

# 人脸识别：从“认脸”到“读心”

■本报实习生 徐竟然 记者 卜叶

相由心生、看脸识人，近日，一篇发表在《自然》杂志的研究为这些听起来“很玄”的词提供了科学依据。

12000多位志愿者、30000多张真照片，研究团队利用人工智能(AI)神经网络技术，基于志愿者的静态面部图像预测其人格特征，完成了对他们开放性、外向性、宜人性、神经质与责任感大五人格的测量。研究发现，AI预测的准确率达58%，其中对责任感的预测比其他四个特质更为准确。

“面相”和个性究竟有哪些联系？越来越“智慧”的人脸识别将如何进一步影响人类的现实生活？技术背后的伦理底线又该如何“化虚为实”？为此，《中国科学报》对多位相关专家进行了采访。

## 脸上有太多秘密

外貌与性格，本就有千丝万缕的联系。

基因对脸蛋和性格都有影响，雌激素、雄激素等激素既会影响面部形状，也会影响个性。从后天来看，“我好不好看”等个人对自身面部特征的感知也影响着人的个性发展，“大笑”“皱眉”等习惯性的情绪表达也能塑造面部纹路与肌肉的发育特征。

“早期的人脸识别技术对图像的光照要求很高，且只能识别照片；目前跨年龄段的人脸识别已经实现，如识别儿童和童成年后照片。”中科院心理研究所副研究员王甦菁介绍，“未来的人脸识别要从读脸向读心发展，根据照片、视频、微表情等，判断人的差异，再结合心理学的知识让计算机不仅能够知道你是谁，还知道你在想什么。”

在王甦菁看来，该研究已能证明人脸图像中包含与人类个性相关的信息。其研究结果可以基于人脸图像分析出大五人格特征，将AI预测与志愿者们的自我报告结果进行对比，形成纠错机制，修正预测的结果，确保预测的准确性。

研究人员还归纳了一些更“接地气”的研究结果——从长相预测人的身体健康情况、人际关系、刺身的可能性以及收入高低。研究的逻辑是，外貌能预测人的性格，而性格越开放的人，往往收入越高。



可以利用人脸识别技术分析挖掘人眼难以察觉的信息。

图片来源：Unsplash

## 应用何往

“人脸蕴含着无数有价值的信息。”论文的共同作者、HSE大学科研人员Evgeny Osin表示。

论文指出，AI系统将能够在客户服务、在线辅导等领域发挥重大作用。如帮助老板面试员工，甚至帮助单身人士在相亲网站上更好地选择另一半。同时，考虑到该技术的效率高、成本低，其应用潜力将被大量挖掘。

北京邮电大学信息与通信工程学院教授邓伟洪也对人脸的信息价值表示认可。除身份、表情、性别、年龄等众所周知的属性外，驾驶者的疲劳、撒谎者的微表情、自闭症儿童的表情、某些疾病患者的长相甚至是血流流动都会在脸部留下细微的痕迹。

正如2019年1月《自然—医学》所刊登的一项研究成果，该研究应用了一种深度学习算法，用以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。当将这项辅助诊断方法搭载到智能手机App上，就可以通过扫描儿童的面部照片以检测儿童是否患有罕见遗传病。目前，该技术的准确率已高达91%，超过了人类医生，倘若加以持续改进和推广，就可帮助全球儿童获得更好的医疗条件。

“如何利用图像和视频去分析挖掘这些人眼都难以察觉的信息，是人脸识别的新问题。”邓伟洪告诉《中国科学报》。不仅在安防事业中应用，结合人脸识别所获得的新信息，该技术还可以支撑传统的心理和健康治疗设施的转型升级，并在新基建中落地。

## 伦理底线亟待化虚为实

该研究利用神经网络技术评估了包括人嘴巴宽度、厚度和眼睛高度等在内的

128个人脸特征，以确保数据多样性。但即便如此，面部特征也数量庞大，难以量化，变量多，联系复杂且难以分离。

即使不讨论深层次的“读脸”与“读心”的关系，人脸识别也还面临着若干挑战。

“虽然识别精度提升了，但光线对人脸识别的影响问题还不能完全解决。此外，双胞胎如何准确识别？不同成长阶段对面部造成影响后如何识别？如何在姿态变化较大的情况下识别？如何排除遮挡对识别造成的影响？如何克服室外远距离对识别精度的影响？”中科院自动化所模式识别国家重点实验室研究员雷震一口气指出了一连串问题。

数据是人脸识别的核心。据雷震介绍，因数据库里中青年的信息更多，机器学习训练更充分，也就导致了对中青年的面部识别准确率更高。此外，还因女性的发型、妆容等影响，易呈现男高女低的准确率差异。

因非理想场景下人脸信息收集成本较高，未来还可能需要研究在极端情况下，如在样本量极少的情况下如何完成人脸识别。据邓伟洪介绍，目前他的研究方向正是破解数据偏差条件下的人脸识别难题，用新型的深度学习算法保证不同种族、不同性别和不同地区的均衡识别率，克服文化和地区差异对人脸识别的影响等。

与此同时，因数据导致的人脸识别技术的伦理问题也日益引起人们重视。

王甦菁表示，志愿者信息使用的知情权尤为重要。雷震建议，数据交换、流通的前提是对个人信息脱敏，使关键信息互相独立，防止个人信息被滥用。应当制定相关法律法规，并将数据提取或清洗落实到责任单位。

“虽然人脸识别技术的广泛落地应用涉及到肖像权、隐私保护、种族和性别公平性等伦理问题，但大家不必过于担忧，目前可以通过技术和法律手段进行双重保障。”邓伟洪表示，目前已有面向加密照片的人脸识别新方法，新技术也可以在保护隐私的条件下进行照片处理。民法典中的人格权编草案对此也有明确要求——不得利用信息技术手段伪造等方式侵害他人的肖像权。

相关论文信息：  
<https://doi.org/10.3886/E109082V1>

## 编者按：

5月30日是第四个“全国科技工作者日”，习近平总书记给25位科技界代表回信。5月30日上午，中国科协召开“全国科技工作者日”座谈会，组织科技界代表深入学习贯彻习近平总书记重要回信精神。全国政协副主席、中国科协主席万钢出席会议、宣读回信并讲话；中国科协党组书记、常务副主席、书记处第一书记怀进鹏主持座谈会。叶培建、薛其坤、李兰娟、黄璐琦、陈薇、谭旭光、吴燕生、樊会涛、巩金龙、王海峰、付英波、海星朔十二位科技工作者代表作了发言。

在座谈会上，百度首席技术官王海峰发言称，我们所从事的人工智能事业，正在党和政府的科技战略引领下蓬勃发展。而今国家提出的“新基建”也为科技创新提供了良好的机遇，为科技工作者注入了坚持创新、科技为民的动力。

以下为发言实录(有删节)。



百度首席技术官 王海峰

非常荣幸参加“全国科技工作者日座谈会”，与大家一起学习习近平总书记回信精神，与各位科技界同仁分享个人的一些思考和感想。

“全国科技工作者日”是广大科技工作者的节日。在座的各位专家、全国的科技工作者们一定同我一样，感到由衷的光荣和自豪。同时，感到责任在肩，心中充满了奋进的力量。一直以来，党和政府高度关切我国科技事业发展和广大科技工作者心声，激励和引导科技工作者追求真理、勇攀高峰，弘扬爱国、创新、求实、奉献、协同、育人的新时代科学家精神。

2017年第一个“全国科技工作者日”，我有幸荣获首届全国创新争先奖。这是党和政府，以及科技界同仁对我多年来在人工智能领域坚持创新发展、积极推动技术成果产业化应用的认可和鼓励。结合今年“科技为民、奋斗有我”的节日主题，我想从“坚持创新、科技为民”的角度，谈谈我的思考和感想。

习近平总书记在2016年“科技三会”的讲话中指出，“实现中华民族伟大复兴的中国梦，必须坚持走中国特色自主创新道路，面向世界科技前沿、面向经济主战场、面向国家重大需求，加快各领域科技创新，掌握全球科技竞争先机。这是我们提出建设世界科技强国的出发点。”我们所从事的人工智能领域，正是在党和政府的科技战略引领下蓬勃发展。一方面，从技术自主创新做起，直面领域前沿，逐渐在某些方面达到甚至超越国际领先地位；另一方面，我们也始终坚持科技服务经济社会发展和广大人民群众的宗旨，持续推动人工智能技术产业化进程，加快推进我国经济智能化转型升级。

以飞桨开源深度学习平台为例。大家都知道，深度学习是新一代人工智能的基础和核心技术，而深度学习平台则相当于智能时代的操作系统，是人工智能技术发展和产业化的牢固根基。经过多年的技术积累和自主创新，飞桨在2016年开源，是我国首个开源的深度学习平台，其核心技术自主可控。飞桨的开源开放，使得我国产业界和学术界能够在自己的基础平台上进行创新和技术应用，创新人工智能产业生态。截至目前，飞桨累计汇聚了超过190万的开发者，超过8万家企业在飞桨平台上创建了近23万个模型。在疫情防控的过程中，飞桨平台联合科技工作者、广大开发者等，快速落地智能应用，在疫苗研究、防控管理、知识传播、复工复产等方面发挥科技的作用。这些都成为促进我国产业智能化升级、提升人民生活便捷度和幸福感的关键力量。

当今世界正处于百年未有之大变局，人类正迎来科技文明前所未有的大发展。我国经济正在由高速增长阶段转向高质量发展阶段，迈进经济新旧动能转换的关键时期。国家提出“新基建”，通过科技端基础设施建设，加快新兴科技突破和落地应用速度，为科技创新创造了良好的机遇，也为科技工作者注入了坚持创新、科技为民的动力。

我们将谨记习近平总书记对广大科技工作者的嘱托和鼓励，秉持中国科技人的优良传统，不断凝聚中国科技界团结奋进、振兴中华的磅礴力量，咬定青山不放松，勇攀科技珠穆朗玛峰，为把我国建设成为世界科技强国而接续奋斗。

(作者系百度首席技术官)

坚持创新，科技为民  
在全国科技工作者日座谈会上的发言

王海峰



机器人在外界导航信息缺失的情况下自主导航。  
图片来源：Unplash

识别等任务。

在斯白露看来，类脑智能为人工智能研究提供了新的可能。类脑智能就是要根据大脑运行的内在神经机制，来启发新的人工智能算法和系统。“类脑智能是推动人工智能突破现阶段许多瓶颈的可行途径。”斯白露说。

多伦多大学计算机科学系教授杰弗里·辛顿也曾表示，要解决人工智能的问题，需要在计算机科学和生物学之间架设桥梁。

目前，类脑科学的研究主要集中在理论层面，结合实际并落地的应用较少。

中国科学院脑科学与智能技术卓越创新中心副主任王佐仁认为这项工作很有意义。

王佐仁解释说，首先，该研究实现了

无人机、机器人领域方位与运动信

息的融合，降低了计算复杂度，增强了

模型的可解释性。其次，在外界导航信

息缺失的情况下，实现了一边精确定位

导航，一边自主构建地图。

另外，使用单相机就可以实现对角速度与平移速度的

测量和实时定位，无需实时定位和全方

位图测图。

王佐仁认为，高级人工智能的典型特征是能够把感知到的信息抽象成概念，并组织成结构化的认知地图，实现理解型学习、抽象思维和推理。认知地图是逻辑推理、想象和创造的基础，也是类脑智能达到通用智能必须具备的能力。

“目前，我们已经掌握了一些知识和

方法，但大脑是如何进行高效编码的依

然需要继续深入研究。”斯白露表示，接

下来，研究人员还需要在理论和算法层

面弄清楚大脑是如何将多种感知信息

融合处理，并在海马和相关脑区形成抽象

表达的，以启迪后续的类脑研究。

相关论文信息：<https://doi.org/10.1016/j.jneurosci.2020.02.023>

## 声音

在信息技术高度发达的今天，企业、社会机构等已经收集了海量数据。这些数据在用于国家公共服务时却面临两难，一方面要推进大数据的高效、合理使用，另一方面也要确保个人隐私数据安全和企业经济效益。

为解决这些问题，国内外的信息技术领军企业正致力于研发“数据可用不可见”的数据安全融合技术。该技术主要包括多方安全计算、数据脱敏、差分隐私、可执行环境、联邦学习等内容。使用该技术可有效突破当前数据使用的两难困境，确保数据高效、安全使用。

目前，我国已有一批代表性、原创性、自主性的技术达到世界领先水平，为信息基础设施资源池建设奠定了坚实的技术基础。

为此，我们认为接下来，首先，相关部门应论证“数据可用不可见”的可行性，协调社会化数据进入公共服务基础设施资源池，推动融合应用，让我国大数据发展迈上新台阶。

其次，应扶持和引导国内相关技术先进、发展快速的企业机构，保持我国在数据安全融合技术领域的原创性、领先性，为我国自主建设信息基础设施资源池提供技术支撑，并积极推动社会数据优先用于“防灾”“减灾”“救灾”等领域。

再次，要鼓励数据基础条件好的行业和城市，开展基于数据安全融合技术的资源试点项目，使社会大数据通过技术创新实现按用量和用途使用，且在使用过程中保障数据不被泄露，达到数据“可用不可见”的效果。

(作者系中国气象局公共气象服务中心气象服务首席专家，本报见习记者池涵采访整理)

## 《2019年深圳市机器人产业发展白皮书》发布

本报讯 5月29日，中国科学院深圳先进技术研究院指导发布了《2019年深圳市机器人产业发展白皮书》(征求意见稿)(以下简称《白皮书》)。《白皮书》显示，2019年深圳市工业机器人产值增速虽有所放缓，增长依然稳健；非工业机器人总产值持续高速增长，其产值已经超过深圳市机器人产业工业总产值的三分之一。

据了解，深圳非工业机器人已出现多个应用场景分化，包括家用服务机器人、商用服务机器人、医疗康复机器人等。其中，商用服务机器人产值增长较快，占非工业机器人总产值一半以上；物流机器人呈爆发式增长，后续发展空间较大。

《白皮书》提到深圳市机器人产业发展有四大趋势：一是贸易环境为本土品牌带来更多进口替代机遇，产业链的核心环节自主化程度提高；二是产业链支持机器人更快进入细分市场；三是机器人有望率先进入5G时代；四是5G、工业互联网、人工智能等新基建领域将为机器人产业发展带来技术支持。(刁斐onte)

## 本报讯 《2019年深圳市机器人产业发展白皮书》发布

本报讯 5月29日，中国科学院深圳先进技术研究院指导发布了《2019年深圳市机器人产业发展白皮书》(征求意见稿)(以下简称《白皮书》)。《白皮书》显示，2019年深圳市工业机器人产值增速虽有所放缓，增长依然稳健；非工业机器人总产值持续高速增长，其产值已经超过深圳市机器人产业工业总产值的三分之一。

据了解，深圳非工业机器人已出现多个应用场景分化，包括家用服务机器人、商用服务机器人、医疗康复机器人等。其中，商用服务机器人产值增长较快，占非工业机器人总产值一半以上；物流机器人呈爆发式增长，后续发展空间较大。

《白皮书》提到深圳市机器人产业发展有四大趋势：一是贸易环境为本土品牌带来更多进口替代机遇，产业链的核心环节自主化程度提高；二是产业链支持机器人更快进入细分市场；三是机器人有望率先进入5G时代；四是5G、工业互联网、人工智能等新基建领域将为机器人产业发展带来技术支持。(刁斐onte)