人工智能助力科学研究,多家实验室发布医学AI模型——

当AI开始追问“为什么”

■ 本报记者 杨千莹 陈宁

读懂人类基因组中全部30亿个碱基对,需要多久?

在约半个多世纪的时间里,科学家们解码了30亿中的2%;而在人工智能的加持下,只用短短几天,便顺利读完了剩余的98%。这个惊人的对比,向我们展示了AI的强大力量。就在上个月,谷歌深层思维公司(DeepMind)研发的AI模型AlphaGenome登上《自然》期刊封面。这个能一次性读入100万个DNA(脱氧核糖核酸)碱基对的模型,让人们惊觉:生命科学的研究范式已然改变。

事实上,在“人工智能+科学研究”的趋势下,AI与生命健康领域的基础研究早已频繁“牵手”。从2024年诺贝尔化学奖授予能预测蛋白质结构的AlphaFold(阿尔法折叠,谷歌深层思维公司开发的人工智能模型),到谷歌研发出AlphaGenome,全球范围内,AI的角色已从回答“是什么”,转向开始追问“为什么”,甚至能够预判“会怎样”。

这股浪潮之中,浙江的实验室也纷纷借“AI之手”,竞相开发基因组解读、药物靶点开发、疾病预断等医学AI大模型。

走访良渚实验室、中国科学院杭州医学研究所、杭州华大生命科学研究院等科研机构,我们惊喜地发现,在创新的源头环节,医学基础研究已与AI碰撞出诸多火花,触发科学家新的思考。

郭国骥(左一)团队正在进行实验。 受访者供图

“无人区”的探索者

搜索、整合、加速……AI的进行时远不止于此。在医学的源头,AI不仅是辅助提效的工具,更能探索人所不能及之处。

2024年,北京大学刘君课题组及清华大学杨雪瑞课题组与中国科学院杭州医学研究所合作,在《细胞》期刊上发表论文,揭示了癌症发生的新机制,这一发现,离不开AI的强大助力。

细胞的癌变,由一系列基因突变积累而成。RNA(核糖核酸)作为DNA和蛋白质之间的信息传递者,如果携带突变基因,就会助长癌细胞的分裂增殖。

研究团队正是通过AI,精准识别出RNA中突变碱基上特别的“修饰物”,从而阻断携带“修饰物”的碱基与蛋白质接触,进而阻止癌细胞的增殖。

找到这个“修饰物”有多难?

中国科学院杭州医学研究所医学人工智能中心副主任张亮研究员参与了该研究AI算法的开发,他为我们打了个通俗的比方:“就像在地上撒一把沙子,传统的方法是用肉眼观察,找到带有磁性的微粒,而AI就像一块吸铁石。”

张亮说,比如一条30个单位长度的蛋白质短序列,只在其中4种氨基酸中寻找,也有近3万种可能。

张亮用“撞”来形容这个过程——面对如此庞大且无规律可循的数据量,人只能随机试验,而AI却能在算力支撑下精准快速地找到位置。

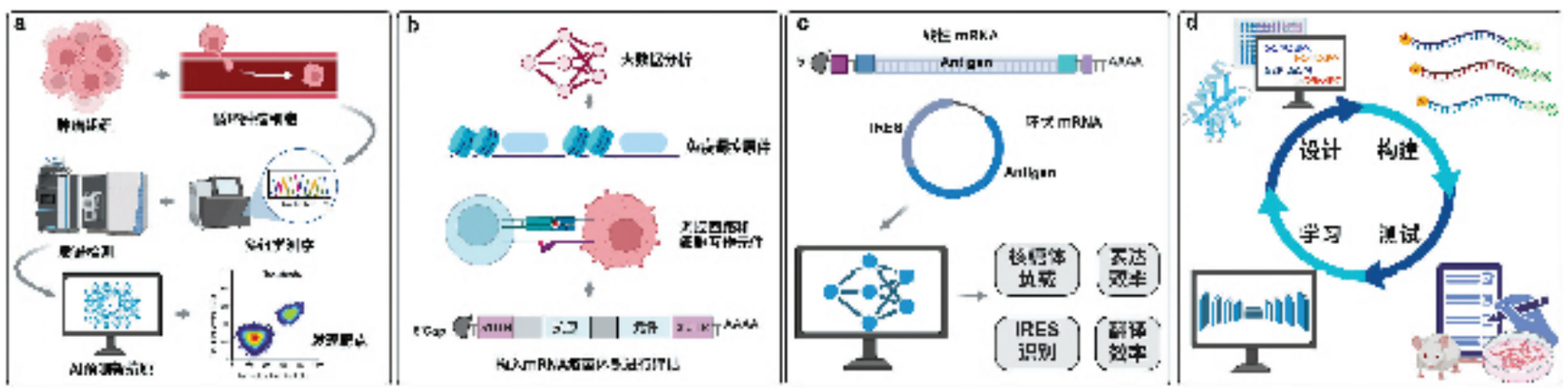
穷举所有可能性,是科研人员以往不会去做的事,而AI耐心高效地补上了生命科学领域一块块细碎的拼图。

这些“无人区”,不仅有人类不会探索之地,还有人类不能探索之地。

在良渚实验室,浙江大学血液学研究所副所长、浙江大学医学院教授郭国骥团队开发的多任务深度学习模型“女娲CE”正一遍遍地运行,演算着更多可能性。

黑底的电脑屏幕上,团队成员、浙江大学良渚实验室百人计划研究员王晶晶轻点鼠标,导入模型,“女娲CE”便开始运转。一行行代码迅速从眼前闪过,不到2秒,模型就给出了19个突变的所有可能结果。

“有些突变在自然或实验中发生几率极低,通过AI可以模拟这类型的突变。”王晶晶打开一张表格,这是模型刚刚算出的预测值、细胞类型和突变种类,接下来,只需按预测值从高到低排序,便能初步得知这些突变对于基因调控的影响。



AI助力肿瘤药物开发原理图。 受访者供图

预测突变影响、读懂“基因调控”,是什么概念?

人类基因组中,编码序列仅占1%至2%,剩余98%都是包含调控序列的非编码序列。这剩下98%,在很长一段时间内无法被人理解。

“2003年完成的人类基因组计划,也只是绘制出一个模糊粗糙的草图。”郭国骥解释,传统的研究方法,是将表型还原到某个基因,但在调控序列中,有无数细小的“开关”,其规则极其庞杂,很难倒推。

而不读懂这些“开关”如何工作,就意味着许多疾病的成因无法被解释。“简单来说,比如一些人的基因上某个位点是腺嘌呤(A),就容易导致老年痴呆,而位点是鸟嘌呤(G)的人群就不容易患病。”王晶晶说,AI能由果溯因,从疾病表型倒推回是哪个位点出了问题。

基于一维DNA序列,“女娲CE”已经能预测其在任意脊椎动物单细胞中的染色质可及性水平。基因表达遵循“DNA—RNA—蛋白质”的顺序,解析表达就像参考“二手资料”,而直接解析DNA开放片段,就像研读“一手资料”,更能找到本质规律。

快速计算“最优路径”

基础生命科学研究领域,答案往往藏在无垠混沌之中。

人类生命,这个高度复杂的系统,包含着无法估量的“数据”。成年男性体内平均有大约36万亿个细胞,而成年女性体内平均有28万亿个细胞。深入细胞之中,蛋白质、RNA、碱基对等更加微观的物质,更是数不胜数。

“以前做一个物种的解读,需要多家实验室一起研究4至5年,但在‘女娲CE’的帮助下,现在只需要1天。”郭国骥感慨。

2025年10月,杭州华大生命科学研究院发布的人类基因组基础模型Genos,其参数已经达到了惊人的百亿规模。

杭州华大生命科学研究院领域首席科学家刘石平告诉我们,当时,Genos的训练数据只用了全球范围内636个人类基因组,而现在,他们正在朝着千亿规模进发。

在极其庞大的数据量面前,“观察—假设—验证”的传统科研范式已经逐渐显露出局限性。AI的介入,为那些过去无法解决的重大科学难题提供了新的方案。

采访时,刘石平拿起一株花解释Genos训练数据的方式:“传统数据分析,就像把茎切成一段段,逐段分析后拼凑成一个图,这些信息是断裂不完整的。而我们用以训练Genos的‘端粒到端粒’水平的基因组,就像直接把整根茎扔进AI模型,得出信息的完整度大大提升。”

更重要的是,人体作为一个多维系统,往往超出人类认知的处理极限。而AI提供了一种对高维度解读的可能性。

“比如磁共振仪扫描人体后,能得出一个静态三维人体图像,但对更高维的动态图像刻画存在明显局限。”张亮解释,高维的信息是无法被低维采集和理解的,除非将高维信息投射到低维——比如三维的球投在二维平面上,就成了一个圆。但这个过程会损失很多信息。要想直接看见肌肉层、血管细节或血流结构等,目前单一数据采集的精度还远远不够。

在纷乱复杂的碎片中,AI可以综合多个维度的信息,在理论上成立、现实中无法模拟的理想环境中,进行推理分析,并得出规律。

上海人工智能实验室双聘青年科学家孙思琦认为,这一能力,让AI在“模式识别”(如看懂病理切片)、“高维搜索”(如在海量化合物中筛选药物苗头)和“生成预测”(如预测RNA三级结构)上已经展现出超越人类的效率,极大缩短试错周期。

2025年,张亮团队以靶点发现和核酸分子设计算法为核心,把AI设计直接导入真实药物研发与临床转化,将传统方法需要数月的分子设计周期,降低到以周为单位。

在肿瘤药物研发中,找到合适靶点是关键步骤。这个过程,首先要快速;其次要精准。

“如果没有AI的帮助,几乎不可能做成这件事。”张亮说。在这项研究中,AI发挥了极关键的作用——精准地在患者的肿瘤组织中预测出“重要的突变”,并据此预测新抗原。

突变,把模糊的疾病症状定格成具体的“锁”,而药物研发就是为这把“锁”配出合适的“钥匙”。

什么叫“重要的突变”?张亮说,肿瘤组织中有大量突变,但只有少部分既能被免疫系统识别,又具备成药潜力。“重要的突变”,意味着它可以被当作潜在治疗靶点或新抗原来源,进而导向新药的研发。“比如甲类型的突变,只有几位患者体内有,那么据此制成的新药只对这几个人有效。而乙类型的突变是通用的,据此制成的新药可以给患有某一种疾病的患者使用。我们的目标就是找到后者。”

传统流程中,从患者数据中人为筛选出合适靶点,需要耗费巨量时间精力。且不同组织、个体间差异极大,这其中包含成万上亿种可能。

浙江大学医学院附属第一医院余杭院区副院长及病理科主任章京教授为我们解释了这个繁复过程:“大约需要十几到二十名研究人员连续工作一个月,才能系统梳理一个小分子的可能路径,且最终选出的靶点未必可用。”而有了AI的帮助,原本将近十年的研发过程,已经被缩短至一至两年。

“它就像现在的地图软件,输入目的地和出发点,AI就能自动计算出几条最优路径。”张亮说。

“模仿游戏”的未来式

“你想来场游戏吗?判断对方到底是机器还是真正的人?”

1950年,“人工智能之父”图灵提出了“模仿游戏”,即著名的“图灵测试”。一直到现在,它都是评估AI智能程度的经典方法。

70多年过去,随着AI不断迫近甚至超越人类智力,图灵的预言似乎已经实现,这让那个不可避免的问题再次出现——

人,还能做什么?

“AI的性能强弱,很大程度上取决于‘喂’给模型的数据质量高低。而判别并挑选出高质量数据,目前只有人能做到。”采访中,多位专家表达了这一相同观点。

数据、算力和算法,是决定AI能力的基础因素。最新数据显示,我国智能算力规模已超过1590EFLOPS(每秒百亿亿次运算),位居全球前列。而目前多数算法均为开源(可公开访问),只有高质量的数据仍然稀缺。

“高质量的数据,对于目前AI性能的影响极为关键。”在训练“女娲CE”的过程中,郭国骥团队自主研发出单细胞百万级测序技术UUATAC-seq。“数据质量提升后,困扰我们很久的瓶颈骤然松动。”

郭国骥说,相比传统细胞系中杂乱的分子信息,单细胞水平的数据,分子信息更多、不容易丢失,且不带人为偏见,非常适合AI理解学习。

多模态数据的对齐、高质量数据的开放共享,已经成为未来的发展方向。

多模态,意味着数据充分、完整、准确、真实、覆盖多个维度。目前,不同实验室、不同医院、不同研究目标下产生的数据,往往缺乏统一标准。即便在临床场景中,数据的记录方式、完整程度也存在巨大差异。这意味着,在数据转化为AI的“燃料”之前,仍需要大量科研人员进行处理。

“从海量数据中筛选高质量数据、将数据转换成AI能理解的语言、在训练中不断调整模型架构、再进行不断评测与优化,这个过程中的每一步都仍然需要人力牵引。”刘石平说。

某种程度上,当前的AI并非“自动驾驶”,而更像是需要人类不断喂养优质养料的超级外脑。

也许,AI在许多已知的问题上已经超越人类能力,但在基础科研领域那些完全未知的地图上,AI仍不擅长完成“从0到1”的过程。

“AI本质上是基于数据的概率模型。”孙思琦说,它擅长解答“How”(如何优化路径),但往往无法独立提出“Why”(科学机制的解释),因此,科学家的直觉和洞察力是不可替代的。

在上海人工智能实验室主任、首席科学家周伯文看来,在全新复杂的科研问题中,AI的预测能力将会遭遇瓶颈,比如AlphaFold能预测蛋白质结构,但尚不能通过分析模型本身来揭示蛋白质折叠的原理。

“AI可以在既有数据框架内判别哪种模式更优,但如果让AI确定一种新方法好不好、可不可行,这还远远不够。”章京认为,跨界联想与创造性假设——这种“跳出数据”的能力,仍然高度依赖人类思维。

而在医学领域,人类的决策能力更显现出不可取代的特性。

“当前的AI模型,的确能做到无限趋近高准确率,但医学不容许‘接近正确’,只要这个数字不是100%,我们就必须慎重,因为承担风险的是患者。”章京说。

目前,大部分AI模型仍属于弱人工智能(Narrow AI),在特定的任务和领域中表现出色,但缺乏跨场景泛化能力,无法自主思考、决策或创新。当前AI界流行的“世界模型”(World Models)、“强人工智能”(General AI)、“通用人工智能”(AGI)等概念,其实都在朝着同一个目标努力——让AI能像人类一样理解世界。

图灵的回旋镖,在70多年后又飞回我们面前。

记者手记

我们始终掌握“驶向何方”

■ 陈宁

在中国科学院杭州医学研究所采访张亮的时候,他很形象地用“车载导航”形容AI之于基础研究的意义。两个地点之间也许有30种路线,“但AI能告诉你‘最优路径’”。

我们顺势发问:“但不管哪种路径,司机得是人,对吧?”他完全同意我们的比喻。

从策划AI助力生命健康领域基础研究的选题,到着手采访、写作,我们不断“提醒”自己:在呼啸而来的人工智能浪潮中,尤其在严肃的科学研究领域,人始终是主导。

但与科研人员对话时,我们发现,这种“提醒”显得有些多余。当他们在实验室里,与这些瞬息万变、看似无所不能的算法打着交道时,自始至终能清醒地认识到AI目前尚存的短板:无法独立决策、创新思维有限、精准度达不到100%……

回顾现代科学的发展历程,我们一次次感叹人与技术的“牵手”给世界带来无限惊喜,但人与技术的辩证关系也几乎贯穿始终——一个多世纪来,汽车从问世到普及,从“蹒跚起步”到日行千里,即便现在新型电车不断颠覆人们的认知,但我们清醒地知道,速度的上限、驾驶体验的边界,始终掌握在人的手上;20世纪初期,抗生素的问世令人惊呼现代医学的伟大,但也正是科学家们及时发现了它的“耐药性”危机,医学才朝着更有利于人类健康的方向稳步前行;今天,各行各业都深度依赖计算机和网络技术的突飞猛进,我们深知,只有人能守住信息安全,防住网络暴力的“漩涡”。

用辩证的眼光看待技术,是源于科学研究“用之于人”的天然属性。在加快打造人工智能创新发展高地的进程中,掌握了方向,我们便能坦然拥抱“AI+”:只有对人工智能的短板、可能产生的风险有足够了解,才能滋养一方足够包容的创新土壤;也只有清醒认识到人的主导作用,才有足够的底气打开更为广阔的创新空间。

本报社址:杭州市体育场路178号 问询电话:0571-85310114 邮政编码:310039 电子信箱:zjrb@zjnews.com.cn 印刷:浙江日报报业集团印务有限公司 定价每月43.50元 零售每份1.50元