

人工智能也可输出视频“观后感”

■本报见习记者 赵利利

“对视频的智能理解是人工智能的最大战场。”说出这一观点的上海交通大学教授倪冰冰坦言这一说法“不会有反对意见”。作为深耕计算机视觉中智能视频分析的专家,他甚至愿意称之为人工智能的“体量之最”。

这一称呼不无道理。在近日召开的全球科技青年论坛上,倪冰冰将其中的“简单道理”娓娓道来,“人类每天接触到的‘听说读写’信息里有70%实际上是视觉信息。而在信息领域,大多数的信息也是视频图片数据。”

“体量之最”当然还包括视频智能识别的“计算量之巨”。

“全世界每秒钟传到 Youtube 的视频要达到几千个小时。”倪冰冰表示,即使使用现在流行的神经网络、深度学习技术,加上高性能硬件设备,计算时间之久而超乎想象。

“视频一直被视为计算机视觉里的暗物质。”谷歌云人工智能和机器学习前首席科学家李飞飞曾如是说。即便如此,科学家及互联网公司视频智能理解技术的“征服欲”也没有减退之意,纷纷投入巨大的财力物力人力去攻坚克难,IBM、谷歌都曾为视频智能理解的深入研究发布视频数据集,脸书 CEO 扎克伯格在最新一季财报分析师沟通会上透露,视频识别已成为旗下视频在线产品、VR 产品的核心竞争力之一。

因何具有如此动力?倪冰冰一语道破其中缘由:智能视频理解技术在安防监控、自动驾驶,以及在社交媒体短视频、直播里,都是最核心的技术。

而随着视频在人们生活中所占的分量越来越重,视频智能理解技术有了越来越深厚的现实基础。经过过去数年的努力,视频智能理解研究领域已经有了一些突破性进展,变得更具有实际应用价值。

增加时间尺度建模

打开一些视频网站,用户经常可以看到与自身历史浏览痕迹相关的视频推荐,或者“以视频搜索视频”的内容搜索功能,这其中首先涉及到的技术其实就是视频的行为识别。通俗而言,就是给正在运行的程序一段视频,让它告诉你“里面发生了什么”。

识别智能理解技术的研发过程并不像描述起来这么简单,其面临的技术挑战巨大。除了视频体量,视频的质量、视频内容中有效信息的数量也都是很大的技术难题。



对视频的智能理解是人工智能的最大战场。

图片来源:百度图片

以安防领域的视频数据为例,倪冰冰解释,摄像头的视频质量有时候比较差,人动一下会有些模糊,使得技术对所捕捉到的动作的理解变得非常困难;另外,不同的人做同样动作的样式非常不同,“有人快有人慢”,这也造成智能理解的难点;不同位置的物体在视频中的尺寸相异,要把“大的东西、小的东西”同时“抓”出来是很有难度的。

倪冰冰所在的团队提出了时序金字塔算法,该算法通过对时间轴进行分尺度建模,使不同尺度的数据特征都能得到很好的表达。

“我们又配合流行的深度学习里的 LSTM(长短期记忆网络)、深度递归网络,使得不同尺度的信息在不同时间进行流动,最后达到很好的识别效果。”倪冰冰进一步阐释道。

这一研究在由谷歌、斯坦福等国际顶尖人工智能研究机构主办的 THUMOS 行为检测国际竞赛中,获得视频检测小组国际第一名。

“找人”不再是难题

“找人”是理解视频中行为的关键所在。在计算机视觉领域中,“找人”行为的专业术语叫做“行人重识别”,专指利用计算机视觉

技术判断图像或者视频序列中是否存在特定行人的技术。

而在人工智能产业化的过程中,“找人”也恰恰是视频智能理解可以快速投入的应用场景的形象描述。倪冰冰讲述道,“如果我们能把人和人在不同摄像头下的影像追踪起来的话,就可以得到这个人在整个时间轴和空间轴上的分布,这可以实现很多的应用,比如找到一些非常敏感的人物。”

过去,倪冰冰团队在进行视频行为里的行人重识别时,通常利用单帧信息,比如在某个时间点上选一幅图像与旁边的其他帧图像进行对比,但这样的做法存在一定的问题。

“我们发现单帧图像的信息是丢失的,摄像头所拍摄的是行人整个‘动’的过程,取得单帧的话,整个‘动’的行为就损失掉了。”倪冰冰说。

另外,利用单帧图片信息很有可能会遇到一些特殊姿势、遮挡、运动的模糊等问题。对此,倪冰冰团队提出的一个方法使得识别技术能够对整个视频进行时序上的个人信息加以汇总,形成定量表征,从而使行人重识别的性能达到最优。

基于多角度摄像机的 3D 定位跟踪模块是当前业界比较新的技术。倪冰冰表示,过去,单个相机对人在 2D 场景进行定位,现在,技术可以在相机和相机之间形成匹配,用

重识别以及跟踪技术可以在 3D 空间里对入进行追踪,对人的动态过程进行画像。

“这些技术产生了很多应用,新零售、智能零售系统就是很好的例子。”倪冰冰自豪地说道,“新零售里,无非是有多人进入我的商店、多少男生多少女生、什么年龄段、他们在哪个商品面前停留了多少时间,这些通过技术都可以解决。”

倪冰冰表示,抓到这些数据之后,可以反映在大数据端,有了客流分析、轨迹、略读图的数据,店家就能够对行销展开布局,对行销策略做一定的改变和优化。

在小型终端运行

“用深度学习网络进行视频智能分析所需的网络是很大的,它的参数量非常多,甚至超过几十、几百兆。”倪冰冰表示,以这样大的网络去做视频端,显然使用时间是非常久的,但新一代人工智能希望很多智能的工作在端上进行,比如手机、平板电脑等,而如此大的模型显然无法在一个小的运算设备上运行。

针对这个问题,倪冰冰团队对网络进行了二值化量度的简化,使得过去用浮点数据表达的网络,用现在的二进制数据网络来表达,这使得整合各个模型的过程变得简单,计算变得更加高效。

“有了这个技术之后,实际上,很多原来一定要在‘云’上、在服务端去做的一些人脸识别功能,现在可以在很小的手机上进行。”倪冰冰表示,“智能刷脸闸机、智能访客机以及智能门禁等有了这个网络压缩技术,都能在小型设备上运行。”

倪冰冰认为,对一个公司团队来说,从算法到技术落地还有很长的路要走,“这里面相当多的功夫用在了工程化的落地上”。对此,其团队提出的人、物检测的算法也做了工程化的努力。

“现在对人的定位有轻量级的人脸检测模块,这样,一个几兆的小模块可以在手机上运行;也有轻量级的人体骨骼检测模块,实际就是抓住人的一个动作,过去它是一个体量非常大的计算,现在,通过一些新算法和工程实践可以把它做到很小的手机模块上,而且能够达到实时运行的效果。”倪冰冰表示。

“定位了人以后要进行识别,现在在小设备上,人脸识别也能做到在百万底座下 95% 以上的通过率,这是非常难得的一个过程。”倪冰冰说。

腾讯宣布致力于打造「救命的 AI」

本报见习记者 赵广立

11月4日,2018 腾讯 WE 大会在北京举行。腾讯首席探索官刘大为介绍了腾讯在“人工智能+医疗”领域的进展——打造“救命的 AI”。他表示:“我们在很多领域开发人工智能,但是我们的激情在于,能够用人工智能推进医疗的发展。我们非常关心人的健康,我们有很大的激情希望能够优化大家的生活方式。”

网大为表示,在健康领域、医学领域使用人工智能,如果做得好的话,能清晰掌握心脏、脑、眼睛等重要器官的状态,综合分析数据,从而提升诊断和治疗的准确性,更早发现癌症和肿瘤,也能向病人提供更适合的治疗方案。

腾讯 AI 技术医疗领域已有多款产品落地,发布于 2017 年 8 月的“腾讯觅影”是腾讯首款 AI 医学影像产品。腾讯相关负责人介绍,目前腾讯觅影已构建起两项核心能力:利用 AI 医学影像分析,实现食管癌、糖尿病视网膜病变、乳腺癌等疾病的早期筛查;利用 AI 辅助引擎,对 700 多种疾病风险进行预测诊断。截至 2018 年 7 月,腾讯觅影已累计辅助医生阅读医学影像超 1 亿张,服务超百万患者,提示高风险病变 15 万例。

“大家想象一下农村地区,或一些医疗资源比较匮乏的国家。在那里即便不具备专业医疗知识,你还是可以用成像技术,获得相当数量的数据去支撑诊断手段的运用。”网大为在演讲中说,“通过尽可能低的成本获得尽可能多的洞见,形成尽可能大的成效。这个领域让我们倍感兴奋。”

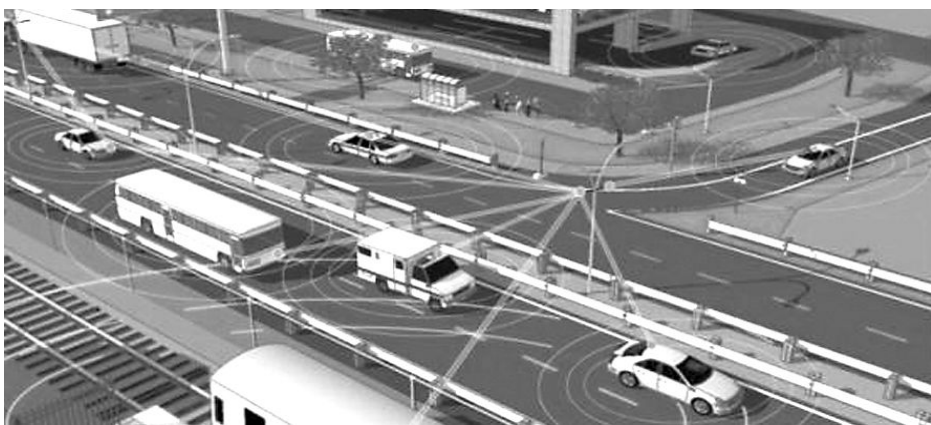
当天,网大为还宣布其 AI+ 医疗领域研究已从影像筛查进入病理分析阶段。他介绍说,融入 AI 及 AR(增强现实)技能的智能显微镜,可根据医生的语音指令,自动识别、检测、定量计算和生成报告,并将检测结果实时显示到目镜中,提高诊断效率和准确度。

目前,“脑卒中”(即“中风”)成为我国居民死亡首要原因,具有发病率高、死亡率高和致残率高的特点。腾讯与天坛医院展开合作,借助人工智能技术,开发出能早期快速判断脑出血病因的辅助算法与系统,以降低医生的漏诊和误诊率。与此同时,腾讯还致力于推动这套系统未来在基层医院的使用,推动普惠医疗的发展。

结合人工智能和脑科学的最新研究成果,腾讯还推出了帕金森病运动功能智能评估系统。据介绍,该系统可在患者不穿戴传感器的情况下,仅通过普通智能手机摄像头拍摄,就能完成帕金森病的运动功能日常评估,全程只需 3 分钟,诊断速度相比传统评估方式提升 10 倍。

铺平智能网联汽车之“道”

■本报见习记者 赵利利



很多海归人才,但是还不够。”

需要更多的政策包容

即使是基础设施建设的优势,也不可以“坐享其成”。百度公司董事长兼 CEO 李彦宏表示,“当我们的汽车越来越智能,当我们在憧憬以后可以取代人类驾驶员的时候,我们必须要考虑,道路的基础设施也必须进行改造。”

以自动驾驶为例,汽车需要依赖昂贵的激光雷达辅助。李彦宏表示,激光雷达如果每辆车都配一个,那么,单车成本就会增加 10 万元,但如果把激光雷达配在道路上,很多车就可以共享激光雷达,成本大大降低。

因此,创造有利于创新发展的环境是汽车产业升级的重要条件,政府在其中扮演着重要角色。

“国有企业现在管得很严,虽然(一定程度上)允许失误,但是新技术发展及创新一定会有更多的失误,需要更大的包容。”董扬举例道,“《道路交通安全法》应该修改,不修改我们(无人驾驶或其他智能汽车)上不了路;《测绘法》需要修改,不修改我们做不出很好的高精度三维地图。”

李彦宏对此表示认同。他说,“一定要有相应的政策跟上,智能驾驶才能真正在开放道路上跑起来。”

董扬表示,这就要求政府做好顶层设计和规划;要促进产业间的合作;还要支持产业创新基金的发展。

安全问题不能忽视

董扬认为,虽然智能网联汽车和电动汽车都是新技术,但两项技术并不一样。他

表示,电动汽车关系到环境、关系到能源,可以由政府主导来推动,但是智能网联汽车主要涉及到安全效率,应该由市场来推动。

“过去我们 30 年发展主要是引进技术追赶型的发展,在追赶型的发展中互相合作可以少一些,但未来是创新发展,企业家必须加强合作。”马化腾说。

马化腾表示,“今天,闭门造车已经不可能,没有一个企业能够完全拥有智能网联汽车所需要的全部资源和能力,开门造车正在成为现实。大家需要通过跨产业协作共建新生态。”

马化腾认为,未来汽车产业的发展需要传统汽车企业和互联网企业共同沉下心来打磨三个基本功:网络连接能力、数据处理能力和安全。

“对于汽车企业来说,在这次产业变革中,如果要摆脱沦为一个生产车间的命运,就要尽快实现生产和消费的连接,打通从造车到用车的价值链,真正触达终端用户,而不仅仅满足于一个制造产业链。”马化腾表示,“更重要的是,如何借助硬件的优势把软件和服务三位一体连接起来,做成一个生态。”

“汽车会成为一个数据密集型产业,每辆车每天产生大量的运行和交互数据。”马化腾认为在此背景下,服务中心怎么作出实时响应、研发部门如何洞察海量数据并获得启示都是重要的问题。

此外,智能化提升生活品质的同时也带来新的风险和挑战。马化腾表示,“汽车成为万物互联的重要节点,系统安全、信息安全就相当于所有的 0 前面的 1,如果没有 1 什么都没有。而且和智能手机终端不一样,智能网联汽车的安全问题事关生死,更加不能忽视。”

简讯

第二届“面向未来的车联网”研讨会在沪举行

本报讯 日前,第二届“面向未来的车联网”研讨会在沪举行。本次研讨会由上海先进通信与数据科学研究院等主办、上海交通大学协办。与会专家,在 LTE-V 商用以及 5G V2X 标准化的关键时间节点上深入交流了车联网领域的政策措施、发展经验和最新研究成果,探讨了车联网产业未来发展之路的重大意义。

与会专家认为,实现全自动驾驶需要网联和智能协同发展,且网联与智能协同发展的路线与中国国情贴近,能够充分调动社会各界的鲜活力量。“智慧的车”“聪明的路”以及“无处不在的高速无线网络”将最终成就高度可靠与安全的全自动驾驶,给人们的生活带来“质”的提升。

“路线清晰,上下齐心,产学研用协力,中国必将走出一条有中国特色的智能网联汽车之路,成为智能网联汽车工业强国。”中国通信标准化协会常务副秘书长兼副理事长长晓慧详细解读了 2018 年国家最新发布的《车联网标准体系指南》,并指出该指南由国家多个部委联合发布,标志着我国车联网标准化工程进入了新的阶段。(黄辛)

机器人与智能制造高峰会议暨企业家论坛在天津举办

本报讯 以“拥抱触手可及的机器人时代”为主旨,由河北省自然科学基金委主办、河北工业大学承办的“燕赵科学论坛之机器人与智能制造高峰会议暨企业家论坛”11月4日在河北工业大学举办。

“中国工业机器人高速增长,高端工业机器人严重依赖进口的局面没有明显改观;中国机器人在汽车零部件等领域有巨大的发展空间;‘软能力’成为中国工业机器人的‘痛点’,也是中国机器人向高端发展的主要抓手之一。”南开大学教授、中国科学院沈阳自动化研究所客座研究员韩建达在论坛上作《工业机器人现状、机遇、挑战》演讲时说,复杂环境下工业机器人面临着三难:互联互通难、动态重构难、协同操作优化难。

“推动机器人技术与产业的发展是一项系统工程,要从基础前沿技术、共性关键技术、核心部件、核心软件、核心器件、应用工艺及系统解决方案等多个方面全面突破,协同发展。”哈尔滨工业大学教授、机器人研究所所长赵杰表示,中国机器人市场秩序有待进一步规范。(高长安)

发力 AI+ 媒体 第四范式发布“先荐”智能推荐系统

本报讯 11月6日,人工智能(AI)明星创企“第四范式”在其举办的“2018 人工智能+新媒体论坛”上,发布了其进军“AI+ 媒体”的产品——先荐智能推荐系统(下简称“先荐”)。第四范式先荐业务团队负责人周开拓介绍说,先荐的定位是“媒体智能转型的专属工具”,理念是利用人工智能算法和技术帮助媒体快速、精准匹配用户的阅读需求。目前,先荐已与新华社、CSDN 等三百余家媒体开展合作。

目前,先荐可提供个性化推荐、关联推荐、热门推荐三种服务。“先荐是集内容上传、内容管理、内容分发、推荐干预、前端渲染于一体的一站式推荐服务可视化平台,用户可通过简单的界面操作,仅需三步,即可完成自建推荐系统的全过程。”周开拓介绍称,先荐支持 PC 端、WAP 端、APP 端全平台接入,并提供场景管理、内容管理、规则配置、上线渲染、报表统计等推荐配置和数据统计功能。用户可随时了解推荐状态,精确控制推荐服务。

“AI 提升了用户获取信息的效率,让信息获取进入个性化时代。第四范式推出‘先荐’,就是希望通过机器学习技术,实现内容与用户间的精准匹配,使内容价值最大化。”第四范式创始人、首席执行官戴文渊在上述论坛上表示,AI 不仅是媒体连接用户的桥梁,它还将成为媒体提升内容和运营效率的重要生产力。(赵广立)

北科院设立城市生态环境研究中心

本报讯 近日,北京市科学技术研究院(以下简称“北科院”)在京举行城市生态环境研究中心成立大会。

北科院院长郭广生介绍说,城市生态环境研究中心整合了北科院属轻工业环境保护研究所、北京市劳动保护科学研究所、北京市理化分析测试中心、北京市辐射中心、北京麋鹿生态实验中心等科研资源,未来将以城市为研究对象开展生态环境领域科学研究、关键共性技术研发、成果示范应用和转化,“主动融入生态文明建设的‘主战场’”。

北科院相关负责人告诉《中国科学报》记者,城市生态环境研究中心实行主任负责制,第一届主任由北科院副院长谢威研究员担任。同时,该中心还决议聘请中科院院士王桂斌为中心第一届学术委员会主任委员,聘请中国科学院院士彭永臻、中国科学院院士张锁江为副主任委员,中心首席科学家张忠国研究员为常务副主任委员;清华大学、北京师范大学、北京林业大学、华北电力大学、中国国际工程咨询公司、中国环境保护产业协会、北京市环境保护研究院、北京市水利规划设计研究院的 10 多位国内外专家受聘为学术委员会委员。(赵鲁)

金山云携手鞍钢共建“精钢云”

本报讯 近日,辽宁鞍山钢铁集团(以下简称“鞍钢”)与中国云计算企业金山云共建的工业云平台——“精钢云”正式落地。双方将在这一平台运用“云+大数据+人工智能”技术,助力企业降低生产成本,大幅度提高生产率。

鞍钢集团自动化有限公司董事长刘凯、金山云合伙人宋伟代表双方签署合作协议。刘凯表示,鞍钢将利用“精钢云”作为科技转型的载体,将自身工业制造和数字化能力转化为服务能力,构建工业智能制造的行业云平台。同时,依托信息基础设施和平台,将服务能力迁移到云端构建服务平台,输出可运营的服务。(赵鲁)

英特尔傲腾数据中心级持久内存测试版程序发布

本报讯 11月5日,英特尔于北京发布了英特尔傲腾数据中心级持久内存测试版程序,该程序使各大原始设备制造商和云服务提供商能够让客户抢先体验英特尔革命性的内存技术,并为英特尔傲腾数据中心级持久内存存在 2019 年上半年的全面上市奠定基础。目前,全球众多原始设备制造商和云服务提供商已公布了各自用于早期客户试用和部署的测试服务和系统。

据介绍,英特尔傲腾数据中心级持久内存通过两种特殊的运行模式来实现独特的能力。利用 App Direct 模式,经过专门调试的应用程序可以从产品固有的持久性中充分获取价值并获得更大的容量。而在内存模式下,在受支持的操作或虚拟环境中运行的应用程序可将该产品用作易失性存储,从而在无需重写软件的情况下有效利用最高达 512 GB 的内存空间所带来的附加系统容量。(贡晓丽)