

一种脑控四足机器人：概念验证

吴可*

通号通信信息集团有限公司北京轨道交通分公司，中国·北京 100071

摘要：将脑机接口（Brain-Computer Interface, BCI）与机器人系统相结合，有望在未来实现无缝的日常个人助理系统，使各类请求仅依靠个体的脑活动以离散方式完成。这类系统对闭锁综合征（Locked-in Syndrome, LIS）或肌萎缩侧索硬化（Amyotrophic Lateral Sclerosis, ALS）人群尤其具有潜在意义，因为他们可通过脑信号感知接口与机器人助理进行交流从而获益。在本项概念验证工作中，我们探索了一种无线可穿戴 BCI 设备如何控制一台四足机器人——波士顿动力 Spot 四足机器人（Boston Dynamics Spot quadruped robot，以下简称 Spot）。该设备通过嵌入眼镜镜框的电极测量用户的脑电图（Electroencephalography, EEG）与眼电图（Electrooculography, EOG）活动。用户通过执行心算类脑动（Brain Teaser）任务来回答一系列问题的“是/否”，从而作出响应。每一个“问题—答案”对都为 Spot 机器人预先配置了一组动作。例如，当一段序列被判定为“是”时，Spot 机器人会被指令穿过房间、拾取物体并为用户取回（如带来一瓶水）。我们的系统达到了 83.4% 的成功率。据我们所知，这是首个在个人助理场景下，将无线、非视觉范式的 BCI 系统与 Spot 机器人集成的实例。尽管该 BCI 四足机器人系统尚处于早期原型阶段，未来的迭代有望具备类似常规服务犬那样友好且直观的提示方式。因此，本项目旨在为在真实日常生活条件下，依托无线可穿戴 BCI 系统驱动的当代个人助理机器人之未来发展铺设道路。

关键词：脑机接口；机器人；个人辅助；人工智能；脑电图；眼电图；可穿戴设备

A brain-controlled quadruped robot: proof of concept

Wu Ke*

Beijing Rail Transit Branch, China Railway Signal & Communication Co., Ltd., China Beijing 100071

Abstract: Combining brain-computer interfaces (BCI) with robotic systems holds the potential to achieve seamless daily personal assistant systems in the future, enabling various requests to be completed solely through an individual's discrete brain activity. Such systems are particularly meaningful for populations with Locked-in Syndrome (LIS) or Amyotrophic Lateral Sclerosis (ALS), as they can benefit from interacting with robotic assistants via brain-signal perceptual interfaces. In this proof-of-concept work, we explored how a wireless wearable BCI device can control a quadruped robot—the Boston Dynamics Spot quadruped robot (hereafter referred to as Spot). The device measures the user's electroencephalography (EEG) and electrooculography (EOG) activity through electrodes embedded in the glasses frame. Users answer a series of questions with 'yes/no' responses by performing mental arithmetic tasks, thereby providing input. Each 'question-answer' pair pre-configures a set of actions for the Spot robot. For instance, when a sequence is determined to be 'yes,' the Spot robot is instructed to move across the room, pick up an object, and deliver it to the user (such as bringing a bottle of water). Our system achieved a success rate of 83.4%. To our knowledge, this is the first instance of integrating a wireless, non-visual BCI system with the Spot robot in a personal assistant scenario. Although this BCI quadruped robot system is still an early prototype, future iterations are expected to feature friendly and intuitive prompts similar to those of conventional service dogs. Therefore, this project aims to pave the way for the future development of contemporary personal assistant robots driven by wireless wearable BCI systems under real daily living conditions.

Keywords: Brain-computer interface; Robots; Personal assistance; Artificial intelligence; Electroencephalography; Electrooculography; Wearable devices

0 引言

机器人产业正在兴起，机器人已在诸多环境中积极参与：博物馆^[1]、时装秀^[2]、酒店^[3]、课堂^[4]、仓库^[5]以及军事应用场景^[6]。尽管军事与仓储等用例正伴随数百万美元投资被积极探索并报道，但能够专门满足用户个体需

求、并可通过搬运日用品、给植物浇水或打开门来辅助用户的个人助理移动机器人几乎不存在。此外，这些机器人的大量控制方式几乎完全依赖语音或触摸界面，而无法通过这些方式操控机器人的用户，则被提供一种利用自身脑活动来控制目标系统的选项。这些系统更常见的称谓是脑

机接口 (Brain - Computer Interface, BCI)。在许多情况下,患有闭锁综合征 (Locked-In Syndrome, LIS) 或肌萎缩侧索硬化 (Amyotrophic Lateral Sclerosis, ALS) 等疾病的用户,最能从使用 BCI 控制机器人这一能力中获益。然而,为这些用户设计的大多数系统均涉及轮椅或外骨骼控制^[7,8]。此外,在多数情况下,为了控制轮椅,用户仍需依赖注视或视觉刺激的存在——要么通过眼动追踪或眼电图 (Electrooculography, EOG) 直接测量,要么通过事件相关电位 (Event-Related Potentials, ERPs) 这种对刺激 (事件) 在大脑特定区域产生的激活进行间接利用^[9]。同时,在多数情况下,这些 BCI 系统配套的是承载电极的凝胶帽,或属于植入式方案,因此由于涉及手术成本,使得这种交互在真实世界中鲜有尝试,或难以轻松规模化^[8]。

本文旨在通过展示一套由可穿戴、无线 BCI 控制的助理机器人概念验证 (proof-of-concept, PoC) 系统来弥合上述差距。PoC 系统采用 (1) 非视觉范式的心理任务作为控制方式,以及 (2) 一台能够为用户搬运物品、并可搬来椅子或打开门的移动机器人。与外骨骼不同,此类方案可部署于偏好 / 需要保持不移动的用户。此外,我们希望确保系统的 (1) 与 (2) 两部分均具备移动性、无线性与自主性。因此我们选择使用一副在镜框内嵌入脑电图 (EEG) / 眼电图 (EOG) 电极的无线眼镜^[10,11]。我们认为这种外形规格既新颖又有前景,同时对用户而言也是侵入性最低的方式。我们还选择了波士顿动力 Spot 四足机器人 (Boston Dynamics Spot quadruped robot, 以下简称 Spot) 作为原型^[12]。佩戴该眼镜时,用户通过 iPhone,对一系列问题以“是 / 否”作答,作答方式为进行“脑动”式的心算活动。每个“问题—答案”对都为 Spot 机器人预先配置了一组动作。例如,当某次序列判定为“是”时,Spot 机器人会被指令穿越房间到达房间的另一处、拾取一个物体并将其取回给用户 (例如带来一瓶水)。系统实现的整体准确率为 83.4%。尽管在撰写本文时 Spot 机器人的价格为 130 万美元,但就我们所需的机动性而言,它是唯一能够代表用户有效执行重载任务、并能在人类中心化环境中承担其他任务的敏捷机器人。在机动性需求方面,它也是一个独特的解决方案。据我们所知,这是首次在个人助理用例背景下,将无线、非视觉范式的 BCI 系统与 Spot 机器人集成的实例。

本文其余部分的组织如下:我们将简要描述 BCI 与机器人领域的当前研究现状及最常用用例,随后对系统基础架构进行深入概述,介绍任务与研究的初步结果,最后强调面临的挑战、局限与未来工作。

1 研究现状

1.1 机器人应用中使用的 BCI 范式

在过去 20 年中,研究者积极探索利用非侵入式脑电图 (EEG) 来控制不同的机器人与神经假体设备的可能性,EEG 通过放置在头皮上的电极捕获脑信号^[13-17]。迄今为止,EEG 仍然是最实用、最具应用性的非侵入式 BCI 方法之一,因为其他成像手段 (如功能性磁共振成像 (fMRI)、脑磁图 (MEG) 与正电子发射断层成像 (PET)) 价格昂贵^[18]、技术要求高且不易便携,尽管近期已有大量关于降低其成本并引入便携式方案的进展报道。

BCI 系统可由内源性 (自发性) 或外源性 (诱发性) 信号进行控制。在外源性 BCI 中,当个体注意于外部刺激 (如视觉或听觉线索) 时会产生诱发信号。这种方法的优势包括极简训练以及可达 60 bits/min 的高比特率^[19,20]。然而,用户需要始终关注刺激,这限制了其在真实生活条件下的适用性。此外,使用外源性 BCI 时用户可能很快产生疲劳。典型例子包括稳态诱发电位 (SSEP) 与 P300 的调制^[18]。相对地,在内源性 BCI 中,控制信号独立于任何外部刺激而产生,用户可按需完全自主执行。这种方式对具有感觉障碍的用户同样有用,并提供更自然直观的交互途径,因为用户可以自发地向系统发出指令^[18]。其通常需要更长的训练时段,且据报道比特率更低。示例包括运动想象 (MI) ^[21]。然而,运动想象被认为是一项复杂且在认知上负荷较大的任务^[22],因此研究者正积极探索用于 BCI 控制的其他心理策略,如音乐想象^[23]、音素想象^[24]、视觉想象^[25]、心理旋转、心算以及词语联想^[26]。

在本文中,我们探索心算——这是一种需要针对问题进行特定心理加工、但不同于外源性任务无需刺激存在的心理任务,因此有时被称为“脑动”^[26]。对于该任务,易度量表在多次会话中呈上升趋势,这表明与 MI 等策略不同,练习也可能降低该任务的负荷^[26]。据我们所知,我们尚未发现一种在训练与控制任务中均采用“脑动”策略且两者都无需视觉刺激的脑控机器人系统 (见图 1)。大多数论文采用被认为认知负荷较高的 MI 范式,而另有若干工作采用 SSVEP 或 P300 这类视觉范式。因此,我们认为探索其他非视觉范式可能有利于脑控机器人控制,因为这将提升设备本身的可穿戴性 (无需将电极定位在前运动皮层与运动皮层上、无需帽式设计,头戴式设计即可);“脑动”类任务通常不需要任何视觉训练,且对参与者的认知负荷更低^[26]。

1.2 由 BCI 赋能的机器人应用

最常见的利用 BCI 的机器人应用通常面向需要辅助的人群,常见形式包括轮椅与外骨骼。图 1 展示了机器

人与 BCI 的最新研究进展概览。我们进一步推荐读者参考 Zhang 与 Wang 的综述^[27], 该文系统回顾了由运动想象 BCI 控制的机器人, 包括移动机器人与机械臂。

1.2.1 轮椅

智能轮椅通常具备从辅助到康复的不同功能。脑控机器人系统可进一步扩展这些用例, 使其也能为 LIS、ALS 等人群所用。Galán 等人^[28]提出了一种基于 64 通道 EEG 的 BCI, 将左手想象运动、静息与词语联想等心理任务相结合, 以沿预设路径驱动真实与仿真的轮椅。部分患者也可以使用混合 BCI 直接操纵轮椅转向, 而不依赖自主导航。在 Wang 等人^[9]的研究中, 研究者将运动想象、眨眼与 P300 组合成混合 BCI, 与 4 名健康被试一起用于驱动轮椅前进或后退。Tonin 等人^[21]亦表明, 脊髓损伤参与者可以通过训练, 在 MI 范式下控制轮椅。尽管这些系统有潜力帮助 ALS 或 LIS 人群, 但目前仍多局限于实验室或医院环境。

为适配真实世界部署, 多个实验室现正转向使用电极数量最小化、并将信号处理托管在移动应用上的 BCI 系统, 以在移动化场景中提供与以往相同的轮椅控制功能。在 Banach 等人^[7]的研究中, 一名健康用户佩戴含 4 个 EEG 电极的系统来操纵轮椅。用户尝试按一个由“码本”确定的连续序列进入注意集中状态; 该码本根据达到的集中次数指示轮椅向前、向后、向左或向右移动。

所选案例具有代表性, 体现了研究方向, 其中 (A)^[29]与 (F)^[8]——采用植入体在 MI BCI 范式下控制肢体与外骨骼; (E)^[30]——迄今为止最新的论文, 采用混合 MI + SSVEP + EMG 系统; (B)^[16]、(D)^[31]与 (G)^[32]——均使用相同的帽式设备服务于不同的机器人用例, 并采用 MI BCI 范式; (C)^[33]——采用 P300 接口; 最后, (H)^[34]——据我们所知唯一采用四足机器人并使用 SSVEP BCI 范式的案例。在上述所有系统的训练中, 所有用户都需要视觉范式的训练方案。该图还说明了佩戴 BCI 的用户在便携性与舒适性方面的挑战, 而本文提出的方案 (I) 是唯一真正可穿戴的形态, 搭建时间不超过 2 min。

Tang 等人设计了一种基于 P300 的脑控全向轮椅, 并配备机械臂, 旨在用于脊髓损伤的康复。基于实时识别模块, 患者可以对场景中的目标物体触发预设的解决方案。例如, 当患者打算进入下一间房间时, 轮椅可以靠近门并展开机械臂打开门^[31]。该系统使用图像跟踪、Kinect 相机与 P300 来检测患者在居家环境中希望执行的动作。该系统具有日常使用的潜力, 因为它不受限于预编程动作, 并使用目标检测来响应患者环境中的刺激。类似的系统还包括 Na 等人^[35]的工作, 他们提出了一种带有混合刺激器的

嵌入式轻量化 SSVEP-BCI 电动轮椅。

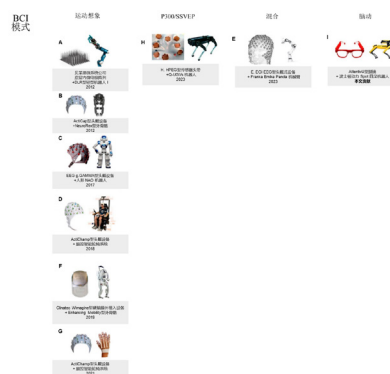


图1 截至 2023 年 BCI 与机器人领域最新进展的概览

1.2.2 外骨骼

相较于轮椅, 外骨骼有望提供更灵巧的功能, 近年来各类外骨骼被积极开发与研究^[28]。外骨骼是与人类用户进行紧密物理交互的可穿戴机器人。其中, 上肢与下肢动力外骨骼被视为可辅助截瘫与四肢瘫痪人群的潜在技术。Benabid 等人^[8]报道了一套概念验证系统, 采用半侵入式硬脑膜外 ECoG 植入与神经假体外骨骼, 利用躯体感觉-运动皮层输入来增强全身运动能力。例如, 两套 ECoG 记录器 (各 64 个电极) 被植入一名受试者大脑的上肢感觉运动区。该受试者首先能够通过游戏控制虚拟化身, 随后在执行心理任务时能够控制四肢外骨骼。这类系统当前的主要缺点之一是外骨骼本身的成本, 以及患者术后所需的大量康复时间。此外, 与轮椅控制类似, 许多研究实验室正专注于探索更廉价、可扩展的非侵入式解决方案。Lee 等人^[36]提出了一种 BCI 系统, 用于基于记录用户感觉运动皮层区域上的 EEG 信号实现对下肢动力外骨骼的实时控制。作者训练了一个二分类解码器, 以 MI 区分两种不同的心理状态, 并以级联方式将其应用于三种方向的外骨骼控制: 向前行走、向左转与向右转。五名受试者在穿戴外骨骼的情况下, 能够通过脑信号完成三路导航任务。

1.2.3 类人机器人

脑控类人机器人被作为帮助 LIS 患者的潜在解决方案进行探索。Spataro 等人^[33]报道邀请了四名 ALS 患者与四名健康对照, 使用 NAO 类人机器人以实现取放一杯水的目标。使用基于 P300 的“奇异范式”, 用户所选命令由机器人执行。尽管 NAO 属于类人机器人, 但其体型较小, 限制了其可达的姿态与位置。此外, 它无法搬运较重物体、打开门等。Chae 等人^[37]同样使用 NAO 构建了由异步 BCI 控制的类人导航系统。用户可以将机器人在室内迷宫中导航至目标位置。在穿越迷宫各阶段的过程中, 用户会看到由机器人视角拍摄的照片。用户佩戴 EEG 帽, 通过左手、右手或前进的 MI 命令控制其运动。Daeglaue 等人^[38]强调,

当在用于控制机器人运动的 MI 任务中纳入个体差异、动机与社会情境时，类人机器人有助于优化神经反馈。Choi 与 Jo^[39] 提出了混合 BCI 协议，融合 P300、SSVEP 与事件相关去同步化（ERD）以协调机器人控制。Sorbello 等人^[40] 展示了四名患有闭锁综合征的 ALS 患者与四名健康对照在类人机器人导航上的初步结果。NAO 类人机器人可以沿六个方向行走并执行抓取 / 递交物体。所有用户都能够导航该机器人，而 ALS 患者在视觉注意与意图任务上的表现优于健康对照。

1.2.4 机械臂

Collinger 等人于 2012 年^[41] 与 Wodlinger 等人于 2015 年^[42] 分别报告了两项研究，受试者为同一名四肢瘫痪女性，在其左侧运动皮层植入了两组 96 通道皮层内电极阵列。第一项研究于 2012 年开展^[30]，报告了对假肢（7 个自由度）的连续平移、姿态与一维抓取控制。第二项研究于 2015 年发表^[42]，通过同时提取手型指令扩展了控制信号的范围。研究发现对手型的稳健神经调谐，导致在所有测试中的十维（10D）表现均显著高于随机水平。

Hockberg 等人^[29] 还展示了四肢瘫痪者在其中枢神经系统损伤或脑卒中发生多年后，如何通过植入的 96 通道皮层内微电极阵列来控制一只假臂。患者能够在指定的机器人工作空间中操控机械臂用吸管喝咖啡。为使 BCI 控制的手部外骨骼运动训练方案更易获得，Araujo 等人^[32] 以轻量低成本的 3D 打印手部模型构建了概念验证系统 HERO。一名健康受试者被提示执行以伸展与屈曲五指为目标的 MI 任务来想象手部动作，并以从其初级运动皮层（M1）非侵入记录到的 EEG 控制符合人体工学的假手。

Zhang 等人^[30] 提出了 NOIR（Neural Signal Operated Intelligent Robots），这是一种通用脑—机器人接口，使人类能够利用脑信号控制机器人完成日常活动，例如做三明治、熨衣服以及抚摸机器狗。三名健康受试者使用 128 通道的 EGI 系统对该系统进行了测试。研究以 SSVEP + MI + EMG 实现了从 EEG 信号解码人类意图目标的模块化流程：（a）操纵哪个物体，由 SSVEP 信号解码；（b）如何与物体交互，由 MI 信号解码；（c）在何处交互，由 MI 信号解码。系统使用捕捉牙关紧咬所致肌肉张力的安全机制来确认或否决议码结果。实验使用了两台机器人：Franka Emika Panda 机械臂与 PAL Tiago 机器人。

1.2.5 四足机器人

Jia 等人^[43] 提出了一种脑控四足运动的协议，理论上可结合非侵入式 EEG 与 MI 使用。用户通过想象五种类型的运动，为机器人的相关肢体创建类似本能的运动轴。该

BCI 协议旨在帮助潜在开发者将 EEG 信号与受指令集约束的机器人实时屈伸运动进行映射，因为该领域仍处于早期阶段。

总体而言，四足机器人常用于在复杂工作环境或国防应用中为用户提供支持。最著名的四足机器人之一是波士顿动力的 Spot 机器人，其可携带最高 15 kg 的载荷，并可对隧道等维护场所进行迭代建图^[12]。地产与采矿行业也在采用类似 Spot 的四足机器人，以协助在物流复杂的作业现场开展监测^[12]。

本研究选择 Spot 作为机器人平台，原因在于其能够执行带有重载的多样任务，并可开 / 关门等。我们认为，Spot 机器人是目前最接近用于 ALS 或 LIS 患者助理机器人的候选者，而且此类场景迄今尚未被探索。

据我们所知，迄今关于用 BCI 控制四足机器人的报道仅有一例。Faisal 等人^[34] 报道了使用来自受试者枕叶区域的 8 通道 EEG 来控制 Spot 机器人。作者将 HoloLens 2 作为既承载传感器又作为显示屏的设备，以显示代表八种可能命令选项的闪烁视觉标签，从而使用户能够在 SSVEP 范式下进行选择。

作者报告了一名用户接受 24 次试验的实验结果。尽管这项工作颇具趣味，但其基于 HoloLens 2 与 SSVEP，而在我们的研究中，我们探索了使用内源性 BCI（心算）进行机器人控制的机会，以便为无法使用眼动的用户提供潜在的控制方式。

下面我们介绍系统的体系结构，我们称之为 Ddog：一种基于心算任务的移动 BCI 方案所控制的 Spot 机器人。

2 Ddog 的基础架构

该系统的目标是设计一个自主应用，使用户能够基于通过 BCI 提供的输入控制 Spot 机器人，并通过应用以语音向用户及其照护者提供反馈。

该系统被设计为可完全离线或完全在线运行。在线版本具备更先进的一组机器学习（ML）模型及更优的微调模型，同时保护本地设备的电池续航。该系统被设计为可在生产环境中运行，同时允许对大多数组件进行快速迭代。系统概览见图 2 与图 3。

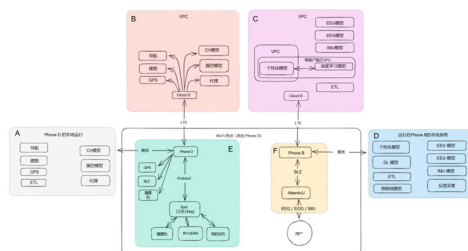


图2 Ddog 系统的体系结构



图3 整个系统的概览

在客户端，用户通过一款移动应用与 BCI 设备交互（本文场景下为 AttentivU 设备（Cambridge, MA, USA）[10,11]，其详细介绍见第 4 节），该应用通过低功耗蓝牙（BLE）协议与设备通信。用户的移动设备（Phone B）与另一部用于控制 Spot 机器人的手机通信，以实现代理、操控、导航，并最终为终端用户提供辅助。两部手机之间的通信可以通过 Wi-Fi 或长期演进（LTE）协议进行。在 Phone D 上建立一个 Wi-Fi 热点，Ddog 通过 Wi-Fi 连接到该热点。Phone B 亦通过相同的 Wi-Fi 连接到 Phone D。同时，也可以在线方式运行系统，并独立连接到云端运行的模型。

对于用于敏感生物数据的人工智能（AI）模型与用于机器人和代理系统的 AI 模型，系统在关注点上做了分离。

2.1 服务器端

系统以与云平台无关的方式设计。它运行在 Kubernetes（K8S）集群中，集群内部为称作 pod 的专用容器。每个集群部署在其各自的虚拟私有云（VPC）中，且 Cloud B 为每个租户（tenant）设置独立的 VPC。

2.2 Cloud D

该云在专用 VPC 中工作，通常部署在同一可用区，尽可能靠近终端用户，从而最小化各服务的响应时延。

该集群由一组容器（pod）构成，每个容器被设计为服务单一目的。

在某种意义上，这是微服务架构：每个服务都是一个正在运行的 AI 模型，其中多数是 ML 模型，亦包含少量 DL 模型，用于基于每位终端用户的地点与情境特定数据进行模型微调。

这些 pod 负责：（1）导航；（2）建图；（3）计算机视觉；（4）操控；（5）定位；以及（6）代理。

与此同时，多数服务被设计为可在“边缘”运行，在此语境中，“边缘”指部署在现场的设备（移动电话、Spot

机器人、BCI 设备 AttentivU）。

下面简要介绍各服务及其功能。

建图——从不同来源收集有关机器人周围环境的信息。它既对静态不可移动的数据（树、建筑、墙）进行建图，也收集随时间变化的动态数据（车辆、行人）。

导航——基于前述服务所收集与增强的建图数据，负责在时空中构建从 A 点到 B 点的路径。它还负责在先前路径失效或发生变化时构建备选路径，并评估与各路径相关的成功率以及机器人在给定载荷（机器人可能携带的额外物体）下通过该路径所需的预计时间。

计算机视觉——处理来自机器人相机并由移动电话数据增强的视觉数据，用于生成时空表示。该服务还尝试对各视觉点进行分割并识别目标。

操控——依赖建图、导航与计算机视觉，使用户关于抓取与搬运特定物体（如水瓶）或采取特定姿态（不仅是位置，还包括下文所述机械臂的姿态）的意图得以实现。

定位——将全球定位系统（GPS）与室内定位数据进行唯一映射，并结合来自机器人传感器的补充数据：深度点云、激光雷达（LIDAR）、红外（IR）发射等。定位服务对建图服务起到补充作用，并提供一组通用知识，用于将导航与房间名称、街道名称与地标等条目关联起来。

代理——基于主体性原则，并最终使 Ddog 成为一个主动助理。每个 agent 会基于用户意图收集并分析数据，同时利用公共知识中的可用数据，也依托于整个系统聚合的私有数据。agent 负责与用户交互，并调用不同的情景，如“去厨房”“拿起一瓶水”“带来一份报纸”。

2.3 Cloud B

Cloud B 负责与 BCI 相关的模型训练，具体包括若干不同模型：EEG、EOG 与惯性测量单元（IMU）。

与 Cloud D 的示例相同，Cloud B 的前提是复刻在 Phone B 上运行的系统离线版本，并以更快的训练管线与节能来加以补充。

以下是各模型流水线的简要概览：

EEG 模型——基于 ML 模型聚合特征。本文所用的模型在第 4 节中描述。

EOG 模型——作为每位用户生物数据的附加侧面，与 EEG 模型配合使用。

IMU 模型——聚合头部与设备运动的数据，其用途与 EOG 模型相同，用于增强 EEG 模型。

深度学习模型——聚合上述三类预训练模型（EEG、EOG、IMU），以为每个用户个性化提升准确率与精度。

ETL 流水线——持续地、并依据现行地区关于生物数据存储与处理的法规，将数据聚合并分发至各对应模型。

推理——当各模型完成训练后，即部署用于执行推理并处理来自用户的实时请求。每个租户的推理均在独立的 VPC 中部署，并针对时延与准确率进行优化。

2.4 客户端

客户端实际上运行在两部手机上（Phone B 与 Phone D；见图 4）。每部手机分别以更低速、节能的方式复刻与 Cloud B 和 Cloud D 相对应的功能。

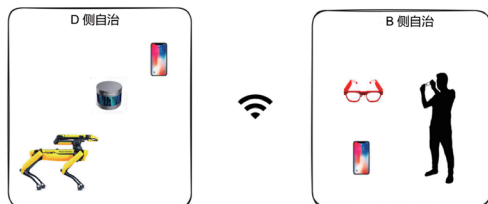


图4 Cloud B 与 Cloud D 的可视化表示

Phone D 还会采集 GPS 等相关数据，并通过其 Wi-Fi 热点功能提供网络承载，充当 Ddog 与 Phone B 之间的中介。

Phone B 负责使用 BLE 与 AttentivU 设备通信，并依据正在运行的 agent 的状态解析并处理用户意图。

在离线模式下，由于硬件限制，代理功能非常有限，而当部署在 Cloud D 上时性能最佳。

下面是我们实现的两个 agent：

导航代理——执行导航类意图，如前往厨房或客厅。

操控代理——负责识别物体、拾取物体，并将该物体带回给用户。

2.5 计算机视觉

部署在手机本机的离线模型负责数据采集与汇聚，并使用针对较小 RAM 占用与基于 ARM 的 CPU 优化的 TensorFlow 移动端模型进行实时推理。

我们在 Cloud D 中运行了更大的 TensorFlow 模型，以更高的准确率为目标进行优化，但代价是更大的 RAM 占用与显著更高的能耗。

2.6 数据采集

部署在移动电话上的离线系统可利用部署在 Phone D 上的 ML 模型，自动完成预定义情形的导航。

我们还探索了波士顿动力（Waltham, MA, USA）团队提供的自治功能，具体为基于 Spot 栈之上的 Map Data 与 Graph Nav API，用以基于现有算法构建决策树。尽管其针对保护设备与尽最大能力完成目标做了优化，但开箱即用的 API 与功能仍落后于移动电话与云端所能提供的功能。因此，在牺牲部分可靠性的前提下，我们决定默认采用比波士顿动力所提供方案更先进的技术栈，以服务工业类用例。

2.7 分割

我们最初部署分割模型所用的版本依赖于单一的

TensorFlow 3D 模型，利用 LiDAR 数据。随后，我们将模型扩展为小样本模型，并辅以运行神经辐射场（Neural Radiance Fields, NeRF）与 RGBD 数据的补充模型。我们还采用了 TensorFlow 与 PyTorch，并编写了性能更优的 C++ 与 Rust 自定义优化以部署在 Phone D 与 Cloud D 上。

2.8 深度、IR、LiDAR

Ddog 采集的原始数据由五个相机汇聚而成。每个相机都可提供灰度、鱼眼、深度与红外数据。另外在机械臂夹持器内部还配有第六个相机，具备 4K 分辨率与 LED 功能，以便使用 TensorFlow 的现有预训练模型在足够分辨率下进行目标检测。

这也是我们用来自 Phone D 的图像数据进行补充的原因之一，Phone D 也可提供最高 4K@60 FPS 或 HD@280 FPS 的分辨率。

同时，由于我们在手机上采用了 Apple 硬件，我们原生利用了 LiDAR 数据采集来获取深度与点云。最低硬件需求为 iPhone 12 Pro、iPad Pro 11（第三代）或 iPad Pro 12.9（第五代）（Apple, Cupertino, CA, USA）。

2.9 建图

在 Phone D 上离线提供了基础建图功能；然而，Cloud D 版本会汇聚全部可用数据，并为预定义位置构建地图。

2.9.1 点云

点云由 Ddog 与 Phone D 生成的 LiDAR 数据与 RGBD 数据构成。数据收集完成后，会在单一坐标系下归一化，并与聚合全部成像与三维定位数据的全局状态对齐。

2.9.2 成像

如上所述，最佳图像来自 Phone D 与 Ddog 的机械臂区域，因为它们能够以 1080p@30 FPS 提供实时数据，因此我们的模型针对这一特定画质做了优化。

当 Ddog 未能执行导航与操控指令时，我们会采集 4K 质量的附加数据，以改进特定位置的模型。

2.9.3 GPS

得益于 Phone D 的硬件，我们能够采集维护与全局定位所必需的 GPS 坐标。

在公寓场景中它的作用较小，但在足够分辨率（1 m 精度）下，我们也能够将 GPS 数据用于建图。

2.10 操控

操控完全依赖于安装在 Ddog 上的机械臂夹持器的品质（见 Figure 5）；我们使用的是波士顿动力生产的夹持器，其针对管道与通风口等工业用例进行了优化，尽管如此，我们仍找到方法将其用于民用的个人助理场景。

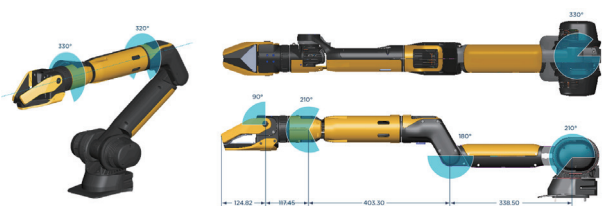


图5 Spot机器人的机械臂与夹持器

2.10.1 抓握

在不同环境下，针对超过 40 种不同物体测试夹持器与操控功能后，我们意识到，要构建可靠的夹持器，唯一方法是预训练模型，使其能够处理我们预期用户会交互的熟知物体。

为满足本文目标，我们将一处大型实验室空间映射并布置为“公寓”，包含“厨房”区域（放置有不同杯子与瓶子的托盘）、“客厅”区域（带靠垫的小沙发与小茶几）以及“窗边休闲区”。这三个区域见图 6。我们使能了对椅子的操控序列，也使能了对毛绒玩具的操控序列。最后，我们使能了 Spot 机器人的拍照功能。

关于夹持器使用的最大经验体会是：针对不同检测到的物体，在灵敏度与施力参数上采用不同设置，并记录由机械臂执行的操作序列。我们敦促读者在尝试复现此类功能时务必谨慎，因为不可避免地会损坏不同物体（塑料瓶、杯子、玩具）；因此，必须确保构建反馈回路，并在 UI/UX 中始终包含便捷的“紧急停止”按钮。这一点在波士顿动力的 API 中已可开箱即用，因为它是基础的机器人安全措施，我们强烈建议任何情况下都不要忽视。



图6 一名佩戴 AttentivU 眼镜的用户即将执行心算类心理任务，以便将 Spot 机器人从“客厅”空间发送至“厨房”空间

得益于代理功能，我们可以更好地理解用户意图；通过更深入地把握用户预期要做的事情，我们构建了与各类物体交互所需的多种操作序列。

举一个实际的例子：思考如何移动一把椅子以便坐下，或将其挪开，或者抓取一瓶水并在不损伤瓶体的情况下搬运它。系统的最终目标是在水瓶被掉落时能够找到或拾起它，或依据用户的状态在恰当的時刻与时间递送水瓶。

用例数量在持续增长，因此覆盖多数用例的唯一方式，是部署一个可持续运行的系统，并利用数据来优化此类操作序列与使用体验。

还需强调另一点：务必关注所使用设备的温度工作模式。如果在手机及 Spot 机器人本体上持续运行 ML 模型，会发现处理质量可能显著下降，因为各自的操作系统（OS）会尝试限制 CPU/GPU 的周期以降低温度。

在理想条件下，Spot 机器人的电池应能持续约 90 min。在我们的实际实验中，在实验室降温环境且电池充满的情况下，我们获得了 60 min 的有效性能。在未开启空调的房间内，我们获得了约 30-40 min 的有效性能。

成像传感器的质量会影响我们训练的计算机视觉模型的基线，这些模型使用了波士顿动力提供的现成 LiDAR 与 RGBD 相机，以及 Apple iPhone 与 iPad 提供的相机（RGB、IR、LiDAR）和 Raspberry Pi。

尽管构建环境的高清地图对实现本文目标至关重要，但我们最终发现，在空间内构建适当数量的闭环成为更大挑战，尤其当机器人必须在未知环境中首次自主完成该过程时（例如我们更换了房间），且有时用户难以及时提供情境信息，最终导致机器人出现最初的错误，并通过来自人类反馈的强化学习（Reinforcement Learning from Human Feedback, RLHF）闭环进行自我纠正。

2.10.2 操控模型

我们构建了一个受 RT-1 模型启发、基于 Transformer 架构的模型。然而，我们必须对该模型进行调整，以考虑机械臂基座——亦即 Ddog 本体——的机动性。

2.11 通信

我们在系统设计中预设：系统在任何时刻都可能离线，且任一部件都可能发生故障或错误。Spot 机器人的 API 允许我们在 Spot 机器人无法与任何外部系统通信，或需要其自主得出所需结果时，较为容易地构建安全措施。

构建安全的自治能力是任何涉及机器人的项目的目标，但这同时也是一项挑战。我们强烈建议确保系统的所有组件都能在任何时刻停止，并可从任一现有状态恢复运行。

2.11.1 Wi-Fi 与 LTE

大多数数据在 Ddog、Phone D 与 Phone B 之间通过 Wi-Fi 进行交换，该方式可为多路 1080p 流提供足够带宽。

LTE 用于以高效视频编码（HEVC）编解码器传输优化后的 1080p 流，以减少流量，但代价是 Phone D 的能耗更高。

Phone D 的 LTE 接入是与 Cloud D 的 VPC 通信的主通道，Phone B 的 LTE 接入用于与 Cloud B 的各 VPC 通信。

2.11.2 BLE

在整个系统中关注能耗至关重要。BLE 对 Phone B 本身的能耗负载最小，且仅在系统需要理解用户意图时使用。

2.12 能耗

在理想条件下，Spot 机器人的电池可运行最长 90 min。配合充电系统，续航可延长至 4 h，这也是 AttentivU 设备当前电池系统的上限。手机的续航更长——在我们的场景中，iPhone 12 与 13 Pro Max（Apple, Cupertino, CA, USA）在质量与温度达到显著阈值前，能够运行 ML 与 DL 模型约 8 h。

2.12.1 离线模式

系统的各组件可不依赖其他部件独立工作,从而为实现多种失效保护提供更高的安全性,原因可能包括环境、硬件、网络、用户相关等。

各组件都会检查网络上其他组件的存在情况，因此当任一组件重新出现（例如重启、更换电池或连接中断后恢复）时，主要组件会尝试重新连接，并与各组件与设备建立新的会话。

自治控制器部署在 Phone D 之上，因为它可能成为整个系统的单点故障。若系统面向用户连续使用部署，应避免这一风险。最简单的解决方案是将网络层（Wi-Fi 热点）与应用本身分离，同时为部署的手机数量提供冗余，这在手机形态下较易实现。

2.12.2 遥测

鉴于连接性、带宽与能耗方面的既有限制，我们不会总是立刻将遥测数据回传至 Cloud B 或 Cloud D。所有敏感数据均采用端到端（end-to-end, E2E）加密，链式密钥对云端 VPC 可用。在生产部署中，我们需要设计“降采样”的遥测，以汇聚边缘用例所需的数据，并支持对系统故障的调试。

3 BCI 实施与结果

3.1 AttentivU

EEG 数据采集自 AttentivU 设备。AttentivU 眼镜在国际 10 - 20 电极放置系统中于 TP9 与 TP10 位置配置由天然银制成的电极。该眼镜还包含两个位于鼻托处的 EOG 电极以及一个位于 Fpz 位置的 EEG 参考电极（见 Figure 7）。关于该眼镜架构的更多信息，请参见[10,11]。



图7 AttentivU 眼镜 (左) 与 EEG 电极位置组合图 (右)

AttentivU 眼镜包含 2 个 EEG 通道, 分别为 TP9 与 TP10, 并在 Fpz 位置设置参考电极。其还配备 EOG 通道以及内置音频与触觉反馈。

该设备的主要优势在于其可携性与易用性：为用户佩戴该眼镜所花费的时间少于 2 min。该设备还具备两种模式（EEG 与 EOG），二者对捕捉当前现象（例如一项主动的心理任务）都极为有用。多篇研究论文报道，EEG 与 EOG 传感器可作为实时捕捉注意、投入、疲劳与认知负荷的信号。EEG 已被用作觉醒与睡眠之间转换的神经生理学指标^[44-46]，并已在需要注意与警觉性的不同环境中得到测试^[47,48]。EOG 基于测量由于角膜-视网膜偶极特性在眼动过程中诱发的生物电位信号。研究者表明，眼动与完成特定任务所需的记忆访问类型相关，并且是衡量视觉投入、注意与困倦的良好指标^[49,50]。多项研究指出，不同模态的信号可表征隐性心理状态的不同侧面^[51-54]。EEG 与 EOG 分别能够表征内部认知状态与外部潜意识行为。这两种模式包含互补信息，结合使用时，可基于最新方法构建对警觉、注意与投入更为稳健的估计模型^[10]。在我们的任务情境下，这些传感器既可提供所需信息，也可在需要时支持实时的闭环干预。

此外,若干论文报告了使用侧后区电极(即我们眼镜的 EEG 电极位置)与熵度量进行警觉评估。例如, Hu^[55]对单导 EEG 通道进行特征提取时采用了熵度量,报道单通道性能约为 90%,其方法为在 T6、TP7、T5、TP8、T4 与 T3 之中选择单通道,并结合特征模糊熵(FE)与随机森林分类器(RF)。Mu 等人^[56]亦对静息与疲劳状态下采集的 EEG 信号采用熵度量以提取特征。对 EEG 的回顾性分析显示,来自 T5、TP7、TP8 与 FP1 电极的特征提取得到更优表现。

T8 通道是检测困倦 / 疲劳状态变化的公认区域^[45,53,57]。因此,我们认为位于 T8 及其他侧后区电极近邻位置的 TP9 与 TP10 电极非常适合用于检测并复现上述发现以及专注状态。EOG 信号提供关于多种眼动的信息,因其易于搭建与信噪比 (SNR) 高,常被用于估计警觉性[58,59]。与 EEG 相比,EOG 的幅值显著更高,这使得 EOG 相较于 EEG 更具抗噪性。Kanoh 等人^[60]提出仅放置三个电极进行 EOG 测量,这与我们的设置相似。在其方法中,三个电极用于测量 EOG: 一个置于鼻梁,另两个分别置于鼻托的左侧与右侧。

拥有两种生理模态自然引出了融合二者以获得更佳分类结果的可行性问题。例如, Khushaba 等人^[45]开发了一种特征提取算法,通过分析来自三个 EEG 与一个 EOG 通道以及心电图 (ECG) 的相应生理信号,提取最相关特征以

表1 所提模型在 MA、WA 与 ME 任务上的结果

受试者/任务	心算（MA）	词语联想（WA）	旋律（ME）
受试者1	73%	74%	60%
受试者2	72%	60%	59%
受试者3	64%	58%	56%

表2 使用 LSTM 模型得到的结果

受试者/任务	MA-训练	MA-测试	MA-实时	WA-训练	WA-测试	WA-实时	ME-训练	ME-测试	ME-实时
受试者1	90%	88%	65%	94%	92%	66%	86%	84%	62%
受试者2	92%	91%	62%	94%	93%	62%	93%	92%	59%
受试者3	91%	89%	66%	83%	82%	64%	87%	85%	59%

预测驾驶员的困倦与疲劳状态。采用仅 EEG 通道，或 EEG + ECG、EEG + EOG 的组合，在使用非线性特征投影方法时均表现出很高的分类准确度。Huo 等人^[61]表明，融合来自 4 个电极的前额 EOG 数据与来自 12 个电极的 EEG 数据可提升疲劳检测性能。作者还报告，相较于单一信号模态（EEG 或 EOG），EEG 与 EOG 的模态融合可得到具有更高预测相关系数与更低均方根误差（RMSE）的模型。

关于在眼镜形态下使用 EEG 与 EOG 的更深入综述与更多参考，请参见 Kosmyna 等人^[10]。

我们使用自研的数据管线从该眼镜进行数据流传。

3.2 预研实验

本预研实验共招募 3 名受试者（2M/1F）。其中 1 名受试者患有 ALS。实验范式（例如一次“运行”）是一种主动且无需用眼的活动，受试者要么执行某项心理任务（见下文），要么进行心理放松（静息）。上述两种状态以级联方式排列，并由间隔段分隔，间隔段被视为缓冲性的中性状态（见图 8）。心理任务与心理放松各持续 2 min，而每个间隔段持续 40 s。我们将一次“运行”定义为按顺序执行以下任务的序列：任务、间隔、静息、间隔、任务、间隔、静息。每次运行以 40 s 的初始化期开始，在此期间受试者需尽量减少动作并尝试保持心境中性。



图8 一次运行的任务结构。此处，任务与静息的持续时间各为2 min，而校准与间隔各为40s

3.2.1 任务

我们为预研实验选择了三项任务，即心算（MA）、词语联想（WA）与旋律（ME）^[26]。在心算任务中，受试者需对随机数字快速进行一系列心算（中等到较高难度——取决于受试者的舒适程度），例如（128 × 56）、（5689 + 7854）、（36 × 12）。在词语联想任务中，受试者需要想出词语。他们需以一个随机词语开始，随后依次生成以下一个词语，其首字母与前一个词语的末字母相同，例如大象、老虎或浣熊。在旋律任务中，他们需在脑中回想熟悉的曲调，并进一步想象聆听该曲调，而不是发出词

语，而是专注于旋律本身。

3.2.2 方法与分类

首先，将 EEG 数据划分为历程 / 窗口。我们将每个窗口定义为时长 1 s 的 EEG 数据，并与其前一窗口有 75% 的重叠。下一步是数据预处理与清理。对历程化数据采用 50 Hz 陷波滤波器与通带为 0.5 Hz ~ 40 Hz 的带通滤波器组合进行滤波，以确保去除电源线噪声与不必要的高频成分。随后，我们构建了自有的伪迹剔除算法。该伪迹剔除算法的工作方式是：若相邻两个历程的绝对功率差超过预设阈值，则剔除该历程。在分类的最后一步，我们使用不同频带功率比的组合，依据任务来跟踪每位受试者的心理活动。对于 MA，所用比值为（alpha/delta）。对于 WA，所用比值为（delta/low beta），而对于 ME，所用比值为（delta/alpha）^[26]。随后，我们采用变点检测算法来跟踪这些比值的变化。这些比值幅度的突然升高或降低表明用户心理状态发生了变化。每当检测到可能的变化时，我们使用一种基于当前变点与上一变点处比值均值差异的算法来检查该变化是否指示了新的类别。若某次变化导致出现新类别或改变了该比值的趋势，则认为该变化有效。所有其他变化被视为无效并予以剔除。图 9 展示了运行一次 MA 任务的该比值。我们还探索了其他监督式方法，这些方法在测试与训练中取得了较高的准确率，但在实时数据中表现不佳。所提出的非监督方法从不对任何数据进行“训练”，并持续适应频带功率的变化与漂移；因此，预期在实际条件下具有良好表现。

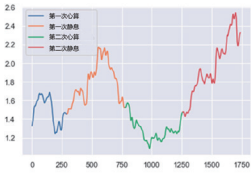


图9 上述曲线为alpha/delta比值，用于对MA任务进行分类
该曲线的变化用于判定受试者何时发生心理状态转换

3.2.3 结果

所提出的模型在 ALS 受试者上分别取得：MA 准确率 73%，WA 准确率 74%，ME 准确率 60%。相同模型在其

表3 初步结果，数据表示为 1 中位数（四分位距）；2 均值 ± 标准差

会话名称	训练	机器人会话	无监督
正确指令数 ¹	10（10 - 10）	30（30 - 30）	12（12 - 12）
成功率 ² （%）	100 ± 0.0	100 ± 0.0	100 ± 0.0
准确率 ² （%）	74.5 ± 6.9	100 ± 0.0	83.4 ± 9.2

他受试者上分别取得：MA 72%，WA 60%，ME 58%。需要指出的是，另外两名受试者为健康人，与 ALS 受试者相比，其身体和眼部运动更多，这可能是影响数据质量与最终结果的一个因素。表 1 给出了对每位受试者采用一种在实时数据上效果最佳且分类时间最短的非监督方法所得的结果。表 2 则给出了我们所探索的众多监督方法之一的结果；此处结果由 LSTM 模型^[54]生成。因此，我们决定在主研究中采用 MA 作为任务，但使用其他脑力挑战类任务（如 WA）也显示出较大潜力。

3.3 主要研究

两名健康受试者（1M/1F，中位年龄 32 岁）参与了本研究。两名受试者均具有 BCI 的既往使用经验。

如前所述，BCI 范式被设定为一种主动的心理活动，受试者要么执行一项心理任务，要么静息（REST）。我们选择心算（MA）作为主动心理任务。

如本文基础架构部分所述，我们为处理 EEG 数据构建了一条精心设计的数据管线，它构成了本研究工作的骨干。该管线由两个基本部分组成：前端与后端。前端包括一台 EEG 设备（AttentivU）以及安装在 iPhone 上的客户端应用（见图 10）。AttentivU 记录大脑的电活动并将这些数据回传给客户端应用。该应用旨在提供友好的界面，使用户能够开始数据采集录制、查看已录数据，并启动推理模式，以便通过 YES/NO 选项提出与回答问题（详见下文）。该管线的后端涵盖了云端组件，用于实现高效的数据流传、长期存储与实时预测。我们使用了 Google Cloud Platform 的 Cloud Storage，从而以安全且具成本效益的方式保存海量数据。驻留于后端的机器学习（ML）模型是另一关键组件。这些模型对缓存数据执行实时预测，为用户提供多种通信选项。为保证 BCI 系统的可扩展性与可靠性，除云存储外的所有后端组件均托管在 Kubernetes 集群中。这一整套基础设施使我们能够从多名用户处收集大量 EEG 信号。

3.4 任务与 BCI 会话

每名受试者共完成 30 次监督式会话，我们称之为“机器人会话”，在会话中他们需对六个问题之一给出预设的心理响应，问题为二元选择“是 / 否”，例如“你希望我去厨房吗？”。系统使用文本转语音以及屏幕文本来指示各时间段与过渡时段（见图 10）。要求每位参与者在“是或否”中选择其一，当听到自己所选项时执行 10s 的心算任务；若

非其所选项，则被告知执行 10 s 的静息。一次运行的示例如下：受试者听到“请对以下问题以“是”作答：‘你希望我去厨房吗？’”。该指令播放后，3s 后开始播放针对“是”的倒计时，持续 10 s，在此期间受试者应进行心算。随后有 3s 的间隔，然后播放“否”一词 10s；在这种情况下，受试者应处于静息状态。若分类输出为“是”，则 Spot 机器人将移动至厨房区域，而不论其当前所处位置。机器人会话包含所有需由 Spot 机器人执行的响应。



图10 BioSignal 应用的“是/否”选择界面

自左至右：选择“是”的倒计时；选择“否”的倒计时。系统输出为响应“是”。

尽管这种机器人运行的监督式设置可能使人感到疲劳，但这样做是为了评估 BCI 的输出。每个问题重复 5 次（共 30 次运行），所提问题包括：“你希望我去厨房吗？”“你希望我去窗边吗？”“你希望我去客厅吗？”“你希望我拍一张照片吗？”“你希望我把一个玩具带来吗？”以及“你希望我把一把椅子带来吗？”。在第 15 次运行之后安排 5 min 的休息。需要特别说明的是，在机器人会话之前，我们向用户演示了当每个问题选择 YES 时机器人将执行的“预期响应”动作，例如机器人在面对“厨房、客厅、窗边”时会去往何处以及如何移动。最后，我们还演示了“拍照、带来玩具、带来椅子”的动作，以便用户除看到 Spot 机器人在房间内移动外，不会对其即将出现的动作与声响感到意外或害怕。当选择“否”时，Spot 机器人不执行任何移动或动作。我们将此类会话称为“陪伴会话（pet session）”。在陪伴会话期间，受试者不执行任何活动，仅观察机器人。

在正式的“机器人”会话之前进行了包含 10 次运行的简短训练会话。训练会话中的每次运行同样对每个选项（MATH/REST）各记录 10 s，并通过文本转语音与屏幕文字向用户提供反馈，例如在屏幕与语音中指示“是”或“否”（“是 / 否”由系统事先请求）。训练会话与机器人会话之间有 5 min 的休息。训练会话期间未使用 Spot 机器人。

最后，以一次无监督会话结束任务。该会话在机器人

会话结束后休息 5 min 之后进行。无监督会话对每个问题包含两次运行，但由受试者自行决定如何回答每个问题。与机器人会话类似，此处也使用 Spot 机器人来执行用户的响应。我们要求受试者在每次运行之后告知其意向响应，以便为响应加标签。最后，我们与受试者进行了非正式访谈，以收集他们对 Ddog 的看法，尤其是关于易用性、舒适度与有效性相关的。

3.5 结果

结果见表 3。

由于我们仅有两名受试者，这些结果需被视为初步结论。

在主观意见方面，两名受试者均报告系统在易用性、舒适度与总体满意度方面表现良好。

他们表示享受控制机器人的过程，并且在实现控制成功方面具有很高的动机，尤其是在无监督会话期间。

4 挑战、局限与未来工作

在机器人与人机交互领域，多数挑战来自安全隐患与技术约束，以及 AI 代理在自行决策并与环境和人类交互时的约束。

随时间推进将被解决的技术挑战取决于能耗，以及电池的外形与成本。

成像传感器的质量会随时间提升，以更高的控制度与质量实时采集环境信息，从而使机器人系统对情境与环境具有更高感知能力。

如前所述，Spot 机器人是为工业用例而设计的，将相同硬件应用于助理用例及家庭 / 办公环境是可行的，但潜在上并非理想。

因此，针对居家导航与交互开发指南仍需进一步研究与试验，尤其是在机器人直接与人进行交互时。

总体而言，我们的研究应被视为概念验证（PoC），所有发现均为初步结论，因为我们仅有两名健康受试者；因此无法开展任何统计分析。

参与研究的受试者指出，他们在控制机器人以及使用 BCI 进行控制方面均感到高度投入。

我们认为，应开展更多同时包含健康对照与 LIS/ALS 患者的研究。

所采用的 BCI 范式亦应进一步评估；尽管其他使用心算的研究^[26]强调似乎未出现由习惯化导致的性能下降，我们仍应进一步探索个性化模型，以及词语联想等其他“脑力挑战”范式。

正如 Spataro 等人所指出的^[29]，除了在经济上可能减少辅助需求之外，控制一台助理机器人在心理层面也可能具有重要意义，可恢复基本的独立性形式。

除帮助日常生活任务外，激发积极情绪的可能性还可作为新兴机器人的康复与照护提供一系列潜在应用^[29]。

鉴于 Ddog 可被编程以执行极其广泛的功能，我们的早期结果有望为脑驱动机器人改善用户自主性开辟新应用。

随着探索机器人在人类任务中的应用迅速扩展且成本持续下降，本研究或可引领先进助理机器人的发展。

作者贡献：构想，N.K.、E.H. 与 Y.H.；方法学，N.K.、E.H. 与 Y.H.；软件，N.K. 与 E.H.；验证，N.K. 与 E.H.；形式分析，N.K.；调查研究，N.K.；资源，N.K.；数据整理，N.K.；撰写一初稿准备，N.K.、E.H. 与 Y.H.；撰写一评阅与修订，N.K.；可视化，N.K.、E.H. 与 Y.H.；监督，N.K.；项目管理，N.K.；经费获取，N.K. 全体作者已阅读并同意本文发表版本。

经费：本研究未获得外部资助。

机构审查委员会声明：本研究遵循《赫尔辛基宣言》开展，并经 MIT 机构审查委员会批准（方案编号 210700428，2021 年 9 月 15 日）。

知情同意声明：所有参与研究的受试者均已签署知情同意。

数据可用性声明：请联系通讯作者。

致谢：我们感谢 Edward Hartley 在本项目早期建图与早期目标检测迭代中的帮助。

利益冲突：作者 Eugene Hauptmann 受雇于 Reactive Lions Inc. 其余作者声明，开展本研究时不存在可被视为潜在利益冲突的任何商业或金融关系。

参考文献：

- [1] Bandom, R. 波士顿动力的 Spot 正走出实验室；The Verge: 美国纽约州纽约市，2019 年。
- [2] Harrison, M. 机器狗为模特“脱衣”的视频热传：对不祥科技的戏仿；Futurism: 美国纽约州纽约市，2023 年。
- [3] Etemad-Sajadi, R. 认识 Pepper: 类人机器人礼宾。EHL Insights, 2018 年 12 月 17 日。
- [4] Newton, D.P.; Newton, L.D. 作为教师的类人机器人及一份实践规范建议. Front. Educ. 2019,4,125. [CrossRef].
- [5] 在线可得: <https://www.fanucamerica.com/solutions/industries/warehouse-robots> (访问日期: 2023 年 12 月 14 日)。
- [6] 在线可得: <https://www.defenseone.com/technology/2023/03/soldiers-can-now-steer-robot-dogs-brain-signals/384338/> (访问日期: 2023 年 12 月 14 日)。
- [7] Banach, K.; Małeck, M.; Ros ó ł, M.; Broniec, A. 基

于 EEG α 波的电动轮椅脑机接口. *Bio.-Algorithms Med.-Syst.* 2021,17,165-172. [CrossRef].

[8] Benabid, A.L.; Costecalde, T.; Eliseyev, A.; Charvet, G.; Verney, A.; Karakas, S.; Foerster, M.; Lambert, A.; Morinière, B.; Abroug, N.; 等. 由硬膜外无线脑机接口控制的四肢瘫痪患者外骨骼: 概念验证示范. *Lancet Neurol.* 2019,18,1112-1122. [CrossRef][PubMed].

[9] Wang, H.; Li, Y.; Long, J.; Yu, T.; Gu, Z. 基于混合 EEG-EOG 的异步轮椅控制. *Cogn. Neurodyn.* 2014,8,399-409.[CrossRef][PubMed].

[10] Kosmyrna, N.; Morris, C.; Nguyen, T.; Zepf, S.; Hernandez, J.; Maes, P. *AttentivU*: 在车载环境中用于生理感知与反馈的 EEG 与 EOG 兼容眼镜的设计. 载于: 第 11 届汽车用户界面与交互式车载应用国际会议 (AutomotiveUI 2019) 论文集, 荷兰乌得勒支, 2019 年 9 月 21-25 日; Association for Computing Machinery: 美国纽约州纽约市, 2019; 355-368 页.

[11] Kosmyrna, N.; Morris, C.; Sarawgi, U.; Nguyen, T.; Maes, P. *AttentivU*: 一副用于实时生理处理的 EEG 与 EOG 可穿戴眼镜. 载于: 2019 年 IEEE 第 16 届可穿戴与植入式人体传感器网络国际会议 (BSN), 美国伊利诺伊州芝加哥, 2019 年 5 月 9-22 日; 1-4 页.

[12] 在线可得: <https://www.bostondynamics.com/products/spot> (访问日期: 2023 年 12 月 14 日).

[13] Carlson, T.; Millan, J.d.R. 脑控轮椅: 一种机器人体系结构. *IEEE Robot. Autom. Mag.* 2013,20,65-73.

[14] Tavella, M.; Leeb, R.; Rupp, R.; Millan, J.d.R. 走向日常生活的自然非侵入式手部神经假体. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2010,2010,126-129. [PubMed].

[15] Müller-Putz, G.R.; Scherer, R.; Pfurtscheller, G.; Rupp, R. 基于 EEG 的神经假体控制: 迈向临床实践的一步. *Neurosci. Lett.* 2005,382,169-174. [CrossRef].

[16] Contreras-Vidal, J.L.; Grossman, R.G. *NeuroRex*: 面向下肢机器人外骨骼的基于 EEG 的脑机接口临床路线图. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2013,2013,1579-1582.

[17] Onose, G.; Grozea, C.; Anghelescu, A.; Daia, C.; Sinescu, C.J.; Ciurea, A.V.; Spircu, T.; Mirea, A.; Andone, I.; Spânu, A. 在慢性四肢瘫痪者中使用基于运动想象 EEG 的脑机接口进行助理机械臂控制的可行性: 临床测试与长期随访. *Spinal Cord.* 2012,50,599-608. [CrossRef].

[18] Kosmyrna, N.; Lecuyer, A. 基于 EEG 的脑机接口概念空间. *PLoS ONE* 2019,14,e0210145. [CrossRef].

[19] Ramos-Murguialday, A.; Schürholz, M.; Caggiano, V.; Wildgruber, M.; Caria, A.; Hammer, E.M.; Halder, S.; Birbaumer, N. 本体感觉反馈与基于脑机接口 (BCI) 的神经假体. *PLoS ONE* 2012,7,e47048. [CrossRef][PubMed].

[20] Polich, J. P300 更新: P3a 与 P3b 的综合理论. *Clin. Neurophysiol.* 2007,118,2128-2148. [CrossRef][PubMed].

[21] Tonin, L.; Carlson, T.; Leeb, R.; Millán, J.d.R. 由运动障碍者控制的远程存在机器人. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2011,2011,4227-4230. [PubMed].

[22] Blankertz, B.; Losch, F.; Krauledat, M.; Dornhege, G.; Curio, G.; Müller, K.R. 柏林脑机接口: BCI 新手在首个会话中的高精度表现. *IEEE Trans. Biomed. Eng.* 2008,55,2452-2462. [CrossRef][PubMed].

[23] Halpern, A.R.; Zatorre, R.J. 当那段旋律在你脑中回响: 熟悉旋律听觉表象的 PET 研究. *Cereb. Cortex* 1999,9,697-704. [CrossRef][PubMed].

[24] Pei, X.; Barbour, D.; Leuthardt, E.C.; Schalk, G. 利用人类皮层电图信号解码口语与想象词语中的元音与辅音. *J. Neural Eng.* 2011,8,046028. [CrossRef][PubMed].

[25] Kosmyrna, N.; Lindgren, J.T.; Lécluyer, A. 关注视觉刺激与执行视觉意象: 基于 EEG 的脑机接口控制策略比较. *Sci. Rep.* 2018,8,13222. [CrossRef].

[26] Friedrich, E.V.C.; Scherer, R.; Neuper, C. 不同心理策略对脑机接口分类性能的影响. *Int. J. Psychophysiol.* 2012,84,86-94. [CrossRef].

[27] Zhang, J.; Wang, M. 基于运动想象脑机接口控制机器人的综述. *Cogn. Robot.* 2021,1,2-24. [CrossRef].

[28] Galán, F.; Nuttin, M.; Lew, E.; Ferrez, P.; Vanacker, G.; Philips, J.; Millán, J.d.R. 脑控轮椅: 用于连续机器人控制的异步、非侵入式脑机接口. *Clin. Neurophysiol.* 2008, 119,2159-2169. [CrossRef].

[29] Hockberg, L.R.; Bacher, D.; Donoghue, J.P.; Jarosiewicz, B.; Masse, N.Y.; Simeral, J.D.; Vogel, J.; Haddadin, S.; Liu, J.; Cash, S.S.; 等. 四肢瘫痪者使用神经控制机械臂实现伸手与抓握. *Nature* 2012,485,372-375. [CrossRef][PubMed].

[30] Zhang, R.; Lee, S.; Hwang, M.; Hiranaka, A.; Wang, C.; Ai, W.; Tan, J.J.R.; Gupta, S.; Hao, Y.; Levine, G.; 等. NOIR: 面向日常活动的神经信号操控智能机器人. *arXiv* 2023,arXiv:2311.01454. [CrossRef].

[31] Tang, J.; Liu, Y.; Hu, D.; Zhou, Z. 面向脑机接口驱动的智能轮椅系统. *BioMed. Eng. OnLine* 2018,17,111.

[CrossRef][PubMed].

[32] Araujo, R.S.; Silva, C.R.; Netto, S.P.N.; Morya, E.; Brasil, F.L. 纺织品 3D 打印的低成本 EEG 控制手部外骨骼的开发. *Front. Neurosci.* 2021, 15, 661569. [CrossRef][PubMed].

[33] Spataro, R.; Chella, A.; Allison, B.; Giardina, M.; Sorbello, R.; Tramonte, S.; Guger, C.; La Bella, V. 闭锁综合征 ALS 患者通过 BCI 控制的类人机器人实现触及并抓握一杯水. *Front. Hum. Neurosci.* 2017,11,68. [CrossRef].

[34] Faisal, S.N.; Do Nguyen, T.T.; Torzo, T.; Leong, D.; Pradeepkumar, A.; Lin, C.T.; Iacopi, F. 基于微图案外延石墨烯的非侵入式脑机接口传感器. *ACS Appl. Nano Mater.* 2023,6,5440 – 5447. [CrossRef].

[35] Na, R.; Hu, C.; Sun, Y.; Wang, S.; Zhang, S.; Han, M.; Yin, W.; Zhang, J.; Chen, X.; Zheng, D. 带混合刺激器的嵌入式轻量化 SSVEP-BCI 电动轮椅. *Digit. Signal Process.* 2021,116,103101. [CrossRef].

[36] Lee, K.; Liu, D.; Perroud, L.; Chavarriaga, R.; Millán, J.d.R. 采用级联 ERD 分类器的脑控外骨骼. *Robot. Auton. Syst.* 2017,90,15 – 23. [CrossRef].

[37] Chae, Y.; Jeong, J.; Jo, S. 迈向脑控类人机器人：基于 EEG 的 BCI 异步直接控制. *IEEE Trans. Robot.* 2012,28,1131 – 1144. [CrossRef].

[38] Daeglau, M.; Wallhoff, F.; Debener, S.; Condro, I.S.; Kranczioch, C.; Zich, C. 挑战接受？在考虑个体差异、动机与社会情境的前提下，面向类人机器人 EEG 神经反馈的运动想象练习个体性能提升. *Sensors* 2020,20,1620. [CrossRef].

[39] Choi, B.; Jo, S. 基于低成本 EEG 系统的混合脑机接口用于类人机器人导航与识别. *PLoS ONE* 2013,8,e74583. [CrossRef][PubMed].

[40] Sorbello, R.; Tramonte, S.; Giardina, M.E.; La Bella, V.; Spataro, R.; Allison, B.; Guger, C.; Chella, A. 通过 BCI 的人 - 类人机器人交互：面向闭锁综合征 ALS 患者的神经-生物反馈融合. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2018,26,487 – 497. [CrossRef][PubMed].

[41] Collinger, J.L.; Wodlinger, B.; Downey, J.E.; Wang, W.; Tyler-Kabara, E.C.; Weber, D.J.; McMorland, A.J.; Velliste, M.; Boninger, M.L.; Schwartz, A.B. 一名四肢瘫痪个体实现的高性能神经假体控制. *Lancet* 2013,381,557 – 564. [CrossRef][PubMed].

[42] Wodlinger, B.; Downey, J.E.; Tyler-Kabara, E.C.; Schwartz, A.B.; Boninger, M.L.; Collinger, J.L. 人体脑机接口中的十维仿生手臂控制：难点、解决方案与局限. *J. Neural*

Eng. 2015,12,016011. [CrossRef][PubMed].

[43] Jia, W.; Huang, D.; Luo, X.; Pu, H.; Chen, X.; Bai, O. 基于 EEG 的本能式脑控四足行走机器人. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2012,2012,1777 – 1781. [CrossRef][PubMed].

[44] Berka, C.; Levendowski, D.J.; Lumicao, M.N.; Yau, A.; Davis, G.; Zivkovic, V.T.; Olmstead, R.E.; Tremoulet, P.D.; Craven, P.L. 在警觉性、学习与记忆任务中任务投入与心理负荷的 EEG 相关性. *Aviat. Space Environ. Med.* 2007,78 (增刊 5) , B231 – B244.

[45] Khushaba, R.N.; Kodagoda, S.; Lal, S.; Dissanayake, G. 基于模糊小波包特征提取算法的驾驶员困倦分类. *IEEE Trans. Biomed. Eng.* 2011,58,121 – 131. [CrossRef].

[46] Shi, L.-C.; Duan, R.-N.; Lu, B.-L. 基于 EEG 的警觉性估计：极限学习机方法. *Neurocomputing* 2013, 102,135–143. [CrossRef].

[47] Shi, L.-C.; Duan, R.-N.; Lu, B.-L. 一种用于基于 EEG 的警觉性估计的稳健主成分分析算法. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2013,2013,6623–6626.

[48] Shi, L.-C.; Duan, R.-N.; Lu, B.-L. 用于基于 EEG 的警觉性估计的差分熵特征. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2013,2013,6627–6630.

[49] Schleicher, R.; Galley, N.; Briest, S.; Galley, L. 眨眼与扫视作为困倦预警中疲劳指标：看起来很累？ *Ergonomics* 2008,51,982–1010. [CrossRef][PubMed].

[50] Skotte, J.H.; Nøjgaard, J.K.; Jørgensen, L.V.; Christensen, K.B.; Sjøgaard, G. 通过眼电图量化不同计算机任务中的眨眼频率. *Eur. J. Appl. Physiol.* 2007,99,113 – 119. [CrossRef][PubMed].

[51] Bulling, A.; Ward, J.A.; Gellersen, H.; Tröster, G. 基于眼电图的活动识别之眼动分析. *IEEE Trans. Pattern Anal. Mach. Intell.* 2011,33,741–753. [CrossRef][PubMed].

[52] Morris, T.L.; Miller, J.C. 模拟飞行中疲劳的眼电与绩效指标. *Biol. Psychol.* 1996,42,343–360. [CrossRef][PubMed].

[53] Sahayadhas, A.; Sundaraj, K.; Murugappan, M. 基于传感器检测驾驶员困倦：综述. *Sensors* 2012,12,16937 – 16953. [CrossRef].

[54] Spampinato, C.; Palazzo, S.; Kavasidis, I.; Giordano, D.; Souly, N.; Shah, M. 用深度学习解读人类思维以实现自动视觉分类. 载于：IEEE 计算机视觉与模式识别会议论文集，美国夏威夷州檀香山，2017 年 7 月 21 – 26 日；6809 – 6817 页.

[55] Hu, J. 基于单通道 EEG 的驾驶员疲劳检测: 不同特征与分类器的比较. *Comput. Math. Methods Med.* 2017,2017,5109530. [CrossRef].

[56] Mu, Z.; Hu, J.; Min, J. 基于组合熵特征的 EEG 驱动疲劳检测系统. *Appl. Sci.* 2017,7,150. [CrossRef].

[57] Jap, B.T.; Lal, S.; Fischer, P.; Bekiaris, E. 使用 EEG 频谱成分评估检测疲劳算法. *Expert Syst. Appl.* 2009, 36, 2352 - 2359. [CrossRef].

[58] Ma, J.X.; Shi, L.C.; Lu, B.L. 通过眼电特征进行警觉性估计. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2010,2010,6591-6594. [CrossRef].

[59] Papadelis, C.; Chen, Z.; Kourtidou-Papadeli, C.; Bamidis, P.D.; Chouvarda, I.; Bekiaris, E.; Maglaveras, N. 车载电生理记录监测困倦以预防睡眠不足致交通事故. *Clin. Neurophysiol.* 2007,118,1906-1922. [CrossRef].

[60] Kanoh, S.; Ichi-nohe, S.; Shioya, S.; Inoue, K.; Kawashima, R. 一种用于测量眼部与身体运动的眼镜式设备的开发. 载于: 2015 年 IEEE 医学与生物工程学会第 37 届年度国际会议 (EMBC), 意大利米兰, 2015 年 8 月 25 - 29 日; 2267 - 2270 页. [CrossRef].

[61] Huo, X.-Q.; Zheng, W.-L.; Lu, B.-L. 融合 EEG 与前额 EOG 的驾驶疲劳检测. 载于: 2016 年国际神经网络联合会议 (IJCNN), 加拿大不列颠哥伦比亚省温哥华, 2016 年 7 月 24 - 29 日; 897 - 904 页. [CrossRef].

免责声明 / 出版者注: 所有出版物中的陈述、观点与数据仅代表各作者与贡献者个人, 而非 MDPI 及 / 或编辑的观点。MDPI 及 / 或编辑对内容中所涉任何思想、方法、操作或产品导致的人身或财产损害不承担责任。

通讯作者: 原著: nkosmyna@media.mit.edu (Nataliya Kosmyna)

译文: wuke3983@163.com (吴可)

原文信息:

原期刊: *Sensors* 2024, 24, 80. <https://www.mdpi.com/journal/sensors>

原文链接: <https://doi.org/10.3390/s24010080>

Kosmyna N, Hauptmann E, Hmaidan Y. A Brain-Controlled Quadrupe Robot:

A Proof-of-Concept Demonstration. *Sensors*, 2024, 24: 80. DOI: 10.3390/s24010080

期刊主页: *Sensors* (MDPI) — <https://www.mdpi.com/journal/sensors>

原文链接 (DOI): <https://doi.org/10.3390/s24010080>

译者说明:

本译文选择发表于 *Sensors* 的《A Brain-Controlled Quadrupe Robot: A Proof-of-Concept Demonstration》作为对象, 旨在降低国内读者获取前沿脑机接口 (Brain - Computer Interface, BCI) 与四足机器人交叉研究成果的门槛, 促进工程实现与临床 / 康复场景的落地交流。原文以波士顿动力 Spot 机器人为平台, 探索在可穿戴、无线、非视觉范式 (以心算为代表) 下, 由眼镜式 EEG/EOG 设备对移动与操控任务进行意图判别与执行, 作为个人助理方向的概念验证; 对 LIS/ALS 等人群的潜在助残价值与人机共生形态具有参考意义。

为保证学术可读性与可复现性, 本译文遵循以下原则:

(1) 逐句对齐: 仅输出中文, 但严格“一句英文对应一句中文”, 不合并、不删改; (2) 术语一致: 术语首现采用“中文 (English, 缩写)”, 后文统一; 变量 / 单位 / 上下标 / 斜体等排版属性与符号保持原样; (3) 编号与引文: 保留图 / 表编号与文内交叉引用; [n]、[n - m] 等引用标记与 DOI/URL 原样保留; (4) 图表与公式: 表格逐格翻译保留结构; 公式仅译周边文字, 不改动数学内容; (5) 一致性审校: 抽检术语与引用、单位与变量书写、链接可用性、页眉 / 脚注等, 确保与原文一致。

版权与许可声明 (CC BY 4.0):

本译文作为原文的衍生作品, 遵循 CC BY 4.0 协议 (允许包含商业用途的共享与再创作), 使用者须署名原作者与译者并注明改动; 第三方素材 (商标 / 图片 / 数据) 如另有权利声明, 可能不受 CC BY 4.0 覆盖。译文仅进行语言与版式转换, 学术结论以原文为准。期望此译文为国内科研与产业同仁在 BCI- 机器人融合、助理机器人、人机协作与康复工程 等方向提供可读、可复用的参考。

Confirmation re: Chinese translation under CC BY 4.0 (Sensors 2024, 24:80)

Dear Mr. Ke Wu,
Thank you for your message and for your interest in our paper:
Kosmyna N, Hauptmann E, Hmaidan Y. A Brain-Controlled Quadrupe Robot: A Proof-of-Concept Demonstration. *Sensors*, 2024, 24:80. DOI: 10.3390/s24010080.
We confirm that the article is published under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. Under CC BY 4.0, no additional authorization from the authors is required to produce and distribute a Chinese translation (a derivative work), provided that you:
1. Give appropriate attribution to the original authors and journal and include the DOI/link;
2. Provide a clear statement of changes (e.g., that this is a Chinese translation and any formatting/localization edits);
3. Respect the rights status of any third-party materials (figures/tables/images/data) that may not be covered by CC BY and obtain separate permissions where needed;
4. Do not imply author or institutional endorsement of the translation unless explicitly agreed in writing.
For your convenience, you may include the following license notice in the translation:
"This Chinese translation is based on the original article published under CC BY 4.0: <https://creativecommons.org/licenses/by/4.0/>. The translation provides language and formatting conversion only; scientific conclusions remain as in the original publication."
We appreciate your courtesy in notifying us and wish you success with the dissemination of the translation.
Best regards,
On behalf of the authors.

Nataliya Kosmyna
Sep 25, 2024