

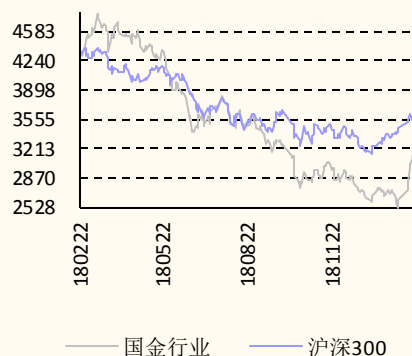
创新技术与企业服务研究中心

半导体/电子/通信行业 买入（维持评级）

行业深度研究

市场数据(人民币)

市场优化平均市盈率	18.90
国金半导体指数	3252.34
沪深 300 指数	3439.61
上证指数	2755.65
深证成指	8440.87
中小板综指	8379.04



相关报告

- 1.《全年衰退，最糟已过，进入另一波上行周期-全年衰退，最糟已过，...》，2019.2.18
- 2.《5G 东风已至，VR/AR 乘势起航-5G 东风已至，VR/AR 乘...》，2019.2.15
- 3.《芯片封测行业 2018 年业绩遇冷，2H19 下半年恢复可期-封测...》，2019.2.1
- 4.《【半导体周报】下半年半导体有望复苏，可折叠手机带动 OLED-...》，2019.1.31
- 5.《【国金计算机周报】政策的快速推动和实质性改善，板块吸引力和发...》，2018.4.1

樊志远 分析师 SAC 执业编号: S1130518070003
(8621)61038318
fanzhizhuan@gjzq.com.cn

罗露 联系人
luolu@gjzq.com.cn

人工智能无所不在

投资建议

- **人工智能是未来无所不在的工具：**人工智能平台（包括芯片，模组，算法，云训练/推理）不算新型应用，而是整合半导体，通信，软件及云/边缘运算/设备电子后成为各种提升物联网应用效能的人工智能工具平台，这就像我们常用的微软 Office 软件。云端芯片巨头英伟达早在 2016 年，公司就累计了 7 大应用领域及 19,439 客户使用其深度学习的服务工具，配合半导体，算法和之前在云端大数据的深度学习训练和推断的数据库，来执行更佳智能推理。
- **投资建议：**国内重点关注公司华为海思（半导体设计），寒武纪（设计），海康威视（安防 AI 系统），商汤（算法软件），伊图（算法软件）

行业观点

- **AI 半导体 10 倍数增长可期一半导体篇：**目前人工智能芯片仍多是以 GPU，张量处理器，或 FPGA+CPU 为主，但未来 ASIC 将在边缘运算及设备端遍地开花，及逐步渗透云端市场，预估全球 AI 云端半导体市场于 2018-2025 年 CAGR 应有 37%，边缘运算及设备端半导体市场于 2018-2025 年 CAGR 应有 249%，远超过全球半导体市场在同时间 CAGR 的 5%，占整体份额从 2018 年的 1% 到 2025 年的 10%，超过 10 倍数增长可期。
- **谁能引领国内人工智能芯片产业突围？**有今年中国科创板融资平台的加持，国内的半导体公司将陆续推出人工智能 ASIC 抢先机。而华为海思因为强大的财力来开发 10 纳米及以下产品，负担 EDA 软件，验证，光掩膜成本的蹿高，加上强大的设计团队及对系统的认知，预计将引领国内 AI 芯片行业突围；而比特大陆因为深谙 IP，芯片，模块，到生态系的竞争，不排除其反而率先推出有竞争力的云端人工智能的解决方案模块；寒武纪目前有较佳的设计团队及较充裕的估值融资能力来陆续推出边缘运算端及云端推理的人工智能芯片。
- **AI 让电子行业插上腾飞的翅膀—电子篇：**近年来，AI 技术得以大幅度提升，应用场景也越来越多。2017 年人工智能市场中，计算机视觉位居第一，占比达到 37%，语音第二，达到 22%，而在计算机视觉应用中，安防占比高达 68%，基于计算机视觉开发的人脸解锁、人像美颜功能在 AI 手机应用中渗透率分别高达 75% 与 90%。我们看好 AI 未来在安防、智能手机、可穿戴设备、智能家居等方面的应用爆发，优质龙头公司为海康威视及大华股份。
- **政策，技术，资本三轮驱动—系统平台战略篇：**全球主要国家均将 AI 作为未来主导性战略，在国家竞争版图中，中国在 AI 领域强在应用层，基础和技术层是相对短板。AI 将重塑 ICT 产业生态，芯片和软件平台是产业制高点，但从应用看，AI 在安防、自动驾驶等重点场景应用加速，云化是 AI 发展前提，计算机视觉、智能语音是 AI 市场两大应用方向，但 AI 视觉将最具商业化价值赛道。行业已进入群雄逐鹿阶段，头部算法企业如商汤、伊图横向建立平台，垂直领域龙头如海康、大华纵向深耕行业，科技巨头如阿里、华为立体式打造全产业链生态。

风险提示

- 除了国内算法软件公司及美国 AI 芯片大厂外，未来 24 个月，我们看不出国内 AI 芯片大厂能摆脱亏损；美国商务部工业安全局可能对其 11 项 AI 和机器学习技术列入出口管制清单；AI ASIC 大厂对各种应用系统认知不足，安防 / 语音 / 自驾系统公司不愿分享其系统设计机密。

内容目录

一、 人工智能无所不在一半导体篇	6
1、人工智能平台到底是工具还是应用?	6
2、人工智能会渗入各领域应用一无所不在	7
3、三种主流人工智能演算法	8
4、到底哪种人工智能芯片将成云计算的主流?	11
5、哪种人工智能芯片将成边缘运算及设备端的主流?	17
6、谁能引领国内人工智能芯片产业突围?	20
二、 AI 让行业应用插上腾飞的翅膀一电子应用篇	27
1、AI 让安防升级换代，催生更多应用场景	27
A、稳健发展中的安防产业	27
B、AI 技术助力安防行业智能化升级	27
C、智能安防才刚进入初级阶段，AI 占比低于 1%，未来大有可为	32
D、安防 AI 面临的三大挑战	34
E、AI+安防使用技术	35
F、AI+安防场景应用	36
G、2018 年安防 AI 发展情况	37
H、AI+安防发展趋势	38
2、智能手机 AI 应用势如破竹	39
A、AI 视觉给手机带来新方向	39
B、解放双手，AI 语音助手发展在路上	41
3、智能穿戴“钱”途无量	42
A、可穿戴设备保持快速成长	42
B、可穿戴 AI 市场增速更快	42
C、蓝牙耳机有望会成为私人“AI 小秘”	43
4、安防+AI 让智能家居正式迈入智慧家居	43
A、智能音箱，智能家居设备的入口	44
B、家庭智能安防渐成“刚需”	44
三、 AI 信息服务的蓝海已至一系统平台战略篇	45
1、政策、技术、资本三轮驱动	45
A、全球主要国家均将人工智能作为未来主导性战略，美中领跑	45
B、技术变革促使 AI 突破应用瓶颈，推动产业应用爆发	46
C、资本涌入加速产业成熟，AI+、视觉、大数据等领域成为投资热点	47
D、人工智能重塑产业生态，芯片和软件平台是产业制高点	48
2、AI 是全新的生产要素及工具，AI+行业将成为普遍商业模式	49
A、高数字化行业将优先实现人工智能的应用	49

B、安防、自动驾驶是当前和未来人工智能重点应用领域.....	50
3、计算机视觉、智能语音是 AI 时代入口，中国初创龙头企业具备全球竞争力	51
A、计算机视觉：AI 行业最具商业化价值的赛道	51
B、中短期核心竞争壁垒是技术和产品能力，长期是生态构建能力	54
C、头部算法企业横向建立平台，垂直领域龙头纵向深耕行业，科技巨头立体式打造全产业链生态。	55
D、智能语音：产业进入爆发期，看好对话式人工智能发展	55
4. 领先初创企业介绍：商汤、旷视、依图、云从、云知声	57
A、商汤：AI 算法提供商龙头，平台化战略赋能公司和行业	57
B、旷视：优秀的人脸识别平台企业，战略进军 IoT	59
C、依图：注重实战，业务聚焦于安防和医疗	59
D、云从：AI 国家队，银行、机场人脸识别产品第一大供应商	60
E、云知声：“云端芯”战略，重点布局家居、车载及医疗三大领域	61
四、投资建议	61
五、风险提示	62
1、中美贸易及关税战是否趋缓？	62
2、国内 AI 芯片大厂能否摆脱亏损？	62
3、禁售人工智能新兴技术？	62
4、系统设计机密不愿分享：	62

图表目录

图表 1：人工智能的多样性.....	6
图表 2：人工智能云，边缘运算，设备端半导体及行业市场营收预估.....	6
图表 3：人工智能技术工艺的演化.....	8
图表 4：卷积输入及输出特征贴图及最大池.....	9
图表 5：循环神经机器翻译.....	10
图表 6：深度神经网络.....	11
图表 7：深度学习	11
图表 8：各种人工智能半导体优缺点比较.....	12
图表 9：人工智能云端系统图形处理芯片面积.....	12
图表 10：人工智能半导体市场预测以不同芯片种类来分类.....	13
图表 11：英伟达云端人工智能芯片 V100 及系统 DGX-1 规格比较表.....	14
图表 12：赛灵思 / BlackLynx 与 GPU 在机器学习推理解决方案的比较.....	15
图表 13：谷歌张量处理器 TPU 3 vs. TPU 2	16
图表 14：各公司云端人工智能芯片比较表.....	17
图表 15：谷歌的智能推断边缘运算端的解决方案模块 (System-on-module, SOM)	18

图表 16: 谷歌的智能物联网整体解决方案.....	18
图表 17: 智能推断边缘运算端的解决方案模块比较表.....	19
图表 18: 人工智能深度学习从云到设备端.....	20
图表 19: 集成电路设计成本飙升.....	21
图表 20: 华为云, 边缘运算, 及设备端的芯片解决方案.....	22
图表 21: 寒武纪 16nm MLU100 云端 AI 芯片.....	23
图表 22: 地平线 Matrix 1.0 SAE L3/L4 无人驾驶平台解决方案模块.....	24
图表 23: 比特大陆损益表.....	25
图表 24: 比特大陆挖矿机芯片制程工艺.....	25
图表 25: 设备端人工智能芯片每瓦算力功耗比.....	25
图表 26: 中国安防产业稳健成长.....	27
图表 27: AI 融入安防产业链.....	28
图表 28: 智能安防云边结合示意图.....	28
图表 29: 海康威视 AI Cloud 核心理念.....	29
图表 30: 智能安防产业链.....	30
图表 31: AI+安防产业图谱.....	30
图表 32: 人工智能安防涉及的主要算法.....	31
图表 33: 安防监控四类主要芯片.....	31
图表 34: GPU 在智能安防的应用.....	32
图表 35: AI+安防典型落地应用产品.....	32
图表 36: 2018 年中国 AI 摄像机建设逻辑与发展概况.....	33
图表 37: 中国智慧安防市场规模预测.....	33
图表 38: 2017 年安防厂商视频监控业务收入占比.....	34
图表 39: 2018 年安防厂商视频监控业务收入占比.....	34
图表 40: AI 技术在手机应用中的渗透率.....	40
图表 41: 计算机视觉应用全景图.....	40
图表 42: OPPO R17 Pro 的 AR 测量.....	41
图表 43: 17-22 年各类可穿戴设备出货量.....	42
图表 44: 17-22 年各类可穿戴设备出货量占比.....	42
图表 45: 全球可穿戴 AI 市场预测.....	43
图表 46: AirPods 出货量预测.....	43
图表 47: 全球智能家居发展情况.....	44
图表 48: 中美各 AI 板块企业数量分布.....	45
图表 49: 中美各 AI 板块企业融资金额分布 (亿元).....	45
图表 50: 中国在人工智能领域的专利数呈现爆发式增长态势.....	46
图表 51: 人工智能技术不断突破.....	46
图表 52: 2018 年 Gartner 人工智能技术成熟度曲线.....	47
图表 53: 全球人工智能投融资笔数分布 (2013-2018Q1).....	47
图表 54: 中国 AI 企业融资规模持续走高.....	48

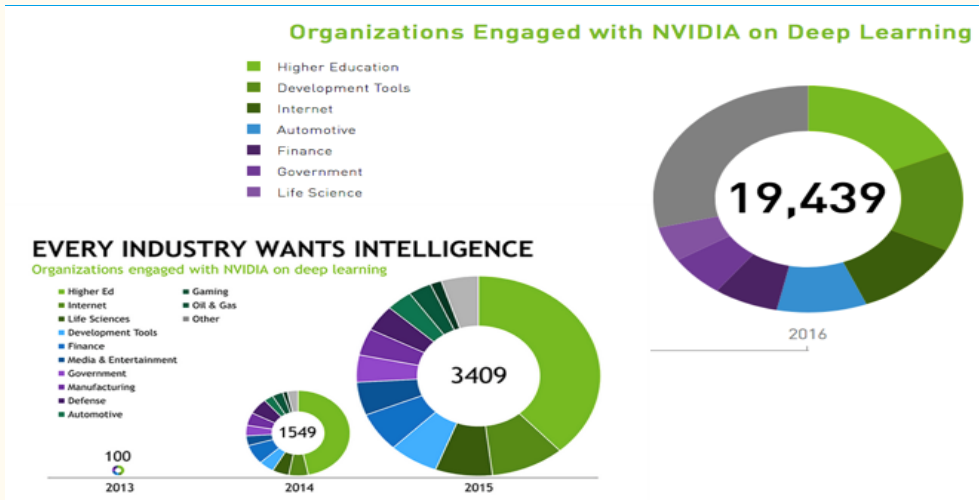
图表 55: 中国 AI 各领域融资金额分布.....	48
图表 56: ICT 产业正在跨界和重构，人工智能芯片成为新战略控制点.....	48
图表 57: 全球 ICT 巨头 AI 布局一览.....	49
图表 58: AI 为行业带来增值.....	50
图表 59: AI 为行业带来利润提升.....	50
图表 60: 不同行业 AI 技术应用程度.....	50
图表 61: 自动驾驶汽车产生海量数据.....	51
图表 62: 华为 MDC 600	51
图表 63: 全球 AI 企业应用技术方向分布	52
图表 64: 中国 AI 企业应用技术方向分布	52
图表 65: 计算机视觉架构及各层级玩家示意.....	53
图表 66: 2017 年中国计算机视觉应用市场份额.....	53
图表 67: 2018 年全球计算机视觉行业市场结构.....	54
图表 68: AutoML 工作原理（红色部分自动完成）	54
图表 69: AutoML 内部工作机制	54
图表 70: 近几年语音识别准确率变化.....	56
图表 71: 头部智能语音初创企业首款专用芯片发布梳理.....	56
图表 72: 对话式人工智能在智能家居、随身设备、智能车载的渗透率.....	57
图表 73: 商汤业务布局梳理.....	58
图表 74: 商汤近年对外重要投资梳理.....	58
图表 75: 2015-2017 三大顶级会议论文收录数.....	58
图表 76: 算法和算力支撑公司应用拓展.....	58
图表 77: 旷视对外投资梳理.....	59
图表 78: FRVT(2018)比赛结果（部分）	60

一、人工智能无所不在一半导体篇

1、人工智能平台到底是工具还是应用？

人工智能平台（包括芯片，模组，软件）在一般人看起来像是一种新型应用，但在我们看来人工智能芯片在整合软硬件后将成为各种物联网应用的提升效能工具平台，这就像我们常用的微软 Office 软件，微软 Office 软件是在办公室应付各种应用的生财工具，因此人工智能平台除了被广泛利用在云端大数据的深度学习训练和推断外，我们认为人工智能平台也将出现在各式各样的应用端的边缘设备，从英伟达公布的数字来看，早在 2016 年，公司就累计了 7 大领域（高等教育，发展工具，互联网，自驾车，金融，政府，生命科学）及 19,439 客户使用其深度学习的工具，配合软件和之前在云端大数据的深度学习训练和推断的数据成果库，来达到帮助使用者或取代使用者来执行更佳智能判断推理。

图表 1：人工智能的多样性



来源：英伟达，国金证券研究所

虽然目前人工智能芯片仍多是以昂贵的图形处理器 (GPU)，张量处理器 (Tensor Processing Unit)，或现场可编程门阵列芯片配合中央处理器 (FPGA+CPU) 为主，来用在云端的深度学习训练和推理的数据中心，但未来特定用途集成电路 (ASIC) 将在边缘运算及设备端所需推理及训练设备遍地开花，及逐步渗透部分云端市场，成为人工智能芯片未来的成长动能，我们预估全球人工智能云端半导体市场于 2018-2025 年复合成长率应有 37%，边缘运算及设备端半导体市场于 2018-2025 年复合成长率应有 249% (请参考图表)，远超过全球半导体市场在同时间的复合成长率的 5%，整体约占全球半导体市场的份额从 2018 年的 1% 到 2025 年的 10%。

图表 2：人工智能云，边缘运算，设备端半导体及行业市场营收预估

	2017	2018E	2019E	2020E	2021E	2022E	2023E	2024E	2025E	CAGR
全球半导体市场 (US\$bn)	412	478	454	477	534	571	588	647	660	5%
全球半导体市场 (同比)	22%	16%	-5%	5%	12%	7%	3%	10%	2%	
AI 半导体 (US\$bn)	2	5	10	17	26	38	48	59	69	46%
AI 半导体 (同比)	180%	129%	108%	70%	53%	46%	26%	22%	18%	
AI IC 佔全球 IC 份额 (%)	1%	1%	2%	4%	5%	7%	8%	9%	10%	
云端 AI 半导体 (US\$bn)	2	5	8	12	18	24	30	37	44	37%
云端 AI 半导体 (同比)	180%	129%	69%	52%	43%	34%	28%	23%	19%	
边缘及设备端 IC (US\$bn)		0	2	5	8	15	18	22	25	249%
边缘及设备端 AI IC (同比)			190%	147%	79%	73%	24%	19%	16%	

AI 行业销售额 (US\$bn)	104	143	192	261	361	490	657	854	1,093	34%
AI 行业销售额 (同比)	76%	37%	34%	36%	38%	36%	34%	30%	28%	

来源: Statista, Tractica, Frost & Sullivan, 国金证券研究所

2、人工智能会渗入各领域应用一无所不在

当大多数的产业研究机构把自驾车(Autonomous drive vehicle), 虚拟/扩增实境(Virtual Reality/Augmented Reality), 无人商店 (Unmanned store), 安防智能监控 (Smart Surveillance System), 智能医疗, 智能城市, 和智能物联网(Internet of Things, IoT) 分别当作半导体产业不同的驱动引擎, 国金半导体研究团队认为其实自驾/电动车, 5G, 虚拟/扩增实境, 无人商店, 安防智能监控, 智能医疗, 智能城市其实都只是人工智能物联网的延伸。

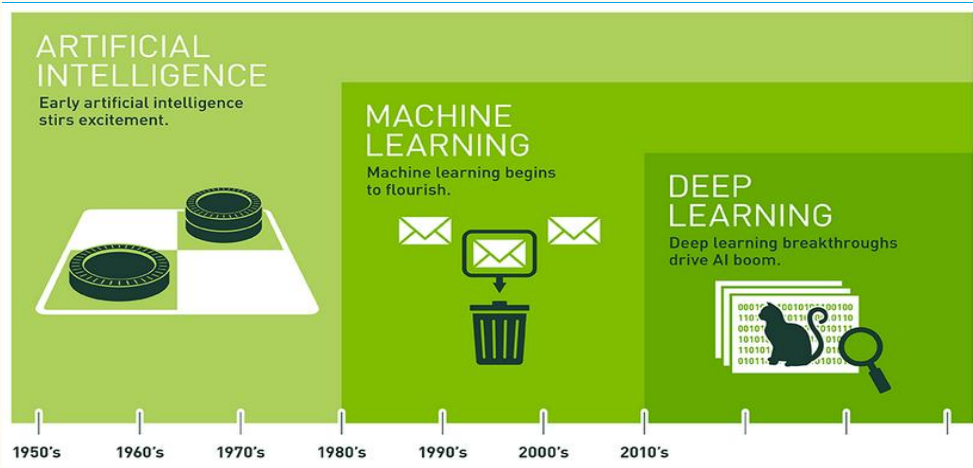
- **无人驾驶/电动/联网车对半导体的需求爆增 10 倍:** 虽然全球汽车半导体产业, 目前由传统的整合元件制造商 (IDM) 掌控, 但未来很难说, 主要是因为先进辅助或自动驾驶系统兴起, 人工智能, 摄像, 传感, 雷达芯片公司的出现。像英伟达图形处理器 (Xavier, Pegasus 320), 谷歌张量处理器 (Tensor Processing Unit), 英特尔 / Altera/Mobileye 的 CPU/FPGA/AI 解决方案, 地平线 L3/L4 自动驾驶 (Matrix 1.0 平台, 征程 2.0 芯片), 高通, 联发科, 华为 / 海思的 5G 无线通信平台及毫米波雷达, 索尼, 豪威的传感器, 博通及瑞昱的以太网芯片。尤其是自动驾驶对雷达, 摄像头, 传感器及芯片的 3 倍增幅; 千倍的 3D 视频数据的上传及云端的存储, 学习, 推理; 因无人驾驶车设备及云端建制成本高昂 (US\$150,000 / Waymo 车), 额外成本必须由广大消费者共同分摊费用较有利, 乘坐共享, 公交服务业, 产品运送的后勤支援业无人驾驶应会领先乘用车市场, 而 Waymo/谷歌将带动出租车 / 公交车自动驾驶市场, 领先英伟达的自驾乘用车市场 (请参阅国金 **电动, 无人驾驶, 车联网的三部曲驱动力** 的深度报告), 依照美国加州 DMV (Department of Motor Vehicles) 最新公布的数据显示 Waymo 于去年测试的 120 万英里中, 每 1000 英里发生解除自驾系统状况频率是 0.09 次, 远低于前年的 0.179 次, 及通用 Cruise 的 0.19 次, 苹果的 872 次, 及 Uber 的 2860 次; 自动泊车、车道偏离检测、无人驾驶的带宽需求, 及车内电线费用和重量的不断增加。为了让增加数倍的电子控制单元 (ECU) 能彼此间快速地沟通, 数倍的以太网路节点和转换器芯片需求便随之而来。
- **无人商店及安防智能监控:** 除了自动驾驶车联网外, 最近风起云涌的无人商店和智能监控, 不也是利用大量监控摄像头, 配合三维人脸辨识系统, 二维码/近场通讯来收集大数据资料, 再透过 WiFi/xDSL/光纤传输, 将资料送到云端人工智能储存与处理来达到无需柜台人员的无人商店和能随时辨识的视频智能监控, 而政府机关可透过此系统来调查人口移动来重新设计城市智能公共交通系统, 协寻通缉犯, 恐怖分子, 失踪人口, 及在展场, 车站, 机场, 学校, 大型活动场地的安全监护; 系统整合业者除可做无人商店外, 也可靠着人口动向预测消费热点 (商家必备), 人口居住热点 (房地产业必备), 如果再配合无人商店, 线上购物系统, 和政府的大数据, 系统业者便可更精准的投送广告, 发展个人信用评级。类似于自驾车联网, 无人店和监控联网系统需要大量并且高清晰度的三维辨识摄像头和芯片, 传输系统和芯片, 和庞大的云端, 边缘运算, 及设备端的储存及智能训练及推理的各式高速芯片及软件。
- **智能医疗:** 智能医疗系统可利用三维脸部个人辨识来挂号, 减少排队时间, 让看诊更顺畅。医院可收集资料并整合个人在不同医疗院所的所有医疗纪录; 医生可利用人工智能数据平台辅助做更精准的医疗判断, 减少重复用药的浪费和对病人的副作用, 医院可利用此大数据资料做更深入的医学研究, 数家政府医院应先抛砖引玉, 带头做整合。
- **智能物联网: 1. 空污, 水污染化学感测物联网:** 政府是否应利用强制安装并定期检验各式气体/液体的物联网化学感测器在每部汽、机车和工厂排污管道上, 再透过大量的低耗能无线通讯将资料上传到云端储存和处理, 并透过人工智能来监测空污, 水污来收取空污或水污税; **2. 身份识别证明联网:** 大型互联网企业像谷歌 (Google US), 百度 (Baidu US), 腾讯 (00700.HK), 或海康威视 (002415.SZ) 为何不能发展三维脸部个人辨识智能系统来整合所有的线上线下购物, 启动驾驶, 银行转帐/汇款/提款, 进入手机/电脑/应用

App, 和政府机关办事所需要的身份证明。您可否想过你现在要纪录多少密码, 多少使用者名称, 身上带着多少付款软体, 银行卡, 信用卡, 钞票, 居民身分证, 交通卡, 驾照, 护照, 居民健康卡, 电梯卡, 加油卡, 金融社保卡, 大卖场会员证, 和各式通行卡; **3. 同步翻译连网:** 目前先进智能芯片/软体的语音辨识速度太慢和不够精准的理解与翻译, 应是语言同步翻译机仍未大卖的主因, 但透过更高速的智能芯片, 无线通讯芯片和更庞大的语音数据库来训练云和端的人工智能的推理反应, 相信未来国际多种语言的零障碍沟通将指日可待; **4. 人工智能教师和消费机器人:** 当把强大的人工智能导入到人工智能教师和消费机器人联网物中, 透过不断的反覆学习, 这不但可提升学生的教育水平和兴趣, 未来都能解决老人及残障人士的健康照顾, 清洁, 饮食, 娱乐, 保全等需求, 减少后代的负担及外佣虐老事件, 也明显能提升老人和残障人士的寿命和生活品质。

3、三种主流人工智能演算法

最早的人工智能出现及运用在 1950—1980 年代, 接着转换到 1980-2010 年机器学习, 从 2010 年以后, 随着各种演算法 CNNs, RNNs, DNNs 等图影像视觉学习, 辨识, 推理的普及, 让深入人工智能深度学习的突飞猛进。深度学习是人工智能和机器学习的一个子集, 它使用多层人工神经网络在诸如对象检测, 语音识别, 语言翻译等任务中提供最先进的准确性。深度学习与传统的机器学习技术的不同之处在于, 它们可以自动学习图像, 视频或文本等数据的表示, 而无需引入手工编码规则或人类领域知识。它们高度灵活的架构可以直接从原始数据中学习, 并在提供更多数据时提高其预测准确性。人工智能的深度学习最近取得的许多突破, 例如谷歌 DeepMind 的 AlphaGo 及更强大的 AlphaZero 陆续在围棋, 西洋棋类比赛夺冠, 谷歌 Waymo, 英伟达的 Xavier/Pegasus 320, 及 Intel/Mobileye 的 Eye 4/5 自动驾驶汽车解决方案, 亚马逊的 Alexa, 谷歌的 [Google Assistant](#), 苹果 Siri, 微软的 Cortana, 及三星的 Bixby 智能语音助手等等。借助加速的深度学习框架, 研究人员和数据科学家可以显着加快深度学习培训, 可以从数天或数周的学习缩短到数小时。当模型可以部署时, 开发人员可以依靠人工智能芯片加速的推理平台来实现云, 边缘运算设备或自动驾驶汽车, 为大多数计算密集型深度神经网络提供高性能, 低延迟的推理。

图表 3：人工智能技术工艺的演化

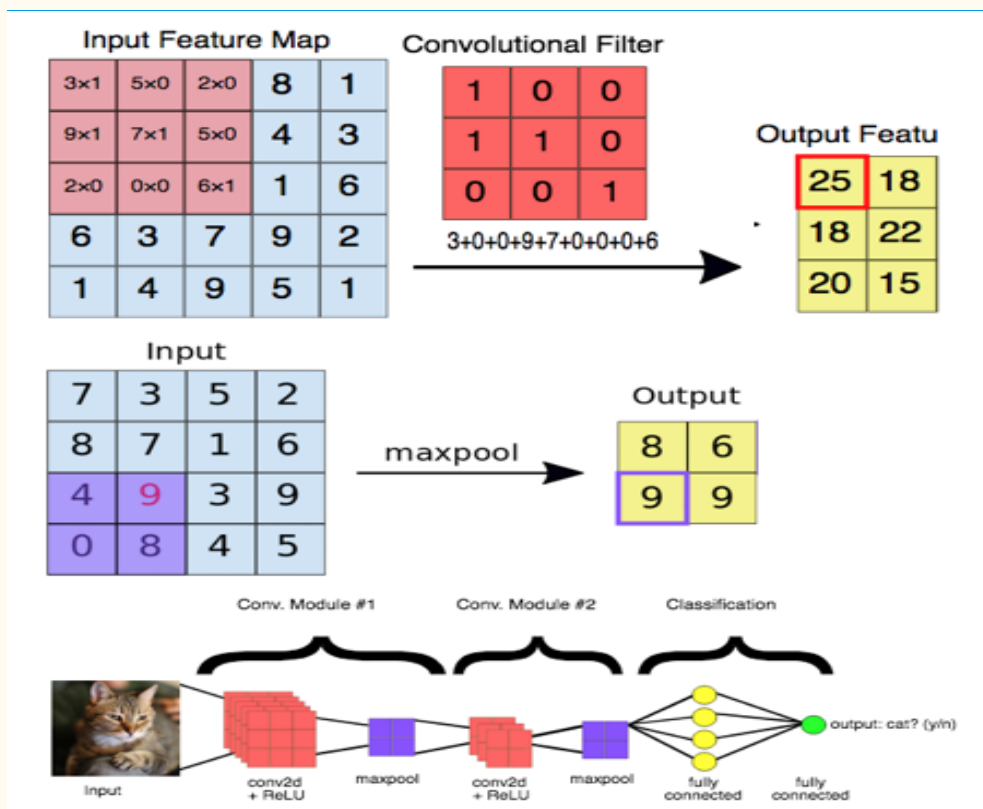


来源：英伟达，国金证券研究所

- **卷积神经网络 CNNs ([Convolutional Neural Networks](#)):** 卷积神经网络 (CNN) 是建立在模拟人类的视觉系统, 并透过图影像分类模型的突破, 也将是, 主要来自于发现可以用于逐步提取图影像内容的更高和更高级别的表示。CNN 是将图像的原始像素数据作为输入, 并“学习”如何提取这些特征, 并最终推断它们构成的对象。首先, CNN 接收输入特征图: 三维矩阵, 其中前两个维度的大小对应于图像的长度和宽度 (以像素为单位), 第三维的大小为 3 (对应于彩色图像的 3 个通道: 红色, 绿色和蓝色)。CNN 包括一堆模块, 每个模块执行三个操作。举例而言, 卷积将 3x3 过滤

贴图的 9 个条件 (0, 1) 套用 (先乘后求和以获得单个值) 在 5x5 输入特征贴图的 9 个像素特征上, 而得出 3x3 全新的卷积输出特征贴图。在每次卷积操作之后, 会采用最大池演算法 (Max pooling), CNN 对卷积特征贴图进行下采样 (以节省处理时间), 同时仍保留最关键的特征信息, 最大池化是要从特征贴图上滑动并提取指定大小的图块 (2x2), 对于每个图块, 最大值将输出到新的特征贴图, 并丢弃所有其他值。在卷积神经网络的末端是一个或多个完全连接的层, 完全连接的层将一层中的每个神经元连接到另一层中的每个神经元。它原则上与多层感知器神经网络 (multi-layer perceptron neural network (MLP)) 类似, 他们的工作是根据卷积提取的特征进行分类, CNN 可以包含更多或更少数量的卷积模块, 以及更多或更少的完全连接层, 工程师经常试验要找出能够为他们的模型产生最佳结果的配置。总之, CNN 专门于图影像处理如自动驾驶汽车, 安防, 人脸辨识, 及疾病图像辨识解决方案。

图表 4: 卷积输入及输出特征贴图及最大池

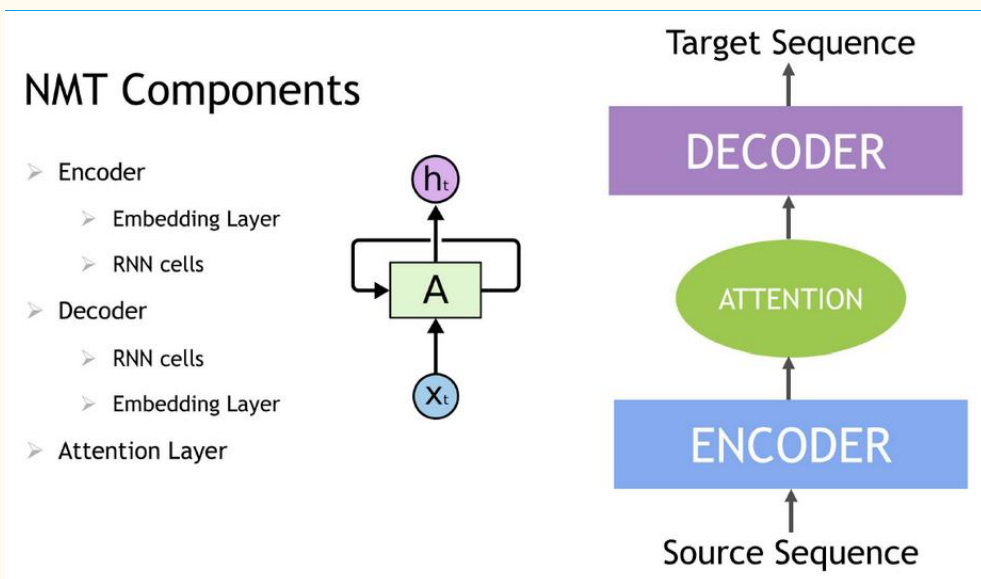


来源: 谷歌, 国金证券研究所

- 循环神经网络 RNNs (Recurrent Neural Network):** RNN 是一类人工听觉及说话的神经网络, 具有记忆或反馈回路, 可以更好地识别数据中的模式。RNN 是常规人工神经网络的扩展, 它增加了将神经网络的隐藏层送回自身的连接 - 这些被称为循环连接。循环连接提供了一个循环网络, 不仅可以看到它提供的当前数据样本, 还可以看到它以前的隐藏状态。具有反馈回路的循环网络可以被视为神经网络的多个副本, 其中一个的输出用作下一个的输入。与传统的神经网络不同, 循环网络使用他们对过去事件的理解来处理输入向量, 而不是每次都从头开始。当正在处理数据序列以进行分类决策或回归估计时, RNN 特别有用, 循环神经网络通常用于解决与时间序列数据相关的任务。不同于 CNN 专门于图影像处理, 循环神经网络的应用包括自然语言处理, 语音识别, 机器翻译, 字符级语言建模, 图像分类, 图像字幕, 股票预测和金融工程。机器翻译是指使用机器将一种语言的源序列 (句子, 段落, 文档) 翻译成相应的目标序列或另一种语言的矢量。由于一个源句可以以许多不同的方式翻译, 因此翻译基本上是一对多的, 并且翻译功能被建模为有条件而非确定性。在神经机器翻译 (NMT) 中, 我们让神经网络学习如何从数据而不是从一组设计规则进行翻译。由于我

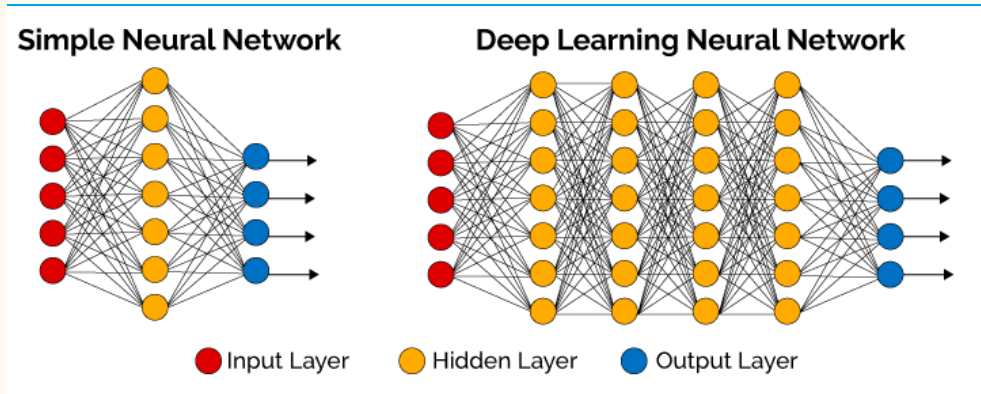
们处理时间序列数据，其中语境的上下文和顺序很重要，因此 NMT 的首选网络是递循环神经网络。可以使用称为注意的技术来增强 NMT，这有助于模型将其焦点转移到输入的重要部分并改进预测过程。举两 RNN 的例子，为了跟踪你的自助餐厅主菜的哪一天，每周在同一天运行同一菜的严格时间表。如周一的汉堡包，周二的咖喱饭，周三的披萨，周四的生鱼片寿司和周五的意大利面。使用 RNN，如果输出“生鱼片寿司”被反馈到网络中以确定星期五的菜肴，那么 RNN 将知道序列中的下一个主菜是意大利面（因为它已经知道有订单而周四的菜刚刚发生，所以星期五的菜是下一个）。另一个例子是如果我跑了 10 英里，需要喝一杯什么？人类可以根据过去的经验想出如何填补空白。由于 RNN 的记忆功能，可以预测接下来会发生什么，因为它可能有足够的训练记忆，类似这样的句子以“水”结束以完成答案。

图表 5：循环神经机器翻译



来源：谷歌，国金证券研究所

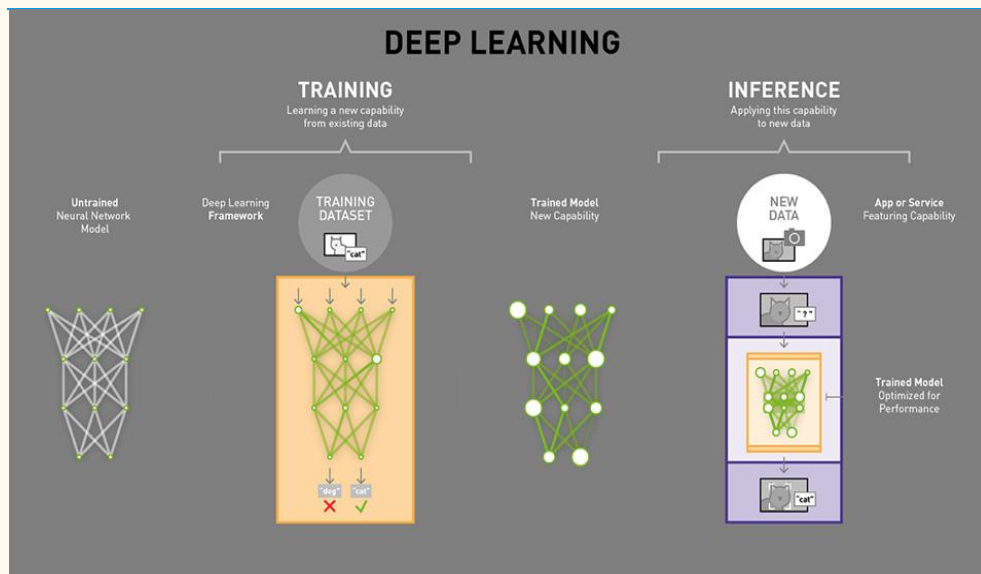
- **深度神经网络 DNNs (Deep Neural Network):** DNN 在视觉，语言理解和语音识别等领域取得了关键突破。为了实现高精度，需要大量数据和以后的计算能力来训练这些网络，但这也带来了新的挑战。特别是 DNN 可能容易受到分类中的对抗性示例，强化学习中遗忘任务，生成建模中的模式崩溃的影响以及过长的运算时间。为了构建更好，更强大的基于 DNN 的系统，是能否有效地确定两个神经网络学习的表示何时相同？我们看到的两个具体应用是比较不同网络学习的表示，并解释 DNN 中隐藏层所学习的表示。设置的关键是将 DNN 中的每个神经元解释为激活向量，神经元的激活矢量是它在输入数据上产生的标量输出。例如，对于 50 个输入图像，DNN 中的神经元将输出 50 个标量值，编码它对每个输入的响应量。然后，这 50 个标量值构成神经元的激活矢量。因为深度神经网络路的规模（即层数和每层的节点数），学习率，初始权重等众多参数都需要考虑。扫描所有参数由于时间代价的原因并不可行，因而小批次训练（微型配料），即将多个训练样本组合进行训练而不是每次只使用一个样本进行训练，被用于加速模型训练。而最显著地速度提升来自 GPU，因为矩阵和向量计算非常适合使用 GPU 实现。但使用大规模集群进行深度神经网络训练仍然存在困难，因而深度神经网络在训练并列化方面仍有提升的空间。

图表 6：深度神经网络

来源：Deep Cognition，国金证券研究所

4、到底哪种人工智能芯片将成云计算的主流？

深度学习是一种需要训练的多层次大型神经网络结构（请参考图表），其每层节点相当于一个可以解决不同问题的机器学习。利用这种深层非线性的网络结构，深度学习可以从少数样本展现强大的学习数据集本质特征的能力。简单来说，深度学习神经网络对数据的处理方式和学习方式与人类大脑的神经元更加相似和准确。谷歌的阿法狗也是先学会了如何下围棋，然后不断地与自己下棋，训练自己的深度学习神经网络，更厉害的阿法零 (AlphaZero) 透过更精准的节点参数，不用先进行预先学习就能自我演化训练学习。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，训练数据的稀缺使得深度学习人工智能在过去没能成为人工智能应用领域的主流算法。但随着技术的成熟，加上各种行动、固定通讯设备、无人驾驶交通工具，可穿戴科技，各式行动、固定监控感测系统能互相连接与沟通的亿物联网，骤然爆发的大数据满足了深度学习算法对于训练数据量的要求。

图表 7：深度学习

来源：英伟达，国金证券研究所

训练和推理所需要的神经网络运算类型不同。神经网络分为前向传播 (Forward algorithm) 其中包括输入层，隐藏层，输出层和后向传播 (Backward algorithm) 主要指的是梯度运算，两者都包含大量并行运算。训练同时需要前向和后向传播，推理则主要是前向传播。一般而言训练过程相比于推理过程计算量更大。云端人工智能系统透过海量的数据集和调整参数优化来负责训练和推理，边缘运算终端人工智能设备负责推理。推理可在云端进行，也可以在边缘运算端或设备端进行。等待模型训练完成后，将训练完成的模型（主要是各种

通过训练得到的参数) 用于各种应用。应用过程主要包含大量的乘累加矩阵运算，并行计算量很大，但和训练过程比参数相对固定，不需要大数据支撑，除在云端实现外，也可以在边缘运算端实现。推理所需参数可由云端训练完后，定期下载更新到应用终端。

图表 8：各种人工智能半导体优缺点比较

	CPU	GPU	FPGA	TPU/ASIC
优点	通用性最高 一般用途 广泛应用于数 据中心的云计 算	适合深度学习训练 适用于云计算 并行运算性能及通用 性高	适合深度学习推理 较低功耗，低延迟， 半定制 适合不断变化的工作 量	可完全定制 耗电率低 性能高
缺点	不适合并行运 算	耗电高，单位成本高	耗电高，浮点速度较 GPU 慢，布线不易	研发期长，初期开发 成本高，通用性差
代表企业	英特尔，超威	英伟达，超威	赛灵思，Lattice， 英特尔/Altera	谷歌，亚马逊，华为 海思，寒武纪，比特 大陆，阿里巴巴，百 度

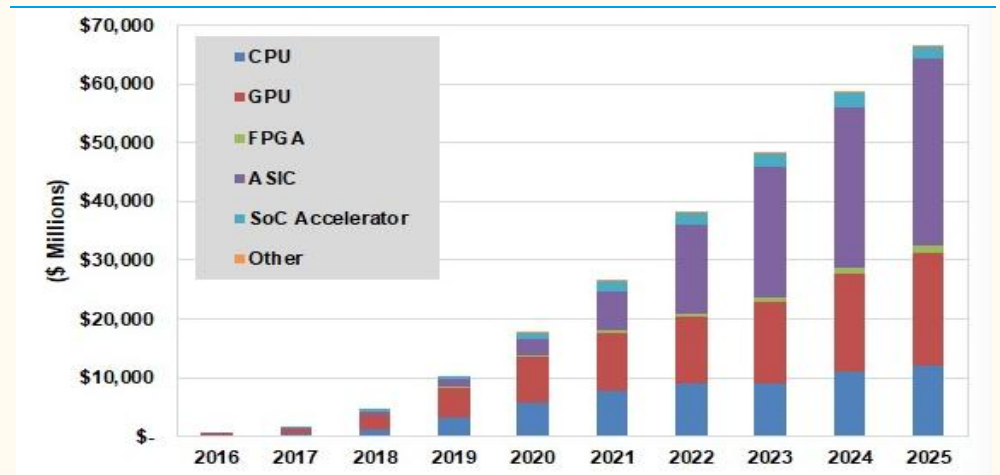
来源：Frost & Sullivan, 国金证券研究所

在深度学习半导体领域里，最重要的是数据和运算。谁的晶体管数量多，芯片面积大，谁就会运算快和占据优势。因此，在处理器的选择上，可以用于通用基础计算且运算速率更快的 GPU 迅速成为人工智能计算的主流芯片，根据美国应用材料的公开资料(请参考图表)，英伟达的人工智能逻辑芯片配合英特尔的中央处理器服务器芯片面积达 7,432mm²，是不具人工智能的企业用和大数据服务器的八倍或谷歌专用张量处理器人工智能服务器的三倍多，存储器耗用面积(32,512mm²) 是其他服务器的三倍以上。可以说，在过去的几年，尤其是 2015 年以来，人工智能大爆发就是由于英伟达公司的图形处理器，得到云端主流人工智能的应用。但未来因为各个处理器的特性不同，我们认为英伟达的图形处理器 GPU 和谷歌的张量处理器仍能主导通用性云端人工智能深度学习系统的训练，可编程芯片 FPGA 的低功耗及低延迟性应有利于主导云端人工智能深度学习系统的推理，而特殊用途集成电路(ASIC) 未来将主导边缘运算及设备端的训练及推理，但因为成本，运算速度，及耗电优势，也会逐步侵入某些特殊应用人工智能云端服务器市场，抢下训练及推理运算的一席之地，以下就先列出各种处理器在云端人工智能系统的优缺点：

图表 9：人工智能云端系统图形处理芯片面积

	ENTERPRISE SERVER	“BIG DATA” SERVER	A.I. SERVER GPU BASED	A.I. SERVER CUSTOM ASICS
	 e.g. Applied Data Center	 e.g. Facebook In-memory DB	 e.g. NVIDIA® DGX-1	 e.g. Google TPU v2
SPEC	Intel® Xeon® E5 (176 GB DRAM)	Intel® Xeon® E5 (256 GB DRAM)	8x Tesla® V100 (128 GB DRAM) + Host CPU (512 GB DRAM)	Custom ASIC (16GB HBM) + Host CPU (256 GB DRAM)
LOGIC AREA	912 mm ²	912 mm ²	7,432 mm ²	2,232 mm ²
MEMORY AREA	7,392 mm ²	10,752 mm ²	32,512 mm ²	16,256 mm ²
TOTAL SILICON	8,304 mm ²	11,664 mm ²	39,944 mm ²	18,488 mm ²

来源：谷歌，微软，脸书，英伟达，应用材料，国金证券研究所

图表 10：人工智能半导体市场预测以不同芯片种类来分类

来源：Tractica，国金证券研究所

- **中央处理器 CPU**: X86 和 ARM 在内的传统 CPU 处理器架构往往需要数百甚至上千条指令才能完成一个神经元的处理，但对于并不需要太多的程序指令，却需要海量数据运算的深度学习计算需求，这种结构就显得不佳。中央处理器 CPU 需要很强的处理不同类型数据的计算能力以及处理分支与跳转的逻辑判断能力，这些都使得 CPU 的内部结构异常复杂，现在 CPU 可以达到 64bit 双精度，执行双精度浮点源计算加法和乘法只需要 1~3 个时钟周期，时钟周期频率达到 1.532~3gigahertz。CPU 拥有专为顺序逻辑处理而优化的几个核心组成的串行架构，这决定了其更擅长逻辑控制、串行运算与通用类型数据运算，当前最顶级的 CPU 只有 6 核或者 8 核，但是普通级别的 GPU 就包含了成百上千个处理单元，因此 CPU 对于影像，视频计算中大量的重复处理过程有着天生的弱势。
- **图形处理器 GPU 仍主导云端人工智能深度学习及训练**: 最初是用于在计算机、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，但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 只需要进行高速运算而不需要逻辑判断。GPU 具备高效的浮点算数运算单元和简化的逻辑控制单元，把串行访问拆分成多个简单的并行访问，并同时运算。例如，在 CPU 上只有 20-30% 的晶体管（内存存储器 DRAM dynamic random access memory，缓存静态随机存储器 Cache SRAM，控制器 controller 占了其余的 70-80% 晶体管）是用作计算的，但反过来说，GPU 上有 70-80% 的晶体管是由上千个高效小核心组成的大规模并行计算架构（DRAM 和微小的 Cache SRAM，controller 占了剩下的 20-30% 晶体管）。大部分控制电路相对简单，且对 Cache 的需求小，只有小部分晶体管来完成实际的运算工作，至于其他的晶体管可以组成各类专用电路、多条流水线，使得 GPU 拥有了更强大的处理浮点运算的能力。这决定了其更擅长处理多重任务，尤其是没有技术含量的重复性工作。不同于超威及英特尔的 GPU 芯片，英伟达的人工智能芯片具有 CUDA 的配合软件是其领先人工智能市场的主要因素。CUDA 编程工具包让开发者可以轻松编程屏幕上的每一个像素。在 CUDA 发布之前，GPU 编程对程序员来说是一件苦差事，因为这涉及到编写大量低层面的机器码。CUDA 在经过了英伟达的多年开发和改善之后，成功将 Java 或 C++ 这样的高级语言开放给了 GPU 编程，从而让 GPU 编程变得更加轻松简单，研究者也可以更快更便宜地开发他们的深度学习模型。因此我们认为目前英伟达价值（6,000/7,500—9,300/10,500 美元）的图形处理器加速卡 Tesla V100 PCIe/SXM2（640 Tensor 核心，5,120 CUDA 核心）或配备 8 / 16 颗 V100 的 DGX-1/H-2 系统（180,000—360,000 美元），配合其 CUDA 软件及 NVLink 快速通道，能达到近 125 兆次深度学习的浮点运算训练速度（Tera FLOPS），以 16bit 的半精度浮点性能来看，可达到 31 Tera FLOPS，32bit 的单精度浮点性能可达到 15.7 Tera FLOPS，及 64bit 的双精

度可达到 7.8 Tera FLOPS，210 亿个晶体管，台积电 12 纳米制程工艺，815mm² 芯片大小，仍然是目前云端人工智能深度学习及训练的最佳通用型解决方案，但未来会受到类似于华为海思 Ascend—Max 910 ASIC 芯片及 Ascend Cluster 系统的挑战。

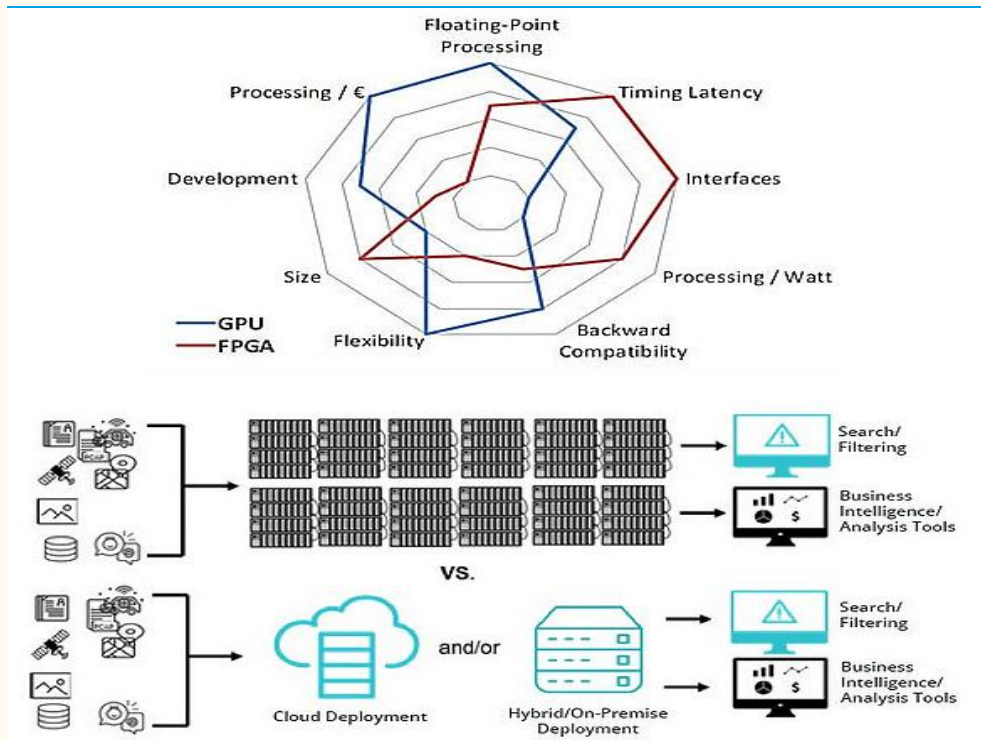
图表 11：英伟达云端人工智能芯片 V100 及系统 DGX-1 规格比较表

			
Tesla V100 PCIe		Tesla V100 SXM2	
GPU Architecture NVIDIA Volta			
NVIDIA Tensor Cores		640	
NVIDIA CUDA® Cores		5,120	
Double-Precision Performance	7 TFLOPS	7.8 TFLOPS	
Single-Precision Performance	14 TFLOPS	15.7 TFLOPS	
Tensor Performance	112 TFLOPS	125 TFLOPS	
GPU Memory		32GB /16GB HBM2	
Memory Bandwidth		900GB/sec	
ECC		Yes	
Interconnect Bandwidth	32GB/sec	300GB/sec	
System Interface	PCIe Gen3	NVIDIA NVLink	
Form Factor	PCIe Full Height/Length	SXM2	
Max Power Consumption	250 W	300 W	
Thermal Solution		Passive	
Compute APIs		CUDA, DirectCompute, OpenCL™, OpenACC	
SYSTEM SPECIFICATIONS			
GPUs		8X Tesla V100	
Performance (Mixed Precision)		1 petaFLOPS	
GPU Memory		256 GB total system	
CPU		Dual 20-Core Intel Xeon E5-2698 v4 2.2 GHz	
NVIDIA CUDA® Cores		40,960	
NVIDIA Tensor Cores (on V100 based systems)		5,120	
Power Requirements		3,500 W	
System Memory		512 GB 2,133 MHz DDR4 RDIMM	
Storage		4X 1.92 TB SSD RAID 0	
Network		Dual 10 GbE, 4 IB EDR	
Operating System		Canonical Ubuntu, Red Hat Enterprise Linux	
System Weight		134 lbs	
System Dimensions		866 D x 444 W x 131 H (mm)	
Packing Dimensions		1,180 D x 730 W x 284 H (mm)	
Operating Temperature Range		5-35 °C	

来源：英伟达，国金证券研究所

- 现场可编程门阵列芯片 FPGA 的优势在低功耗，低延迟性：CPU 内核并不擅长浮点运算以及信号处理等工作，将由集成在同一块芯片上的其它可编程内核执行，而 GPU 与 FPGA 都以擅长浮点运算著称。FPGA 和 GPU 内都有大量的计算单元，它们的计算能力都很强。在进行人工智能神经网络 (CNN, RNN, DNN) 运算的时候，两者的速度会比 CPU 快上数十倍以上。但是 GPU 由于架构固定，硬件原来支持的指令也就固定了，而 FPGA 则是可编程的，因为它让软件与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。虽然 FPGA 比较灵活，但其设计资源比 GPU 受到较大的限制，例如 GPU 如果想多加几个核心只要增加芯片面积就行，但 FPGA 一旦型号选定了逻辑资源上限就确定了。而且，FPGA 的布线资源也受限制，因为有些线必须要绕很远，不像 GPU 这样走 ASIC flow 可以随意布线，这也会限制性能。FPGA 虽然在浮点运算速度，增加芯片面积，及布线的通用性比 GPU 来得差，却在延迟性及功耗上对 GPU 有着显著优势。英特尔斥巨资收购 Altera 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，那就是从现在分立的 CPU 芯片+分立的 FPGA 加速芯片 (20nm Arria 10GX)，过渡到同一封装内的 CPU 晶片+FPGA 晶片，到最终的集成 CPU+FPGA 系统芯片。预计这几种产品形式将会长期共存，因为 CPU 和 FPGA 的分立虽然性能稍差，但灵活性更高。目前来看，用于云端的人工智能解决方案是用 Xeon CPU 来配合 Nervana，用于云端中间层和边缘运算端设备的低功耗推断解决方案是用 Xeon CPU 来配合 FPGA 可编程加速卡。赛灵思 (Xilinx) 于 2018 年底推出以低成本，低延迟，高能效率的深度神经网络 (DNN) 演算法为基础的 Alveo 加速卡，采用台积电 16nm 制程工艺的 UltraScale FPGA，预期将拿下不少人工智能数据中心云端推理芯片市场不少的份额。

图表 12：赛灵思/BlackLynx 与 GPU 在机器学习推理解决方案的比较



来源：赛灵思，BlackLynx，国金证券研究所

- **谷歌张量处理器 TPU 3 强势突围，博通 / 台积电受惠，可惜不外卖：**因为它能加速其人工智能系统 TensorFlow 的运行，而且效率也大大超过 GPU—Google 的深层神经网络就是由 TensorFlow 引擎驱动的。谷歌第三代张量处理器 (TPU, Tensor Processing Unit, 大约超过 100 Tera FLOPS/hp-16bit) 是专为机器学习由谷歌提供系统设计，博通提供 ASIC 芯片设计及智财权专利区块，台积电提供 16/12 纳米制程工艺量身定做的，执行每个操作所需的晶体管数量更少，自然效率更高。TPU 每瓦能为机器学习提供比所有商用 GPU 和 FPGA 更高的量级指令。TPU 是为机器学习应用特别开发，以使芯片在计算精度降低的情况下更耐用，这意味每一个操作只需要更少的晶体管，用更多精密且大功率的机器学习模型，并快速应用这些模型，因此用户便能得到更正确的结果。以谷歌子公司深度思考的阿尔法狗及零 (AlphaGo, AlphaZero/DeepMind) 利用人工智能深度学习训练和推理来打败世界各国排名第一的围棋高手，世界排名第一的西洋棋 AI 程式 Stockfish 8, 世界排名第一的日本棋 Shogi AI 专家，但我们估计 AlphaZero 系统使用至少近 5 大排人工智能主机，5,000 个张量处理器，1,280 个中央处理单元而让云端的设备异常昂贵,且无提供任何边缘运算端设备。

图表 13：谷歌张量处理器 TPU3 vs. TPU2

谷歌 TPU3 的深度学习数据中心			谷歌 TPU2 的深度学习数据中心		
处理器	张量处理器	中央处理器	张量处理器	中央处理器	
制程工艺	16/12nm	Intel/AMD	20nm	Intel	
存储容量	32GB HBM		16GB HBM		
芯片/主机板	4	2	4	2	
主机板/架子	2	1	2	2	
架子/大支架	16	16	8	8	
大支架/一排	8	8	4	4	
芯片/一排	1024	256	256	128	
Petaflops	> 100		11.5		

来源：谷歌, TIRIAS, 国金证券研究所

- ASIC 特定用途 IC 需求即将爆发：**即使研发期长，初期开发成本高，通用性差，但国内芯片业者因缺乏先进 x86 CPU, GPU, 及 FPGA 的基础设计智慧财产权 (IPs)，可完全客制化，耗电量低，性能强的特定用途 IC (ASIC Application-specific integrated circuit) 设计就立刻成为国内进入人工智能云端及边缘运算及设备端芯片半导体市场的唯一途径。但因为起步较晚，除了比特大陆的算丰 (SOPHON) BM1680 及 BM1682 云端安防及大数据人工智能推理系列产品已经上市之外，其他公司在云端人工智能训练及推理芯片设计都还停留在纸上谈兵阶段。举例而言，华为海思使用台积电 7 纳米制程工艺设计的昇腾 Ascend 910 ASIC 系列，号称在 16bit 半精度下能达到 256 兆次的浮点运算，倍数于英伟达目前最先进的 Volta 100 解决方案(台积电 12 纳米) 也要等到 2H19 量产及谷歌最新推出的张量处理器 3 (台积电 16/12 纳米)。而从智能手机端 IP 事业切入设计云端芯片的寒武纪，将于 2019 首发的产品 MLU100 PCIe 智能推理加速卡 (台积电 16 纳米) 在打开稀疏模式时的峰值也可达到 64 兆次浮点运算 (16bit 半精度)。最后可惜的是国内搜寻引擎龙头百度的昆仑芯片(818-300 采用 Samsung 14 纳米)，阿里巴巴的 Ali-NPU, 及亚马逊的 Inferentia 目前都还没有提供更确实的芯片速度，耗电量，应用，价格，量产时点，及软件框架规格让我们做出更好的比较图表。

图表 14：各公司云端人工智能芯片比较表

	英伟达	赛灵思	英特尔	谷歌	华为/海思	寒武纪	比特大陆	百度	亚马逊	阿里巴巴
模块或芯片	V100 SXM2	Alveo U250 PCIe 加速卡	可编程加速卡	TPU3	昇腾 910 云端数据 Cluster	MLU 100 PCIe 智能处理卡	SOPHON	昆仑	Inferentia	Ali-NPU
图形处理器	TSMC 12nm Volta GPU 21bn 晶体管, 815mm2/die	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
机器学习加速器	640 Tensor cores, 5120 CUDA cores	TSMC 16nm UltraScale 20nm FPGA/54MB SRAM	Arria 10GX FPGA	TSMC 16/12nm 张量处理器	Ascend-Max 910 TSMC 7nm 特定 IC ASIC	TSMC 16nm MLU, 1.0/1.3Ghz, 特定 IC ASIC	64x 神经处理单元 NPU BM1686/ TSMC 12nm	SEC 14nm 818-300 训练; 818-100 推理	Inferentia ML ASIC	阿里云公共服务神经网络 (NPU)
速度	125 Tera FLOPS; 31 Tera FLOPS/16bit 半精度/300W	4Kp60, 33.3 INT8 TOPS/ 225W	1.5 Tera FLOPS/66W TDP	>100 Tera FLOPS/16bit 半精度 / 200W	256Tera FLOPS/16bit 半精度 /350W	64 Tera FLOPS/16bit 半精度 /75/110W (打开稀疏模式时峰值)	9 Tera FLOPS/30W	260 Tera OPS/100W	N/A	N/A
中央处理器	N/A	AMD EPYC x86	Intel Xeon CPU	Intel, AMD	N/A	N/A	N/A	语音, 图像, 自动驾驶	N/A	N/A
内存 DRAM	32GB HBM2	64GB DDR 4	8GB DDR4	128GB HBM	32GB HBM	8/16GB	N/A	N/A	N/A	N/A
应用	图像训练及推理	DNN 图像推理	AI 图像推理, 大数据分析, 网络安全, 基因工程, 高速运算	CNN, RNN, DNN 图像训练及推理	云端数据中心, 训练及推理	云端数据中心推理	云端安防, 大数据推理用	自然语言处理、大规模语音识别、自动驾驶和大规模推荐	N/A	图像视频分析、机器学习等 AI 推理计算
价格 (美元), 量产	10,500	12,995	已量产	内部使用/已量产	加速卡 2H19 量产	US\$800-1000; 处理卡 2019 量产	2H19 量产	N/A	2H19 量产	N/A

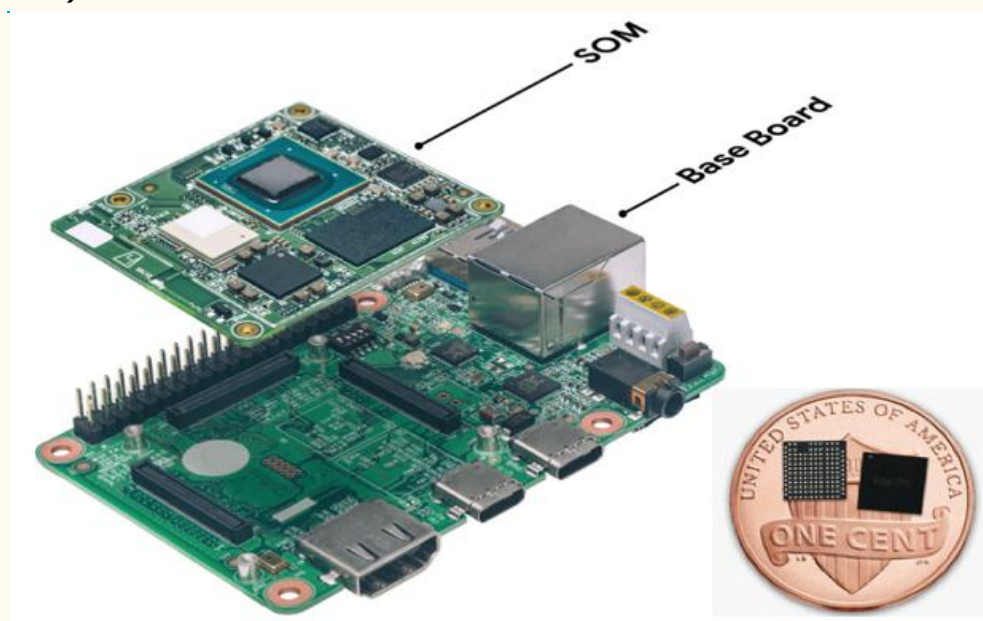
来源：各公司，国金证券研究所

5、哪种人工智能芯片将成边缘运算及设备端的主流？

■ 从算法，IP，边缘运算及设备端芯片转换到模块，平台或生态系的战争

谷歌在 Google Cloud NEXT 2018 活动中，首次推出人工智能推断用边缘运算张量处理器（Edge Tensor Processing Unit, Edge TPU, <30mm²）具有低功耗，低延迟，高运算效率，其开发版系统模块套件中还包括有恩智浦的 CPU，图芯的 GPU，再配合谷歌开源简易版机器学习框架软件（Open-source TensorFlow Lite），设备端物联网核心运作（Edge IoT Core）和边缘运算端张量处理器，来推动各种应用，像是预测性维护，异常检测，机器视觉，机器人，语音识别，医疗保健，零售，智能空间，运输交通等等。

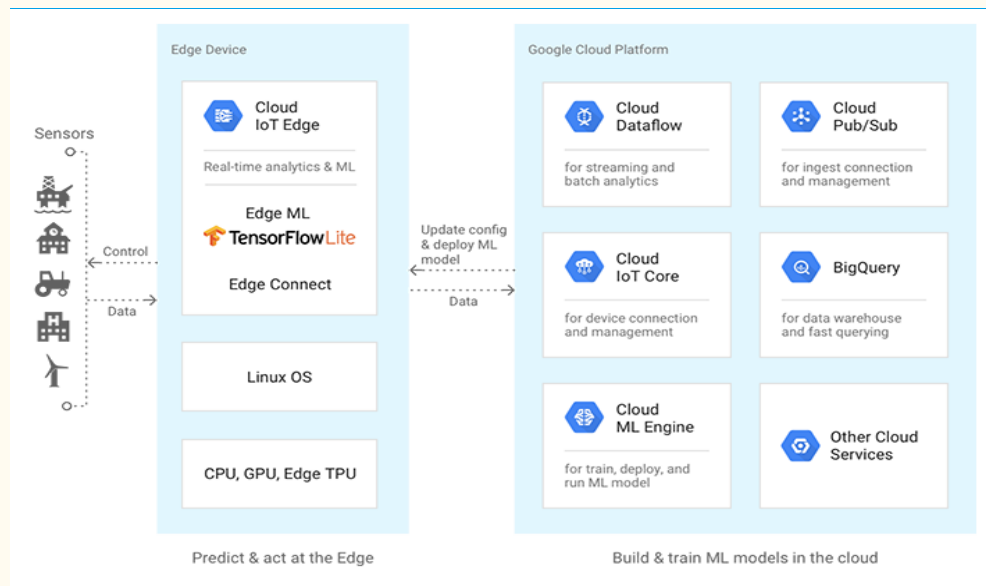
图表 15：谷歌的智能推断边缘运算端的解决方案模块 (System-on-module, SOM)



来源：谷歌，国金证券研究所

在谷歌的 Waymo 建立了全球最大的无人驾驶车队后，谷歌再次利用其在深度学习及云端软，硬件的技术领先优势，提供机器学习边缘运算端软件，固件，安卓物联网作业系统及专用半导体芯片整体解决方案模块，让客户对其智能物联网解决方案的黏着度提升，这不但对中国大陆人工智能芯片公司是利空，也对目前主要提供云及边缘运算物联网服务的竞争厂商亚马逊（AWS），微软（Azure），阿里巴巴造成市场压力。

图表 16：谷歌的智能物联网整体解决方案



来源：谷歌，国金证券研究所

但是，目前谷歌终端型张量处理器目前只能算是个机器学习的加速器，不能独立运作，其解决方案中还要配合其他半导体公司的芯片，像是恩智浦以安谋为核心的中央处理器，图芯（Vivante Corporation）的 GC7000 Lite 图形处理器（请参考图表），我们认为谷歌的解决方案效能比较与量产时间未定。而且，谷歌过去还是主推软件解决方案，自有硬件及半导体的产品上市都是拿来推广其软件及生态系解决方案，其边缘运算型张量处理器硬件规格中，还是使用相对旧的 28 纳米晶圆代工制程工艺，Tera OPS 浮点运算能力，耗电瓦特数，价

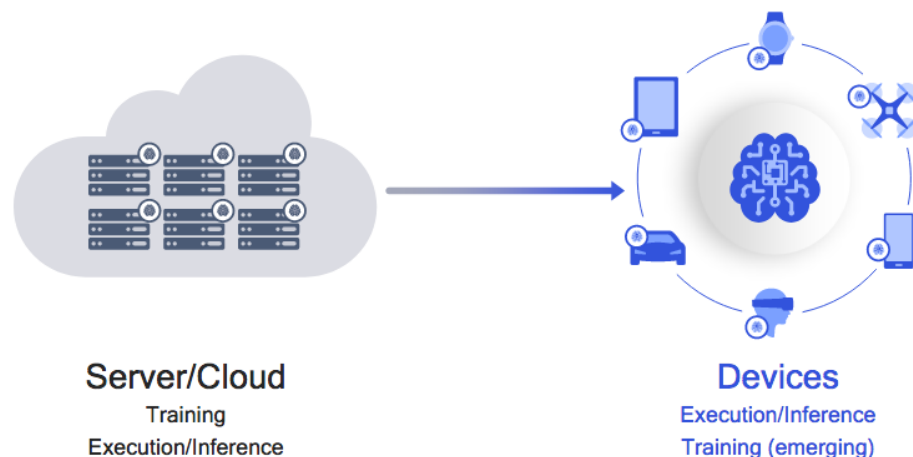
格等指标都还是未知，来与其他人工智能芯片设计公司的产品来做比较。举例而言，英伟达今年推出的 Jetson AGX Xavier 模块也是瞄准物联网应用端的解决方案，但在其高单价下规格明显胜出。华为即将推出的边缘运算模块 Atlas 200 和 Atlas 300 也相当有竞争力。

图表 17：智能推断边缘运算端的解决方案模块比较表

	英伟达		英特尔		谷歌		华为	
模块 (System on Module)	Jetson AGX Xavier	TX2	Movidius Myriad Edge TPU			Atlas 200 加速模块	Atlas 300 加速卡	
图形处理器	512-core Volta 1.38Ghz with 64 Tensor Cores	英伟达 Pascal, 256 N/A CUDA cores	图芯 GC7000 Lite Graphics			N/A	N/A	
机器学习加速卡	12nm Volta GPU 21bn 晶体管, 815mm2/die	16nm Pascal GPU, 12bn 晶体管, 471mm2/die	28nm Myriad X 28nm Edge 张量处理器	28nm Myriad X 28nm Edge 张量处理器	28nm Myriad X 28nm Edge 张量处理器	12nm 昇腾 ASIC	310 12nm 昇腾 ASIC	310 12nm 昇腾 ASIC
视频速度	4Kp60 Encode; 8Kp60 Decode; 22 TOPS/INT8	4Kx2K 60Hz Encode; 4Kx2K 60Hz Decode; 47 TOPS/INT8	N/A	N/A	N/A	2K p60 Encode	2K p60 Encode, 64TOPS INT8/10W	p60
中央处理器	8-core Carmel ARMv8.2 2.2Ghz	64bit HMP Denver 2/2 L2+, Quad ARM processors A57/2MB cache	Dual 16x SHAVE 恩智浦 i.MX 8M SoC and (Quad Cortex-A53, Cortex-M4F)	Dual 16x SHAVE 恩智浦 i.MX 8M SoC and (Quad Cortex-A53, Cortex-M4F)	Dual 16x SHAVE 恩智浦 i.MX 8M SoC and (Quad Cortex-A53, Cortex-M4F)	N/A	N/A	
闪存 NAND	32GB eMMC	32GB eMMC	N/A	8GB eMMC	8GB eMMC	N/A	N/A	
内存 DRAM	16GB 64bit LPDDR4	8GB 128bit LPDDR4	N/A	1GB LPDDR4	1GB LPDDR4	N/A	N/A	
有线, 无线	Gigabit 802.11ac Bluetooth	Ethernet, WLAN, Bluetooth	N/A	Gigabit 802.11b/g/n/ac	Gigabit 802.11b/g/n/ac	N/A	N/A	
应用	机器人, 智慧视频分析, 医学器材, 嵌入式物联网	智慧视频分析, 智慧交通, 人工智能城市	主要用于镜头异常检测, 机器人视觉, 语音识别, 医疗保健, 零售, 智能空间, 运输头、无人机等	主要用于镜头异常检测, 机器人视觉, 语音识别, 医疗保健, 零售, 智能空间, 运输头、无人机等	主要用于镜头异常检测, 机器人视觉, 语音识别, 医疗保健, 零售, 智能空间, 运输头、无人机等	可支持 16 路高清视频实时分析和边缘侧服务场景	面向数据中心和边缘侧服务场景	
价格 (美元)	1300-1400	550-600	N/A	尚未量产	尚未量产	尚未量产	尚未量产	
模块面积	100mm x 87mm	50mm x 87mm	27mm x 72.5mm	56mm x 85mm	56mm x 85mm	35mm x 43 mm	53mm x 87mm	

来源：英伟达，英特尔，谷歌，华为，国金证券研究所

从英伟达，谷歌，英特尔，及华为在深入学习边缘运算端解决方案模块及软件与固件的竞争，而高通及联发科陆续将人工智能专利区块透过嵌入式整合到其手机应用处理器中，可以看出未来在人工智能端的应用领域已经不再单是人工智能算法，IP(智财权)到芯片设计推陈出新的竞争，中国大陆的人工智能算法/IP/芯片龙头公司像寒武纪，地平线为了存活，就必须与应用领域系统公司紧密合作，共同推出更佳的嵌入式或独立式解决方案模块及软，固件，否则就要像谷歌和英伟达一样推出自己整套的解决方案。虽然比特大陆及嘉楠耘智进入人工智能芯片领域较晚，但其在挖矿机业务及挖矿生态系的系统整合经验，反而比只具备算法/IP/芯片的人工智能设计公司还有机会。

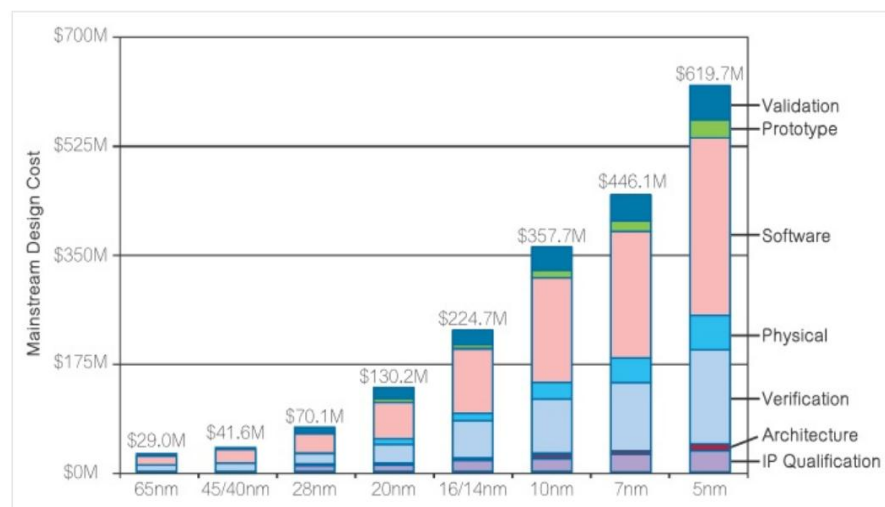
图表 18：人工智能深度学习从云到设备端**Intelligence is moving to the device**

来源：高通，国金证券研究所

6、谁能引领国内人工智能芯片产业突围？

未来因人工智能边缘运算推理端和云端推理（Inferencing）芯片及设备成本，性能，耗电，效率的考量，以及各种处理器的特性不同，我们预期特定用途集成电路（ASIC）或系统集成电路（SoC, system on chip）未来将在设备，边缘运算，及云端推理市场和设备，边缘运算，小部分云端训练市场遍地开花，百花齐放，根据产业链调查，我们认为中国在处理器/芯片领域的投资有加速的迹象，AI 芯片的创业企业目前已达到 40 家左右。

而中国的人工智能半导体公司像华为海思（HiSilicon），寒武纪（Cambricon），地平线（Horizon Robotics），比特大陆（Bitmain），耐能（Kneron），肇观（NextVPU），及大数据系统公司如百度，阿里巴巴及安防龙头海康威视，大华将追随国际大厂陆续推出人工智能特定用途集成电路和系统芯片，从过去的模仿和追赶模式改为提前布局的思路，加上今年中国科创板融资平台的加持，相信一定可以孵育出未来世界级的人工智能芯片设计龙头公司。但就未来 2-3 年来看，因为有强大的财力资源来支持 10 纳米以下先进制程工艺的电子设计自动化软件（EDA tool），验证（Verification），光掩膜（Photo Mask）成本的蹿高（请参考图表），最有实力的半导体设计团队，及其对人工智能深度学习系统的知识及各种设备人工智能化的需求，我们预测华为海思将能引领国内人工智能芯片产业突围；而比特大陆因为深谙 IP，芯片，模块，到平台或生态系的争战，我们不排除比特大陆反而领先其他一些一线设计公司率先推出些云端人工智能的推理/训练芯片及解决方案模块；当然，我们还是认为纯算法及半导体设计公司寒武纪，目前有较佳的设计团队及较充裕的估值融资能力来陆续推出边缘运算端及云端推理的人工智能芯片；最后因为 SAE Level 4/5 自动驾驶平台成本过高及生态系组成庞大及复杂，目前我们对地平线在自动驾驶的解决方案方面的短期营运及获利模式存疑。

图表 19：集成电路设计成本飙升

来源：International Business Strategies，国金证券研究所

- **华为海思后发先至：**在供应全球智能手机第二大厂，中国智能手机第一大厂华为超过近六成的应用处理器及基频芯片，及率先数月于高通及联发科采用台积电的 7 纳米制程工艺推出的麒麟 980 芯片，让华为的 Mate 20 手机大卖，并让华为整体智能手机在国内的市占从 2017 年的 20%，提高到 2018 年的 27%，及 2018 年四季度的 29%，预期这也将同时拉升华为海思 (HiSilicon) 在全球无晶圆设计公司的市占到 6-7% 或是在 2019 年成为前五大。而 7 纳米的麒麟 980 芯片是整合了寒武纪 (Cambricon) 设计的 Cambricon-1M 神经处理单元 (NPU, Neural Processing Unit) 专利区块来让摄像头看得更全，更清，更快，更久，让麦克风听得更清，更广，识别场景，推测用户行为意图，和适时处理高性能或高效率任务，根据华为公布资料，在人工智能性能比上，其 NPU 是一般 CPU 的 25 倍，GPU 的 6.25 倍 (25/4)，能效比上，NPU 更是达到了 CPU 的 50 倍，GPU 的 6.25 倍 (50/8) (请参考图表)。但因为目前其边缘运算端神经处理单元架构简单，我们估计此专利区块芯片应不超过整体芯片面积(基频和应用处理器整合在一起的系统芯片面积)的 10%。但我们预期采用台积电 7 纳米 EUV 制程工艺的华为海思麒麟 990 芯片，将使用海思自行设计的人工智能专利区块 Ascend Lite SKU (请参考图表)，而苹果 (Neural Engine, 5 Tera OPS) 高通 (Neural Processing Engine SDK Software Development Kit, < 8 Tera OPS)，联发科 (NeuroPilot SDK, Android Neural Network NNAPI [联发科 NeuroPilot SDK](#)) 也将陆续整合其人工智能专利区块及软件到其手机应用处理器 (Application processor) 中，这将对寒武纪及其他人工智能算法及芯片公司在智能手机及物联网 (IoT) 设备端的芯片发展造成限制。

华为海思因为有这样强大的财物资资源来支持 7 纳米以下先进制程工艺的电子设计自动化软件 (EDA tool)，验证 (Verification)，光掩膜 (Photo Mask) 成本的蹿高，拥有国内最强的半导体设计团队，及其庞大的系统知识及各种设备对人工智能化的需求，我们预测华为海思将后发先至引领国内人工智能芯片产业突围。华为海思今年将陆续问世的 7 纳米昇腾 Ascend-Max 910 ASIC (整合 8 颗芯片 dies)，及 1,024 颗昇腾 910 芯片的 Ascend Cluster (256 Peta FLOPS)，使用 12 纳米昇腾 Ascend-mini (310)，-Lite，-Tiny，-Nano ASIC 推出的 Atlas 200 加速模块，Atlas 300 加速卡，Atlas 500 智慧小站，Atlas 800 私有云解决方案一体机，MDC 600 移动数据中心 (Mobile Data Center)。我们预期这些 AI 芯片及系统，陆续将对英伟达，赛灵思，英特尔，谷歌在云端及边缘运算端人工智能芯片及平台的地位，带来挑战，但特定用途 IC 的专用性缺点会让华为海思切入像是外部安防等系统公司客户时，碰到些安防公司系统知识领域不愿意外泄的问题。

图表 20：华为云，边缘运算，及设备端的芯片解决方案

	Device 端				Edge 边缘运算		云 AI
	Earphone 耳机 电话	Always-on	Smartphone 智能手机	Laptop 计算机	IPC 工业电 脑	Edge Server 边缘服务器	Data Center 数 据中心
算力	20 MOPS	100 GOPS	1-10 TOPS	10-20 TOPS	10-20 TOPS	10-100 TOPS	> 200 TOPS
功耗 (Watts)	1mW	10mW	1-2W	3-10W	3-10W	10-100W	> 200 W
模型大小	10KB	100KB	10MB	10-100MB	10-100MB	> 100 MB	> 300 MB
延迟 Latency	< 10 ms	10 ms	10-100 ms	10-500 ms	10-500 ms	N/A	N/A
推理	yes	yes	yes	yes	yes	yes	yes
训练	No	No	yes	yes	yes	yes	yes
Ascend 系列	Nano	Tiny	Lite	昇腾 310 Mini	昇腾 310 Mini	昇腾 310 Multi-Mini	昇腾 910 Max

来源：华为海思，国金证券研究所

- **寒武纪从设备端步入云端：**寒武纪 (Cambricon) 是于 2016 年在北京成立，核心成员陈天石及陈云霁兄弟都曾在中科院计算所工作，专攻计算机处理器结构优化和人工智能，而后来寒武纪在拿到中关村科技园区支持资金及上海市政府对神经网络处理器，深度学习处理器 IP 核项目，智能处理器核项目等多项补助近 6,000 万人民币，加上多次拉高估值的融资，于 2018 年 5 月 3 日，发布了使用 TSMC 7nm 工艺 IP 的 1M，每瓦速度达 3.1-3.3 兆次运算，为 10 纳米 1A 智财权专利区块的 10 倍左右，且超越英伟达 V100 的每瓦速度达 0.4 兆次运算，其 8 位运算效能比达 5Tops/watt（每瓦 5 兆次运算）。寒武纪提供了三种尺寸的处理器内核（2Tops/4Tops/8Tops）以满足不同领域下，不同量级智能处理的需求（智能手机、智能音箱、摄像头、自动驾驶）。Cambricon 也介绍以 TSMC 16nm 工艺制程设计的 MLU 100 及 MLU 200 云端服务器 AI 芯片，具有很高的通用性，可满足计算机视觉、语音、自然语言处理和数据挖掘等多种云端推理，甚至训练的任务。在发布会上，联想 (ThinkSystem SR650)，中科曙光 (Phaneron 服务器)，科大讯飞 (翻译机 2.0) 都介绍了使用 Cambricon MLU100 芯片相对应的云端服务器。此外，专为开发者打造的 Cambricon NeuWare 人工智能软件平台，加上支持 TensorFlow, Caffe, MXNet 等主流机器学习框架 (Framework)，让寒武纪在尚未扭亏为盈的情况下（估计 2017 年亏损超过 1,000 万人民币以上），2018 年营收连 1,000 万美金都达不到的状况下（2016/2017 年营收估计约 400 万人民币上下），市值已被拉高到超过 25 亿美元。

图表 21：寒武纪 16nm MLU100 云端 AI 芯片



来源：寒武纪 (Cambricon)，国金证券研究所

- **卖人工智能解决方案模块的地平线：**地平线 (Horizon Robotics) 创办人是前百度深度学习研究院负责人余凯、还有前华为芯片研发架构师周峰，地平线不是只做芯片，在软件方面，地平线已经研发出了自动驾驶的雨果神经网络 OS 平台及智能家居的安徒生平台。地平线的目标是做分支处理单元 (BPU, Branch processing unit) 的人工智能算法架构+嵌入式芯片的 (Embedded ARM, CPU, GPU, FPGA) 自动驾驶 (征程 2.0 处理器)，智能城市，智能商业 (旭日 1.0 处理器) 的人工智能设备终端解决方案模块 (具有感知，识别，理解，控制的功能) 给产品厂商。地平线的整个流程是根据应用场景需求，设计算法模型，在大数据情况下做充分验证，模型成熟以后，再开发一个芯片架构去实现，该芯片并不是通用的处理器，而是针对应用场景跟算法是结合在一起的人工智能算法处理器，得到芯片大小，执行，耗能 (Area、Performance、Power) 的综合解决方案。地平线 BPU 架构的解决方案只可以用在符合车规的 FPGA 或 GPU 等计算平台上，下一步地平线会将自己的 BPU 处理器 IP 授权给国际厂商，让他们生产车规级处理器。智能驾驶方面，基于高斯架构研发的 ADAS 产品 (征程 1.0 处理器) 也会作为重点推进，SAE L3/L4 的无人驾驶 Matrix 1.0 平台，也会持续落实与系统厂商像是博世的技术合作，计划在年中实现特定道路的自动驾驶；智能生活方面，除与美的的合作之外，地平线继续在家电、玩具、服务机器人等领域发力；公共安防方面，地平线去年与英特尔在北美安防展上进行联合展示。类似于寒武纪，地平线于 2018 年 11 月 27 日获得近 10 亿美金的 B 轮融资，持续拉高其市值。

图表 22：地平线 Matrix 1.0 SAE L3/L4 无人驾驶平台解决方案模块

来源：地平线，国金证券研究所

- 从挖矿机转云端人工智能推理模块的比特大陆：**由 CEO 詹克团及创办人吴忌寒于 2013 年联合成立的挖矿机及芯片霸主比特大陆 (BITMAIN) 于 2017 年 11 月，正式介绍其 AI ASIC 芯片品牌 SOPHON (算丰)，宣布全球首款云端安防及大数据人工智能推理系列的张量加速计算芯片 28nm BM1680 的震撼面世，并展示了视频图像分析、人脸人体检测的演示。并同步发布了 SOPHON.AI 官网，并将系列产品在官网中面向全球发售。BITMAIN 致力于通过强大的芯片工程、快速迭代和系统设计制造能力，提供最具性价比、最具性能功耗比的 AI 计算力，同时致力于为行业定制、优化全栈的硬件和系统方案，从而极大降低行业+AI 的难度，促进 AI 普及。BM 1682 在 2017 年 12 月已进入流片阶段，并于 2Q18 量产。12nm 的 BM 1684 是预计于 1Q19 量产及第四代的 12 纳米芯片 BM1686 是预计于 2H19 量产。这两款芯片会拥有 6 / 9 Tera Flops 的能力和 30W 的功耗。BM1880 是比特大陆于 2018 年 10 月发布的一款设备端 AI 芯片，将主要应用于安防、互联网及园区等领域 BM1880 芯片可以作为深度学习推理加速的协处理器，也可以作为主处理器从以太网接口或 USB 接口接收视频流、图片或其它数据，执行推理和其他计算机视觉任务，其它主机也可以发送视频流或图片数据给 BM1880，BM1880 做推理并将结果返回主机。比特大陆将于 2019 年推出第二代产品 BM1882，以及 2020 年的 BM1884，按照规划，BM1882 和 BM1884 的主要应用场景将是智能摄像机、智能机器人和智能家居等。虽然目前比特币跌破 4,000 美元以下，占比特大陆 98% 的矿机销售，自营挖矿业务要是采用两年折旧几乎是无利可图（除了于 4Q18 推出的 7 纳米 BM1391 挖矿芯片及 S15 挖矿机应可获利外），而 AI 芯片导入云端系统又遥不可期，但不同于其他新兴 AI 芯片设计公司大多缺乏现金，比特大陆在手现金（7-8 亿美元现金，4-5 亿美元的加密货币）应该还是有超过 10 亿美元，芯片研发设计资源仍然丰厚，每一代芯片代与代之间的间隔是快于摩尔定律而达到 9—12 个月。摩尔定律是指芯片行业每 18 到 24 个月的周期里，计算能力能翻一倍，或者在相同的单位芯片面积里，晶体管数量翻一倍。

图表 23：比特大陆损益表

美元 (US\$bn)	2015	2016	2017	1H18
营业收入	0.14	0.28	2.52	2.85
营业成本	0.07	0.13	1.30	1.82
毛利率 (%)	52%	55%	48%	36%
营业利润	0.06	0.14	1.14	1.07
营业利润率 (%)	42%	50%	45%	38%
净利润	0.05	0.11	0.70	0.74
净利润率 (%)	35%	41%	28%	26%

来源：比特大陆，国金证券研究所

图表 24：比特大陆挖矿机芯片制程工艺

系列	加密货币/算法	工艺	推出日期	状态
BM1382	比特币、比特币 现金/SHA256	28纳米	2014年6月	停产
BM1384			2014年12月	
BM1385			2015年8月	
BM1387		16纳米	2016年3月	生产中
BM1391	以太坊/Ethash	7纳米	2018年9月	
BM1790		28纳米	2018年3月	
BM1485	莱特币/Scrypt	28纳米	2016年9月	
BM1760	达世币/X11	28纳米	2017年7月	
BM1740	Zcash/Equihash	10纳米	2018年4月	

来源：比特大陆，国金证券研究所

- 耐能专注于低功耗设备端的人工智能芯片：耐能 (kneron)是由一群留美华人于 2015 年成立于圣地亚哥，CEO 刘峻诚博士是毕业于 UCLA，并于 2018 年 7 月延揽前高通多媒体研发部总监李湘村（前展讯，华为，VIVO VP）为其首席科学家，其余团队成员多有 UCLA，清华大学，高通，三星电机，电子，计算机背景，并于 2017 年 11 月，耐能宣布完成超过千万美元的 A 轮融资，阿里创业者基金（Alibaba Entrepreneurs Fund）领投，奇景光电（HIMX, Himax Technologies, Inc.）、中华开发资本（CDIB）、高通、中科创达（Thundersoft）、红杉资本（Sequoia Capital）的子基金 Cloudatlas，与创业邦跟进投资，2018 年 5 月由李嘉诚旗下维港投资（Horizons Ventures）领投的 A1 轮融资，还有最近一轮从 Iconiq capital（Mark Zuckerberg's private fund）拿到的融资。耐能的定位是子系统设备端人工智能的技术提供厂商，现在主打低功耗、轻量级，可压缩/重组（reconfigurable，靠软件重组 CNN，Pooling 运算区块的组合）的 NPU（神经网络处理单元）芯片，专注在智能手机的子系统（NPU IP-KDP300）、物联网（IoT）、智能家居、智能安防（NPU IP-KDP500）设备端市场，机器人，无人机，安防（NPU IP-KDP 700）能耗比可以做到 100mw 到 300mw，最新的一款产品甚至可以到 10mw 以下，但在算力方面可以达到华为海思 Ascend Lite 系列的芯片等级，而耐能另外与 Cadence 的 Tensilica Vision P6 DSP 处理器整合的 KDP 720 NPU 处理器，主要是锁定智能安防与监控。有别于目前市场上主流的云端人工智能，耐能提供创新的设备端人工智能解决方案，可将一部份的人工智能从云端转移到设备端上，进行实时识别与分析推断，不用等到把所有数据经由网络传送到云端后才能处理，并可大幅减轻网络、云端的负担与成本。耐能目前手机加 OEM/ODM 客户可达 6-8 家，主要客户有手机相关的高通，格力，奇景光电，互联网的客户包括搜狗，腾讯，钰创，钰立微，工业计算机客户有研扬，安防客户有大华，苏州科达等。格力已经使用其智财权区块量产，目前一些芯片已经量产。因扩大研发团队及产品线，耐能从 2017 年的获利扭赢转到 2018 年亏损达 400-500 万美元，但 2019 年将有二颗芯片流片（Tape out）。

图表 25：设备端人工智能芯片每瓦算力功耗比

	耐能	耐能	高通	寒武纪	海思	海思	英特尔
产品线	KDP 700	KDP 720	HVX	1M	昇腾 310 Mini	昇腾 Lite	Movidius Myriad X
每瓦算力功耗比 TOPS	8.8	13.2	0.5	5	3	1-5	2-4
时钟频率 Mhz	600Mhz	600Mhz	1.2 Ghz	1.3Ghz	N/A	N/A	N/A
制程工艺	TSMC 16nm	TSMC 28nm	SEC 8nm	TSMC 7nm	TSMC 12nm	TSMC 12nm	TSMC 16nm

来源：各公司，国金证券研究所

- **亿智是有实力又低调的 AI 芯片公司：**的由前全志团队组成的亿智电子科技于 2016 年 7 月在珠海高新区注册成立。同年 10 月在北京设立人工智能（AI）算法研究团队。亿智核心团队是中国最早一批进行 SOC 系统设计的专家，有 20 多年的行业经验，目前亿智在珠海的研发团队已近 100 余人。亿智电子科技已于 2018 年 2 月完成了数千万元天使轮融资，并于 2018 年 8 月由北极光创投领投，达泰资本跟投。亿智的商业模式主要为代理商和大客户提供整套的解决方案。目前，亿智解决方案主要聚焦在视像安防、智能硬件（家电）、汽车电子等方面。2017 年底第一颗 Test Chip 首次流片即成功，2018 年第四季度流片 AI 功能量产版系统级芯片，于 2019 年实现量产出货。亿智在珠海、北京、深圳均设有办公地点，其中珠海为总部，负责芯片设计、算法研究、软件开发等方向。北京负责人工智能 AI 算法的研究。深圳负责方案开发、技术支持、市场与客户拓展等。亿智成立至今，一直坚持 AI 加速、高清显示、音视频编解码、高速数模混合等 IP 的自主研发，这样可以实现更低带宽、更低功耗、更低成本地落地应用产品。特别是 AI 的 IP 的 PPA 指标均优于业界对手。目前已经成长为具备完全自主 AI SOC 产品量产落地能力的人工智能芯片设计公司。亿智凭借在音视频编解码以及 AI 视觉算法方面积累了超过 10 年的领先经验，通过对人工智能需求市场的垂直化、场景化应用研究，在汽车电子应用方面，亿智的产品线具有车牌识别、路牌识别、文字识别的能力，ADAS 智能算法可实现 4 路全景拼接、全景泊车，行车记录仪/智能后视镜/智能中控车机等汽车电子产品应用。目前的夜视后视镜产品，长焦夜视摄像头、短焦行车摄像头，显示车辆油耗、车速、水温等，信息全部手机互联。4G 后视镜提供在线导航、在线音乐、云狗、行车记录，ADAS 安全驾驶辅助系统。亿智的占道抓拍产品，可进行车牌检测，车牌识别，抓拍路段时间规划、黑白名单管理、车辆轨迹显示、后台管理系统多车道实时识别，具备软件能力，团队表示目前的识别成功率达到 95%。志在成为视像安防、汽车电子、智能硬件领域智能化（AI）赋能的全球领导者。

二、AI 让行业应用插上腾飞的翅膀—电子应用篇

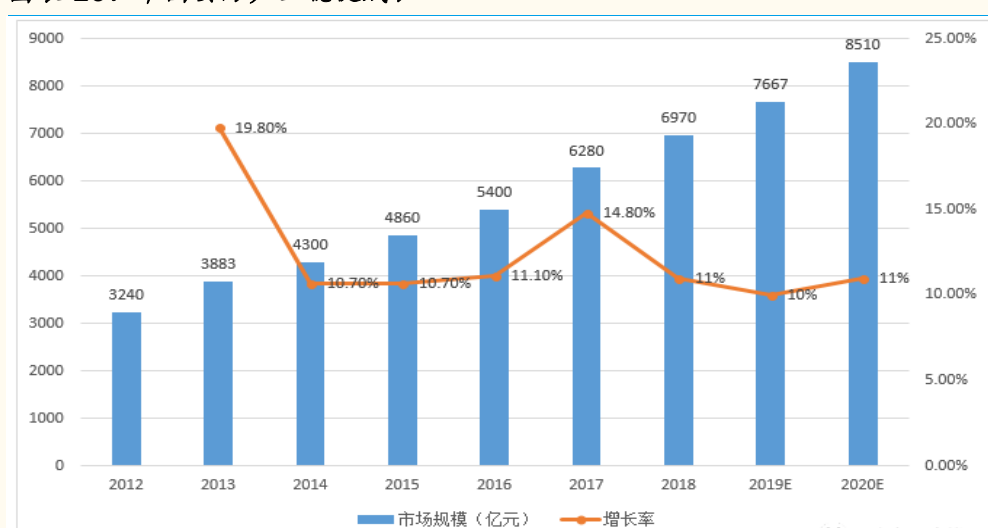
近年来，在计算、大数据、深度学习等技术的综合应用下，AI 技术得以大幅度提升，应用场景也越来越多。2017 年人工智能市场中，计算机视觉位居第一，占比达到 37%，语音第二，达到 22%，而在计算机视觉应用中，安防占比高达 68%，基于计算机视觉开发的人脸解锁、人像美颜功能在 AI 手机应用中渗透率分别高达 75% 与 90%，而语音主要应用在智能手机、可穿戴设备、智能家居中。

1、AI 让安防升级换代，催生更多应用场景

A、稳健发展中的安防产业

根据中安协发布《中国安防行业“十三五”（2016—2020 年）发展规划》， “十三五”期间，安防行业将向规模化、自动化、智能化转型升级，且到 2020 年，安防企业总收入达到 8000 亿元左右，年增长率达到 10% 以上。

图表 26：中国安防产业稳健成长

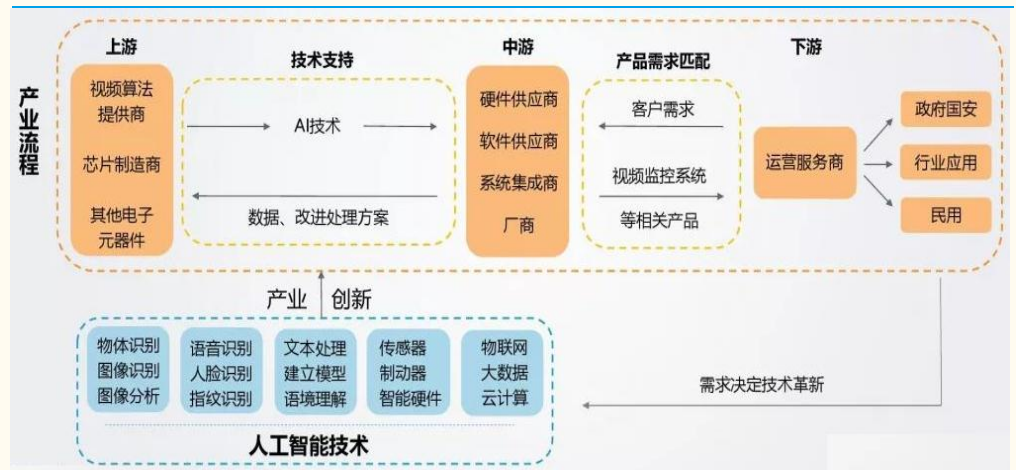


来源：赛智能时代、国金证券研究所

B、AI 技术助力安防行业智能化升级

随着 GPU 等硬件设备的研发和人工智能相关软件算法的成熟，人工智能逐渐被大规模应用各个领域，而安防行业对实时性、准确性要求极高，人工能+安防应运而生。在安防产业链中，硬件设备制造、系统集成及运营服务是产业链的核心，渠道推广是产业链的经脉。未来安防产业的运营升级势在必行，通过物联网、大数据与人工智能技术提供整体解决方案是众多企业的发展趋势。

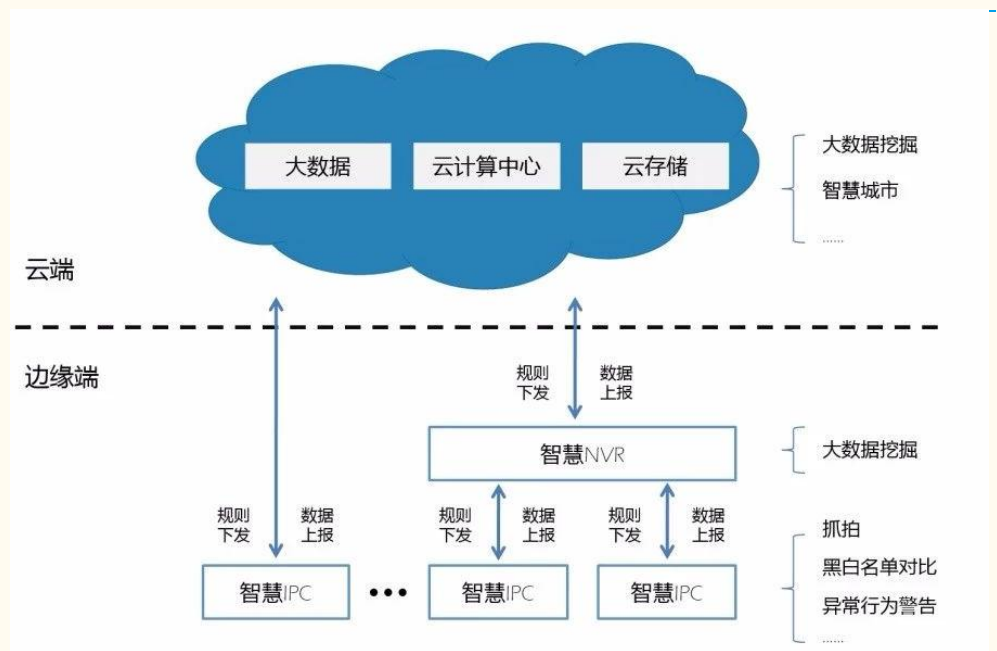
图表 27：AI 融入安防产业链



来源：亿欧智库、国金证券研究所

- **安防是人工智能最理想的落地行业之一。**随着深度学习、大数据等技术持续突破，“人工智能+”已成为我国经济增长的新引擎。人工智能的产业化是“AI+垂直行业”的一场变革。安防行业的海量数据以及事前预防、事中响应、事后追查的诉求与人工智能训练需求及技术逻辑完全吻合，是人工智能最理想的落地行业之一。“AI+安防”已经从概念普及、技术比拼，进入到产品、场景、实战应用和生态构建阶段。
- **智能安防云边结合成为新趋势。**目前安防系统中，常见的中心计算架构问题已经日趋严重，主要体现为网络传输带宽问题、及时性问题得不到有效解决。边缘计算的出现有效缓解了上述问题。云计算聚焦非实时、长周期数据以及业务决策场景，而边缘计算在实时性、短周期数据以及本地决策等场景方面有不可替代的作用。这使得云端及边缘端结合成为新趋势：一些需要集中式处理的计算继续交由大型云计算中心，如大数据挖掘、大规模学习；大量实时的需要交互的计算、分析在边缘节点完成。

图表 28：智能安防云边结合示意图

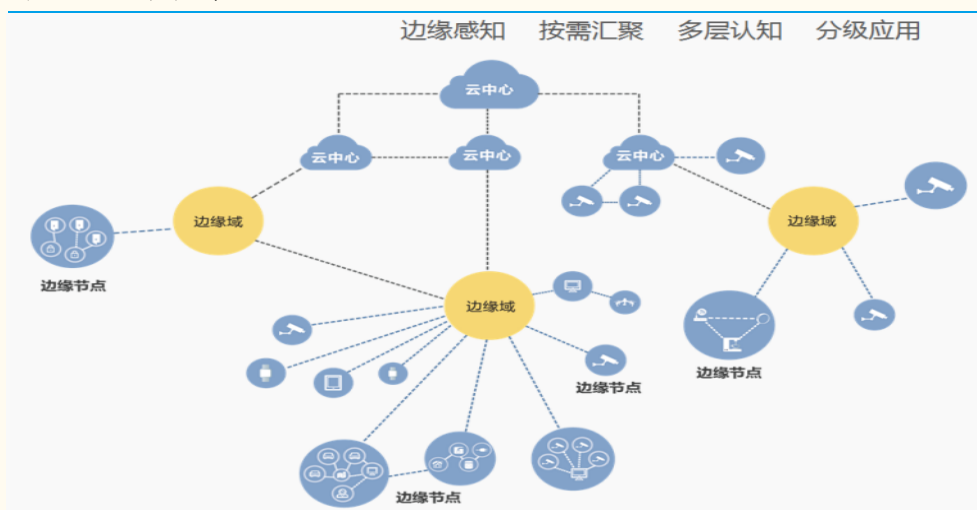


来源：安防+AI 人工智能工程化白皮书、国金证券研究所

- **海康威视提出云边融合 AI Cloud 架构。**海康威视 AI Cloud 架构由边缘节点、边缘域和云中心三个层级构成。边缘节点侧重多维感知数据采集和前端智能处理；边缘域侧重感知数据汇聚、存储、处理和智能应用；云中心

侧重业务数据融合及大数据多维分析应用。数据从边缘节点到边缘域，实现“聚边到域”；从边缘域到云中心，实现“数据入云”。边缘域和云中心可多级多类，根据不同应用，边缘域汇聚的数据和传到云中心的数据在格式和内容上也会不同。边缘域所发挥的作用就像足球“中场”，负责决定在什么时候，将什么样的数据，处理到什么程度，发送到云中心，实现“按需汇聚”。

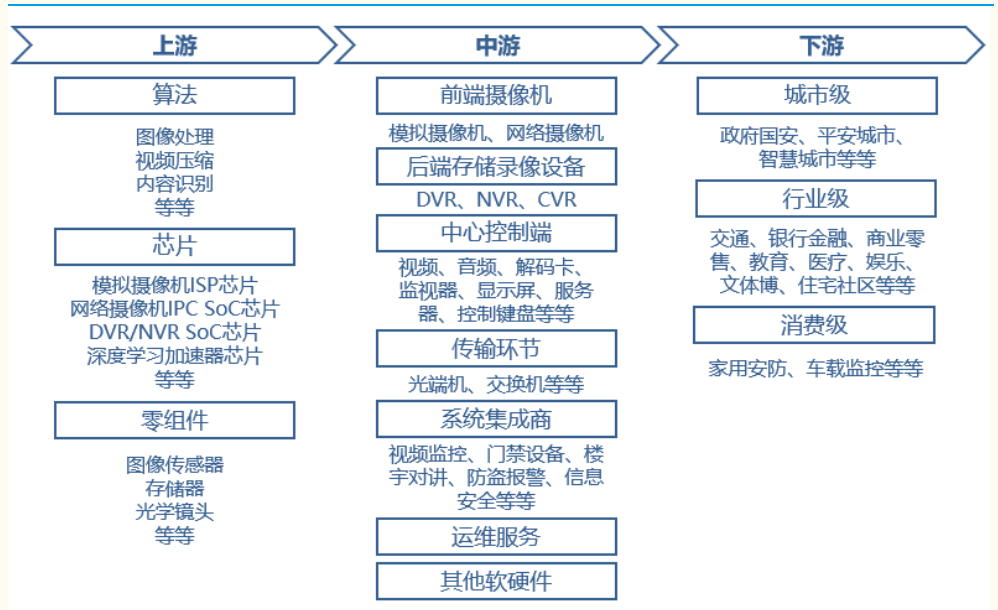
图表 29：海康威视 AI Cloud 核心理念



来源：海康威视、国金证券研究所

- **智能安防产业链：**上游有零部件供应商、算法和芯片供应商等；中游为软硬件设备设计、制造和生产环节，主要包括前端摄像机、后端存储录像设备、音视频产品、显示屏供应商、系统集成商、运营服务商等；下游为产品分销及终端的城市级、行业级和消费级客户应用。上游零部件供应商，代表企业有华为海思、索尼、中星微，还有视频算法提供商 Object Video 等；中游软硬件供应商、系统集成商的主力厂商包括海康、大华等。下游为终端客户，主要涉及到政府、公共行业、民用行业等。在智慧城市的普及下，安防市场容量持续增加，安防产品的智能化程度不断提高，数字监控技术日益成熟，人工智能技术在安防市场上的应用大规模落地，推动传统安防产业进化革新，从而带动了安防产业市场规模的扩大。智慧安防产业链中，上游的算法、芯片和其他零部件供应环节属于技术集成部分，是智能安防产业的发展基础。AI+安防产业链公司众多，系统厂商有海康威视、大华股份，芯片公司有英伟达、华为海思、寒武纪，比特大陆，而 AI 算法公司有商汤科技、依图科技、云从科技等。

图表 30：智能安防产业链



来源：弗若斯特沙利文、国金证券研究所

图表 31：AI+安防产业图谱



来源：艾瑞咨询、国金证券研究所

- **智能安防算法：** 算法环节主要包括图像处理、视频压缩和内容识别三个分类。计算机视觉、深度学习、集成算法等重要的人工智能算法的引入和革新，助力安防产品适应大众不断增长的安防需求。在安防领域智能算法的运用主要体现在，利用无间歇工作的智能算法对视频画面进行监控，以弥补人力无法长时间保持监控状态的缺失；利用生物识别技术解决人脸、指纹、语音等生物特征被盗用的情况；提高安防产品保密和防御级别，确保场所安防布控。此前，算法的基础框架的研发几乎被国外企业垄断，但近年来，伴随人工智能深度学习算法的快速成熟，中国芯片设计公司 and 设备产品厂商都致力于在基础算法上进行改进和优化，培养自己独有的算法技术，部分优质的图像内容识别、算法供应商陆续出现。

图表 32：人工智能安防涉及的主要算法

算法名称	主要内容
CNN卷积神经网络	智能识别技术的基础，模仿人脑对数据抽象化表达，图像识别领域大量应用
DSP算法	数字信号处理，数字时代大部分产品正常运行的基础，典型的有FPGA算法
安全哈希算法	网络安全的基础，防止中间人攻击或网络钓鱼攻击
RSA算法	公开密钥密码体制，确保信息加密和网络安全，视频会议的基础
二叉树算法	H.264和H.265视频编码标准基础，特定码率对高画质数字图像传送

来源：弗若斯特沙利文、国金证券研究所

- **安防监控芯片：**在智能安防监控领域，芯片是硬件设备中成本占比最高的零组件之一，也是安防视频监控设备的核心部件，通过前端摄像机内置人工智能芯片，可实时分析视频内容，检测对象，识别人、车属性信息，并通过网络传递到后端人工智能的中心数据库进行存储。目前，安防视频监控设备中所需要的处理器芯片主要包括网络摄像机中的 SoC 芯片、后端 DVR/NVR 中的 SoC 芯片以及深度学习算法、加速器芯片以及前端模拟摄像机中的 ISP 芯片四种类型。目前，高性能的深度学习算法加速器芯片仍由国外芯片厂商提供，但其余三类处理器芯片已实现了较大程度的国产化替代。

图表 33：安防监控四类主要芯片

监控系统	对应芯片	主要功能	主要供应商
模拟监控系统	前端：ISP芯片	对原始图像信号进行降噪、曝光调整等处理，决定成像质量	富瀚微、NextChip
	后端：DVR SoC芯片	将模拟音视频信号数字化、编码压缩与存储	华为海思、德州仪器、意法半导体
网络监控系统	前端：IPC SoC芯片	主要集成ISP技术和视频编解码技术，同时集成视频分析功能	华为海思、安霸、德州仪器、北京君正、富瀚微、国科微
	后端：NVR SoC芯片	接收网络摄像机的IP码流，进行编解码、存储	华为海思、德州仪器、Marvell

来源：弗若斯特沙利文、国金证券研究所

- **安防用人工智能芯片：**目前安防领域最主流的深度学习芯片方案是 GPU，但 GPU 存在成本、效率、功耗等瓶颈，现已有针对安防产业开发的 FPGA/ASIC 智能芯片，如深鉴科技（去年被 Xilinx 赛灵思以 3 亿美元并购）的 DPU 芯片（FPGA）、北京君正的 NPU 协处理器（ASIC）、寒武纪的 AI 服务器芯片（ASIC 等）。当前安防智能化进程中，算法层面已经接近成熟，宏观架构上云边融合（即云端计算和边缘计算的融合）的理念成为行业共识，虽然具体场景的解决方案层面需要通过大量前期项目进行摸索，但政府部门和大企业客户对于安防智能化的趋势已经有较为充分的认知，而当下对智能化落地的进度最为关键的影响因素便是芯片的成本。为了实现智能化的功能（即运行深度学习算法），安防监控系统的前端和后端设备中需要加入英伟达或是英特尔等国际大厂所设计的 GPU、FPGA 或者 ASIC 加速芯片，与原有的承担图像处理和编解码功能的主处理器芯片一起构成双芯片方案，而采用这些芯片一般要为安防监控设备新增高额成本（2017 年仅前端摄像机中采用的 AI 加速器芯片的成本就高达上百美元），因此导致智能化设备的成本普遍偏高，在很大程度上影响了智能化的大面积应用。未来随着根据应用场景定制的 ASIC 或专用 SoC 智能监控芯片的逐渐成熟（安霸、海思的方案已经在小排量测试，而海康及大华也都在开发自研智能监控 ASIC / SoC 芯片），高性价比的专用芯片对边缘运算及部

分云端通用 GPU 芯片的替代，将使得智能监控设备的成本有望大幅降低，安防监控智能化进程在未来加速落地可期。

图表 34：GPU 在智能安防的应用

Tesla	Tesla 产品主要用于视频监控，作为 PCI 卡形式，安装在 NVR 或者视频服务器，可用于深度学习“训练”或者“推理”之用。 • V100 - 高价格高性能，主要用于深度学习研发之用。 • P100 - 中等价格及性能，用于数据训练。 • P40 - 中等性能及价位，用于数据训练及推断，使用在性能强劲的服务服务器上。 • P4 - 价位中低，主要用于“推理”（Inferencing）之用，在 NVR 服务器或者其他视频分析、深度学习应用中。	
Jetson	Jetson 主要应用于前端摄像机，实现 AI 功能。 • TX1 - Maxwell architecture • TX2 - Pascal architecture	
GTX Quadro	• GTX - Graphics Cards • Quadro - Advanced Graphics Cards	
DGX	The DGX-1 用于复杂情况下大数据训练，深度学习。其性能等价于 1 台至强处理器配置了 8 个 Tesla 的 V100 模块。	

来源：安防+AI 人工智能工程化白皮书、国金证券研究所

图表 35：AI+ 安防典型落地应用产品

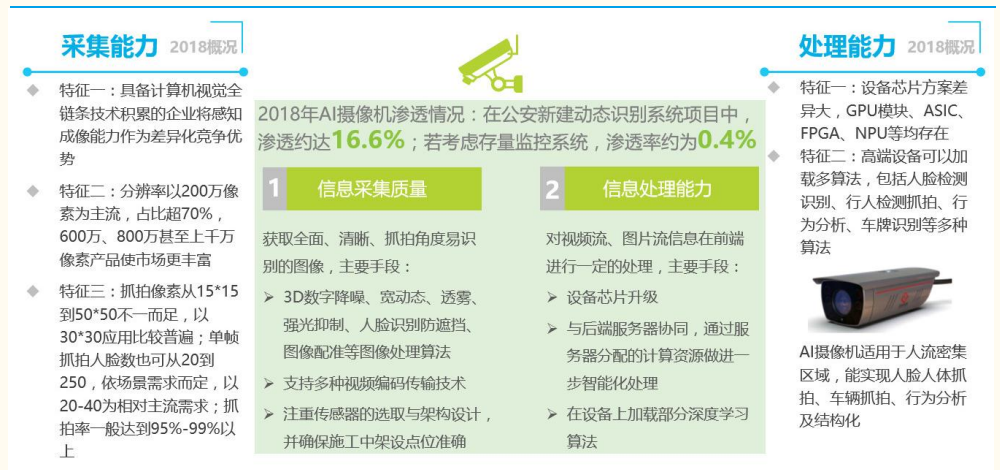
厂商	产品	图片
	海康 深眸 摄像机，基于 NVIDIA Jetson TX1 的最新双目智能摄像机产品。该产品由双目摄像机（基于人眼工作原理）与高性能 GPU 模块组成，实现人员的多种智能分析功能。基于深度学习技术，支持人员行为、客流统计、人体属性、人脸识别等多种智能检测，适用于公共场所，银行，商场，景区等环境。	
	宇视大容量分布式的云结构化智能分析服务器（代号“昆仑”），在 4U 的高度上支持了 80 颗 NVIDIA Jetson TX1 处理器，可以并发处理 640 张/秒的人脸识别、160 路的人员计数、80 路人车物的结构化分析，芯片和算法基于深度学习，准确性很高。	
	“Deep Sense 睿智”系列视频结构化服务器，采用 NVIDIA Tesla P4 GPU 作为核心处理负载，功耗低，性能强。服务器采用深度学习算法，最多可支持 192 路全清视频实时结构化分析，相较于市场上同类别的产品，视频处理能力将提高 50 倍以上。	

来源：公司公告，国金证券研究所

C、智能安防才刚进入初级阶段，AI 占比低于 1%，未来大有可为

- 目前 AI 在安防领域仍处在发展阶段，具有较好的发展前景：《2018 年中国 AI+ 安防行业发展研究报告预告版》的报道中指出，人工智能在安防领域的应用早在 2012 年就已经取得一定成就，但是无论是人脸识别、结构分析等其他定制化服务，也就 2016-2017 年左右才稍有起色。
- 安防领域智能摄像头应用率 1%：智能摄像头无疑在智能安防领域充当着最基础也最重要的眼睛作用，单单就智能摄像头来说，我国一年有 5000 万个摄像头需求。实际上我们只有 50 万不到智能摄像头进入到安防领域应用，用率也就刚刚达到 1%而已。
- AI 摄像机在公安动态识别系统渗透率 0.4%：2018 年市场较为理性，从铺设速度看，AI 摄像机在公安动态识别系统项目中渗透约达到 16.6%，而若考虑全国近 2300 万路现存公安监控摄像头，则渗透率约 0.4%，从设备能力看，符合应用场景可用指标为核心准则。

图表 36：2018 年中国 AI 摄像机建设逻辑与发展概况



来源：艾瑞咨询、国金证券研究所

- **智慧安防 22.6%复合增长**：在安防领域，智能化逐渐成为行业转型升级的方向，智能安防在安防行业占比也逐步提升。据中商情报网数据，2018 年中国安防行业市场规突破六千亿元，其中，智慧安防行业市场规模约 2001 亿元，预计到 2022 年智慧安防行业市场规模将达到 4514 亿元，2018-2022 年均复合增长率达到 22.6%。

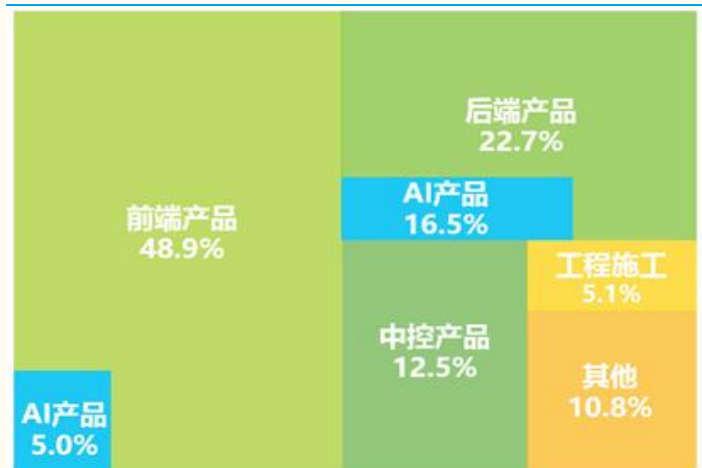
图表 37：中国智慧安防市场规模预测



来源：弗若斯特沙利文、国金证券研究所

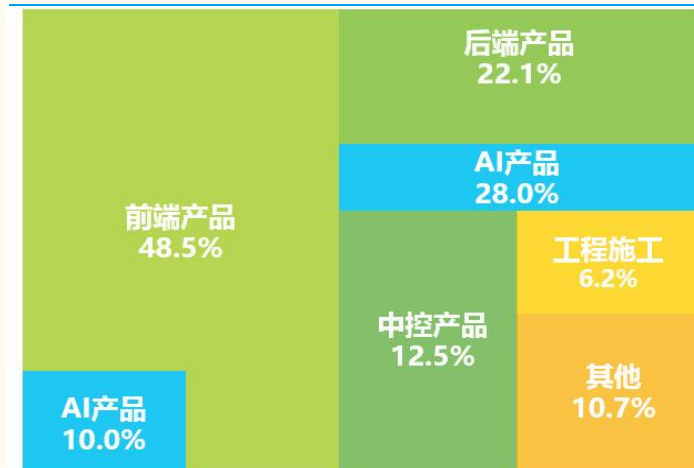
- **2018 年安防 AI 产品营收在前端营收占比中增长 100%，后端增长 70%**：2017 年和 2018 年人工智能开始真正落地安防场景，其中 2018 年更是飞速发展的一年。按前端产品、后端产品、中控产品、工程施工和其他划分的营收结构在这两年中基本没有发生变化，而 AI 产品的营收在前端营收占比中增长了 100%，在后端营收占比中增长了 70%（注：图中蓝色区域 AI 产品占比为前端产品或后端产品中的占比，并非整体营收占比。），说明在符合已有市场需求结构的基础上，AI 产品越来越受到市场的重视与欢迎，虽然整体占比还很小，但在公安部门“新建一批、利旧一批、淘汰一批”的准则下，2019 年 AI 产品在安防厂商视频监控业务中的占比增速不会低于 2018 年。

图表 38：2017 年安防厂商视频监控业务收入占比



来源：艾瑞咨询、国金证券研究所

图表 39：2018 年安防厂商视频监控业务收入占比



来源：艾瑞咨询、国金证券研究所

- **AI、安防企业有望整合：**新兴 AI 企业比部分安防上市企业要更早推出智能安防产品，且在技术方面更领先一些，但 AI 需要通过数据进行机器学习，AI 创企在数据方面稍逊，若没有数据支撑难以实现更大的突破，而大量数据都在政府国安手里，若没有数据支撑也难以有更大的突破；大型安防上市企业尽管早已提出结合深度学习、AI 技术的研发策略，且拥有大数据和综合指挥平台优势，但大部分在 2017 年下半年才真正推出智能安防产品。这也就形成了资源与技术不能很好地匹配，有资源的手里缺 AI 技术，而 AI 产品难以规模化应用。不过，存在问题并不意味着智能安防达到了瓶颈。从 2017 年中国智能安防行业下游需求分析中可以看出，平安城市、智能交通是智能安防崛起的主因，目前这些领域均处于快速发展期，此外，智能安防企业在物流、旅游、能源等领域开辟了新的应用。未来智能安防无疑拥有非常大的发展空间。

D、安防 AI 面临的三大挑战

安防 AI 的发展主要面临以三个方面的挑战：算法的场景适应能力、大规模应用的技术与经济可行性和面向业务应用的解决方案。

- **算法的场景适应能力：**以人脸识别为例，目前绝大多数厂商的算法来源于对人脸静态图片的学习训练——比如标准证件照，对所采集的人脸照片的清晰度、光照、角度、妆容都有着严格的要求，但在实际监控场景中，除了室内个别环境外，绝大多数现场环境采集的人脸图片无法满足这样的要求，造成人脸识别精度的大幅度衰减。如国内某知名厂商的人脸识别算法在静态人脸比对测试中总是名列前茅，但其复杂环境下动态人脸识别的效果就比较差，特别是在偏转角度较大（超过 30°）、化妆、戴墨镜、戴口罩等情况下识别率会大打折扣。
- **大规模应用的技术与经济可行性：**目前计算机视觉的大规模运算还主要依赖 GPU 和 CPU 的算力，以人脸识别和视频结构化分析为例，一般利用 GPU 做视频图像处理与特征解析运算，用 CPU 做人脸特征比对运算。在复杂动态人脸采集环境下，1 台 8 卡 GPU（TESLA P4）+2 个 16 核 CPU 服务器每秒可支持 80 路 1080P 高清视频实时动态解析和 100 万库的实时动态比对运算，硬件成本折合 4000 元人民币/路，电力消耗（服务器自身耗电+散热制冷耗电）折合每年 800 元/路。如果采用 800 万像素的监控图像，硬成本会再增加 3 倍，这还没有计算算法与应用软件系统的费用，如此高昂的建设与运行费用很难得到大规模的应用推广。
- **面向业务应用的解决方案：**AI 是一项技术，从 AI 技术到客户价值，需要一整套的产品与解决方案，向客户最终交付的是应用，尤其是面向业务的应用软件，不同行业、不同场景、不同用途的业务应用都不尽相同，需要针对性的集成与应用软件开发，否则无法规模化推广。而目前情况来看，无

论是 AI 算法厂商、安防产品商还是系统集成商，都不具备全行业应用软件的开发能力。

E、AI+安防使用技术

(a) 视频结构化技术

- **一是目标检测**——从视频中提取出前景目标，然后识别出前景目标是有效目标（如：人员、车辆、人脸等）还是无效目标（如：树叶、阴影、光线等）。在目标检测过程主要应用到运动目标检测、人脸检测和车辆检测等技术。
- **二是目标跟踪**——实现特定目标在场景中的持续跟踪，并从整个跟踪过程中获取一张高质量图片作为该目标的抓拍图片。在目标跟踪过程中主要应用到多目标跟踪、目标融合以及目标评分技术。
- **三是目标属性提取**——对已经检测到的目标图片中目标属性的识别，判断该目标具有哪些可视化的特征属性，例如人员目标的性别、年龄、着装，车辆目标的车型、颜色等属性。目标属性提取过程主要基于深度学习网络结构的特征提取和分类技术。

(b) 情感计算技术

情感计算在安防领域具有广泛的应用前景。计算机通过对人类面部表情、语音表情、姿态表情、生理表情和文本情感的获取、分类和识别，可以及时获取目标对象的情感变化，并对异常危险行为提出预警，实施相应的应对措施。

- **脸部表情**：国际著名心理学家 Paul Ekman 和研究伙伴 W. V. Friesen 对人脸面部表情作了深入的研究，通过观察和生物反馈，于 1976 年描绘出了不同的脸部肌肉动作和不同表情的对应关系，即面部表情编码系统 FACS。FACS 根据人脸的解剖学特点，将人脸划分成若干既相互独立又相互联系的运动单元，分析了这些运动单元的运动特征及其所控制的主要区域以及与之相关的表情，并给出了大量的照片说明。FACS 是如今面部表情的肌肉运动的权威参照标准，也被心理学家和动画片绘画者使用。为满足视频信息传输的需要，人们进一步将人脸识别和合成的工作融入到视频图像编解码之中。典型如 MPEG4 V2 视觉标准，其中定义了 3 个重要的参数集：人脸定义参数、人脸内插变换和人脸动画参数。当前人脸表情处理技术研究的热点多侧重于对三维图像的更加细致的描述和建模。通常采用复杂的纹理和较细致的图形变换算法，达到生动的情感表达效果。在此基础上，不同的算法形成了不同水平的应用系统。
- **语音理解**：目前，国际上对情感语音的研究主要侧重于情感的声学特征的分析。中国科学院自动化研究所模式识别国家重点实验室的专家们针对语言中的焦点现象，首先提出了情感焦点生成模型。这为语音合成中情感状态的自动预测提供了依据，结合高质量的声学模型，使得情感语音合成和识别达到了实用水平。
- **姿态变化**：针对肢体运动，科学家专门设计了一系列运动和身体信息捕获设备，例如运动捕获仪、数据手套、智能座椅等。国外一些著名的大学和跨国公司，例如麻省理工学院、IBM 等则在这些设备的基础上构筑了智能空间。也有人将智能座椅应用于汽车的驾驶座上，用于动态监测驾驶人员的情绪状态，并提出适时警告。意大利的一些科学家还通过一系列的姿态分析，对办公室的工作人员进行情感自动分析，设计出更舒适的办公环境。
- **生理识别**：不同的生理信号的特征模式也是情感识别的重要依据之一。人的生理信号比起面部表情和语音，识别难度更大，所以目前生理模式的情感识别研究还处于初级阶段。哪些信号可以转化为情感参数、信号各个方面的权重、比例应该是多少，这些都还需要进行进一步的研究和探索。
- **文本情感计算**：文本情感计算是自然语言处理的一个研究分支，其工作展开的关键在于情感特征提取和情感分类方法的不断进步优化。尽管经过了大量研究，文本情感计算取得了很大的进展，但整体仍处于探索阶段，存在

一些亟待解决和研究的问题：缺乏规范统一的实验语料和词典。目前针对语言规律和句子语义成分的分析问题，还没有成熟的解决方案。

- **多模态的情感计算**：虽然人脸、语音、姿态、生理、文本均能独立地表示一定的情感，但只有实现多通道的情感信息采集，才能实现完整的情感识别。这通常要求系统集自然语言、语音、手语、人脸、唇读、头势、体势等多种交流通道于一体，进行综合的采集、分析和识别。目前，多模态技术正在成为情感计算的研究热点，实现情感的多特征融合，能够有力地提高情感计算的研究深度。美国麻省理工学院、日本东京科技大学、美国卡内基·梅隆大学均在情感机器人和情感虚拟人的研究领域做出了较好的演示系统。中科院自动化所模式识别国家重点实验室也已将情感处理融入到了多模态交互平台中，结合情感语音合成、人脸建模等技术，构筑了栩栩如生的情感虚拟头像。

(c) 大数据技术

大数据技术为人工智能提供强大的分布式计算能力和知识库管理能力，是人工智能分析预测、自主完善的重要支撑。

- 一是海量数据管理被用于采集、存储人工智能应用所涉及的全方位数据资源，并基于时间轴进行数据累积，以便能在时间维度上体现真实事物的规律。同时，人工智能应用长期积累的庞大知识库，也需要依赖该系统进行管理和访问。
- 二是大规模分布式计算使得人工智能具备强大的计算能力，能同时分析海量的数据，开展特征匹配和模型仿真，并为众多用户提供个性化服务。
- 三是数据挖掘是人工智能发挥真正价值的核心，利用机器学习算法自动开展多种分析计算，探究数据资源中的规律和异常点，辅助用户更快、更准地找到有效的资源，进行风险预测和评估。

F、AI+安防场景应用

安防与人工智能技术结合是大势所趋，“AI+安防”朝着视频结构化、数据可视化、安防移动化、云防立体化演进。但在发展过程中软硬件问题尚待解决，只有妥善解决这些问题，才能满足市场需求，提升整个安防领域的智能化水平，从而推动安防产业的升级换代。

- **公安行业场景应用**：公安行业用户的迫切需求是在海量的视频信息中，发现犯罪嫌疑人的线索。前端摄像机内置人工智能芯片，可实时分析视频内容，检测运动对象，识别人、车属性信息，并通过网络传递到后端人工智能的中心数据库进行存储，再利用强大的计算能力及智能分析能力，对嫌疑人的信息进行实时分析，为案件的侦破节约宝贵的时间。其强大的交互能力，真正成为办案人员的专家助手。
- **交通行业场景应用**：在交通领域，利用人工智能技术，可实时分析城市交通流量，调整红绿灯间隔，缩短车辆等待时间，提升城市道路的通行效率。城市级的人工智能大脑，实时掌握着城市道路上通行车辆的轨迹信息，停车场的车辆信息，以及小区的停车信息，合理调配资源、疏导交通，实现机场、火车站、汽车站、商圈的大规模交通联动调度，提升整个城市的运行效率，为居民的出行畅通提供保障。
- **智能楼宇场景应用**：在智能楼宇领域，人工智能是建筑的大脑，综合控制着建筑的安防、能耗，对于进出大厦的人、车、物实现实时的跟踪定位，区分办公人员与外来人员，监控大楼的能源消耗，使得大厦的运行效率最优，延长大厦的使用寿命。
- **工厂园区场景应用**：在工厂园区场所，利用可移动巡线机器人，定期巡逻，读取仪表数值，分析潜在的风险，保障全封闭无人工厂的可靠运行，真正推动“工业4.0”的发展。

- **民用安防场景应用**：在家庭安防中，当检测到家庭中没有人员时，家庭安防摄像机可自动进入布防模式，有异常时，给予闯入人员声音警告，并远程通知家庭主人。而当家庭成员回家后，又能自动撤防，保护用户隐私。夜间期间，通过一定时间的自学习，掌握家庭成员的作息规律，在主人休息时启动布防，确保夜间安全，省去人工布防的烦恼，真正实现人性化。

G、2018 年安防 AI 发展情况

2015 年以来，旷视、依图、商汤、云从等人工智能算法团队率先将人脸识别技术应用到公安科技中，帮助公安机关侦破了很多旧案积案，抓获了不少长期潜逃的案犯，引起了轰动，安防 AI 应用才始见庐山真面目。经过 2016、2017 两年的认知、试用和推广，2018 年开始进入一轮人脸识别初级阶段应用的小高潮。

(a) 2018 年安防 AI 发展特点

2018 年安防 AI 市场发展的特点可概括为四句话十六个字：“看脸时代”、“风正起时”、“群雄并起”、“两营对垒”。

- **“看脸时代”**：2018 年谈到安防 AI 就离不开人脸识别，一方面是因为一流算法团队的人脸识别技术已经超越人眼的辨别能力且得到了众多现实场景的检验，技术上完全达到了实用化阶段；另一方面人脸识别的应用范围和场景比较宽，如公安追逃、身份核查、案件侦查和所有需要身份鉴别认证的场景都用到人脸识别技术，人脸识别是目前最成熟、最广泛的安防 AI 应用。
- **“风正起时”**：作为全国性的安防行业最大的工程项目——雪亮工程和天网工程的建设重点都已经从原来点位建设转移到 AI 应用上来，以及其它垂直应用领域的安防系统建设也都在关注 AI 应用，从而催生出了未来巨大的安防 AI 产品与应用市场，安防 AI “风”已经吹起，AI 算法团队、AI 芯片厂商、传统安防厂商、系统集成商、传统 IT 厂商都吹响了“集结号”，意欲掘金未来庞大的安防市场。
- **“群雄并起”**：受安防 AI 风口的驱动，除了目前国内几家知名的视觉 AI 算法团队外，传统安防厂商、新 AI 创业团队、垂直行业应用集成商也都在投入人力物力来开发自己的 AI 技术和产品，形成了群雄并起的局面。
- **“两营对垒”**：2018 年以前，率先在安防 AI 市场刮起旋风的是依图、商汤、旷视、云从等 AI 新锐。2018 年，传统安防厂商布局 AI 技术产品，在市场竞争层面主要是 AI 算法专业团队与传统安防厂商两大阵营的对垒，在到底是 AI+安防还是安防+AI 的问题上各自演绎。从当前竞争的初步结果来看，AI 算法团队在算法精度和云端算力上具有较为明显的碾压优势，传统安防厂商在 AI 边缘产品——特别是人脸检测抓拍摄像机产品化和安防 AI 行业应用方面处于领先地位。

(b) 安防 AI 技术、产品应用落地情况

目前安防 AI 技术主要集中在人脸识别与人、车、物视频结构化解析两个方面，其中人脸识别技术已经成熟，人脸静态识别准确率已经达到 98%以上，人脸动态识别（利用监控摄像头进行实时人脸识别）的准确率也已达 90%以上（百万人脸比对库以内），已经满足绝大多数场景下的应用需求。人/车/物视频结构化解析在机动车特征解析上的准确率已达到 95%以上，而在人体特征和非机动车识别上准确度与实战要求还存在一定的差距，特别是在大场景、多目标、复杂光照条件下，准确度还难有保证。

在 AI 产品上，云端产品主要是基于 GPU 运算的人脸识别与视频结构化解析服务器，受限于前端芯片算力的影响，边缘运算产品主要是人脸抓拍摄像机和支持小规模人脸库的人脸门禁终端机，其中以人脸抓拍摄像机和人脸闸机的市场销售已经初步放量。在落地应用场景上，以雪亮工程、天网工程等社会公共安全为主要应用场景，占安防 AI 市场的 80%以上，但目前还仅仅是小规模

试点和少量项目落地应用，还远没形成规模化应用的格局。另外，人脸门禁在政府机关、对安全有较高要求的企事业单位、居民小区、校园等也有试点场景落地。

H、AI+安防发展趋势

(a) 从最初的“谁是你”，到“你是谁”

近些年因为 AI 技术的出现，视频监控层面从最初的数字化、网络化、高清化向智能化快速转变，安防行业达成的共识是从“看得清”到“看得懂”，从“看视频”到“用视频”进行过渡。最初摄像头解决的是图像传输和处理，网络、互联网普及后开始进行万物互联，无论是依托大数据还是云计算技术来判别人与人、人与物之间的关联性。但最近在近期 AI 或安防上市企业发布会或展会上，我们看到一些多功能 AI 产品面世。

比如感知型 IPC，在车辆识别方面精细化程度从车牌、车型、车身颜色、品牌到驾驶员是否系安全带、是否打电话、是否开启遮阳板，再到车辆通过的时间、道路名称、车辆行驶方向进行车辆轨迹分析，包括人员以图搜图，整理你的行为轨迹。AI+安防要解决的将不再是人与人之间、人与车之间的结构联系，而是能自主判断“你是谁”，相信在不久的将来人工智能技术将会取代众多传统的安防技术，整个安防行业的发展已经到了比拼核心技术的关键节点。

(b) 投资并购，完美结合

因为 AI 技术和供应链资源参差不齐的事实客观存在，我们发现，不少传统安防企业通过投资/并购 AI 公司的方式来弥补各自短板，向完美结合体演进。2008 年，全球视频监控领域三大巨头 AXIS、BOSCH、Sony 宣布合作，安防领域内企业开放、合作呼声渐高。2015 年英飞拓收购藏愚科技，高新兴收购创联电子，东方网力收购华启智能、动力盈科等 6 家新技术企业；2016 年海康威视收购英国报警企业 SHL，佳都收购华之源，东方网力再次投资爱耳目、数智源等等。我们看到安防市场格局逐渐明朗，已经形成了“两超多强”的格局，海康威视和大华股份领跑市场，东方网力、佳都、苏州、汉王等第二梯队企业奋起直追。当有了新技术的运作，尤其在 AI 应用正式落地安防之后，投资或收购 AI 技术公司成为传统安防企业最有效创新升级的方式，而 AI+安防企业也将面临行业集中度提高，市场份额向头部集中的洗礼。

(c) AI+安防企业的业务逻辑是先 TO B

- 公安部门属于比较高端的市场，但从历史来看，都是从政府这边去切入，然后成熟之后再往民用方向普及，随着 AI 成熟度的进一步加强，很多安防产品已经开始逐步下沉到更多细分的民用场景，如社区、学校、工业园区、智能家居等。
- 受益于安防领域深度学习算法的快速发展，智能安防已经得到了越来越广泛的应用。在 AI+安防时代，面对安防视频产品下游的需求，运营服务将有较大的市场空间，这也将成为我国安防产业未来的发展方向。以人脸识别为例，可广泛应用于公安、零售、教育、金融、医疗等行业；除此之外，未来也可以尝试新兴的场景，如智慧景区，完成物品遗留检测、客流统计以及智能巡检等；智慧商业，对客流量统计以及人流密度检测等。
- 而对于具体做法，AI 初创公司未来需要选择和深挖垂直行业解决方案。做解决方案的好处就是能够端到端地了解行业应用，从系统化、全局化的角度才能真正理解业务对于技术的需求，重新定义问题才有可能更好解决问题。
- AI 对安防领域的改造才刚开始，这一点从几家传统安防企业的 AI 产品落地情况可以看出。2017 年是 AI+安防企业正式落地应用的第一年，具备深度学习算法或 AI 产品开始在政府国安得到小范围运用，包括提供个性化定制解决方案，随着未来技术的成熟，以及国家政策的推动。在原有安防场景里，AI+安防产品大规模运用最多只有 5 年左右时间，这些年弯道超车的

AI 初创企业，以及积极拥抱 AI 新技术的传统安防企业的行业格局变化值得期待。

(d) 安防 AI 应用场景落地需要算法、算力、产品与应用共同推动

安防 AI 应用场景落地需要算法、算力、产品与应用四个方面的合力才能推动，其中算法仍然是核心。随着算法、算力、芯片的快速发展和泛 AI 业务场景的深度应用，行业将迎来泛 AI 时代，预计到 2022 年，AI 将成为几乎所有安防监控系统的必备能力。

- **算法层面：**除了需要不断提升应用场景的环境适应能力外，还需要在保证算法精度的前提下对模型的持续精炼，以降低对算力的消耗，从而控制硬件资源的投入、系统建设成本和电力消耗，为大规模泛智能化应用创造条件。因此谁掌握了深度学习的核心技术，具备算法的持续优化迭代能力，谁就会在未来的市场中占据主导地位。虽然安防 AI 市场的参与者都宣称具有自主开发算法的能力，但绝大多数是利用开源深度学习框架进行二次开发和模型训练，本身并不掌握核心技术，而开源框架多数来源于美国的公司或研究机构，未来在算法的迭代能力上将面临极大的考验。
- **算力层面：**无论是云端算力和边缘算力将会得到大幅度提升，英伟达不断刷新其推理型 GPU 芯片的运算性能和能耗比，为云端运算提供强大的算力支撑。Intel、NXP、AMD、GOOGLE、苹果、IBM、ARM、高通、博通、三星等国际知名芯片厂商以及华为、寒武纪、比特大陆，深鉴、地平线、中星微等国内芯片厂商都在 AI 芯片上持续发力，未来必将呈现精彩纷呈的局面，让每个安防设备都具备强大的 AI 能力。
- **产品层面：**目前云端产品主要以 GPU 服务器为主，未来会出现集成 ARM、高通、华为海思、FPGA 等芯片架构的云端 AI 运算服务器。随着高性能、低功耗、低成本 AI 芯片的推出，未来最值得期待的将会是边缘节点泛智能产品，如可同时运行多种 AI 算法模型的 AI 摄像机、智能 NVR 等产品，AI 将会是安防监控设备的标配能力。
- **应用层面：**AI 来源于深度学习与大数据，AI 同时也产生大数据，如基于人脸、人体、车辆特征的时空轨迹大数据。AI 的核心是算法，本质是大数据，特别是在公安行业，AI 给警用大数据注入了新鲜血液，有了 AI 才算真正有了警用大数据。因此未来在安防 AI 应用层面，不仅仅是目前简单的人脸比对、黑名单布控、身份甄别等简单化应用，而是具备更大可挖掘价值的 AI 大数据应用，也将会涌现出一批专注于面向各行业的 AI 大数据挖掘和应用的企业。

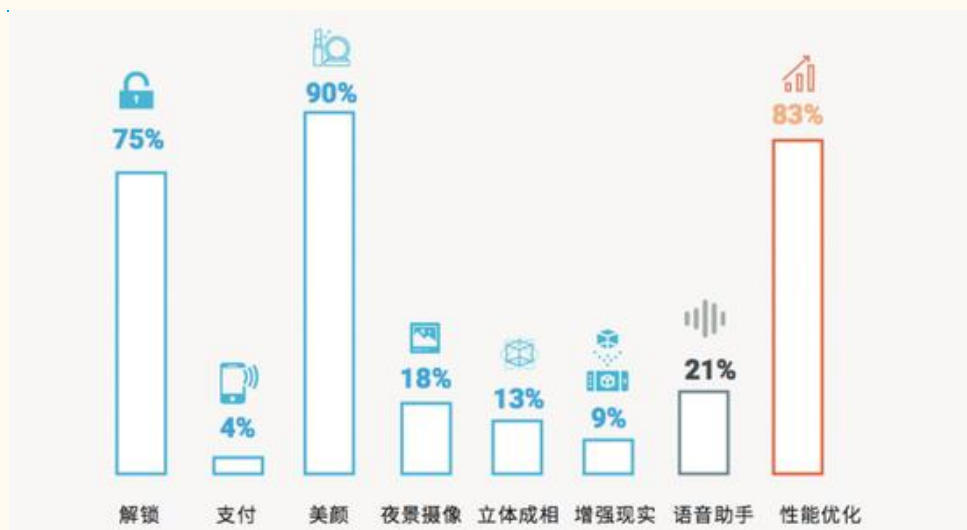
2、智能手机 AI 应用势如破竹

随着移动互联网的发展，智能手机已经从单纯的通讯工具成为了人们生活和工作的入口，智能手机的快速普及也带动了诸多产业的发展，AI 也瞄准了智能手机这个大蛋糕市场，同时手机厂商意识到了 AI 给智能手机行业带来的价值，AI 手机应运而生，从 2017 年开始，“AI 手机”如雨后春笋，2018 年主流手机厂商更是在各自的旗舰机型中全面引入 AI 技术，甚至搭载 AI 芯片的手机正在成为市场的主流。

A、AI 视觉给手机带来新方向

目前手机 AI 应用较多的还是体现在摄像头上，无论是拍照、解锁，还是认知、优化、分析等都是基于这个“AI Camera”。最浅层的 AI Camera 即人脸解锁，再往上则可能是 AI 的智能优化。视觉应用是手机 AI 应用中最主要的诉求，从图中我们可以明显地看到，目前主流的手机应用大多与视觉应用相关（蓝色柱状图），可以说，提升视觉技术在当前是改善用户体验的最直接和最有效的手段。目前，计算机视觉技术在手机种的应用主要可以分为三大方向：识别与认证、AI 摄影、3D 感知。AI 手机基于计算机视觉开发的人脸解锁、人像美颜功能在手机应用中渗透率分别高达 75% 与 90%。与此同时，AI 视觉需求的日益强烈对产业链因此提出了更高的要求，应用、算法、解决方案、硬件将会环环相扣、缺一不可。

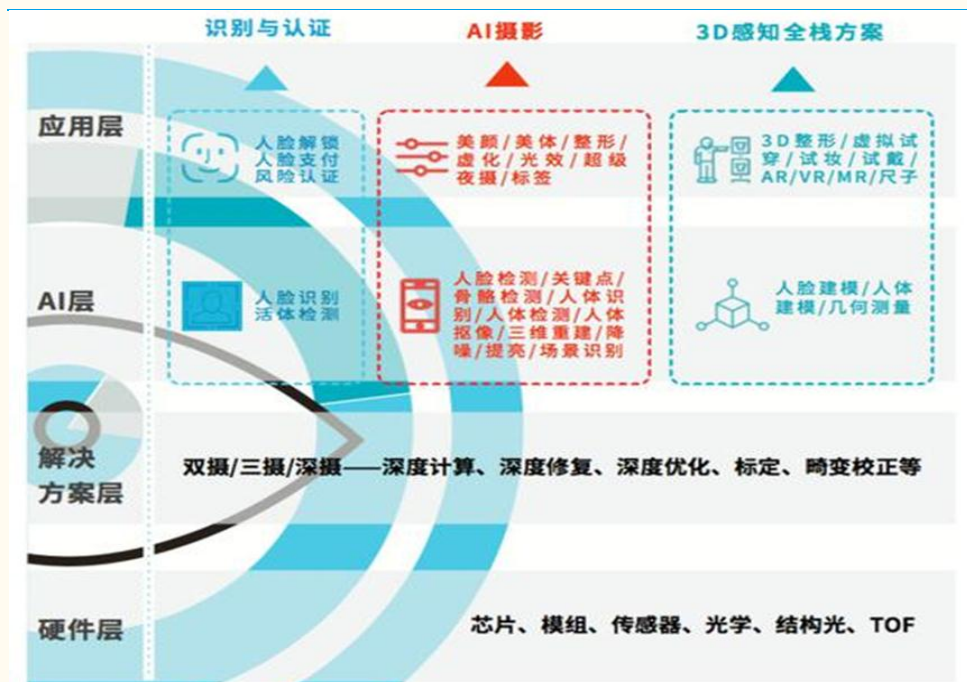
图表 40：AI 技术在手机应用中的渗透率



来源：IDC、国金证券研究所

想要给用户带来超预期的视觉体验，需要产业链的协同发展。视觉技术的推陈出新与产业链的发展紧密相连。IDC 将计算机视觉技术对安手机的影响分为四个层面，每一个层面的创新既相对独立又相互促进。利用新器件、新算法打造新方案探索新应用，改善用户体验，提升用户粘性是 AI 手机快速普及的重要助推力。

图表 41：计算机视觉应用全景图

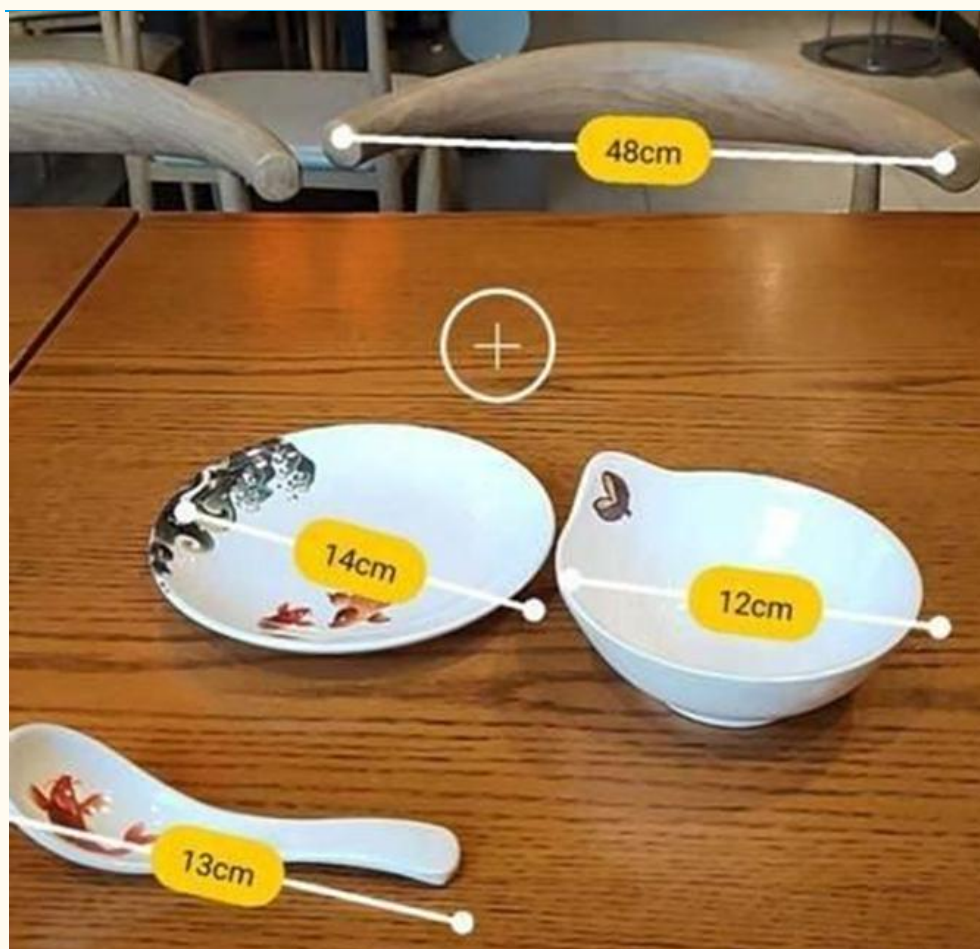


来源：IDC，国金证券研究所

- **AI 视觉方案对手机产业链的革新需求：**随着人工智能视觉需求的日益强烈，AI 视觉解决方案整合将加速，对产业链提出了更高的要求。计算机视觉技术将成为手机产业链的关键环节，可见计算机视觉技术在手机领域的应用正在趋于成熟，但就技术本身来讲还有太多应用场景尚未被开发出来。当前市场关注度、渗透率及技术采用度相对较高的应用场景，如人脸识别、物体识别、物体检测等还停留在较为基础的物体探测阶段，在更具体的事件检测、更灵活的人机交互及更复杂的信息重组、自主行为等方面的应用明显不足。

- **AI 视觉产业生态的整合正在发生：**以高通、联发科、紫光展锐等公司为代表的芯片厂商，以及以奥比中光、艾迈斯、舜宇光学、欧菲科技为代表的光学器件和模组厂商正在通过战略合作资本等形式与上层算法和应用公司进行深度合作。未来，无论是计算机视觉技术本身还是“AI 手机”行业的市场前景，都具有非常大的想象空间。
- **AI 算法加持，手机摄像头实现 AR 测量：**2018 年 11 月，OPPO R17 Pro 正式发售，其搭载的“AR 测量”令人瞩目，成为首款实现景深距离测量的安卓手机。“AR 测量”就是可以直接借助手机 AR 代替尺子、量角器等工具测量实际物体长度、角度、距离、面积的应用。开启“AR 测量”创意功能，用户可以测量物体的角度、长度、测距离以及面积。那么，OPPO AR 测量的精度如何呢？OPPO AR 通过景深距离的测量，透过手机屏幕，用户可以直观看到被标记物体与自己的距离，其绝对精度为 1%，相对精度为 0.5%。

图表 42：OPPO R17 Pro 的 AR 测量



来源：OPPO，国金证券研究所

B、解放双手，AI 语音助手发展在路上

语音助手是“人工智能”应用的一种具体表现方式，“人工智能”最初的形态就是语音助手；特别是 AI 语音助手可以通过不断学习和整合服务，从而变得更智能和更实用。AI 语音助手应用在智能手机中，可以实现智能对话和即时问答等智能交互，帮忙手机用户解决各种问题。因此，智能手机搭载 AI 语音助手，也是希望智能手机变得更加智能。

- **2023 年，智能手机 AI 助手渗透率将达 90%：**AI 语音是仅次于 AI 视觉的在智能手机中第二大应用，2017 年，全球销售的智能手机中，只有 36.6% 的智能手机配置 AI 语音助手，而 2018 年，这个数字将达到 47.7%；

Strategy Analytics 预测到 2023 年前，全球 90% 的智能手机势必会配置 AI 语音助手。这说明智能手机配置 AI 语音助手功能，已经成为一种越来越流行的发展趋势。

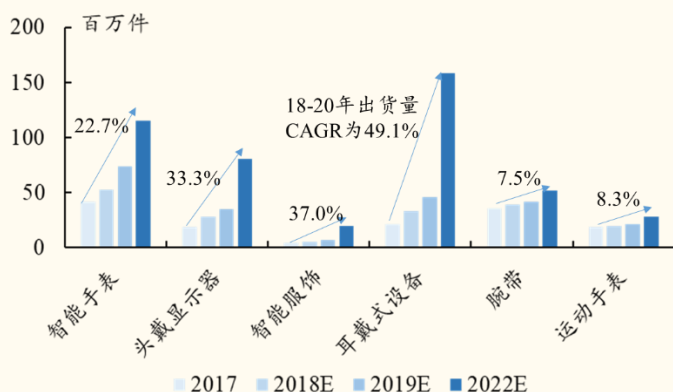
- **谷歌 Assistant 的 AI 技术明显超前：**2017 年，谷歌的 Assistant 语音助手就占了 46.7% 的市场份额，其次是苹果的 Siri 语音助手，占了 40.1% 的市场份额。预计 2018 年，谷歌 Assistant 语音助手的市场份额可能会增长到 51.3%；并且到 2023 年，它的市场份额预计增长到 60.6%。

3、智能穿戴“钱”途无量

A、可穿戴设备保持快速增长

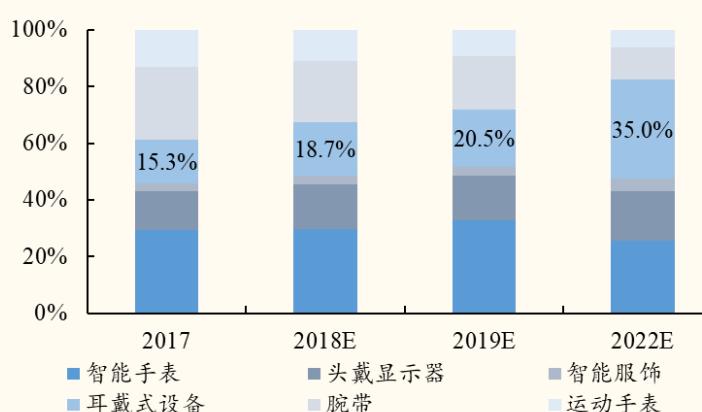
Gartner 预测，2018 年全球可穿戴设备出货量预计将达 1.79 亿部，较 2017 年的 1.41 亿部同比增长 27.1%。新兴市场可穿戴设备的出货量进一步增长，智能可穿戴设备未来五年复合年增长率将达到 26.3%，预计 2022 年出货量将跃升至 4.53 亿部。举例而言，至 2022 年，耳戴式设备、头戴式设备、智能手表的出货量分别将达到 1.58 亿台、8018 万台、1.15 亿台，五年出货量复合增速分别为 49.1%、33.3%、22.7%，耳戴式设备将取代智能手表成为可穿戴设备领域最主要的产品。到 2022 年，苹果 AirPods 之类的耳机设备，将占可穿戴设备市场份额的 30% 以上。

图表 43：17-22 年各类可穿戴设备出货量



来源：Gartner，国金证券研究所

图表 44：17-22 年各类可穿戴设备出货量占比



来源：Gartner，国金证券研究所

B、可穿戴 AI 市场增速更快

AI 助手需求量的不断增加、医疗行业运作量的增加、物联网技术的出现、无线技术的整合、可穿戴组件技术的增长前景，以及消费者对先进可穿戴设备的喜爱和需求的增加，都是全球可穿戴市场不断增长的主要驱动力。MarketsandMarkets 预测，2018 年全球可穿戴 AI 市场规模预计为 115 亿美元，到 2023 年，该数字余预计将达到 424 亿美元，预测期（2018 年~2023 年）内的年复合增长率为 29.75%。

可穿戴 AI 产品主要由显示、处理器、电源管理、集成电路、存储器 / 存储、传感器以及用户界面组成，按照产品种类的不同，可穿戴 AI 产品拥有智能手表、智能耳机、智能眼镜以及其他智能产品（如健身追踪器、智能服装、智能饰品、智能鞋等等）四大类；而根据操作的不同，可分为设备上的 AI 和基于云的 AI。

图表 45：全球可穿戴 AI 市场预测



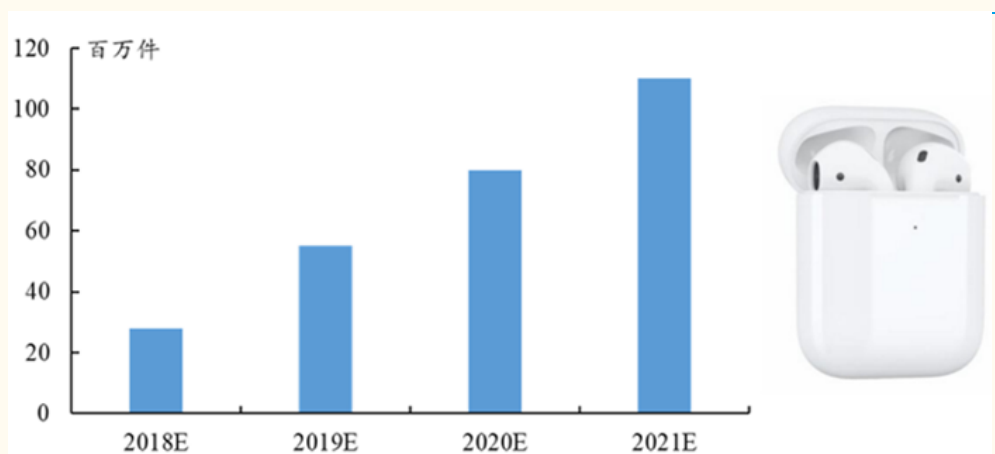
来源：MarketsandMarkets、国金证券研究所

C、蓝牙耳机有望会成为私人“AI 小秘”

而在过去的 2018 年里，无线耳机这个相对大的品类之中，其实也在悄然发生着改变。国内外音频 / 耳机大厂如索尼、Bose、森海塞尔、JBL 和惠威，乃至华为、荣耀和魅族等手机品牌，也都争先恐后的发布 TWS 类耳机（True Wireless Stereo，真无线耳机），TWS 耳机在 2018 年迎来了一次大爆发，结合 2018 年双十一、双十二电商销量数据，无线耳机产品销售量占比超过 50%，无线耳机已正式开始取代传统有线耳机的征程。2019 年，在 AI 技术的引入下，蓝牙耳机的功能将会更加千变万化，这也将是各家厂商竞争的焦点之一，具体有可能碰撞出什么火花？大致可以分为这些：（1）智能降噪将成为标配，（2）运动耳机将加入更多检测功能，（3）耳机也能识别你是谁：声纹 ID 与 AI 语音，（4）私人定制 EQ，（5）更多 AI 功能，如智能实时翻译、听译功能等。

- **AirPods 仍是最受欢迎的无线耳机：** AirPods 是苹果设备中最具备成长动能的产品，预计 2018 年 AirPods 出货量约 2800 万台，占比全球耳戴式设备出货量的 80% 以上。此外，苹果新款 AirPods 获蓝牙技术联盟认证，并支持健康追踪、无需手动可直接语音唤醒 Siri 等功能，其充电盒也有望增加无线充电技术，市场预计全新设计的 AirPods 将于 2020 年发布，18-21 年苹果 AirPods 将是全球耳戴式设备快速成长的主要驱动力。

图表 46：AirPods 出货量预测



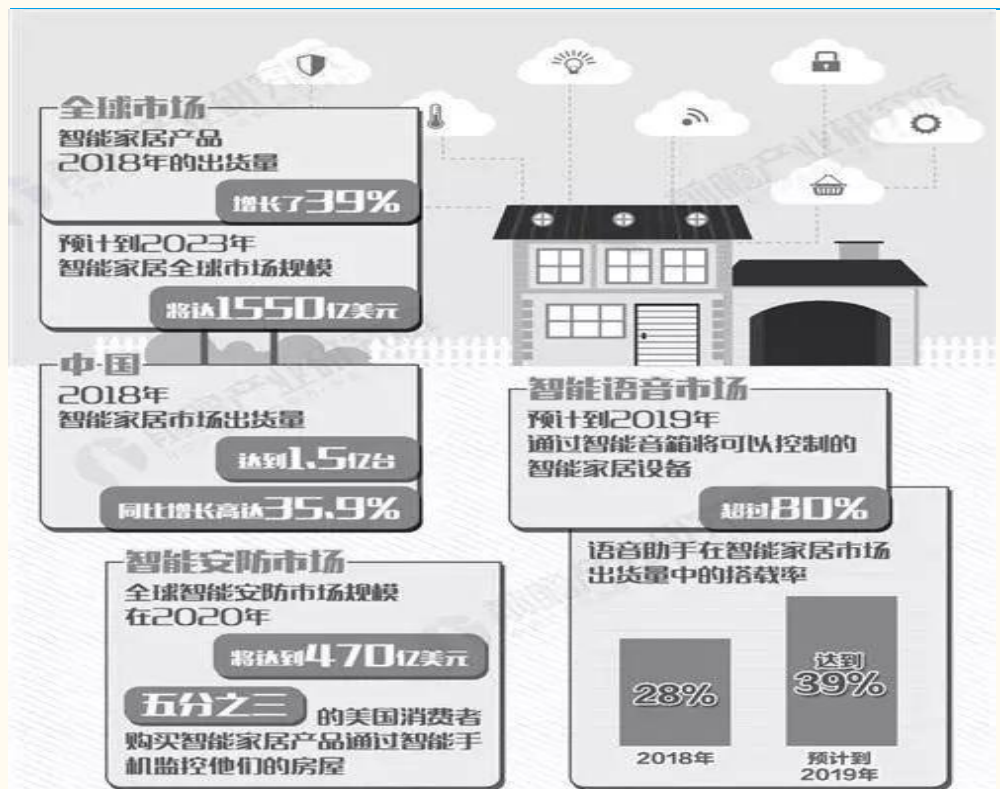
来源：TechWeb、国金证券研究所

4、安防+AI 让智能家居正式迈入智慧家居

智能家居产品已经悄然飞入“寻常百姓家”，如今的智能家居产品早已不只停留在概念阶段，各种神奇的“脑洞”已经纷纷落地成为现实。智能语音助手快速崛起，使其成为串联智能家居设备的重要“入口”；智能安防产品不断发展，明显提升了生活质量”。无论在国内还是全球市场，智能家居在过去一年中都交出了亮眼的“成绩单”。前瞻产业研究院统计数据显示，在全球市场，智能家居产品 2018 年的出货量增长了 39%，预计到 2023 年，智能家居全球市场规模将达 1550 亿美元。在中国，2018 年中国智能家居市场出货量达到 1.5 亿台，同比增高达 35.9%。

在 2019 年的 CES(美国消费电子展)上，智能语音助手成为谷歌和亚马逊两大巨头展台上的主角，智能语音助手连接的摄像头、温度计、空调和门铃占据了最显著的位置。亚马逊方面表示，其语音助手已被安装在超过 1 亿台设备上；谷歌则公布数据称，2018 年包括智能手机、智能手表、智能家居设备等在内，内置谷歌语音助手的产品数量从 4 亿台升至 10 亿台。

图表 47：全球智能家居发展情况



来源：前瞻经济学人、国金证券研究所

A、智能音箱，智能家居设备的入口

搭载了智能语音助手的智能音箱，被厂商们视为能够串联起智能家居设备的“入口”。厂商们争夺智能音箱市场的主导地位，其实就是在争取首次购买人群以及刚刚开始建立自己智能家居设备系统的消费者，预计到 2019 年，通过智能音箱将可以控制超过 80% 的智能家居设备，2018 年语音助手在智能家居市场出货量中的搭载率为 28%，预计到 2019 年将达到 39%，未来将更多应用在智能插座、智能摄像头以及智能网关等产品上。

B、家庭智能安防渐成“刚需”

如果说智能语音助手作为“入口”的价值在于“串联”，那么以智能门锁、智能摄像头等为代表的智能安防设备作为“入口”，其价值则在于“刚需”。全

球智能安防市场规模在 2020 年有望达到 470 亿美元，五分之三的美国消费者购买智能家居产品并通过智能手机监控他们的房屋。

随着家庭场景自动化需求逐渐提升，家庭环境、安全和控制类设备市场将迎来快速增长期，预计 2019 年增速将达到 60%。人脸识别技术作为 AI 最为成功的应用之一，其虽然已经大范围应用于智慧城市、平安城市等大型项目中，但大部分技术并未下放至家居安防产业中，家居安防依然以普通视频监控为主，而一旦家居安防与 AI 相结合，将进一步推动智能家居智能化的发展步伐，让智能家居正式迈入智慧家居。

三、AI 信息服务的蓝海已至—系统平台战略篇

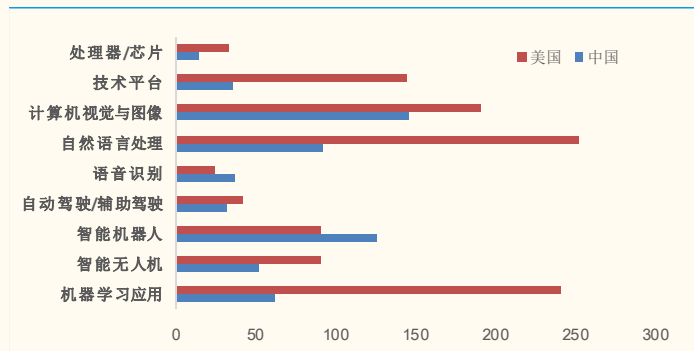
1、政策、技术、资本三轮驱动

A、全球主要国家均将人工智能作为未来主导性战略，美中领跑

美中暂时领先，中国在人工智能领域强在应用层，基础和技术层是相对短板。全球已有超 22 个国家发布 AI 计划，2017 年与 AI 相关的 VC 投资达 140 亿美元。美国和中国在人工智能领域处于领先地位，根据腾讯研究院的统计，全球人工智能企业总数达到 2542 家，其中美国拥有 1078 家，占 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。从中美两国在人工智能版图的产业布局看，中国在基础层（计算硬件及基础设施，核心为处理器、AI 芯片和底层系统、数据等）和技术层（算法理论、开发平台和应用技术，包括开源框架、自然语言处理、计算机视觉与图像等）与美国差距巨大，在应用层（AI 向各传统行业的渗透，包括机器学习应用、智能无人机、智能机器人、自动驾驶辅助驾驶、语音识别等）则与美国平分秋色。

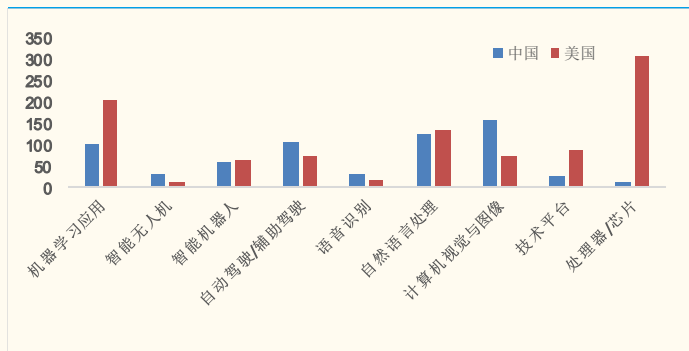
美国在基础层和技术层的 AI 企业数量约为中国的两倍，在应用层两国公司无论是数量还是融资金额差距较小。这说明在人工智能领域，中美两国延续了在互联网时代的竞争格局和发展路线，即美国的产业发展从底层技术开始再向上游应用拓展，而中国则开始更多在应用侧进行创新，逐步向底层技术进行渗透。美国 AI 创业公司中排名前三的领域分别为自然语言处理、机器学习应用以及计算机视觉与图像；中国 AI 创业公司中排名前三的领域分别为计算机视觉与图像、智能机器人以及自然语言处理。在基础和技术层的处理器/芯片，技术平台领域中国与美国存在巨大差距。

图表 48：中美各 AI 板块企业数量分布



来源：腾讯研究院，国金证券研究所

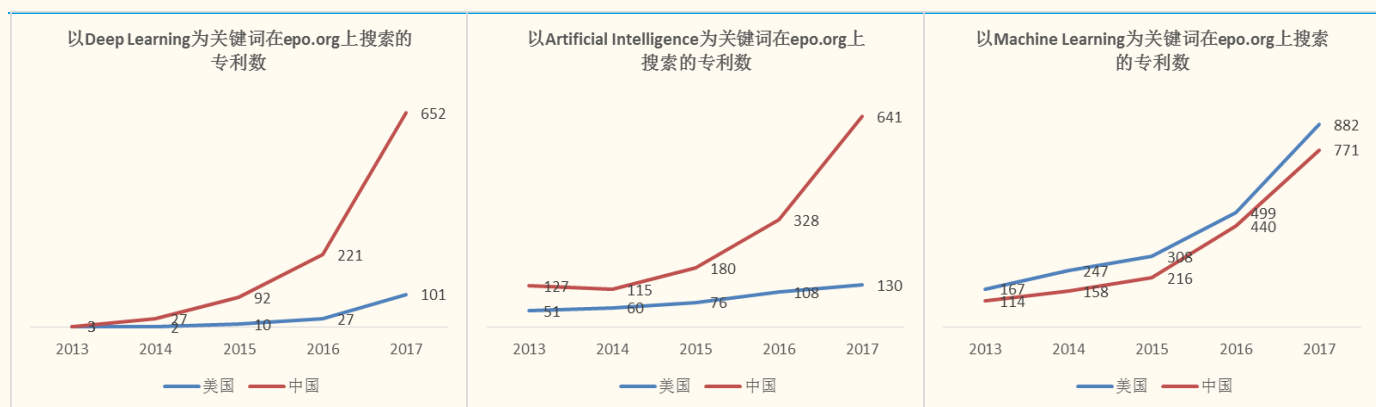
图表 49：中美各 AI 板块企业融资金额分布 (亿元)



来源：腾讯研究院，国金证券研究所

- **人工智能上升至中国国家战略层面，行业发展全面加速：**2017 年 7 月，国务院印发《新一代人工智能发展规划的通知》，首次从国家战略层面对人工智能进行了系统布局，规划提到至 2030 年，人工智能行业将在中国产生 10 兆的产业带动效益；2017 年 10 月，人工智能写入十九大报告；17 年 12 月，《促进新一代人工智能产业发展三年行动计划（2018-2020 年）》；18 年 3 月，人工智能再次被写入政府工作报告。政策密集出台，行业已进入发展黄金阶段。
- **5G+AI 是特朗普政府让美国再次伟大的关键：**美国希望在人工智能领域延续移动互联网时代领先优势，在 5G 领域重新抢占制高点。2019 年 2 月，美国总统特朗普签署行政命令，正式启动美国人工智能计划，这是美国政府首次推出国家层面的人工智能促进计划。美国人工智能计划包括研发领域、开放资源、政策制定、人才培养、国际合作等五个关键领域。5G+AI 将成为支撑川普承诺“再次让美国变得伟大”的关键。

图表 50：中国在人工智能领域的专利数呈现爆发式增长态势

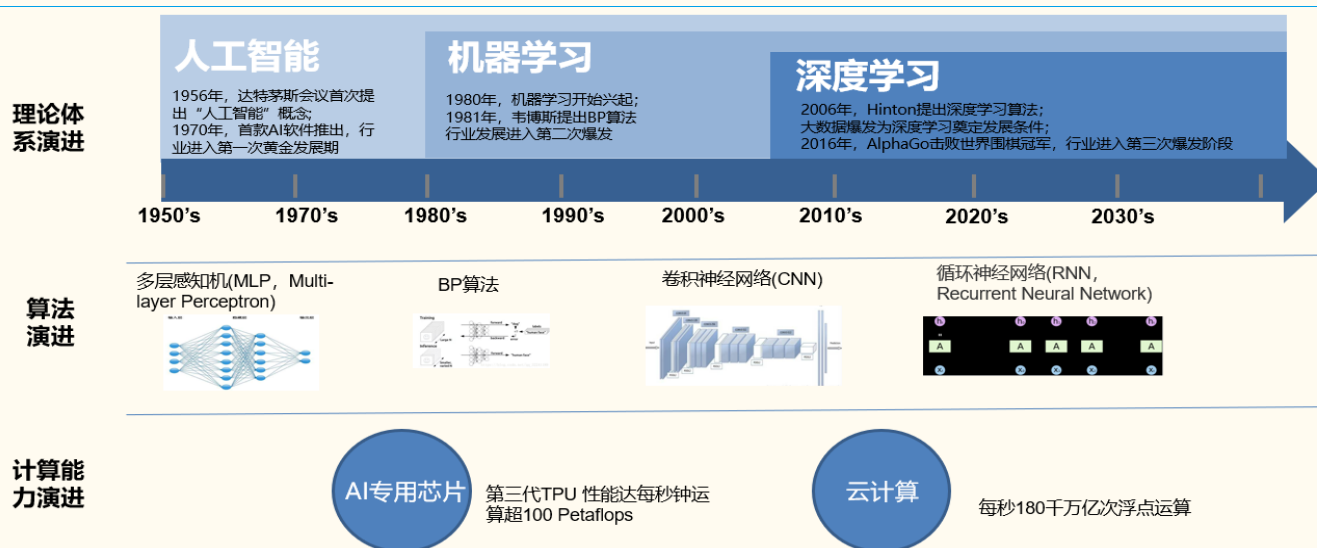


来源：CB Insights，国金证券研究所

B、技术变革促使 AI 突破应用瓶颈，推动产业应用爆发

人工智能的概念形成于 20 世纪 50 年代，诞生于 1956 年的达特茅斯会议。其发展阶段经历了三次浪潮：第一次是 50-60 年代注重逻辑推理的机器翻译时代，在此阶段机器人和智能软件开始出现；第二次是 70-80 年代依托知识积累构建模型的专家系统时代，但由于缺乏实用性行业很快又一次趋冷；第三次是 2006 年起深度学习算法的推出，开始了重视数据、自主学习的认知智能时代。在数据、算法和计算力条件成熟的条件下，本次人工智能的爆发浪潮中技术开始真正实现落地，深入到应用层面，帮助传统行业创造切实经济效果。

图表 51：人工智能技术不断突破



来源：公开资料，国金证券研究所

经过六十多年的发展，人工智能已从实验室（AI in vitro）全面走向了产业与行业的应用和研发（AI in vivo）。根据 Gartner 2018 年发布的人工智能技术成熟度曲线，VR/AR 等技术已逐步进入稳步爬升的光明期（Slope of Enlightenment），语音识别技术甚至已经到达实质生产的高峰期（Plateau of Productivity）。

图表 52：2018 年 Gartner 人工智能技术成熟度曲线

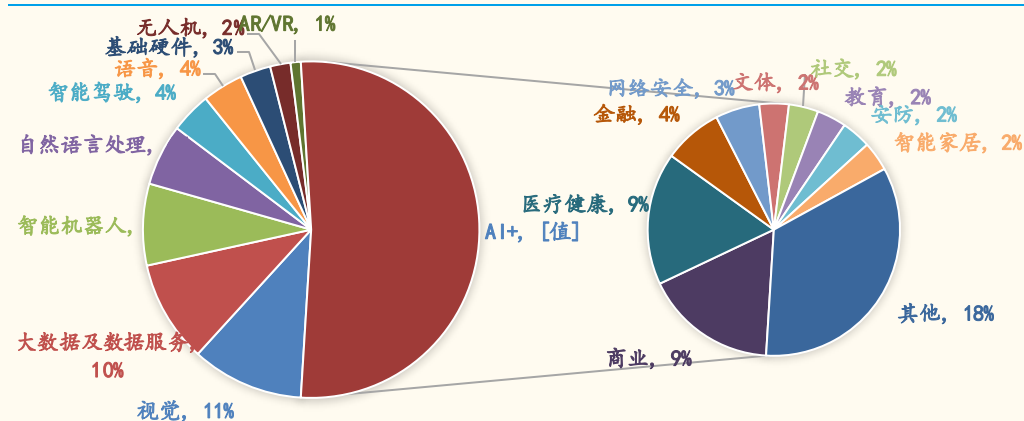


来源：Gartner，国金证券研究所

C、资本涌入加速产业成熟，AI+、视觉、大数据等领域成为投资热点

根据《2018 世界人工智能产业发展蓝皮书》，从 2013-2018Q1 全球人工智能投融资的分布来看，投资热点主要集中在 AI+（垂直行业）、计算机视觉、大数据及数据服务和智能机器人等领域。

图表 53：全球人工智能投融资笔数分布（2013-2018Q1）

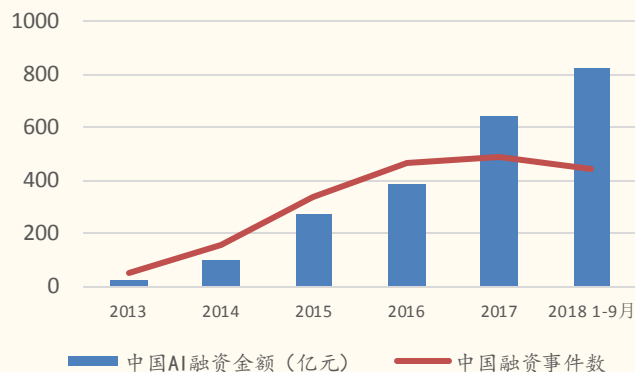


来源：中国信息通信研究院（2018），国金证券研究所

- 资本不断涌入，助推 AI 技术及应用进一步走向成熟：2017 年我国人工智能行业融资金额达 645 亿，融资事件达 492 起；2018 年前三季度我国 AI 发生融资事件 441 起，总金额达 825 亿元。据 CB Insight 数据，2017 年全球人工智能融资规模达到 152.42 亿美元，中国已对美国实现超越，融

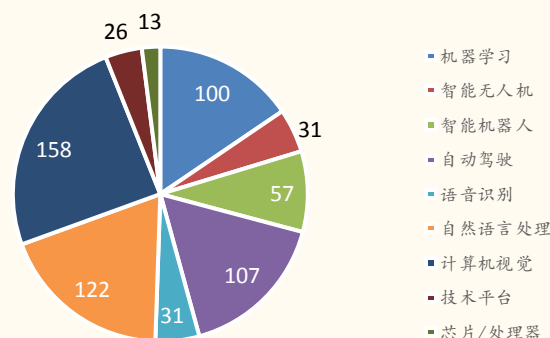
资规模占比 49%。资本的涌入将进一步催熟 AI 技术，加快技术在应用层的下沉。

图表 54：中国 AI 企业融资规模持续走高



来源：IT 橘子，国金证券研究所

图表 55：中国 AI 各领域融资金额分布



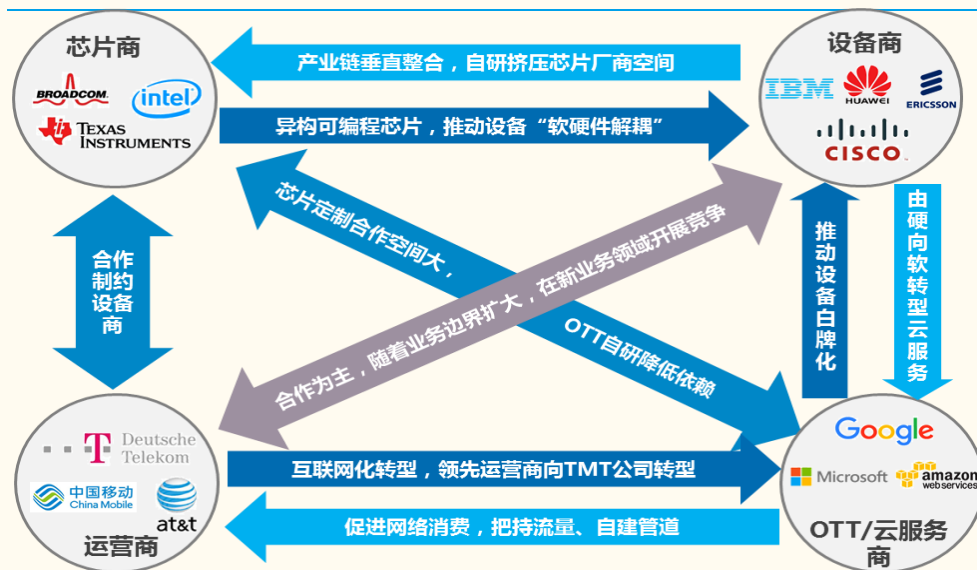
来源：腾讯研究院，国金证券研究所

D、人工智能重塑产业生态，芯片和软件平台是产业制高点

ICT 产业正在跨界和重构，传统的芯片商、运营商、设备商和互联网公司以及云服务商的合作和竞争边界变得越来越模糊，设备商和云服务商巨头进入芯片领域成为普遍趋势。

- **AI 颠覆芯片产业格局，行业进入群雄逐鹿阶段：**不同于传统通用芯片功能，AI 时代，底层芯片的设计目的并非为了执行逻辑复杂的指令，而是为了实现大规模数据训练和推理。随着 AI+成为普遍商业模式，人工智能已成为芯片产业最重要的增长动力，未来将加速普及渗透。芯片市场蛋糕将越做越大，拥有不同功能和定位的芯片将长期共存，百花齐放。全球 AI 芯片产业目前已初步形成五大阵营：一是传统的芯片企业转型升级做 AI 芯片，典型海外公司如 Intel，英伟达，赛灵思，高通，国内公司如杭州国芯、瑞芯微等；二是初创 AI 芯片公司，海外如 SambaNova Systems、Movidius，国内如寒武纪、地平线、深鉴科技，耐能，肇观等；三是互联网公司巨头进入 AI 芯片领域，海外如 Google、Facebook，Amazon 等，国内包括阿里巴巴、百度等；四是传统的 ICT（Information and Communication Technology）巨头进入 AI 领域，海外如 IBM，国内如华为；五是中国的产业生态，即从矿机芯片向 AI 芯片进行转型，典型的如比特大陆、嘉楠耘智等。

图表 56：ICT 产业正在跨界和重构，人工智能芯片成为新战略控制点



来源：公司公告等公开资料，国金证券研究所

- **人工智能的软件技术平台将是巨头的必争之地：**人工智能平台被认为是 AI 领域的“下一代操作系统”，这一领域将成为 Google、微软、Facebook、Amazon 等大玩家的竞争重点。Google 推出了 TensorFlow 跨平台深度学习架构，能够实现云端大规模学习到终端多平台部署的无缝衔接；Microsoft 推出的 CNTK 支持对常用深度神经网络架构的结构化执行，比如卷积神经网络 (CNNs)，循环神经网络 (RNNs) 和长短期记忆网络 (LSTMs)，同时支持多个计算设备以及多个 GPU 的计算；此外 Amazon、Facebook、阿里、腾讯等均已推出了自己的 AI 软件平台。我们认为，人工智能平台的技术壁垒高，马太效应强，且需要长期持续的大规模研发投入，未来这一领域将是国际巨头间博弈。

图表 57：全球 ICT 巨头 AI 布局一览

		Google	Amazon	Facebook	Apple	Microsoft	IBM	Alibaba	Tencent	Baidu	Huawei
应用层	产品应用	1.Waymo 2.Google Home 3.Google Nest	1.Echo 2.Alexa 3.Amazon Go	1.Chat Bot 2.Javis 3.Moments	Siri	1.微软小娜、小冰 2.Cortana 3. Tay 4.A-Eye		1.天猫精灵 2.阿里小蜜 3.家居、零售、出行（汽车）、金融、智能城市、智能工业	1.WeChatAI 2.Dreamwriter 3.绝艺 4.腾讯觅影	1.Raven H 2.度秘 3.Apollo小度 4.阿波龙	1.Atlas 2.MDC
	应用使能	1.Voice Intelligence API 2.Google Translate API	1.Amazon Lex 2.Amazon Polly 3.Amazon Rekognition	1.DeepFace 2.DeepMask 3.SharpMask 4.MultiPathNet		Azure 认知服务	1.Watson 2.BlueMix 3.ROSS	ET 大脑	云搜 文智 优图	1.Apollo 2.DuerOS 3.百度大脑	1.HiAI 2.SoftCOM AI 3.Cloud EI 4.Fusionmind 5.ModelArts
技术层	算法	AlphaGo									
	平台/框架	TensorFlow	Amazon Machine Learning	Torchnet	1.Neural Engine 2.Core ML 3.Create ML	1.CNTK 2.DMTK 3.ONNX 4.ML.NET	SystemML	PAI 2.0	1.Angel 2.NCNN	Paddle-Paddle	MindSpore
基础层	芯片/算力	1.TPU 2.量子计算机 3.Google Cloud	1.Annapurna ASIC 2.Blink ASIC	Big Sur		Azure+FPGA	1.Power 2.TrueNorth	Ali-NPU		1.XPU 2.昆仑 3.DuerOS智慧芯片	1.昇腾 2.麒麟

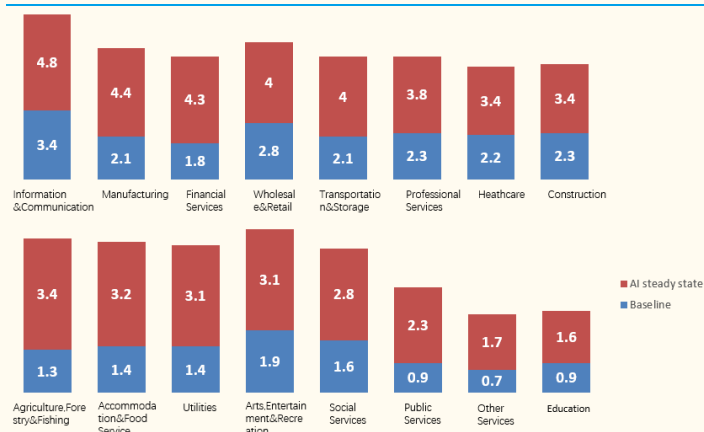
来源：腾讯研究院，华为，各公司公开资料，国金证券研究所

2、AI 是全新的生产要素及工具, AI+行业将成为普遍商业模式

A、高数字化行业将优先实现人工智能的应用

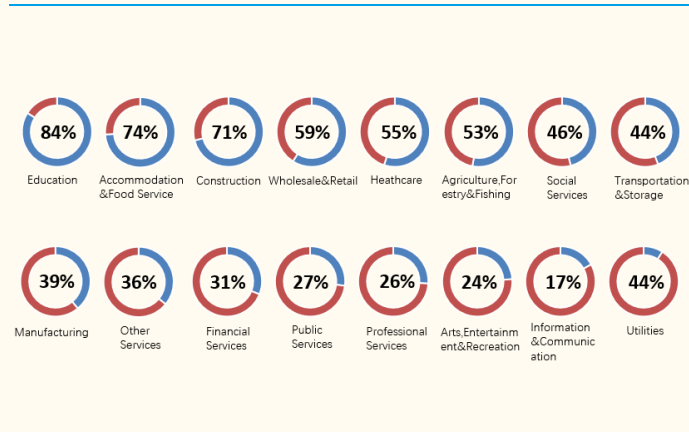
- **人工智能作为全新的生产要素，将改变企业竞争及增长方式：**人工智能作为全新的生产要素，将至少在三个方面推动增长：首先是智能自动化效应，AI 能够自动执行实体环境中对敏捷性及适应性要求较高的复杂任务，同时伴随大规模的重复工作能够进行自我学习以实现持续升级；第二，AI 可以补充和增强现有劳动力及资本，对其赋能，提高资源利用效率；第三，人工智能可以激发大量创新，如自动驾驶，新零售等。故人工智能未来将成为社会发展的基础设施，作为新的生产要素彻底改变企业的竞争及增长方式。据埃森哲数据，未来与 AI 深度融合的企业能够将盈利能力平均提升 38%，同时 AI 将为包括教育、制造、批发、零售等 16 个行业额外带来超 14 万亿美元的总附加值。

图表 58：AI 为行业带来增值



来源：埃森哲，国金证券研究所

图表 59：AI 为行业带来利润提升

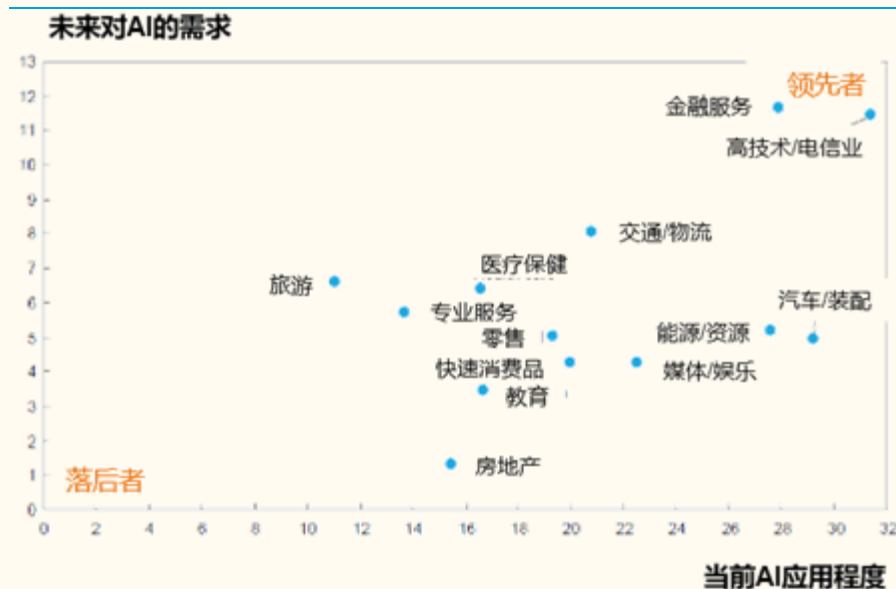


来源：埃森哲，国金证券研究所

- **人工智能的应用进展取决于数字化程度，高数字化行业将优先实现应用：**
AI 时代是数据驱动的时代，深度学习算法的优化需要大规模数据来训练提升，数据越丰富完整，应用效果越完美。高数字化程度的行业拥有较密集数据资源，成为 AI 优先落地的领域。在落地应用场景中，目前来看安防、高技术/电信业、金融、汽车等行业落地最快，医疗、零售、旅游、地产等行业较为滞后。

与以往信息化不同，AI 深入到各行各业需要深入到生产系统，与线下、本地各种场景相结合，采用云服务是必然的。ABC 融合是云服务进入 2.0 阶段的标志，行业数字化还需遵循 C—>B—>A 的路径，即先云化，再有大数据最后实现智能化。未来人工智能将成为基础设施，科技巨头将成为人工智能基础设施的提供者，普通企业拥抱人工智能的前提是先云化，实现核心业务在线才能加入人工智能的大生态圈，这将不是 IT 支出转云服务支出“左右互搏”的零和游戏，而是涉及未来生死存亡的抉择。

图表 60：不同行业 AI 技术应用程度



来源：麦肯锡，国金证券研究所

B、安防、自动驾驶是当前和未来人工智能重点应用领域

人工智能触发的产业变革，将涉及所有行业，从交通、教育、医疗到翻译、运维和自动驾驶等等。我们认为，人工智能在行业的应用进展取决于行业数字化程度，科技行业之外的应用虽然目前整体上还处于早期阶段，但在部分领域

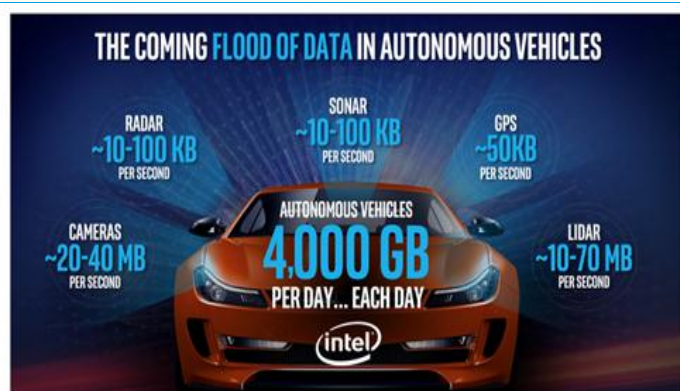
已经取得了巨大的进展。人工智能率先应用的场景，一定是有海量数据和海量计算需求的场景，安防和自动驾驶正是这样的场景之一。

■ **AI 推动安防行业进一步升级，拓宽行业成长空间：**受益于 AI 机器视觉技术的迅速成熟，安防行业亦逐步由“看得见”——“看得清”——“看得懂”发展，智能安防能够自行提取更多细微特征，大幅减少人力及开发成本，将原来的事后追责转变为事前预防及事中识别，大幅提升政府/企业工作效率，如 2017 年商汤科技的人脸识别在重庆落地后，相比人工效率提升 200 倍。安防行业从过去的模拟到数字再到高清化时代，每一次产业升级都伴随着市场需求的进一步释放。我们认为，在智能化时代，安防行业的市场边界将进一步拓宽，未来 AI+视频监控将逐步下沉至自动驾驶、智慧医疗、环境监测等下游新兴应用，市场空间将进一步拓宽。预计 2020 年我国安防行业整体规模将超过 9000 亿元。

■ **自动驾驶是人工智能的重要落地领域：**Intel 曾经预估，由于激光雷达（每小时可产生 36-252 Gigabyte 的影像资料）和摄像头（每小时可产生 72-144 Gigabyte 的影像资料）的大量使用，每辆无人驾驶车每天将产生 4 Terabyte or 4,000 Gigabyte 的路况原始影像资料。

目前各整车厂及科技巨头均在自动驾驶领域有所布局：整车厂方面，奥迪 L3 级 A8 已实现量产上市；特斯拉已宣布覆盖各类驾驶场景的完全自动驾驶将于 19 年年末上市；通用汽车 18 年计划生产数千辆自动驾驶汽车用于 Lyft 出租车服务。科技巨头方面，谷歌 Waymo 18 年已拿到首个商业自动驾驶打车服务执照；百度 Apollo 于 11 月与一汽共同宣布了中国首款 L4 级自动驾驶乘用车的量产计划。

图表 61：自动驾驶汽车产生海量数据



来源：Intel，国金证券研究所

图表 62：华为 MDC 600

RTOS	Sensor data abstraction	Vehicle dynamics Model	Over The Air(OTA)	...	Security module	Management module	Tool Chain
	Communication middleware	AI operation library	Heterogeneous acceleration resource	...			
MDC hardware system							
SOC			MCU				
Host CPU	Ascend	ISP					

来源：华为，国金证券研究所

华为在自动驾驶领域近期也动作频频，去年 10 月华为联合奥迪发布了自动驾驶方案 MDC600。利用 8 颗昇腾芯片+CPU+ISP 所组成的系统，基于昇腾自动驾驶方案的奥迪 Q7 自动驾驶汽车号称达到了 L4 水平。我们认为在 L3 以上的级别，自动驾驶未来的数据处理方式将从单一的自带处理单元本地处理，发展为自带处理单元和网络边缘服务器处理简单计算过程，云端处理大量复杂计算过程的模式。在自动驾驶领域，我们也将看到人工智能多场景的融合，即设备端智能（车）+边缘运算智能（MEC）+云端计算智能的统一。

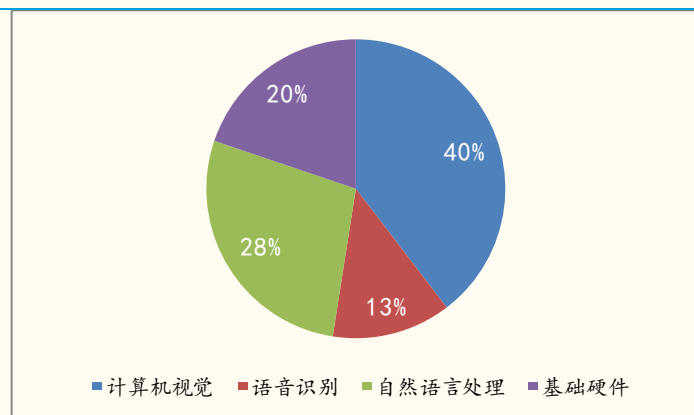
3、计算机视觉、智能语音是 AI 时代入口，中国初创龙头企业具备全球竞争力

A、计算机视觉：AI 行业最具商业化价值的赛道

■ **计算机视觉是国内外 AI 企业最集中的领域，商业成熟度较高：**从 AI 企业的应用技术方向分布来看，计算机视觉技术企业在全球 AI 企业中占比约 40%，在国内占比约 46%；无论国外还是国内，计算机视觉都是 AI 企业最集中的领域。从市场规模来看，2017 年计算机视觉市场占全球 AI 市场总规模的 16.9%，排在语音识别之后；而国内计算机视觉市场占 AI 市场的 34.9%，排名第一。国内外计算机视觉的市场规模差异要远大于企业分布

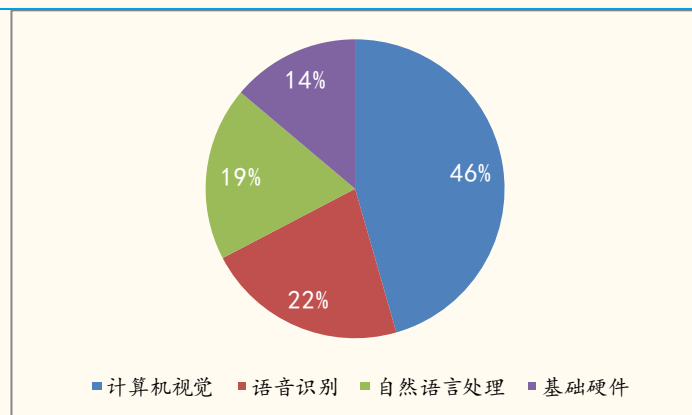
差异，说明国内计算机视觉公司的总体盈利能力较其他 AI 领域的公司较强，商业成熟度较高。

图表 63：全球 AI 企业应用技术方向分布



来源：CISTP，国金证券研究所整理

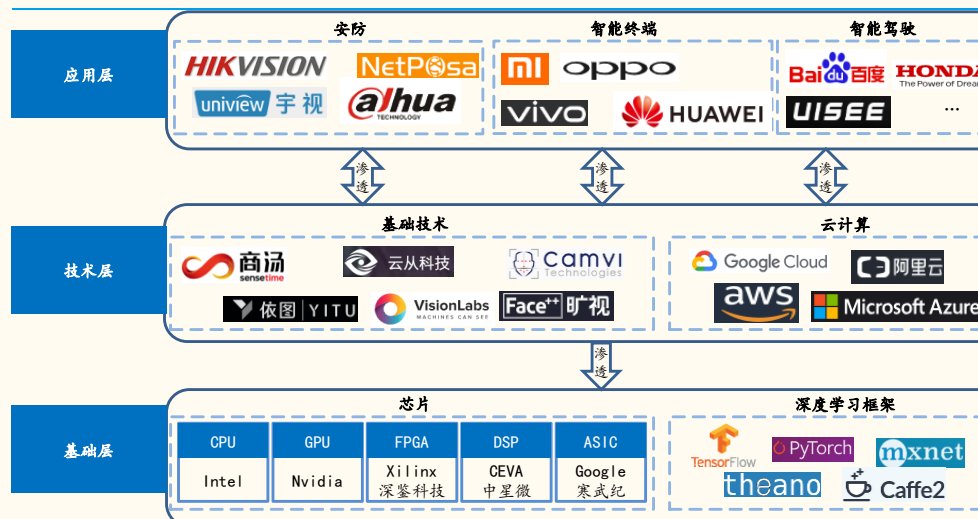
图表 64：中国 AI 企业应用技术方向分布



来源：CISTP，国金证券研究所整理

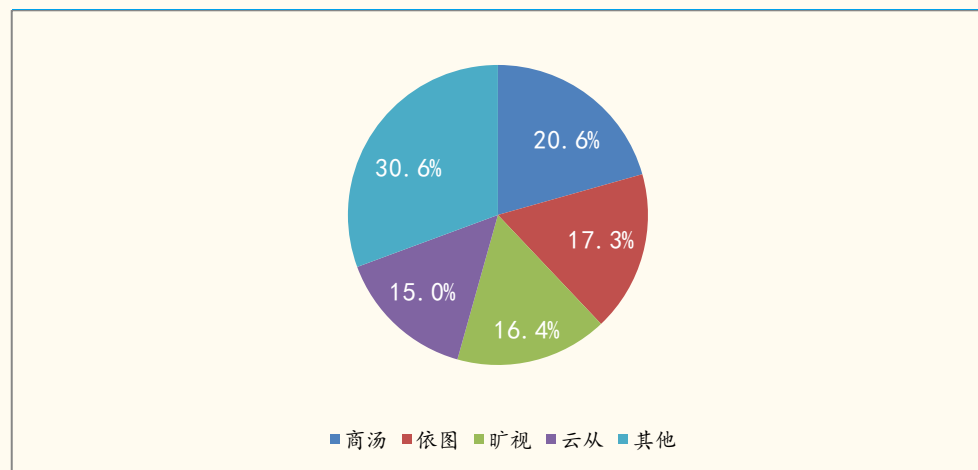
- 计算机视觉是 AI 领域应用场景最丰富、商业化价值最大的赛道：**目前，AI 技术处理的数据类型不外乎四类：文字、语音、图像和视频。从信息维度来看，从文字到视频维度是递增的，文字的信息维度最少、包含的信息量也最少，视频的信息维度最多、包含的信息量也更大。反映在数据量占比上，以线上数据为例，根据 Cisco 的研究，到 2022 年全球线上视频流量占总流量的比例将从 2017 年的 75% 上升到 82%，说明线上数据将越来越被视频数据所主导。信息维度更高加之数据量更大，因此以图像和视频为主要处理对象的计算机视觉要比以文字或语音为主要处理对象的其它 AI 技术具有更加丰富的应用场景和商业化价值。我们认为，当前资本市场也正以其资源配置、资产定价功能充分反映计算机视觉相对其它 AI 领域的优势。例如，根据公开资料，当前计算机视觉行业四家头部初创企业（商汤、旷视、依图、云从）的总估值已经超过 1000 亿人民币，也超过了语音识别行业几家头部初创企业（思必驰、云知声、出门问问、图灵机器人、捷通华声）总估值加上 A 股 AI 语音龙头科大讯飞的市值总和。
- 科技巨头把控基础层，初创企业领跑应用层：**计算机视觉架构从下至上：1）基础层——核心芯片被 Intel、Nvidia 等传统芯片厂商把控，新型芯片厂商尚未崛起，规模应用有待时日；开源平台以谷歌的 Tensorflow、Facebook 的 Caffe 等为主，其他企业的深度学习框架多为二次开发；2）技术层——算法，初创企业占优；云计算，几乎被亚马逊的 AWS、谷歌的 Cloud、微软的 Azure、阿里云等垄断；3）应用层——垂直行业龙头占据场景，技术层初创企业向上渗透。根据 IDC 的统计，2017 年中国计算机应用市场商汤、旷视、依图、云从等算法及软件解决方案公司“四小龙”总体市场份额达 69.4%，其中商汤市场份额 20.6% 排名第一。

图表 65：计算机视觉架构及各层级玩家示意



来源：公司公告等公开资料，国金证券研究所整理

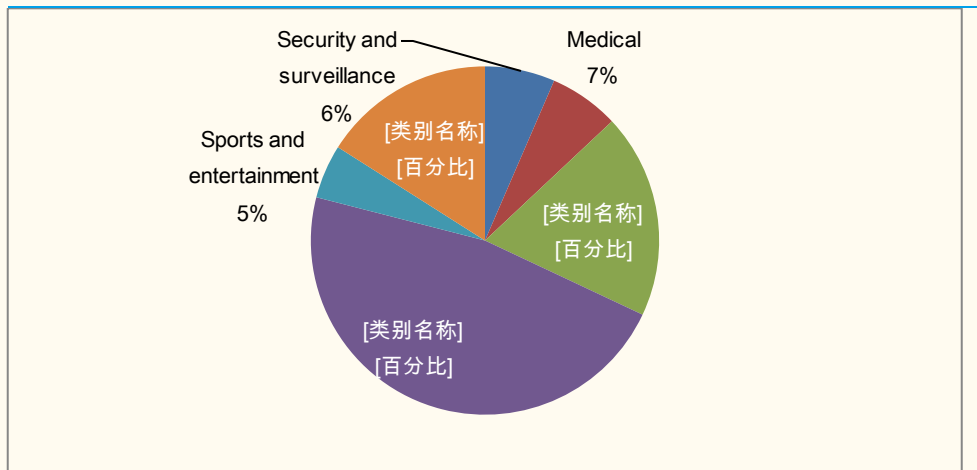
图表 66：2017 年中国计算机视觉应用市场份额



来源：公司公告等公开资料，IDC，国金证券研究所整理

- **国内以安防、金融、互联网为主 vs. 国外消费、机器人（及机器视觉）、智能驾驶领先：**根据 IDC 的市场跟踪，2018 年中国计算机视觉技术输出规模最大的 3 个行业是政府、金融和互联网，最大的两个场景为政府行业中的平安城市以及金融行业中基于人脸识别的身份认证。而根据 Tractica 的预测，2018 年全球计算机视觉技术输出规模最大的 3 个行业分别为消费、机器人（及机器视觉）以及智能驾驶。我们认为，造成国内外市场结构巨大差异的原因，主要是政府对市场的干预程度。政府的干预体现在两个方面：1) 政府作为市场玩家之一参与市场的深度。在国内，基于公共安全效率及成本方面的考虑，政府成为最早和最积极采用计算机视觉产品的市场玩家之一。由于安防行业商业成熟度高、实际买单方（政府）信用和购买力好、图像及视频数据多，因而计算机视觉最容易落地。2) 政府作为监管者出台相关政策干预市场的强度。在欧美，生物特征作为关键的个人信息，在被科技公司收集和使用受到法律法规的严格保护。例如，去年 5 月 25 日，欧盟史上适用范围最广、定则条例最严、处罚金额最昂贵的数据保护法案 GDPR 生效，该法案将指纹、人脸、视网膜等信息全部纳入个人资产范畴，对科技公司利用上述信息盈利进行了严格规定和限制。因此，欧美计算机视觉的主要落地方向多为对个人隐私保护要求较低的行业。政策因素（如实名制、反洗钱等）也是导致国内金融业主动拥抱计算机视觉技术的主要原因。

图表 67：2018 年全球计算机视觉行业市场结构



来源：Tractica 预测，国金证券研究所整理

B、中短期核心竞争壁垒是技术和产品能力，长期是生态构建能力

在这里技术能力是指企业算法的设计能力、迭代速度等，产品能力是指企业技术的落地能力。中短期来看，包括人脸识别、人体识别、图像识别等在内的主要计算机视觉技术的研发将保持基于使用神经网络的深度学习算法，而神经网络种类选择、结构设计以及参数调整等是一连串极其复杂的工作，因此算法的优劣和迭代速度将直接取决于算法设计人员的知识和经验储备，也会进一步决定计算机视觉技术的准确率、可靠性等关键性能。此外，计算机视觉技术在场景中落地时需要企业具备大量的除算法以外的 know how，例如在软硬件结合以及保护终端用户隐私上的知识和经验。现阶段，掌握技术的 AI 人才（尤其是顶级人才）主要分布在高校及科研院所，产业内的人才供应不足，加之多数场景下计算机视觉的落地能力不足，技术（人才）的争夺与产品（落地能力）的竞争将决定企业的发展速度。

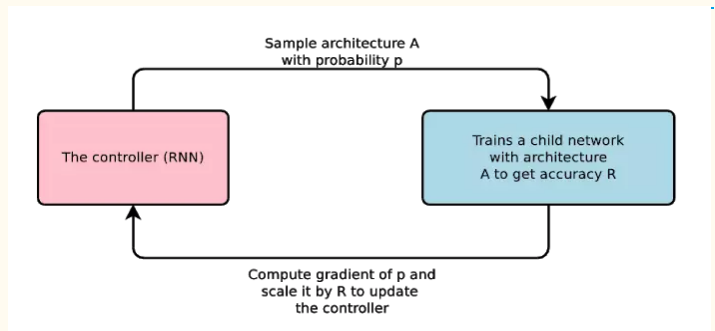
技术层面，随着自动机器学习技术 AutoML 的出现，机器学习模型的设计门槛随之降低。AutoML 的出发点是用强大的算力通过更多次的训练提高模型的准确度，其最大特点是将机器学习模型的设计过程自动化。算法设计人员只需了解模型的基本概念并提供标签数据即可，神经网络的参数及结构调整是自动完成的，无需人工干预。目前 AutoML 已经进入落地阶段。去年 1 月谷歌发布了提供自定义图像识别系统自动开发服务的 Cloud AutoML Vision，用户从导入数据到训练模型都可以通过拖放式界面完成。AutoML 已经被谷歌应用于 CIFAR-10 高度基准测试数据集，并且训练出了与手工设计不相上下的模型。我们认为，随着 AutoML 技术的成熟，算法设计的门槛将越来越低，但是相应的，算力的重要性会越来越明显，长期来看有可能成为计算机视觉企业的核心竞争力之一。

图表 68：AutoML 工作原理（红色部分自动完成）



来源：Google，国金证券研究所整理

图表 69：AutoML 内部工作机制



来源：Google，国金证券研究所整理

产品层面，我们认为，随着竞争的加剧，成功的企业不仅要具备大量的跟技术落地相关的 know how，还必须能够主动地去挖掘甚至创造客户的需求，具备类似苹果在智能手机上的理解力。这就要求企业在纵向上能够主导整个产业链，横向上能够接入更多开发者和场景，因此长期看生态构建能力将成为主要竞争壁垒。

C、头部算法企业横向建立平台，垂直领域龙头纵向深耕行业，科技巨头立体式打造全产业链生态。

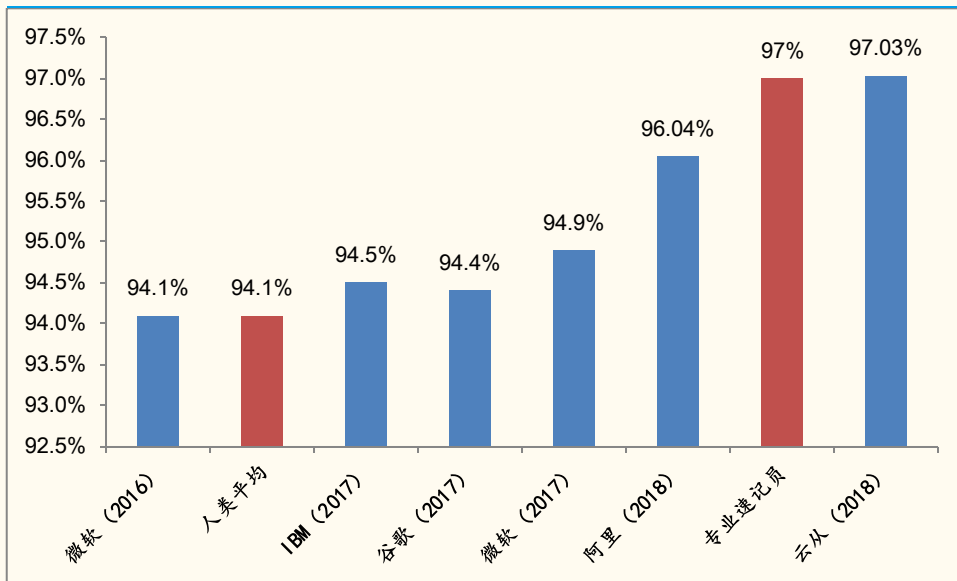
头部算法企业的布局是横向拓展基础技术，建立基础平台。例如，商汤以人脸识别技术起家，逐渐拓展到人体分析、机器人、无人驾驶等领域，加上 GPU 超算中心和云计算平台，逐渐将自身打造成一个基础的人工智能平台，然后探索在应用端的迅速落地。

垂直领域龙头公司的布局是纵向打通计算机视觉框架，深耕所在行业。例如，安防设备龙头海康威视逐渐脱离了与初创算法企业的合作，自主研发关键技术，并在 ILSVRC (2016) 图像分类获得第一名；此外，还推出了 AI Cloud 平台，着力解决算力问题，而大华积极投资其基础层边缘运算芯片的研发。通过纵向打通基础层、技术层和应用层，垂直领域龙头公司可以提出更加智能的行业解决方案。

科技巨头公司的布局是以平台为核心立体式地推进，打造全产业链生态。例如，Google 以 TensorFlow 及 Google Cloud 为核心，横向上进军图像分析、文字识别等技术，推出 Google Lens 等视觉服务产品；纵向上向下拓展基础硬件推出 TPU 张量处理器 1, 2, 3 芯片，向上拓展医疗 (Verily)、无人驾驶 (Waymo) 等垂直领域，逐渐打造包括计算机视觉在内的人工智能大生态。

D、智能语音：产业进入爆发期，看好对话式人工智能发展

- **智能语音技术逐渐成熟，未来几年市场将维持高速增长：**根据 Gartner 发布的 2018 年 AI 技术成熟度曲线，语音识别已经进入“实质生产高峰期”。此外，像自然语言处理、虚拟助理等相关智能语音技术历经淘洗之后即将落入“泡沫化低谷期”，商业模式越来越成熟，也将进一步推动智能语音的落地。根据 ReportLinker 的预测，全球智能语音市场规模将从 2018 年的 75 亿美元增长至 2024 年的 215 亿美元，其中医疗健康、移动银行以及智能终端智能语音技术快速增长的需求将成为主要的驱动因素。
- **国内外市场均已相对集中，头部企业占据大部分市场份额：**根据中商产业研究院的数据，2018 年智能语音全球市场 CR5 达到 88%，中国市场 CR5 达到 84.6%，均已呈现出相对较为集中的市场格局。值得注意的是，国内的市场格局与国外存在着明显的差异，主导国外市场的巨头玩家如 Nuance、谷歌等并没有在中国市场取得同样的市场地位。我们认为，这种现象主要是由两方面原因造成的：一是由于中文与外语在分词及句法的逻辑等方面有着明显的不同，并且中国有大量独具特色的方言，这种语言上的差异使得国内企业在中文处理上独具优势；二是受益于国内互联网和移动互联网的高速发展，智能语音在国内具有更加丰富的落地场景和商业化环境。这种“天时”、“地利”的优势使得中国企业领跑国内市场。
- **现阶段，产品化能力是智能语音初创企业的核心壁垒：**一方面，从技术的角度，在无噪音环境下，机器的语音识别准确率已经达到相当高水平，未来上升空间有限。科研界将错词率 (Word Error Rate, WER) 视为衡量语音识别技术的核心指标 (准确率=1-WER)，其中人类的平均 WER 为 5.9% (即 94.1% 的准确率)，受过严格专业训练的速记员平均 WER 为 3% (即 97% 的准确率)。自 2016 年机器的语音识别准确率首次超过人类平均水平以来，目前最好的算法已经可以将准确率做到 97.03%，超越专业速记员。另一方面，从商业化角度，技术难以成为企业核心壁垒，产品化能力才是成长的关键。这是因为技术的门槛越来越低，单纯依靠提供技术的商业模式将面临越来越大的竞争因而难以持续，这一点从近几年头部智能语音初创企业如云知声、思必驰等纷纷从技术提供商转向产品/服务提供商就可以看出。

图表 70：近几年语音识别准确率变化

来源：公司公告等公开资料，国金证券研究所整理

- **技术全栈化已经成为新的趋势，产业链延伸是当前最好的策略：**对于智能语音而言，多数场景下普遍存在的问题是落地体验不好，而这带来了技术上和产业上的两层影响：技术上，落地体验的改善有赖于从硬件到软件、从算法到产品各个环节的紧密协调，在这种情况下，头部企业多倾向于走技术全栈化路线，即打通全链条上的技术环节；产业上，为了实现更好地商业化，企业倾向于通过上下游延伸创造更多价值。最能说明上述两个趋势的例证是当前几乎所有的头部初创企业都在布局专用语音芯片，主要原因是专用语音芯片的应用性能可以达到通用芯片的 40-50 倍，而同时成本只有通用芯片的 1/4-1/3。

图表 71：头部智能语音初创企业首款专用芯片发布梳理

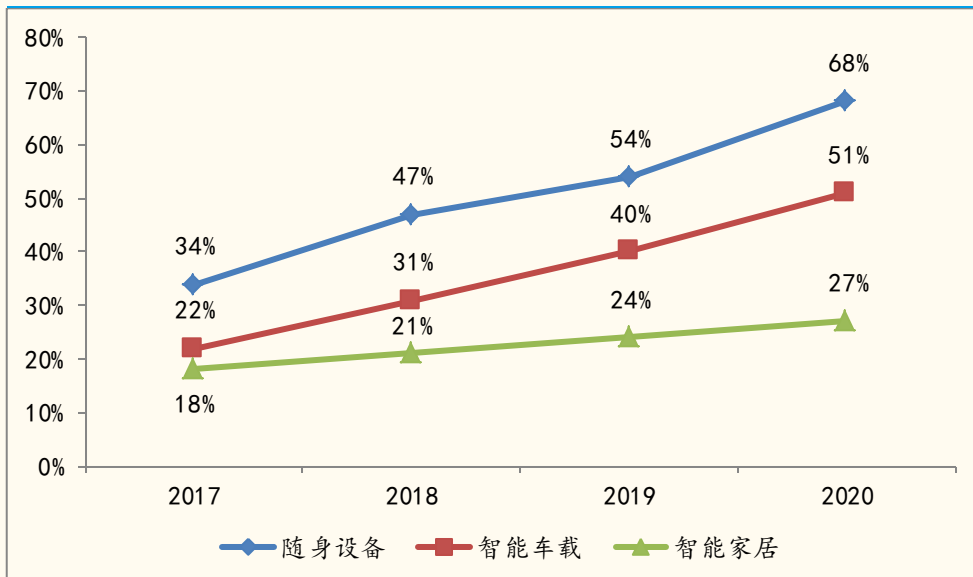
时间	公司	芯片
2018.05.15	云知声	面向 AIoT 的 UniOne “雨燕”
2018.05.24	出门问问	AI 语音芯片模组问芯 Mobvoi A1
2018.07.02	Rokid	AI 语音专用 SoC 芯片 KAMIN018
2019.01.04	思必驰	AI 专用语音芯片 TH1520

来源：公司公告等公开资料，国金证券研究所整理

- **对话式人工智能即将规模化落地，看好智能家居、随身设备、智能车载三个场景：**与一般的智能语音应用相比，对话式人工智能中人与机器之间是双向交互的，机器在追求理解人的意图的同时也会给人以反馈（比如要求人进一步明确意图或者给出选择等）。对话式人工智能的整个人机交互过程可分为唤醒、识别、理解、反馈四个环节。我们认为对话式人工智能将率先在智能家居、随身设备、智能车载三个场景中实现大规模落地，原因是：1）上述三个场景作为家庭 IoT、个人 IoT 以及车联网的重要入口，是头部企业争夺的重点，例如亚马逊的 Echo、谷歌的 Assistant、苹果的 Siri、思必驰的飞歌 GS2 正是分别瞄准上述场景；2）上述三个场景对智能语音技术的要求相对较低，例如智能家居和智能车载都是在相对封闭的环境，语

音唤醒成功率和语音识别的准确率会相对较高；3）在这三个场景里 C 端用户对交互式人工智能的接受程度较高，IDC 调研数据显示 89% 的人有意向在未来两年内采用对话式人工智能终端。另外，根据 IDC 预测，到 2020 年中国对话式人工智能在上述三个关键领域的渗透率将分别达到 27%、68%、51%。

图表 72：对话式人工智能在智能家居、随身设备、智能车载的渗透率



来源：IDC，国金证券研究所整理

- **市场玩家的布局：通用平台商打造开放语音生态，专业应用商抢占重点赛道：**智能语音市场玩家大体上可以分为两种，一种是通用平台商，另一种是专业应用商。通用平台商如谷歌、百度等科技巨头，致力于围绕智能语音开放系统打造智能语音生态。以百度为例，公司在 2017 年全面开放了语音接口，目前依托 DuerOS 打造出较为完整的技术、开发与商业生态系统，截至 2018 年 7 月底搭载 DuerOS 的智能设备激活数量已突破 1 亿台。专业应用商以垂直领域和细分场景为突破口，均有重点布局的赛道。例如，云知声重点布局家居、车载与医疗，思必驰重点布局家居、车载与机器人，两者多有重合但侧重也有差异。

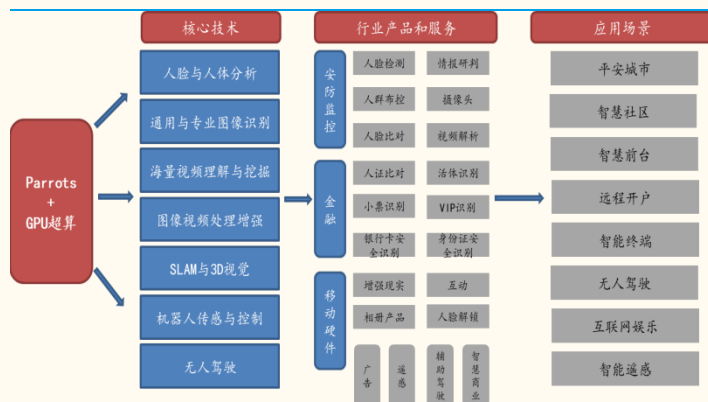
4. 领先初创企业介绍：商汤、旷视、依图、云从、云知声

A、商汤：AI 算法提供商龙头，平台化战略赋能公司和行业

- **AI 算法提供商龙头，“1+1+X”打造核心平台：**公司成立于 2014 年，创始人人为香港中文大学工程学院教授汤晓鸥，业务聚焦于计算机视觉和深度学习领域，拥有 2200 位员工，其中约 150 位拥有世界顶级大学博士学位。2018 年 9 月，公司被科技部选为“智能视觉”国家新一代人工智能开放创新平台建设依托单位。公司采取平台化战略，主要的业务模式是首先进行前期基础技术研究，然后结合当前行业应用热点生成相关产品和服务，最后进行营销和推广，即“1（基础研发）+1（产品和服务化）+X（行业应用）”打法。当前公司战略推进顺利，去年 5 月 31 日公司公告表示 2017 年已实现全面盈利，业务营收连续三年保持 400% 同比增长，2018 年主营业务合同收入同比增长 10 多倍。
- **内生+外延，围绕计算机视觉展开全方位布局：**公司通过内生和外延不断拓展技术和业务边界，提前进行产业布局。内生上，公司以人脸识别技术起家，当前核心技术已覆盖人脸、人体、图像、视频、SLAM 与 3D、机器人、无人驾驶等多领域计算机视觉技术，并且仍在不断横向拓展，进而衍生出智能视频、身份验证、移动互联网、智慧商业等多类产品和服务，覆盖智慧城市、智能终端、互联网娱乐、智慧金融、智慧商业、遥感、移动运营商、无人驾驶、AR/VR 等多个应用场景。外延上，公司通过成立投资部门，

通过直投或者产业基金方式向下游延伸布局新应用和新场景，投资标的以在垂直领域已建立起行业场景、占据一定地位或者已具备一定行业认知及客户资源但技术和产品尚不成熟的初创企业为主，例如医疗健康领域的 AI 初创企业禾连健康、VR 看房企业 51VR 等。

图表 73：商汤业务布局梳理



来源：公司公告等公开资料，国金证券研究所整理

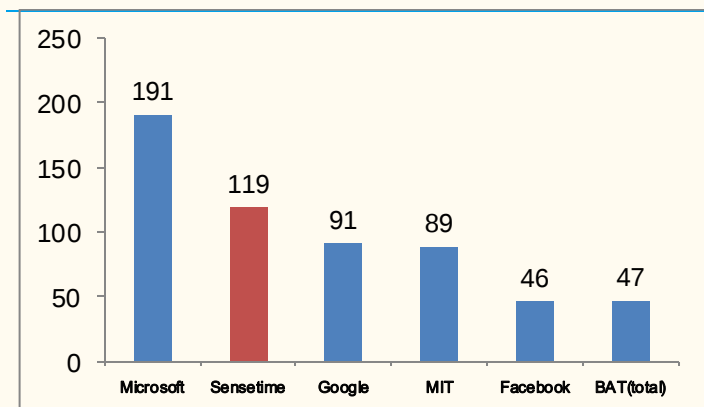
图表 74：商汤近年对外重要投资梳理

时间	标的	方式	标的业务方向
2016	新舟锐视	收购	智能安防摄像头
	Linkface	收购	智慧金融解决方案
2017	51VR	领投	地产数字化、无人驾驶仿真等
2018	禾连健康	领投	互联网医疗
	影谱科技	领投	大文娱领域智能影像生产
	苏宁体育	跟投	体育传媒、体育零售等
	特斯联	跟投	智能物联网

来源：公司公告等公开资料，国金证券研究所整理

- **算法储备丰富、算力供应充沛，长期发展具有深厚底蕴：**根据公开资料统计，公司 2015-2017 年在三大顶级计算机视觉会议（CPVR、ICCV 及 ECCV）上共发表了 119 篇论文，内容涵盖深度学习、无人驾驶、物体检测等诸多领域，数量上仅次于微软和 CMU，远领先于国内其他企业。2018 年，公司以 81 篇 CVPR 和 ECCV 论文创下中国 AI 学术新纪录，展示出越来越强的学术研究能力，也说明公司在算法领域有着丰富的储备。此外，公司自主研发了支持千亿级参数模型、百亿训练样板、亿级类别分类任务的深度学习框架 Parrots，以及拥有超过 6000 块 GPU 并支持千卡并行训练的亚洲最大超算中心，算力供应充沛，保证了模型可以得到快速验证。我们认为，公司重视基础研究，强调技术原创，在算法、算力上都能充分保障自身技术顺畅迭代，长期发展具有深厚底蕴。

图表 75：2015-2017 三大顶级会议论文收录数



来源：公司公告等公开资料，国金证券研究所整理

图表 76：算法和算力支撑公司应用拓展



来源：公司公告等公开资料，国金证券研究所整理

B、旷视：优秀的人脸识别平台企业，战略进军 IoT

- **人脸识别先行者，软硬件结合深入垂直场景：**公司早于“四小龙”其它三家公司成立，三位联合创始人印奇、唐文斌、杨沐均毕业于清华姚班，当前拥有约 1500 名员工。公司早期业务以开发 C 端人脸识别相关互联网应用为主，后转向提供 B 端计算机视觉相关产品和服务，旗下 Face++ 是当前世界上使用量最大的人脸识别引擎。公司主要的业务模式是基于核心计算机视觉技术，深入垂直场景提供行业解决方案。目前，公司旗下的 Face++ 是目前，相关产品和服务已经广泛应用于金融、安防、手机等多个领域，拥有包括阿里、蚂蚁金服、富士康等在内的上千家核心客户。公开信息显示，依靠多领域的业务落地，公司近年来实现了 5 至 7 倍的营收增长，2017 年已实现盈利，主要收入来自于金融和安防。
- **战略进军 AIoT，打造个人、城市、供应链“大脑”：**2019 年 1 月 16 日，公司公布了从 AI 到 AIoT 的全新战略，战略定位从算法提供商全面升级为 AIoT 解决方案提供商，并推出了智能供应链操作系统“河图”。根据我们的梳理，公司目前个人生活大脑、智慧城市大脑、智能供应链大脑三大“大脑”场景业务群，其中：个人生活大脑以连接个人终端为主，聚焦于为 C 端用户提供人脸解锁、人脸支付等服务，应用场景包括手机解锁、身份核验等；智慧城市大脑以连接城市摄像终端为主，聚焦于为 G 端客户提供智能门禁、智能迎宾等产品/服务，应用场景包括智慧社区、教育管理等；供应链大脑以连接工业机器人为主，聚焦于为 B 端客户提供“货到人”、“订单到人”等解决方案，应用场景包括智能工厂、智能仓储、新零售等。近两年公司在外延方面表现活跃，先后投资了机器人、娱乐、零售等多个领域的初创企业，业务布局持续延伸。

图表 77：旷视对外投资梳理

投资时间	投资标的	投资方式	标的方向
2018.04	Video++	投资	AI+文娱
	艾瑞思	收购	智能机器人
2018.11	好邻居	投资	零售，便利店
2018.12	鲜生活	投资	新零售

来源：公司公告等公开资料，国金证券研究所整理

- **强大的软硬件结合能力以及“阿里系”背景是公司核心差异化优势：**公司自 2015 年起涉足硬件，陆续推出了摄像头、人脸识别面板机、人证核验一体机、服务器、搬运机器人等十余款软硬件结合产品，在初创企业中处于绝对领先地位。我们认为，从算法到产品/服务有天然隔阂，优秀的算法并不能保证优秀的产品。公司的历史发展表明其具有强大的软硬件结合能力，能够保障其技术端领先优势在产品端的延续。我们认为，软硬件一体化是打造商业闭环的必要条件，仅依靠算法无法满足行业客户的整体需求，安防行业内 AI 算法提供商可以获得部分项目机会但始终无法撼动安防设备龙头的地位即是最明显的例证。此外，公司还是“四小龙”中唯一一家两次获得“阿里系”投资的企业，我们判断公司是阿里在计算机视觉领域选定的战略抓手，未来“阿里系”可能会有更多资源注入，从支付宝、滴滴、天猫超市等“阿里系”公司相继采用公司产品来看也能部分验证该想法。我们认为，公司软硬件结合能力突出，保障了技术端优势在产品端的延续，具备打造商业闭环的潜力；同时，公司背靠股东“阿里系”，在金融、出行等多个领域有望被注入更多资源，有较强的盈利支撑，因此具备极大竞争优势。

C、依图：注重实战，业务聚焦于安防和医疗

- **实战型 AI 公司，善于从行业里寻找研究课题：**公司成立于 2012 年，创始人 ULCA 统计学博士朱珑及前阿里资深云计算专家林晨曦，当前拥有约 700 名员工。公司与“四小龙”中其他几家企业最大的不同点是其技术并非来自创始团队的科研成果转化，而是从零开始，先选择行业，然后针对企业需求进行课题研究。基于这种打法，公司于 2013 年为苏州公安开发了车辆识别系统，将套牌车的识别率从不足 30% 提高到 90%，在初战告捷

后又推出了人脸识别系统并同样取得成功，吸引了公安客户。我们认为，公司的这种基因和业务模式决定了其更加注重实战，也意味着具有更强的行业痛点挖掘能力。

- **业务聚焦于安防和医疗，积极布局智慧金融：**公司的战略是深耕行业、深挖痛点，业务聚焦于安防和医疗两个行业：安防领域，公司的主要产品为“蜻蜓眼”系列，包括蜻蜓眼人像大平台和蜻蜓眼车辆大平台，目前已在全国二十多个省份、上百个地市公安系统以及海关总署、中国边检等得以应用；医疗领域，公司的主要产品为 care.ai 系列，深入疾病预防、检查、诊断、治疗和科研等各个环节，目前已落地全国 100 多家三甲医院。除安防和医疗外，2015 年公司通过为招商银行提供人脸识别技术切入金融领域，相继提出了、远程核身、刷脸取款、智能网点等多个解决方案，客户以银行和互联网金融公司为主。此外，公司还通过外延方式进一步布局，除继续深化在医疗领域的布局外，还投资了芯片初创企业 ThinkForce，展示了进军芯片的意图。
- **先发优势明显，实战指标领先：**公司是最早进入智能安防领域的 AI 企业之一，2014 年起即已实现智能安防产品的规模化落地。公司实战指标处于国内领先地位。以人脸识别的准确率为例，公司在由美国国家标准局 NIST 组织的人脸识别供应商测试 FRVT (2018) 最新报告中综合排名世界第一，在相同误报率下漏报率是商汤的 30%-45%。不同于由学术机构组织的人脸识别竞赛，FRVT 更贴近安防实战，测试集来自美国国土安全局真实业务场景，数据更加复杂和多样，因此参赛者的成绩很大程度上可以代表其实战水平。我们认为，公司实战指标之所以领先，是由于其人脸识别产品较早实现了规模化落地，因而在使用真实场景数据改善算法方面具有先发优势，这对其技术进步起到了巨大的推动作用。

图表 78：FRVT(2018)比赛结果（部分）

#	Developer	VISA Photos FNMR@ FMR ≤ 0.000001	VISA Photos FNMR@ FMR ≤ 0.0001	MUGSHOT Photos FNMR@ FMR ≤ 0.0001	WILD Photos FNMR@ FMR ≤ 0.00001	CHILD EXP Photos FNMR@ FMR ≤ 0.01	Submission Date
1	yitu-002	0.004 ¹	0.001 ¹	0.013 ⁷	0.052 ¹³		2018_10_19
2	yitu-001	0.007 ²	0.003 ⁷	0.013 ⁸	0.058 ²⁶	0.579 ¹³	2018_06_12
3	sensetime-001	0.009 ³	0.003 ⁶	0.013 ¹¹	1.000 ⁷⁶		2018_10_19
4	sensetime-002	0.010 ⁴	0.003 ¹⁰	0.015 ²⁹	1.000 ⁷⁷		2018_10_19
5	siat-002	0.013 ⁵	0.004 ¹⁵	0.014 ¹⁵	0.055 ²⁰	0.428 ³	2018_06_13
6	ntechlab-004	0.013 ⁶	0.003 ⁴	0.013 ¹²	0.046 ⁶	0.420 ²	2018_06_14
7	ntechlab-005	0.014 ⁷	0.002 ²	0.013 ¹⁰	0.050 ¹⁰		2018_10_19
8	megvii-002	0.014 ⁸	0.004 ¹²	0.030 ⁶³	0.071 ³⁵		2018_10_19

来源：FRVT(2018)，数据截止到 2019 年 2 月 22 日，国金证券研究所

D、云从：AI 国家队，银行、机场人脸识别产品第一大供应商

- **公司是银行和机场人脸识别产品第一大供应商：**公司成立于 2015 年，创始人是 UIUC 博士、中科院“百人计划”人选周曦，当前拥有员工 400 余人。公司聚焦于计算机视觉技术，产品以人脸识别、文字识别、活体检测为主，应用方向集中在金融、交通、安防三大行业，目前是国内银行和机场人脸识别产品的第一大供应商：在银行领域，截至 2018 年 3 月，在 121 家已完成招标、具备自建系统能力的银行中，公司中标了包括四大行之内的 88 家总行平台，市场占有率高达 72.7%；在机场领域，截至 2018 年 7 月，已有 54 家机场选择公司产品，覆盖全国 75% 的枢纽机场。
- **公司深耕三大行业，布局智慧商业：**公司的业务布局思路为在金融、交通、安防三大行业内深挖应用场景，同时通过合作布局商业。例如，在金融行业，公司积极拓展非银机构，向杭州消费金融、分期 GO 等互联网金融公

公司提供实名认证等服务；在交通行业，与中国民航管理干部学院共同成立智慧民航联合研究中心引领机场智慧化等等。在商业领域，公司中标中国电信生物识别平台项目，将联手电信实现全网实名制，杜绝电信诈骗。“国家队”背景有助于公司拓展安全敏感行业。公司是孵化于中科院重庆研究院的高科技企业，股东名单中有多只国家战略投资基金，是“四小龙”中唯一没有外资入股的企业。此外，公司还是唯一一家参与人脸识别国标、部标、行标制定的研发企业，并且承担了发改委牵头的国家人工智能基础平台重大项目。我们认为，上述事实给公司贴上了“国家队”的标签，在国内核心技术自主可控的大趋势下，公司的“国家队”背景有助于其在安全敏感行业的拓展。

E、云知声：“云端芯”战略，重点布局家居、车载及医疗三大领域

- **公司是优秀的智能语音专业应用提供商，多个垂直领域市场地位领先：**公司成立于 2012 年，发展至今历经三个阶段：第一个阶段为以发育云端语音识别能力为主，通过开放语音平台吸引了近一万家企业客户接入；第二个阶段定位做 AIoT，开始在‘端’上发力，探索技术的产品化，例如与乐视合作推出可以进行远距离自然语言交互的乐视第三代 TV；第三个阶段场景落地阶段，通过为家电厂商美的、汽车厂商上海通用、公立医院北京协和等提供智能语音解决方案，落地智能家居、智能车载、智能医疗等场景。公司官网显示，公司目前的合作伙伴已经超过 2 万家，覆盖用户达 2 亿，其中开放语音云覆盖的城市超过 470 个，覆盖设备超过 9000 万台。此外，公开资料显示，在智慧教育领域，公司的教育评测平台的高峰调用量达到每天 3 亿次；在智能车载领域，已经落地 30 多个车型；在智慧医疗领域，已与 500 多家医院达成合作，均处于同行业领先地位。
- **“云端芯”全产业链布局，重点经营家居、车载、医疗三大领域，抢占教育、机器人赛道：**公司 2014 年正式定位做 AIoT 服务商，确立了“云端芯”战略，即将基于云端的智能语音技术与终端设备和芯片紧密联合，通过云端提供服务、设备端实现算法、芯片提供算力的全栈方式打造生态闭环。公司近一年在芯片端的发力迹象比较明显，于 2018 年 5 月制定了 UniOne 系列芯片计划，并推出了面向智能家居和智能音箱的第一代芯片“雨燕”，同时计划于 2019 年和 2020 年推出面向智能车载领域的第二代芯片“雪豹”和面向智慧城市的第三代芯片“旗鱼”。
- **公司重点经营家居、车载以医疗三大领域：**家居领域，面向家电厂商推出了 UniHome 智能语音解决方案芯片；车载领域，面向车联网产品开发商以及整车厂商推出了 UniCar “云端芯”一体化解决方案；医疗领域，面向医生推出了智能语音录入软硬件一体化解决方案。除上述三大领域外，公司还积极抢占教育和机器人赛道：教育领域，推出了面向教育机构的智能语音评测系统；机器人领域，推出了软硬件一体的行业服务机器人和面向儿童早教的 UniToy 机器人解决方案，此外还有面向客服的 SaaS 模式的 SAMANTHA 智能电话机器人。
- **人才优势、先发优势铸就公司核心竞争力：**根据 IDC 的调研，对话式人工智能落地面临的两个最大问题是人才及落地案例的缺乏。公司的核心团队来自 IBM、摩根大通、中科院、剑桥等全球顶尖公司和院校，人工智能研究院研究员中博士占比超过 45%，产品和运营团队来自 Nuance、阿里巴巴等国内外知名公司，人才优势明显；另外，公司较早进入家居、车载、医疗领域，在跟 B 端企业合作上以及终端用户需求的理解和隐私保护等方面都具备较多 know how，后进者难以在短时间内完成追赶。我们认为，人才优势结合先发优势助力公司在场景拓展和深耕上都具备极大竞争优势，看好公司发展。

四、投资建议

国内重点关注公司：华为海思（半导体设计），寒武纪（设计），海康威视（安防 AI 系统），商汤（算法软件），伊图（算法软件）；

全球重点关注公司：英伟达（云端 AI 训练芯片），谷歌（半导体及 AI 云，边缘运算系统整合），及赛灵思（AI 推理芯片）及台积电（AI 芯片制造），新思（AI 专利权区块）。

五、风险提示

1、中美贸易及关税战是否趋缓？

中央及地方政府是否应美国要求减少或停止对 AI 芯片设计公司研发的补贴，华为孟晚舟女士是否被引渡及华为产品是否持续被各国抵制。

2、国内 AI 芯片大厂能否摆脱亏损？

除了国内算法软件公司及美国 AI 芯片大厂英伟达及赛灵思外，未来 24 个月因为国产替代效应不够快及飘高的设计成本，我们看不出来国内 AI 芯片大厂能摆脱亏损。

3、禁售人工智能新兴技术？

美国商务部工业安全局可能对其 11 项 AI 和机器学习技术((i) 神经网络和深度学习（例如，大脑建模，时间序列预测，分类）；(ii) 进化和遗传计算（例如遗传算法，遗传编程）；(iii) 强化学习；(iv) 计算机视觉（例如，物体识别，图像理解）；(v) 专家系统（例如决策支持系统，教学系统）；(vi) 语音和音频处理（例如，语音识别和制作）；(vii) 自然语言处理（例如机器翻译）；(viii) 规划（例如，安排，游戏）；(ix) 音频和视频处理技术（例如，语音克隆，深度伪造）；(x) AI 云技术；(xi) AI 芯片组）列入出口管制清单？

4、系统设计机密不愿分享：

在调研安防 AI 公司多次提到 AI 半导体 ASIC 厂商对其各种安防应用系统设计认知不足，所以采用其 AI ASIC 芯片缓慢，而另一方面，安防 / 语音 / 自驾系统公司也不太愿意分享其系统设计机密于 AI ASIC 厂商用以优化其芯片设计，以防竞争者切入。

公司投资评级的说明：

买入：预期未来 6—12 个月内上涨幅度在 15%以上；

增持：预期未来 6—12 个月内上涨幅度在 5%—15%；

中性：预期未来 6—12 个月内变动幅度在 -5%—5%；

减持：预期未来 6—12 个月内下跌幅度在 5%以上。

行业投资评级的说明：

买入：预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上；

增持：预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%；

中性：预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%；

减持：预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。

特别声明：

国金证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证，对由于该等问题产生的一切责任，国金证券不作出任何担保。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整。

本报告中的信息、意见等均仅供参考，不作为或被视为出售及购买证券或其他投资标的邀请或要约。客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告反映编写分析员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，且收件人亦不会因为收到本报告而成为国金证券的客户。

根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级（含 C3 级）的投资者使用；非国金证券 C3 级以上（含 C3 级）的投资者擅自使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

此报告仅限于中国大陆使用。

上海

电话：021-60753903

传真：021-61038200

邮箱：researchsh@gjzq.com.cn

邮编：201204

地址：上海浦东新区芳甸路 1088 号

紫竹国际大厦 7 楼

北京

电话：010-66216979

传真：010-66216793

邮箱：researchbj@gjzq.com.cn

邮编：100053

地址：中国北京西城区长椿街 3 号 4 层

深圳

电话：0755-83831378

传真：0755-83830558

邮箱：researchsz@gjzq.com.cn

邮编：518000

地址：中国深圳福田区深南大道 4001 号

时代金融中心 7GH