

释放“想象力” 让AI更加理解人类世界

■ 成澄

目前,南加州大学的研究团队正在研究如何让AI模仿人类的想象过程,令其产生与人类类似的“想象力”,以合成或了解不同属性、前所未见的事物。研究人员认为,这项技术不但可以协助人们合成新药,还能提高自动驾驶汽车的安全性,并且让人工智能消除数据算法带来的潜在偏见。

当前AI难以分辨事物属性

对于人类而言,想象出一只黑色猫咪在长城上快速奔跑的脑内图像显得十分简单。然而,尽管深度神经网络(机器学习领域中的一种技术)在某些任务上取得了跟人类相当或超过人类的表现,但计算机在人类特有的“想象力”技能方面仍然薄弱。

研究人员们的长期目标之一,即“让AI创建可以推断的模型”。比如说,如果研究人员想让某个AI自行生成一张汽车图像,那么在理想情况下,可以通过为这个AI提供更多的汽车图像,然后这个AI就可以从颜色、类型等角度生成多种类型的汽车,如红色的保时捷、绿色的皮卡等。

如果这个目标实现,就意味着当AI接收到人们给出的几个样本时,就算它在此之前从未见过这样的事物,也能立即提取出这些样本的基本共性,并将这些共性应用于大量的新事物中。然而,目前的AI并不能清楚分析出事物的属性特征,比如一块长方形的木板,具有长方形、木质、板等属性,但AI并不清楚这些属性,只能从给定的木板图片中提取出类似长度、颜色相同的共性,这些共性并不一定是事物的属性。

通过“解耦”让AI展开“想象力”

南加州大学的研究团队试图使用“解耦”来克服AI模型难以分辨研究对象不同属性的难题。

解耦,即当你直接替换事物的某个特性时,对原

来事物的理解不造成影响。比如,当你拥有一台黑色电脑,并且从未见过黑色水杯,却知道水杯的样子时,你就会从黑色电脑上知道什么是黑色,然后将这个属性应用于水杯上,这时你就知道了黑色水杯是什么样子了。这个过程,就是实现了“解耦”。

与传统算法一次只分析一个样本不同,AI学习拆解事物的属性需要通过解耦技术分析一组样本图像,并挖掘它们之间的相似性,实现“可控的解耦表征学习”。然后,AI将重新组合这些属性信息,以实现“可控的新图像合成”,或者可以叫作“想象力”。

这样的想象方式在某些方面类似于人类的推断行为,如同当人们看到一个物体的颜色时,可以用新颜色替换原始颜色,并轻松推断出这个颜色的其他物体会是什么样。

解耦还可以应用于深度伪造技术。例如将人的身份和动作分开,再由AI模型将A的身份替换为B,但保留A的动作,以此合成为新的图像和视频。

此外,在进行该项研究时,研究团队还生成了一个156万张的图像数据集,可以帮助相关领域的研究团队开展研究。

在众多领域拥有广泛前景

研究人员表示,他们的设计框架可以让AI适应于大多数的数据和知识领域,这将扩大AI“想象”的应用范围。此外,AI还可以从相关信息中完全抹除种族、性别等容易造成歧视的属性,让人们的信息在被数据分析时,得到相同的处理方式。

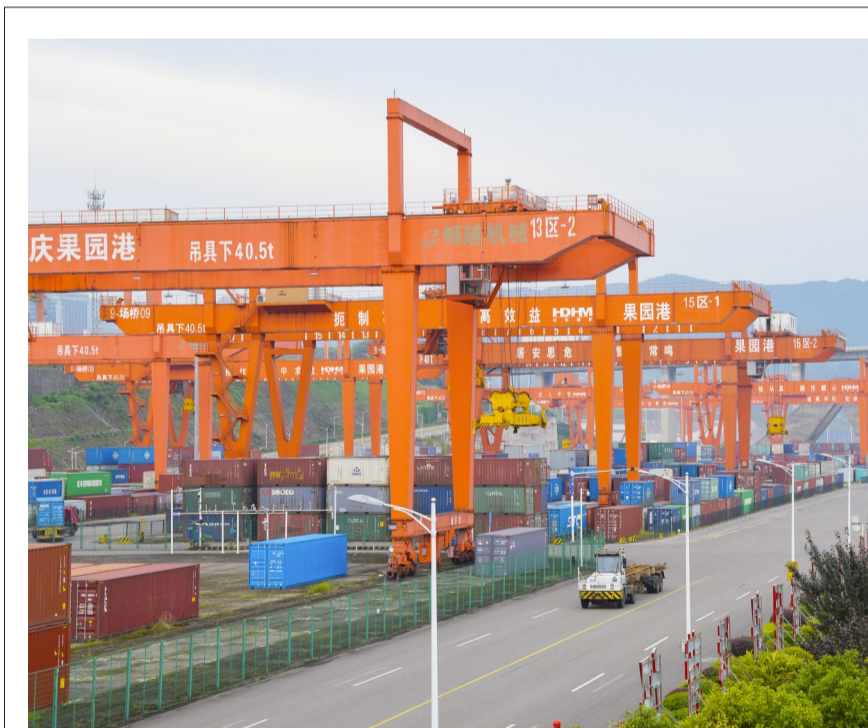
在医学领域,学会分辨属性的AI可以帮助医生和生物学家发现更多有用的药物。AI将药物的功能与药物的其他特性分开,然后将不同药物的特性重新组合,以合成新药。

不仅如此,当AI具有“想象力”后,它还可以帮助创建更安全的人工智能。比如说,它让自动驾驶汽车



“想象”如何避开训练中从未遇见的危险场景。

“深度学习在很多领域都展示了较大的前景,但这些研究往往都没有深入去了解每个研究对象的独特属性特征。”计算机科学教授Laurent Itti说,“这是第一次真正地释放了人工智能的想象力,使它们能更加理解人类是如何看待世界的。”



7月19日拍摄的重庆果园港。

今年上半年,重庆果园港国家物流枢纽完成货物吞吐量超1000万吨,同比增长55.8%。其中,渝新欧在果园港到发151班次,西部陆海新通道班列到发120班次,较去年同期均呈现出增长态势。

新华社记者徐钦 摄

通用AI平台 实现图文音“无缝转换”

新华社北京电(记者董瑞丰)从检索图像、描述视频,到吟诗作赋、续写文章,再到识别语音、双语翻译,虚拟人“小初”日前亮相2021世界人工智能大会,展示了图、文、音三种模态的智能转换和生成。

“小初”具备这样的能力,得益于一个名为“紫东太初”的跨模态通用人工智能平台。该平台由中国科学院自动化研究所(以下简称“中科院自动化所”)研发,基于国产化基础软硬件,仅采用一个大模型,即可“锻炼”人工智能在视觉、文本、语音多个场景的理解能力。

中科院自动化所所长徐波介绍,“大数据+大模

型+多模态”将改变当前单一模型对应单一任务的人工智能研发范式,多模态大模型将成为不同领域的共性平台技术,是迈向通用人工智能的重要路径探索。

“‘紫东太初’实现了图、文、音三种模态的统一表达,以图生音,以音生图,将开拓AI在视频配音、语音播报、标题摘要、海报创作等更多元场景的应用。”徐波说。

据了解,中科院自动化所构建了中文预训练模型、语音预训练模型、视觉预训练模型,并通过跨模态语义关联,构建了多模态预训练大模型。

哈勃空间望远镜完成修复 已恢复科学运行

今年6月13日,哈勃空间望远镜因科学仪器指令和数据处理单元(SI C&DH)中的有效载荷计算机故障而停止运行,整体进入安全模式。经过一个多月的排查,美国国家航空航天局(NASA)终于将故障点定位在一个电源控制模块中。该电源控制模块原本应负责给有效载荷计算机稳定地提供5伏电压,但当时,分析显示出它并没有按照预设指令工作。

目前,美国国家航空航天局将该设备切换到备份设备后,终于在7月17日,哈勃空间望远镜恢复到了科学运行状态,大部分因故障无法进行的科学观测也会在近期重新排期。

哈勃空间望远镜至今已经工作了31年,进行过150万次观测,超过18000篇论文采用了它的数据。(本报综合)

故宫门锁上也有黑科技

北京故宫是中国明清两代的皇家宫殿,旧称为紫禁城,位于北京中轴线的中心,占地面积72万平方米,建筑面积约15万平方米,有大小宫殿七十多座,房屋9000余间。是世界上现存规模最大、保存最为完整的木质结构古建筑之一。

故宫每一扇门上的锁正面都印有“故宫专用”的字样,并标明了使用房间的名称及编号,却没有锁孔。其实,这种锁是专门为故宫开发的一款集各种功能于一体的智能电子挂锁,故宫共有这样的专用锁2750把。

该智能电子挂锁体积小、功能多,可供电、防潮、防摔、防砸、防翘,还有无线通信等作用,在低功耗、防干扰方面做到了世界领先,续航能力可高达两年,机械寿命达到60万次。

在使用上,它不仅不需要与文物接触,也不需要任何布线。它可以联动周围摄像头,实时抓拍现场画面,并能实现实时监控门禁的开闭状态、自动识别开门者的身份,对非法开锁等异常情况进行自动报警。(本报综合)