

视觉中国

# Siri、小度和小冰

## 它们是怎么做到和你相谈甚欢的

本报记者 陆成宽

你知道小冰吗?你呼叫过小冰吗?你和小冰互动过吗?你使用过Siri吗?如果都没有,那你就out了。它们都是当下很火爆的智能语音机器人,很多人都和它们聊过天。

日前,美国投资机构Mangrove Capital Partners发布了2019年《语音技术报告》。报告分析了语音助理服务的日益普及,并预测苹果2020年将推出Siri操作系统。那么,究竟什么是语音技术?它有哪些应用?语音操作系统又是什么?为此,记者采访了相关专家。

编码(LPC)技术的特征提取方法和动态时间规整(DTW)技术,并且使用模板匹配的方法做一些简单的语音识别(小词汇量、特定人、孤立词)。”

从20世纪70年代中期到80年代,语音识别的框架有了突破,统计模型逐步取代模板匹配的方法,隐含马尔科夫模型成为语音识别系统的基础模型。同时,也采用高斯混合模型作为声学模型的主要建模方法,连接词识别和中等词汇量连续语音识别得到了较大发展。

“到90年代的时候,基础的神经网络语音识别模型已经提出。但当时神经网络语音识别模型之所以没能取得较大的突破,主要是因为当时服务器的计算能力不够强以及训练语音数据的量不够多。”赵庆卫说,90年代时的神经网络

语音识别模型没能替代传统方法,此时语音技术还是以隐含马尔科夫模型和高斯混合模型为基本框架。

从20世纪90年代到21世纪初,非特定人、大词汇量、连续语音识别系统的研究成为国际语音界研究方向的主流。1997年,IBM首个听写产品Via Voice问世,用户只要对着话筒说出想要输入的文字,系统就会自动识别并输出文字。

2002年,中科院自动化所推出了“天语”中文语音系列产品——Pattek ASR;2005年,中科院声学所推出国内第一个自主研发的电信级语音识别平台,首次实现了国产语音识别软件的规模应用,在中国移动23个省的增值业务上线应用,占据了国内80%市场份额,使美国公司对中国语音识别市场的垄断成为历史。

### 通过纯语音信息实现与机器交互

与图像识别、机器学习一样,智能语音是人工智能的一个分支。在人工智能异常火热的当下,从Siri到小冰,从小冰到小娜,智能语音正在融入人们的生活之中。

所谓智能语音技术,就是研究人与计算机直接以自然语音的方式进行有效沟通的各种理论和方法,涉及语音识别、内容理解、对话问答等。一般来说,智能语音就是利用计算机对语音信息进行自动处理和识别的技术。

“从引擎模块的角度讲,智能语音技术包括语音前端处理(含语音增强)、语音识别、语音合成、语义理解对话管理和声纹识别等模块。其中,语音识别就是将语音信息通过计算机自动处理转化成文字的过程,也叫语音转写,它包括语音分段、端点检测、特征提取、解码以及后处理等过程。”中国科学院声学研究所(以下简称中科院声学所)研究员赵庆卫告诉记者。

目前,智能语音技术主要应用于智能家居、虚拟助手、可穿戴设备、智能车载、智能客服、智能医疗、陪伴机器人等方面。所谓虚拟助手,就是智能语音助手,它的核心在于人类通过纯语音信息实现与机器的交互,让智能机器“助手”帮忙完成指派的任务。

在赵庆卫看来,语音操作系统是一个比较大胆的想法,基于语音的人机交互有很大的发展潜力,所以不少互联网企业都看好这个方向。目前,亚马逊已经打造了一个智能语音平台(Alexa),平台上有各种智能语音应用(8万种技能),在这个平台上,用户可以通过语音发出一系列指令,比如购物、搜索、听音乐、讲故事等。

### 智能语音技术的前世今生

事实上,智能语音技术的研究起源于20世纪50年代。1952年,美国贝尔实验室制造了一台6英尺高的自动数字识别机“Audrey”,它可以识别数字0—9的发音,且准确度高达90%以上。并且它对熟人的精准度高,而对陌生人则偏低。1958年,中科院电子所的声学研究室利

用电子管实现了10个元音的识别。“由于那时计算能力很弱,智能语音只能做一些特别简单的字母或数字的识别。”赵庆卫说道。

20世纪60年代到70年代初,语音识别的研究取得了一定进展。“此时,智能语音技术开始形成系统的框架,提出了基于线性预测

### 深度神经网络框架成为主流

2010年,随着服务器的计算能力大幅提高(受益于GPU的应用)和训练语音数据的大幅度增加(受益于移动互联网和云计算的发展),微软基于深度神经网络的语音识别研究取得较大进展,“识别错误率相对下降20%以上”。此后,深度神经网络的建模优势被许多国际和国内知名语音研究机构所验证,业界开始认识到基于深度神经网络的建模框架比原来的框架识别效果明显要好,“现在大家基本都采用了基于深度神经网络的建模框架。”赵庆卫说道。

“我们利用AI图像识别技术实现生产防错与纠正,例如:轮胎机运链上线时,发现和用户选配需求不一致或者轮胎安装过程中没有按照设定的安装定位参数安装,就会向制造监控系统预警并驱动机械臂进行相应的替换或安装调整。”邵鑫说,此外在焊接过程中,也会结合大数据分析+机器学习来提升焊接质量。“通过不断的自我学习,智能化调整焊接参数,最终消除缺陷焊点。”

人研究,一些系统也已经上线,中科院声学所将其最新研究成果实际应用于中国移动通信集团和中国电信集团的客服热线,智能技术直接服务了数以亿计的客户。

据了解,中科院声学所长期致力于语音识别核心技术研究。针对实时语音识别的需求,研究人员提出一种基于混合神经网络(延时神经网络+输出投影门循环单元)的低延时声学建模技术,可处理长时信息,网络结构简洁,计算速度快,易于并行化训练。该模型结构已作为一种新型的反馈神经网络结构被国际主流语音识别开源软件Kaldi采纳。在非实时语音识别方面,提出基于BLSTM-E(双向长时记忆扩展)的深度学习神经网络结构,提升了现有主流BLSTM的性能,并解决了序列化训练条件下LSTM(长时记忆网络)对不同长度语音输入的鲁棒性差的问题。

## 从选车、造车到驾车 “AI+汽车”碰撞出这些新可能

### 第二看台

随着人工智能产业和实体经济的加速融合,落地场景的不断丰富,AI正在悄然改变着各行各业的发展模式。汽车行业拥有庞大的产业链,其中包括汽车的设计、生产、销售、服务等业务,在经济活动中占有重要地位。

面对日益激烈的竞争,汽车行业希望借助人工智能技术降低成本,增强竞争力。时值世界人工智能大会前夕,记者走进已经被人工智能所深刻改造的汽车行业,为你揭秘“AI+制造”将会为汽车行业产业上下游带来哪些改变。

#### 智能客服不断丰富汽车用户画像

在用户运营环节,过去一个客服只能对应几十个用户,有了智能助手后,运营人员不仅对应的用户数量翻倍,而且还能跨车型服务,实现了人机互相协作的局面。

“未来利用AI技术还可以进一步优化用户画像,更好地打通线上线下的用户数据。”上汽大通数字技术及转型中心高级产品经理邵

鑫说,用户什么时候去过哪家4S店,在哪个场景逗留时间比较长,和销售人员的对话记录,线上曾经和智能助手做过哪些沟通,都可以进行完整的数据打通,形成价值链闭环,使用AI技术可以不断地丰富用户画像,更精细地服务用户。

#### 线上线下无缝打通智能选车体验

在上汽大通自主研发、不断迭代的智能选配平台“蜘蛛智选”上,已经有多个车型、单车型多达60多项个性化选择供消费者定制。消费者可以在线上选择组合心仪车型的各类配置及功能,座椅颜色、前格栅、中控屏、主动安全功能等都可以定制,PC端、手机端、店面多触点覆盖,线上线下无缝打通智能选车体验。

“利用知识图谱这一技术,我们其实针对不同用户会进行非常个性化的推荐。”邵鑫说,比如说用户过去在运营平台留下过地域信息,“蜘蛛智选”就会据此标签进行一些相应的车辆配置推荐。例如,就轮胎而言,对南方用户会推四季胎,对北方用户推荐购买雪地胎作为备胎。

#### 图像识别技术实现生产防错与纠正

走进上汽大通的“灯塔工厂”,借助基于网络物理系统、物联网、云计算以及人工智能技术的综合性制造技术,例如行业领先的工程数据智能分析、数字化生产技术、数字化质量管理系统和数字化供应链等,驱动生产线的智能化转型,令生产更加透明、高效和智能。

“我们利用AI图像识别技术实现生产防错与纠正,例如:轮胎机运链上线时,发现和用户选配需求不一致或者轮胎安装过程中没有按照设定的安装定位参数安装,就会向制造监控系统预警并驱动机械臂进行相应的替换或安装调整。”邵鑫说,此外在焊接过程中,也会结合大数据分析+机器学习来提升焊接质量。“通过不断的自我学习,智能化调整焊接参数,最终消除缺陷焊点。”

麦肯锡全球董事合伙人、中国区汽车与先进制造业咨询业务负责人王平表示,服务全球和国内车企的经验表明,数字化结合物联网和自动化技术可大幅提升收入,缩短产品开发周期10%—20%,提高劳动生产效率20%—30%,减少库存30%。70%以上的后台工作可实现自动化。

从宏观面上看,数据显示,上海智能制造平均生产效率提升50%以上,最高提高3.8倍以上,运营成本平均降低30%左右,最高降低79.4%。目前,上海已成为国内最大的智能制造系统解决方案供应商输出地之一。

#### 人工智能助力无人驾驶快速进步

“未来中国很可能成为全球最大的无人驾驶市场。”麦肯锡未来出行研究中心报告称,由于本土及跨国公司的蓬勃发展,中国现已成为全球最大的车辆及出行服务市场,无人驾驶若能在中国落地生根,前景将十分广阔。

王平认为,目前无人驾驶技术还在快速进步之中,但离技术成熟到应用还有很长的路。预计2026年之后,无人驾驶技术在全生命周期成本上可能取得优势。

“人工智能对汽车行业的赋能,是一个逐渐深入的过程,现阶段应用的深度也不同。”上汽大通首席数字官吴钢说,汽车行业的智能化总体而言是两个大的方向:一是把汽车打造成一个智能产品,也就是无人驾驶等;另一个是实现汽车制造和销售等环节的智能化,提升效率。(据新华社)

### 产业界

## 云服务呼入或将解决中小企业AI客服发展痛点

本报记者 操秀英

1956年,人工智能的概念诞生。60多年后的今天,人工智能被频频提及,并赋能到各个行业,包括金融、医疗、交通、制造业,以及客服行业。

在人工智能发展迅猛的过程中,呼叫中心也受到了不小的影响,人工智能在赋能客服行业的过程中,角色已经从工具变成助手。随着国内首款轻量级客服电话机器人的诞生,人工智能或将给客服行业带来更深刻的变化。

众所周知,客户服务是企业运营业务过程中较为重要的工作,客户的去留在很大程度上直接取决于客户服务的质量以及体验。但在客服工作过程中,有许多简单的、重复性的劳动,为企业增加了不少人力成本。人工智能客服系统可以解决高成本的重复劳动。

“目前中国的呼叫中心业务大概有500亿的市场规模,从业人员达500多万。”天津深思维科技有限公司CEO张曦介绍道。

呼叫中心是企业必不可少,但又存在诸多痛点的尴尬存在。有一组数据显示,客服行业解决问题的重复率是83.2%,100条核心知识的场景覆盖率达90%,而行业离职率却是37%。也就是说,客服每天要解决的问题有八成是相同的,而且经过短期培训,快速掌握100条核心知识,便能够应对九成客户的咨询,但公司每培养出10位成熟的客服,就会有4位跳槽离职。

“人员成本越来越高且流动性大,服务及培训难以标准化,质量难以保证,信息过滤难度大。”张曦认为,这些痛点正是人工智能要着力解决的问题。

在传统机器人赋能客服行业的最初阶段,的确极大地解放了人力,加快了响应速度。但随着客服业务场景的不断拓展丰富,传统机器人客服系统便有些力不从心。传统机器人只能机械提供回复,提醒客户转人工服务,体验不尽如人意。

随后,随着人工智能及深度学习技术的发展,客服机器人也越来越聪明。一个优秀的智能客服机器人,要在长期的交互过程中不断学习和自我完善,达成对接收到的语句进行更精准的语义分析,能够通过上下文关联、场景管理、个性化推理等过程对自然语言进行准确理解,同时更需要积累庞大的知识库,特别是在相关专业知识方面进行长期学习。

近年来,我们看到,银行、运营商等行业的客服机器人越来越智能。“但是,这个行业的现状是,头部人工智能公司服务头部客户,大公司通常会耗费上千万建立一个智能客服体系,而中小企业则还没有享受到人工智能客服的便捷。”张曦说,云服务呼入客服电话机器人正好解决了中小企业人工智能客服发展中存在的痛点。该技术针对人工客服场景,通过AI+SaaS的形式,为企业提供轻量级自助式智能电话客服解决方案。

张曦说,面向中小企业,云服务呼入客服电话机器人考虑到移动办公场景下的用户需求,将内呼产品触达至个体经营者和小商户人群。用户只需扫码进入小程序进行简单设置,便可自助一键开通,无需布线 and 硬件投入。同时,使用费用相对低廉,大大降低了使用门槛。

业内专家认为,轻量级客服电话机器人或将引领智能语音客服产品标准化、SaaS化的新模式。

### 新鲜事

## 斯坦福AI模型 炼成小鼠“读心术”

读心术真的能够实现吗?近日,据新智元报道,Two Six Labs和斯坦福研究团队利用神经网络实现了对小鼠的“读心”,他们利用神经网络读取小鼠脑内的电信号,预测小鼠的行为和在迷宫中的位置,平均预测误差仅为4厘米。

大脑由相互连接的神经元组成;神经元可以响应输入信息并处于激活状态,反过来激活其他神经元。这些系统的“简化版”就是第一个人工神经网络的灵感来源。斯坦福Schnitzer实验室的研究人员制作了一个数据集,用于监控实验室的小鼠在“竞技场”中移动时的神经活动。

所谓“竞技场”其实是一个带有地标贴纸的小盒子。研究人员通过将一个微型显微镜连接到小鼠的头部,并记录荧光染料轨迹,这种染料会在单个神经元放电时发出绿光,从而实现记录神经活动的目的。这项技术可以同时跟踪数百个、甚至数千个神经元的活动。

研究人员还训练了一个神经网络,根据最近的神经元放电模式预测小鼠的位置,并使用实验观察结果的前80%作为训练数据,仅给出神经元的活动,来预测后20%观察结果的小鼠位置。在尝试了许多模型体系结构后,研究人员发现具有回归输出层的简单密集神经网络表现最好,平均预测误差仅为4厘米。

据介绍,目前实验室人员正在制作更复杂的行为数据集,以便更好地应用这些方法。比如可以在小鼠通过迷宫时对其进行映射,预测左右转弯,并量化小鼠在学习走迷宫时的不确定性。或识别小鼠展示哪些主题的图像会刺激到它。实验室研究人员表示,使用小鼠作为研究模型,目的是更多地了解我们自己,希望我们的人工神经网络有助于更好地理解生物的神经网络。