

# 高通量AI一体机架起数据高速路

## 好机友

本报记者 姜婧

目前我国人工智能企业总量位居全球第二,但支撑AI计算的基础架构几乎全靠进口昂贵的GPU来支撑。尽管大量的AI算法已被开发,但鉴于AI的高技术门槛,研发用于支撑AI计算的基础架构十分匮乏。

近日,中科睿芯联手高通正式发布全球首款基于ARM的高通量人工智能一体机。打破了x86架构在服务器领域的霸主地位,填补了ARM架构服务器的空白,并省去用户搭建服务器环境等诸多复杂步骤。

如何降低人工智能领域的门槛,让更多“外行人”有机会开发出自己的人工智能服务,是AI真正发挥商业价值的关键。“睿芯正是瞄准了AI

的应用门槛,通过软硬一体的方式,搭起开箱即用的基础平台,用户只需要在此基础上做个性化应用开发。”中科睿芯董事长、中科院计算所高通量中心主任范东睿博士说。

范东睿介绍,睿芯高通量人工智能一体机通过软硬件结合深度调优,充分发挥Cen-triq处理器高通量和低功耗的优势,实现大数据和人工智能快速、高效、稳定地交付。可提供从AI计算平台、管理调度与分析平台到深度学习框架的端到端解决方案,并在语音识别、图像识别、无人驾驶等AI领域实现规模化应用。

据互联网数据中心预计,未来全球数据总量年增长率将维持在50%左右,到2020年全球数据总量将达到40ZB。庞大体量的数据,对算力也有严苛的要求。“人工智能进入快速发展产生新模式,并对算力提出新需求。当前

监督式学习的神经网络计算量是可控的,但自动学习模型,现有的计算机能力还达不到。”中科睿芯CTO、北京智能计算产业研究院院长张浩博士解释道。

故而就需要设计AI服务器,使得模型设计与数据训练不再受限于计算能力的限制。“高通量计算提供的正是触发人工智能爆发的强大算力,为人工智能的实现架起了真正的高速公路,使得数据信息海量通过,并能及时消化、避免堵车。就像跑步,高通量是追求在一定时间内跑到终点的人越多越好,而不是跑出冠军。”张浩解释道。

功耗问题已成为数据中心的核心理念,据统计,谷歌的数据中心能耗占其总能耗的80%以上,其中服务器能耗占40%,另外40%是为消除服务器产生热量的空调系统能耗。而低功耗正是ARM架构最核心的优势。目

前国内外不少企业,比如华为海思、飞腾、Cavium、高通都相继推出了ARM高端服务器芯片。

“事实上,ARM架构天然的结构简单、低功耗等优势,与人工智能应用场景更契合。而且,ARM平台统治着智能时代的终端设备,在未来,端与云会有更深层次的融通,如果端与云都能统一在ARM架构上,一定能有效提升全系统效率。”范东睿解释说,这也是中科睿芯与高通深度合作,研发世界首款基于ARM架构的高通量AI一体机的初衷。

“我们既要在移动端的低功耗AI芯片领域完成应用场设计,也要在数据中心的单位空间,部署具有高密度计算资源的超级AI服务器。随着互联网应用深层次演进和变革,高通量计算将覆盖到生活中的方方面面。”范东睿对高通量计算的前景很乐观。

## 新鲜事

### Enaex研发矿工机器人 或有望实现采矿自动化



据网易智能报道,近日,为矿山运营商提供爆破服务的智利公司Enaex与研发机构SRI International合作,尝试创造一种矿工机器人,并把这种自动化机器人解决方案称为RoboMiner。

该人形机器人由仿生化躯干、头部和手臂组成,机器上的有效载荷包括可以实现立体视觉、力反馈、场景渲染和定位的传感器。它被安装在一个带有四个坚固轮子的移动平台上,可以通过远程控制被派往危险的矿井环境下工作,适应各种地形情况,能够在5—10公斤的承重范围内像人类一样移动。这使得运营商能够到达那些威胁人类安全的高风险采矿领域。这款机器人最初计划在露天矿山运营,最终会实现地下来采。

### 亚马逊推出AI摄像头 面向人工智能开发者



据新浪科技报道,近日,亚马逊正式推出DeepLens摄像头,它是一个面向AI开发者的定制摄像头,开发者可以在AWS(Amazon Web Services)上开发AI系统,让系统在摄像头上传运行。

DeepLens摄像头可以识别对象、搜寻脸部图像,或者给人的动作贴上标签,完成上述任务时不需要回访问亚马逊服务器。

智能手机流行后,移动处理器的制造成本越来越低,亚马逊认为未来世界上大多数计算不会交给大型服务器处理,而是由摄像头、智能手机及其他电子设备的内部组件处理,这些产品会收集数据。而即使开发者的AI不在云端运行,亚马逊也有软件供DeepLens摄像头用户使用。

但是,DeepLens摄像头仍有一些限制,比如摄像头本身可以侦测视野内是否有脸部图像,但要判断脸部是谁的,还需要回访问亚马逊云服务。

### 英特尔与华纳兄弟合作 开发无人驾驶车内娱乐系统



据新浪科技报道,近日,英特尔在洛杉矶车展上公布消息,与华纳兄弟围绕自动驾驶汽车达成合作,目标是在汽车行驶过程中为乘客提供娱乐体验。

英特尔年初收购了汽车传感器制造商Mobileye,如今又与华纳兄弟合作进行车内系统开发。英特尔方面称,未来汽车不再需要人来驾驶,也不需要车中人盯住路面,所以人在车中坐着,就可以看电影、电视或进行其他娱乐。汽车车窗可以变成AR屏幕,汽车本身也可以变成移动广告牌。英特尔CEO科再奇说:“我们与华纳兄弟合作,将这种体验变成现实。”

华纳兄弟首席数字官托马斯·格维克认为,自动驾驶汽车可以延长娱乐时间。

(本版图片除标注外来源于网络)

# 智能交互赛事升级 移动终端成AI热点赛道

本报记者 刘艳

11月30日,三星人工智能语音Bixby中文(普通话)版推送到三星S8系列及三星Note8手机端。

这条看似波澜不惊的消息预示着,随着人工智能交互功能的升级,手机不仅能听懂、看

## 交互助手大战升级

虽然智能终端的AI助手早就不是什么新奇事物,但AI助手战情的胶着程度却在不断升级。

最早出现于智能终端上的AI助手是苹果的Siri,随后,在PC领域追着苹果打的微软推出了小娜、小冰,现在,小娜每月的活跃用户已达1.48亿。亚马逊的Alexa虽然与微软小娜同年发布,但大家更喜欢拿它与2016年出生的Google Assistant进行对比。

2017年才随着三星S8推出的Bixby看起来有点姗姗来迟了。

然而,打着人工智能交互概念的“助手”出现这么多年,并没有获得消费者的一致好评,原因在于人们始终觉得这些“助手”离智能还差那么一点。

通信行业专家柏松说:“Bixby打破了‘助手’的概念,变成了一个真正意义上的人工智能交互平台,它打开的是AI多模交互方式,无论是硬件底层支持还是语音语义识别都领先其他语音助手不止一个level。关键是,其他产品多数只有智能语音控制功能,但语音只是

## 芯片领域竞争加剧

可以预见的是,未来的AI交互产品将渗透所有智能电子设备,成为物联网智能的“基础设施”,移动终端将成为全球最普遍的人工智能平台,AI芯片的问世加速了这一进程。

事实上,相关厂商和资本向终端聚拢的趋势已很明显,终端芯片逐渐成为AI领域的热

Bixby的功能之一。”

也就是说,Bixby不仅仅是又一个挤入市场的语音助手,它可以通过语音、拍摄等操作方式完成相对复杂的操作流程。从简单的设定闹钟到自动打开并发布微博,从自动清除系统缓存到智能识物、拍照购物,甚至根据预设场景让多个APP组合执行指令。而目前其他语音助手只能执行简单的常规操作。

柏松说:“智能手机是与用户关系最密切的智能化产品,语音功能是用用户最喜欢的交互方式。当硬件同质化越来越严重时,提升用户体验的软件改造任务落在了人工智能身上。但是,走出语音和手机,AI交互还有更多可能。”

当亚马逊、谷歌、苹果首先将AI交互产品延伸到音箱时,三星似乎正在用一种全新的方式与未来接轨,跳出为智能手机增添功能的范畴,将Bixby“作为用户设备生态的控制中枢”,指向物联网智能功能系统。

据了解,被设计成完整体系的Bixby现在能控制三星手机,未来将成为三星全系产品的标配。

点赛道。

从互联网巨头及活跃的风险投资、股权投资不断对中国AI领域加码的资本布局看,与AI芯片相关的中国公司正成为投资热点。

10月24日,深鉴科技对外宣布完成约4000万美元A+轮融资,此前一直保持神秘的

人工智能新创公司耐能(Kneron)11月15日宣布完成超过千万美元的A轮融资。

与此同时,全球IT巨头包括英伟达、英特尔、高通都在加紧AI芯片的布局。

10月20日,国内最大的人工智能创业公司商汤科技与高通宣布,双方将围绕移动终端和物联网领域,在人工智能和机器学习方面展开合作。商汤科技的算法将针对高通的骁龙平台进行针对性优化,高通的骁龙芯片则向商汤开放“能力”,“算法+芯片的融合优势”将被释放。

“多少被看好的芯片厂家出局,多少厂家试图超越,高通移动芯片的霸主地位现在无人能替代。”通信行业专家赵宇说。

下一个十年谁也不敢说,至少在这轮竞赛中,中国公司并不落后。

今年8月18日,中科院计算所孵化的的寒

## 手机成AI应用前沿阵地

Gartner数据显示,未来5年智能手机出货量将累计达到85亿部,这让手机成为AI应用落地智能终端的前沿阵地。

去年12月,荣耀推出了全行业第一款人工智能手机Magic。荣耀总裁赵明在接受记者采访时表示,这款被定义为“未来手机”的产品,灌注了华为和荣耀积蓄多年的人工智能技术和理解,几百人的团队研发了四年,这款对人工智能手机进行前期探索的产品,不以利润和销售额来衡量它的成败。

11月28日,荣耀发布了搭载麒麟970芯片的第二代人工智能手机V10,打出“秒懂、秒杀、秒拍”的旗号,强调一个“快”字。

赵明表示,“传统意义上手机的快,是不断加强手机硬件的处理速度,优化软件的流畅度,但是,现在速度提升已经遇到瓶颈。AI加速相较传统操作有大幅提升。”

不仅“手机加速”遇到了瓶颈,整个手机行

业也在发展瓶颈期,产品创新从特色走向平庸,基于传统思维建立的产品竞争优势被拉平,人工智能成手机行业破局关键点。苹果、华为、荣耀,已引发了人工智能手机的整体热潮,也引来争吵,焦点是,“全面屏”的热度和竞争力要让位于人工智能了?

但是,11月28日,360集团创始人兼CEO周鸿祎为360旗下首款全面屏手机N6Pro站台。周鸿祎和360手机总裁李开新都表现出对这款手机的自信。

李开新强调,这是一款为游戏而生的手机,在“游戏加速器”运行时,手机处理器八核全开,游戏将始终获得最多的性能资源。

“在手游风靡的今天,这绝对是一大卖点。看似人工智能抢了全面屏的风头,但并不意味着它是唯一竞争力,包括全面屏在内的其他性能指标的综合提升才是智能手机真正的升级。”柏松说。

# 中国深度学习创业的长板在哪里

## 产业界

本报记者 马爱平

在前不久召开的AI WORLD 2017世界人工智能大会上,硅谷知名企业家、美国电子电气工程师协会士克里斯·罗恩分析了中国深度学习创业市场,他认为中国的深度学习公司只关注少数几个领域,比如嵌入式系统、监控、人机互动界面、无人驾驶,做云工作的相对较少。而就全世界来说,三分之二的深度学习创业公司都在做云软件,在中国只有20%的公司在做;在全世界大概有一半深度学习创业都在做视觉,而中国只有17%。

“在23个我所追踪的中国创业公司中,没有一个做制造业,没有一个做金融,也没有一个做

人力资源或者客户关系管理方面的相关工作。也就是说,对中国的创业公司来说,这方面是一个空白。”罗恩说。

国内AI创业生态是否真如罗恩所说?

“这样说并不全面。从整体体量来看,中国AI创业公司数量可能不及美英等国家,但做云工作的并不在少数。云端应用对于大多数中国AI公司来说都是基础设施。比如旷视科技Face++在业务上更多侧重云+端的感知技术在金融安全、城市大脑等一些实际的应用场景,而我们最底层的一些工作还是通过云端AI开放平台去实现的。”11月28日,旷视科技Face++品牌与市场中心总经理谢忆楠告诉科技日报记者。

谢忆楠认为,视觉、语音和服务机器人应该是中国大多数AI创业公司在做的,罗恩的统计

可能不尽全面或者并不了解中国AI行业现状。“但是制造业的确是AI创业的一个短板,这和中国长期以来产能过剩、处于世界产业链低端的大环境有关。”谢忆楠说。

实际上,中国深度学习创业市场的长板非常显著。比如人才方面,尤其是计算机视觉领域,中国拥有众多杰出的CV学者和专家,在国际范围的顶级会议和竞赛中有着非常高的活跃度和竞争实力,比如在今年IEEE国际计算机视觉与模式识别会议接收的论文中有将近一半都有华人学者参与。

在数据方面,中国的深度学习创业市场也优势明显。“数据都是基于人产生的,而中国拥有全球最多的人口,且中国移动互联网发达程度也在全球处于领先地位,所以我们有海量的数据可以去挖掘。另外,人脸识别技术和应

用来讲,中国拥有最全的人脸数据库,但是在国外基本没有哪个国家能有比较全的人脸数据。”谢忆楠说。

在AI WORLD 2017世界人工智能大会上,业内人士也纷纷表示,相比海外,中国“互联网+”战略为人工智能的进一步应用打造了非常好的基础,同时也衍生出了众多需求场景,比如互联网金融,用智能的方式实现转账、支付等业务操作,不管是技术上还是策略制定上中国比一些西方国家走的更快得多。

另外,人工智能将成为未来国际竞争中的关键,在政策上,从近两年开始,中国政府已经把人工智能上升到国家战略层面,这也是很多西方国家还没有意识到的。而政策的倾斜则意味着人工智能发展的基础设施平台建设有了保障。