

给人工智能装上一双“人的眼睛”

● 本报记者 陈彬

近年来,随着人工智能的飞速发展,以无人驾驶、具身智能等为代表的智能无人系统得到不断推广和应用。人们希望这些系统可以在现实世界中更加自如地活动,更加准确地给人类提供帮助。但要达到这一目的有一个前提,那就是这些智能无人系统能更真切地“看”到这个世界,并能够更准确、迅速地对周围环境变化作出反应。

换句话说,它们需要一双更加敏锐的“眼睛”。

不久前,清华大学精密仪器系类脑计算研究团队聚焦类脑视觉感知芯片技术,提出了一种基于视觉原语的互补双通路类脑视觉感知新范式。在此基础上,该团队成功研制出世界上首款类脑互补视觉芯片“天眸芯”,大大提升了人工智能的信息采集效率。

近日,该研究成果以封面文章形式发表于《自然》。

“鱼”与“熊掌”不可兼得

清华大学类脑计算研究团队成立于2013年。自成立之初,该团队便将发展人工通用智能作为科研目标。

“这个目标我们从来没有改变过,但路还要一步一步走。”团队成员、清华大学精密仪器系教授赵蓉告诉《医学科学报》。最初,他们想到人工通用智能首先需要对数据进行高效计算。团队经过多年刻苦攻关,于2019年成功研制出世界首款异构融合类脑芯片“天机芯”,这是该团队的成果第一次登上《自然》封面。

“天机芯”的研发成功,给人工通用智能提供了一个足够强大的“大脑”,但仅有“脑子”,而没有准确且及时的数据传入是不行的,人类对外界的信息采集,90%都依赖于视觉。于是,研究团队将目光对准了人工智能的视觉感知领域,他们想知道,目前的视觉芯片是否存在一些尚待解决的问题。

问题很快就被他们找到了。

“人工智能在复杂环境下的信息采集主要靠传感器完成。具体到视觉感知领域,几乎所有的传感器在采集信息时,都希望将目标区域内的所有信息‘一网打尽’,似乎信息收集越全,

人工智能对周围环境的感知力越强。”受访时,团队带头人、清华大学精密仪器系教授施路平表示。

虽然现有的传感器可以凭借越来越高的分辨率海量地采集数据,但从传感器到数据终端的传输网络带宽是有限的。于是,数据的激增和网络的局限之间形成了一种明显的矛盾。

“这就像我们可以购买很多辆汽车,但道路不拓宽,这些汽车就会在道路上拥堵,无法快速到达目的地。”施路平说。

此外,与信息大量采集和大量传输相伴的,还有软硬件能耗的激增以及信息采集动态范围(最大可测光强与最小可测光强的比值)的限制。

“分辨率、速度、精度/动态范围,这三者就像‘鱼’和‘熊掌’一样,在传统技术条件下几乎不可能兼得,一方性能的提升会导致其余两方性能的下降,进而导致视觉出现失真、失效或高延迟的问题,严重影响系统的稳定性和安全性。”赵蓉说。

对于这个问题,施路平团队是如何解决的?

做两套视觉系统

据施路平介绍,现阶段发展人工通用智能的方法主要有两种,一种基于电脑思维,另一种基于人脑思维。两种方法各有优缺点,但都代表人脑处理信息的部分模式。

“在研发‘天机芯’时,我们最大的收获便是创造了一种将两种思维方式相互融合类脑计算思维架构。”施路平说。

研究团队研发视觉芯片的出发点,同样来自将计算思维与人脑思维相结合。

赵蓉解释说,与机器类似,人类的视觉系统同样存在“信息量”与“带宽”不可兼得的问题,但在漫长的演化过程中,人类找到了该问题的最佳解决方式——双通路互补的视觉系统。

“人体的视觉细胞分为两类——视锥细胞和视杆细胞,这两类细胞具有不同的特点。”赵蓉说,视杆细胞拥有很高的感光度,对弱光反应灵敏,但只能提供

粗糙的视觉质量,不能感受颜色;视锥细胞拥有较低的感光度,但可以提供较高的视觉质量,使我们看到颜色。大脑通过功能互补的双通路进行视觉信息处理,

其中一个通路主要负责处理颜色、细节等高精度信息,但处理速度相对较慢,另一个通路主要负责处理突发事件,能够作出快速反应,但精度不高。

“当一个足球快速向你飞来时,你会迅速发觉有物体运动,但完全不会注意到足球的任何细节;等足球停止运动,足球的所有细节都会呈现在你眼前。”赵蓉说,正是双通路视觉系统的相互配合,才使得人眼可以既准确又快速地观察到物体,并在极端场景下完成信息处理、进行应对。

施路平团队所提出的类脑视觉感知新范式便借鉴了人类视觉系统这一基本原理。

具体而言,他们将开放世界的视觉信息拆解为不同表达类型,并通过对这些信息的组合,模仿人类视觉系统的特征,形成两条优势互补、信息完备的视觉感知通路。

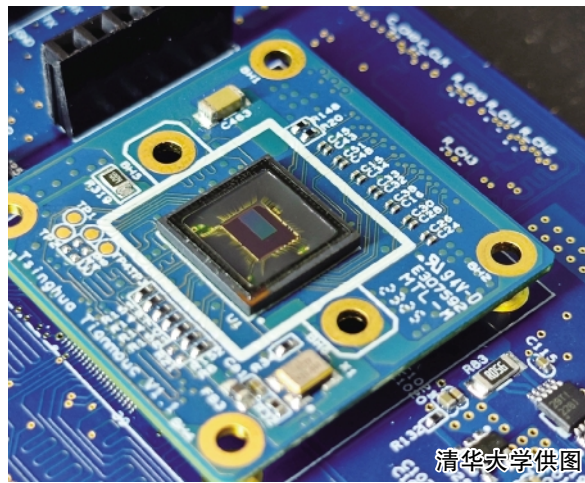
“这套系统首先会模仿人眼,将收集到的周围场景分解成很多元素,如物体运动、轮廓、色彩以及不同区域的对比度等,之后模仿人脑的处理机制,将这些信息分配到不同的‘处理通路’中,这就像工厂的多条流水线。”赵蓉说,系统再通过模仿人脑响应机制,对相关信息进行连接和协同,最终实现信息的高效、稳定传输。

“这相当于给机器人装上了一双人的眼睛。”赵蓉说。

未完成的工作

基于上述新范式,施路平团队几经攻关,成功研制出世界首款类脑互补视觉芯片——“天眸芯”。

据介绍,该芯片可在极低的带宽(降低90%)和功耗代价下,实现每秒10000帧的高速、10比特的高精度、130分贝的高动态范围的视觉信息采集,不仅突破了传统视觉



清华大学供图

感知范式的性能瓶颈,而且能够高效应对各种极端场景,确保系统的稳定性和安全性。

在“天眸芯”的研发基础上,团队还自主设计了高性能软件和算法,并在开放环境车载平台上进行了性能验证。在多种极端场景下,该系统均实现了低延迟、高性能的实时感知推理。

据介绍,研究人员还开发了一套集成“天眸芯”的汽车驾驶感知系统,以评估其在开放道路上行驶时间面对极端情况——如强光干扰、光线强烈变化、突然出现异常物体时的反应。结果表明,“天眸芯”可以有效适应极端光环境并提供先进的多级感知能力。

“‘天眸芯’的成功研制,不仅为智能革命的发展提供了强大的技术支持,还为自动驾驶、具身智能等重要应用开辟了新道路。”施路平表示,将这一科研成果进一步推向应用,是目前该团队的一个重要工作。

赵蓉表示,一项技术与实际应用相结合,甚至产生商业价值,才称得上真正完成了研发的全流程。接下来,团队将推动“天眸芯”的落地应用。

同时,该团队也在开展相关工作,以将“天眸芯”与“天机芯”进行结合,并最终将其成功“植入”机器人的身体。“人工通用智能肯定是要在机器人身上展现的。”赵蓉说。

当然,在机器人视觉传输领域,现有的“天眸芯”虽然在“类人”的层面有了很大进步,但跟真正的人眼视觉系统相比仍有很大差距。弥补这些差距将是该团队的重要工作。

“总之,要实现团队成立之初立下的目标,我们还有太多工作要做。”施路平说。

相关论文信息:<https://doi.org/10.1038/s41586-024-07358-4>