

核心观点

❖ 华为发布全球首款内置 NPU 人工智能处理器-麒麟 970

华为在 IFA2017 大会上公布了麒麟 970 处理器，是全球首款内置 NPU 的处理器。麒麟 970 采用异构架构完成人工智能处理，另一创新之处在于端云结合实现人工智能。麒麟 970 芯片的发布使得人们再度聚焦人工智能芯片的发展。

❖ 人工智能芯片的崛起之路

传统 CPU 的架构在解决深度学习任务时效率低、成本高，催生新硬件。人工智能芯片的出现不是为了执行指令，而是为了大量数据训练和应用的计算。其中，GPU 因其海量数据并行运算能力被首先引入深度学习。人工智能是互联网诞生以来的第二次技术社会形态在全球的萌芽，中国和海外国家也分别从国家层面对人工智能产业进行战略布局。

❖ AI 芯片三种技术路线，ASIC 是终端应用趋势

目前适合深度学习的人工智能芯片主要有 GPU、FPGA、ASIC 三种技术路线。三类芯片代表分别有英伟达 (NVIDIA) 的 Tesla 系列 GPU、赛灵思 (Xilinx) 的 FPGA 和 Google 的 TPU。GPU 最先被引入深度学习，技术最为成熟；FPGA 具有硬件可编程特点，性能出众但壁垒高。ASIC 由于可定制、低成本是未来终端应用的趋势。

❖ 科技巨头加紧布局 AI 芯片，寒武纪跻身国际前列

全球科技巨头都在加紧布局 AI 芯片，希望走在科技变革时代的前线。NVIDIA 是 AI 芯片的市场领先者，占据了全球 GPU 70% 的市场份额；Intel 接连收购 Altera 等公司，全方位布局 AI 产品；Google 发布两代 TPU，从 ASIC 方向进军 AI 芯片市场；而国内的寒武纪科技是中科院计算所孵化的一家独角兽公司，2016 年推出了国际首个深度学习专用处理器芯片 (NPU)，技术全球领先。

❖ 该领域主要龙头企业

AI 芯片是人工智能领域的上游，基于 AI 芯片的下游场景应用极为丰富，包括安防、消费电子、自动驾驶、可穿戴设备等。由于 A 股市场没有像英伟达、Google 这样的 AI 芯片企业，可以从 AI 芯片产业的下游精选个股。相关标的包括富瀚微 (国内领先的视频监控芯片设计商，产品包括安防视频监控多媒体处理芯片和数字接口模块等)、中科曙光 (国内 HPC 龙头，与寒武纪战略合作)、科大讯飞 (智能语音领导者，技术全球领先)、东方网力 (安防行业国内领先的视频监控管理产品与解决方案提供商) 等公司。

❖ 风险提示：半导体行业景气度不及预期；技术创新对传统产业格局的影响

📌 证券研究报告

所属部门 | 股票研究部

报告类别 | 行业深度

所属行业 | 电子

行业评级 | 增持评级

报告时间 | 2017/10/11

📌 分析师

杨欧雯

证书编号：S1100517070002

010-66495651

yangouwen@cczq.com

📌 联系人

王睿

证书编号：S1100117090008

0755-25332321

wangrui@cczq.com

📌 川财研究所

北京 西城区平安里西大街 28 号
中海国际中心 15 楼，
100034

上海 陆家嘴环路 1000 号恒生大厦
11 楼，200120

深圳 福田区福华一路 6 号免税商务大厦 21 层，518000

成都 中国 (四川) 自由贸易试验区成都市高新区交子大道
177 号中海国际中心 B 座 17
楼，610041

正文目录

一、华为发布麒麟 970，人工智能行业热潮再起	4
1.华为发布全球首款内置 NPU 人工智能处理器-麒麟 970	4
2.人工智能热潮再起	5
二、人工智能芯片的崛起之路	6
1.什么是人工智能芯片？为什么需要它？	6
2.现在处于弱人工智能、感知智能的初步阶段	7
3.人工智能已成为各国战略布局的重要一环	8
三、AI 芯片三种技术路线，ASIC 是终端应用的趋势	9
1.GPU:最先被引入深度学习，技术成熟	10
2.FPGA:具有硬件可编程特点，性能出众壁垒高	11
3.ASIC:未来移动端人工智能硬件的方向	12
四、科技巨头加紧布局 AI 芯片，寒武纪跻身国际前列	13
1.NVIDIA:GPU 龙头，AI 芯片的市场领先者	13
2.Intel: 全领域布局人工智能产品	14
3.Google: 发布两代 TPU，从 ASIC 方向进军 AI 芯片市场	16
4.寒武纪：全球 AI 芯片领域第一个独角兽初创公司	17
5.其他 AI 芯片参与企业	19
五、该领域主要龙头企业	19
1.富瀚微：国内领先的视频监控芯片设计商	19
2.中科曙光：国内 HPC 龙头，与寒武纪战略合作	20
3.科大讯飞：智能语音领导者	21
4.东方网力：国内领先的视频监控管理产品与解决方案提供商	21
风险提示	22

图表目录

图 1:	华为麒麟 970 芯片.....	4
图 2:	华为麒麟 970 芯片性能参数.....	4
图 3:	人工智能应用领域.....	5
图 4:	2035 年的实际经济总增加值增速 (%).....	6
图 5:	人工智能芯片的诞生之路.....	7
图 6:	人工智能发展进程.....	7
图 7:	中美人工智能初创企业总量占全球比.....	8
图 8:	中美人工智能产业累计融资额对比 (元).....	8
图 9:	NVIDIA 高性能 GPU-TESLA 概况.....	10
图 10:	CPU 与 GPU 架构对比.....	11
图 11:	XILINX KINTEX 7 ULTRASCLE 芯片和 ALTERA CYCLONE IV 芯片.....	11
图 12:	海康威视搭载 MOVIDIUS MYRIAD 2 处理器.....	12
图 13:	MOBILEYE EYE Q1 芯片.....	12
图 14:	NVIDIA 股价表现.....	13
图 15:	GPU 是 NVIDIA 营收主要来源.....	13
图 16:	NVIDIA 全方位布局人工智能芯片业务.....	14
图 17:	INTEL 收购 ALTERA 后的技术路线图.....	15
图 18:	谷歌第一代 TPU.....	16
图 19:	TPU 内部架构.....	16
图 20:	TPU/CPU 与其他组合对比.....	16
图 21:	谷歌第二代 TPU—CLOUD TPU.....	17
图 22:	CLOUD TPU POD 内含 64 颗 TPU.....	17
图 23:	DIANNAO 结构.....	18
表格 1.	华为麒麟 970 芯片重要技术参数.....	4
表格 2.	世界各国人工智能相关战略、规划.....	9
表格 3.	FPGA & ASIC 对比.....	12
表格 4.	INTEL 人工智能布局.....	15
表格 5.	其他国内外参与 AI 芯片产业的科技公司及初创企业.....	19
表格 6.	富瀚微主营业务收入.....	20

一、华为发布麒麟 970，人工智能行业热潮再起

1. 华为发布全球首款内置 NPU 人工智能处理器-麒麟 970

华为在 IFA2017 大会上公布的麒麟 970 处理器引起了业内关注，它的到来彻底打破了唯频率和核数崇拜的传统观念。

根据其公开的信息，麒麟 970 处理器采用了 10nm 工艺制程（与苹果即将发布用在 iPhone 8 上的 A11 处理器工艺相同），搭载了 Cortex-A73（CPU）、Mali-G72（GPU）和麒麟 NPU（神经网络处理单元）。

表格 1. 华为麒麟 970 芯片重要技术参数

类别	技术指标
工艺制程	台积电 10nm
晶体管数目	55 亿颗晶体管，面积 100 平方毫米
CPU	4 个 A53+4 个 A73
GPU	12 核 Mali G72MP12
通信单元	LTE Cat.18 mit bis zu 1.2 Gbit/s
图像处理	双 ISP 图像处理

资料来源：ofweek 物联网，川财证券研究所

图 1： 华为麒麟 970 芯片



资料来源：华为，川财证券研究所

图 2： 华为麒麟 970 芯片性能参数



资料来源：华为，川财证券研究所

创新在于异构与端云结合。麒麟 970 采用异构架构完成人工智能处理，芯片包括了 CPU（中央处理器）、GPU（图形处理器）、NPU（神经网络处理单

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

元)、ISP (图像信号处理)、DSP (数字信号处理)。NPU 在实现 AI 上承担主要工作,但 CPU、GPU 等也参与处理。未来以异构架构实现人工智能将成为主流,这个异构 SoC 将成为一个 AI 计算平台。

麒麟 970 的另一创新之处在于端云结合实现人工智能,云端智能的优势在于强大的运算能力,基于大数据给出大智慧,在策略性输出上有极大优势。而端侧在具备了强大的感知能力之后,手机就会成为帮助用户思考的分身和伙伴。两者达成平衡与协调,才是未来终端发展的出路。

华为麒麟 970 芯片的发布再次点燃了市场对人工智能领域的关注。

2. 人工智能热潮再起

1946 年,计算机诞生;1956 年人工智能 (AI) 诞生;1969 年互联网诞生……近期无论是华为的麒麟 970 芯片,还是 AlphaGo 与柯洁的人机大战,人工智能领域再度风起云涌。

人工智能正席卷安防、人脸识别、自动驾驶等多个领域。就当前的现状来看,大部分业内人士都认为最先被人工智能“革命”的行业非金融业莫属。曾经,高盛在纽约总部的美国现金股票交易柜台的交易员曾一度高达 600 人,而现在偌大的交易大厅却只有两个人值守。

图 3: 人工智能应用领域



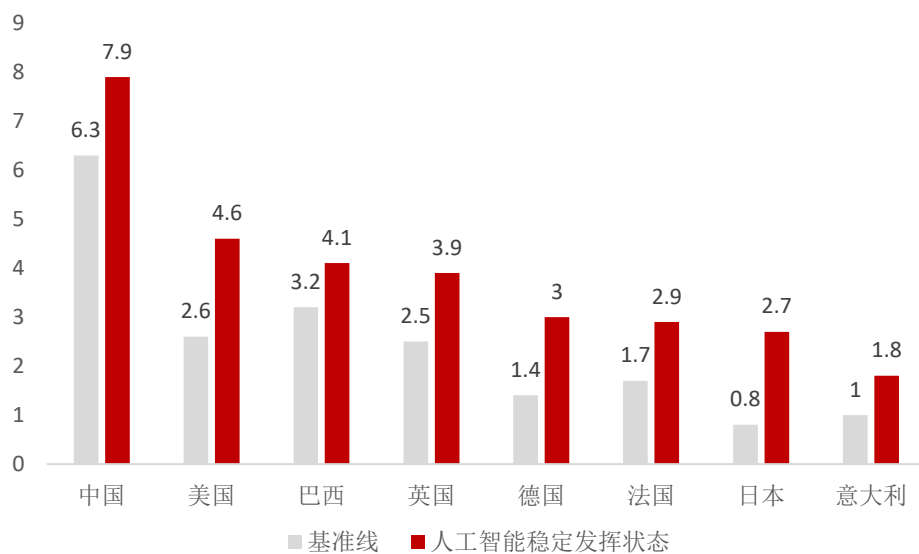
资料来源: 腾讯研究院&IT 桔子《2017 年中美人工智能创投现状与趋势研究报告》, 川财证券研究所

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

根据埃森哲今年发布的报告《人工智能：助力中国经济增长》预测，到 2035 年人工智能有潜力拉动中国经济年增长率上升 1.6 个百分点。人工智能将成为一种全新生产要素，与资本、劳动力拥有同等重要地位，将成为拉动中国经济增长的新动力。

而人工智能芯片作为人工智能的核心要素，已成为各国未来发展布局的战略性一环。

图 4：2035 年的实际经济总增加值增速（%）



资料来源：埃森哲《人工智能：助力中国经济增长》报告，川财证券研究所

二、人工智能芯片的崛起之路

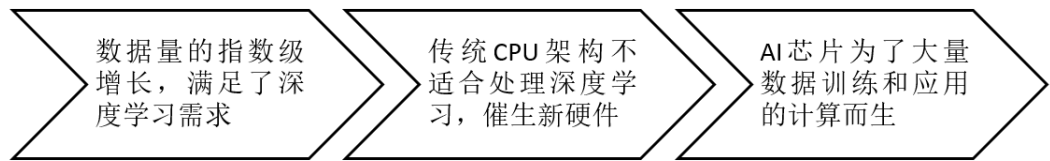
1. 什么是人工智能芯片？为什么需要它？

人工智能的实现依赖三个要素：算法是核心，硬件和数据是基础，芯片就是硬件的最重要组成部分。它其实包括两个计算过程：1、训练（Train）；2、应用（Inference）。

为什么需要人工智能芯片？深度学习作为机器学习的分支，是当前人工智能研究的主流方式。简单说就是用数学方法模拟人脑神经网络，用大量数据训练机器来模拟人脑学习过程，其本质是把传统算法问题转化为数据和计算问题。所以对底层基础芯片的要求也发生了根本性改变：人工智能芯片的设计目的不是为了执行指令，而是为了大量数据训练和应用的计算。

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

图 5：人工智能芯片的诞生之路



资料来源：川财证券研究所

深度学习模型需要通过大量的数据训练才能获得理想的效果。随着大数据时代的到来，数据量呈指数级别增长，这一客观环境满足了深度学习算法对于训练数据量的要求，但是算法的实现还需要相应处理器极高的运算速度作为支撑。

CPU 的架构在解决深度学习任务时效率低、成本高。当前流行的传统 CPU 架构需要数百上千条指令完成一个神经元的处理，但对于不需要太多的程序指令，却需要海量数据运算的深度学习的计算需求，这种结构就显得相对笨拙。

GPU 作为图像处理芯片，其海量数据并行运算的能力与深度学习需求不谋而合，因此被最先引入深度学习。以 GTX 200 (GPU) 为例，在 CPU 上有大约 20% 的晶体管是用作运算之用的，而 GPU 上有 80% 的晶体管用作运算。

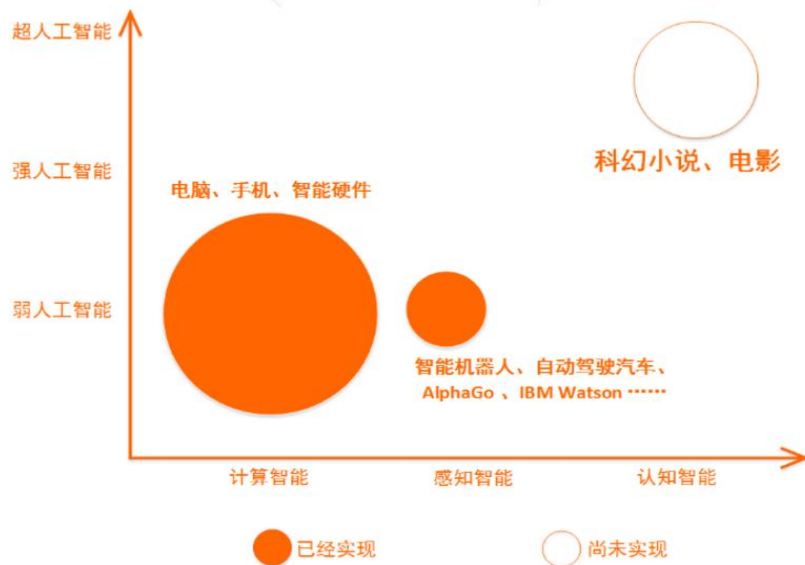
2. 现在处于弱人工智能、感知智能的初步阶段

人工智能的阶段可分为弱人工智能、强人工智能、超人工智能；也可分为计算智能、感知智能、认知智能三个阶段。

计算智能是最初级的，主要是计算能力的进化，这有赖于算法的优化和硬件（CPU 芯片）的技术进步。感知智能有赖于数据获取技术，目前主要有语音识别和机器视觉两种技术。认知智能是最高级的形态，也是未来需要突破的方向。

目前阶段，人工智能处于弱人工智能，刚刚进入感知智能时代。

图 6：人工智能发展进程

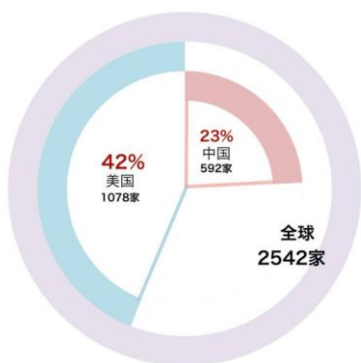


资料来源：腾讯研究院&IT 桔子《2017 年中美人工智能创投现状与趋势研究报告》，川财证券研究所

3. 人工智能已成为各国战略布局的重要一环

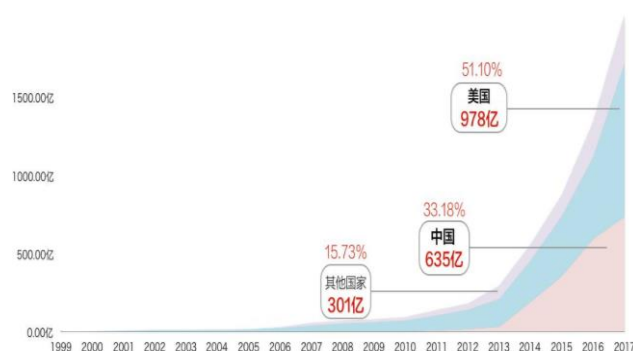
根据腾讯研究院发布的《中美两国人工智能产业发展全面解读》，截止到 2017 年 6 月，全球人工智能企业总数达到 2542 家，美国拥有 1078 家，占据 42%；中国其次，拥有 592 家，占据 23%。其余 872 家企业分布在瑞典、新加坡、日本、英国、澳大利亚、以色列、印度等国家。

图 7：中美人工智能初创企业总量占全球比



资料来源：腾讯研究院《中美两国人工智能产业发展全面解读》，川财证券研究所

图 8：中美人工智能产业累计融资额对比（元）



资料来源：腾讯研究院《中美两国人工智能产业发展全面解读》，川财证券研究所

人工智能是当前人类所面对的最为重要的技术社会变革，是互联网诞生以来的第二次技术社会形态在全球的萌芽，因此也成为各国战略布局的重要一环。

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

2017年7月，国务院印发《新一代人工智能发展规划》，明确了我国发展人工智能的战略目标：到2030年，人工智能核心产业规模超过1万亿元，带动相关产业规模超过10万亿元。其他各国也分别从国家层面对人工智能产业进行战略布局。

表格 2. 世界各国人工智能相关战略、规划

国家	时间	战略、计划	备注
中国	2015.05	《中国制造 2025》	明确提出“加快发展智能制造装备和产品”
	2015.07	《国务院关于积极推进“互联网+”行动指导意见》	明确提出人工智能作为重点布局的 11 个领域之一
	2016.05	《“互联网+”人工智能三年行动实施方案》	到 2018 年，打造人工智能基础资源与创新平台，人工智能产业体系、创新服务体系、标准化体系基本建立
	2017.07	《新一代人工智能发展规划》	到 2030 年，人工智能核心产业规模超过 1 万亿元，带动相关产业规模超过 10 万亿元
美国	2016.10	《为人工智能的未来做好准备》	由美国总统办公室发布
	2016.10	《美国国家人工智能研究与发展策略规划》	
	2016.12	《人工智能、自动化与经济报告》	白宫发布，深入考察人工智能驱动的自动化对经济的影响并提出了国家的三大应对策略
日本	2015.01	新机器人战略	通过发展机器人技术，推动工业生产力的提高
	2015 年	人工智能研究中心	前期投入 10 亿日元
	2015.12	第五个科学与技术基础五年计划	提出名为“超级智能社会”的未来社会构想，发展信息技术、人工智能技术及机器人技术
韩国	2013.05	Exobrain 计划	开发专业领域人机交流的自然语言对话系统
	2014 年	第二个智能机器人总规划（2014-2018）	将机器人产业与其他制造业和服务业相结合，保持在机器人技术及相关重点产业的优势
	2015 年	AI Star Lab	人工智能是项目五大关键领域之一

资料来源：Google，新华社，川财证券研究所整理

三、AI 芯片三种技术路线，ASIC 是终端应用的趋势

目前适合深度学习的人工智能芯片主要有 GPU、FPGA、ASIC 三种技术路线。三类芯片代表分别有英伟达（NVIDIA）的 Tesla 系列 GPU、赛灵思（Xilinx）的 FPGA 和 Google 的 TPU。

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

1. GPU: 最先被引入深度学习，技术成熟

图形处理器 GPU 最初是用在个人电脑、工作站、游戏机和一些移动设备上运行绘图运算工作的微处理器，可以快速地处理图像上的每一个像素点。其海量数据并行运算的能力与深度学习需求不谋而合，因此，被最先引入深度学习。

GPU 不是完全代替 CPU，而是两者分工合作。在 GPU 计算中 CPU 和 GPU 之间是相连的，而且是一个异构的计算环境。这就意味着应用程序当中，顺序执行这一部分的代码是在 CPU 里面进行执行的，而并行的也就是计算密集这一部分是在 GPU 里面进行。

图 9： NVIDIA 高性能 GPU-Tesla 概况

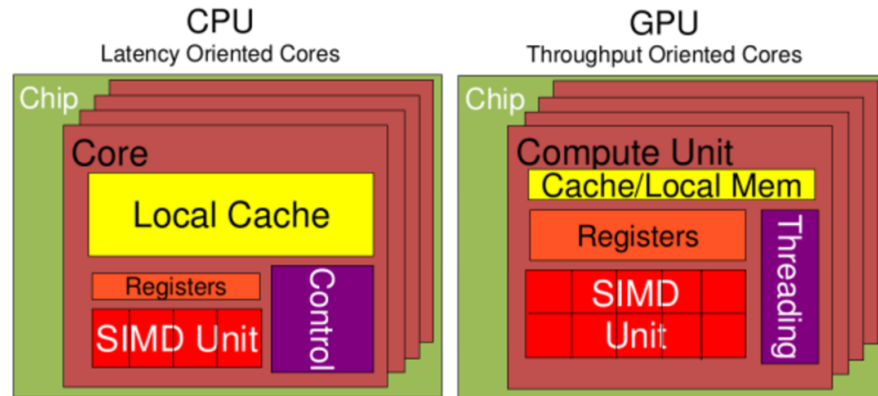


资料来源：NVIDIA，川财证券研究所

GPU 是基于大吞吐量的设计。GPU 的特点是有很多的 ALU 和很少的 cache，这点和 CPU 不同，是为 thread 提高服务的。压缩缓存空间 Cache，从而最大化激发内存吞吐量，可以处理超长的流水线。缓存的目的不是保存之后需要访问的数据，而是担任数据转发的角色，为线程提高服务。

目前，GPU 芯片市场主要由 NVIDIA 把控，占据了 70% 以上的市场份额。而应用在人工智能领域的可进行通用计算的 GPU 市场则基本被英伟达公司垄断。据英伟达官网数据显示，2016 年，有近两万家机构将英伟达产品用于深度学习加速计算，相比 2014 年翻了 13 倍。

图 10: CPU 与 GPU 架构对比



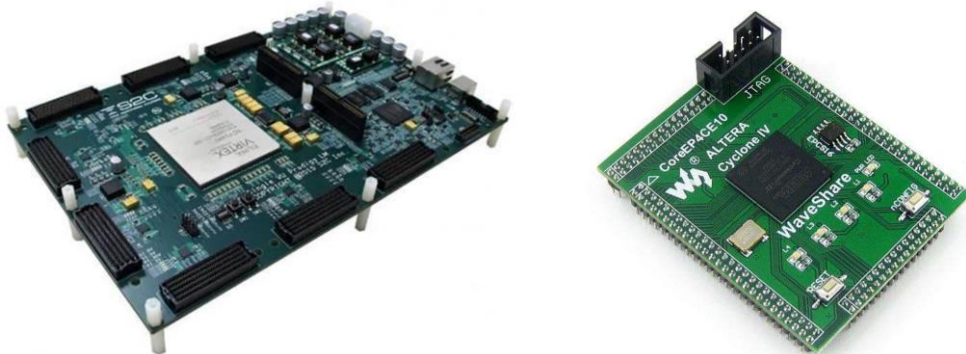
资料来源: NVIDIA, 川财证券研究所

2. FPGA: 具有硬件可编程特点, 性能出众壁垒高

FPGA 即场效可编程逻辑阵列, 一个出厂后的成品 FPGA 的逻辑块和连接可以按照设计者的需要而改变。

FPGA 是可编程的, 而 GPU 由于架构固定, 硬件原生支持的指令也就固定了。其可编程性是关键, 因为它让软件与终端应用公司能够提供与其竞争对手不同的解决方案, 并且能够灵活地针对自己所用的算法修改电路。

图 11: Xilinx Kintex 7 Ultrascale 芯片和 Altera Cyclone IV 芯片



资料来源: Google, 川财证券研究所

FPGA 市场前景诱人, 但是门槛之高在芯片行业里无出其右。IBM、德州仪器、飞利浦、东芝、三星等 60 多家公司先后斥资数十亿美元投入这个行业, 但是最终成功的只有位于美国硅谷的两家公司: Xilinx 与 Altera (被 Intel 收购)。这两家公司共占有近 90% 的市场份额, 专利达到 6000 余项之多, 如此

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

之多的技术专利也构成了极高的技术壁垒。

赛灵思 (Xilinx) 表示, 相较于仅使用 CPU 的服务器, 使用百度 FPGA 云端服务器在赛灵思 FPGA 所驱动的服务器中, 提供 10~80 倍的每瓦效能比。

目标市场方面, FPGA 成本太高, 所以适合对价格不是很敏感的地方, 比如企业应用, 军事和工业电子等等。

3. ASIC: 未来移动端人工智能硬件的方向

ASIC(Application Specific Integrated Circuit)是专用集成电路, 是为专门目的而设计。它是应特定用户要求和特定电子系统的需要而设计、制造的集成电路。ASIC 的特点是面向特定用户的需求, ASIC 在批量生产时与通用集成电路相比具有体积更小、功耗更低、可靠性提高、成本降低等优点。

FPGA 上市速度快, 但性能较低。ASIC 上市速度慢, 需要大量时间开发, 而且一次性成本 (光刻掩模制作成本) 远高于 FPGA, 但是性能远高于 FPGA 且量产后平均成本远低于 FPGA。FPGA 可以完全重配置, 但是 ASIC 也有一定的可配置能力, 只要在设计的时候就把电路做成某些参数可调的即可。

表格 3. FPGA & ASIC 对比

	上市速度	性能	一次性成本	量产成本	可配置	目标市场
FPGA	快	差	低	高	完全	企业、军工
ASIC	慢	好	高	低	有限	消费电子

资料来源: ofweek 电子工程网, 川财证券研究所

目标市场方面, **ASIC** 由于低成本适合消费电子类应用, 是未来移动端人工智能硬件的发展方向。像大疆无人机的 Movidius Myriad 芯片、Tesla 汽车自动驾驶曾用的 Mobileye 芯片和 Google 的张量处理器 TPU 都是 ASIC 的代表。

图 12: 海康威视搭载 Movidius Myriad 2 处理器



资料来源: 海康威视, 川财证券研究所

图 13: Mobileye Eye Q1 芯片



资料来源: 电子工程世界网, 川财证券研究所

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

四、科技巨头加紧布局 AI 芯片，寒武纪跻身国际前列

全球科技巨头都在加紧布局 AI 芯片，希望走在科技变革时代的前线。

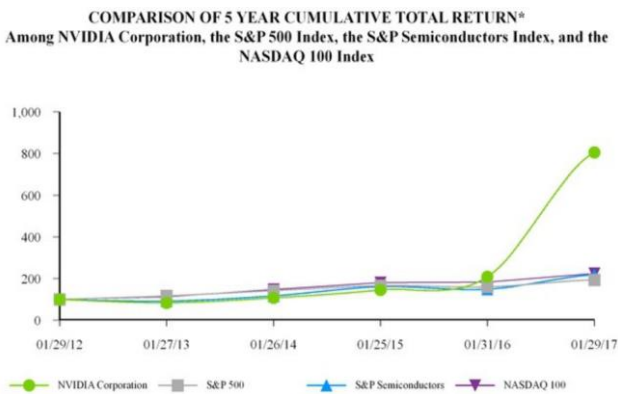
NVIDIA 是 AI 芯片的市场领先者，占据了全球 GPU 70% 的市场份额；Intel 接连收购 Altera、Nervana、Movidius，全方位布局 AI 产品；Google 发布两代 TPU，从 ASIC 方向进军 AI 芯片市场；寒武纪科技是中科院计算所孵化的一家独角兽公司。2016 年推出了国际首个深度学习专用处理器芯片(NPU)，技术全球领先。

1. NVIDIA: GPU 龙头，AI 芯片的市场领先者

目前全球 GPU 行业的市场份额有超过 70% 被英伟达公司占据。而应用在人工智能领域的可进行通用计算的 GPU 市场则基本被英伟达公司垄断。

自 1999 年发布第一款 GPU 以来，GPU 就成为了英伟达最为核心的产品，占到了英伟达总营业收入的八成以上。英伟达的股价表现也是十分惊人，2016 年英伟达的股价上涨了 228%，过去的 5 年内累计上涨 500%。

图 14: NVIDIA 股价表现



资料来源: NVIDIA 2016 年报, 川财证券研究所

图 15: GPU 是 NVIDIA 营收主要来源

	Year Ended			
	January 29, 2017	January 31, 2016	\$ Change	% Change
(\$ in millions)				
GPU	\$ 5,822	\$ 4,187	\$ 1,635	39%
Tegra Processor	824	559	265	47%
All Other	264	264	—	—%
Total	\$ 6,910	\$ 5,010	\$ 1,900	38%

资料来源: NVIDIA 2016 年报, 川财证券研究所

目前英伟达 GPU 芯片主要应用方向为数据中心芯片、自动驾驶芯片和嵌入式芯片。主要包括采用 Pascal 架构的 Tesla P100 和 Tesla P10 芯片、采用 Volta 架构的 DGX-1 芯片、自动驾驶的 Driver PX2 芯片、Jetson TX2 芯片等。

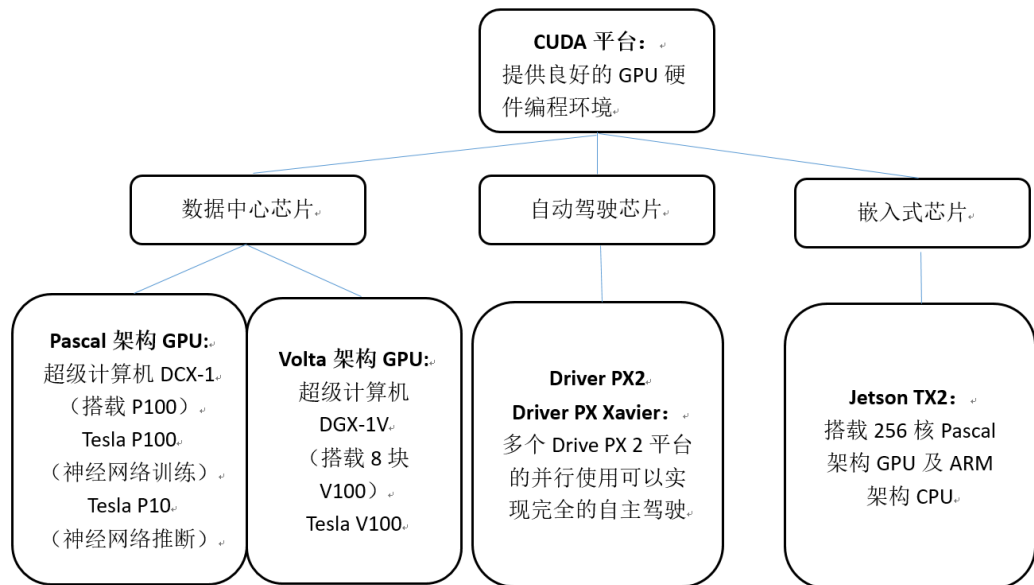
NVIDIA 用于深度学习的 Tesla 芯片经历了 Kepler、Maxwell、Pascal、Volta 四代架构:

(1) **Kepler 架构 (2012 年):** 相较 GPU 第一代 Fermi 架构计算速度高出三

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

- 倍。代表芯片为 K40，被用于深度学习的应用。
- (2) **Maxwell 架构 (2014 年)**: 较 Kepler 架构计算速度提高一倍。代表芯片为 M40。
 - (3) **Pascal 架构 (2016 年)**: 专门针对每瓦性能优化的新架构, 采用 16nm 工艺。代表芯片为用于深度学习训练的 Tesla P100 和神经网络推断的 Tesla P10。
 - (4) **Volta 架构 (2017 年)**: 加入 Tensor (张量) 运算单元, 主要针对 AI 人工智能、DL 深度学习。代表芯片是 Tesla V100。

图 16: NVIDIA 全方位布局人工智能芯片业务



资料来源: NVIDIA, 川财证券研究所

如今英伟达已经不再是一家单纯的显卡技术厂商, 而是一家“人工智能计算公司”。据英伟达官网数据显示, 2016 年, 有近两万家机构将英伟达产品用于深度学习加速计算, 相比 2014 年翻了 13 倍。医疗、生命科学、教育、能源、金融、汽车、制造业以及娱乐业等诸多行业均将得益于海量数据的分析。

2. Intel: 全领域布局人工智能产品

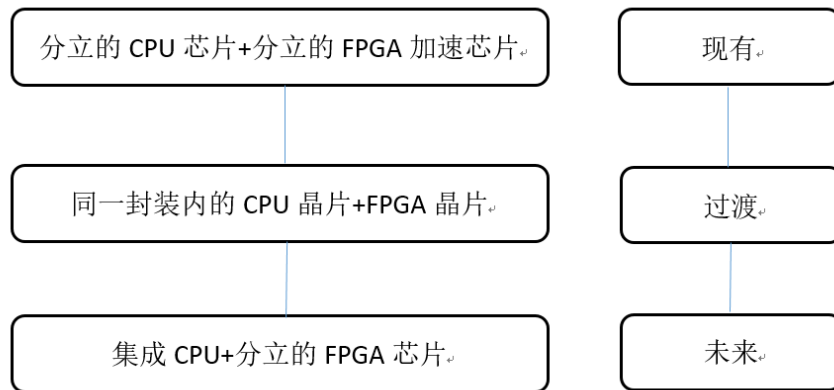
2015 年, 英特尔以 167 亿美元收购了 FPGA 制造商 Altera。英特尔斥巨资

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

收购 Altera 不是来为 FPGA 技术发展做贡献的，而是要让 FPGA 技术为英特尔的发展做贡献。

表现在技术路线图上，就是从现在分立的 CPU 芯片+分立的 FPGA 加速芯片，过渡到同一封装内的 CPU 晶片+FPGA 晶片，到最终的集成 CPU+FPGA 芯片。预计这几种产品形式将会长期共存，因为分立器件虽然性能稍差，但灵活性更高。

图 17: Intel 收购 Altera 后的技术路线图



资料来源: ofweek 电子工程网, 川财证券研究所

2016 年 8 月，Intel 宣布收购创业公司 Nervana Systems。Nervana 准备推出深度学习定制芯片 Nervana Engine，相比 GPU 在训练方面可以提升 10 倍性能。与 Tesla P100 类似，该芯片也利用 16-bit 半精度浮点计算单元和大容量高带宽内存（HBM，计划为 32GB，是竞品 P100 的两倍），摒弃了大量深度学习不需要的通用计算单元。

简单地总结一下 Intel 对于人工智能的布局，大体如下表：

表格 4. Intel 人工智能布局

产品	用途
Xeon Phi + Nervana	用于云端最顶层的高性能计算
Xeon + FPGA	用于云端中间层 / 前端设备的低功耗性能计算
Core (GT)	用于消费级前端设备的性能计算、图形加速
Euclid	提供给开发者 / 创客的开发板，集成 Atom 低功耗处理器、RealSense 摄像头模块、接口，可用做无人机的核心开发部件
Curie	提供给开发者 / 创客的模块，其内置 Quark SE 系统芯片、蓝牙低功耗无线电、以及加速计、陀螺仪等传感器，可用做低功耗可穿戴

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

设备的核心部件

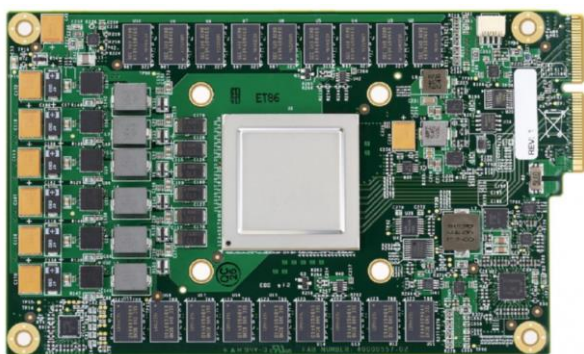
资料来源: Intel, ofweek 电子工程网, 川财证券研究所

此外，2016 年 Intel 还兼并了人工智能芯片厂商 Movidius，提供低能耗计算机视觉芯片组。Movidius 芯片可以应用在可穿戴设备、无人机和机器人中，完成目标识别和深度测量等任务。除了 Google 之外 Movidius 与国内联想和大疆等公司签订了协议。Movidius 的 Myriad 2 系列图形处理器已经被联想用来开发下一代虚拟现实产品。

3. Google: 发布两代 TPU，从 ASIC 方向进军 AI 芯片市场

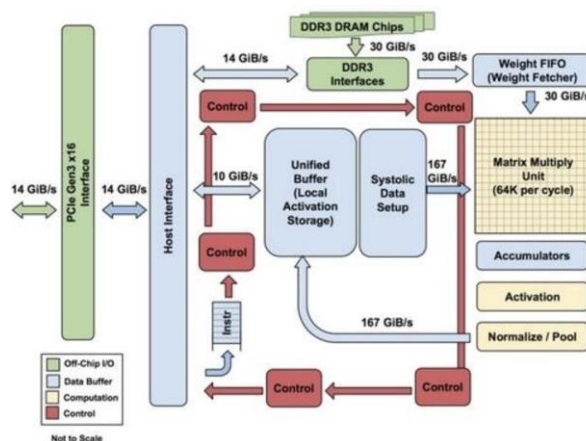
第一代 TPU 展现出 ASIC 在神经网络推断方面的优秀性能。谷歌在 2016 年的 I/O 大会上推出了自己的 AI 芯片——张量处理器 TPU（第一代）。

图 18: 谷歌第一代 TPU



资料来源: Google 官网, 川财证券研究所

图 19: TPU 内部架构



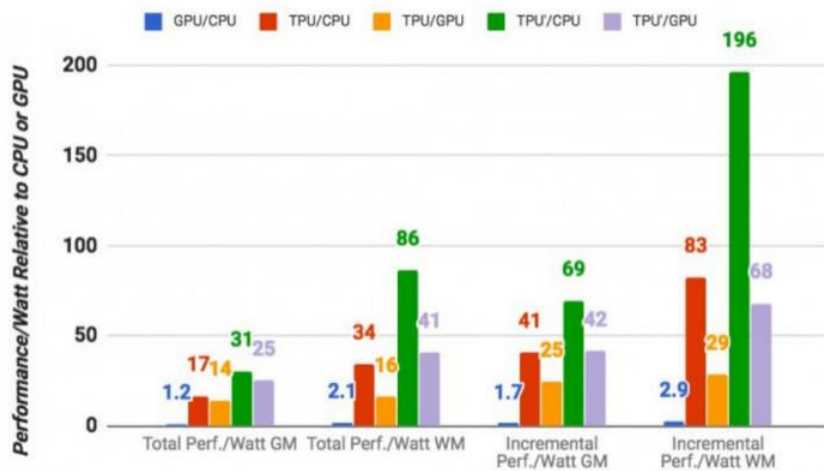
资料来源: Google 官网, 川财证券研究所

谷歌资深硬件工程师 Norman Jouppi 刊文表示，初代 TPU 平均比当时的 GPU 或 CPU 快 15~30 倍（和 TPU 对比的是英特尔 Haswell CPU 以及 Nvidia Tesla K80 GPU），性能功耗比（TOPS/Watt）高出约 30~80 倍。

TPU 是一个推理芯片，它并非是要取代 GPU，可以确定的是，TPU 与 CPU 一起使用对训练分析更加有益。

图 20: TPU/CPU 与其他组合对比

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

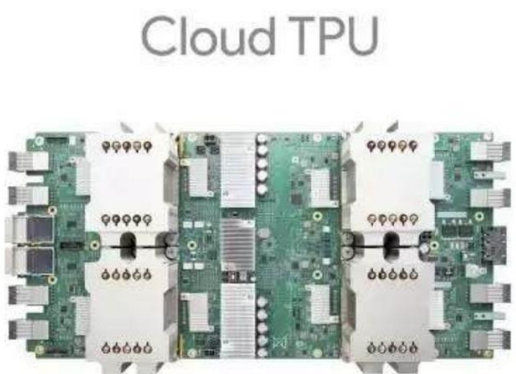


资料来源: ofweek 电子工程网, 川财证券研究所

第二代 TPU 兼具推理+训练。2017 年 5 月 I/O 大会上, 谷歌发布了第二代 TPU—Cloud TPU, 峰值性能达到 180 TFLOPS/s。第一代 TPU 只加速推理, 但第二代 TPU 新增了训练的功能。不仅如此, 谷歌的用户还能通过专门的网络, 在云端利用 TPU 构建机器学习的超级计算机。

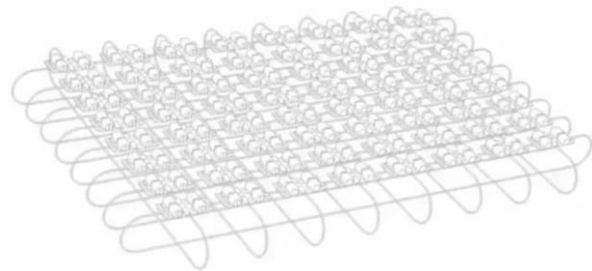
在第二代 TPU 里, 每个 TPU 都包含了一个定制的高速网络, 构成了一个谷歌称之为“TPU 舱室”(TPU POD)的机器学习超级计算机。一个 TPU 舱室包含 64 个第二代 TPU, 最高可提供多达 11.5 千万亿次浮点运算, 内存 400 万兆字节, 4 倍快于当时市面上最好的 32 台 GPU。

图 21: 谷歌第二代 TPU—Cloud TPU



资料来源: Google 官网, 川财证券研究所

图 22: Cloud TPU POD 内含 64 颗 TPU



资料来源: Google 官网, 川财证券研究所

4. 寒武纪: 全球 AI 芯片领域第一个独角兽初创公司

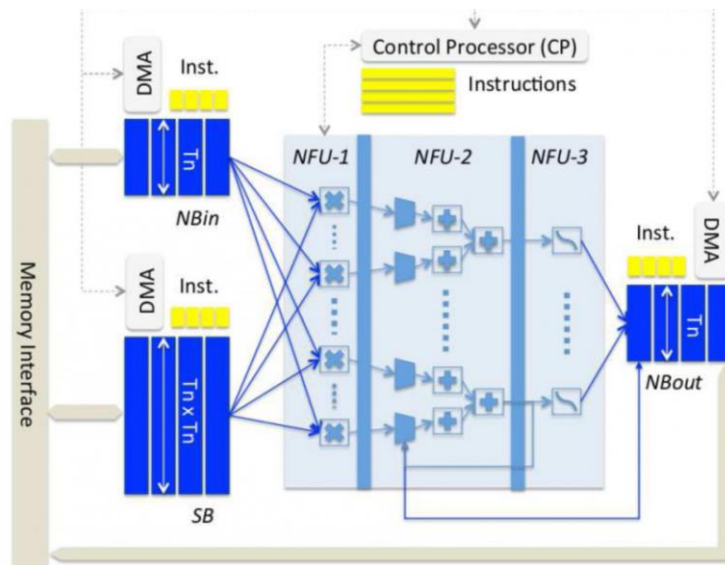
寒武纪科技 (Cambricon) 是中科院计算所孵化的一家独角兽公司。2016

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

年推出的寒武纪 1A 处理器 (Cambricon-1A) 是世界首款商用深度学习专用处理器, 其搭载了国际首个深度学习专用处理器芯片(NPU) (属于 ASIC), 面向智能手机、安防监控、可穿戴设备、无人机和智能驾驶等各类终端设备, 并于 2017 年 8 月获得了包括阿里在内的 1 亿美元 A 轮融资。

CPU、GPU 与 NPU 相比, 会有百倍以上的性能或能耗比差距。在若干代表性神经网络上的实验结果表明——寒武纪的 DianNao 处理器的平均性能超过主流 CPU 核的 100 倍, 但是面积和功耗仅为 1/10, 效能提升可达三个数量级; DianNao 的平均性能与主流 GPGPU 相当, 但面积和功耗仅为主流 GPGPU 百分之一量级。

图 23: DianNao 结构



资料来源: 雷锋网, 川财证券研究所

目前寒武纪主要包括三条产品线:

(1) **智能终端处理器 IP 授权:** 智能 IP 指令集可授权集成到手机、安防、可穿戴设备等终端芯片中, 客户包括国内顶尖 SoC 厂商。

(2) **智能云服务器芯片:** 作为 PCIE 加速卡插在云服务器上, 客户主要是国内的知名服务器厂商。

(3) **家用智能服务机器人芯片:** 从智能玩具、智能助手入手, 使服务机器人独立具备看听说的能力。客户是各类下游机器人厂商, 产品的推出将比智能云服务器芯片更晚一些。

中国正在发力人工智能芯片, 以缩小与全球科技巨头的差距。

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

5. 其他 AI 芯片参与企业

其他国内外参与 AI 芯片产业的科技公司及初创企业如下表。

表格 5. 其他国内外参与 AI 芯片产业的科技公司及初创企业

	公司	简介
国外	微软	将 FPGA 与云计算服务结合，推出 Project Brainwave 低延迟深度学习云平台
	高通	与 Facebook AI 研究所合作研制 AI 芯片 收购 NXP 发展智能驾驶芯片
	IBM	类脑神经芯片 TrueNorth
国内	深鉴科技	着眼于智慧城市与智能数据中心两大市场，通过包括板卡模组、FPGA、编译器、深度压缩等在内的完整解决方案
	地平线机器人	BPU 是地平线机器人自主设计研发高效的人工智能处理器架构 IP，支持 ARM/GPU/FPGA/ASIC 实现。
	西井科技	自主研发的深度学习类脑神经元芯片（deepwell）和可模拟 5000 万级别的“神经元”的类脑神经元芯片深南（deepsouth）产品
	云之声	云知声正加大 AI 可定制化芯片 UniOne 的研发

资料来源：Google，Ofweek 电子工程网，川财证券研究所整理

五、该领域主要龙头企业

AI 芯片是人工智能领域的上游，基于 AI 芯片的下游场景应用极为丰富，包括安防、消费电子、自动驾驶、可穿戴设备等。由于 A 股市场没有像英伟达、Google 这样的 AI 芯片企业，可以从 AI 芯片产业的下游精选个股。相关标的包括富瀚微（国内领先的视频监控芯片设计商，产品包括安防视频监控多媒体处理芯片和数字接口模块等）、中科曙光（国内 HPC 龙头，与寒武纪战略合作）、科大讯飞（智能语音领导者，技术全球领先）、东方网力（安防行业国内领先的视频监控管理产品与解决方案提供商）等公司。

1. 富瀚微：国内领先的视频监控芯片设计商

公司主营业务为数字信号处理芯片的研发和销售，并提供专业技术服务。主要产品为安防视频监控多媒体处理芯片及数字接口模块。

公司主要产品和服务具体情况包括：

- （1）安防视频监控多媒体处理芯片：包括图像信号处理（ISP）芯片、网络摄像机（IPC）SoC 芯片、数字硬盘录像机（DVR）SoC 编解码芯片。

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

- (2) 数字接口模块：公司数字接口模块主要是接受客户委托，为客户定制开发的电路模块。
- (3) 专业技术服务：公司在研发和销售芯片及模块产品的同时，也为客户提供专业技术服务，主要包括：相关算法和 IP 核的开发服务及授权、集成电路设计服务等。

表格 6. 富瀚微主营业务收入

项目	2016 年度		2015 年度	
	金额	比例	金额	比例
1、芯片及模块产品	32168.88	100%	18160.84	100%
其中：视频监控多 媒体处理芯片	29226.90	90.85%	14668.53	80.77%
数字接口模块	2922.87	9.09%	3492.21	19.23%
其他	19.11	0.06%	0.09	0.00%
2、技术服务	0.72	-	-	-
合计	32169.60	100%	18160.84	100%

资料来源：富瀚微招股说明书，川财证券研究所

未来，富瀚微将继续巩固和持续提升公司在模拟摄像机 ISP 市场的占有率，并积极拓展在安防视频监控高清网络摄像机 SoC 芯片、家居安防等市场的份额，拓宽公司营收渠道，实现营业收入、市场占有率及竞争地位的显著提高。

2. 中科曙光：国内 HPC 龙头，与寒武纪战略合作

公司以高端计算机、存储及其相关设备的研发、生产制造为基础，主要向社会输出云计算及大数据综合服务。

国内高端计算机（HPC）龙头：根据 2015-2016 年《中国高性能计算机性能 TOP100 排行榜》显示，公司连续八年获得数量份额第一名。公司在硅立方架构高性能计算机、100Gbps 高速互连网络技术、6D-Torus 高速网络技术、TC4600E-LP 液冷刀片服务器、M-Pro 架构刀片服务器、XSystem 深度学习系统、高性能计算云管理和运维平台等领域开展了大量研发工作，并形成了一系列软硬件产品。在存储方面，根据 IDC 报告，2016 年前三季度公司 NAS 存储产品在中国市场按厂商销售额排名第一位。

与寒武纪战略合作，开拓人工智能领域：2016 年公司与北京中科寒武纪科技有限公司签署战略合作协议，双方将在人工智能领域展开深度合作。这是中

本报告由川财证券有限责任公司编制 谨请参阅尾页的重要声明

科曙光继 2015 年联手中科院计算所与 NVIDIA 共同建立深度学习联合实验室之后，在人工智能领域的又一次重要布局。两家公司同属中科院计算所，有利于发挥协同效应。

3. 科大讯飞：智能语音领导者

科大讯飞在智能语音及人工智能行业深耕十八年，拥有国际领先的源头技术。科大讯飞是我国唯一以语音技术为产业化方向的“国家智能语音高新技术产业化基地”、“国家规划布局内重点软件企业”、“语音及语言信息处理国家工程实验室”。

公司确立起主导产业方向、探索性方向和第三方创业的三层人工智能产业生态体系，实现人工智能技术与多行业、多领域的深度结合推进，市场拓展高歌猛进，讯飞语音深入各行各业，产业链话语权进一步提升。公司已与三大运营商、金融机构、手机、电视、汽车等行业领先企业实现广泛战略合作。讯飞开放平台在线日服务量超 30 亿人次，合作伙伴达到 25 万家，过去 18 个月总用户数（独立终端数量）超 9.1 亿，以科大讯飞为中心的人工智能产业生态持续构建，形成了从源头技术创新到产业技术创新的全生态创新格局，进一步巩固和加强了公司在智能语音及人工智能市场的引导者地位和市场竞争优势。

2017 年 6 月 27 日，《麻省理工科技评论》（MIT Technology Review）全球 50 大最具技术创造力企业榜单发布，科大讯飞首次上榜名列全球第六，在同期上榜的中国公司中位居第一，科大讯飞作为中国人工智能产业领导品牌获得广泛共识，成为国际人工智能竞争格局中的中国力量代表。

4. 东方网力：国内领先的视频监控管理产品与解决方案提供商

公司成立至今主要业务集中于安防行业，是国内领先的视频监控管理产品与解决方案提供商；同时，公司积极布局轨道交通领域，提供系统、专业的轨交领域自动化、信息化系统行业解决方案。

注重技术研发，沉淀核心技术：公司自成立以来一直致力于视频监控管理平台的研究和开发，并紧跟国内外行业发展趋势，在视频联网、视频结构化、视频大数据等领域拥有众多核心技术积淀，并投入资源开展人工智能技术的基础研究，结合安防行业需求研发产品应用，“视云天下”系列视频监控管理平台产品得到行业市场的高度认可。2016 年公司研发人员人数占公司全体员工总数的 44.09%，共计研发投入 2 亿元，占公司营业收入比例为 13.63%。2016 年，公司及子公司新增专利申请 64 个，其中发明专利 56 个，实用新

型专利 3 个，外观设计专利 5 个；新增获得软件著作权 58 个。

产品优势明显，围绕轨道交通自动化、信息化系统打造产品地图：公司“视云天下”系列视频监控管理应用产品体系包括视频图像联网共享解决方案、视云数据中心解决方案、视图解析综合应用解决方案、公安大数据解决方案、视频侦查解决方案、智能交通解决方案、物理安防解决方案、人证核验解决方案等。公司根据技术发展路径和行业需求的变化，持续加大自主研发力度，新形成了警用图像处理系统、人脸系统、人群系统、集群云存储、PVG+、公安大数据平台等多个产品，并实现销售落地。公司在轨道交通业务中，围绕轨道交通自动化、信息化系统打造产品地图，在保持除车载系列产品优势同时，还设计开发了多个地铁地面系统产品和自动化新产品，包括地面 PIS、地面 CCTV、现代有轨电车自动售票系统、弓网视频监控系统、车载 WIFI 系统、门控器、便携式摄像记录仪、人证闸机等产品。

风险提示

半导体行业景气度不及预期

半导体行业景气度下滑，可能会带来电子全行业需求疲软

技术创新对传统产业格局的影响

技术创新可能会使得原有的产业格局发生变化，行业龙头替换甚至淘汰个别传统行业

分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉尽责的职业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告。本人薪酬的任何部分过去不曾与、现在不与、未来也不会与本报告中的具体推荐意见或观点直接或间接相关。

行业公司评级

证券投资评级：以研究员预测的报告发布之日起6个月内证券的绝对收益为分类标准。30%以上为买入评级；15%-30%为增持评级；-15%-15%为中性评级；-15%以下为减持评级。

行业投资评级：以研究员预测的报告发布之日起6个月内行业相对市场基准指数的收益为分类标准。30%以上为买入评级；15%-30%为增持评级；-15%-15%为中性评级；-15%以下为减持评级。

重要声明

本报告由川财证券有限责任公司（已具备中国证监会批复的证券投资咨询业务资格）制作。本报告仅供川财证券有限责任公司（以下简称“本公司”）客户使用。本公司不因接收人收到本报告而视其为客户，与本公司无业务关系的阅读者不是本公司客户，本公司不承担适当性职责。本报告在未经本公司公开披露或者同意披露前，系本公司机密材料，如非本公司客户接收到本报告，请及时退回并删除，并予以保密。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司对该等信息的真实性、准确性及完整性不作任何保证。本报告所载的意见、评估及预测仅为本报告出具日的观点和判断，该等意见、评估及预测无需通知即可随时更改。在不同时期，本公司可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。本公司不保证本报告所含信息保持在最新状态。对于本公司其他专业人士（包括但不限于销售人员、交易人员）根据不同假设、研究方法、即时动态信息及市场表现，发表的与本报告不一致的分析评论或交易观点，本公司没有义务向本报告所有接收者进行更新。本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供投资者参考之用，并非作为购买或出售证券或其他投资标的的邀请或保证。该等观点、建议并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。根据本公司《产品或服务风险等级评估管理办法》，上市公司价值相关研究报告风险等级为中低风险，宏观政策分析报告、行业研究分析报告、其他报告风险等级为低风险。本公司特此提示，投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素，必要时可就法律、商业、财务、税收等方面咨询专业财务顾问的意见。本公司以往相关研究报告预测与分析的准确，也不预示与担保本报告及本公司今后相关研究报告的表现。对依据或者使用本报告及本公司其他相关研究报告所造成的一切后果，本公司及作者不承担任何法律责任。

本公司及作者在自身所知情范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。投资者应当充分考虑到本公司及作者可能存在影响本报告观点客观性的潜在利益冲突。在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，也可能为之提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。本公司的投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

对于本报告可能附带的其它网站地址或超级链接，本公司不对其内容负责，链接内容不构成本报告的任何部分，仅为方便客户查阅所用，浏览这些网站可能产生的费用和风险由使用者自行承担。

本公司关于本报告的提示（包括但不限于本公司工作人员通过电话、短信、邮件、微信、微博、博客、QQ、视频网站、百度官方贴吧、论坛、BBS）仅为研究观点的简要沟通，投资者对本报告的参考使用须以本报告的完整版本为准。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许范围内使用，并注明出处为“川财证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。如未经川财证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

本提示在任何情况下均不能取代您的投资判断，不会降低相关产品或服务的固有风险，既不构成本公司及相关从业人员对您投资本金不受损失的任何保证，也不构成本公司及相关从业人员对您投资收益的任何保证，与金融产品或服务相关的投资风险、履约责任以及费用等将由您自行承担。

本公司具有中国证监会核准的“证券投资咨询”业务资格，经营许可证编号为：11080000

本报告由川财证券有限责任公司编制 谨请参阅本页的重要声明 C0001