

人工智能芯片产业技术发展研究

商惠敏

(广东省科学技术情报研究所, 广州 510033)

摘 要: 人工智能芯片是人工智能技术的核心硬件基础, 对人工智能技术的发展影响巨大。当前人工智能芯片的市场由英伟达、英特尔、赛灵思、谷歌等国际巨头占据主要份额, 国内企业近年来加速布局, 在边缘推理芯片等领域发展较快, 但整体上距国际先进水平还有较大的差距。本文阐述了人工智能芯片的概念和分类, 深入分析了人工智能芯片产业的发展现状、发展趋势等, 并提出我国在人工智能芯片领域应加大布局 and 研发力度。

关键词: 人工智能; 芯片; 芯片产业

中图分类号: G311; F124.3 **文献标识码:** A **DOI:** 10.3772/j.issn.1009-8623.2021.12.005

人工智能是新一轮科技革命和产业变革的重要驱动力^[1]。当前, 人工智能技术已全面渗透到制造、医疗、交通、金融、教育、安防等众多领域。算力、算法、数据是人工智能发展最重要的三大要素。其中, 算力主要由人工智能芯片支撑, 是承载人工智能核心技术的硬件基础。

1 人工智能芯片概述

1.1 人工智能芯片的概念内涵

广义而言, 所谓人工智能芯片, 是指专门用于运行人工智能算法且做了优化设计的芯片。为满足在不同场景下的人工智能应用需求, 人工智能芯片逐渐表现出专用性、多样性的特点。

1.2 人工智能芯片的分类

1.2.1 按技术架构分类

GPU: 即图形处理单元, 是一种由大量运算单元组成的大规模并行计算架构芯片, 主要用于处理图形、图像领域的海量数据运算。GPU 上集成了规模巨大的计算矩阵, 从而具备了更强大的浮点运算能力和更快的并行计算速度, 与 CPU 相比, 更

加适用于解决人工智能算法的训练难题。英伟达的 GPU 目前在人工智能计算市场上占据了主导地位。

半定制化 FPGA: 即现场可编程门阵列。与 GPU 的固定电路不同, 使用者可以根据不同的应用需求, 使用硬件描述语言对 FPGA 芯片上集成的基本门电路和存储器进行重新定义。按照新的定义完成烧录后, FPGA 芯片内部的电路就固化成了实际的连线, 从而具备了使用者所需要的功能。此类芯片非常适合在芯片功能尚未完全定型、算法仍需不断迭代完善的情况下使用。使用 FPGA 芯片需要通过定义硬件去实现软件算法, 对使用者的技术水平要求较高, 因此在设计并实现复杂的人工智能算法方面难度较高。赛灵思和英特尔在 FPGA 领域具有较大的优势。

全定制化 ASIC: 即专用芯片, 是一种根据特殊应用场景要求进行全定制化的专用人工智能芯片。与 FPGA 相比, ASIC 芯片无法通过修改电路进行功能扩展; 而与 CPU、GPU 等通用计算芯片相比, 其性能高、功耗低、成本低 (见表 1), 也很适合应用于对性能功耗比要求极高的移动设备

作者简介: 商惠敏 (1981—), 女, 副研究员, 主要研究方向为科技情报与科技战略、产业技术分析。

项目来源: 广东省重点领域研发计划“人工智能技术标准研究及公共支撑服务平台建设” (2018B010112001)。

收稿日期: 2021-10-18

端^[2]。谷歌公司发布的 TPU 芯片是当前最知名也最有实用价值的 ASIC 芯片。

神经拟态芯片：即类脑芯片，是一种对人脑的神经网络结构进行物理模拟的新型芯片架构，通过模拟人脑的神经网络工作机理实现感知和认知等功能^[3]。IBM 研发的 TrueNorth 芯片就是一种典型的

类脑芯片，其逻辑结构颠覆了经典冯·诺依曼架构，把定制化的数字处理内核当作神经元，把内存当作突触，CPU、内存及通信元件等完全集成在本地^[4]，实现了算存一体，突破了冯·诺依曼架构中 CPU 与内存之间的内存墙瓶颈，但目前多数仍是实验室产品。

表 1 人工智能芯片特点对比

芯片类型	适用场景	主要优点	主要缺点
GPU	应用于高级复杂算法和通用性人工智能平台	通用性强、浮点运算能力强、速度快	性能功耗比较低
半定制化 FPGA	适用于具体行业	灵活性强、速度快、功耗低、可编程性强	价格高、编程复杂、整体运算能力不高
全定制化 ASIC	适用于全定制场景	性能高、功耗低	电路需定制、开发周期长、难以扩展、风险高
神经拟态芯片	实验室、自研平台	功耗低、可扩展性强、算存一体	算力低、对主流算法支持差、难以商用

1.2.2 按部署位置分类

云端人工智能芯片：这类芯片运算能力强大，功耗较高，一般部署在公有云、私有云、混合云或数据中心、超算等计算基础设施领域，主要用于深度神经网络模型的训练和推理，处理语音、视频、图像等海量数据，支持大规模并行计算，通常以加速卡的形式集成多个芯片模块，并行完成相关计算任务。

边缘端人工智能芯片：这类芯片一般功耗低、体积小、性能要求不高、成本也较低，相比于云端芯片来说，不需要运行特别复杂的算法，只需具备少量的人工智能计算能力，一般部署在智能手机、无人机、摄像头、边缘计算设备、工控设备等移动设备或嵌入式设备上。

1.2.3 按功能任务分类

人工智能训练芯片：训练是指向人工智能算法模型输入大量已标注好的数据和素材，进行“学习”，对模型的参数不断进行优化调整，最终形成一个具备某种特定功能、结果最优的神经网络算法模型；人工智能训练芯片即是指专门对人工智能训练算法进行优化加速的芯片，由于训练所需的数据量巨大，算法复杂度高，因此，训练芯片对算力、能效、精度等

要求非常高，而且还要具备较高的通用性，以支持已有的多种算法，甚至还要考虑未来的算法的训练。由于对算力有着极高要求，训练芯片一般更适合部署在大型云端设施中，而且多采用“CPU+GPU”“CPU+GPU+加速芯片”等异构模式，加速芯片可以是 GPU 或 FPGA、ASIC 专用芯片等。人工智能训练芯片的市场主要被英伟达的 GPU 和谷歌的 TPU 所占据，英特尔和 AMD 都在积极进入该领域^[5]。

人工智能推理芯片：推理是指向已经训练好的人工智能算法模型输入新的数据和素材，经过计算后获得符合人们预期的相应的输出；人工智能推理芯片即是指专门对人工智能推理算法进行优化加速的芯片，其更加关注能耗、算力、时延、成本等综合因素。其可以部署在云端和边缘端，实现难度和市场门槛相对较低，因此，这一领域的市场竞争者较多。云端推理芯片领域，英伟达、谷歌、AMD、赛灵思等传统芯片厂商是主要的领导者，国内的寒武纪、燧原科技、比特大陆也推出了性能较高、市场反响不错的自研芯片。终端推理芯片领域，应用场景丰富，市场集中度不高，产品有一定的多样性，英伟达、英特尔、高通、ARM 等传统芯片大厂在该领域布局较早，国内的寒武纪、地平

线、阿里平头哥、云天励飞等新兴企业在垂直行业也有不俗表现。

1.3 人工智能芯片的发展历程

第一阶段（2006 年以前）：在这一阶段，尚未出现突破性的人工智能算法，且能够获取的数据也较为有限，传统通用 CPU 已经能够完全满足当时的计算需要，学界和产业界均对人工智能芯片没有特殊需求，因此，人工智能芯片产业的发展一直较为缓慢。

第二阶段（2006—2010 年）：在这一阶段，游戏、高清视频等行业快速发展，同时也助推了 GPU 产品的迭代升级。2006 年，GPU 厂商英伟达发布了统一计算设备架构（CUDA），第一次让 GPU 具备了可编程性，让 GPU 的核心流式处理器既具有处理像素、顶点、图形等渲染能力，又同时具备通用的单精度浮点处理能力，即令 GPU 既能做游戏和渲染，也能做并行度很高的通用计算，英伟达称之为 GPCPU。统一计算设备架构推出后，GPU 编程更加易用便捷，研究人员发现，GPU 所具有的并行计算特性比通用 CPU 的计算效率更高，更加适用于深度学习等人工智能先进算法所需的“暴力计算”场景。在 GPU 的助力下，人工智能算法的运算效率可以提高几十倍，由此，研究人员开始大规模使用 GPU 开展人工智能领域的研究和应用。

第三阶段（2010—2015 年）：2010 年之后，以云计算、大数据等为代表的新一代信息技术高速发展并逐渐开始普及，云端采用“CPU+GPU”混合计算模式使得研究人员开展人工智能所需的大规模计算更加便捷高效，进一步推动了人工智能算法的演进和人工智能芯片的广泛使用，同时也促进了各种类型的人工智能芯片的研究与应用。

第四阶段（2016 年至今）：2016 年，采用 TPU 架构的谷歌旗下 DeepMind 公司研发的人工智能系统阿尔法围棋（AlphaGo）击败了世界冠军韩国棋手李世石，使得以深度学习为核心的人工智能技术得到了全球范围内的极大关注。此后，业界对于人工智能算力的要求越来越高，而 GPU 价格昂贵、功耗高的缺点也使其在场景各异的应用环境中受到诸多限制，因此，研究人员开始研发专门针对人工智能算法进行优化的定制化芯片。大量人工智能芯片领域的初创公司在这一阶段涌现，传统互

联网巨头也迅速入局该领域争夺市场，专用人工智能芯片呈现出百花齐放的格局，在应用领域、计算能力、能耗比等方面都有了极大的提升^[6]。

2 人工智能芯片产业发展现状

2.1 传统芯片企业在人工智能芯片领域优势地位明显

英伟达、英特尔、AMD、高通等传统芯片厂商凭借在芯片领域多年的领先地位，迅速切入人工智能领域，积极布局，目前处于引领产业发展的地位（见表 2），在 GPU 和 FPGA 方面则基本位于垄断地位。英伟达推出了 Tesla 系列 GPU 芯片，专门用于深度学习算法加速；推出了 Tegra 处理器，应用于自动驾驶领域，并提供配套的研发工具包。AMD 于 2018 年推出了 Radeon Instinct 系列 GPU，主要应用在数据中心、超算等人工智能算力基础设施上，用于深度学习算法加速。当前，GPU 作为业界使用最为广泛、人工智能计算最成熟的通用型芯片，成为数据中心、超算等大型算力设施的首选，占据了人工智能芯片的主要市场份额。在效率和场景应用要求大幅提升和变化之前，GPU 仍将是人工智能芯片领域的主要领导者。

2.2 互联网及 IT 巨头纷纷加大人工智能芯片自研定制力度

2015 年以来，谷歌、IBM、Facebook、微软、苹果、亚马逊等国际互联网及 IT 巨头纷纷跨界开展人工智能芯片研发，力图突破算力瓶颈，并把核心部件掌握在自己手中。如谷歌于 2016 年发布了专门针对开源框架 TensorFlow 开发的芯片 TPU，并帮助 AlphaGo 击败李世石；近年，谷歌还推出了可在 Google Cloud Platform 中使用的云端芯片 Cloud TPU 以及用于边缘端推理的 Edge TPU，打造闭环生态。微软于 2017 年发布了基于 FPGA 芯片组建的 Project Brainwave 低时延深度学习系统，让微软的各种服务可以更迅速地支持人工智能功能。2018 年，亚马逊发布了高性能推理芯片 AWS Inferentia，支持 TensorFlow、Caffe2 等主流框架。

2.3 国内人工智能芯片行业发展迅猛，但仍处在起步阶段

目前，在 CPU、GPU 等高端通用芯片领域，

表2 国外典型人工智能芯片产品

序号	公司	典型人工智能芯片	发布年份	技术架构	功能任务
1	英伟达	Tesla V100	2017	GPU	云端训练、云端推理
		Tesla A100	2020	GPU	云端训练、云端推理
2	英特尔	Nervana NNP-T	2019	NNP-T1000	云端训练
		Nervana NNP-I	2019	NNP-I1000	云端推理
3	IBM	TrueNorth	2015	类脑芯片	边缘端推理
4	谷歌	TPUv3	2018	ASIC	云端训练、云端推理
		Edge TPU	2018	ASIC	边缘端推理
5	苹果	A14	2020	ARM 架构 SoC	边缘端推理
6	AMD	EPYC2	2019	Zen2 架构	云端推理
7	ARM	ARM Cortex-M55	2020	ARM Helium	边缘端推理
		ARM Ethos-U55	2020	micro NPU	边缘端推理
8	高通	骁龙 888	2020	ARM 架构 SoC	边缘端推理
		Cloud AI 100	2020	ASIC	云端推理
9	三星	Exynos 2100	2021	ARM 架构 SoC	边缘端推理
10	赛灵思	Versal ACAP	2019	SoC	云端推理等

我国的设计能力与国外先进水平仍然差距较大，部分自研芯片采用了 ARM 架构等国外成熟芯片架构和 IP 核等进行设计，自主研发能力较弱。但是，随着人工智能技术大规模应用于安防、金融、政务、自动驾驶、智能家居等领域，促进了各类专用人工智能芯片的发展，我国的一些初创型企业，如寒武纪、地平线、云天励飞、深鉴科技等也开始在人工智能芯片领域有所建树^[7]（见表3）。我国人工智能芯片企业基本都围绕边缘端语音、视觉芯片进行开发，从事云端芯片研发尤其是云端训练芯片的企业较少，仅华为、百度等有产品推出，我国云端芯片与国外技术水平差距仍然较大。此外，我国还尚未形成有影响力的“芯片—算法—平台—应用—生态”的产业生态环境，企业多热衷于追逐市场热点，缺乏基础技术积累，研发后劲不足^[8]。

2.4 类脑芯片领域呈现异军突起之势

IBM 公司率先在类脑芯片领域取得突破，2014 年推出了 TrueNorth 类脑芯片，采用 28nm 工艺，

集成了 54 亿个晶体管，包括 4096 个内核、100 万个神经元和 2.56 亿个神经突触。2019 年，清华大学施路平教授团队发布了类脑芯片“天机芯”，使用 28 纳米工艺流片，包含约 40 000 个神经元和 1 000 万个突触，支持同时运行卷积神经网络、循环神经网络以及神经模态脉冲神经网络等多种神经网络，是全球首款既能支持脉冲神经网络又可以支持人工神经网络的异构融合类脑计算芯片^[9]。西井科技发布的 DeepSouth 芯片，核心是用 FPGA 模拟神经元以实现脉冲神经网络的工作方式，包含约 5 000 万个神经元和高达 50 多亿个神经突触，可以直接在芯片上完成计算，并在“无网络”情况下使用，处理相同计算任务时，DeepSouth 芯片的功耗仅为传统芯片的几十至几分之一。浙江大学与杭州电子科技大学共同研发了“达尔文”芯片，集成了 500 万个晶体管，包含 2 048 个硅材质的仿生神经元和约 400 万个神经突触，可从外界接受并累积刺激，产生脉冲信号，处理和传递信息。

表3 中国典型人工智能芯片产品

序号	公司	典型人工智能芯片	发布年份	技术架构	功能任务
1	中科寒武纪	思元 270	2019	ASIC	云端推理
		思元 220	2019	ASIC	边缘推理
2	地平线机器人	征程二代	2019	ASIC	边缘推理
		旭日	2019	ASIC	边缘推理
3	华为	昇腾 310	2018	达芬奇架构	边缘推理
		昇腾 910	2019	达芬奇架构	云端训练
4	中星微电子	星光智能一号 VC0758	2019	ASIC	边缘推理
5	云天励飞	DeepEye1000	2019	SoC	边缘推理
6	灵汐科技	天机芯	2019	类脑芯片	边缘推理
7	比特大陆	算丰 BM1684	2019	ASIC	云端推理
		算丰 BM1880	2018	ASIC	边缘推理
8	百度	昆仑	2018	FPGA	云端训练、云端推理
9	平头哥	含光 800	2019	FPGA	云端推理
		玄铁 910	2019	RISC-V	边缘推理
10	依图科技	求索	2019	ARM 架构	云端推理

3 人工智能芯片产业发展趋势

3.1 物联网领域不断涌现的应用需求将催生超低功耗人工智能芯片

随着以人工智能、物联网、5G 等为核心的新一代信息技术的高速发展,涌现出越来越多新的应用场景和需求。未来物联网领域将需要体积更小、功耗更低、能效比更高的人工智能芯片。常见的边缘端芯片如手机中的人工智能芯片,其功耗一般在几百毫瓦至几瓦,云端训练芯片的工作功耗则更高,通常要达到数百瓦,而超低功耗人工智能芯片的工作功耗一般是几十毫瓦甚至更低。同时,由于芯片的计算模块在大多数时间里均处于休眠状态,只有在发生相关事件时才会在事件驱动技术的支持下被激活为工作状态,这样就进一步降低了平均功耗。如在以智能手表为代表的智能可穿戴设备领域,设备的电池容量因尺寸等原因受到极大限制,而此类设备需要具备心率检测、手势识别、语音识别等智能生物信号处理功能,因此需要集成体积

小且能效比超高的人工智能加速芯片,降低对电池的消耗;在智能家居等领域,具备人脸识别、指纹识别等功能的智能门锁须由电池供电,而且不能经常更换电池,否则会降低用户体验,这就对门锁中执行人脸识别等功能的智能模块提出了极高的能效比要求。除消费电子之外,制造业等工业应用场景中也需要使用超低功耗人工智能芯片,如安装在机械臂、管道等重要设备和环境中的智能传感器须由电池供电,使用超低功耗人工智能芯片则可以有效减少电池消耗,大幅降低此类设备的维护成本^[10]。

3.2 开源芯片有望开辟新的赛道

当前,摩尔定律已逼近极限,传统通用芯片的性能提升也逐渐走向瓶颈^[11],通用处理器架构越来越难以适应需求各异的人工智能算法和广泛的应用场景,对新型架构人工智能芯片的需求日益增长,为各类初创型中小企业带来新的市场机遇。然而,芯片领域过高的技术门槛和知识产权壁垒,严

重阻碍了人工智能芯片的进一步技术创新和发展。开源芯片的兴起有望突破这一瓶颈。开源芯片大幅降低了芯片设计领域的门槛,为企业节省了芯片架构和IP核等方面的授权费用,可以有效降低企业的研发成本。同时,由于开源社区的开发者们会持续不断地对开源芯片进行更新迭代,企业可以免费获取到最新、最优化的版本,并向社区贡献自己的力量,不断提升行业整体发展水平,有效促进人工智能芯片产业的繁荣。2014年,美国加州大学伯克利分校的研究团队正式发布了“RISC-V”开源精简指令集架构,具有灵活简洁、模块化、扩展性强、易实现等优点,可以较好地适应专用硬件设备、高性能计算设备、低功耗嵌入式设备等众多应用领域的需求,而且“RISC-V”完全免费,可以被任何人自由地用于任何目的,因此,“RISC-V”也成为目前推广度、普及度最高的开源芯片项目。此外,伯克利创建了开源服务社区,向开发者们提供完善的软件工具链。目前,“RISC-V”已有大量的开源实现和流片案例,如西部数据公司于2018年发布了基于“RISC-V”的自研处理器架构SweRV;阿里平头哥于2019年正式发布了基于“RISC-V”的处理器玄铁910;中科院计算所于2021年6月发布了国产开源高性能“RISC-V”处理器核“香山”,其首版架构“雁栖湖”即将流片。

3.3 人工智能芯片将从确定算法、场景的加速芯片向通用智能芯片发展

近几年,人工智能技术在语音识别、视频图像识别等应用领域取得突破性的进展,但要实现从单点突破走向全面开花,需要人工智能领域产生像CPU一样的通用人工智能计算芯片,适用于任意人工智能应用场景。从目前来看,短期内人工智能芯片仍以“CPU+GPU+AI加速芯片”的异构计算模式为主,中期会重点发展可自重构、自学习、自适应的人工智能芯片,未来将会走向通用的人工智能芯片。通用人工智能芯片就是能够支持和加速任意人工智能计算场景的芯片,即通过一个通用的数学模型,最大程度概括出人工智能的本质,其在经过一定程度的学习后,能够精确、高效地处理任意场景下的智能计算任务。通用人工智能芯片发展的主要难点在于通用性和实现的复杂度,同时,还面临着传统冯·诺伊曼架构的技术瓶颈以及摩尔定律接近物理

极限这两大挑战。未来,随着芯片的制程工艺、新型半导体材料和物理器件等出现新突破,以及人类对于大脑和智能本身形成更深层次的认知,将有望最终实现真正意义上的通用人工智能芯片。

4 对我国的建议

我国人工智能应用市场规模不断增长,在5G、大数据、物联网、云计算等先进技术的支撑下,已经不断累积,形成了海量的数据资源,为人工智能芯片提供了广阔的发展空间^[12]。我国应在人工智能芯片领域持续加大关注力度,围绕产业布局、技术攻关、企业扶持、构建人工智能创新生态等方面进行部署,构筑人工智能芯片产业发展新优势。

(1) 加强人工智能芯片产业布局。加强规划引导、分类指导,积极发挥“国家新一代人工智能创新发展试验区”和“国家人工智能创新应用先导区”的示范引领作用,开展人工智能技术示范、政策试验和社会实验,加速人工智能技术和产品的应用推广,持续推进与实体经济深度融合。探索打造人工智能芯片设计领域的产业园区,因地制宜有针对性地开展招商引资,培育更多优质项目和人才团队。

(2) 强化核心技术攻关,优化产业技术创新体系。围绕基础前沿理论、新型体系架构、算法框架、云端训练芯片、类脑芯片、开源芯片等方面组织科技攻关,力争形成一批重大标志性科技成果和一批自主可控的核心软硬件产品。充分发挥华为、寒武纪、平头哥、云天励飞、全志科技等行业龙头骨干企业的示范引领作用,鼓励其在部分领域实现底层软硬件系统的开源开放,持续输出核心研发能力,提升产业技术创新能力。

(3) 支持人工智能芯片企业做大做强。培育人工智能芯片领域的创新型标杆企业,支持创新型企业在科创板、创业板、新三板上市融资,打造一批人工智能芯片细分领域的“独角兽”“瞪羚”“隐形冠军”企业。降低人工智能芯片企业创新创业成本,提高芯片设计、流片等的补贴力度。鼓励有条件的企业建设芯片设计领域的人工智能开放创新平台,打造知识共享和经验交流社区,鼓励人工智能芯片硬件层面的开源开放,降低人工智能芯片技术研发和进入门槛,支持创新创业者、研发团队和中小微企业投身人工智能芯片技术创新。

(4) 培育以市场应用为牵引的产业氛围, 实现“芯片—算法—平台—应用—生态”协同发展。充分发挥我国人工智能数据资源丰富、应用场景广阔的优势, 以创新应用牵引人工智能芯片发展, 推动人工智能芯片、算法、平台、应用、整机等领域的企业之间建立有机互动和合作研发战略联盟关系, 开展深度合作和协同创新。支持国资企业、大型科研机构采用国产芯片、设备等建设自主可控的人工智能算力平台, 为各类应用场景提供强大的算力支撑。■

参考文献:

- [1] 新华社. 习近平在中共中央政治局第九次集体学习时强调——加强领导做好规划明确任务夯实基础 推动我国新一代人工智能健康发展 [EB/OL]. [2021-07-13]. http://www.gov.cn/xinwen/2018-10/31/content_5336251.htm.
- [2] 张蔚敏, 蒋阿芳, 纪学毅. 人工智能芯片产业现状 [J]. 电信网技术, 2018 (2): 67-71.
- [3] 赵崇文. 人工神经网络综述 [J]. 山西电子技术, 2020 (3): 94-96.
- [4] 张军阳, 王慧丽, 郭阳, 等. 深度学习相关研究综述 [J]. 计算机应用研究, 2018 (7): 1 921-1 928, 1 936.
- [5] 北京未来芯片技术高精尖创新中心. 人工智能芯片技术白皮书 [R/OL]. [2021-07-13]. <http://www.icfc.tsinghua.edu.cn>.
- [6] 亿欧智库. 2019 年中国 AI 芯片行业研究报告 [R/OL]. [2021-07-13]. <http://www.199it.com/archives/859776.html>.
- [7] 清华 - 中国工程院知识智能联合实验室. 人工智能芯片研究报告 [R/OL]. [2021-07-13]. AMiner.org.
- [8] 施羽暇. 人工智能芯片技术体系研究综述 [J]. 电信科学, 2019 (4): 114-118.
- [9] 李钢, 李繁荣, 程健. 应用场景需求: 驱动人工智能芯片设计发展 [J]. 前沿科学, 2018 (4): 37-40.
- [10] 半导体行业观察. AI 芯片的新风向 [EB/OL]. [2021-07-13]. <https://www.jiqizhixin.com/articles/2020-07-14-6>.
- [11] 史入文. 美国人工智能芯片研发动态 [J]. 上海信息化, 2019 (11): 80-82.
- [12] 丛瑛瑛, 陈丝. 人工智能芯片发展态势分析及对策建议 [J]. 信息通信技术与政策, 2018 (8): 65-68.

Research on Technology Development of Artificial Intelligence Chip Industry

SHANG Hui-min

(Guangdong Institute of Scientific and Technical Information, Guangzhou 510033)

Abstract: Artificial intelligence (AI) chip is the core hardware foundation of AI technology, which has a great impact on the development of AI technology. At present, the market of AI chips is dominated by international giants such as NVIDIA, Intel, Xilinx and Google. In recent years, domestic enterprises have accelerated their layout and developed rapidly in the field of edge inference chips, but there is still a big gap from the international advanced level. This paper expounds the concept and classification of AI chip, deeply analyzes the development status and trend of AI chip industry, and puts forward that China should increase the layout and R&D efforts in the field of AI chip.

Keywords: artificial intelligence; chip; chip industry