

编者按:对于中国2780万聋人来说,即将过去的2021年,是无障碍技术进入生活的“大年”。几家互联网“大厂”和媒体平台不约而同地在今年推出自己的AI(人工智能)手语数字人,希望为听障人士提供及时准确的手语内容播报,就连专业性较高的冬奥赛事,也会有AI手语主播登场,提供在线实时解说。这些会“打”手语的数字人,可以帮助聋人更好地理解健听人所言,完成信息“输入”的过程。

而在天津理工大学,有一个小小的“手语实时翻译系统”研究团队,他们人数不多,由听障大学生和健听人共同组成。他们各司其职,致力于教AI“认”手语,将聋人想“输出”的思想、想表达的情感,借助科技的力量让更多人“听见”“听懂”。他们想证明,在聋健融合的路上,听障者一直在和健听人“相向而行”。

本报记者雷琨、李亭、梁婷

今年5月,世界智能大会现场,一个“手语角”吸引了不少人驻足。只要在摄像头前打出手语,身后的显示屏上就会出现语意连贯的文字翻译。这套由天津理工大学聋人工学院和计算机学院共同参与研发的“复杂场景下中国手语实时翻译系统”(以下简称“手语实时翻译系统”),已经覆盖了教育、法律咨询、餐饮、交通几大应用场景,在光线充足稳定的室内,识别率最高能达到95%,部分场景下可以实现“秒翻”。

王建源是研发团队的成员之一。他是聋人家庭长大的聋儿,听力损失严重,吐字发声非常困难,手语是他和这个世界打交道最初也最有效的方式。一路长到22岁,聋人的身份不曾让他感到自卑;口不能言,也没耽误他长成一个妥妥的“学霸”;但如果你问他,因为听力障碍,从小到大遇到过哪些方面的不便,他会平静地在手机上打出四个字:方方面面——因为认识手语的人太少了。

一个最典型的例子是看病,如果没有手语翻译陪着,即便只是头疼脑热,像王建源这样的“手语族”也很难独自就医——绝大多数医生不会手语,你不能要求一个病人在急诊室里,用手写或打字的方式,条分缕析地快速描述自己的症状。

袁甜甜是王建源的老师,也是“手语实时翻译系统”项目团队的负责人。2006年研究生一毕业,她就在聋人工学院计算机系工作。15年来,袁甜甜记不清有多少次,凌晨一两点钟接到学院的电话,让她陪学生去看病,帮忙做手语翻译。如今,她已经是学院的副院长,带听障生上医院依然是她和学院所有老师的职责。

她不怕辛苦,就怕自己的手语水平耽误事。袁甜甜是健听人(听力健全人的简称),手语是走上教职之后自学的,她常常自嘲说,在语言方面天赋有限,“我是个‘半颤子’(天津方言,指半吊子、不熟练)手语翻译。学生描述的情况一复杂,或者打手语的速度太快,我就只能看懂一部分了。”

听障大学生自立自强、想方设法克服身体上的不便,特教老师尽职尽责、苦练手语为残疾人事业无私奉献——如果要讲一个励志故事,这样的情节就足够了。但袁甜甜和王建源们显然更想讲一个“科幻故事”:他们想教计算机学手语,让机器代替人,成为随叫随到、及时准确的“手语翻译”。

袁甜甜学计算机出身,王建源的专业是网络工程,他们笃信,依靠科技的力量可以帮助中国2780万聋人,乃至更多需要的人,克服生活上的种种不便。他们想让手语实时翻译系统化作一道桥梁,无论是听障者还是健听人,都可以通过它无障碍地交流。

这个目标听起来宏大,想实现也很难,但他们还是决定动手试试看。

新的桥梁

帮助聋人与健听人交流,前一座已经建成的“技术之桥”,是语音识别。

2007年,曾任中国聋人协会副主席的富志伟,发表了一篇题为《我看“聋人信息无障碍”》的学术文章。他在文中写道:“我期望将来能研制出一种机器,它可以和现在的微型摄像机一样大小,一端输入口语,另一端小屏幕上即显示文字……等到这种机器问世,聋人信息无障碍环境将大大改善。”

十几年后回看这段话,会发现富志伟期待中的机器,和如今智能手机上常见的语音识别应用,几乎一模一样。语音识别技术的受益者,也已不限于听障群体——看看身边有多少人在用语音输入法就知道了。

袁甜甜带领团队研发的“手语实时翻译系统”,就类似于手语版的“语

教AI认手语的年轻人

他们借助科技力量,试图让更多聋人被“听见”“听懂”

音输入法”。只不过后者是输入语音、机器识别、转成文字,前者是输入手语动作和表情、机器识别,再翻译成文字——“老桥”与“新桥”,相似又不同。

袁甜甜没想到,“老桥”的存在竟然会成为项目团队建设“新桥”的“阻力”。

“为什么一定要搞手语翻译呢?聋人听不到,现在有语音识别啊!有的聋人不会说话,拿手机打字不就行了?”曾经有一些企业向袁甜甜了解过正在开发中的手语实时翻译系统,想谈投资。但谈来谈去,他们却对项目的必要性产生了质疑,觉得手语翻译除了“做善事”没有其他意义。

类似的话听得太多,快人快语的袁甜甜有点儿急了,“那要这么说,当初搞什么微信呢?就用短信不一样吗?搞什么语音识别呢,就直接打字呗?很多创新都有传统形式上的替代品,那科技就别进步了呗?”

袁甜甜说,直到今天,手语仍然是许多听障人士最自然、高效的表达方式,“比写字、打字都快。”但为了促进无障碍交流,要求健听人都去学手语,显然不现实。人类要掌握一门语言,总会受到认知水平、记忆力等各方面的限制。“我们的手语实时翻译系统,就是要解决这个问题。”袁甜甜有着典型的工科思维,遇到难题,第一反应就是思考能不能从技术层面突破。“我感觉如果技术到了那儿,机器的记忆力和它的学习能力,要比人强很多。所以要是能让机器学会手语,它肯定比我这种‘半颤子’手语翻译管用多了。”

目前,手语实时翻译系统还处于试用阶段。理想状态下,等到研究成果真正落地,健听人只要打开手语翻译程序,就能和“手语族”直接交流了。

但对袁甜甜和年轻的研发团队来



天津理工大学“手语实时翻译系统”研究团队主要成员合影。受访者供图

说,想达到这个“理想状态”不容易。

作为深度参与项目的听障生,王建源和聋人工学院计算机专业的乌力杰在手语识别、翻译技术上投入的精力越来越多。他俩开过一个向全社会推广、普及手语的微信公众号,如今已经一年半没顾上更新了。当初办公众号的时候,他们的想法是,要打开一扇窗,让聋人被“看见”。而现在,他们忙着造桥,想让听障者被“听懂”。

手语之美

除了听障生,“手语实时翻译系统”项目团队中也有健听人。无论是不是手语的日常使用者,他们之间有个共识:手语是一门很美的语言。

袁甜甜对手语之美的直观感受,都来自她的学生。袁甜甜的母校是天津师范大学,当老师是她唯一的职业目标,但她没有特教专业背景,平时打的手语,一部分是自己照着书学的,另一部分是在和听障生的日常交流中逐渐掌握的。刚入职的时候,学院的老教师还夸她,“挺厉害啊,一上讲台就敢跟学生‘比画’!”

袁甜甜有着典型的老天津人性格,心肠热、说话直,她坦言当初苦练手语,心里也没有存着“为残疾人教育事业贡献力量”这样的远大理想。只是觉得,既然要当老师,教的又是听障生,手语就是上课必备的工具,“你不敢‘比画’就永远学不会,永远没法跟孩子们无障碍沟通,对吧?”

就这样一边用一边学,她发现聋人日常使用的自然手语,远不止“照着手比画”这么简单。它有自己的语序和语法,口语里说“灭火”,灭在火之前,而学生们打手语时,会先比出“火”——起了“火”,再去“灭”;它有独特的空间感,同样一个掌心向下,五指由聚到散,模拟光源的手势,在室内打意思是灯,在室外打就可以指代太阳;同样一句话,来自五湖四海的学生,可能打出好几种不同的“方言”版本;手语的表达方式更是丰富,要打出一个意思完整的句子,不光靠手势,也要配合相应的表情,还需要加上肢体动作……

“真的是一门很美的语言!”袁甜甜说,直到今天,她和学生用手语交谈,有时候还会愣神,觉得他们“身上有光”。而亲眼看过王建源、乌力杰打手语的人,就会承认袁甜甜的感受并不夸张。

王建源生在青岛一个聋人家庭,父母都是听障人士。2018年考进聋人工学院网络工程专业之前,他一路都在聋校接受教育。

小时候,父亲就像其他家长教孩子说话一样,带着王建源认认真真地练习手语。父亲并不认为那是一门“特殊”的语言,“他觉得普通话能字正腔圆、抑扬顿挫,手语同样能够大开大合、诙谐幽默。”父母对手语的坦然态度,影响了王建源。他从不避讳在公开场合使用手语,也从不隐藏自己的聋人身份。在他看来,手语就是手语,不是某种“残疾的外显”。聋人可以用它表达内心的想法,而健听人,只要掌握了这门语言,同样可以用它来交流——手语和其他任何语言一样,是破除障碍、沟通彼此的工具,它可以是桥梁、是纽带,唯独不该是障碍本身。

乌力杰是来自青海的蒙古族小伙,“家在茶卡盐湖边上”——除了精通手语,他也可以用口语交流,只是声音有些沙哑低沉。

比起同年入学的王建源,乌力杰

的求学经历要更复杂些。小学阶段,他靠助听器和读唇的本领,与健听孩子一道读过三年普校,“三年都是全班第一”。那段经历锻炼了他的适应性和口语表达能力。后来,他离开青海,在武汉第二聋校完成高中学业,通过单考单招来到天津理工大学。这样的成绩在家族同辈的孩子中“笑傲群雄”,更让他坚信“聋人也不比谁差”。他一度直接把“Deaf-无音”用作自己的微信昵称——在英文中,Deaf 就有聋人的意思。

无论是学专业课、推广手语还是做科研,这个古铜色皮肤的蒙古族青年有股“一马当先”的闯劲儿,“海伦·凯勒能做到的事情,我们为什么做不到?”

技术之难

王建源和乌力杰是袁甜甜邀请加入项目团队的。2019年,他俩才上大二,就被委以重任,负责收集手语语料,撰写符合自然手语语法、语序的句子。

手语是一门视觉性语言,基于这一特点,“撰写”的过程不是通过手写或打字输入,而是以录制视频的方式完成的——王建源、乌力杰以及团队中其他听障生的重要工作之一,就是对着摄像头反复打手语。

为什么是他们?因为手语识别和翻译技术领域,各国学界都走过弯路:比如尝试通过数据手套建立手语语料库。作为一种常见的传感器,数据手套在手势识别技术上的应用已经相对成熟——尽管成本高昂,但用手套来收集“手势”,看起来理所当然,准确率也该有保证。只可惜手语不是简单的手势:抛开表情和大动作不谈,即便是同样的手势,指向稍有不同,也可能意思迥异,这样微妙的差异,手套是识别不出的。

意识到这一点,近年来各国推进的手语识别、翻译项目,大多采用计算机视觉的方法,利用摄像头收集数据、建立语料库,袁甜甜和她的团队也是如此。现阶段更常见的问题,一是收集的样本太少,数据集不够大。二是数据集起来了,但没能对语料进行高质量筛选和标注。说白了,就是语料“不好使”。

袁甜甜说,人工智能深度学习的过程,有点像教小孩学说话,一个单词先得全家上阵,翻来覆去地重复,次数够了,孩子才能建立印象,“哦,这个单词叫妈妈,不管是从爸爸嘴里说出来,姥姥嘴里说出来,还是舅舅、阿姨说出来,都叫妈妈。”同样的,要让机器识别一个手语句子,也要有足够的面对摄像头,以不同的风格重复打同一套动作,计算机才能“记住”。

很多国外团队建立的手语数据集无法支持复杂场景下的手语翻译,一个直接的原因,就是无法找到足够多的人在自然状态下采集手语。“而我们的团队背靠聋人工学院,很多成员本身就是手语使用者,在自然手语的采集上有优势。”袁甜甜说。

“但是跟您说实话,我们也建过‘不好使’的数据集。”袁甜甜直接给团队2018年建立的手势汉语数据集下了定义,“花了钱、费了功夫,最后挺失败的”——失败之处在于,那一次他们收集的语料是手势汉语而不是自然手语。手势汉语是按照健听人习惯的汉语语序,而不是手语语序连词成句的。

“比如说,‘爱是我们共同的语言’。这个句子用手势汉语来打的话,就是按顺序一个字一个字地比画,‘的’也有对应的手势,也要打出来。但聋人日常使用手语的时候,惯用的语序是:爱、我们、共同、语言、是,‘的’不用打。我这个‘半颤子’手语,讲课的时候也不会把‘的’‘了’都打出来,那不符合自然手语的表达习惯。”袁甜甜解释。

“你把手势汉语翻译得再精准也没用,聋人平时打的不是这个啊!”汲取了教训,这一次再建手语数据集,袁甜甜说,他们不求速成,只求每条语料都是原汁原味的自然手语。先按手语语序识别、再按口语习惯翻译,虽然多了一个技术环节,团队研发的难度也就长了一截,袁甜甜们却坚持要选择那条“难走但正确”的路径。

“咱做这个研究,不想跟人吹我们的数据集有多大,也不想说我们发了多么高大上的论文,我们就一个目的:

能用。”袁甜甜斩钉截铁,“这个系统出来了,聋人必须真的能用。”

收获之年

对于袁甜甜和她的团队来说,2019年和2021年是两个关键的时间节点。2019年,是他们的初创之年,“手语实时翻译系统”入选国家工信部新一代人工智能产业创新揭榜项目,获得了2000万的资金支持。袁甜甜很高兴,一方面是因为国家对无障碍建设的重视让她欣慰,另一方面,数据采集、技术研发、成果落地也确实是处处都要花钱。

2021年,则可以说是团队的收获之年。

闫思伊到现在还记得,今年年初她第一次跑通手语翻译系统代码框架时的心情。“就好像你在组装一台机器,零件都已经拼全了,但是因为各种各样的小毛病,比如说有几个螺丝没拧紧,它就是不能正常运转。我每天在实验室,一个一个地把螺丝给上紧。都调完之后,忽然有一天,我一按开关,机器就轰轰地转起来了。那种成就感啊……”

闫思伊在天津理工大学计算机学院读研二,是手语实时翻译项目团队的健听人成员,主要负责“后台”工作——构建手语识别算法框架模型。如果沿用教孩子学说话的比喻,学前阶段完成了反复加深印象的步骤,一上小学,孩子就要开始学习拼音、语法,学会按照教材上总结的规律,把学过、听过的那些单词组成句子和课文,写在作业本和卷子上——闫思伊就是编“教材”,帮AI这位“同学”总结手语规律的“老师”之一。

“这部分技术要求更高的工作,主要由我们学校计算机学院的师生来完成。”袁甜甜说,给人工智能“编教材”的过程很艰辛。在手语识别、翻译领域,可借鉴的经验太少,同行的进度也都差不多。项目团队参考了德国天气预报自然手语翻译系统的算法,也通过各种渠道寻找类似系统在网络上公开发布的源代码,然后就是一遍遍地尝试和修正。

闫思伊他们每天早上8点半就扎进机房,等意识到天黑该回宿舍,再看表已经是晚上10点了。“我觉得我们的项目有意义啊,就想快点出成果,能落地的成果。”她说自己是个典型的理工女,就喜欢计算机,导师让周一到周五做项目,她周六日也忍不住去跑代码。就这样一点点摸索,一点点“紧螺丝”,第一次跑通代码的那天早上,计算机“同学”在它的人类老师们面前,写了一份不算完美,但成绩合格的手语翻译试卷。

今年5月,王建源、乌力杰带着研究成果上了世界智能大会;10月,他俩又和聋人工学院其他几位同学组成创业团队,代表天津理工大学,拿下了第七届中国国际“互联网+”大学生创新创业大赛总决赛高教主赛道的金奖。

他们的“夺冠”项目叫“鲸可语”,全称“鲸可语多模态连续手语自动标注识别系统”。自动标注识别就是手语翻译的前期步骤,而“鲸可语”正是依托“手语实时翻译系统”而生的——他们把两年来“教机器学手语”的经验转化成自己的科创项目,亲手为他们想象中那座“聋人融合之桥”建起了“桥基”。

决赛夺金的那天,袁甜甜在朋友圈给她的学生们一个点赞。作为“鲸可语”的指导老师,她很喜欢这个温柔又有力量的名称。

“鲸可语”来源于“52赫兹鲸”的故事,在海洋中,这只因为叫声频率与众不同而无法与同类交流的鲸鱼,被称为“最孤独的存在”。但事实上,如果你能破解52赫兹的密码,就会发现它并非一座孤岛,它也会歌唱,也有自己的语言。王建源觉得这只鲸鱼很像听障群体,“沉寂在人海中,无时无刻不在渴望着与外界沟通,渴求着回应。”

几个听障年轻人给“鲸可语”设计的标志,像一只蓝色海洋中漂浮的白色鲸鱼,也像人们把拇指和食指搭在一起“比心”时的手形。“我们建手语语料库的时候,就是要把手语的动作手型概括为像这样简洁的线条,正是一个个线条,支撑起了‘鲸可语’系统,实现了对手语的识别。”王建源说。

而他们开发手语识别系统的初心,就是帮助那只“52赫兹鲸”,获得言语的能力和“被听懂”的权利……

(参与采写:吴泽运)