新技术应用中的隐私保护

AI“换脸”打开潘多拉魔盒，如何应对？

■本报见习记者 卜叶 记者 赵广立

近日，一款名为“ZAO”的换脸APP引发舆论关注。用户上传照片，就可以把影视剧主角的脸替换成用户自己的脸。在过把“明星瘾”的同时，也让大众关注到“换脸”带来的个人信息泄露风险。

信息时代，“脸面”不仅仅是一个人的名片，刷脸打卡、刷脸登录、刷脸进站、刷脸付款……面部信息成为个人财产保险箱的一把钥匙。而一旦钥匙掌握在别人手中，将会对个人财产安全等构成极大威胁。

为此，《中国科学报》走访了科研院所、人工智能（AI）科技企业，请技术专家解析以刷脸为代表的生物识别技术背后的安全风险，同时进一步探索：应如何设置保障措施，来应对AI技术打开的潘多拉魔盒？

图片来源：视觉中国

人眼难辨，造假越来越容易

“人工智能带来的安全和隐私问题早已引起国内外的关注，AI换脸只是最近出现的一个新例子而已。”中国科学院院士、清华大学人工智能研究院院长张钹开诚布公地说。

早在20多年前，基于三维模型的AI换脸技术已经走进研究领域，但换脸的效果不尽如人意。近年来，随着生成对抗网络技术的引入，换脸效果大幅提升，人眼已经很难发现换脸痕迹。

换脸的本质是面部生物信息在人工智能领域的应用，此外，人类早已实现指纹、虹膜、掌纹、手形、静脉、笔迹、步态、语音等生物信息的应用。

但是，个人生物信息泄露引发的安全问题从未如此备受关注。谈及原因，清华大学孵化的人工智能公司RealAI（瑞莱智慧）算法科学家萧子豪表示，一方面，近年来大众越来越关注个人隐私问题；另一方面，AI技术发展迅速，技术越来越成熟、效果越来越逼真、操作越来越简单，使得应用领域和场景越来越多。

“以前世界顶尖科研团队才能完成的‘换脸术’，现在寻常人躺在沙发上动动手指就能轻松实现，大大降低了造假成本。”萧子豪告诉记者。

北京邮电大学模式识别实验室教授邓伟洪认为，“ZAO”收集的人脸图像和人脸生成技术如果被滥用，将会带来极大的风险。

此前，鉴于对个人隐私获取的担忧，美国旧金山、萨默维尔市先后禁止人脸识别的应用。

过度攫取用户信息，滥用风险加剧

“ZAO”引发关注的另外一个重要原因是，该应用中含有涉及个人隐私保护的“霸王条款”。

该APP最初的用户协议规定，用户自行上传的内容，一旦发布，平台即默许已获得用户和影片原主人公的肖像授权，且“ZAO”及其关联公司有权在全球范围内“完全免费、不可撤销、永久、可转授权和可再许可”地使用、修改、编辑这些内容素材。用户如果不同意协议，便无法使用该APP。

该规定很快引发公众对个人信息安全及隐私保护的质疑。

与此同时，相关生物识别技术应用需要用户上传越来越多的个人信息。以“ZAO”为例，该APP要求用户提供个人高清图片外，还要求上传点头、眨眼等活体检测的视频。

萧子豪对记者说，软件应用开发方获取大量的个人生物信息后，能够丰富训练数据，用以模型优化，可大大提高应用精度和准确率，进而增强用户体验，但这也意味着个人信息的完全暴露。

今年8月，中国信息通信研究院发布的《人工智能数据安全白皮书》显示，AI应用可采集的人脸、指纹、虹膜等信息具有强个人属性，且具有唯一性。个人数据的过度采集，将加剧隐私泄露风险。

针对“ZAO”APP用户隐私协议存在不规范行为、存在数据泄露风险等网络数据安全问题，9月3日，工业和信息化部网络安全管理局对北京陌陌科技有限公司（以下简称陌陌科技）相关负责人进行了问询约谈。

该局要求陌陌科技“组织开展自查整改，依法依规收集使用用户个人信息”，同时责令其“规范协议条款，强化网络数据和用户个人信息安全保护”。

换脸或能突破支付，数据管理仅靠自律？

“ZAO”引发大量关注之后，“支付宝安全中心”8月31日通过官方渠道发布声明称，不管换脸有多逼真，都无法突破刷脸支付。

支付宝安全中心解释说，刷脸支付采用的是3D人脸识别技术，在进行人脸识别前，会通过软硬件结合的方式进行检测，来判断采集到的人脸是不是照片、视频或者模拟生成的，能有效避免各种人脸伪造带来的身份冒用情况。

话虽如此，但考虑到刷脸支付还没有得到大规模应用，潜在的安全风险可能并未浮出水面。

邓伟洪也告诉记者，尽管目前还没有利用个人面部信息突破刷脸支付等功能的案例，但这并不代表刷脸功能的绝对安全。并且，造假生成技术不断升级，未来存在突破现有刷脸支付功能的可能性。

他介绍了一种可能性：AI换脸软件收集了大量人脸图像，这些信息如果被滥用，技术上可以准确地恢复用户的三维面部信息，结合三维打印模型和面具，足以对刷脸支付等系统进行攻击。

邓伟洪提出，管理部门和学术界应该重视这些风险问题，并制定或研发高效的防御策略。

应如何保障个人生物信息安全？张钹表示，至少应该加强管理和约束，包括数据的管理和合理使用。

他举例，最近一些公司的语音合成技术已经达到足可乱真的水平，这与“换脸”技术一样很容易被滥用。从这点出发，这些公司已经制定了自律规则，对这类AI技术加以限制使用，以保障用户个人生物信息的安全。

“目前，个人生物信息获取多少，又该如何使用，缺乏行业标准。如何制定规则有效管控个人生物信息的使用是当务之急。”张钹告诉《中国科学报》。

今年4月，清华大学成立了战略与安全研究中心，开展国际战略与安全相关的学术研究和政策研究工作。其中，AI安全政策的制定正是该中心的一项重要内容。

“真相”“假脸”难辨，防反之争或将旷日持久

目前有没有技术手段，能够鉴别哪个是“真相”、哪个又是“假脸”呢？

答案是既肯定又否定。邓伟洪告诉记者，由于原始人脸图像特征已经完全被覆盖，目前没有特别可靠的技术可以反推出原始头像。但计算机视觉技术比人眼更精密，可以有效地检测出因为换脸留下的图像痕迹。

他举例说，去年美国国防部发布了一款AI侦测工具，对换脸的检测精度可达到99%以上。加州大学伯克利分校的科研人员也提出了通过特定人物的面部行为细节，准确地“揪出”模仿各种名人的假视频。

邓伟洪所在的科研团队也据此进行了初步的实验。实验结果显示，假视频的检测率可达到98%以上。

不过，在见识了美国国防部研发的“反AI变脸”刑侦工具后，达特茅斯大学的数字取证专家Hany Farid认为，目前的一个关键问题是，机器学习系统可以接受更先进的训练，然后超越当前的反变脸工具。

这也就意味着，这些“以AI攻AI”的反变脸工具仅仅是个开始——AI视频伪造者和数字刑侦人员之间或将展开一场长期持久的“AI军备竞赛”。

技术滥用频频，何以解忧引人深思

引发个人信息安全及隐私保护问题的AI技术，不只有换脸。

近日，一组课堂上学生行为分析的视频截图在网上引发热议。照片中，摄像头通过计算机视觉算法获取学生的课堂表现，并以标签的形式计算出“听讲、阅读、举手、趴桌子、玩手机、睡觉”等动作次数。截图显示该技术方案来自旷视科技。

尽管旷视科技随即在9月3日回应称此系该公司“技术场景化概念演示”，目前“仅停留在技术展示阶段，尚未落地应用”，仍引得教育界人士直呼，智慧课堂行为分析已变成“全景敞视监狱”。

AI技术发展至今，以“赋能”之姿为安防、医疗、教育、工业等各行各业带来了巨大变化，甚至扮演着颠覆行业的角色。然而，面对一波未平一波又起的技术滥用问题，AI应用背后的个人隐私保护问题、信息安全问题，如何规范使用AI技术，的确发人深思。

美国一些城市以及欧盟的做法是严字当头，甚至出台禁止法案。正如前文提及，在旧金山，政府机构擅自使用人脸识别技术、“靠人脸识别技术获得某些信息”都是被禁止的违法行为；在瑞典，一所高中因试验用人脸识别系统统计学生出勤率，而被瑞典数据监管机构处以20万瑞典克朗（约合14.8万元）的罚款，理由是“数据的所有者及数据的管理者之间存在明显的信息不对称”。

“围绕个人隐私、群体安全问题，对于人脸识别等技术的使用需要一个平衡。”商汤科技联合创始人徐冰在近期的一次采访中对记者说：“这个平衡，就是一个好的规则规范——不仅要去规范人脸识别技术企业，还要去规范使用这些技术的政府。”

“实际上，人脸识别技术所创造的价值是被认可的。”徐冰不希望诸如人脸识别这样的新技术只是简单粗暴地被一禁了之，“作为发明和使用新技术的人们，都要去认真思考这些问题，并且一起努力促使新规范的形成，而不是直接禁止。”

速递

华为与太原市共建60万台台式机生产基地

本报讯 记者近日从华为技术有限公司与山西百信信息技术有限公司（以下简称山西百信）联合主办的山西鲲鹏计算产业发展峰会上获悉，山西百信将与华为携手，在太原市建设年产60万台台式机的大型生产基地，预示着“太原智造”的拥有自主知识产权国产电脑将成为全国公务电脑的主流产品。

据了解，山西百信是太原市的高科技企业，是一家专注于自主可控、可信计算等领域的信息技术企业。

山西百信相关负责人介绍，他们与华为携手合作以来，共同研发的基于鲲鹏架构的“恒山”服务器和“太行220s”台式机，代表了国家信息技术应用创新的最新成果。据了解，“太行220s”台式机是地道的国产品牌，拥有多项自主知识产权，采用的CPU是华为公司鲲鹏920s，内部集成4核或者8核，属于目前较为高端的CPU。而且该台式机支持最大64G内存，拥有独立显卡，在信息处理、技术保障等方面性能优越。此外，“太行220s”台式机安装的操作系统有“深度”和“中标麒麟”两种，都是我国具有自主知识产权的操作系统。在应用终端方面，“太行220s”台式机预装了办公学习、编程开发、图形图像等软件，完全可以满足公务及家庭范围使用者。

记者了解到，该型号台式机目前已在太原市建成一条年产10万台的生产线，随着双方的深度合作和产品升级与产能扩大，将为太原市乃至全国信息技术创新应用注入崭新而强大的发展能量。（程春生 邰丰）

10万家公立机构官网纳入搜索显示保护

本报讯 在百度嵌入“××医院”等关键词，担心上当受骗？以后可稍放宽心了。9月11日，百度在京宣布升级“公立机构官网保护计划”，加强对公立机构官方网站在百度搜索的结果显示保护。通过该计划，网友在百度搜索政府、医院等公立机构时，百度将优先展示经过认证的公立机构官网或相关信息，并对搜索结果标注官方认证标识。

百度将通过“基础数据+搜索策略+前端展现”等手段，来保证公立机构官网的优先呈现。百度搜索产品部负责人表示，截至目前，百度“公立机构官网保护计划”已引入超过10万家公立机构官网，涉及700万个搜索词，覆盖了政府机关、事业单位、医院、殡仪馆、学校、博物馆、景区等公立机构，未来覆盖范围还将持续增加。

其中，在公立医院方面，百度基于2018年上线的公立医院品牌保护计划，截至今年5月已对超过14.5万个公立医院名称（含简称、别称、俗称）进行保护。当用户搜索这些受保护的公立医院名称时，将不会出现其他医院的商业推广。

此外，当网友搜索到还没有认证的公立机构官网时，百度在结果页会给予“暂无认证的官方网站”的提醒，并提示用户登录“百度公立机构官网保护共建平台”进行反馈。（赵广立）

科大讯飞新款智能录音笔瞄准“新青年”

本报讯 9月10日，科大讯飞在京发布讯飞智能录音笔SR301青春版新品。这是继SR701旗舰版和SR501标准版后，讯飞专为青年学生群体打造的支持语音转文字的智能录音笔。

科大讯飞副总裁兼消费者BG副总裁李传刚介绍说，作为讯飞智能录音笔家族中的“新青年”，这款智能录音笔搭载了科大讯飞行业领先的语音转写引擎，具备语音秒转文字、中英文边录边译、重点标记、语音搜索、多平台同步等强大功能，同时更具小巧轻便，便携易用等新特点。

取道消费者端、发力智能硬件，是科大讯飞在2019年的战略新动向。对此，李传刚在会上说：“技术革新与消费升级双轮驱动，将出现巨大的颠覆式创新机会，并催生新的领导品牌。与此同时，AI距离消费者也不再遥远，各种AI产品开始变得触手可及。”

据悉，讯飞智能录音笔SR301青春版从8月26日在京东开启预约，到9月10日正式售卖，预约量已突破10万台。（赵鲁）

中关村政策“组合拳”促第三代半导体产业发展

本报讯 支持企业研发创新、支持先进工艺成果转化和产业化、支持建设产业孵化体系、支持产品示范应用、给予最高不超过2000万元资金支持……近日，中关村管委会联合北京市顺义区政府发布4方面17项举措，以促进第三代半导体等前沿半导体产业在中关村顺义园集聚发展。

中关村管委会主任翟立新表示，第三代半导体等前沿半导体产业是战略性、基础性、前瞻性的先导产业。随着第五代移动通信等产业的迅猛发展，第三代半导体等前沿半导体产业具有巨大的应用空间和发展潜力。

中关村拥有清华大学、北京大学、中国科学院半导体研究所等全国第三代半导体领域一半以上的科技资源，顺义园在第三代半导体衬底、关键装备和配套材料及产业服务等环节均有企业布局，初步形成了一定的产业基础。为促进产业集聚发展，中关村管委会与顺义区政府共同出台了《关于促进中关村顺义园第三代半导体等前沿半导体产业创新发展的若干措施》（以下简称《若干措施》）。

中关村管委会副巡视员刘航介绍，《若干措施》主要包括四方面内容：一是降低企业生产经营成本，包括研发投入资助、成果转化及产业化资助、科研成果补贴、首购首用奖励等措施，鼓励企业自主创新；二是通过高层次人才奖励、购房租房补贴等措施，吸引产业高级人才；三是通过支持企业上市、融资贷款补贴等措施，提高企业上市积极性；四是通过支持搭建公共服务平台、孵化器、组建成果转化基金等措施，完善前沿半导体产业生态。

《若干措施》显示，为支持第三代半导体等前沿半导体设计企业研发创新，可根据上一年器件验证流片或掩膜版制作等实际发生的费用，给予最高不超过2000万元的资金支持；为支持前沿半导体先进工艺成果转化和产业化，可根据上一年零部件采购或原材料购置等实际支出，给予最高不超过2000万元资金支持。

此外，为支持前沿半导体领域发明专利布局、产业协同创新平台建设、新产品示范应用、顶级人才和创业人才吸引、高精尖项目落地等，《若干措施》也安排了相关资金支持。

《若干措施》还将支持组建第三代半导体成果转化基金。中关村管委会和顺义区政府将发挥北京市科技创新基金作用，加强与知名投资机构、高校院所等主体合作，吸引社会资本设立成果转化子基金，重点投向第三代半导体等前沿半导体企业和成果转化项目。

顺义区区长孙军民表示，该《若干措施》精准指向顺义区企业生产、研发、应用等多个环节，实现从设计、晶圆加工、衬底及应用的前沿半导体产业全链条政策覆盖。（郑金武）

CNNIC再次当选新通用顶级域名应急托管机构

本报讯 近日，互联网名称与数字地址分配机构（ICANN）官网宣布，中国互联网络信息中心（CNNIC）、加拿大互联网注册局（CIRA）、英国非营利域名注册机构Nominet成为新一轮新通用顶级域名应急托管机构（EBERO）。此次是CNNIC连续第二次中标EBERO项目。

CNNIC作为亚太地区唯一一家ICANN认证的EBERO服务提供方，将为我国及亚太地区注册局、互联网行业从业机构及广大域名注册人提供安全稳定、强健有力的应急保障服务。与此同时，CNNIC也成为全球新通用顶级域名行业唯一一家具备四大资质（EBERO、DEA、RDE TPP、TMCH）的服务提供方。

据悉，常见的顶级域名包括“.com”“.net”“.cn”“.de”等，世界上只有少数国家拥有，数量非常有限。2012年，ICANN正式启动新通用顶级域名申请工作，以丰富互联网多样性，为大众提供更多选择以及推动创新。

新通用顶级域名计划的推出，使得全球可注册的新通用顶级域名暴增，达到1000余个，如“.site”“.link”“.zone”等。同时，出现了越来越多新通用顶级域名注册管理机构。

当某域名注册管理机构遇到技术和运维等问题时，就需要“头冠机构”快速接管并恢复该顶级域名的运营。为此，ICANN推出了新通用顶级域名应急托管机制。

2013年，ICANN公布首批EBERO提供方，CNNIC在列。2018年，ICANN开放新一轮EBERO项目遴选，经过评审小组一年严谨漫长的评选过程，最终选定CNNIC、CIRA、Nominet三家机构成为EBERO提供方。（卜叶）

前沿扫描

谷歌自动重建果蝇大脑

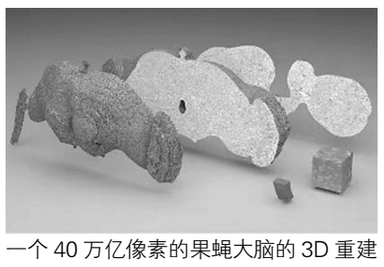
作为“培养”诺贝尔奖得主的“明星昆虫”——果蝇，被认为是人类研究最彻底的生物之一。近日，谷歌与霍华德·休斯医学研究所（HHMI）以及剑桥大学合作，发布了一项最新研究成果——自动重建整个果蝇的大脑。

该研究一共有16位研究人员参与，他们希望通过绘制果蝇大脑完整的神经网络以了解神经系统是如何运作的，最终绘制出人类大脑图像。

为何选择果蝇？该论文的几位合著者在一篇博客文章中指出，与青蛙大脑（超过1000万个神经元）、老鼠大脑（1亿个神经元）、章鱼大脑（50亿个神经元）或人类大脑（1000亿个神经元）相比，果蝇大脑相对较小，它具有10万个神经元，因而更容易作为一个完整的回路进行研究。

实际上，这并不是第一次对果蝇大脑完整绘制。早在今年1月，麻省理工学院（MIT）和HHMI科学家们就已经成功对果蝇的完整大脑进行了成像，并且清晰度达到了纳米级，但是，由于使用了显微镜进行成像，这仍然属于人工方法。

在此项实验中，研究人员向果蝇的脑细胞和突触中注入标记物，以标记每个神经元及其连接处的轮廓。为了产生图像，他们用一束电子束击中大约7062个大脑切片，这些电子束穿过除注入标记物部分以外的所有部分。


一个40万亿像素的果蝇大脑的3D重建

要想完整绘制果蝇大脑，首先需要将果蝇大脑切割成成千上万个40纳米的超薄切片，然后，再用透射电子显微镜对其进行成像。这便产生了超过40万亿像素的大脑图像，再将这些2D图像排列成整个大脑的3D图像。

紧接着，由谷歌定制的人工智能加速器芯片——谷歌云张量处理单元（TPU）会运行一种名为洪水填充网络（FFN）的特殊算法，进而自动追踪果蝇大脑中的每个神经元。

果蝇大脑的重建之路并不顺利。FFN当遇到连续切片中的图像内容不稳定或缺少多个连续切片时，则表现不佳。为保持精度和准确性，研究团队估算了3D大脑图像中的切片一致性，在FFN突出显示每个神经元的同时，保持了局部内容的稳定。

除此之外，为了使FFN更好地追踪果蝇大脑的神经元，研究团队还使用了一种名为分割增强循环（SEC-GAN）的人工智能模型，它是一种专门用于分割的生成性对抗网络，以此来计算并填充图像体积中缺失的切片。基于这两种新程序，研究团队发现FFN能够更加稳健地追踪多个缺失切片的位置。

如何实现包含数万亿像素和形状复杂的物体的3D图像可视化？为此，在大脑完全成像的情况下，研究团队使用NeuroGleener解决了可视化问题，NeuroGleener是一个在github上的开源项目，目前被艾伦脑科学研究所、哈佛大学、HHMI、马克斯普朗克研究所、麻省理工学院、普林斯顿大学等合作者广泛使用。

论文的共同作者指出，他们绘制的大脑图像并不完美，因为它仍然包含一些错误，并且跳过了对大脑突触的识别。但是他们预计，分割方法的发展将进一步改进大脑重建。

谷歌表示，HHMI和剑桥大学的合作者已经开始使用这种重建来加速他们对果蝇大脑学习、记忆和感知的研究。

目前，他们正在与珍妮亚研究院（Janelia Research Campus）的FlyEM团队合作，利用聚焦离子束扫描电子显微镜技术获取的图像，创建一个高度验证和详尽的果蝇大脑连接组。（田瑞颖）

相关论文信息：
http://dx.doi.org/10.1101/605634