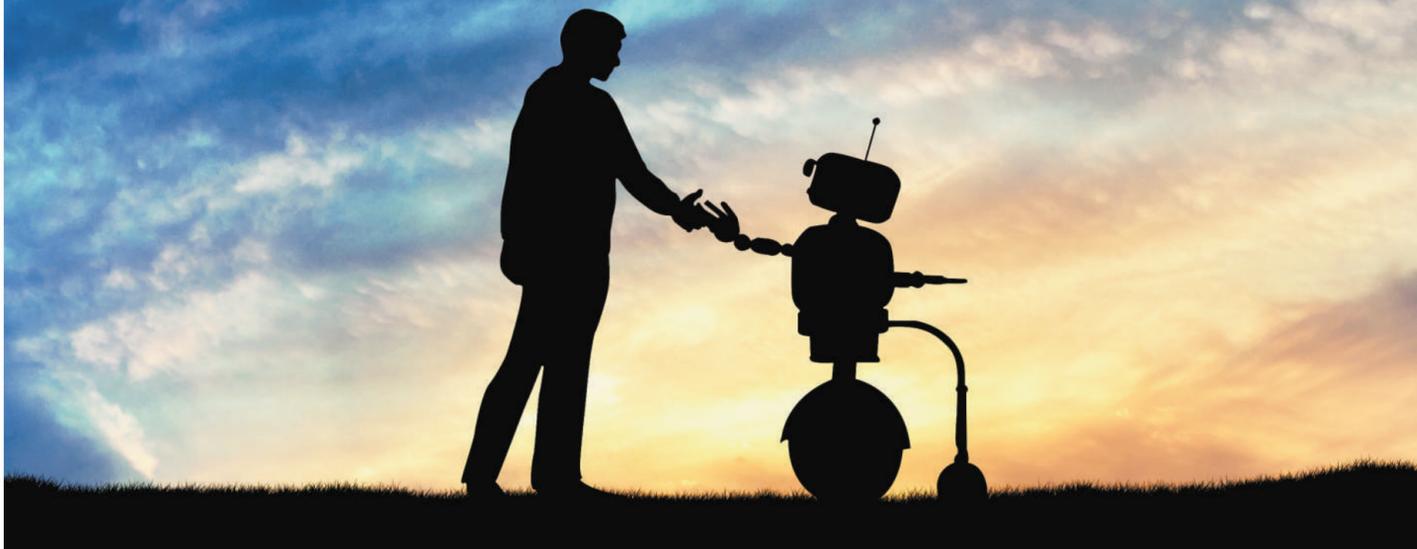


阿尔法狗打败了人类棋手 “阿尔法折叠”们成了科学家助手



本报记者 刘园园

人工智能不仅仅在改变人们的生活方式,也在改变科学家做研究的方式。近日,北京大学人民医院在其官网宣布,该院研究团队利用人工智能方法协助确定中国及

美国初步诊断糖尿病人群的糖尿病分型,为糖尿病的精准治疗提供了理论依据。人工智能学会下围棋、冲咖啡、打乒乓球就已赚足人气,干嘛要做协助确定糖尿病分型这种深奥的事情呢?因为,科学家需要它们。

正在成为得力助手

人工智能感兴趣的深奥领域,不光是协助医学专家确定糖尿病分型。

DeepMind公司之前公布的“阿尔法折叠”人工智能系统,就引起不少关注。与“阿尔法狗”不同,“阿尔法折叠”的特长是通过基因序列来预测蛋白质的3D结构。

这可绝非易事。因为DNA信息只告诉科学家蛋白质的基础构成即氨基酸残基的序列。氨基酸残基会形成长长的链状结构,预测这些链状结构如何折叠成蛋白质成了生物学领域的大难题——蛋白质折叠问题。

偏偏预测蛋白质折叠对科学家而言非常有用:不但有助于理解蛋白质形状在人体中扮演的角色,还有助于诊疗与蛋白质错误折叠有关的疾病,如阿尔兹海默症、帕金森综合征等。

传统做法是靠大量实验来确定蛋白质结构。DeepMind公司的解决之道是,对深度神经网络进行训练,使“阿尔法折叠”能够根据基因序列数据来预测蛋白质的物理特征,包括蛋白

努力拥抱人工智能

其他领域的科研人员也在努力拥抱人工智能。南京大学现代工程与应用科学学院教授李涛正尝试将深度学习算法应用在超光子技术

领域。“超光子技术是通过每一个纳米结构单元对光的调制来实现特定的光学功能。”李涛介

质内部两个氨基酸之间的距离,以及连接氨基酸化学键的角度。最终实现精准预测蛋白质3D结构。

与北京大学人民医院的科研团队一样,不少科研人员开始把人工智能当做得力助手。

北科院北京市计算中心副研究员裴智勇告诉科技日报记者,他已经与一些医院合作,运用人工智能算法进行了几项医学领域的研究。

其中一项研究是判断肾病人是糖尿病肾病还是非糖尿病肾病。因为二者虽然都是肾病,但致病机制不同,治疗方法也不同。传统的判断方法是做肾穿刺,但这种方法比较痛苦,成本也高。裴智勇介绍,他们希望通过凭借一些医学检查指标来直接判断。

“我们构建了一个机器学习模型,运用人工智能算法对病人的几十种检查指标进行大数据分析,来预测是糖尿病肾病还是非糖尿病肾病。”裴智勇介绍,在此基础上,他们又筛选出关键指标,最后实现仅凭8个检查指标就达到95%的预测准确率。

绍,以往在设计纳米结构时,需要人工设计一系列复杂的参数。

李涛现在尝试,将已知的一套参数输入到深度学习网络中进行训练,使其学习参数的规律,然后不断调试,最终通过深度学习网络来设计更多的参数。

在李涛看来,当前基于纳米结构设计的超

构材料光子学提供了调控光场的强大手段,它正逐渐从物理演示走向技术应用。而对于实用化的光学器件,需要满足诸如工作效率、工作带宽、成像分辨率、像差和色差等一系列性能参数要求,且不同使用场景对相关参数有不同要求。人工智能算法能大大提高设计效率,在多参数空间的优化上具有巨大的优势。

“人工智能算法的引入将对超光子技术发展,乃至变革光学技术的开发,如无标记超分辨成像、无透镜成像等起到不可估量的推动作用。”李涛说。

南京大学化学化工学院副教授李承辉前不久看到一篇科研论文,介绍如何用人工智能算法来推荐分子的合成路线,这给了他不少启发。他正在考虑如何用人工智能帮助他做化学

帮助科学家在大数据中“采矿”

不过,目前来看,像“阿尔法折叠”这样针对特定科研领域的人工智能应用凤毛麟角。普通科学家要想得心应手地用上人工智能这个工具,还是有一定门槛。

人工智能开源算法,是李涛实现想法的途径。他和学生已经找到一种开源算法,但这种算法对超光子技术研究而言并非最优,因此需要不断调试参数才行。

李承辉则打算找人工智能领域的专业人士聊一聊,学习一下怎么将人工智能算法应用到自己的研究当中。

裴智勇则幸运得多。他的研究领域是生物信息学,是将计算机技术应用于生物学领域的交叉学科,因此裴智勇对人工智能算法的关注比较早。在具体应用中,开源的人工智能算法如支持向量机(SVM)和随机森林等为他提供了不少帮助。

“这些算法本身是开源的,我们只需要根据自己的研究去修改其中的关键参数。也就是说,参数训练是我们团队自己来做。”裴智勇说。

至于人工智能在科研领域的应用前景,不少人非常看好。寒武纪公司总裁陈天石在接受科技日报记

研究。

李承辉告诉科技日报记者,他最近发现一种新的分子内成环反应,希望了解这种反应是否在其他分子内也存在。“要检测这种化学反应是否具有普适性,需要用不同分子做大量实验才行。”李承辉说,假如能利用人工智能缩小目标范围,将大大减少他的工作量。

在知道精确结果的前提下,可以通过输入关键字在化学数据库去搜索相关分子。问题在于,眼下并不知道精确结果。只知道满足这种化学反应的大致条件,比如分子具有某种特殊的功能团等。因此李承辉期待将人工智能算法引入化学数据库中:只告诉化学数据库这种分子的大致特征,让数据库去智能搜索、筛选匹配

的分子,缩小目标范围。

者采访时说,人工智能的本质是提供求解问题的方法,比如在非常大的选择空间当中做出最优选择。这是科学研究经常会遇到的问题,而人工智能恰恰可以帮助科研人员更快更好地解决此类问题。

“未来的人工智能可能会替代科学家的部分思考过程。”李承辉说,比如在化学领域,人工智能可以基于大数据帮助科研人员分析化学材料的性质、特征,向科研人员提供多种合成路径作为参考,并推荐低融合成某种化合物的最优路径等等。

裴智勇的判断是,人工智能算法在科研领域的应用会形成一个产业。以基因领域为例,未来的基因组数据积累和增长速度越来越快,靠人工去处理、计算海量数据是远远不够的。这就需要人工智能算法帮助科研人员在大数据中“采矿”,催生有价值的发现。

“人工智能技术迅速更新迭代,仅靠科研人员自己摸索开源算法是不够的。”裴智勇认为,今后可能会出现不同科研领域的团队与人工智能领域的团队合作,前者实现业务层工作,后者实现技术层工作。同时,科研领域可能会涌现出更多“阿尔法折叠”这样的人工智能应用。

瞭望站

提高AI“智商”院士给出路线图

本报记者 张晔

“目前一些公司研发的人工智能芯片主要还是基于GPU、CPU、FPGA、ASIC等,面向图像、语音识别,以及神经网络、深度学习算法等专项人工智能。”在近日举行的东南大学人工智能学院、人工智能研究院成立仪式暨人工智能高峰论坛上,中国工程院院士赵沁平认为,要提高计算机系统的智商,应从硬件和软件两方面同时发力,其中,研发更为通用的相似计算单元SPU,是跨越逻辑思维与非逻辑思维鸿沟的核心单元。

计算机系统的类人思维能力

“思维能力是整个智力的核心,人类思维可以分为逻辑思维与非逻辑思维。”赵沁平说,其中,比较判断、概念形成、演绎推理、归纳推理、类比推理和辩证思维被看作逻辑思维,而形象思维、直觉思维和灵感顿悟被认为是非逻辑思维。

计算机科技工作者的三大永恒目标是使计算机系统运行更高效、使用更自然、处理更聪明。其中,运行更高效的目标带动了过去几十年计算存储能力的大幅提升,使用更自然的要求则带来了VR技术的大发展,而处理更聪明就意味着利用AI技术提升计算机系统的智商。

然而,目前已实现的机器思维,仍然没有跳出逻辑思维的范畴。“计算机系统的智商可以分为三级。”赵沁平说,一级智商是比较与判断,二级智商是演绎,而三级智商是在二级智商支撑下的各种人工智能程序,比如概念形成、机器推理、专家系统、机器学习、神经网络等,使其具有类人思维能力。

“由此可以看出,人工智能系统的类人思维能力是归纳、类比。”赵沁平介绍说,要提高计算机系统的智商,有硬件和软件两种路线。其中,软件路线是研究知识操作、研发知识操作系统。而赵沁平着重提到的人工智能硬件提升路线,是研发拥有相似计算能力的SPU。SPU的应用,或将使计算机系统能够跨越逻辑思维与非逻辑思维鸿沟。

人工智能进入2.0时代

“1956年,科学家在AI诞生初期时描绘了一个愿景:制造出能模拟人类学习或智能的机器。”东南大学人工智能学院院长与人工智能研究院院长(兼)联想集团首席技术官与高级副总裁芮勇说,目前这一愿景正在以两步跨越的方式向前发展,一是“心无旁骛”的专用人工智能,二是“融会贯通”的通用人工智能。

如果以一辆汽车作比较,芮勇把人工智能算法比作汽车的引擎,数据可以看作是燃油,而计算则是决定汽车跑得快不快的轮胎,行业应用是决定人工智能发展路径的方向盘。他着重提到,目前人工智能的算力正在面临挑战,非GPU AI专用芯片生态系统仍待完善。

“如果讲60年前的人工智能是用计算机模拟人的智能行为。经过60年的发展越来越多人工智能专家已经认识到机器智能和人的自然智能毕竟是不同的,他们是两种本质上完全不同的智能。”中国工程院院士潘云鹤认为,机器智能某些方面比人更聪明,比如说AlphaGo。

“但是人工智能的某些方面是机器智能很难替代的。最好的办法是两者融合起来,结合在一起变成一个更强大的智能体为人服务。”潘云鹤说,这是人工智能2.0希望可以解决的问题。

潘云鹤介绍,人工智能2.0已出现很多技术端倪,包括大数据智能、群体智能、跨媒体智能、人机混合智能和自主智能等。

“在人工智能2.0时代,新的信息流、新的信息空间会给我们的生活带来很多新的变化。”潘云鹤认为,相关企业和机构应该根据自身的需求制定起新一代人工智能发展计划。

新鲜事

宁夏将用AI助推教师队伍建设

为助力宁夏“互联网+教育”发展,教育部近日批复了《人工智能助推教师队伍建设工作宁夏试点工作方案》(以下简称《方案》)。

根据《方案》,宁夏将在宁夏大学和宁夏师范学院创建人工智能应用创新中心,与区内外骨干企业及高水平大学联合建立人工智能助推师范教育实验室;推动人工智能、信息化等新技术与教师专业发展深度融合,开展分层分级的教师智能素养培训;开展宁夏优质学校及全国高水平学校与宁夏贫困地区学校结对帮扶,建立远程同步智能课堂等。

到2022年,全区各地教育决策、教师工作决策大数据系统基本建成,教师教学行为评价、进修提升需求和教学研究评价推介等实现智能化;人事、工资和流动管理等实现信息化;所有公办高中学校全部建成具备人工智能特点的学生职业生涯规划、选课走班和相应教学管理功能的平台;所有农村学校、移民学校与城区学校结对远程同步智能课堂系统全覆盖。

今年,宁夏将先期遴选基础较好的石嘴山市、吴忠市以及银川市西夏区、贺兰县作为实验地区;遴选宁夏大学和宁夏师范学院作为教师培养实验学校;遴选银川一中、银川市实验小学、银川市第一幼儿园等15所条件较好的学校建立人工智能应用实验校。通过编制人工智能等以新技术知识应用为主要内容的校本教材,促进教师积极开展人工智能教育教学。(王迎霞 通讯员陈勇)

(本版图片来源于网络)

AI都能伪造指纹了,生物识别还安全吗

第二看台

本报记者 马爱平

“人工智能技术能够将海量的指纹数据作为‘原材料’,学习到他们的结构特征和细节信息,并且根据一定的规则进行重组,生成仿真度极高的伪造数据。”人工智能行业资深人士孙立斌告诉科技日报记者。

最近,美国纽约大学和密歇根州立大学发表的一篇文章详细介绍了深度学习技术如何削弱指

纹识别的安全系统。

AI能伪造指纹,指纹解锁还安全吗?

生成对抗网络可伪造指纹

“指纹识别,即通过识别手指纹路确认身份。指纹识别虽应用广泛但存在一定的弊端。因为触摸式的验证方式对环境要求高,对手指的湿度和清洁度更有要求,指纹磨损也会造成识别困难;另外一些人天生没有指纹,或者指纹特征少导致无法成像;不容忽视的是,指纹痕迹容易留存,存在被复制的可能性,造假成本低。”旷视科技研究院研究人员范浩强告诉科技日报记者。

此次论文显示,研究人员使用神经网络数据训练基础软件,创建出令人信服的伪造指纹,其图像甚至优于原始指纹素材。“团队使用神经网络技术变体,即生成对抗网络伪造指纹。”论文作者之一、纽约大学副教授朱利安·托吉留斯说。“生成对抗网络是当下非常火爆的一种深度学习算法,它本质上是一种生成式模型,通过对抗性训练,制造带有数据噪声的深度伪造的图片,可用于数据增强,也可用于攻破特定的识别系统。”范浩强说。孙立斌解释,人工智能技术还能够利用人眼和计算机认知方式的不同,在指纹图像中嵌入某些隐藏属性,虽然肉眼看不到,但计算机可抓取这些信息,达到利用伪造指纹图像进行身份识别的目的。并且很多系统没有活体检测模块,无法判定获取的图像是否来自于真实的人体,这一漏洞使得伪造的指纹图像可以通过系统验证。

指纹、人脸、虹膜识别各有特色

范浩强介绍,就生物识别来说,目前常见的应用有指纹识别、人脸识别、虹膜识别等。虹膜识别,目前主要见于部分高端智能手机的虹膜识别解锁。相较于指纹识别,虹膜识别技术通过人体独一无二的眼睛虹膜特征来识别身份,虹膜识别的准确性是各种生物识别中较高的,但相较于其它生物识别技术,虹膜识别硬件造价高,识别过程需配合,大范围推广较为困难,镜头可能产生图像畸变而使可靠性降低。因此,虹膜识别的图像获取和模式匹配都相对不便,实现大规模商用还有许多技术难关亟待攻克。“人脸识别,以计算机图像处理技术从人的面部提取关键特征点,利用已建成的人脸特征模板与被验证者的特征进行对比分析,根据分析的结

果给出相似值。通过这个值即可确定是否为同一人。相较于指纹需要接触,虹膜需要配合的识别特性,人脸识别可自动抓取验证,非配合式的识别更加便捷,适用场景更加丰富。”范浩强说。

多模态融合的识别方式更安全

范浩强分析,人脸识别在评估安全性上有一个维度不可忽略。

“是否易获取。这包括获取应用场景的人脸数据和人脸数据,仅拥有任何单一数据都无法完成识别对比。目前人脸识别商业应用场景中不管是数据采集、调用还是比对等任一环节都需要在用户知情、并且同意的情况下进行。而人脸生物样本的核心数据库是由公安、央行等核心机构掌握,并非一般商业运营者能够有权获得的。”范浩强说,其次是,是否易攻破。这不仅考验算法实力,更重要的是抗攻击能力。“综合来看,人脸识别是目前生物识别领域安全性较高的,当然也并非万无一失,想要实现人脸识别安全,规模化落地需要技术水平、法律法规和行业标准逐步完善。”

孙立斌认为,生物特征识别迅猛发展,不同模态的生物特征有其自身的特性,抗伪造能力也不同。若要更高级别抗击假体攻击的风险,可以采取多模态融合的识别方式,例如步态和人脸一体化识别,能够大幅度增加伪造数据的难度,提升识别系统的安全性。



扫一扫 欢迎关注 AI瞭望站 微信公众号

