

“动动嘴皮子”，AI就能知道你在说啥

虽然电视剧《狂飙》已经迎来了大结局，但热度仍然没有丝毫减弱，有的网友利用剧中人物创作娱乐视频，也有的网友剪辑其中精彩片段。不过，还有一些“较真”的网友，发现《狂飙》中的一些人物对话，配音和嘴型对不上，于是便想通过人工智能进行唇语识别，还原最初的剧本情节。

然而，AI读唇语并非只能被用于破译“隐藏剧情”。据统计，我国听力残障人口超过2054万，除了主要的手语交流外，读唇语也是他们重要的沟通方式。但人工解读唇语容易受到个人经验、视觉感受能力、语言理解能力等因素影响，在正确率方面差强人意，于是人们开始尝试利用AI技术来解读唇语。

比唇语专家更懂唇语

“所谓AI读唇语，即人工智能唇语识别，其核心技术框架为视觉识别和自然语言处理。”北京理工大学计算机学院副教授、网络与安全研究所所长闫怀志介绍道，具体而言，就是利用机器视觉技术，将人脸从图像中连续识别出来，并提取其中连续的口型变化特征，输入至唇语识别模型，辨识出该人物口型对应的发音，进而输出可能性最大的表达语句。

“视觉识别和自然语言处理分别有着庞大的技术体系和不同的技术路线，但究其本质，都是通过大量的唇语数据来训练AI模型，力求文本输出的准确性。”闫怀志补充道。

近几年，不断有AI巨头开始在唇语识别赛道上作出尝试。谷歌旗下Deep Mind公司就与英国牛津大学合作，研发出了一款AI读唇语软件，通过让AI读唇语软件“收看”数千小时的电视节目来训练其唇语识别能力。有意思的是，在随机抽取200个视频片段的读唇语测试中，AI读唇语软件的准确率达到46.8%，而经过专业训练的人类读唇语专家，准确率仅为12.4%。

为何AI读唇语能够悄然兴起？闫怀志给出了自己的分析：一是强烈的需求牵引，二是巨大的技术推动。从需求牵引来说，唇语识别不仅可为部分残障人士提供方便，更可以在公共安防等诸多领域发挥巨大作用；从技术推动来说，由于AI算法、算力以及数据瓶颈被不断突破，使得AI技术在唇语识别领域取得较大成功成为现实。

众多难题有待突破

不过，闫怀志也表示，目前我国人工智能唇语



识别技术尚处于起步阶段，若想利用人工智能准确地识别唇语，还有很长的路要走。

从语言本身来看，人类语言具有较高的复杂性，在人类话语所涉及的所有音标中，仅有30%左右是直接由人类嘴唇来控制的，70%是难以通过肉眼，甚至是机器视觉区分的齿音、舌音以及喉音。而且，不同人说话的语气、方言、连词、口音，乃至胡须遮盖等因素，都会导致嘴型的细微变化，而恰恰是这种细微变化，会严重影响人工智能对于唇语的识别和判断。

从技术层面来看，人工智能采集唇语的环境通常较为复杂，若想精准识别难度很高。以目前的人工智能技术而言，对于长句、复杂句式等的识别水平不尽如人意，更不用说还存在着多场景识别、多人像唇语识别等问题。

闫怀志表示，只有解决了上述问题，AI读唇语才能得到突破性提升，迈向成熟发展阶段。

人类不同语种之间千差万别，AI能读懂每个语种的唇语吗？

闫怀志介绍，此前较为成功的AI读唇语系统大多仅限于英语模型，这是因为多数AI模型都是基于英语数据训练而得。但是，从技术框架上来说，不同语种的训练模型是基本一致的，或者说可以依赖于同一类技术手段来实现。

当然，为了适应不同语种的唇语识别，也需要作一些适应性调整：一方面要选择对应语种的数据进行有针对性的训练；另一方面，还需要对AI模型进行调整，比如纳入时间屏蔽、优化语言模型以及改进超参数等。

此外，同一语种也会有不同口型，即便口型类似，也可能代表着完全不同的意思。因此，成熟的AI读唇语系统需要大量的唇语特征样本数据，并尽

可能地覆盖多种应用场景、多类型的说话人群，借此来提升训练后的唇语识别模型的泛化能力，提高AI读唇语对于不同口型和不同表意语言的识别准确率。

亟须监管的技术双刃剑

尽管存在种种难题，但仍有越来越多的AI企业开始涉足并计划深耕人工智能唇语识别赛道。目前来看，各大AI巨头的选择不尽相同，具体可分为唇语数据、唇语视频识别、唇语理解等。

闫怀志也表示，目前许多人工智能唇语识别技术领域已实现初步突破，全链条集成前景可期，产业集群正在逐步形成。

从应用场景来看，AI读唇语在社会公益、公共安全等领域都已开始崭露头角。从目前各大巨头的布局以及相关技术的发展趋势来看，AI读唇语预期可在身份识别、国家安全、智慧系统等方面具有广阔的应用前景。“可以想见，由于公益、公共安全、国家安全等领域巨大的潜在需求的牵引，以及AI技术飞速发展的强力推动作用，在不久的将来，AI读唇语有望实现快速推广与深度普及，产业前景十分可期。”闫怀志说。

例如，在安防安监领域，很多安监场景噪音较大或仅有视频信号，无法准确捕捉声音，人工智能唇语识别技术就能派上用场；在身份识别领域，可以利用AI读唇语来实现口型支付密码输入，“动动嘴唇”就能实现身份识别和支付交易；在公共安全领域，利用AI读唇语，可以在各类视频中分析案件当事人的唇语信息，辅助案件侦查工作；在智慧系统领域，可利用AI读唇语来实现“无声胜有声”——只依靠口型来控制智能设备，比如智能家电等。

当然，技术应用是把双刃剑。很多人担心，AI读唇语会使人们对话中的隐私内容遭到泄露，无论当事人是公开发言、窃窃私语或是自言自语。“张张嘴”就被别人窃取聊天内容，仔细想来确实可怕。

闫怀志表示，这种担心并非杞人忧天。AI读唇语导致的隐私泄露，一方面可能是有人恶意进行唇语获取识别，另一方面也可能是正常使用的AI读唇语系统，但其中的存储、使用等环节保护不当，导致相关数据被窃取或滥用，进而对个人权益造成损害。而且，由于涉及到当事人的对话内容，具有明显的方向性，这种隐私泄露的危害性可能要比普通的个人信息泄露更为严重。

因此，闫怀志建议，应从隐私安全保护的角度，在管理层面加强相关法律法规的制定，严格规范和约束AI读唇语的应用场景、范围和目的，加大对技术恶意利用的监管和惩戒力度。此外，还要在技术层面加强AI读唇语系统的安全保护体系建设，以技术手段提高系统的识别精准度，避免技术滥用，切实保障用户对话的内容安全。

来源:科技日报

数字时代要注重适老化

据媒体报道，有老人因使用老年机，无法在线上购买电影票；还有老人因不熟悉故宫博物院网上预约方式，到了门口却无法进入。类似事件时有发生，网友呼吁：数字时代，民生服务应多些“温度”，别让老年人掉队。

随着移动互联网技术普及，民生服务搭上了数字化快车。从生活缴费、看病就医到交通出行、用餐娱乐，智能化操作几乎成为绝大多数生活场景中的“标配”。线上登记取代纸质材料，提前预约免去排队等待，识别人脸就能办理业务……这些新技术、新业态在方便百姓办事的同时，降低了服务成本、提高了服务效率。

然而，民生服务的创新不应忽略老年人的感受。很多对年轻人而言司空见惯的操作，却给老年人带来不小烦恼，甚至成为难以逾越的“数字鸿沟”。“十四五”期间，我国60岁及以上老年人口总量将突破3亿，我国将进入中度老龄化。在此背景下，适老化应成为民生服务的

“标配”。

近年来，我国大力推进数字技术适老化改造，通过优化界面、简化程序、增加功能等多种方式帮助老年人实现信息无障碍，共享互联网发展成果。但适老化不是一时一地、查漏补缺的简单动作，而是持续不懈、日益完善的长期过程。要打破这一道道看不见的“数字藩篱”，需要全社会成员设身处地站在老人的角度，感受他们可能遭遇的困境和梗阻，让适老化理念贯彻民生服务始终。

对于在科技浪潮中步履稍慢的老年人，我们不妨多保留些传统方式，让线下服务担负起兜底保障作用，如景区保留现场售票、餐厅保留纸质菜单、乘公交保留现金投币、看病就医保留人工窗口。这些看似“过时”的服务方式，却直接关乎老年人晚年生活质量，与建设老年友好型社会有着千丝万缕联系。对老年人来说，民生服务多些适老化改造，不仅意味着办事更方

便、出门更舒心，还可以让老年人更加感受到社会的温暖，更体现了全社会对老年人的尊重和关爱。

来源:经济日报

