

搜索

写

登记

登录



会员专享故事

对话式无人机：在 Jetson Thor 上集成视觉语言模型以实现边缘 AI

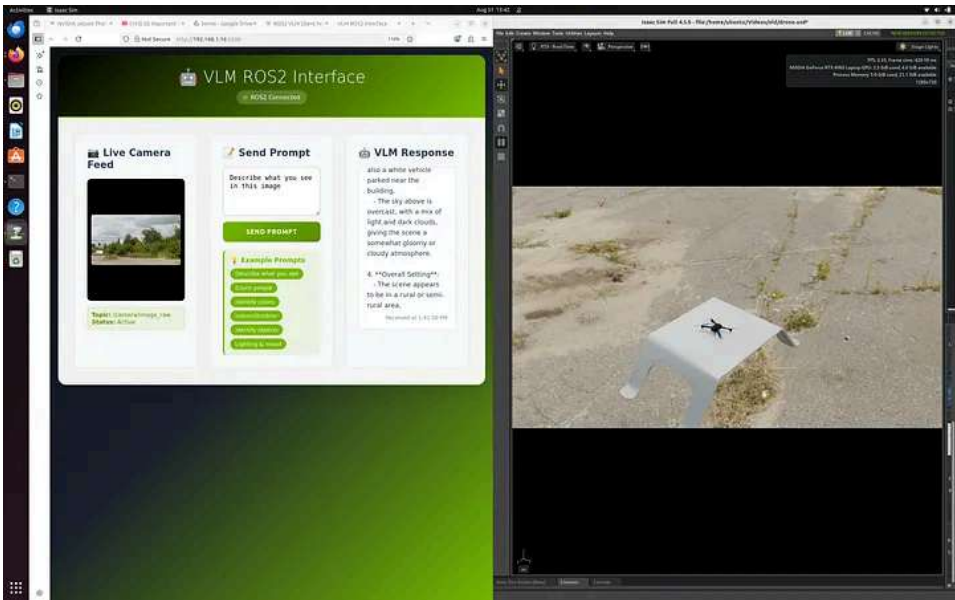


卡比兰克

跟随

5分钟阅读 · 10月 4, 2025

63



艾萨克·西姆：雷神上的无人机和 VLM 特工

想象一下这样一个世界，您可以与无人机交谈，它可以实时理解、观察和响应——所有这些都由边缘的尖端人工智能提供支持。这正是我们使用 **Jetson Thor** 和 **Qwen/Qwen2.5-VL-32B-Instruct-AWQ** 视觉语言模型（VLM）在这个硬件在环设置中构建的内容。

实时自然交互

在我们的演示中，无人机不仅仅是盲目飞行。你可以问它这样的问题：

→ “有森林火灾吗？”

无人机使用其机载传感器观察模拟环境，然后利用 **Qwen2.5-VL** 模型分析它所看到的内容并生成详细的、类似人类的描述——所有这些都 **在 Jetson Thor 上本地处理**。没有云连接，没有延迟 - 只有边缘的实时智能。

这个怎么运作

- 1. 硬件在环模拟
无人机在模仿真实世界条件的 Isaac Sim 模拟环境中运行。这样可以安全地测试高级人工智能功能。
- 2. 视觉-语言模型分析
Qwen2.5-VL模型解释无人机拍摄的图像并生成自然语言响应，将视觉和语言无缝连接。
- 3. 边缘 AI 处理Jetson
Thor 在本地处理所有计算，提供低延迟、高性能的推理，适用于毫秒级的现实场景。

为什么这很重要

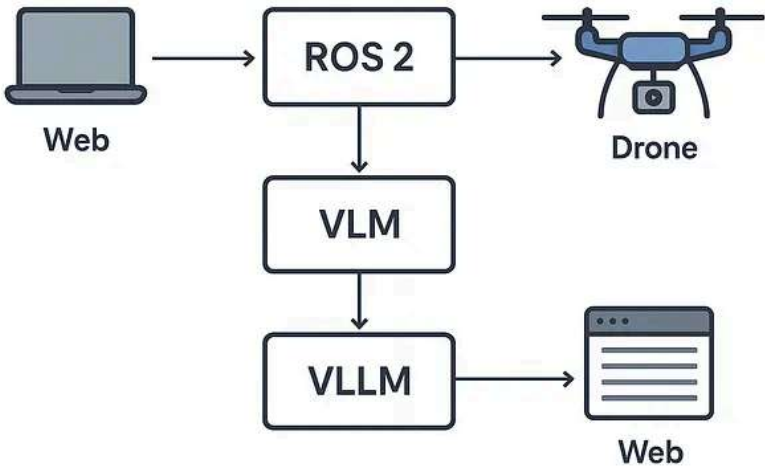
该演示不仅仅是一个技术展示，它还指向了自主系统的未来：

- 灾难响应：快速识别火灾、洪水或其他危险并传递可作的信息。
- 环境监测： 实时跟踪野生动物、森林砍伐或污染。
- 安全作：增强通信延迟可能代价高昂的关键任务的态势感知能力。

推出边缘 AI VLM 无人机系统

要运行由 Jetson Thor 提供支持的无人机与 Qwen2.5-VL-32B-Instruct-AWQ 视觉语言模型交互的演示，您需要在单独的终端中启动多个组件。以下是分步指南：

■



蓝图

在 Jetson Thor 上启动边缘 AI VLM 无人机系统（带 Docker）

要运行由 Jetson Thor 提供支持的无人机与 Qwen2.5-VL-32B-Instruct-AWQ 视觉语言模型交互的演示，您需要启动多个组件。您可以本机执行此作，也可以将 Docker 用于隔离环境。

0 (可选) 启动 Docker 容器

这可确保正确设置所有依赖项。替换为您的 Docker 镜像名称： `vllm-ros2-image`

```
sudo docker run --gpus all --rm -it \  
  --ipc=host \  
  --net=host \  
  -v ~/ros2_ws:/root/ros2_ws \  
  --name vllm-ros2-container \  
  vllm-ros2-image:dev
```

进入容器后，您可以像往常一样运行以下所有命令。

1 启动 ROSBridge (WebSocket 接口)

这允许 Web 客户端与 ROS2 主题进行通信：

```
ros2 launch rosbridge_server rosbridge_websocket_launch.xml
```

2 启动相机桥

该节点将模拟图像转换为 Web 界面可用的格式：

```
ros2 run vlm_ros2_package simple_camera_bridge
```

3 启动 VLM 客户端节点

VLM 客户端处理发送查询和处理文本响应：

```
ros2 run vlm_ros2_package vlm_client_node
```

4 启动 Web 界面

启动基于 Web 的界面，以便您可以通过浏览器与无人机交互：

```
cd ~/ros2_ws/src/vlm_ros2_package/web_interface  
python app.py
```

5 启动 VLLM VLM 服务器

最后，使用 Qwen2.5-VL 模型启动视觉语言模型服务器：

```
./run_vllm_vlm_serve.sh Qwen/Qwen2.5-VL-32B-Instruct-AWQ
```

完善！基于此，以下是我们为什么使用 **vLLM** 进行推理的 **Medium** 就绪解释，以清晰简洁的方式集成您的详细信息：

5 启动 VLLM VLM 服务器

```
./run_vllm_vlm_serve.sh Qwen/Qwen2.5-VL-32B-Instruct-AWQ
```

我们使用 **vLLM** 来服务视觉语言模型（VLM），因为它是一个快速、灵活、高效的 **LLM** 推理和服务库，专为大型模型设计。这就是为什么它是边缘人工智能机器人的理想选择：

1. 高性能推理：

vLLM 使用优化的 CUDA 内核、FlashAttention 和推测解码提供最先进的吞吐量和低延迟执行，非常适合实时无人机交互。

2. 内存效率：

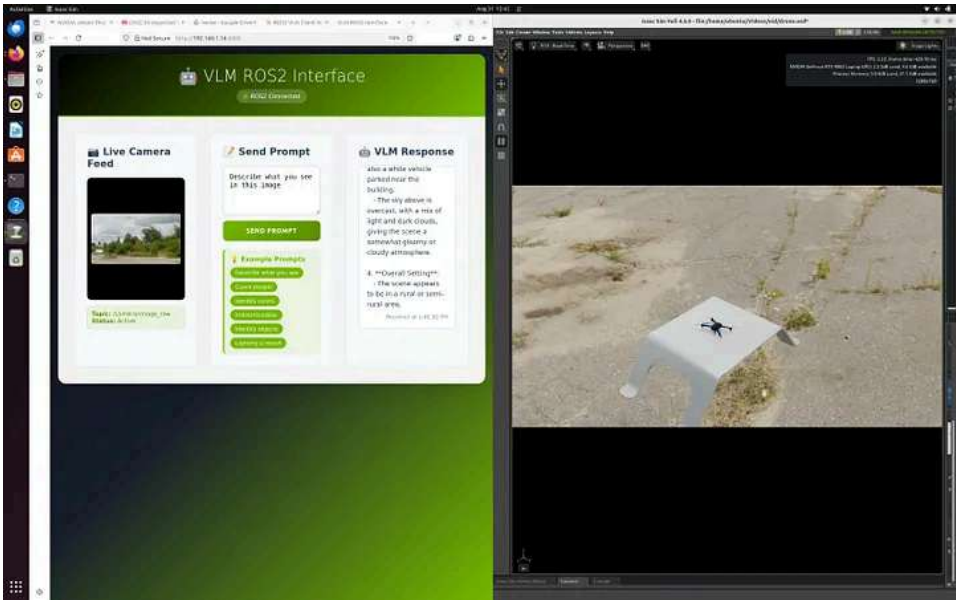
PagedAttention 和量化（GPTQ、AWQ、INT4/INT8/FP8）等先进技术允许像 **Qwen2.5-VL-32B** 这样的大型模型在有限的 GPU 内存上运行，例如 Jetson Thor。

3. 灵活性和兼容性：

它与 **HuggingFace** 模型无缝集成，支持流媒体输出、OpenAI 兼容 API，并可跨各种硬件平台工作，包括 NVIDIA GPU、AMD GPU、Intel CPU/GPU，甚至 TPU。

vLLM 使您的无人机（或任何机器人）能够处理视觉输入、生成自然语言响应并实时智能地行动，同时在边缘设备上高效运行。

✅ 在所有终端（或容器步骤）运行后，您的无人机就可以完全运行了——它可以观察环境，使用 **VLM** 处理查询，并实时提供自然语言响应。



启动系统后

一旦所有终端（或 Docker 步骤）都运行完毕，您的无人机就可以完全运行了——它可以观察环境，使用 **VLM** 处理查询，并实时提供自然语言响应。

💡 超越无人机：此演示不仅限于无人机。同样的方法可以应用于不同类型的机器人，使它们能够看到世界、理解世界并根据观察采取智能行动。

结论

通过将 **Isaac Sim** 中的硬件在环仿真与 **Jetson Thor** 和视觉语言模型集成在一起，该演示展示了用于自主无人机的逼真、安全和高性能的测试环境。该系统允许无人机感知其环境、理解复杂场景并实时智能响应——所有这些都在边缘本地处理。

这种方法不仅加速了先进机器人算法的开发和验证，还凸显了可扩展、智能自主系统的潜力。无论是无人机还是其他机器人平台，将仿真、边缘人工智能和 **VLM** 智能相结合，为更智能、更安全、更自主的机器人带来新的可能性。

- 英伟达
- 杰森
- 机器人
- 机器人
- 英伟达 GPU

63



作者：Kabilankb
427 位粉丝 · 12 位关注者
英伟达大使

跟随

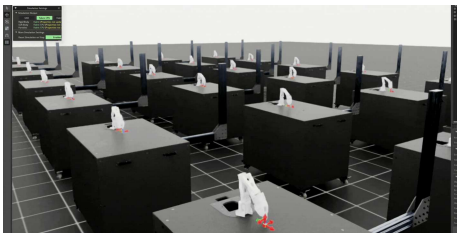
还没有回复



写一个回应

你有什么想法？

更多来自 Kabilankb

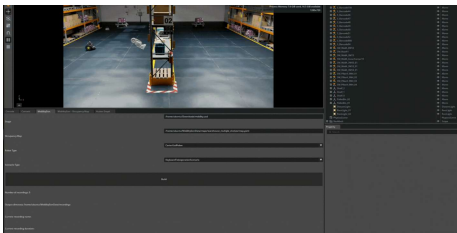


卡比兰克

在艾萨克实验室训练SO-100立方体提升机器人：从仿真到智能控制...

艾萨克实验室和 SO-100 立方体提升任务简介

5月1日 25



卡比兰克

在NVIDIA Isaac Sim的MobilityGen中对自定义机器人进行子类化和注册

开发自主移动机器人（AMR）需要大量高保真训练数据——合成数据.....

9月21 3 2

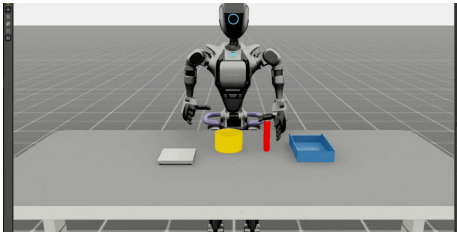


卡比兰克

NVIDIA Jetson Nano 配备英特尔实感深度摄像头，使用 ROS2 谦虚

在本教程中，我们将探讨如何使用 ROS2 Humble 将 NVIDIA Jetson Nano 与英特尔实...

5月17, 2024 69 2



卡比兰克

Jetson AGX Thor 上的 GROOT N1: 物理 AI 的硬件在环


介绍

8月29日 4

查看卡比兰克的所有产品

从中推荐




 在 编码关系 由 代码包

IBM Granite 4.0：可以在本地或浏览器中运行的小型语言模型（SLM）

Granite-4.0-Micro 模型是一个紧凑的、34 亿参数的指令模型，效率很高。

10月3日 4




 在 人工智能进步 由 Himanshu Soni

我在 8GB GPU 上运行了一个 800 亿参数的 AI 模型 - 我对 oLLM 的体验

作为一个为我的项目创建人工智能相关内容并尝试新工具的人，我想探索大型语言模型...

5天前 157 3




 在 SDG集团 由 贡萨洛-萨巴拉

在事件驱动场景中使用 MCP 进行代理文档分类 — 架构概述

在过去的几个月里，我一直在做一个实验项目：将事件驱动处理（Kafka）与 MCP 服务...

10月3日 2





 Linux 指南


设计人工智能和机器人技术以及未来的可能性

人工智能和机器人技术的融合代表了工程领域的变革前沿，突破了工程领域的界限.....

9月7

 彻底改变大上下文 LLM 推理：深入探讨 oLLM Python 库


 为什么银行和区块链在暗中押注 Rust

 在 口袋里的数据科学 由 赛迪拉吉-古马迪

Revolutionizing Large-Context LLM Inference: A Deep Dive into...

In the rapidly evolving world of AI, running large language models (LLMs) with massive...

Oct 1 64 2

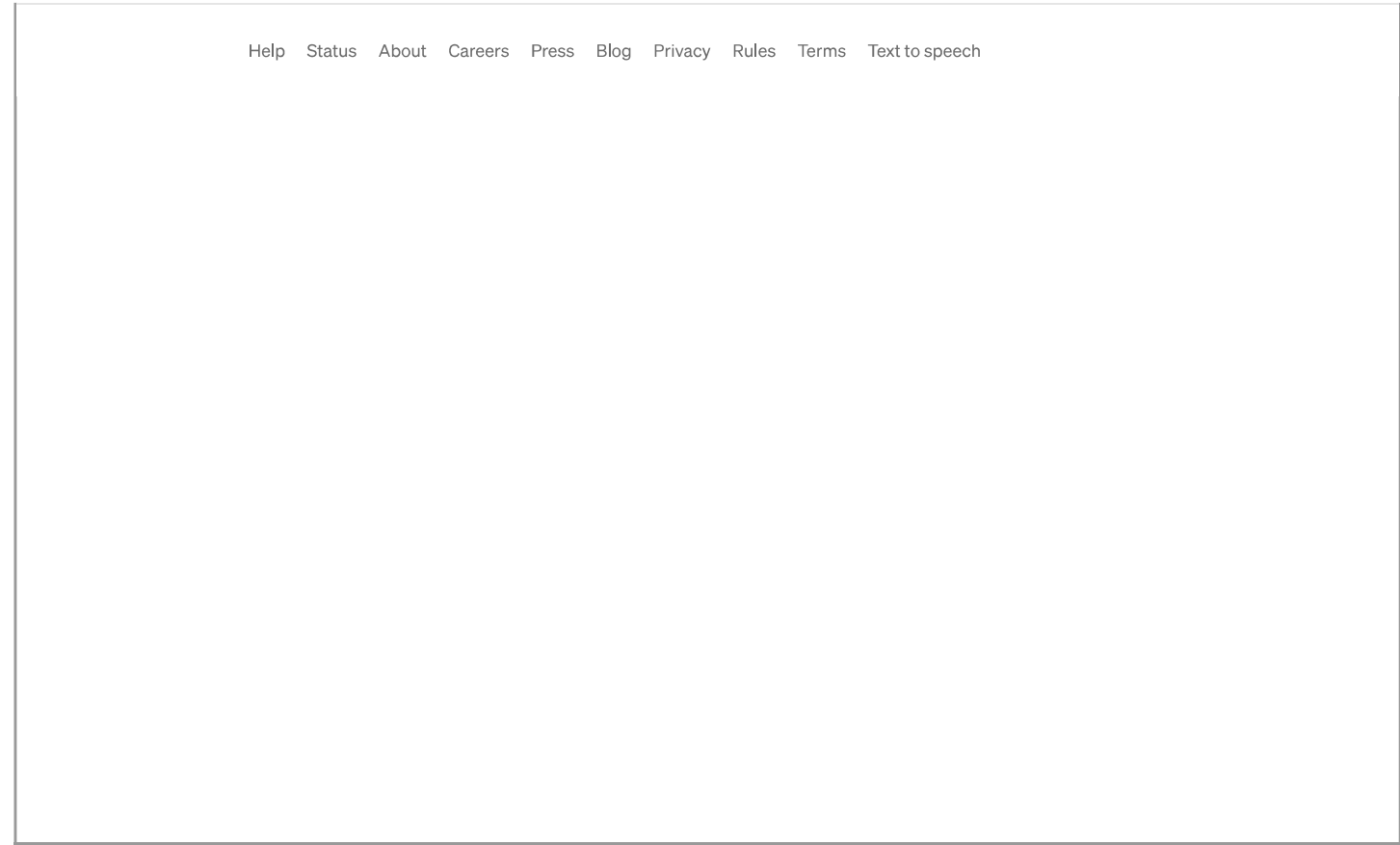
 语法野蛮人

Why Banks and Blockchains Are Secretly Betting on Rust

And why the language you thought was “just for systems programming” is quietly rewritin...

Sep 4 49 2

See more recommendations



[Help](#) [Status](#) [About](#) [Careers](#) [Press](#) [Blog](#) [Privacy](#) [Rules](#) [Terms](#) [Text to speech](#)