

CIC 灼识



全球端侧 AI 推理芯片 行业报告

© 2026 CIC 灼识版权所有。本文件包含高度机密信息，仅供我方客户专属使用。
未经 CIC 灼识书面许可，严禁以任何形式传阅、引用、复制或转载本文任何内容。

摘要

智能设备部署的快速扩张使得感知数据呈指数级增长。这一激增大幅提升了对前端处理和实时本地计算的需求。受时延与算力限制，传统云端架构愈发难以满足当前应用需求。因此，AI 驱动的设备必须在高精度感知与高效计算两方面同步提升。

在此背景下，端侧 AI 推理芯片的重要性愈发凸显。这类芯片将 AI 模型与智能设备感知技术相融合，构建起感知、计算、执行的实时闭环系统。其可利用文本、图像、视频、音频等实体数据，直接在设备上完成 AI 驱动的分析与决策，大幅降低对云端资源的依赖。

目录

1. 行业概览

1.1 端侧 AI 推理芯片行业定义

1.2 端侧 AI 推理芯片行业规模与出货量

2. 快速增长的产品细分领域：视觉端侧 AI 推理芯片

2.1 视觉端侧 AI 推理芯片的行业定义

2.2 视觉 AI 推理芯片的行业规模及出货量

2.3 视觉端侧 AI 推理芯片的核心驱动因素及发展趋势

1. 行业概览

1.1 端侧 AI 推理芯片行业定义

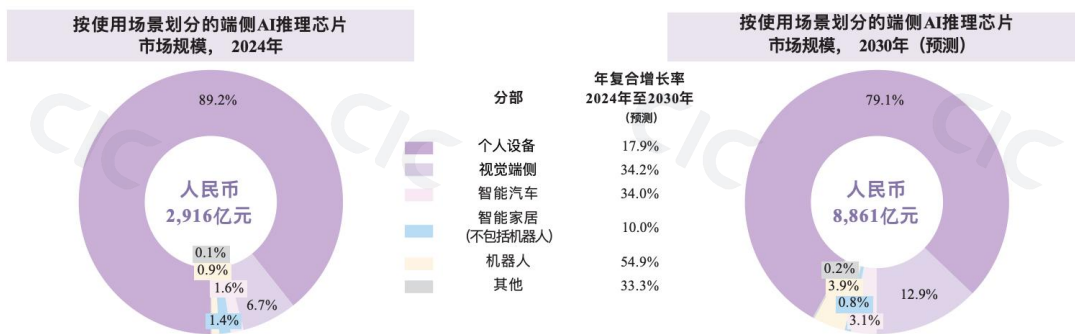
端侧 AI 推理芯片直接应用于终端用户设备，包括智能手机、智能汽车、智能家电等消费电子产品。

感知数据量的激增，推动了对端侧 AI 推理芯片的需求增长。通过将 AI 模型与智能设备感知技术相结合，这类芯片构建起感知、计算、执行一体化的实时闭环系统，可直接在设备端利用文本、图像、视频、音频等实体数据完成 AI 分析与决策，大幅降低对云端资源的依赖。

1.2 端侧 AI 推理芯片行业规模与出货量

端侧 AI 推理芯片行业应用广泛。全球各细分领域端侧 AI 推理芯片行业规模均预计实现快速增长，其中机器人、视觉端侧产品、智能汽车、个人设备及智能家居领域的年复合增长率分别为 54.9%、34.2%、34.0%、17.9%和 10.0%。

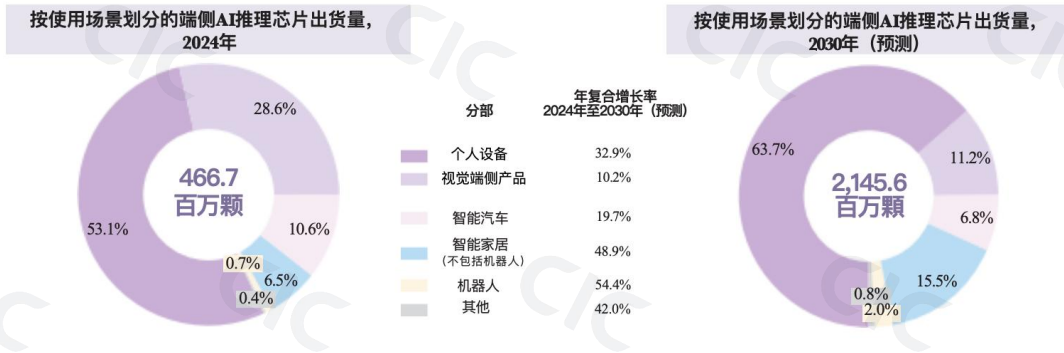
按使用场景划分的全球端侧AI推理芯片市场规模，2024年至2030年（预测）



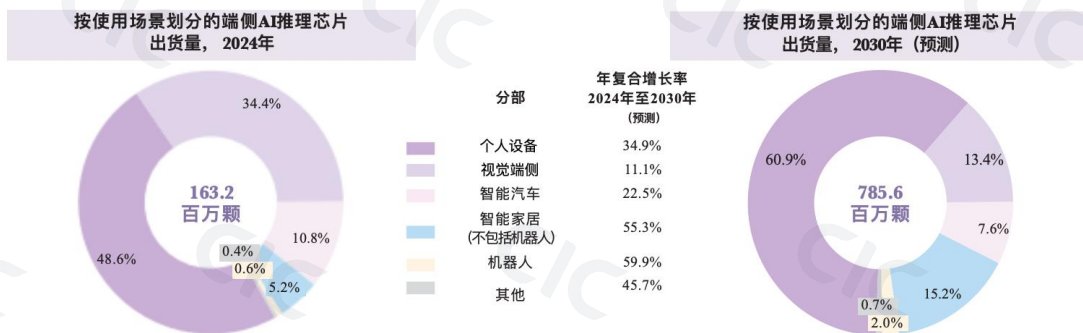
2024年至2030年，全球机器人、智能家居、个人设备、智能汽车及视觉端侧产品的端侧 AI 推理芯片出货量预计将分别实现 54.4%、48.9%、32.9%、19.7%

和 10.2% 的年复合增长率，而中国市场的对应细分领域预计将分别以 59.9%、55.3%、34.9%、22.5% 和 11.1% 的增速增长。

按使用场景划分的全球端侧 AI 推理芯片出货量，2024 年至 2030 年（预测）



按使用场景划分的中国端侧 AI 推理芯片出货量，2024 年至 2030 年（预测）



附注:

个人设备包括手机、可穿戴设备及 XR（指为沉浸式交互体验而设计的头戴式可穿戴设备，

涵盖虚拟现实(VR)及增强现实(AR)中的应用) 等消费电子产品。

视觉端侧产品包括工业、城市、家庭等不同场景中的视觉感知设备。

智能汽车指配备智能驾驶功能的车辆。

智能家居指电视、音箱及空调等家用产品。

机器人包括工业机器人、扫地机器人、送餐机器人等。

其他包括工业检测设备和智能工业控制系统。

资料来源：CIC 灼识、市场参与者访谈、行业出版物、政府统计资料、上市公司公开文件、

新闻

2. 快速增长的产品细分领域：视觉端侧 AI 推理芯片

2.1 视觉端侧 AI 推理芯片的行业定义

视觉端侧产品专为处理单一类型的视觉输入而设计，常见产品包括公共安防摄像头、行车记录仪、智能门锁、机器视觉检测设备等。视觉端侧 AI 推理芯片通常按性能分为低端和中高端两类，其中低端芯片的算力通常低于 1 TOPS（万亿次运算每秒），中高端芯片则超过这一阈值。

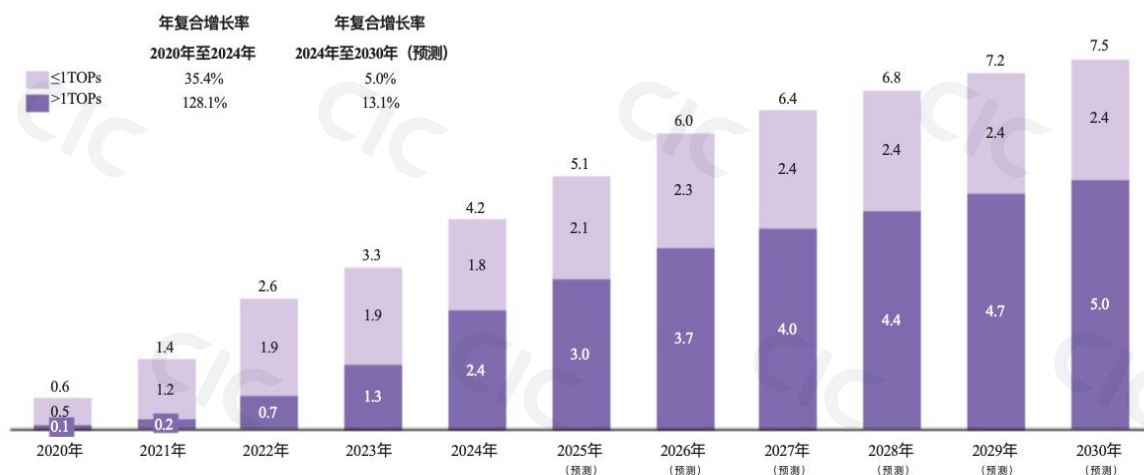
2.2 视觉 AI 推理芯片的行业规模及出货量

随着视觉应用对高分辨率、智能化及低延迟处理的需求持续攀升，市场正快速向中高端芯片倾斜，这类芯片已成为增速最快的产品细分领域。

全球中高端视觉端侧 AI 推理芯片市场规模预计将从 2024 年的 24 亿元人民币增长至 2030 年的 50 亿元人民币，年复合增长率为 13.1%。

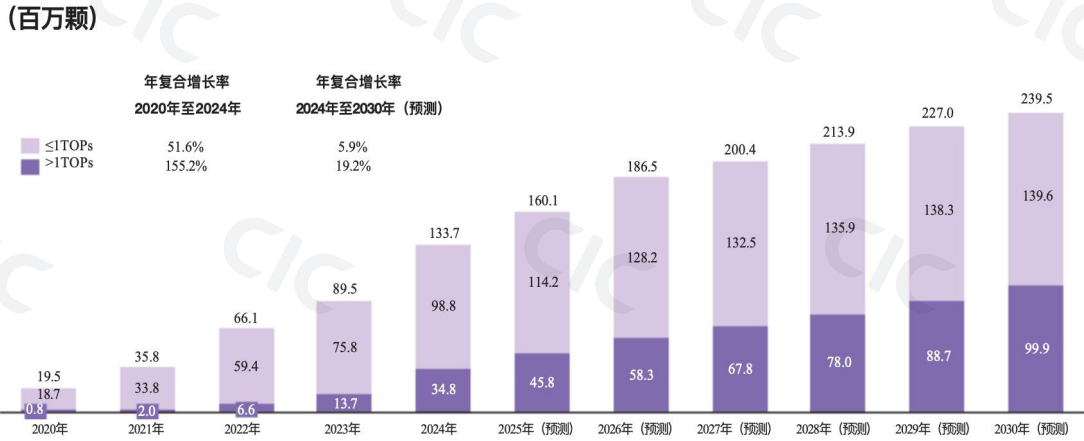
按算力 TOPS 划分的全球视觉端侧 AI 推理芯片市场规模，2020 年至 2030 年（预测）

（人民币十亿元）



相应地，2024 年全球中高端芯片出货量达 3480 万颗，约占总出货量的 26.0%；
 预计到 2030 年，这一数字将增至 9990 万颗，市场份额也将扩大至 41.3%。

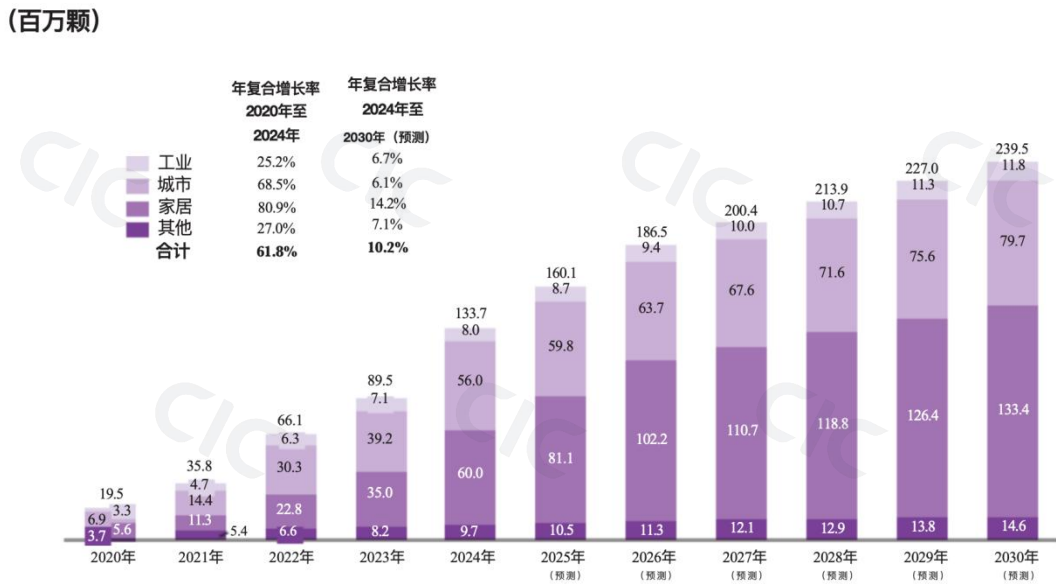
按算力TOPS划分的全球视觉端侧AI推理芯片出货量，2020年至2030年（预测）



按应用场景划分的市场细分提供了另一个关键的测算维度。

全球视觉端侧 AI 推理芯片出货量预计将从 2024 年的 1.337 亿颗大幅增长至 2030 年的 2.395 亿颗，年复合增长率为 10.2%。

按应用场景划分的全球视觉端侧AI推理芯片出货量，2020年至2030年（预测）



预计到 2030 年，中国视觉端侧 AI 推理芯片出货量将达到 1.054 亿颗。这类芯片已在家庭、城市、工业及其他应用领域实现广泛应用。

按应用场景划分的中国视觉端侧 AI 推理芯片出货量，2020 年至 2030 年（预测）
（百万颗）



资料来源：CIC 灼识

2.3 视觉端侧 AI 推理芯片的核心驱动因素及发展趋势

智能转型：传统图像与视频系统已无法满足现代家庭、城市及工业场景对实时性、精准度和智能化的要求。为实现实时人脸识别、行为分析、异常检测等功能，高性能且高能效的 AI 芯片必不可少，这推动了视觉端侧计算芯片的普及应用。

数据处理需求日益增加：设备对高分辨率、多模态感知的需求日益增长，进而带动了集成神经网络处理器 (NPU)、图像信号处理器 (ISP) 及数字信号处理器 (DSP) 的异构芯片架构需求，这类架构为图像增强、语义理解、语音交互等复杂 AI 任务提供算力支撑。

政策支持：中国《关于推进移动物联网“万物互联”发展的通知》明确鼓励芯片及模组技术创新与产业化，《新基建规划》、《“十四五”智能制造发展规划》等其

他举措更强化了支持端侧 AI 推理芯片大规模部署的政策体系。

关于 CIC 灼识

CIC 灼识咨询（简称“CIC 灼识”）是一家专业咨询机构，围绕投融资全生命周期，提供定制化一站式全流程服务。公司在全球各大市场主导打造多个行业首创的标杆 IPO 项目，业绩稳居世界前列。同时在全类专业细分赛道中，拥有无可匹敌的资源触达能力与深度全覆盖研究实力。

CIC 灼识助力企业优化具备规模化潜力的商业模式，塑造极具说服力的资本市场价值叙事，畅通对接全球资本市场的路径。同时作为投资机构信赖的尽职调查合作伙伴，输出精细化行业研判视角，并直通各领域权威专家资源，助力客户精准锁定高价值机遇、有效规避核心重大风险。

CIC 灼识团队深耕金融服务、人工智能、大数据、互联网、高新技术、医疗健康、教育、文娱、消费品、交通运输物流、能源电力、环境与建筑科技、化工、工业制造、农业等多元领域，实时掌握深度一线市场动态，能够为客户独家输出贴合细分行业、可落地执行的专业洞察结论。

CIC 灼识报告 & 行业概览

CIC 灼识搭建了一套严谨的多元化研究框架，整合一手调研与二手资料，为所有分析研判筑牢根基。一手调研主要深度对接行业权威专家与一线从业者，重点深耕供应链金融领域。二手研究则汇总梳理各大权威机构的公开数据，数据来源包括：中华人民共和国国家统计局、国家金融监督管理总局（SAFR，原中国

银行业监督管理委员会)、中国证券监督管理委员会, 以及上市公司公开披露文件。

我们运用自研专属数据分析体系对收集到的信息进行加工处理, 并通过多渠道研究数据交叉比对验证研究结论, 确保分析过程严谨、结果真实可靠。

本报告中展示的所有统计数据均可核验追溯, 全部基于报告出具当日可获取的有效信息整理而成。

本篇内容摘编提炼自 CIC 灼识深度行业研究报告精华, 聚焦各细分赛道的供需走势、核心增长驱动因素、研发创新趋势与行业未来发展前景等核心内容, 同时融合专家访谈、市场实地调研、行业数据解析等多维度专业研判成果。

免责声明

本报告由 CIC 灼识依据截至出具当日可获取的信息编制。本报告仅作参考之用, 内容不具备最终定论效力, 亦不得被解读为确定性结论。

本报告所载全部内容, 均不构成且不得视作投资建议、投资推荐, 亦非开展任何投资活动的要约、招揽或劝导。

凡因使用或依赖本报告所载信息, 直接或间接引发任何损失、损害及各类索赔诉求的, CIC 灼识特此明确免除一切相关责任。



CIC 灼识 | 全球端侧 AI 推理芯片行业报告

联系我们

如需了解本报告更多详情,或咨询 CIC 灼识的各项专业服务,欢迎访问 [CIC 灼识官方网站](#),亦可发送邮件至: marketing@cninsights.com。