

车载人工智能计算芯片 白皮书 (2021 年)

中国软件评测中心

智能网联驾驶测试与评价工业和信息化部重点实验室

赛迪（浙江）汽车检测服务有限公司

二〇二一年十二月

顾问

尚进 清华大学 客座研究员

国汽智控（北京）科技有限公司 总经理

指导专家

安晖 中国电子信息产业发展研究院 副总工程师

金颖 华为技术有限公司 昇腾 AI 芯片产品领域总经理

李星宇 北京地平线机器人技术研发有限公司 生态发展与战略规划副总裁

王平 寒武纪行歌（南京）科技有限公司 执行总裁

邓堃 黑芝麻智能科技有限公司 应用工程副总裁

王超 南京芯驰半导体科技有限公司 高级项目总监

王云 中国科学院微电子研究所 汽车电子研发中心主任

何丰 北京赛目科技有限公司 总经理

褚文博 国家智能网联汽车创新中心 计算平台部部长

丛炜 国汽智控（北京）科技有限公司 战略发展副总裁

特别鸣谢

刘寻 刘喜喜 雷蕾 刘妍

编写单位（按笔画排序）

中国软件评测中心（工业和信息化部软件与集成电路促进中心）

北京地平线机器人技术研发有限公司

华为技术有限公司

国汽智控（北京）科技有限公司

国汽（北京）智能网联汽车研究院有限公司

南京芯驰半导体科技有限公司

黑芝麻智能科技有限公司

参研单位（按笔画排序）

一汽(南京)科技开发有限公司

上海蔚来汽车有限公司

中国科学院微电子研究所

北京智行者科技有限公司

北京赛目科技有限公司

安谋科技（中国）有限公司

汽车电子产业联盟

郑州宇通集团有限公司

湖北芯擎科技有限公司

寒武纪行歌（南京）科技有限公司

撰写专家

王 伟 程智锋 刘艳玲 余 云 关 贺 夏 萌 陶 圣

邹博松 林志杰 贺龙龙 钟 薇 薛 雷 乌尼日其其格

版权声明

本白皮书版权属于中国软件评测中心，并受法律保护。转载、摘编或利用其它方式使用本白皮书文字或者观点的，应注明“来源：中国软件评测中心”。违反上述声明者，本单位将追究其相关法律责任。

CSSTC 中国评测

目 录

编制概要	1
(一) 编制方法	1
(二) 特别说明	1
一、研究背景	2
(一) 自动驾驶是汽车行业转型升级的主要方向，人工智能极大地推动了自动驾驶的发展	2
(二) 车载人工智能计算芯片成为行业竞争热点，国内外企业竞相发力 ..	3
(三) 车载人工智能计算芯片缺少共识，亟待行业专家联合研究	4
二、车载人工智能计算芯片概述	5
(一) 车载人工智能计算芯片及技术概览	5
(二) 关键指标	5
三、车载人工智能计算芯片关键技术分析	10
(一) 架构设计	10
(二) 软硬件协同设计	10
(三) 互联通信和存储	11
(四) 可靠性与安全性	12
(五) 算力与制程工艺选择	14
(六) 基准测试	14
四、车载人工智能计算芯片产品及应用	16
(一) 产业链分析	16
(二) 车载人工智能计算芯片典型产品及应用	18
五、发展建议	26
(一) 完善产业布局 促进技术创新	26
(二) 促进企业交流 推动产业发展	26
(三) 完善配套保障 提升技术实力	26
附件：缩略语	28

编制概要

（一）编制方法

一是研究学习国内外相关文献，充分参考借鉴国内外最新产业动态和研究成果。

二是调研国内外车载人工智能计算芯片相关企业，汇集整理和分析来自实践应用的相关素材。

三是邀请行业专家进行技术研讨和咨询评审。

（二）特别说明

1) 研究聚焦车载人工智能计算芯片

近年来，人工智能技术广泛应用于汽车行业，其车内应用主要包括智能座舱和智能驾驶两部分。本白皮书重点关注人工智能芯片在自动驾驶计算中的应用。为凝聚共识、形成合力，加强车载智能计算平台的研发和应用，推动智能网联汽车产业高质量发展，在《车载智能计算平台白皮书（2018年）》《车载智能计算基础平台参考架构 1.0（2019年）》《车载智能计算平台功能安全白皮书（2020年）》研究的基础上，本白皮书聚焦车载人工智能计算芯片，重点分析其关键指标、关键技术、产业链现状及典型产品的应用。

2) 研究内容仍有待进一步丰富完善

本白皮书的主要观点和内容仅代表编制组目前对车载人工智能计算芯片的研判和思考，欢迎各方专家学者和企业代表提出宝贵意见，共同推动研究的及时更新和纠偏。后续中国软件评测中心将会继续推出《汽车智能计算平台白皮书（系列）》。

一、研究背景

(一) 自动驾驶是汽车行业转型升级的主要方向，人工智能极大地推动了自动驾驶的发展

智能网联汽车产业规模迅速增长，自动驾驶成为主要研发方向。智能网联汽车为满足出行安全、舒适、高效、环保提供了新的解决方案，已经成为各国必争的战略高地。国内智能网联汽车销量不断增加，2020年，我国L2级智能网联乘用车销量为303万辆，同比增长达到107.5%，渗透率达到15%（数据来源：中国汽车产业发展年报2021）。智能驾驶产品及解决方案提供商、整车企业纷纷通过自研、合作等方式布局自动驾驶功能的研发，以期在未来行业竞争中占据有利位置。

人工智能技术的应用推动了自动驾驶的发展。目前深度学习算法在自动驾驶感知领域应用广泛。自动驾驶感知系统利用传感器（如摄像头、毫米波雷达、激光雷达）采集的数据，为规划和决策环节识别出目标的关键信息。该过程主要通过应用深度学习算法加以实现。深度学习算法实现了利用计算机对图像特征进行提取，保证了提取特征的丰富性与全面性，提高了目标检测、预测的准确性。目前应用于自动驾驶感知领域的深度学习算法包括目标检测与识别算法、图像分割算法等。此外，深度学习算法正在逐渐向自动驾驶规划、决策环节渗透。

（二）车载人工智能计算芯片成为行业竞争热点，国内外企业竞相发力

车载智能计算平台是实现自动驾驶的必要解决方案，其中人工智能芯片是其核心硬件之一。自动驾驶运行过程中需要完成大量高速计算以处理海量、多源、异构的数据，并需要运用人工智能、信息通信、大数据和云计算等技术进行实时感知、决策、车辆运动规划及控制。高性能车载智能计算平台成为必需。车载智能计算平台的硬件架构采用异构芯片集成化设计，为自动驾驶提供算力支持和安全保障，其硬件主要包括 AI（Artificial Intelligence，人工智能）单元、计算单元、控制单元等。其中，AI 单元主要用于提供车载智能计算平台的 AI 算力支撑，通常包含若干颗人工智能芯片。随着自动驾驶的不断发展，自动驾驶功能的计算需求持续提升，客观上对人工智能芯片的算力提出了更高的要求。

先进制程工艺半导体产业链垂直化分工愈发明确，芯片设计成为国内外企业的发力点。半导体行业具有生产工序多、技术更新快、投资风险高等特点，尤其是 AI 芯片所需的先进制程工艺的制造厂投资巨大。为了降低投资压力、维持竞争力，许多集成化芯片企业对其半导体业务实施重组或剥离，形成了芯片设计、晶圆代工、封装测试等类型的企业。芯片设计企业可以将主要精力专注于芯片设计，而无需面对大规模的工厂投资。目前，无论是国外的英伟达、赛灵思等，还是国内华为、地平线、芯驰、黑芝麻等，都将精力集中于芯片设计方面，而将晶圆代工及封测等业务交给其他企业代工完成（晶圆代工企

业，如台积电等；封测企业，如日月光等）。国内外整车企业也积极布局开展芯片设计工作，如特斯拉、吉利、蔚来等。

（三）车载人工智能计算芯片缺少共识，亟待行业专家联合研究

当前车载人工智能计算芯片存在关键指标要求及关键技术标准缺失、滞后等问题，梳理其关键指标、关键技术有利于形成行业共识。

“硬件趋同、软件定义、数据驱动”成为汽车行业的发展方向。车载人工智能计算芯片是车载智能计算平台的“大脑”，对其关键指标、关键技术进行梳理，有利于明确其发展的技术路线，加速软硬解耦，为标准研制提供参考，并促进车载人工智能计算芯片产业生态的持续健康发展。

梳理车载人工智能计算芯片产业链现状有利于完善产业布局。车载人工智能计算芯片产业链包括了半导体材料、半导体设备、芯片设计、晶圆制造、封装测试、芯片应用等多个环节，且各环节之间相互联动、相互依存。车载人工智能计算芯片对智能网联汽车的发展至关重要，对其产业链展开分析，有利于完善产业布局，并对保障智能网联汽车的供应链稳定发挥积极作用。

二、车载人工智能计算芯片概述

(一) 车载人工智能计算芯片及技术概览

车载人工智能计算芯片是符合车规级要求的、用于运行自动驾驶人工智能算法的芯片。

车载人工智能计算芯片涉及的技术主要包括半导体制造及封装技术、芯片设计、人工智能算法设计、车载人工智能计算芯片应用四个层级，车载人工智能计算芯片的技术概览图如下图 1 所示。各个层级之间紧密联系，一方面，半导体制造及封装技术、芯片设计、人工智能算法设计的发展使芯片性能得到不断提升，从而促进了车载人工智能计算芯片自动驾驶方面的应用，另一方面，车载人工智能计算芯片的实际应用及应用需求的不断提升，推动着半导体制造及封装技术、芯片设计、算法设计的不断发展。



图 1 车载人工智能计算芯片技术概览图

(二) 关键指标

车载人工智能计算芯片通常以 SoC (System on Chip, 片上系统) 的形态呈现, 其组成包括 CPU (central processing unit, 中央处理器)、AI 处理器等, 具有涉及技术面广、应用场景复杂多变等特点。车载人

工智能计算芯片设计过程中存在多项指标要求，掌握这些指标对芯片的设计及应用过程具有良好的指导意义。本部分重点对其关键指标展开分析。

1) 算力

特定场景下对芯片计算能力评价的一个维度。算力大小代表芯片数字化信息处理能力的强弱。自动驾驶场景需要标量、矢量、矩阵三者结合的异构算力，通常可以将算力的综合评价分为两方面，即 AI 算力和 CPU 算力。

2) AI 算力

AI 处理器在特定场景下提供的矢量和矩阵计算能力。AI 算力常用的单位是 TOPS (Tera Operations Per Second) 或 TFLOPS(Tera Floating-point operations per second)，1TOPS 代表 AI 处理器每秒可进行一万亿次 (10^{12}) 定点操作，1TFLOPS 分别代表 AI 处理器每秒可进行一万亿次 (10^{12}) 浮点操作。

3) CPU 算力

CPU 主要提供的标量算力。CPU 算力常用的单位是 DMIPS (Dhrystone Million Instructions executed Per Second)，其含义为每秒钟执行基准测试程序 Dhrystone 的次数除以 1757(这一数值来自于 VAX 11/780 机器，此机器在名义上为 1MIPS 机器，它每秒运行 Dhrystone 次数为 1757 次)。

4) AI 算力利用率

也称 MAC (Multiplier and Accumulation, 乘累加器) 利用率，指

特定负载(神经网络)下 AI 处理器中用于矩阵计算的 MAC 阵列利用率。AI 算力利用率的计算通常采用特定网络的单次计算量与 AI 理论算力(OPS)和推理时延(s)乘积的比值。

5) 功耗

芯片或处理器在单位时间中所消耗的能量。常用单位为瓦(W)。实际中功耗的度量一般采用典型功耗,典型功耗是指芯片或处理器在典型负载工况下(如基于特定测试数据集,运行特定神经网络,采用特定数据精度、批量及图像分辨率的工况)获取的实际功耗。典型功耗的测试场景与实际生产环境更为接近,具备一定的代表性。

6) 能耗比

用于度量在单位功耗下芯片或处理器的 AI 计算能力。常用单位为 TOPS/W 或 TFLOPS/W。对于车载人工智能计算芯片来说,能耗比至关重要,能耗比高的芯片或处理器可以用更少的能量完成 AI 计算。能耗比的计算通常采用芯片或处理器单位时间内的推理次数除以对应功耗。根据功耗对象选择的差异,能耗比通常可分为 AI 处理器能耗比(功耗采用特定负载工况下获得的 AI 处理器功耗)与 SoC 能耗比(功耗采用特定负载工况下获得的 SoC 功耗)两类。

7) 时延

芯片完成一次完整的输入、处理、输出过程的时间。常用单位为毫秒(ms)。对于自动驾驶场景,不仅需要关注运行感知神经网络的时延,还应关注端到端时延。端到端时延主要包括三点:一是传感器通过芯片接口将传感器数据输入 SoC,通过 ISP(Image Signal

Processor，图像处理器）等做信号处理并将处理后的数据写入 DDR（Double Data Rate Synchronous Dynamic Random Access Memory，双倍速率同步动态随机存取内存）的时延；二是 AI 处理器从 DDR 读入数据并进行感知，融合及定位处理的时延；三是基于感知及定位的结果进行路径规划及控制，并完成输出信号发送的时延。

8) 计算精度

芯片的计算结果相比预期结果的精确程度。芯片的计算精度越高，说明计算结果越接近真值。提升计算精度通常需要增加芯片的计算资源，相应的会带来成本和功耗的增加。实际中，计算精度还受到模型、数据集等多重因素的影响，不同的数据集对计算精度的要求不同。

9) 数据精度

芯片支持的数据计算精度类型，如 INT8，INT16，FP16，FP32。

10) 带宽

芯片支持访问内存数据的速率。该指标代表了内存总线所能提供的数据传输能力。目前，带宽要求主要包括两个方面，一是片内缓存的访问速率，二是片外内存的访问速率。带宽会对 AI 处理器的利用率、端到端时延、功耗等产生影响。

11) 泛化能力

芯片采用特定数据集训练的特定模型对不同测试场景的适应能力。这种适应能力通常会选用多种模型，测试每个模型的算力时延比及精度，综合评价芯片的泛化能力。

12) 最大感知能力

芯片最大支持的感知类传感器输入数据的处理能力。通常需要按照芯片最大支持的感知类传感器数目及其性能要求，应用给定的测试数据和测试模型（与传感器数目对应），完成数据处理过程，获取芯片输出的时延、速率、精度等参数，实现对芯片最大感知能力的评定。

13) 安全性

主要包括功能安全、信息安全等安全要求。

14) 可靠性

描述芯片可以正常、准确、稳定地发挥其功能和性能的能力和程度。

15) 开放性

提升芯片的开放性即从开发层面开始为用户（如应用开发工程师、中间件开发工程师、算法工程师等）提供一个具有高度可扩展性、兼容性、灵活性的开发平台，降低不同框架适配迁移可能产生的不可预知成本，实现提高开发效率、缩短开发周期等要求。开放性主要体现在三个方面：一是提供统一开放的编程接口，便于自动驾驶开发人员调用，有效实现 SoC 资源的调用及异构算力的编排；二是提供统一开放的神经网络及算子开发接口；三是提供不同框架下网络模型的转换、量化、压缩工具，以及算子、整网的调优工具，便于快速开发、调试及调优工作的实施。

三、车载人工智能计算芯片关键技术分析

(一) 架构设计

车载人工智能计算芯片采用多核异构的集成化设计,其硬件主要包括计算核心、系统外设、数据外设三部分。一是计算核心,主要包括 CPU 核心、AI 处理器等。CPU 主要用于提供标量算力。AI 处理器主要用于提供 AI 算力支撑,其类型主要包括神经网络计算单元、图像处理单元、数字处理单元等一种或多种组合。二是系统外设,主要用于维持芯片基本的正常运转,主要包括电源、复位、计数器、看门狗等。三是数据外设,主要用于芯片和外界进行数据交换,主要包括模数转换、总线控制器、总线收发器、通用输入输出接口等。

架构设计在安全性方面应考虑功能安全、信息安全的协同设计。功能安全方面,需要采用系统化设计的思想,即目标应用中涉及安全的模块,都需要满足功能安全设计要求。信息安全方面,通常采用集成的硬件安全模块,为系统提供全方位的数据保护。此外,在车载人工智能计算芯片的安全设计过程中,由于功能安全以及信息安全不是孤立的,需同时考虑这两个维度的要求,通过对二者进行协同设计来保证系统的安全性。

(二) 软硬件协同设计

软硬件协同设计可以实现车载人工智能计算芯片的综合优化。车载人工智能计算芯片的算法与硬件架构深度耦合,二者的协同设计是实现高效人工智能计算的重要手段。软硬件协同设计通过软件设计与

硬件设计的并行开发与相互反馈，协调软件和硬件之间的制约关系，有利于缩短开发周期并降低开发成本，其重点和难点在于软件任务与硬件任务的划分与调度，实现在满足目标约束的条件下提升芯片的整体计算能力。此外，软硬件协同设计可采用计算芯片作为中心载体，通过构建完备的异构计算体系，支持数据驱动人工智能系统进化，即利用实际驾驶场景获取数据，通过系统开发、测试、安全评估等自动化迭代过程，促进计算芯片架构的优化设计。

软硬件协同设计需要综合考虑多个因素的影响。软硬件协同设计需要根据自动驾驶的应用场景设计深度学习算法，制定合理的算法优化与硬件架构调整策略，通过反复迭代达到硬件加速的需求。在开发策略方面，应考虑平衡精度和性能、数据链路的完整性（前后处理和网络加速相互配合）、多模型灵活调度、网络加速与应用直接配合等。在用户应用方面，需要考虑软件开发工具链与车载人工智能计算芯片的适配程度，并保证软件开发工具的开放性与便利性。此外，为实现兼容算法设计的快速迭代，硬件架构设计应具备一定的前瞻性以适配汽车电气架构的设计要求。

（三）互联通信和存储

互联通信主要包含外部通信总线和内部通信总线。外部通信总线主要用于外部通信和外部扩展，既包含高速总线，如千兆/万兆以太网、PCIe3.0/4.0 等；又包含传统总线，如 MIPI、CANFD、CAN、SPI、UART、I2C 等；还包括存储扩展总线，如 LPDDR、eMMC、UFS 接口等。为满足不同级别自动驾驶开发的灵活性要求，单个芯片设计算

力通常可以满足当前辅助驾驶的发展要求，多个芯片通过总线互联的方式（如使用 PCIe3.0/4.0 总线）满足更高的算力需求。内部通信总线用于实现芯片内部多个组件之间的数据交互。随着自动驾驶对数据需求的日益提高，芯片内部各组件间的数据交换需求不断提升，内部总线通信通常采用增加片内带宽、应用低功耗管理技术、优化内部空间布局等方法进行统筹设计，以满足内部数据传输的快速性、稳定性等要求。此外，未来应用缓存一致性总线可以实现多个芯片的缓存一致性互联，从而为自动驾驶提供跨片缓存无缝扩展能力。

内存布局通常采用片外内存加多级片内缓存的方式，存内计算等新型内存技术亟待突破。内存带宽是限制车载人工智能计算芯片计算性能提升的一个关键要素，数据的存取速度直接影响着 AI 芯片的处理速度。为了提升数据的存取速度，当前 AI 芯片常采用片内缓存加片外内存的方案。对于片外存储，目前常采用高带宽存储器或低功耗双倍速率同步动态随机存储器。对于片内存储，目前常采用静态随机存取存储器，且可使用增加片内缓存容量方法提高缓存带宽。在内存新技术研究方面，存内计算将计算和存储合二为一，通过电路设计使存储器具有计算能力，将大部分计算放在存储器内完成，从而减少了处理器访问存储器的频率，如基于 NOR 闪存的存算一体人工智能推理芯片。应用该技术可以使芯片实现低能耗、高效率、低延时等优点。目前，受限于技术原因，存内计算尚未得到大规模应用。

（四）可靠性与安全性

车载人工智能计算芯片需满足严苛的设计标准、规范的体系要求

以及生产测试要求，以实现高可靠性和长期（>10 年）的质保承诺。车载芯片的可靠性要求指芯片设计的运行环境应覆盖车辆及其电控单元的应用环境要求，如温度、湿度等要求。可靠性设计需考虑晶体管老化导致晶体管性能变化，以及片上金属导线在大电流与高温下长期工作带来的可靠性问题。需要结合车规级芯片的规格要求进行电子迁移等规则检查，避免局部电流过大引起导线的可靠性问题。同时为满足部分电路大电流负载需要，对导线做加宽设计。在生产制造阶段还应考虑选择符合要求的工艺和材料。目前可靠性测试的主要参考标准是 AEC-Q100。AEC-Q100 规定了一系列可供参考的可靠性测试，同时定义了应力测试驱动型认证的要求以及 IC 认证的参考测试条件；根据工作温度、耐久性与可靠度等工作要求差异，AEC-Q100 将汽车电子产品分为四个等级，并提出了相应的测试要求。

车载人工智能计算芯片安全性主要考虑功能安全和信息安全。功能安全主要考虑复杂的车载系统中芯片可能出现的失效，通过合理的安全机制的监控和诊断，及时发现潜在的失效，并及时进行失效处理，功能安全的主要参考标准是 ISO 26262 《道路车辆 功能安全》。信息安全主要考虑安全启动、用户认证、访问控制、数字签名、数据加密与解密等要求，常通过支持硬件加密与解密（含国密算法）、密钥管理及验签、安全/可信启动、安全升级、数据安全存储、安全与非安全域隔离等方式加以实现。目前信息安全的主要参考是 ISO/SAE 21434 《道路车辆 网络安全》及密码安全相关的标准。此外，在芯片软件过程改进及能力评定方面，需参考 A-SPICE 相关标准。

（五）算力与制程工艺选择

峰值算力并不能完整体现出 AI 处理器的真实处理能力，该能力同时受到多种因素的影响。汽车自动驾驶的发展对芯片算力的要求不断提升，通常业界会将峰值算力选做衡量车载人工智能计算芯片的主要指标，但峰值算力只反映其理论上的最大 AI 计算能力，实际应用场景中 AI 处理器的处理能力受到软件 SDK（Software Development Kit，软件开发工具包）、模型、环境、模型输入等多重因素的影响，仅使用峰值算力单个指标来判断芯片的 AI 计算能力具有很大的局限性。此外，由于芯片算法演进速度远超硬件改进速度，使得 AI 性能的评估方法与不断演进的算法之间存在脱节的问题。

芯片制程工艺不断提升，其选择需综合考虑成本、功耗、车规级要求等多个因素。芯片性能需求的不断提升驱动着芯片制程工艺的不断发展。制程工艺的选择需要考虑车规级芯片的特殊要求，如代工厂是否符合 IATF16949 质量管理体系的要求、制造工艺本身是否已经通过 AEC-Q100 可靠性验证并达到目标产品的可靠性等级要求、是否可以采购针对该工艺节点的成熟 IP 等。制程工艺的选择还需要综合考虑制程发展、成本、供应链等因素，国内芯片厂商主要以实现量产化为目的，而非单纯追求先进性。

（六）基准测试

基准测试可作为评价人工智能芯片性能的参考依据之一，目前缺乏面向车用领域的专用基准测试方法。各大公司针对芯片关键性能

（尤其在算力方面）展开了激烈的竞争，基准测试成为展现其芯片性能的重要途径。国内外多个行业组织机构都推出了相关 AI 性能评测方法或标准，表 1 举例展示了四种基准测试方法。现有的 AI 芯片基准测试评价对象以偏通用领域人工智能芯片为主，兼顾个别细分领域。关注的内容主要包括测试指标选择、测试场景选择、以及测试结果的表示等。目前 AI 芯片基准测试常选用典型神经网络模型的测试结果来评价 AI 芯片整体性能，该方法面临的主要挑战在于需要不断提升芯片的泛化能力，通常芯片应用典型网络模型可以获得良好的性能表现，但当典型神经网络模型被更换为其他类型或其自身出现更新时，芯片性能可能表现出某种程度的下降。车载人工智能计算芯片属于 AI 芯片的一个细分领域，有其特殊的场景、模型、性能等要求，目前亟待提出面向车载人工智能计算芯片的专用基准评测标准。

表 1 国内外基准测试标准或方法简介

标准或方法	发起方	简介
MLPerf	由谷歌、百度等联合发布	适用于从智能手机到数据中心服务器的系统，目前推理最新版本为 v0.7
AI benchmark	由苏黎世联邦理工学院开发	主要面向手机测试，目前最新版本 V4
ITU-T F.748.11 (2020)	由中国信通院和华为联合牵头发布	评测场景以视觉、语音、推荐为主，主要测试指标包括时延、吞吐、能耗比和准确率等
基于 MAPS 的性能测试方法	由地平线提出	针对智能网联汽车视觉感知计算应用场景的特点，重点关注在精度有保障范围内的平均处理速度，采用可视化、量化表达，便于横向比较。采用该方法的团体标准正在制定中。

四、车载人工智能计算芯片产品及应用

(一) 产业链分析

车载人工智能计算芯片产业链主要分为上中下游三部分。上游主要包括半导体材料、半导体设备、IP核（Intellectual Property core）设计、EDA(Electronic Design Automation, 电子设计自动化)工具四个方面，中游主要包括芯片设计、晶圆制造、封装测试三个方面，下游主要包括车载人工智能计算芯片在自动驾驶行业的应用。产业链简图如下图 2 所示。



图 2 车载人工智能计算芯片产业链简图

半导体材料方面，晶圆制造材料主要包括硅片、光掩模、光刻胶、光刻胶辅助材料、工艺化学品、电子特气、靶材、化学机械抛光材料（抛光液和抛光垫）等，封装材料主要包括引线框架、封装基板、陶瓷基板、键合丝、包封材料、芯片粘结材料等。2020 年全球硅片市场

前五的企业为日本信越化工、日本 Sumco、德国 Siltronic、中国台湾环球晶圆、韩国 LGSultron，五家企业的市场份额总和已超过 95%（数据来源：赛迪顾问）。

半导体设备方面，目前全球半导体设备市场的优势企业主要包括美国应用材料、荷兰阿斯麦、日本东电电子、美国泛林集团等。2021 年第二季度，美国应用材料收入为 50.71 亿美元，荷兰阿斯麦收入为 48.34 亿美元，日本东电电子为 40.56 亿美元，美国泛林集团收入为 41.45 亿美元，四家合占全球前十大半导体设备厂商总收入的 78.5%（数据来源：CINNO Research）。

IP 核设计方面，目前面向全球高端市场的优势企业主要包括 Arm、Synopsys、Cadence 等。国内 IP 核设计方面的优势企业是神州龙芯、芯原微电子、和芯微电子、苏州国芯、芯动科技等。目前 Arm 公司设计的 CPU 核凭借其在开发生态、功耗、性能等方面的优势在国内外车载人工智能计算芯片中的应用占主导地位。

EDA 工具方面，目前全球优势企业主要包括 Synopsys、Cadence 和 Siemens EDA 等。2020 年，Synopsys 中国市场市占率为 29.1%；Cadence 中国市场市占率为 32%；Siemens EDA 中国市场市占率为 16.6%（数据来源：赛迪智库）。国内 EDA 发展相对国外较为落后，EDA 软件企业数量较少，主要有华大九天、苏州芯禾科技、杭州广立微、概伦电子等。

芯片设计方面，英伟达、特斯拉、谷歌等国外企业纷纷斥巨资加速车载人工智能计算芯片的设计；国内车载人工智能计算芯片的设计

也展现出百花齐放的格局，华为、地平线、芯驰、黑芝麻等企业纷纷推出了面向不同等级自动驾驶要求的车载人工智能计算芯片。

晶圆制造方面，2021 年全球半导体 40nm 及以下制程市场，台积电市场占比为 28%，联电占比为 13%，中芯国际作为占比 11%，三星电子占比 10%（数据来源：Counterpoint Research）。国内半导体企业目前正致力于先进制程车规级产线研发，且在产线成熟度方面有带进一步完善。

芯片封测方面，2021 年第二季度，全球芯片封测市场中，日月光占有全球 23.7% 的市场份额，安靠占有全球 17.9% 的市场份额，长电科技占有全球 14.0% 的市场份额，通富微电占有全球 7.5% 的市场份额，天水华天占有全球 5.9% 的市场份额（数据来源：TrendForce）。国内芯片封测方面的优势企业是长电科技、通富微电、天水华天和晶方科技等。

芯片应用方面，多家供应商推出了面向自动驾驶的计算平台，如国外英伟达推出的 Xavier 平台和 Orin 平台，国内华为推出的 MDC 平台，地平线推出的 Matrix 平台，黑芝麻推出的 FAD 平台，芯驰推出的 UniDrive 平台等。多家整车厂通过与供应商合作，在供应商提供的计算平台基础上开展自动驾驶车辆研发工作。

（二）车载人工智能计算芯片典型产品及应用

国内外企业积极布局车载人工智能计算芯片研发，本部分重点对相关企业的典型产品及应用展开介绍。

1) 华为

华为推出的车载人工智能计算芯片产品主要为昇腾系列，包括昇腾 310、昇腾 610 两款产品。昇腾 610 芯片不仅具有强大的 AI 计算能力，其 CPU、CV (Computer Vision, 计算机视觉) 处理能力及内存带宽同样处于业界领先，保证了 SoC 对自动驾驶应用负载的 Scalar (标量)、Vector (向量)、Cube (矩阵乘) 算法具备很高的异构处理效率及最佳的算法适配性。结合芯片的车规加强 IP 及功能安全设计，昇腾 610 芯片的功能安全等级可以达到 ASIL-D。昇腾自动驾驶软件栈可提供确定性调度及多计算域协同特性，为感知、融合、定位、规控提供端到端低时延、高确定性的执行环境；结合算子开发、网络调优与图优化开发工具，以及自动驾驶 AI/CV 算子加速库支持，为合作伙伴提供易用、好用的开发生态，并降低合作伙伴开发与集成的成本。

2019 年，华为基于昇腾 310 芯片推出第一代 MDC 300 域控制器；2020 年至 2021 年，基于昇腾 610 芯片推出 MDC 310, MDC 610, MDC 610 Pro, MDC 810 多系列域控制器，AI 算力可实现从 100TOPS、200TOPS 到 400TOPS 的全覆盖，满足 L2+/L3/L4/RoboTaxi 等智能驾驶场景需求。

2) 地平线

地平线推出的车载人工智能计算芯片产品主要为征程系列。分别于 2019 年、2020 年和 2021 年发布了征程 2、征程 3、征程 5 三款产品。三款产品在算力、制程、传感器接入、功耗比等方面逐渐提升，满足面向不同等级自动驾驶的要求。基于征程 2 推出了自动驾驶计算

平台 Matrix2，具备强大的感知计算能力，为高级别自动驾驶提供了稳定可靠的高性能感知系统。基于征程 5 推出了面向下一代软件定义汽车的全场景整车智能中央计算平台 Matrix5，满足辅助驾驶、自动驾驶、智能座舱等全车多场景运算需求。

在应用方面，地平线与国内车厂开展了广泛合作。长安汽车 UNI-T 搭载了基于地平线征程 2 芯片开发的智能驾驶舱 NPU（Neural Network Processing Unit，特指一类专用化设计的神经网络计算单元）计算平台，并于 2020 年 6 月在国内实现量产上市。2021 款理想 ONE 使用了两颗地平线征程 3 芯片，实现了自动辅助导航驾驶等功能。

征程系列产品的详细性能参数如下表 2 所示：

表 2 征程系列产品的性能参数表

指标		征程 2	征程 3	征程 5
关键指标	AI 算力	4TOPS	5TOPS	128TOPS
	SOC 功耗	2W	2.5W	30W
	能耗比	2TOPS/W		4.3TOPS/W
	时延	< 100ms	/	60ms
	DDR 内存	提供接口，支持扩展 32-bit LPDDR4 @ 2667 MHz (MAX 2GB)	/	
	功能安全	/		ASIL-B (D)
	可靠性	AECQ-100 Grade2		
支持开发框架	适配主流的训练框架 Caffe、MXNet、TensorFlow 和 PyTorch，支持 ONNX			
其他产品特性	CPU	双核 Arm Cortex-A53	四核 Arm Cortex-A53	八核 Arm Cortex-A55 CPU
	制造工艺	28nm	16nm	
	封装	FCBGA388 17 × 17	FCBGA 15 × 15	HFCBGA 27 × 27
	以太网	1x GigE Ethernet	/	2*GigE with TSN

3) 芯驰科技

芯驰科技推出的车载人工智能计算芯片产品主要为 V9 系列。从 2019 年到 2020 年，芯驰科技先后发布了 V9 系列的三款芯片：V9L、V9F 和 V9T。V9L 和 V9F 用于满足市面上主流 ADAS 应用场景，V9T 具有更好的扩展性，支持 L3+更高级别的自动驾驶功能。基于 V9 系列芯片芯驰开发了自动驾驶平台 UniDrive，并提供可供评估的传感器及实车组合套件、系统框架、参考算法、工具链等支持。

2022 年，芯驰科技计划发布算力在 10TOPS-200TOPS 之间的自动驾驶芯片 V9P/U，该产品拥有更高算力集成，可支持 L3 级别的自动驾驶。此外，未来芯驰科技计划推出具有更高算力的 V9S 自动驾驶芯片，该芯片面向中央计算平台架构研发，算力高达 500TOPS-1000TOPS，可支持 L4/L5 级别的自动驾驶。

V9 系列产品的详细性能参数如下表 3 所示：

表 3 V9 系列产品的性能参数表

指标		V9L	V9F	V9T
关键指标	AI 算力	0.4TOPS		1TOPS
	SoC 功耗	3W	4W	10W
	数据精度	INT8		
	SRAM	1MB		
	DDR 内存	16-bit LPDDR4/4x2GB		Dual 32-bit LPDDR4/4x16GB
	可靠性	AECQ-100 Grade2		
其他产品特性	GPU 算力	/	40GFLOPS	80GFLOPS
	CPU	单核 Arm Cortex-A55		八核 Arm Cortex-A55
	制造工艺	16nm		
	封装	FCBGA 19 × 19		FCBGA 31 × 31
	工作温度	-40-125℃		
	SDIO/eMMC	1x SDID3.0+1xeMMC5.1		2x SDID3.0+2xeMMC5.1
	CAN-FD	4x		8x
	USB3.0	/		2x
	以太网	1*GigE with TSN		2*GigE with TSN
	GPIO	81		120

4) 黑芝麻智能科技

黑芝麻智能科技推出的车载人工智能计算芯片产品主要为华山系列。分别于2019年、2020年、2021年推出了A500芯片、A1000L和A1000芯片、A1000 Pro芯片。基于A1000芯片，黑芝麻提供了四种智能驾驶解决方案：单颗A1000L芯片适用于ADAS辅助驾驶，单颗A1000芯片适用于L2+自动驾驶，双A1000芯片互联适用于L3等级自动驾驶（算力可达到140TOPS），四颗A1000芯片适用于L4及以上的自动驾驶需求。此外还可根据客户需求提供定制化服务。

基于A1000芯片黑芝麻推出了FAD全自动驾驶计算平台，其中双芯片FAD平台可达到超过200TOPS的算力，满足L2+与L3级别智能驾驶场景的需求；四芯片FAD平台方案，主要面向L4级的自动驾驶场景，如码头、矿山的无人驾驶。

华山二号系列产品的详细性能参数如下表4所示：

表4 华山系列产品的性能参数表

指标		A1000L	A1000	A1000 Pro
关键指标	AI 算力	16TOPS	58TOPS	106TOPS
	SOC 功耗	10W	15W	25W
	能耗比	1.6TOPS/W	3.8TOPS/W	4.2TOPS/W
	时延	/		60ms
	数据精度	INT8		
	DDR 内存	2x 32b LPDDR4@3200,	2x 32b LPDDR4@4266,	4x 32b LPDDR4@4266,

指标		A1000L	A1000	A1000 Pro
		25.6GB/s	37GB/s	74GB/s
	功能安全	ASIL-B(D)		
	可靠性	AEC-Q100 Grade 2		
	开发 SDK 和 NN 工具链	山海系统		
其他产品特性	GPU 算力支持	30 GFLOPS	36 GFLOPS	60 GFLOPS
	逻辑算力	24K DMIPIS	32k DMIPS	60k DMIPS
	CPU	六核 Arm Cortex-A55	八核 Arm Cortex-A55	十六核 Arm Cortex-A55
	制造工艺	16nm		
	封装	FCBGA 25x 25		FCBGA 47.5 x 47.5
	工作温度	-40°C ~ 105 °C		
	ISP	HDR(140dB), Low Light, LFM 2x ISP Engine@400MHz pipelines, 0.8Gpps	HDR(140dB), Low Light, LFM 3x ISP Engine@400MHz pipelines, 1.2Gpps	HDR(140dB), Low Light, LFM 5x ISP Engine@400MHz pipelines, 2.0Gpps
	PCIe3.0	2x PCIe Gen2 2L	1*PCIe Gen3 4L	2*PCIe Gen3 4L
	CAN-FD	2		4
	以太网	2x GigE Ethernet	2*GigE with TSN	4*GigE with TSN

五、发展建议

（一）完善产业布局 促进技术创新

布局车载人工智能计算芯片产业链，系统化促进车载人工智能计算芯片技术创新。围绕车载人工智能计算芯片产业链的内容，建议重点开展如下工作：一是鼓励国产车载人工智能计算芯片的规模化应用，促进国产芯片的升级与发展。二是加快产业链薄弱环节自主研发，如光刻胶、光刻机、IP核、EDA设计工具等，为国内车载人工智能计算芯片发展提供良好的支撑。三是凝聚产业力量，布局先进制程的车规级晶圆生产线研发，提升高性能车载人工智能计算芯片的制造水平，保障芯片供应链的稳定性。

（二）促进企业交流 推动产业发展

促进企业间的交流协作，推动车载人工智能计算芯片产业链的合作与发展。一方面，政府可以通过加强协同创新、开展合作研究项目等方式促进企业间的交流合作，以进一步明确行业痛点，推动建立良好的供需对接关系，增进行业内协作，并实现企业间的相互赋能。另一方面，促进产业链头部企业充分发挥其对产业高质量发展的核心引擎功能和辐射带动作用。支持头部企业联合上下游共同开展车载人工智能计算芯片的研发和产业化，提升头部企业产业链引领带动作用。

（三）完善配套保障 提升技术实力

完善产业链配套保障，提升车载人工智能计算芯片行业的技术水平。一方面，增加车载人工智能计算芯片人才培养。建议通过区域导

向措施促进人才的合理分配，通过高校人才建设促进基础人才培养。另一方面，完善车载人工智能计算芯片测试标准，支持专业检测机构的发展。鼓励行业内优势企业联合推动车载人工智能计算芯片可靠性、功能安全、信息安全、基准测试、软件过程改进及能力评定等相关测试标准的制定。加强车载人工智能计算芯片检测能力，注重车载人工智能计算芯片测评认证能力建设，发挥以测促研，以测促用的良好作用。

ESTC 中国评测

附件：缩略语

缩略语	英文名称	中文名称
AI	Artificial Intelligence	人工智能
CPU	Central Processing Unit	中央处理器
CV	Computer Vision	计算机视觉
DDR	Double Data Rate Synchronous Dynamic Random Access Memory	双倍速率同步动态随机存取内存
DSP	Digital Signal Processing	数字信号处理
EDA	Electronic Design Automation	电子设计自动化
eMMC	Embedded Multi Media Card	内嵌式多媒体卡
GPU	Graphic Processing Unit	图形处理器
IP	Intellectual Property core	知识产权核
ISP	Image Signal Processor	图像处理器
LPDDR	Low Power Double Data Rate SDRAM	低功耗内存
MAC	Multiplier and Accumulation	乘累加器
NPU	Neural Network Processing Unit	特指一类专用化设计的神经网络计算单元
SDK	Software Development Kit	软件开发工具包
SoC	System on Chip	片上系统
UFS	Universal Flash Storage	通用闪存存储

注意事项：

- 未经中国软件评测中心同意，不得用于其它商业运作。
- 未经中国软件评测中心书面批准，不得部分复印报告。

-
- 单位地址：北京市海淀区紫竹院路 66 号
 - 邮政编码：100048
 - 通信地址：北京市海淀区紫竹院路 66 号赛迪大厦 12 层
 - 电话/传真：010-88559296/88559333
 - 电子信箱：xiameng@cstc.org.cn
 - 投诉电话：010-88559315