

在一年前那场经典的人机围棋大战中，人类战力最强的棋手柯洁以0:3败于阿尔法狗，人类终于被自己的模仿品——人工智能(AI)超越，赛后柯洁的眼泪给了世界深深的震撼。

随后，业界普遍产生猜测，接下来，AI会在哪个领域战胜人类？答案很可能就是医疗。

近日，备受关注的首场神经影像

领域的“人机大战”在国家会议中心落下帷幕，在脑肿瘤和脑血管影像判读比赛中，医疗AI“BioMindTM天医智”最终以高出20%的准确率战胜了来自全球神经系统疾病诊断的25名业界“最强大脑”。据悉，如果这款AI产品投入使用，核磁检查的出片速度将从现在的几天缩短至几分钟。

医疗领域『还是赢了』

1. 电脑“提前交卷”

当场内大屏幕上的倒计时还有14分钟30秒的时候，答题完毕的提示音响了。现场数百名观众不禁发出唏嘘声。原来，需要独立完成225道题目的“BioMind天医智”提前了将近15分钟“交卷”，这时的“人类战队”仍在紧张作答。

正式比赛分为AB两组。A组的一方是15名“人类战队”医生成员，由全国线上预赛产生的优胜者6名、国内神经疾病排名前列的专家7名及国外知名医院专家2名组成。每位医生需要在30分钟内对15张颅内肿瘤的CT、MRI影像进行判读及血肿预测。另一方人工智能BioMind同样需要在30分钟内完成15名医生的工作，即完成总计225道题目的判读。B组共有10名医生，全部是副主任级以上，他们进行的是脑血管疾病CT、MRI影像判读，B组每名医生需要在30分钟内判读3张片子，允许医生们进行判读结果讨论，而BioMind需要在30分钟内判读30张片子。

2. 瞬时掌握大量样本

到底是“最强大脑”胜出，还是这个刚刚问世不久的医疗AI胜出？

此次比赛中，A组试题为高培毅从天坛医院脑肿瘤病例库中随机挑选，B组为北京天坛医院常务副院长王拥军教授从国家神经系统疾病临床医学研究中心颅内出血病例库随机挑选，两组试题均非AI训练试题。

最终现场比赛结果是：A组的225例判读，AI用时15分钟准确率87%、15位医生用时30分钟准确率66%；B组：AI用时15分钟准确率83%，10位医生用时30分钟准确率63%，无论时间还是准确率，AI完胜。

荆利娜对它的速度是有心理准备的，天坛医院多年的病例汇总分析都录入了它的系统里，样本量超级大，而且是瞬时掌握，速度方面肯定比不过它。但真的没想到在准确率方面会输。

面对“BioMind天医智”两轮胜出的成绩，

“人类战队”选手面前都是一台显示器、一张纸质版的答题卡以及一支笔，答题时选手们紧紧盯着屏幕，拨动鼠标查看影像图，不时托着下巴思考。另一侧，“BioMind天医智”的屏幕上，飞快地扫过一张张片子，并快速显示出在它看来每组片子反映的脑部肿瘤疾病。

参赛的人工智能机器，由北京天坛医院与北京安德医智科技有限公司合作研发。拜师北京天坛医院神经影像学中心主任高培毅教授后，“BioMind天医智”已成为天坛医院神经影像团队最“年轻”的内室弟子。

北京天坛医院放射科副主任荆利娜是14号选手。她答题与平时看片子状态一样。15道题目里有一两道拿不准，平时如果遇到这种情况，她会写出两个答案，然后向高年资的专家请教，或者大家一起讨论一下。比赛的时候只能写一个答案，当时想了比较久。

他的老师高培毅说，通过对北京天坛医院近十年来接诊的数万神经系统相关疾病病例影像的系统学习，“BioMind天医智”在脑膜瘤、胶质瘤等常见病领域的磁共振影像诊断能力相当于一个高级职称医师级别的水平，实力不容小觑。每种肿瘤背后，它都学习了1000个病例，目前基本上已经掌握了50种颅内肿瘤，这是任何一名医生都难以实现的。

“对它的比赛成绩不够满意，我认为它的准确率应该在90%以上。”高培毅说，接下来他们将对AI的“丢分”原因进行研究分析。

“机器训练时间不够，如果能够再多给AI一些学习时间，它将会表现更好。”王拥军说，如果AI对于血肿预测的准确率能够超过85%，国家神经系统疾病临床医学研究中心就计划启动临床验证研究。

2018年，“人工智能+医疗”火了。在医疗健康行业，人工智能的应用场景越来越丰富，“准确率超过90%”、“速度击败医生”、一个个关键词仿佛让AI成为了医疗领域的“铁拳”。然而，蓬勃发展的背后，人工智能在医疗领域的应用和推广也面临着诸多问题和挑战。比如，人工智能应用要落地，就需要优质的数据土壤，虽然中国的医疗数据整体量很大，但是具体到某一类医疗问题时还存在数据量不够大的问题等等。

医疗本身是一个非常专业的领域，人工智能技术在医疗应用上的突破离不开医学界的深度参与。所以重要的是，在国家层面有意识地整合资源，只有鼓励医学界、科研单位、企业等多方深度合作，只有积极应对人工智能在医疗面前的问题与挑战，才能进一步推动“人工智能+医疗”的发展，这不仅是社会各界的机遇，也是责任。



IBM开发的AI选手Project Debater(左侧)与人类辩手Dan Zafir(右侧)举行辩论。

3. 电脑诊断能实现吗？

事实上，在这次“人机大战”之前，AI就已经在多个医疗细分领域曾与人类医生交手。2016年5月以来，比分结果是AI 6胜、3平、2负，医生已然落于下风。

2017年岁末，斯坦福大学教授吴恩达领导的机器学习小组开发出一种名为CheXnet的算法，能够更敏锐地捕捉胸部X光片中的肺炎迹象，在诊断肺炎的比拼中，也一举击败四名放射科医师，而CheXnet只经历了一个月的诊断学习。

“AI已经在预测中风和心脏病发作、预测婴儿自闭症的风险上表现出明显优势；在外科手术和阿尔茨海默病预测中略胜一筹；在治疗脑肿瘤、先天性白内障诊断和皮肤癌诊断上，跟人类医生打平。”亿欧智库医疗产业分析师尚轲介绍。

速度快、准确率高，医疗AI表现令人赞叹的同时，更多人关心的是它未来将给神经医学带来哪些改变，人脑疾病可以实现电脑诊断吗？

高培毅介绍，天坛医院一个影像大夫每天读片诊断的时间达到18个小时。按照“BioMind天医智”目前的速度，一个大夫一天的工作量，它只需要400—500秒，也就是不到10分钟的时间。现在患者到医院做核磁，结果都要等第二天以后才能拿到。如果让AI来做，那么核磁结果基本上立等可取，几分钟就行。

“医生工作强度非常大，希望未来AI

能把医生解放出来，让医生有更多时间做研究、给病人温暖，也让患者节省看病成本。”高培毅说。

对于“BioMind天医智”在神经影像辅助诊断领域取得的飞速发展和惊人成绩，王拥军说，它在短短几个月的时间内，不断提升疾病诊断效率和准确率；可以学到很多医院多年都见不到的罕见、疑难病例，在神经领域的研究开发和学习深度上，拥有先天优势和大数据基础。目前已经向国家药监局提交申请，希望在临床中应用这款AI产品，提高基层医院影像诊断准确率，同时也提高影像判读速度，为患者节约时间。

本次决赛的评委之一、重庆医科大学神经科学中心主任谢鹏教授表示：“对已有知识的诊断来说，从大数据深度学习的角度看，AI获胜的几率要大得多，它应该是了了于胸的。但在一些新的、特定的、目前医学界也还没有太多共识的疾病领域，AI可能还‘搞’不过人类。”

“我觉得将来它赢不了我。”荆利娜说，它不了解医学的复杂性，影像判读不是单纯看片子，对患者的病史也要有所了解，包括实验室检查等。不同疾病也会有相同表现，需要医生通过多年的积累综合判断倾向于哪一个，而不是单纯通过核磁片子、通过某个征象来判断。

“我个人并不是很在意这场比赛谁输谁赢。”王拥军表示，本次神经影像人工智能大赛并非意在挑起人类医生和AI之间的战火，希望能够通过这个比赛，能让医生们体验到人工智能的魅力，特别是让部分抱有怀疑态度的医生，进一步对人工智能进行了解。

王拥军表示，具备“天坛标准”的AI技术渐趋成熟后，将成为辅助基层医生，特别是偏远地区基层医生如何阅读、诊断、预测片子的学习和培训工具，方便脑病患者在“家门口”就能获得高品质、个性化的诊疗方案。此外，它能够帮助医生完成初筛和评定，最终由医生进行印证判断，提高工作效率，节省重复机械工作，特别是在判断结果不一致时，可提醒医生避免漏诊误诊。

4. 筑牢大数据“地基”是关键

“离开临床数据，AI没法思考。”北京大学肿瘤医院信息部主任衡复修在很多会议上强调这一金句。

AI的开发很像教孩子，需要花时间训练它，给它喂大量数据，同时告诉它什么是错的，什么是对的。通过这种有监督的学习，AI才能成长。

就像早期阿尔法狗的训练一样，医疗AI的训练也得有“棋谱”——以医学影像为例，就是大量由医生标注出重要信息的影像数据集。不过，围棋有统一规则，而人的病例复杂得多，因此，获得高质量的、经标注的影像大数据集，需要大投入。

万里云医疗信息科技(北京)有限公司CEO黄家祥认识一位AI医疗创业公司的创始人，刚融到几千万投资时十分开心，但不到一年就发现，差不多一半的资金得用在数据标注上。

相对于基因、病理等的数据，获取医学影像数据更容易一些，且本身就是结构数字化的，加之原来就有一些公开的标注数据

集，所以一大批创业公司才蜂拥进入影像领域。

基于同样的逻辑，在AI医疗技术的开发中，最重要的不是AI技术哪家强，而是看谁能与医院建立良好合作，因为医院手中既有医疗数据，又有能对数据进行标注的医生资源。

实际上，中国的医疗大数据一直存在应用障碍，信息孤岛现象明显，国内95%医院的电子病历还未全院流通。换句话说，医疗大数据的地基尚未打好。

在医疗过程中，很多最基本的医学术语尚不能统一，如阑尾炎和盲肠炎或食管癌和食道癌，说的都是一个病，但录入数据库后，计算机就会把它分成两种病。

医疗数据不准确、不完整，增加数据挖掘难度的同时，也降低了数据本身的价值。河南省安阳市肿瘤医院每年完成2200台—2500台的食管癌手术，稳居世界第一。但该院院长徐瑞平教授坦陈，“我们做了这么多手术，在国内食管癌的(学术)地位并不

高。”原因就是数据质量不高，后期对病人的随访不够，导致数据不完整。

要想让AI深入，就需要协调电子病历、化验和影像系统、医生记录和医疗保险材料等多方的大数据，这明显是个难上加难的任务。

即使在先行者美国，也有同样困境。《数字美国》报告显示，美国有近四分之一的医院和超过40%的医生尚未采用电子健康记录系统。即便有电子记录系统，也没有与病人或其他提供者无缝共享数据，因为这些系统无法互通操作，病人需要反复讲述他们的病史。

况且，医疗AI在全球都面临着一些独特的高难度障碍：医疗数据的敏感性和严格的隐私规定，限制了AI医疗所要求的高质量聚合数据的收集。如美国医院对患者隐私有很多保护，医院数据不能轻易开放给AI公司。

(本版文字综合科普中国网、财经杂志、新浪科技等相关报道。图片来源于网络，请作者与本报联系，以奉稿酬)



抓住机遇抢滩登陆

□ 张妮

智能问诊、“刷脸”就医、医疗影像辅助诊断、疾病风险预测……当前，人工智能已日渐渗透到问诊、分诊、支付、影像诊断等医疗服务的多个环节中，它在医疗领域的应用是目前最火热，含金量也是较高的一个领域。有机构调研数据显示：目前国内有近200家从事医疗人工智能的公司，累计融资金额超过180亿元人民币。



那些形形色色的“人机大战”

人工智能 VS 著名作家

今年，谷歌公司研制的AI机器人Champion参加了高考，这次考的是作文。为了吸引更多的关注，谷歌举办围棋人机大战一样，也给机器人找了一个人类对手，内地著名青年作家张一一。

结果，Champion的高考作文《绿水青山图》获得了曾参加过高考作文阅卷的专家评审给出的100分满分，而张一一的高考作文《新时代新青年》仅获85分。

这是高考作文史上的首次“人机大战”，被美联社等海外媒体认为人类历史上第三次现象级的“人机大战”巅峰对决。

人工智能 VS 人类辩手

日前，IBM公司发布了被称为Project Debater的人工智能系统，这个系统与经验丰富的人类辩手之间进行了两场辩论。

参与者必须准备一份四分钟的开幕词，然后是四分钟的反驳和两分钟论证总结。开幕辩论的主题是“我们应该资助太空探索”，以及“我们应该扩大远程医疗的应用”。

辩论过程中，Project Debater严格遵守辩论形式，概述论点，并通过从科学研究和全球事件中获取的事实来直接回应人类对手提出的论点。它还一度比较准确地指出了对手所犯的事实错误。

在这两场辩论中，观众们都投票赞成票给Project Debater系统，认为其在传达的信息量方面表现更好，更具说服力。

人工智能 VS 德扑高手

2017年4月6日，李开复创办的创业孵化机构创新工场发起中国首次德州扑克人机大战，此次比拼由中国顶尖德州扑克对战人工智能程序“冷扑大师”。最终“冷扑大师”最终以792327总记分牌的战绩完胜。

“冷扑大师”是由卡内基梅隆大学Thomas Sandholm教授和他的学生Noam Brown博士开发的，是扑克类博弈程序中最强的人工智能程序。Bridges超级电脑作为冷扑大师的幕后支持，每一天的赛事结束后，位于美国匹兹堡的数据中心会对对决的资料进行分析，优化冷扑大师的策略。