

# 泰山盘道上机器狗引网民关注 速度很快! 能驮人吗?

“泰山真是越来越有‘科技范儿’了!”“操控很灵敏,平衡感特别好。”“速度很快啊!”“如果能驮人,我第一个报名体验!”日前,泰山盘道上出现了一道独特的爬山身影——身负重物和清运垃圾的机器狗,引发网民关注。

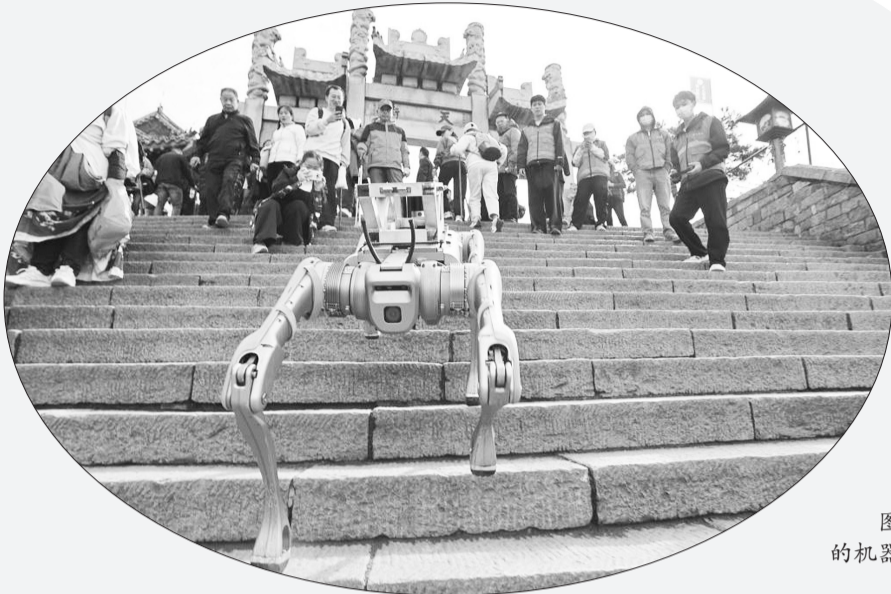
通过视频可以看到,一群机器狗正背着物资在泰山盘道上奋力攀爬,这些机器狗驮着沉重的物资,却丝毫没有显露出吃力的模样,仿佛是一支训练有素的“科技驼队”。

据介绍,这是泰山文旅集团物业公司测试的用来清运垃圾和搬运货物的机器狗。今年以来,负责泰山垃圾清运工作的泰山文旅集团物业公司积极探索科技赋能环卫工作,这次测试有望找到一条解决山岳型景区垃圾清运难题的新路径。

位于山东省中部的泰山被称为“五岳之尊”,每年吸引数以百万计的海内外游客,随之产生的垃圾也逐渐增多。

泰山文旅集团物业公司负责人马玉飞告诉记者,2023年泰山景区游客量达到创纪录的862万人次,年产生垃圾近2.4万吨,垃圾清扫、运输难度、运输成本大大提高。

长期以来,由于泰山景区地形复杂,且均为登山盘道,无法实现大规模机械化作业,清扫垃圾只能通过人工挑运。客流高峰时段,上下山游客占用了全部盘道,只能将垃圾暂存盘道两侧,待游客减少时再进行突击集中清运,这已经成为困扰景区多年的老问题。



图为正在泰山进行测试的机器狗。 新华社发

“我们结合目前科技发展的最新成果,借鉴外地景区的先进经验,最终选定引进机器狗在泰山景区开展垃圾清运测试,以此提高景区货物、垃圾运输效率。”马玉飞说。

据了解,参加测试的机器狗是目前市场上奔跑速度最快、技术最为先进的工业级四足机器人之一,具有翻越障碍能力强、稳定性和平衡能力突出、负重能

力和续航能力出色等性能优势。

业内人士指出,机器狗的出现不仅为泰山垃圾清运提供了新的思路,也为这座世界名山增添了新的魅力。“现代科技与泰山的古老文化相互交融,形成了一种独特的景观,不仅为泰山的旅游发展带来了新的机遇,也为泰山文化的传承和创新提供了新的思路。”山东大学教授王忠武说。 来源:新华网

## 多模态之重新定义人机交互方式

未来的人工智能(AI)什么样?想象一下,只需简单一个指令,它们便能领悟并执行复杂的任务;它们还能通过视觉捕捉用户的表情和动作,判断其情绪状态。这不再是好莱坞科幻电影中的场景,而是正逐步走进现实的“多模态AI”。

据美国《福布斯》网站近日报道,元宇宙平台公司、OpenAI以及谷歌公司等巨头,都推出了各自的多模态AI系统,正不遗余力地加大对这类系统的研发投资,力求提高各种模态内容输出的精确度,从而改善AI与用户的交互体验。

多模态AI标志着一种范式变革。它将深刻改变很多行业的面貌,并重塑数字世界的格局。

### 赋予AI“多重感官”功能

人类是如何了解世界的?我们依赖视觉、听觉和触觉等多种感官,从无数来源接收信息。人脑将这些纷繁复杂的数据模式融合,绘制出一幅生动的现实“画卷”。

IBM公司官网这样定义多模态AI:能集成和处理来自多种模态(数据类型)的机器学习模型,这些模态包括文本、图像、音频、视频等形式的输入。就像赋予AI一整套感官,使它能从多个角度感知并理解输入的信息。

这种跨越不同模态理解和创建信息的能力,超越此前侧重于集成和处理特定数据源的单模态AI,赢得了各大科技巨头的青睐。

在今年的移动通信大会上,高通公司将其开发的多模态大模型首次部署在安卓手机上。用户无论是输入照片,还是语音等信息,都能与AI助手顺畅交流。例如,用户可以拍一张美食照片向AI助手提问:这些食材都是什么?能做出什么菜?每道菜的热量是多少?AI助手能基于照片信息,给出详细的回答。

今年5月,OpenAI发布了多模态模型GPT-4o,其支持文本、音频和图像的任意组合输入和输出。随后,谷歌也于第二天推出了自己的最新多模态AI产品Gemini 1.5 Pro。

9月25日,元宇宙平台公司发布了其最新的开源大语言模型Llama 3.2。公司首席执行官马克·扎克伯格在主题演讲中表示,这是该公司首个开源多模态模型,可同时处理文本和视觉数据,标志着AI在理解更复杂应用场景方面取得了重大进展。

### 悄然推动各领域变革

多模态AI正悄然改变着多个领域的面貌。

在医疗保健领域,IBM旗下“沃森健康”正对病人的影像学数据、病历文本和基因数据进行综合分析,帮助医生更准确地诊断疾病,有力支持医生为病人制订个性化治疗方案。

创意产业也正在经历一场变革。数字营销专家和电影制片人正借助这一技术打造定制内容。试想,只需一个简单的提示或概念,AI系统

就能编撰出引人入胜的剧本,生成故事板(即一系列插图排列在一起组成的可视化故事)、创作配乐,甚至制作出初步场景剪辑。

教育和培训领域也在多模态AI助力下向个性化学习迈进。美国纽约顿公司开发的自适应学习平台能利用多模态AI,深入分析学生的学习行为、表情和语音,实时调整教学内容和难度。实验数据显示,这种方法能将学生的学习效率提高40%。

客户服务也是多模态AI系统令人兴奋的应用之一。聊天机器人不仅能回应文本查询,还能理解客户的语调,分析客户的面部表情,并用适当的语言和可视化线索作出回应。这种更接近人类的交流有望彻底改变企业与客户互动的方式。

### 仍需克服技术伦理挑战

但多模态AI发展也面临诸多挑战。

AI咨询公司“隐空间”创始人亨瑞·艾德尔表示,多模态AI的强大之处在于能够整合多种数据类型。然而,如何有效整合这些数据仍是一个技术难题。

此外,多模态AI模型在运行过程中往往需要消耗大量算力资源,这无疑增加了其应用成本。

更值得注意的是,多模态数据包含更多个人信息。当多模态AI系统能轻松识别人脸、声音乃至情绪状态时,如何确保个人隐私得到尊重与保护?又该如何采取有效措施,防止其被用于创建“深度伪造”或其他误导性内容?这些都是值得深思的问题。

来源:科技日报

