



视觉中国供图

在未经充分训练的情况下,让AI关注人是容易的,但关注某个特定的人是困难的。

图灵奖得主发表机器学习最新观点——AI进化必须修复“注意力Bug”

本报记者 谢开飞 通讯员 许晓凤 王忆希

在美剧《西部世界》中,AI主人公觉醒,并意识到这个世界是人类杀伐决断的乐园,于是开启了反抗之路;电影《黑客帝国》中,AI将人类养起来,控制了整个世界……那么在科幻世界中拥有了意识的AI,在未来生活中能否实现?

日前,在2020年国际学习表征会议(ICLR)

上,图灵奖得主、蒙特利尔学习算法研究所主任约书亚·本吉奥(Yoshua Bengio)对AI的未来提供了最新见解。他认为未来机器学习完全有可能超越无意识,向全意识迈进。而注意力机制正是实现这一过程的关键要素。

人类的注意力机制和人工智能的注意力机制是否一样?现在应用于人工智能的注意力机制还要解决哪些bug,才能让AI真正具备注意力?就此,科技日报记者采访了有关专家。

已经使用了注意力机制来提升翻译质量。在文本问答、对话系统、知识图谱抽取等自然语言处理任务中,注意力机制基本成为了模型标配。在计算机视觉中,注意力机制的使用也成为了一种趋势,包括细粒度检索、图像分类、行人重识别等在内的多个研究任务中都有使用注意力机制的视觉模型。

既然说注意力机制来源于人类的视觉注意力,那么人类的注意力机制和人工智能的注意力机制是否一样?

柯道认为,人工智能的注意力在一定程度上是人类视觉注意力机制的仿生,在功能上两者是近似的,但实际上有很大的不同。

“从过程上来看是非常相似的,都是在给定任务(查询)内容时,帮助人(或模型)去关注最相关的信息。”纪荣嵘举例说,如在被问及“桌子上有什么东西?”的时候,人类就会聚焦到桌子的区域。同样,在视觉问答任务中,人工智能模型也会聚焦到包含有桌子的视觉区域,然后回答“有什么”。

纪荣嵘指出,从原理来说,人工智能的注意力机制主要是通过计算查询信息与给定特征的语义关联程度,来获得模型的注意力分布,而人类的认知系统则可能会更加复杂。

“人工智能运用注意力的过程实际上是一个记忆查询的过程,它很快很精确,但无法发现异常。”柯道举例说,如让AI的注意力多次从同一张图片上获取关注区域信息,结果永远是一样的;但如果让人类来做这件事,人类看第一次的时候关注自己感兴趣的区域,但随着观察图片次数增加,人类可能会关注这张图片的其他区域。

柯道指出,此外人类的注意力很容易就可以做到去关注某个特定的事物,比如人群中的某个人,但是AI注意力没办法在未经充分训练的情况下做这件事,让AI关注人是容易的,但关注某个特定的人是困难的。另外还有一点最大的差别是人类注意力会关注从未见过的新鲜事物,而AI注意力机制不会。

目的在于减少对无用线索的关注

什么是注意力机制?注意力机制来源于人类的视觉注意力,即人类在进化过程中形成的一种处理视觉信息的机制。“中国科学院自动化研究所研究员张兆翔说。

张兆翔解释,人类视觉系统以大约每秒8.96兆比特的速度接收外部视觉信息,虽然人脑的计算能力和存储能力都非常有限,但却能有效的从纷繁复杂的外部世界中有选择地处理重要的内容,在这个过程中选择性视觉注意发挥了重要的作用。如我们在看一个画面时,会有一处特别显眼的场景率先吸引我们的注意力,这是因为大脑对这类东西很敏感。

视觉注意力机制的理论研究自20世纪80年代以来一直是神经科学和认知科学的热点研究问题,目前研究者一般按照两类标准对注意力机制进行分类。“从注意产生的方向划分为:自下而上数据驱动的注意、自上而下任务驱动的注意、两者结合共同产生的注意;从关注的对象划分为:基于空间的注意、基于特征的注意、基于对象的注意。”中国科学院自动化研究所副研究员王威介绍说。

“而对于人工智能来说,注意力机制是机器学习中的一种数据处理方法,广泛应用在自然语言处理、图像识别及语音识别等各种不同类型的机器学习任务中,其主要功能是减少对无用线索的关注。”福州大学数学与计算机科学学院院长、福建省新媒体产业技术开发区副主任柯道博士说。

厦门大学科技处副处长、人工智能系教授纪荣嵘说,基于约书亚·本吉奥的这套理论后续又发展出了多步注意力机制、跨模态交互注意力以及时下非常火热的Transformer网络中的自注意力机制。

一般认为,深度学习中的注意力机制最早是约书亚·本吉奥等人在2014年提出的软注意力机制。基于约书亚·本吉奥的这套理论后续又发展出了多步注意力机制、跨模态交互注意力以及时下非常火热的Transformer网络中的自注意力机制。

“目前研究的AI注意力机制各有各的特点,但总的来说就是构建输入与输出,或者构建输入数据在不同层面上的关联关系,从而达到重点突出部分区域或者数据间的部分关系,提升任务性能。”张兆翔说。

未经训练难以注意新鲜事物

最近几年,注意力机制已被应用到了人工智能的许多细分领域中。“人工智能是可以被设计出不同程度的注意力的。当需要不同程度的

注意力时,我们可以设计多个注意力网络。”柯道说。

据了解,从2016年开始,谷歌的翻译系统就

与人类注意力相比差距明显

据了解,认知神经科学对意识的定义是“对一个人内在思想的感知,或者是对外部事物或内在事物的察觉”。意识形成的过程是一个信息不断筛选、处理以及演化的过程。

“约书亚·本吉奥在2017年的时候曾尝试用机器学习的方法来证明这个理论,其中一个关键环节就是使用注意力机制去选择‘意识状态’。所以,从信息提炼的角度来说,注意力机制对形成机器意识是非常重要的。”纪荣嵘说。

类比人类思维,AI注意力机制靠直觉还是靠推理?

“目前来说,人工智能的注意力机制更接近直觉。”柯道说,大部分注意力机制,是在训练过程中重复告诉AI应该注意哪些地方,哪些东西是有关联的。如在训练AI的时候,不断让AI学习关注猫在河边吃鱼的区域,AI再看到有猫特征和鱼特征的图片时,就会关注猫和鱼区域,而不会看到猫,看到河推理出可能有鱼,然后再去

找鱼,再关注的区域。

“近期的注意力模型研究也有倾向于对推理能力的提升。”纪荣嵘举例说,如近年热门的Transformer模型就利用了多层的注意力网络来不断提炼和处理输入的信息。在这一过程中,每一层的注意力结果都是在变化,从某种程度上来说,可以认为模型是在推理。

现在应用于人工智能的注意力机制还要解决哪些问题,才能让AI真正具备注意力?

“首先需要增加模型的知识储备。当模型有充足的知识后,才知道哪些信息需要去关注,例如,模型如果没有见过飞机的话,就不能很好地执行关于飞机描述的查询。”纪荣嵘说,另外还需要对注意力模型的结构进行改进,提升模型的推理能力,这样模型才能够在复杂的语境下运用注意力完成复杂操作。

柯道认为,就目前来说,AI注意力和真正的人类注意力距离仍然很遥远。

谷歌医疗在泰国“翻车” AI产品落地考验的不只是技术

本报记者 唐芳

近日,谷歌健康团队发布最新报告,显示其2019年开始在泰国测试的一款人工智能系统表现出强烈的“水土不服”。这款人工智能系统通过观察视网膜照片,来判断患者是否患上糖尿病性视网膜病变,然而在泰国11家诊所落地后被排斥。

报告显示,超五分之一的图像因为清晰度问题

被系统拒绝识别,网络不畅造成图像迟迟无法上传,还有更多的时候系统无法给出明确的诊断结果……由于上述原因护士们不得不重拍照片,患者也辗转至其他医院就诊。号称90%的准确度和几秒钟内给出专业结论的谷歌人工智能系统,显然在泰国不灵了。

远距离制约沟通、磨合效率

对此,清华大学附属北京清华长庚医院眼科主任胡运韬认为:“美国工程师很难每次都到泰国考察出状况的原因,在实验室环境下结果不错,到泰国可能就不是那么回事了。”他分析道,眼底相机性能、人员拍摄能力、网络速度,都是导致人工智能系统“水土不服”的原因。好比开一辆豪车到乡下去,到山坡上颠两下容易颠坏。



在近期广州某行业展会上,展出的应用在医疗领域的智能机器人。视觉中国供图

表示,目前人工智能系统在眼科领域最适宜的场景是体检和门诊及社区筛查,通过初筛把患者尽早分诊至二级以上医院进行下一步治疗,这将是减少糖尿病致盲的有力措施。

“不清晰的图像会对医生诊断带来挑战,对人工智能来说也一样。”胡运韬认为,图像质量是AI医疗落地的关键因素,眼底图像首先应让专业医生能看清、能判读,这也是人工智能去识别和判断的前提。要拍出清晰的图像牵涉到一系列现实问题,如眼底相机性能、拍摄环境及技师操作水平,而真实世界中病人情况多样也是一大难题,年龄大小、是否患过眼病、是否受过眼伤、是否做过相关手术等,都会导致所拍摄眼底图像的清晰程度有差异。这要求在产品设计中工程师要到现场和医生反复沟通、测试和不断磨合。

胡运韬还强调,人工智能系统对图像的接受“门槛”值得关注。对图像质量要求太严苛,稍微模糊一点就给拒绝了,那么人工智能系统的落地就比较困难;对图像接受度较为宽泛并且识别能力强,落地自然更为顺畅。

“水土不服”是AI医疗落地通病

10年前,胡运韬致力于糖尿病防治工作时就期待出现一款读片系统,通过初筛减少人力超负荷工作实现快捷诊断。近几年,医院与各大科技公司开始联合研发人工智能检测系统,让这一想法得以实现。

而谷歌遇到的问题,对于国内大多数人工智能医疗研发人员来说并不陌生。“谷歌在泰国面临的问题,我们在一线试点时基本上都碰到过。”国内人工智能创业公司Airdoc医学总监王斌对科技日报记者说,适配各种相机,培训各种水平操作人员,面临各种网络环境,是人工智能系统落地绕不开的难题。但相比谷歌的远程操控,国内研发人员通过不断的磨合与测试,这些落地困难基本克服了。目前,其开发的系统已通过临床试验进入最后的审评阶段。

王斌表示,一些落地问题看起来简单,却需要实地考察才能发现。如在一些医疗条件下,由于没有很好的暗环境而导致拍照时容易产生不合格的眼底图像,对于这个问题,其团队设计制造了一块小挡板,在眼睛和镜头之间的局部位置形成暗环境,成功拍摄合格图片。

此外不断推进医生和护士等对系统的理解也是落地需要解决的问题之一。王斌表示,对于这个问题,可以在系统中设计一个自动的质量控制预警模块,当系统发现质量不理想的图像时就能追踪原因,便于工程师和医护人员的交流和讨论。

在王斌看来,国内外AI医疗的发展趋势良好,只是谷歌开发团队和试用医生未能处于同一环境,交流改进的效率随之受到影响,再加上泰国当地的网络环境、设备环境、现场培训条件不够理想,多种因素叠加导致了AI系统暂时的“水土不服”。

瞭望站

23年前“深蓝”落子 体育棋局AI走向何方

新华社记者 王恒志 王浩明 王楚捷

23年前,IBM公司研制的人工智能系统“深蓝”在6局比赛中3.5:2.5击败国际象棋世界冠军加里·卡斯帕罗夫,将“人机大战”的话题推上一个高潮,要知道,仅在一年之前,它上个版本还是卡斯帕罗夫的手下败将。

虽然这在当时已经是爆炸性新闻,但绝大多数人应该想不到,仅仅20年后,随着李世石、柯洁先后负于“AlphaGo”,被视为“人类智慧最后堡垒”的围棋,也倒在了人工智能面前。

“人机大战”引爆了关于人工智能的舆论场,但在体育领域,“人机对抗”并非人工智能的归宿和目的。时至今日,人工智能已经融入体育领域的各个方面,并在继续以难以想象的速度发展、“进化”。

从体育竞赛本身来说,人工智能指导下的运动训练和比赛已经成为不少运动员和运动队的“标配”。基于计算机视觉的人工智能在运动员和球队技战术方面的分析,开始让训练和比赛中的决策更为高效和科学。近年来,美职篮金州勇士、英超莱斯特城等不少职业俱乐部都在基于大数据的人工智能助力下,取得了令人瞩目的成功。竞赛中的另一个应用场景是人工智能的辅助判罚,譬如近年来发展迅猛也引起很大争议的VAR(视频助理裁判)。

此外,人工智能对赛事的媒体传播正产生越来越深远的影响。众所周知,无论是奥运会、世界杯这样的国际大赛,还是各大职业足球联赛,媒体版权都是体育和商业连接的最核心枢纽。基于图像处理的人工智能快速发展,让精彩片段能够自动批量生成。除此之外,马拉松等大众参与的赛事,基于人脸识别等人工智能技术,为跑者定制的视频回放也可以在比赛结束后马上传送到跑者的手机上。

人工智能与体育器材装备的结合,让智能硬件成为体育的“黑科技”。从可穿戴设备到各种身体数据的智能穿戴设备,到与VR和AR结合的健身娱乐设备,都史无前例地改变了人们参与和观赏体育的方式。5G时代即将来临,VR和AR在体育上的应用更加引人遐想。

人工智能的高速发展让很多人始料未及,而它对体育的传统伦理也在产生巨大冲击,譬如VAR,自诞生以来就一直伴随争议。显然,随着科技的发展,体育的内涵与外延也在变化,双方都需要适应与磨合,这些问题终究会成为人类前进的影子。

虽然“人机大战”引人瞩目,VAR引发争议,但人类开发人工智能的目的自然不是在棋盘或电竞场上击败人类选手,或是引发足球场上的“伦理之争”,归根结底,是希望人工智能成为人类的助手。

工欲善其事必先利其器。对体育而言,现在的人工智能就是“术”,最终这些“术”都要为“道”服务——也就是回归体育的本源。

就如同“AlphaGo”掀起围棋热一样,人工智能的介入可以令很多普通人更便捷、更清晰、更有热情接触体育。比如:更全方位、多角度的直播点播功能会催生更多体育粉丝群体;马拉松赛事、智能篮球场的精彩个人秀视频,足以令一个爱好者将其作为“社交资本”,进而可能带动更多人参与体育运动;更多的“黑科技”装备,基于大数据的科学健身指导也能让更多人真正享受到运动带来的“正收益”……

人工智能和体育从来没有背道而驰,人工智能也不会让人们变得更懒。相反,越来越强大的人工智能可以帮助我们更顺利地打通“运动—健康”这条通道,或许会成为体育这盘大棋上的“神之一手”,真正实现体育野蛮体魄、健全人格的应有之义。

情报所

诗词“混搭”AI 精准找出专业朗诵者

科技日报讯(记者俞慧友 通讯员朱德军)吟诗作赋不罕见,但在包含了5G、人工智能、物联网等技术的“天心诗吧”中吟诗作赋,是什么感受?近日,为了促进消费、助力复工复产,湖南长沙市天心区区委区政府脑洞大开,将诗词歌赋“混搭”人工智能等技术,让人们体验了一把高科技加持下的诗情画意。

记者现场看到,天心区将诗吧设置在了热点消费区域。诗吧外观颇像电话亭和自助K歌吧。传统文化与人工智能的结合,吸粉力满满,市民排队入诗吧参赛的热情高涨。这其中最吸引人的就是AI诗词评分系统,该系统可以通过背诵完整性、情感流畅度等指标,对诗词进行识别、评级、打分,在众多线上线下参赛者中,精准找出高水准、乃至专业的诗词朗诵者。

因诗词评分涉及复杂的云端算法和大量多媒体数据传输,诗吧也同步进行了“5G+文旅示范”应用,将诗词评分过程的规则数据、结果数据、云端协同计算能力,与高带宽、低延时的5G充分结合,为市民提供数十秒录制、一秒传输、数十毫秒运算的体验。此外,诗吧还建立了一套智能引擎,该引擎能通过开放式接口将商户促销、消费者偏好、商业供应链等信息有机结合,从而完成从文娱到消费的全过程服务。

垃圾分类AI帮你



近日,在中关村街道垃圾分类宣教体验中心,参观者通过与智能讲解机器人互动了解垃圾分类知识。

新华社记者 任超摄