

观五官知罕见病,AI“望诊”靠谱吗

■本报记者 赵广立

将人工智能(AI)用于医疗辅助诊断早已经不是什么新鲜事,仅AI医学影像辅助诊疗一项就有非常多的应用场景。然而,近日来自美国波士顿一家名为FDNA的数字医疗公司的研究人员大开脑洞,提出用深度学习技术仅仅通过“看面部”,就能帮助医生识别罕见基因疾病,并付诸实践。他们的研究于1月7日登上了《自然—医学》杂志,引起业内关注。

中医讲究“望闻问切”,其中望即有“望五官知五脏”“望痰辨病”等含义。如果AI能“学会”这本领,相信很容易获得人们青睐。不过,深度学习是一项比较“吃”数据的技术,而罕见病恰恰是临床医学上数据贫乏的领域。FDNA公司研究人员试图依靠AI望五官辨识基因异常引发的罕见病,行得通吗?

猜测准确率达65%

发表在《自然—医学》上的这篇文章提出了一种深度学习算法DeepGestalt,它通过一款名为Face2Gene的App在智能终端发挥作用。据论文介绍,该App依靠深度学习算法和类脑神经网络来区分人类照片中与先天性和神经发育障碍有关的独特面部特征,然后利用从照片中推断出的模式,定位出可能的诊断结果,并提供可能的综合征遗传病选项。

在实验开始之前,FDNA公司首席技术官Yaron Gurovich带领研究者们首先训练人工智能系统来区分“Cornelia de Lange综合征”和“Angelman综合征”,这两种疾病患者都有区别于其他疾病的明显面部特征;同时,研究人员还训练该模型对“Noonan综合征”的不同基因形式进行分类。接下来,研究者们给该算法模型输入了涵盖216种不同综合征的17000多张确诊病例的图像。

论文显示,在对502张不同图像上的新面孔进行测试时,Face2Gene的最佳诊断猜测准确率达到65%。如果考虑多个预测结果,则Face2Gene的top10准确率可以达到约90%。

这个结果显然有其一定的医学意义。FDNA首席医疗官、该论文的共同作者之一Karen Gripp使用该算法帮助诊断了一名4岁女童的Wiedemann-Steiner综合征——这是由一种叫作KMT2A的基因变异引起的罕见疾病。Gripp在研读了相关病例报告之后得知,该疾病会导致儿童牙齿过早生长。

由于年龄较小,这名小患者除了掉落大部分乳牙并长出多个恒牙,很多典型的躯体症状尚未显现。Gripp将该女童的照片上传至Face2Gene后,软件上出现了“Wiedemann-Steiner综合征”;接



新型人脸识别软件可助力罕见病诊断

着,Gripp用靶向DNA测试进一步确诊了这一结果。

虽然Gripp最终依靠的是DNA检测进行确诊,但这名医生认为,该AI方法可以帮助医生缩小病症范围,节约了昂贵的多基因检测费用。

目前,Face2Gene这款App处于开放下载状态,全世界的医疗专家都可以免费使用它。当然,这也是FDNA公司想要获取更多数据的策略,它需要更多数据来训练模型,使Face2Gene能够熟练地帮助医生辨别患者可能的基因遗传病。

Yaron Gurovich介绍,现在该项目的数据库中已有大约15万张照片。而随着更多医疗专家将病人的照片上传到该App,该项目的准确率也得到略微提升。

质疑:结论并不令人信服

并非没有质疑。在记者的采访中,长年致力于罕见病治疗研究的中科院生物物理研究所研究员刘光慧告诉记者,《中国科学报》罕见病种类太多,AI难以达到100%识别。

一般的罕见病都有特殊面容,比如儿童早衰症及成人早衰症患者,面容都很特殊,通过AI识别有一定道理,这是一个很好的方向。”刘光慧转而说道,“但也有很多罕见病没有特殊面容或面部特征不明显,那么AI可能还不能做到100%识别。”

言外之意,如果在上述4岁女童案例中,换一种面部特征不那么明显的罕见病来测试,Face2Gene则可能会失灵;而如果Face2Gene一开始就判断错了,那么最终到底能不能省下多

基因检测费,则又另当别论了。

更重要的是,该论文中的结论并不十分令人信服。

“502个图的测试集,分了200个类,top10的准确率达到90%——假设数据平均,每个类仅两张图片可测试,这样的结果我不认为能得到‘识别遗传病准确率91%’的结论。”有质疑者认为,该论文结论稍显草率。

更有专业人士指出,该结论有“过拟合”之嫌,在实测中不一定能达到这样的准确率。通常,在一些统计中,为了得到一致假设而使假设变得过度严格,就会出现“过拟合”,而要避免过拟合的发生,通常需要使用增大数据量和测试样本集的方法对分类样本进行评价。

对此,中科院自动化所研究员、生物识别与安全技术研究中心主任李子青(注:该文引用了李子青团队的一篇深度学习研究论文)对《中国科学报》记者谈了他的看法。他指出,这篇文章提供了三个实验结果:其中两个实验是从已确诊综合征的人脸图像中,识别出一种特定的综合征;另一个是区分Noonan综合征的两个子类型。可喜的是,三个都超过了门诊专家的准确性。但这篇文章所解决的问题,是从已确诊的人脸图像进一步分类或区分子型;从深度学习和模式识别角度来看,都不算难。

李子青认为,真正有挑战的是综合征筛查,即从普通人群发现有综合征的人,并对大规模人群诊断达到一定的准确度。这个问题更加有临床诊断意义,并且技术上比该文所处理的问题的难度要大很多。

对于深度学习算法而言,李子青认

为“只要有数据,总可以训练出一个东西”,而至于效果如何,首先是取决于算法——在合理算法基础上,就是拼数据和计算资源。至于如何判断实效?则要看其推广性及泛化性。他表示认同记者的这一观点:罕见病病例少,数据难获取,模型的推广性不容易保证,这种方法的大众普查中应用还不成熟。

技术风险仍存

一般而言,只有训练数据集够好,算法才足够好用。从这个评价尺度上,Face2Gene的技术存在风险。对此,牛津大学计算生物学家Christoffer Nel-lkner很是认同。他认为,尤其是涉及到那些全球患者人数极少的罕见疾病时,数据集的封闭化、商业化“会威胁到这项技术的主要潜在优势”。

此外,训练数据过于集中导致的种族偏见(大部分为白人)也是一大问题。2017年的一份儿童智力障碍研究表明,Face2Gene对唐氏综合症的识别率在比利时白人小孩中为80%,而在刚果黑人小孩中仅为37%。对此,Yaron Gurovich意识到“这个问题需要解决”,但他也认为随着训练数据集变得更加多样化,算法对非洲面孔的识别准确率会随之提升,“偏见会越来越少”。

除了数据量的问题外,数据维度也非常重要。汇医慧影创始人兼CEO柴象飞在接受《中国科学报》记者采访时说,有时更多维度的数据比简单更大的数据更能帮助算法模型更精准地预测结果。

不过,科大讯飞智慧医疗事业部总经理陶晓东向《中国科学报》提出了类似算法在临床上常被忽略掉的另一个技术风险,“这一波人工智能过度依赖于数据,忽略了很多数据之外的信息”。

“深度学习解决问题的基本思路,到目前为止还没有什么太大的突破。”陶晓东对记者说,在医疗领域里尤其如此,“你不可能像ImageNet那样获得几百万、上千万的训练数据”。因此,最基础的医学理论,应该被考虑在内。

“连基本的解剖信息都没有用在深度学习的框架里。”陶晓东说,这是值得目前如火如荼的“AI+医疗”深思的。

柴象飞对此也表示赞同,他对记者说,一些基于经验或常识的推理对于模型更精准地接近真相也非常重,这或许对于数据样本偏少的罕见病模型意义更大。

(本报记者李惠钰对本文亦有贡献)

论文相关信息:

DOI:10.1038/s41591-018-0279-0

计算机体系结构国重首届前沿高峰论坛举行,院士专家指出“春天”里的系统结构应得到足够重视

评选中能兼顾各个专业的特色。

“现在是计算机系统结构的黄金时期。”李国杰说,“如果说人工智能目前处于收获的秋天,那么系统结构研究则处于明媚的春天。在强烈的现实需求推动下,未来系统结构研究一定会得到蓬勃的发展。”

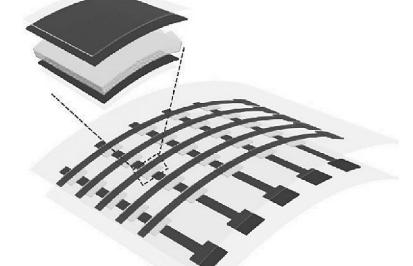
中国工程院院士倪光南、清华大学微电子系主任魏少军以及其他八位来自科研院所和高校的专家在论坛上作了主旨报告,围绕“软硬件安全”“金融计算与图数据体系结构”“人工智能体系结构”“开源芯片”等四大主题展开了研讨,介绍了各自领域前沿的研究成果。国重常务副主任李晓维主持了会议。

该论坛由计算机体系结构国家重点实验室主办。该实验室于2011年10月批准成立,2013年5月通过验收,2017年在科技部信息领域评估中获评优秀类国家重点实验室,主要从事计算机体系结构和系统设计方法领域的应用基础研究。

前沿扫描

柔性透明的可穿戴式触摸键盘系统诞生

本报讯 近期,中国科学院半导体所沈国震研究课题组成功地研制了一种基于特殊微结构银纳米线/PDMS复合电介质层材料的柔性透明电容式压力传感器,相关文章刊登于《中国科学—材料科学》2018年第12期封面。



与采用纯PDMS(一种高分子有机化合物)平面结构的电介质层器件相比,这种有微结构的传感器具有更高的灵敏度、更低的检测范围、更好的稳定性和耐久性。研究人员对导电填料含量和微结构的增强传感机理也进行了讨论。

此外,他们还利用这种材料制备了一个5×5厘米的传感器阵列并用于柔性透明的可穿戴式触摸键盘系统。该键盘系统可以非常灵敏地输出各种信号,这预示着未来这种压力传感器在电子皮肤领域具有良好的应用前景。

随着信息科技的迅猛发展,电子器件开始往柔性可穿戴方向发展。对可穿戴电子设备来说,采用简易制备技术得到具有高灵敏度、快速响应的压力传感器至关重要。



压力传感器至关重要。 (赵广立)

论文相关信息:

DOI:10.1007/s40843-018-9267-3

简讯

工信部苗圩:今年将发放5G临时牌照

本报讯 1月10日,工业和信息化部部长苗圩在接受央视记者专访时表示,今年国家将在若干个城市发放5G临时牌照,使大规模的组网能够在部分城市和热点地区率先实现,同时加快推进终端的产业化进程和网络建设。

苗圩表示,预计到今年的下半年,真正能够具备商业使用的产品将会投放市场。除了消费领域,今年国家还将加快5G技术在教育、医疗、养老等领域应用。

苗圩还表示,特别值得一提的是车联网

网,“将来在路网上,我们也要进行数字化、信息化的改造,将来的红绿灯不光发出‘红绿黄’,也同时发出一个使用5G无线的信号”,将来的智能网联汽车也都可以通过传感器接收到该无线信号,构建起一个车、路、人互连互通的大网络体系。

此前,工信部向中国电信、中国移动、中国联通发放了5G系统中低频段试验频谱使用许可,这预示着三大运营商5G赛道明确,将加速各个环节产品开发,促进产业链成熟。

(赵利利)



位于汉丰湖下游的高阳平湖的“浮标”——水生态环境原位感知系统。

智能技术“把脉”三峡库区生态

■本报记者 甘晓

不久前,位于重庆开州的汉丰湖迎来了600名冬泳者,一场冬泳比赛如火如荼进行。这个因三峡工程而形成的人工湖也成为风景优美的景区,深受游客喜爱。

在科学家眼里,作为三峡库区支流,汉丰湖也是一个条件优良的实验室。日前,《中国科学报》从中科院重庆绿色智能技术研究院(以下简称“中科院重庆研究院”)获悉,科研人员成功研发“三峡库区水生态环境感知系统及平台”,实现了气象、水文、水质等常规参数指标,以及藻毒素、二氧化碳、水质综合毒性等特征污染和生物毒性指标等20多个指标的监测。

中科院重庆研究院院长、该系统研发负责人袁家虎指出:“科研人员采用智能技术研制自动测量传感器和大数据分析平台,为三峡库区水生态安全‘把脉’。”“从地方现实需求出发凝练关键科学问题,并用前沿科学技术解决这些问题。”这是袁家虎的逻辑,也是“面向国民经济主战场”的科学践行。

向环保部门报告“水华”警情

2018年8月,重庆市开州区环保部门收到一封有关汉丰湖夏季水华预警的报告。中科院重庆研究院的科研人员在这份报告中写道:“经过现场两个点位数据测定和样品分析,显示此次暴发水华为蓝藻水华。”水华生物量浓度、水体总氮浓度等数据以及现场观测照片、藻类显微镜照片均详细地记录在报告中。

这次“实战”获得了地方政府的“点赞”。“这份报告背后,是高科技的支撑。”袁家虎指出。

三峡库区蓄水以来,水库次级支

流水体富营养化不时发生,进而藻类大类繁殖形成“水华”现象。袁家虎介绍,过去对水质进行监测,大部分是依靠人工采集到实验室进行测试,已有在线监测设备尚无法实现“水华”预警及实时监控。

2014年起,袁家虎带领科研团队在国家水专项课题“三峡库区水生态环境感知系统及平台业务化运行”支持下,自主研发了多种传感器,并在三峡库区水生生态监测浮标上安装,用无线传输的方式提供监测数据。

据了解,监测指标包括原位藻毒素、水质综合毒性、藻细胞显微观测及水体二氧化碳变化速率等。2017年9月起,科研人员在大宁河、澎溪河、高阳平湖、草堂河及香溪河等四条水华易发支流部署在线监测系统,开展示范应用。

袁家虎介绍,6次数据比对显示,自动监测系统和传统监测系统各项指标结果均具有显著的相关性,监测数据为决策提供了技术支撑。

突破多项实时检测技术

实时、稳定、可靠,是科研人员对该系统监测情况的评价。为实现这一目标,科研人员在实时检测技术上取得了多项突破。

例如,在水质综合毒性检测上,已有水质综合毒性检测仪普遍采用的冻干粉在常温下需要每天更换一次,这不适合野外无人值守。为此,科研人员采用低温技术研发出新型冻干粉试剂盒,更换周期长达10至15天,适用于三峡库区水体现场检测的生物连续培养。

袁家虎介绍,从浮标中获取的大量监测数据传输到这一平台后,借助大数据平台和数据驱动的模型方法,

可实现重点示范区域数据管理、可视化和水环境安全预警。

生态环境局不久前,该平台的应用通过重庆市环保局组织的专家评估。评估中,专家认为,该平台“提高了水质预测的准确性”,“有利于促进三峡水库水生态管理的科学化、系统化”。

此外,科研人员还通过上述监测数据开展了基于水库生命特征演变研究,提出构建耦合生态过程动态表征、数据适时传输、数据分析的三峡库区感知技术与方法体系,为大型水库水环境综合管理提供支撑。

本报讯 1月16日,上海市浦东新区人民政府将与微软通过在传统产业智能制造转型升级、人才培养、商业模式创新等领域不断开拓,打造产业、企业、人才集聚的创新生态系统。

浦东新区政府将支持张江集团构建基于微软前沿技术的“人工智能及物联网服务和技术平台”。借助政府的政策和产业支持,结合浦东产业数字化转型及国际化发展的需求,此平台将为浦东新区企业提供人工智能、大数据、云计算、物联网、敏捷开发、协同办公解决方案等服务。

(计红梅)

中原工学院成立前沿信息技术研究院

本报讯 1月11日,中原工学院前沿信息技术研究院揭牌成立,中国计算机学会区块链专委会主任斯雷明受聘担任研究院院长。该研究院的成立旨在进一步整合学校信息技术领域的学科资源,探索以学科建设为牵引,科学研究、人才培养、军民融合四位一体的机制,建立学科大平台,更好地开展前沿信息技术领域的科学研究和人才培养。

据介绍,前沿信息技术研究院下设先进计算研究中心、网络舆情研究中心、人工智能研究中心和区块链研究中心四个科研机构,实行理事会领导下的院长

负责制。中原工学院党委书记崔世忠介绍说,中原工学院赋予研究院人财物支配权和技术路线决策权,以充分调动科研人员的积极性和创造性,提升教师和科研人员的创新活力,加快推进学校特色骨干大学建设步伐。

斯雷明表示,将借鉴国内外一流院校和研究机构的管理模式,打造河南一流、全国先进的研究院,打造中原工学院科研品牌,培养具有核心竞争力的拔尖科研人才,努力将中原工学院信息技术领域的学科建设、科学研究及人才培养水平提升一个台阶。

(史俊庭 毕明理)