvs>。

1. 引入具身智能领域工业数据集，按四大类标准化任务分类并重新整理现有数据集：感知（如多视角表面质检）、移动（如料箱搬运、堆拆码垛）、操作（如精密装配、精密插装、多工序柔性总装、多品类柔性装配、卷盘转运、设备入盒包装、混合物品分拣）、复合（如电子设备柔性测试、电力巡检），输出数据集调研报告；
2. 在 KubeEdge-Ianvs 中选择至少一种上述场景提供标准化测试套件，包括数据集、测试环境、测试指标，以标准化统一数据格式梳理数据集；
3. 在 KubeEdge-Ianvs 中基于标准化测试套件实现具身智能基线算法。

三、项目详细方案

1、具身智能工业数据集

数据集名称	所属场景类型	总结	链接
Agibot World	操作， 移动， 复合	AgiBot World 数据集诞生于智元自建的大规模数据采集工厂与应用实验基地，空间总面积超过 4000 平方米，包含 3000 多种真实物品，一	链接

		方面为机器人大规模数据训练提供场地，另一方面真实复刻了家居、餐饮、工业、商超和办公五大核心场景，全面覆盖了机器人在生产、生活中的典型应用需求。	
ARIO (All Robots in One), 2024	操作, 移动	<p>ARIO 是一个全面的基准数据集，旨在统一不同实施例和任务类型的机器人数据。它涵盖了 20 多个真实和模拟的机器人平台，包括从基本运动到复杂工具使用和操纵的任务。每个机器人都配备了 RGB 摄像头、IMU 和关节编码器等传感器，使代理能够在平台之间推广和转移技能。</p> <p>在多个机器人必须协同操作的工业环境中，例如运输零件的轮式基地和执行装配的手臂，ARIO 的结构反映了这些系统的异质性以及对它们之间统一智能的需求。</p>	链接
Open X-Embodiment, 2023	操作, 移动	100 万片段 × 22 种机器人类型，覆盖 500+ 技能（含工业分拣、装配）。	链接
RH20T-P (Robotic Hands Dataset with Primitive Skills), 2024	操作	对于专注于微装配或小零件操作的行业，如齿轮、螺钉和 PCB，RH20T-P 为教导机器人精细准确的交互提供了宝贵的数据。	链接
ALOHA 2, 2024	操作	ALOHA 2 扩展了原始的	链接

		ALOHA 数据集，以包括更复杂的双臂协调任务。凭借改进的物理真实感和对齐的 RGB-D 视觉流，它模拟了堆叠、折叠和对象对齐等任务。该数据集捕捉了双手操作的复杂性，这在许多包装和码垛环境中至关重要。	
Baxter_UR5_95_Objects_Dataset, 2023	操作	该数据集包括使用 Baxter 和 UR5 机械臂捕获的 95 个工业和家用物体的 RGB-D 图像。它具有 3D 抓取注释、分割蒙版和对象姿势。包含真实的机器人和类似工业的零件使其在物理现实条件下训练操纵模型特别有价值。 其多样性和设置模拟了仓库自动化系统中经常面临的实际拣选、组件分拣和掌握计划挑战。	链接
YCB Object and Model Set	操作	YCB 对象和模型集是由耶鲁-卡内基梅隆-伯克利合作创建的，旨在为机器人操作、假肢设计和康复研究提供基准测试。该数据集包含 73 个日常生活中的物体，具有不同的形状、大小、纹理、重量和刚性，以及一些广泛使用的操作测试。数据集提供了高分辨率的 RGBD 扫描、物理属性和物体的几何模型，便于集成到操作和规划软件平台中。此	链接

		外，数据集还包括一系列标准任务协议框架和示例协议，以及实验实施，旨在量化评估包括规划、学习、机械设计、控制等多种操作方法。	
Human-Robot Collaboration Dataset	操作, 移动	这个合成数据集旨在模拟人类和机器人并肩工作的工业或服务机器人环境中的真实数据。	链接
SmartAssemblySim-V2	操作	SmartAssemblySim-V2 是 BridgeData V2 数据集的概念子集，旨在模拟与工业制造相关的目标条件机器人操作任务。任务包括插入零件、放置物体和与抽屉交互，所有这些都由 RGB 视频、状态数据和目标指令支持。	链接
FPC_Assembly_Project	操作	场景：将柔性印刷电路（FPC）电缆精密组装到智能手机主板上。“精密组装薄软组件”是工业制造领域中极具挑战性和相关性的一个场景。由于组件的可变形特性，这一场景面临重大障碍，因此成为构建专注于具身智能的新数据集的理想选择。使用 PyBullet 作为模拟环境也是一种实用的方法	链接
NEU Surface Defect Database	感知	NEU 表面缺陷数据库是钢铁行业广泛采用的评估表面缺陷分类的基准。它包括 1800 张热轧钢带的灰度图像，分为六类缺陷，如银纹、夹	链接

		杂物、斑块、麻面、轧制氧化皮和划痕。该数据集紧凑高效，适用于快速原型制作和算法测试。它通常用于开发质量保证系统中的模型，在这些系统中，快速缺陷识别至关重要。	
ISDD – Industrial Surface Defect Detection Dataset	感知	该数据集解决了螺母、齿轮和螺钉等小型工业部件中多视图缺陷检测的挑战，该数据集基于MANTA 数据集构建，为每个对象实例提供了五个不同视角的图像。	链接
Severstal: Steel Defect Detection	感知	Severstal 数据集提供了大量 12568 张钢表面灰度图像，并为四种缺陷类型提供了详细的像素级注释。它专为 Kaggle 竞赛而设计，促进了先进分割算法的发展，这些算法不仅可以识别缺陷的存在，还可以精确地定位缺陷。其注释良好的结构使其特别适用于训练在空间精度至关重要的自动化钢铁检测线上使用的深度学习模型。	链接
BSDData: Ball Screw Surface Defect Dataset	感知	BSDData 是一个特定领域的数据集，专注于检查滚珠丝杠传动——数控机床和自动化设备中的关键部件。该数据集拥有 21000 多张 RGB 图像和点蚀缺陷的详细注释，支持分类和检测模型的监督训练。它已被	链接

		应用于评估计算机视觉系统如何识别局部和微妙的表面异常，为工业环境中的预测性维护系统提供了有价值的基准。	
Kolektor Surface-Defect Dataset	感知	KolektorSDD 及其后续 SDD2 是由电换向器的高分辨率图像组成的数据集。每张图像都有注释，以指示是否存在划痕、凹痕和污染等表面缺陷。这些数据集主要用于二元和多类分类，以模拟现实世界的工业约束而闻名，如数据不平衡和微妙的缺陷模式。它们支持对快速检测解决方案的研究，这些解决方案需要对组件表面的最小偏差具有高灵敏度。	链接
Gear Inspection Dataset (GID)	感知	GID 包含带有缺陷标签的工业齿轮灰度图像。该数据集支持工业机器视觉应用中的对象检测和语义分割任务。它对于旨在实时评估齿轮质量的训练模型特别有用，例如在汽车或机械装配线上的在线检查期间。齿轮的聚焦范围使其成为研究高精度环境中特定零件缺陷检测的理想选择。	链接
Real-IAD: Real-world Multi-View Industrial Anomaly Detection Dataset	感知	Real IAD 是一个全面的基准，旨在评估现实工业环境中的异常检测系统。它包括 30 种不同类	链接

		<p>型组件的 150000 张高分辨率图像，每张图像都是从五个不同的视点拍摄的。该数据集涵盖了八种常见的缺陷类型，包括裂纹、凹痕、污染和错位。Real IAD 的设计考虑了无监督的异常检测，反映了高通量生产线的挑战，在这些生产线上，异常很少见，但对于准确识别至关重要。它支持异常定位、检测鲁棒性和多视图检测策略的研究，特别适用于质量保证管道。</p>	
ISP-AD: Industrial Screen Printing Anomaly Detection Dataset	感知	<p>ISP-AD 是一个针对丝网印刷过程中异常检测的领域特定数据集。它包括合成和现实世界中的缺陷示例，如墨水污迹、缺失打印、重影和配准错误。该数据集的结构支持分类和像素分割，使其适用于各种工业缺陷检测管道。其主要应用场景包括印刷电子产品、标签和包装生产中的自动视觉检测，确保一致的印刷质量对产品性能和品牌至关重要。</p>	链接
MVTec AD	感知	<p>MVTec 数据集包含 5354 张不同目标和纹理类型的高分辨彩色图像。它包含用于训练的正常（即不包含缺陷）的图像，以及用于测试的异常图像。异常有 70 种不同类型的缺陷，例如划</p>	链接

		痕、凹痕、污染和不同结构变化。	
KolektorSDD (Kolektor Surface-Defect Dataset)	感知	该数据集由 Kolektor Group doo 提供和注释的有缺陷的生产项目的图像构建。这些图像是在真实案例中的受控工业环境中捕获的。该数据集由 399 张图像组成，大小为 500 x ~1250 px。	链接
RoboMIND	复合	该数据集包含了在 479 种不同任务中涉及 96 类独特物体的 10.7 万条真实世界演示轨迹。 RoboMIND 数据集汇集了多种机器人平台的操作数据，包括 52,926 条 Franka Emika Panda 单臂机器人轨迹、19,152 条 "天工" 人形机器人轨迹、10,629 条 AgileX Cobot Magic V2.0 双臂机器人轨迹、以及 25,170 条 UR-5e 单臂机器人轨迹数据。	链接

目前，根据已经得到的具身智能工业数据集，可以看到部分数据集中所涉及到的场景纷繁复杂，因而在使用相关的数据集时需要提取其中对本项目研究设计工业场景的部分进行项目的进一步实施。

总之，现有的数据集提供了有价值的起点，但工业制造的独特需求以及大型实体模型的计算需求，突显了开发定制化、全面且可能多模态的数据集的关键需求，而这些数据集通常利用现实世界的收集和先进的合成数据生成技术得到。

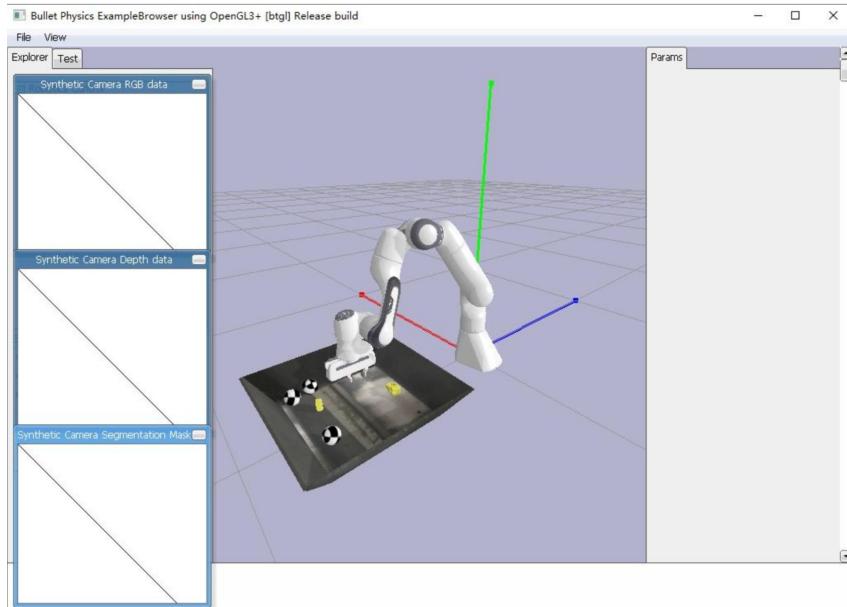
在本项目中的数据集是在仿真平台 PyBullet 上进行仿真以及数据集的获取。

2、实施细节

2.1 PyBullet 仿真平台

PyBullet 基于著名的开源物理引擎 bullet 开发，封装成了 Python 的一个模块，用于机器人仿真和学习。

PyBullet 支持加载 URDF、SDF、MJCF 等多种机器人描述文件，并提供正/逆向运动学、正/逆向动力学、碰撞检测、射线相交查询等功能。除此之外，还提供了不少机器人的例程和实用的调试工具（滑动条、按钮、文本）等。



在本项目中在 PyBullet 中创建了一个全新的数据集，该数据集基于操作类别中薄软组件的精密组装。

2.2 场景构建

场景：将柔性印刷电路（FPC）电缆精确组装到智能手机主板上。

在工业制造领域，“薄软部件的精密组装”是一种极具挑战性的典型场景。这类部件的显著可变形特性，为自动化组装流程带来了一定的障碍——传统刚性物体的建模与控制逻辑难以直接复用，需结合柔性体动力学特性设计针对性算法。

这一类场景的复杂性与工程价值，使其成为“具身智能领域全新数据集的理想候选”：通过构建包含多物理场耦合、接触力学特性的高质量数据集，可有效推动机器人触觉感知、柔顺控制等关键技术的突破。

在仿真验证环节，“基于 PyBullet 的物理引擎仿真平台”为此类高质量数据集的获取提供了低成本、可复现的解决方案。该平台支持柔性体建模、接触力实时计算及传感器数据模拟，能够高效支撑算法开发与策略优化，为薄软部件组装的智能化研究提供重要技术支撑。

2.3 数据集

数据生成策略：使用了一个模拟的机器人手臂（例如，PyBullet 中提供的 UR5 或熊猫机器人模型），该手臂配备了一个简单的夹具和手腕或夹具处的模拟力/扭矩传感器。虚拟 RGB-D 相机将放置在头顶或一定角度以捕捉场景。

数据集内容（针对每个记录的帧/步骤）：

1. 视觉数据：

- `rgb_image_front.png`: 来自观察装配区域的虚拟相机的 rgb 图像。
- `depth_image_front.png`: 来自同一相机的对应深度图像。

2. 机器人状态数据：

- `joint_positions.json`: 机器人手臂当前关节角度的列表。
- `end_effector_pose.json`: 机器人末端执行器的笛卡尔位置 (x, y, z) 和方向 (四元数)。
- `force_torque_sensor_readings.json`: 来自末端执行器传感器的模拟力和扭矩读数（例如，Fx、Fy、Fz、Tx、Ty、Tz）。

3. 地面实况数据（对基准测试至关重要）：

- `fpc_pose_world.json`: fpc 模型在世界坐标系中的真实 6 自由度姿态（位置和方向）。
- `motherboard_pose_world.json`: 主板模型在世界坐标系中的真实 6 自由度姿态。
- `fpc_connector_pose_relativer_motherboard.json`: fpc 连接器相对于主板配合连接器的接地相对变换（位置和方向）。这是对齐的直接目标。
- `assembly_tstage.txt`: 一个离散的标签，指示该框架的组装过程的当前阶段（例如，“预取”、“已拾取”、“接近连接器”、“对齐_解析”、“校准_行”、“插入”、“冲突失败”、“失配失败”）。

4. 场景元数据：

- `scene_id.txt`: 每个程序集尝试/序列的唯一标识符。
- `initial_fpc_pose_variant.txt`: fpc 初始随机姿势变体的标识符。
- `lighting_condition.txt`: 虚拟照明设置的标识符（例如，“正常”、“低照度”、“眩光”）。

确保稳健性的数据集变化：

- 初始 FPC 姿势：通过在工作台上的各种随机位置和方向（在定义的工作区内）启动 FPC 来生成数据。这模拟了取货地点的变化。
- 主板姿势：在工作台上引入主板位置和方向的微小随机变化。
- 照明条件：模拟不同的照明条件（例如，均匀、定向、有眩光源），以测试感知鲁棒性。
- 模拟变形性参数：如果 PyBullet 允许，稍微改变 FPC 模型的刚度或阻尼参数，以表示不同的材料特性。
- 故障模式：有意模拟机器人执行导致错位或碰撞的动作的场景，并将这些数据捕获为“失败”的尝试，这对学习错误恢复很有价值。

数据集大小：

- 目标是数百到数千个独特的组装序列。
- 每个序列将由多个帧组成（例如，每次成功组装 50-200 帧）。
- 这将产生数万到数十万个单独的数据点（帧）。

数据格式：

- 图像：PNG 格式。
- 数值数据（姿势、关节状态、力/扭矩）：JSON 文件（易于解析）。
- 元数据：文本文件或包含在每个序列的摘要 JSON 中。

至此，数据集已经获取完成，为本项目场景“薄软组件的精确组装”定义“PyBullet FlexConnect 组装数据集（PFCA 数据集）”。

2.4 KubeEdge Ianvs 评估算法

在这一步骤中，重点是如何使用具身智能基准测试工具（Ianvs）来测试和评估在 2.3 中创建的 PyBullet FlexConnect 组装数据集（PFCA 数据集）。

使用 KubeEdge Ianvs，这是一个支持边缘/云环境人工智能评估的协作基准测试平台。

架构和相关概念如下图所示。ianvs 设计为在单个节点内运行。关键组件包括：

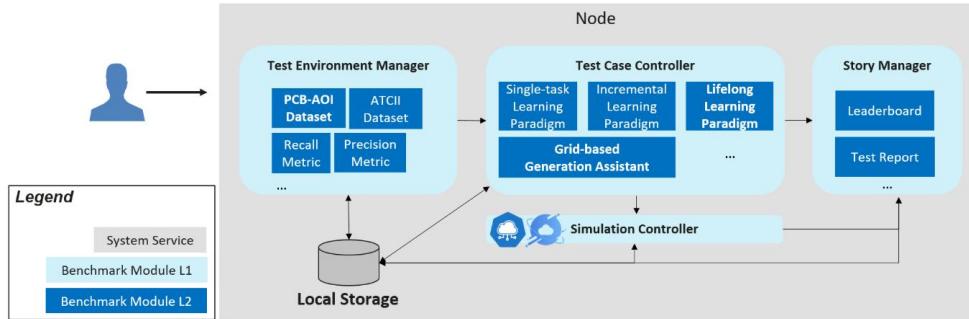
Test Environment Manager: 供全球使用的测试环境的 CRUD

Test Case Controller: 控制测试用例的运行时行为，如实例生成和消失。

Generate Assistant: 协助用户根据一定的规则或约束条件生成测试用例，例如参数范围。

Simulation Controller: 控制边云协同 AI 的仿真过程，包括仿真容器的实例生成和消失，

Story Manager: 测试用例的输出管理和演示，例如排行榜。



具体实施步骤如下：

- 克隆 Ianvs 存储库。
- 创建一个 Docker 容器。
- 安装依赖项和下载 LLM（例如 Qwen2-1.5b、BERT）。
- 将生成好的数据集放置在 ianvs/datasets/ 目录中。
- 准备一个指向测试数据的 YAML 配置文件（fpc_beach.yaml）。
- 在容器内运行 ianvs-f examples/fpc_benchmark.yaml 来评估性能。

四、项目研发计划规划

6 月至 7 月中旬对目前可以得到的具身智能数据集进行调研，输出相应的报告，同时学习 PyBullet 这一仿真环境，学会如何在 PyBullet 仿真环境下搭建工业场景（“薄软部件的精密组装”）并进行数据集的获取。

7 月中旬至 8 月中旬在 PyBullet 仿真环境下，得到相应的数据集，在 KubeEdge Ianvs 构建测试环境、测试指标，并以标准化统一数据格式梳理数据集，同时在 KubeEdge-Ianvs 中基于标准化测试套件实现具身智能基线算法。

8 月中旬至 9 月底，对前面两个阶段的进行总结，思考可以进一步改进或者补充的地方，并输出相应的文档。如果时间和精力允许，考虑在智元数据集 Agibot World 以及智元仿真平台 Genie Sim 进行标准化测试套件，包括指标和示例。