



视觉中国供图

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
北京理工大学计算机学院副教授、网络与安全研究所所长

读唇语软件识别准确率远超人类

“动动嘴皮子”,AI就能知道你在说啥

◎本报记者 翟冬冬

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,我国听力残疾人口超过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个视频片段的唇语测试中,AI读唇语软件的准确率达到了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。

工业检修领域,智能机器人或将迎来“用工潮”

◎本报记者 李禾

2月24日,由中国特种设备检测研究院园区智能检测监测科团队牵头研发的国际上首台原油储罐在油内检测机器人“海豚号”在第二届特种设备检验检测仪器装备国际博览会发布。据悉,原油储罐每隔4—6年就要进行人工检测、维修复原等检修工作。此次发布的“海豚号”,可以在原油储罐不卸油的情况下检测,实现精准开罐检修,避免无差别开罐检修。

目前,在工业生产领域,传统的人工检测存在工作效率低、劳动量大、检测工人不足、检测环境恶劣等诸多问题,无论检测安全性还是作业效率都有待提升。随着人工智能技术的发展和成熟,以及物联网、云计算和5G技术的普及,智能化的检测机器人逐渐成为工业检测的重要手段之一,广泛应用于各类不同领域。

广泛应用于各个行业

检测机器人是应用在工业生产过程中的产品质量检测环节,以及产品使用过程中的质量检测环节的机器人。随着科技进步,各种设备自动化、智能化水平不断

提高,与传统检测设备相比,检测机器人利用人工智能技术、人机交互技术等,可自动感知、识别、分析、判断产品生产及使用过程中的各种情况,帮助工作人员进行决策,或者直接对出现的问题进行解决。今年1月,工信部、教育部等十七部门印发《“机器人+”应用行动方案》,提出应深化重点领域“机器人+”应用,助力智能建造。研制基础设施建设、巡检、操作、维护、应急处置等机器人产品,推动企业突破复杂环境下的运动、感知、作业等关键技术。

在重庆、苏州等地,智能巡检机器人通过自身搭载的红外测温仪、气体监测系统及巡检系统等,对电缆隧道内的电力设备、变电站等进行巡检;为保障高速铁路“零误差、零故障”运行,动车组智能检测机器人利用深度学习和特征识别相融合的图像算法,采用机器视觉等技术对动车组进行全自动检测,由于在车下作业,不受车顶接触网供电等限制,机检作业效率是人检效率的2.5倍以上……随着各行各业智能化转型,在电力、交通、医疗、能源等领域,逐渐开始出现智能检测机器人的身影。

据悉,在工厂、医院等工作任务较为明确或已流程化的作业环境中,智能检测机器人技术和产品的应用已经相对成熟。但

在对智能检测机器人有特定功能设计要求的,或是作业环境较为复杂的领域,检测型智能机器人的理论研究以及实际应用仍处于起步阶段。

香港中文大学信息工程系教授邢国良说,由于不同行业及不同检测任务的多样性、复杂性,智能检测机器人还没有通用的固定形态,但随着相关技术的不断创新与发展,“我们很快就会看到高度定制化的、模块化的检测机器人方案”。

人机协同才能更高效

对高空的高压电线进行检修等传统的的基础设施检测和维修,依赖人工作业,存在作业量大、难度大、风险性高等问题。邢国良认为,以智能检测机器人代替传统人力进行基建检测,正成为基础设施运维的一大发展趋势。

高速发展的智能检测机器人会不会完全取代人类作业?事实上,人类在面对不同任务时的熟练度、灵活性及多任务组织管理方面具有不可替代的优势,智能检测机器人则适合在较为危险的环境中,执行较为常规且耗时的工作任务。邢国良说,更高效的检测任务应当由人机协同合作完成。目前,人机协同作业模式正在实际检测工作中被不断实践及应用。

上海艾瑞市场咨询有限公司发布的《2022年中国智能机器人行业研究报告》指出,协作机器人的出现使得机器人与人可共享工作空间,人机协作方式更进一步。安全性是人机协作的前提,智能控制是安全性的保障。智能技术的加持,不断推进人机协作模式的演进,未来有望在兼顾成本、产品、性能等因素的同时,实现人机的深度合作。

业内人士认为,在传感器和机器学习等创新技术的支持下,更加智能的检测机器人应用将能帮助人们克服基建故障检修等难题。这些技术的创新发展将在很大程度上提高智能检测机器人作业时的自主性,以及检测工作的整体安全性。

尤其在一些较为紧急的检测任务中,智能检测机器人能够帮助人们在更短时间内检测更广泛的区域。比如应用传感器融合技术,智能检测机器人能将不同波长的电磁波信号拼接在一起,提供由不同信号检测出的、更全面的数据信息;无线电波的反向散射技术,使智能检测机器人在受到物体遮挡的情况下,仍可以完成检测任务。

邢国良说:“在某些高度专业化、较为依赖设备的行业,比如金属疲劳检测领域,智能检测机器人能够做到更准确、更全面、更高效。”

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音,即便口音类似,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读唇语对于不同口音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,在安防安防领域,很多安防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。

家电产品不断升级 智能化更受消费者青睐

◎本报记者 李禾

家电的智能功能备受消费者青睐。根据中国质量协会发布的热水器、油烟机、空调、电视等多个家电产品的2022年用户满意度监测报告,我国家电产品智能化不断升级,智能功能受到用户认可。

根据此次报告的检测数据,电热水器产品智能化不断升级,智能电热水器用户满意度水平高于非智能产品。其中,具有语音控制、手机App控制、自动杀菌等功能的产品用户满意度更高。在用户对产品质量的评价中,智能功能的稳定性、使用方便性得分高于其他产品质量指标。

燃气热水器产品的用户满意度创近五年来的最高水平,其智能功能的稳定性和使用方便性得到用户认可。用户表示,最常用的燃气热水器智能功能是智能调温功能、手机App控制和语音控制等。用户对智能功能的稳定性、使用方便性的评价比上一年提高了2分。与电热水器相同,燃气热水器的智能功能优化,持续提升了用户的使用体验,购买智能功能燃气热水器的用户满意度比非智能的高。

对于油烟机产品来说,烟感应功能成为用户最满意的智能功能。监测结果显示,智能油烟机的用户满意度比非智能油烟机高,用户较常用的智能功能有烟感应、手势控制、温度感应、智能清洗提醒等。

报告监测结果显示,具有视频通话功能的电视用户满意度得分最高,其次是智能语音交互、智能推荐功能电视,用户满意度均高于行业平均水平。用户推荐具有语音交互、视频通话等智能功能电视的意愿也更加强烈。

此外,空调的智能化为用户带来更多便利、舒适。分别有67.6%、51.8%、38.8%和30.8%的用户经常使用空调的手机App控制、自调节控温祛湿调风、自清洗和语音控制等智能功能。其中,得分最高的智能体验指标是“智能功能的稳定性、使用方便性”,其次是“操作简便舒适性”。

中国质量协会表示,目前手机移动端的视频功能不断增强,传统的电视视频功能被分流得越来越多。发挥电视的大屏优势,进一步优化、完善电视的语音交互、视频通话、与其他设备的互联互通等智能功能,弥补交互、应用等短板,适应用户更多需求,是突破目前电视行业困境的关键,智能功能不断完善应是电视场景创新的核心。

“数字员工”上岗 助力全场景化“智慧地铁”建设

◎本报记者 华凌

“您好!北京地铁‘数字1号员工’!很高兴为您服务……”未来,当你在首都机场线某站进站需要帮助时,打开北京地铁App或者小程序,即可享受到站内精准导航、规划最优路线、智能客服帮助等服务。

2月26日,记者从北京地铁创新技术研究院实验室获悉,“数字1号员工”是国家重点研发计划项目——“超大城市轨道交通系统高效运营与安全服务关键技术”示范应用工程基于语音识别、知识图谱、语义理解等技术打造的动态自主学习智能客服。

“该项目的启动建设对于北京地铁的科技创新和管理创新具有里程碑式的意义。”北京地铁技术创新研究院院长魏运介绍。

“数字1号员工”有哪些功能?能为乘客带来哪些体验升级?北京地铁公司技术创新研究院基础设施技术研究中心副主任白文飞介绍:“围绕乘客核心诉求,建设站内导航、多元信息发布、快速进站、智能召援交互、爱心预约等服务功能,北京地铁‘数字1号员工’将实现乘客全时交互精准服务,让乘客出行更加便捷。”

比如,通过“数字1号员工”乘客可享受快速进站、票务安检一体化等服务,进一步缩短进站时间,提高出行效率。

据介绍,“数字1号员工”的上岗,也将深入推动北京地铁服务细节优化。比如依托智能客服中心平台,汇聚线上、线下服务,为乘客提供更加精准化、个性化服务。

此外,“数字1号员工”还可以及时提供预警信息。例如,当遇到极端天气,该系统可接入气象数据,车站的智能检测设备可对站外雨量和水位进行监测,遇到险情时,还能够自动启动相应设备,解决现场突发情况,并且让乘客可以实时接收天气预警信息,避开突发状况。

白文飞介绍:“为了全方位高质量推进此项工程,我们联合了高校、企业、科研院所等,牵头开展这项关键技术研发。面对首都机场线全线即将迎来的应用示范,项目组依托北斗、EUHT(5G)、空间数字化、人工智能、区块链等新兴技术,计划将该线路打造成为全国首条全场景化的‘智慧地铁’示范线路。”



在北京地铁创新技术研究院实验室内,研究人员模拟智慧调度指挥和运营维护。受访单位供图