下载日志已记录，仅供内部参考

计算机

发布时间：2026-07-09

证券研究报告 / 行业深度报告

大模型推动视觉 AI 向产业渗透

优于大势

首次覆盖

报告摘要：

1、视觉 AI 应用场景持续拓展，工业生产成为核心落地方向。基于 AI 算法、多模态模型与行业软件平台形成的智能视觉解决方案，正在从传统安防监控、城市治理等场景，逐步向工业制造、能源巡检、仓储物流、自动驾驶和机器人等生产领域渗透。工业质检与缺陷检测是当前商业化最成熟的应用方向，安全生产监控、仓储物流和智慧工厂监控进一步推动视觉 AI 从单点识别走向生产过程可视化。随着制造业智能化改造推进，视觉 AI 逐步参与质量控制、流程优化和生产管理，成为工业智能化升级中的重要基础能力。

2、全球视觉 AI 市场空间持续扩张，国内企业在场景落地方面具备优势。全球视觉 AI 行业仍处于快速扩张阶段，全球视频分析市场规模预计将由 2026 年的 146.5 亿美元增长至 2031 年的 413.9 亿美元，CAGR 约 23.1%。2025 年中国视觉 AI 市场规模预计同比增长 11%，在安防、城市治理和工业视觉等领域具备较强商业化基础。未来，具备硬件入口、算法能力、行业数据积累和项目交付经验的企业，有望在视觉 AI 产业化过程中占据更有利位置。

3、视觉 AI 未来将从识别工具走向生产要素，商业模式由硬件销售向服务化延伸。随着深度学习、多模态大模型和 VLM 视觉语言模型持续发展，视觉 AI 的技术路径正在从“看见”走向“理解”，并进一步向辅助决策升级。未来，视觉 AI 有望从单点质检、安全监控等环节，进一步嵌入生产流程、设备运维、物流调度和机器人执行等场景，成为制造业智能化升级的重要基础能力。

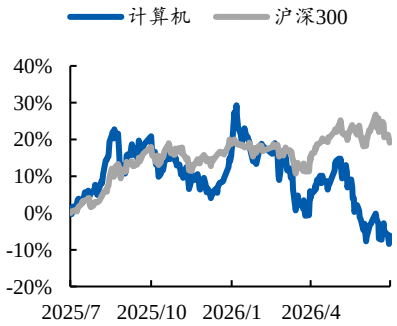
推荐板块：视觉 AI 硬件应用领域：硬件是视觉 AI 在向产业渗透过程中的必要条件，推荐标的：海康威视、大华股份。

风险提示：视觉 AI 技术应用程度低于预期；基础大模型的迭代对原有视觉 AI 大模型的冲击导致的估值偏差；业绩预测和估值不达预期。

重点公司主要财务数据

重点公司	现价	EPS			PE			评级
		2026	2027	2028	2026	2027	2028	
大华股份	16.96	1.22	1.39	1.65	13.90	12.20	10.28	买入
海康威视	34.35	1.86	2.21	2.62	18.47	15.54	13.11	买入

历史收益率曲线



涨跌幅（%）	1M	3M	12M
绝对收益	-2%	-11%	-6%
相对收益	-3%	-15%	-25%

行业数据

成分股数量（只）	334
总市值（亿）	41443
流通市值（亿）	23140
市盈率（倍）	151.19
市净率（倍）	3.70
成分股总营收（亿）	13355
成分股总净利润（亿）	277
成分股资产负债率（%）	45.46

相关报告

《联科科技(001207)：高压电缆用纳米炭黑进入放量期》

--20250728

《大华股份：海外收入持续扩张、视觉 AI 逐步扩散》

--20260629

证券分析师：赵丽明

执业证书编号：S0550521100004

010-63210892 zhaolm@nesc.cn

下载日志已记录，仅供内部参考

目 录

1. 视觉 AI 发展历程 4
1.1. AI 的概念界定与发展 4
1.2. 视觉 AI 的概念与核心技术 4
1.3. 视觉 AI 技术助力行业发展 5
2. 视觉 AI 行业分析 6
2.1. 行业发展与市场规模 6
2.2. 行业竞争格局 10
2.3. 视觉 AI 产业链结构 12
2.4. 硬件与算力：视觉 AI 的基础支撑 13
2.4.1. 感知采集硬件：视觉 AI 的数据入口 13
2.4.2. 云端算力硬件：GPU 与 AI 芯片支撑大模型训练 15
2.4.3. 端侧 AI 芯片与边缘计算硬件：推动实时化、本地化部署 17
2.5. 软件层与算法平台发展现状 18
2.5.1. 算法模型：从识别检测走向多模态理解 18
2.5.2. AI 视频分析软件：软件层商业化最成熟的方向之一 20
2.5.3. 软件私有化：中国市场的重要竞争领域 21
3. 行业应用与商业化现状 22
3.1. 质检与缺陷检测：当前最成熟的工业视觉 AI 应用 22
3.2. 安全生产监控：从人员监管到现场风险识别 22
3.3. 仓储物流：参与物品识别、分拣和调度 23
3.4. 智慧工厂监控：从单点检测走向生产过程可视化 25
3.5. 自动驾驶：驱动交通出行的智能化变革 26
3.6. 机器人：赋能机器感知与自主决策 28
4. 视觉 AI 潜在空间分析 30
4.1. 汽车：视觉 AI 高强度应用的代表性生产场景 30
4.2. 全球固定资本形成提供较大的潜在空间 31
5. 视觉 AI 未来趋势：从识别工具走向生产要素 31
5.1. 技术演进：从“看见”到“理解”再到“决策” 32
5.2. 场景深化：从单点替代到全流程嵌入 33
5.3. 产业渗透率提升：从头部制造向更广泛产业扩散 34
5.4. 商业模式演进：从卖硬件到卖服务 34
6. 典型公司分析 36
6.1. 海康威视：行业步入上行周期、公司业绩加速兑现 36
6.2. 大华股份：海外收入持续扩张、视觉 AI 逐步扩散 36
7. 风险提示 37

图表目录

图 1：我国人工智能产业模态结构占比 4
图 2：全球视频分析市场规模 8
图 3：国内视觉 AI 市场规模 9
图 4：中国视觉 AI 应用领域占比 9
图 5：国内视觉 AI 竞争格局 11

下载日志已记录，仅供内部参考

图 6：视觉 AI 产业链.....13
图 7：视觉 AI 软件私有化市场份额.....21
图 8：全球固定资本形成总额.....31

下载日志已记录，仅供内部参考

1. 视觉 AI 发展历程

1.1. AI 的概念界定与发展

人工智能（Artificial Intelligence，AI）是指利用计算机技术模拟人类智能活动的一类技术体系，核心目标是机器具备感知、学习、推理和决策能力。AI 涵盖机器学习、自然语言处理、计算机视觉等多个领域。根据《人工智能标准化白皮书（2018 版）》，AI 被定义为利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

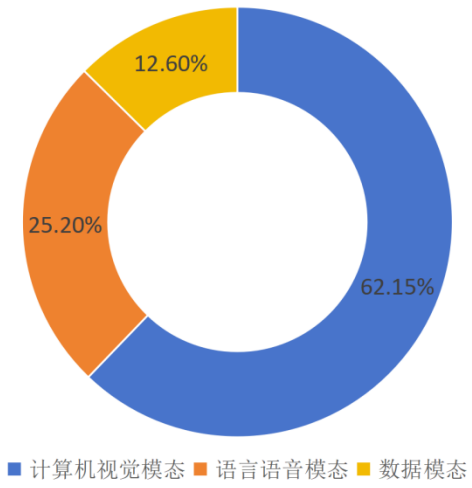
AI 产业链上游为硬件设备和数据设备，其中包括芯片、传感器、大数据、云计算服务等，为 AI 提供数据服务和算力支持；中游是 AI 产业的技术核心，包括通用技术、算法模型、开发平台三个方面，涵盖了机器感知、自然语言理解、机器学习、图谱、基础开源框架等核心技术；下游是 AI 的应用产品和场景，涉及交通、医疗、安防、金融、家居、制造等多个领域。

随着大数据、云计算及深度学习技术的发展，AI 逐渐从理论研究阶段进入产业化应用阶段，并广泛应用于金融、医疗、安防、交通等领域。初步测算 2025-2030 年期间，我国 AI 行业市场规模将进一步扩大，实现《新一代人工智能发展规划》中 2030 年 AI 产业规模超过一万亿元。

1.2. 视觉 AI 的概念与核心技术

视觉 AI，也被称为“计算机视觉”，是 AI 领域的一个重要分支。中商产业研究院发布的《2025-2030 全球与中国高效率人工智能解决方案市场现状及未来发展趋势》显示在我国人工智能产业中，计算机视觉模态（图像识别、图像生成）占比 62.1%，语言语音模态（对话式 AI、智能语音、NLP 文本、知识图谱）占比 25.2%，数据模态（机器学习产品）占比 12.6%。

图 1：我国人工智能产业模态结构占比



数据来源：中商产业研究院、东北证券

视觉 AI 是指用摄影机和计算机代替人眼对目标进行识别、跟踪和检测等机器视觉处理过程，通过图形处理技术，使计算机处理生成更适合人眼观察或适用于仪器检测的图像。视觉 AI 为跨学科的科学领域，使得计算机可分析数字图像或视频，以便提取数据、执行分析及自动化若干任务。视觉 AI 研究相关的理论和技术，试图建立能够从图像或者多维数据中获得可以用于决策的信息的 AI 系统。因为感知可以看作是从感官信号中提取信息，因此计算机视觉也可以看作是研究如何使人工系统从图像或多维数据中感知的科学。当前视觉 AI 正向多模态与 VLM 方向发展。

下载日志已记录，仅供内部参考

AI 行业飞速发展，中国视觉 AI 进入迅猛发展阶段，视觉 AI 的市场规模是 AI 行业市场规模占比最高的细分领域。近年来随着 AI 技术加速变革计算机视觉市场应用，计算机视觉赋能城市安防、交通、医疗、金融、零售等领域，计算机视觉行业下游商用价值不断拓宽和深化。未来，随着网络基础设施逐步成熟，站在自动驾驶和工业互联网高速发展的风口上，汽车领域和消费端应用将是市场增长最快的领域。随着算法和精度的提升，计算机视觉与行业场景融合将更加深入。

1.3. 视觉 AI 技术助力行业发展

视觉 AI 萌芽阶段

与其他 AI 领域一样，计算机视觉的早期发展始于基础研究和理论工作。视觉 AI 的起源可以追溯到 20 世纪 50 年代。1959 年，神经生理学家 David Hubel 和 Torsten Wiesel 通过猫的视觉实验发现，视觉初级皮层神经元对移动边缘敏感，并提出了视功能柱结构，为后来的计算机视觉研究奠定了神经科学基础。同年，Russell 等人研制了第一台数字图像扫描仪，使得图像可以被计算机处理。Lawrence G. Roberts 在 20 世纪 60 年代初的论文“三维固体的机器感知”中记录的关于 3D 目标识别的开创性工作。

早期的计算机视觉研究侧重于边缘检测和特征提取等图像处理技术。20 世纪 60 年代末开发的索贝尔算子等算法，是最早通过计算图像强度梯度来检测边缘的算法之一。诸如 Sobel 和 Canny 边缘检测器之类的技术在识别图像中的边界方面发挥了关键作用，这对于识别对象和理解场景至关重要。

在 20 世纪 70 年代，模式识别成为计算机视觉的一个关键领域。研究人员开发了识别图像中的形状、纹理和对象的方法，这为更复杂的视觉任务铺平了道路。早期的一种模式识别方法涉及模板匹配，即将图像与一组模板进行比较，以找到最佳匹配。这种方法受到其对尺度、旋转和噪声变化的敏感性的限制。早期的计算机视觉系统受到当时有限的计算能力的制约。20 世纪 60 年代和 70 年代的计算机体积庞大、价格昂贵，且处理能力有限。

深度学习推动行业突破

1970-80 年代，计算机视觉理论体系逐渐形成。1980 年代，卷积神经网络（CNN）开始出现，标志着视觉 AI 进入深度学习时代。CNN 能够自动提取图像特征，显著提升了图像分类和识别的准确性。同时，多层感知机和反向传播算法的发展，使得神经网络训练更加高效，为现代视觉模型奠定了算法基础。CNN 的运作主要需要卷积层、激活函数、池化层和全连接层。卷积层是一种深度学习模型，主要是通过自动学习分层模式来处理结构化的网格状数据，如图像或序列。这些过滤器通过在图像上滑动并计算点积来探测边缘、纹理和颜色等各种特征。每个过滤器都会激活图像中的特定模式，使模型能够学习分层特征。卷积之后，使用修正线性单元（如 ReLU）的激活函数，ReLU 是深度学习中一种流行的激活函数，如果输入为正，则直接输出输入，否则输出零，从而帮助神经网络有效地学习数据中的非线性关系。这有助于网络学习复杂的模式。池化层提供了一种降采样操作，可降低特征图的维度，从而有助于提取最相关的特征，同时降低计算成本和过拟合。CNN 的最后一层是全连接层，它解释卷积层和池化层提取的特征以进行预测。这些层与传统神经网络中的层相似。

1989 年的 LeNet 是最早的 CNN 架构之一，主要用于识别手写支票中的数字。它的成功为更复杂的 CNN 奠定了基础，证明了深度学习在图像处理中的潜力。1990 年代，特征提取与模式识别成为研究重点，SIFT、Harris 角点等算法相继出现。GPU 的普及也为大规模视觉计算提供了硬件支持。

2000 年代，高质量标注数据集（如 Pascal VOC）和特征工程（如 HOG、SPM）推动了目标检测与分类的发展。2006 年，Hinton 提出深度信念网络（DBN），开启了深度学习在视觉领域的探索。

下载日志已记录，仅供内部参考

进入 2010 年代后，深度学习迎来突破，视觉 AI 开始处理更复杂的任务，如实时目标检测、面部识别和情绪分析。代表性模型包括 Faster R-CNN、Mask R-CNN 以及 YOLO 系列（YOLOv1-YOLOv8），这些模型在精度和速度上都有显著提升，其中 YOLO 将物体检测作为一个单一的回归问题，在一次评估中直接预测完整图像中的边界框和类概率，这种方法能以前所未有的速度和准确性进行实时物体检测，适用于即时处理的应用领域，比如自动驾驶和监控。现代视觉 AI 的发展依赖于大规模数据集（如 ImageNet）、强大的计算能力以及深度学习算法的优化。2012 年，AlexNet 模型在 ImageNet 竞赛中的表现明显优于现有模型，展示深度学习的先进性。它利用 ReLU 激活、丢弃和数据扩增，为图像分类设定新的基准。2014 年，通过使用较小的卷积滤波器（3x3），VGGNet 在图像分类任务表现十分优异，也加强了网络深度在实现更高准确性方面的重要性。

多模态大模型驱动的新发展阶段

2020 年以来，多模态大模型与视觉语言模型（Vision-Language Model, VLM）快速发展，视觉 AI 进一步从图像识别向场景理解升级，行业应用场景持续拓展。Transformer 架构（如 Vision Transformer）和多模态学习推动了计算机视觉在自动驾驶、医疗影像、智能监控等领域的广泛应用。多模态视觉-语言模型是一种能够同时理解图像或视频与文本，并建立两者的关联关系。它突破了传统单一模态模型的局限，实现跨模态的联合推理、生成与分析能力。

2025 年是 VLM 研究的一个重要年份，涌现出多个新模型和技术，显著提升了模型性能和应用范围。2025 年 1 月 28 日推出 QWen2.5-VL，核心能力为视觉问答、图像描述，模型对外开源；2025 年 3 月 26 日上线 Qwen2.5-Omni，具备全方位多模态理解与生成能力；2025 年 4 月 4 日发布 LLaMA4-Scout/Maverick，主打视觉语言推理与内容生成，依托 LLaMA4 架构搭建；2025 年 4 月 14 日问世 InternVL-3，可实现图像描述、视觉问答以及多模态生成相关功能；2025 年 5 月 14 日推出 BLIP3-o，融合图像理解与生成能力，采用扩散变换器结构，代码开源；2025 年 6 月 23 日发布 OmniGen2，支持文生图、图像编辑、上下文生成。

相比于传统视觉 AI，VLM 能够同时处理视觉和语言信息，适用于需要跨模态理解的任务，如图像描述、视觉问答和多模态推理。VLM 的发展推动视觉 AI 从“感知智能”向“认知智能”升级，扩大了大模型向多领域深度融合，VLM 大模型已应用于医疗、安防、教育和机器人等许多领域，如在自动驾驶领域，VLM 能够同时理解道路环境、交通标识及语义信息；在智慧安防领域，系统可进一步分析异常行为与事件逻辑；在工业场景中，视觉 AI 开始从单一缺陷检测向生产流程分析升级。VLM 还具备自学习能力，通过大规模预训练，VLM 可以在未见过的数据上表现出色，减少对标注数据的需求。

与此同时，VLM 的发展对算力与数据提出更高要求。大规模模型训练需要高性能 GPU、云计算平台及海量数据支持，推动 AI 芯片、数据中心和边缘计算等产业链快速发展。未来，随着多模态技术持续成熟，VLM 有望成为视觉 AI 行业的重要发展方向。

2. 视觉 AI 行业分析

2.1. 行业发展与市场规模

基于视频感知、AI 算法、多模态模型与行业软件平台形成的智能视觉解决方案发展起来的视觉 AI，应用场景已经由传统安防业务逐步拓展至工业制造、能源巡检、智慧医疗与物流运营等生产领域。相比于传统安防行业，视觉 AI 领域对于研发、数据与软件生态的投入要求显著提升。因此，持续的研发投入已成为行业龙头维持新市场竞争优势的重要前提。未来视觉 AI 相关投资将集中于多模态大模型、行业数据训练、边缘 AI 算力与行业软件平台建设方向。

2025 年 8 月，国务院印发《关于深入实施“人工智能+”行动的意见》，明确提出加快实施六大重点行动，推动人工智能与经济社会各行业各领域广泛深度融合，促进

下载日志已记录，仅供内部参考

生产力革命性跃迁和生产关系深层次变革，加快形成人机协同、跨界融合、共创分享的智能经济和智能社会新形态。其中“人工智能+产业发展”是与制造业和实体经济结合最紧密的方向。政策提出四个具体发展方向：第一，要培育智能原生新模式新业态，鼓励企业将人工智能融入战略规划、组织架构中，推动产业全要素发展，助力传统产业改造升级，开辟战略性新兴产业和未来产业发展新赛道，同时加快培育一批底层架构和运行逻辑基于人工智能的智能原生企业，催生智能原生新业态；第二，推动工业全要素智能化发展，并进一步强调推进工业全要素智能化发展，加快人工智能在设计、中试、生产等全环节落地应用，着力提升全员人工智能素养与技能，推动各行业形成更多可复用的专家知识，并且深化人工智能与工业互联网融合应用，增强工业系统的智能感知与决策执行能力；第三，加快农业数智化转型升级，支持人工智能在农业领域的智能应用，提高农业生产和加工工具的智能感知、决策、控制、作业等能力；第四，创新服务业发展新模式，加快服务业从数字赋能的互联网服务向智能驱动的新型服务方式演进，拓展经营范围，推动现代服务业向智向新发展，并探索无人服务与人工服务相结合的新模式，推动新一代智能终端、智能体等广泛应用。同时，文件在基础支撑能力部分提出，要提升模型基础能力、加强数据供给创新、强化智能算力统筹、优化应用发展环境，并加大人工智能领域金融和财政支持力度。

2026 年 1 月 28 日，国务院新闻办公室举行新闻发布会，介绍 2025 年国资央企高质量发展情况。国务院国资委表示，人工智能是引领新一轮科技革命和产业变革的战略性、基础性技术，国资委已连续两年在年初召开专题会议，推进中央企业“AI+”专项行动。目前，中央企业已围绕应用、算力、数据和模型等重点环节开展布局。在应用端，聚焦能源、制造、通信等重点行业，联合相关企业打造超过千个应用场景；在算力端，持续推进万卡集群、全国一体化算力网络和国产智算集群建设；在模型端，推动通用大模型向能源电力和工业制造等领域落地；在生态端，通过开放算力、模型和数据资源，降低人工智能应用门槛。

对于下一阶段工作，国资委提出进一步强化投资牵引、深化场景培育和优化数据供给。具体包括谋划中央企业人工智能“十五五”战略规划，加快信息通信网络、算力网络及国产智算集群建设；聚焦具身智能、能源电力等重点领域，探索组建“AI+”产业共同体，扩大重大场景、行业集成场景及高价值小切口场景开放；同时推动交通物流、智慧能源、绿色低碳等领域数据资源依法合规开发利用，为模型训练、场景应用和产业化落地提供支撑。

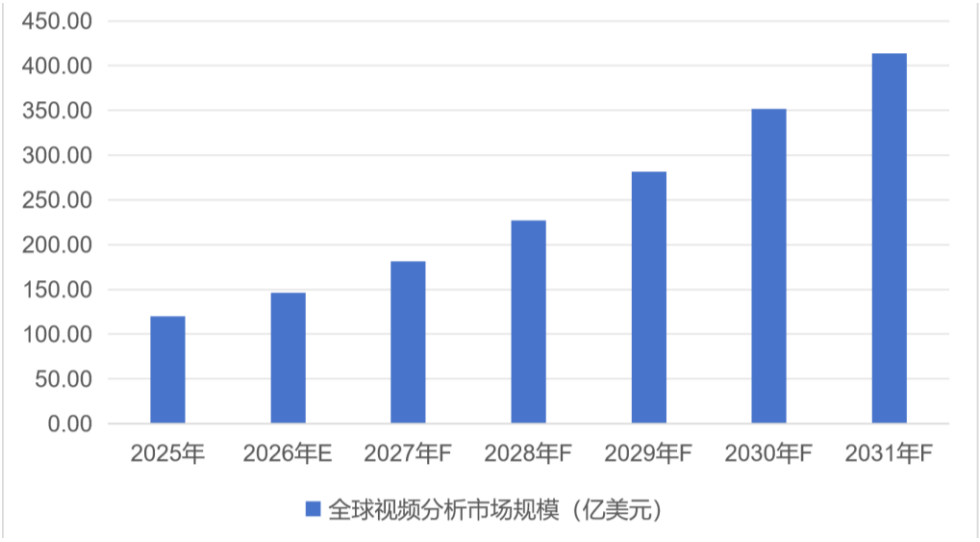
从短期业务影响看，政策在 2026 年一季度发布，政策部署有望在二季度逐步向项目对接和场景合作传导。央企扩大人工智能场景开放，强调与头部技术企业联合建设应用项目，有利于具备政企客户基础、行业解决方案能力和私有化交付经验的视觉 AI 厂商参与项目试点、联合研发及生态合作。视觉 AI 与能源巡检、工业质检、安全生产、交通物流和机器人感知等央企重点应用方向具有较高契合度，因此相关企业在二季度可能迎来更多客户交流、项目验证、战略合作及招投标机会。

从技术影响看，政策在短期内对视觉 AI 企业底层算法水平和核心技术性能的直接推动相对有限，但政策能够通过开放行业数据和算力资源，为企业提供模型训练和持续迭代的条件。中长期看，央企高价值场景的持续开放有助于视觉 AI 企业积累能源、制造和交通等垂直行业数据，提升解决方案的标准化和复制能力，推动视觉 AI 进一步走向规模化部署。

根据 Markets and Markets 的分析，全球视频分析市场规模预计将从 2026 年的 146.5 亿美元，以 23.1%的复合年增长率增长至 2031 年的 413.9 亿美元，显示视觉 AI 行业仍处于快速扩张阶段。

下载日志已记录，仅供内部参考

图 2：全球视频分析市场规模



数据来源：Markets and Markets、东北证券

智能城市支出的增加、边缘 AI 芯片价格下降以及更严格的隐私规则使推断从公共云转移，正在重塑采购标准。软件依然是收入的锚点，但随着集成复杂性和合规审计给内部 IT 团队带来压力，服务增长速度最快。混合型和边缘架构正在弥合云端的差距，因为流量管理、工业检测和医疗监控要求不到 100 毫秒的响应时间。厂商正通过联邦学习模块和符合 ISO/IEC 42001 的工作流程来平衡准确性与隐私要求。

平台许可、分析引擎和模型库构成了大多数部署的核心。视觉 AI 服务市场规模可能在预测期末与软件竞争，利润池将从硬件转移。硬件依然不可或缺，尤其是在专业热成像或多光谱相机价格高昂的情况下，但商品化抑制了增长。行业意识到技能短缺，企业正转向服务行业。与此同时，ISO/IEC 42001 作为责任保护的基础认证正逐渐获得关注。随着客户越来越偏好基于结果的合同，经常性收入模式正在获得动力，进一步巩固了服务在视觉 AI 市场中**的主导地位。**

北美 2025 年收入得益于 12 亿美元的联邦智慧城市资助和零售商试点项目，这些项目将计算机视觉与销售点数据结合起来。加拿大为市政部署提供了 6300 万美元的拨款。到 2025 年底，墨西哥扩展的 C5 中心同时处理了超过 5 万条信息。

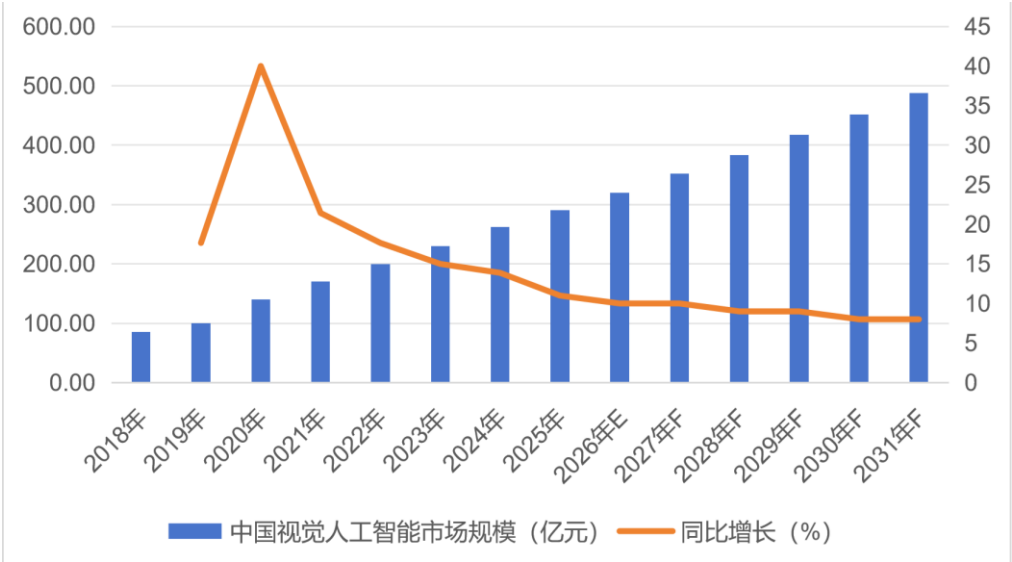
2025 年底，中国已经拥有与国家身份数据库实现实时警报的 AI 支持摄像头。印度智慧城市使命在 82 个城市部署了 AI 视频分析，强调边缘推断以实现本地化合规。日本致力于在主要铁路枢纽实施基于 AI 的人群监测，以防止跌倒事故。该地区的绿地项目和庞大的城市人口支撑着视觉 AI 市场的持续扩张。

前五大供应商约占 2025 年收入的 40-45%，显示出适度集中度。海康视界、大华和安克西斯利用摄像头占用面积捆绑分析，但随着芯片商品化，利润率压缩降低了进入门槛。纯游戏软件公司如 AxxonSoft 和 Vintra 通过零售、医疗和工业模式实现差异化，常与集成商合作以扩大影响力。

随着 AI 深度学习算法的快速成熟，国内的视觉 AI 行业也迅速发展，根据《2025-2031 年中国视觉人工智能行业市场现状分析及发展前景展望报告》，预计 2025 年中国视觉 AI 市场规模同比增长 11%。视觉 AI 的驱动因素呈现多元化、交叉化特征，技术、市场、政策、资本等力量共同塑造行业格局。未来，视觉 AI 将从感知世界向认知世界跃迁，技术融合、场景深化与伦理重构将重塑产业生态，2025 年中国视觉 AI 产值同比增长约 10.4%。

下载日志已记录，仅供内部参考

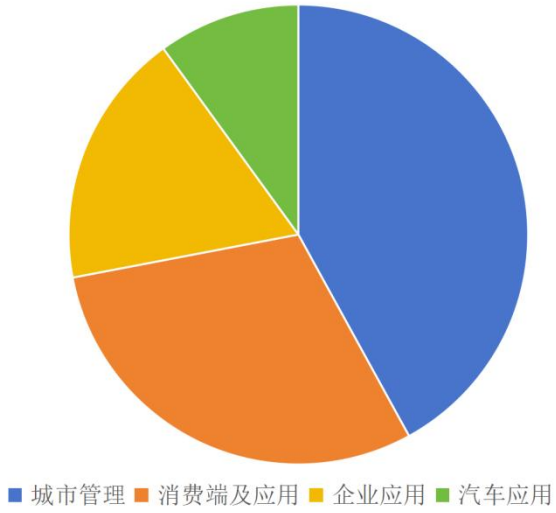
图 3：国内视觉 AI 市场规模



数据来源：智研咨询、东北证券

随着 AI 技术的日渐成熟和扩散，通用化、模块化的算法框架降低 AI 技术的使用门槛，更多行业和企业能够通过 AI 技术对现有的产业和业务进行优化，AI 技术和传统产业的融合程度逐步深入。在中国视觉 AI 应用领域中，以安防为主要场景的城市管理市场规模占比最大，达到 42.2%。

图 4：中国视觉 AI 应用领域占比



数据来源：共研产业咨询、东北证券

随着中国网络基础设施逐步成熟，视觉 AI 技术发展不断完善，中国视觉 AI 行业市场规模不断扩大。同时，国家重视关键核心技术的发展突破，不断出台措施强调 AI 领域的重要性，以及推动 AI 领域的发展。视觉 AI 行业是 AI 行业占比最大的细分领域，受 AI 行业发展的推动，发展十分迅速，孵化出了商汤、云从、依图、旷视四家独角兽企业。同时互联网巨头厂商纷纷布局计算机视觉业务，行业不断繁荣发展。未来，随着网络基础设施逐步成熟，站在自动驾驶和工业互联网高速发展的风口上，汽车领域和消费端应用将是市场增长最快的领域。随着算法和精度的提升，计算机视觉与行业场景融合将更加深入。

下载日志已记录，仅供内部参考

2.2. 行业竞争格局

全球视觉 AI 行业格局呈现海外企业主导底层算力与基础模型、国内企业强化场景落地与商业化应用的特征。其中，以英伟达、OpenAI 为代表的海外企业在 GPU 算力及大模型生态方面占据领先优势；而中国企业则在智慧安防、工业视觉和城市治理等应用场景中具备较强的商业化能力。

在全球视觉 AI 产业链中，算力层有海外芯片与计算机平台企业占据主导地位。算力层是全球视觉 AI 最核心的底层，NVIDIA 自成立以来构建以 CUDA 生态为核心，将底层 GPU+DPU+CPU 进行整合，提供数据中心的全领域加速计算方案，在持续推进的 AI 浪潮中，成为了软硬件一体化的云计算解决方案领导者，也成为全球视觉 AI 行业算力层的龙头企业。其技术优势和生态布局使它在中美算力竞争中成为核心变量，同时通过量子计算、物理 AI 等前沿领域布局，拓展长期增长曲线。Global Market Insights 提到，NVIDIA 在 2024 年拥有超过 9%的市场份额，IBM、Intel、Microsoft、NVIDIA、Technologies 等前五大企业合计占 26%左右。随着视频分析模型复杂程度的上升，算力层在产业链中的战略地位进一步上升。

在多模态技术发展下 VLM 正迅速成为连接图像与文本理解的核心技术，是视觉 AI 基础模型层的最重要部分。OpenAI 自成立以来专注于通用人工智能（AGI）研发，其核心产品覆盖自然语言处理、图像生成、语音识别、视频生成及编程辅助等多模态应用 。GPT-5.2 及其变体在文本、图像、代码处理能力上处于行业领先水平，性能基准测试显示其在专业任务上超越多数行业专家 。OpenAI 在技术和生态上具有明显优势，但面临 Meta、Anthropic 及中国初创公司 DeepSeek 等竞争压力。Meta 的 Llama 系列和 DeepSeek 的低成本模型对开发者生态和市场份额形成挑战，OpenAI 市场份额从 2023 年的 65%下降至 2025 年初的 48%。此外，人才流失和高额研发投入也是潜在风险 。随着 VLM 及多模态大模型快速发展，全球视觉 AI 行业对高性能算力需求提升，进一步强化了海外企业在 AI 基础设施领域的领先优势。

海外科技巨头正在构建 AI 基础设施生态，包括云计算、GPU 集群和 AI 训练平台。全球云计算市场由几大主要企业主导，AWS 以 30%的市场份额位居第一，其可靠性和安全性广受认可。Microsoft Azure 紧随其后，占 21%，得益于其与生产力工具的整合。Google Cloud 以 12%的份额位居第三，特别是在 AI 和数据分析领域表现强劲。此外，Alibaba Cloud 在 4%的市场份额中表现出色，尤其在中国市场，而 Oracle、SalesForce、IBM Cloud 和 Tencent Cloud 各占 2%至 3%，服务于特定行业和地区。OpenAI、Google DeepMind 和 Meta 等科技巨头持续推动多模态 AI 与 VLM 发展，推动视觉 AI 由传统图像识别向场景理解与认知分析升级。与此同时，微软、谷歌和亚马逊等云计算厂商则依托 AI 云平台构建大模型训练生态，进一步提升行业壁垒。相比于传统视觉企业，这类公司更强调模型能力、云端部署能力和生态开放能力，是全球视觉 AI 技术迭代的重要推动者。

在工业机器视觉领域，海外企业具有较强先发优势。Mordor Intelligence 的计算机视觉市场报告提到，2025 年计算机视觉市场前五大企业包括 Intel、NVIDIA、Cognex、Keyence、Sony，合计占市场收入约 38%。Keyence、Cognex、Teledyne、Basler、Sony 等企业长期深耕工业自动化、工业相机、图像传感器和视觉检测系统，形成了较完善的产品体系和客户基础。该领域商业化程度较高，客户主要集中于半导体、汽车、新能源、电子制造等高端制造行业。

在视频分析与智能安防领域，全球市场主要由安防硬件厂商、视频管理软件厂商和 AI 分析平台厂商共同竞争。Motorola Solutions/Avigilon、Axis Communications、Genetec、Bosch、Honeywell、Cisco、Verint 等企业在网络摄像机、视频管理系统、智能监控和 AI 视频分析方面具有较强商业化能力。相比中国市场以海康威视、大华股份为代表的硬件与场景一体化模式，海外市场更强调软件平台、云服务和垂直行业解决方案。

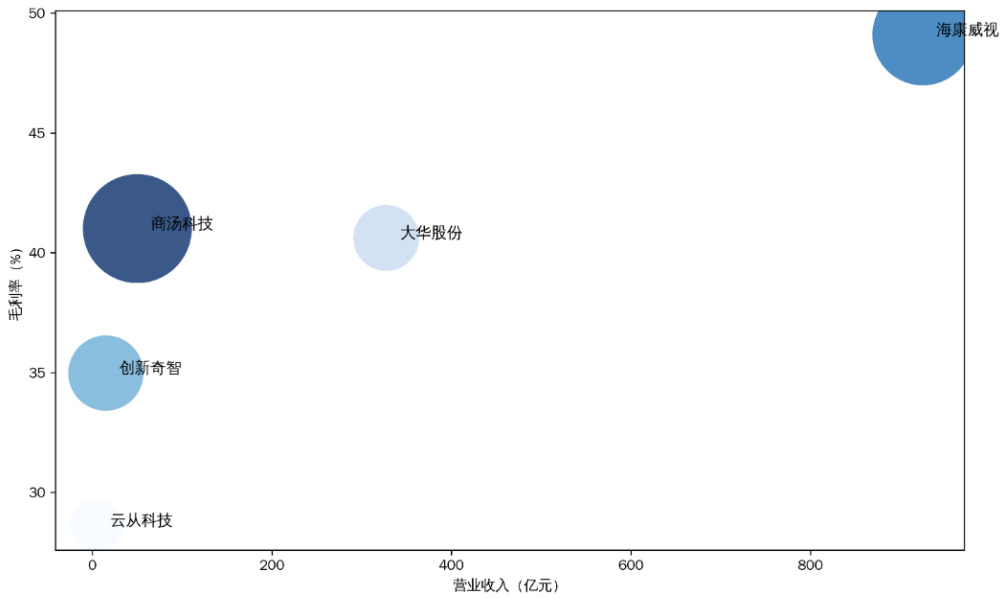
在自动驾驶视觉领域，Tesla、Mobileye、NVIDIA、Waymo 等企业代表了全球视觉 AI 在交通场景中的商业化方向。该领域竞争重点集中在视觉感知算法、车端计算平台、数据闭环和实际道路测试能力。

下载日志已记录，仅供内部参考

整体来看，全球视觉 AI 行业竞争重点已经逐步由传统硬件竞争升级为“算力、模型、数据与场景”的综合生态竞争。

当前国内视觉 AI 行业呈现较为明显的龙头集中趋势。海康威视凭借较大的营收规模、较高的毛利率及领先的市场份额，占据行业主导地位，显示出较强的平台化与规模化优势。大华股份作为行业的主要参与者，在营收规模和市场份额方面位居行业前列，具备较强的渠道和场景落地能力。相较于行业龙头海康威视，大华股份整体规模仍存在一定差距，但在智慧安防和视觉 AI 应用领域已形成较强的竞争力。商汤科技的毛利率较高但营业收入较低，符合其作为算法技术平台的特征，这一点不同于偏硬件商业化的海康威视和大华股份。

图 5：国内视觉 AI 竞争格局



数据来源：公司公告、东北证券

上图中气泡大小视觉 AI 行业市占率，气泡色深表示市占率排名。竞争格局维度包括企业营业收入、毛利率、计算机视觉行业市占率。选取这 3 个维度的原因是首先企业营业收入代表了企业的营利能力，可以看出企业发展状况，未来发展能力。而且，毛利率可以看出企业的盈利能力，能够反映企业的产品竞争力情况。最后，市占率可以看出企业在视觉 AI 行业的市场份额、品牌认知度等关键因素。

相比海外市场以算力、基础模型和云平台为主要竞争优势，国内视觉 AI 市场更突出场景落地和行业解决方案能力，呈现“安防龙头领先、算法平台企业分化、垂直场景厂商突围”的竞争格局。

以海康威视、大华股份为代表的安防与视频物联龙头企业依托摄像头硬件、视频平台、渠道网络和政企客户资源，在智慧安防、城市治理、交通管理等场景中具备较强商业化能力，但面临着硬件增长放缓，需要提升软件与 AI 能力的挑战，因此，海康威视和大华股份分别构建多模态大模型，构建起 AIoT 和物联感知技术平台，支撑它们在安防与场景数字化领域更多的创新应用，从传统安防硬件供应商向“AI 视觉+物联平台+行业解决方案”服务商升级。

商汤科技、云从科技、旷视科技等这类 AI 算法平台企业，主要依托计算机视觉算法、多模态模型和 AI 平台能力，向智慧城市、金融、零售、交通等领域提供软件平台及解决方案。这类企业是国内早期视觉 AI 企业代表，它们最早靠人脸识别、图像识别、城市视觉平台等能力起家，但随着竞争环境的改变，这些单点算法逐渐失去竞争力，行业正在向大模型、多模态、平台化转型。商汤科技在年报提到将继续深

下载日志已记录，仅供内部参考

耕视觉 AI 成熟行业，同时拓展海外市场，并且强调通用视觉大模型、生成式 AI 和 AI 基础设施能力。商汤前瞻性打造新型 AI 基础设施 SenseCore 商汤 AI 大装置，打通算力、算法和平台，大幅降低 AI 生产要素价格，实现高效率、低成本、规模化的 AI 创新和落地，进而打通商业价值闭环，解决长尾应用问题，推动 AI 进入工业化发展阶段。

工业视觉正在成为国内视觉 AI 企业差异化竞争的重要方向。与安防视觉相比，工业视觉更强调检测精度、生产流程适配能力和行业 Know-how，应用场景包括缺陷检测、尺寸测量、智能分拣、质量控制等。海康机器人、凌云光、奥普特、天准科技等企业主要围绕工业相机、光源、视觉检测装备和机器视觉系统展开竞争；创新奇智则更偏向将 AI 视觉能力嵌入工业制造和质量检测场景。该领域竞争的核心不只是算法，而是“算法+工业数据+设备适配+交付经验”的综合能力。

除此之外，还有云、ICT、运营商生态企业，代表企业有华为云、阿里云、百度智能云和电信 AI 公司，以云计算、政企客户、算力资源和私有化交付在视觉 AI 行业表现突出。阿里云、百度智能云、华为云等企业依托云计算、算力资源和 AI 平台能力，为政企客户提供视觉 AI 开发、部署和管理服务。运营商则凭借政企客户资源、云网融合能力和本地化交付体系，切入智慧城市、公共安全、园区治理等私有化部署场景。由此，国内视觉 AI 竞争边界正在从算法厂商之间的竞争扩大为算法公司、安防龙头、工业视觉厂商、云厂商和运营商共同竞争。

整体来看，国内视觉 AI 行业尚未形成单一企业完全主导的格局，而是呈现多赛道竞争特征。海康威视、大华股份等安防视频物联龙头在场景落地、渠道和工程交付方面优势明显；商汤、旷视、云从、依图等算法平台企业在模型和软件能力方面具备技术积累；海康机器人、凌云光、奥普特、创新奇智等企业则在工业视觉和智能制造场景中寻求差异化突围；云厂商和运营商则通过算力、云平台和政企客户资源进入市场。国内视觉 AI 竞争已经从早期的算法能力竞争，逐步升级为“算法、硬件、数据、场景、算力和交付能力”的综合生态竞争。

2.3. 视觉 AI 产业链结构

视觉 AI 产业链上游为基础支持厂商，中游为技术服务厂商，下游为应用领域。基础支持包括硬件支持、数据集和算法支持。硬件支持包括传感器、芯片。目前 CMOS 图像传感器占据中国市场的主流地位，芯片由美国厂商 INTEL 等占主导地位。算法和数据集可以衍生出深度学习框架厂商、服务器厂商和计算平台厂商。这类厂商可以通过开源算法框架快速开发计算机视觉算法及计算机视觉应用。中游技术服务厂商可以分为互联网巨头企业、传统安防转型企业、AI 技术领先企业、AI 垂直领域企业，这四类厂商各自有独特优势。下游应用领域广泛，其中占比最大的是安防视频，为 68%。

上游基础支持持包括硬件支持、数据集和算法支持。硬件支持包括传感器、芯片。目前 CMOS 图像传感器占据中国市场的主流地位，代表厂商有海康威视。芯片为应用算法提供高性能算力服务。目前 CPU、GPU 芯片由美国厂商 INTEL 等占主导地位，中国科技厂商加速研发 ASIC 芯片，代表厂商有寒武纪、商汤科技。算法和数据集可以衍生出深度学习框架厂商、服务器厂商和计算平台厂商。这类厂商可以通过开源算法框架快速开发视觉 AI 算法及视觉 AI 应用。服务器代表厂商有华为、浪潮；计算平台代表厂商有商汤科技、腾讯云、百度智能云。

中游视觉 AI 技术服务商分为四类：第一类为互联网巨头厂商，这类厂商具有强大的数据采集功能和丰富的平台生态，代表企业有阿里巴巴、百度、腾讯。第二类为传统安防转型厂商，依托于政府、安防等项目，布局全产业链，具有极强的工程能力、丰富的项目经验，代表厂商有海康威视、大华股份。第三类为 AI 技术领先厂商，这类厂商独立设计算法，特定场景算法算法和模型训练能力强，代表厂商有商汤科技、云从科技。第四类为 AI 垂直领域厂商，他们基于开源算法框架开发垂直领域产品，专注于特定领域，契合行业需求，代表厂商有小马智能。

视觉 AI 下游应用领域广泛，包括城市服务、政务、交通、教育、零售、物流、医疗、建筑、消费电子等多个领域。2020 年，中国计算机视觉下游应用占比结构分布为：

下载日志已记录，仅供内部参考

安防视频占比最大，为 68%；广告营销占比 17%；泛金融服务占比 9%；互联网娱乐占比 4%；消费电子占比 2%。中国计算机视觉行业各类技术厂商呈现差异化布局态势，其中海康威视、商汤科技的解决方案落地场景最全面。海康威视、大华股份两家传统安防转型厂商通过全产业链布局，覆盖各领域提供软硬件一体化的解决方案。商汤科技通过打造 AI 平台快速落地于各应用场景。

图 6：视觉 AI 产业链



数据来源：公司网站、东北证券

2.4. 硬件与算力：视觉 AI 的基础支撑

视觉 AI 硬件层可以概括为“前端采集+云端训练+边缘推理”，摄像头、工业相机和 CMOS 解决图像数据采集问题；GPU、AI 服务器解决大模型训练和大规模推理问题；端侧 AI 芯片和边缘计算设备则解决实时化、本地化部署问题。

2.4.1. 感知采集硬件：视觉 AI 的数据入口

在感知采集端，摄像头、工业相机、CMOS 图像传感器、镜头和光源构成的视觉 AI 的数据入口。智慧安防、工业检测、自动驾驶等场景都依赖前端设备获取图像或视频数据，因此采集硬件的清晰度、稳定性、低光照能力和实时性会直接影响算法识别效果。随着 AI 摄像机、工业相机和多传感器融合设备的发展，视觉 AI 前端硬件正从单一采集设备向智能感知终端升级。

CMOS 图像传感器：视觉采集的核心元器件

从底层元器件看，CMOS 是摄像头里的核心元件，CMOS 技术是现代电子设备中最常用的集成电路技术之一，广泛应用于微处理器、存储器、传感器等领域。CMOS 传感器相机细分市场预计在 2026 年将占据机器视觉相机市场收入的 55.00%，成为领先的相机类型。这种优势归因于 CMOS 技术所提供的高速性能、低功耗和卓越的画质。主要应用于需要快速图像采集的场景，如自动检测、机器人技术和高通量制造环境。此外，CMOS 摄像机具有良好的性价比，适合在多样化工业应用中大规模部署。随着制造运营不断向智能工厂模式发展，基于传感器的摄像机细分市场因其适应性、效率及与下一代机器视觉软件的兼容性，预计将保持领先地位。

根据 Fortune Business insights 报告，2025 年全球 CMOS 图像传感器市场规模为 243.2 亿美元，预计将从 2026 年的 265.8 亿美元增长到 2034 年的 542.9 亿美元，预测期内复合年增长率为 9.34%。亚太地区主导了全球市场，2025 年占比 43.94%。

互补金属氧化物半导体（CMOS）图像传感器被用于智能手机、单反相机及其他成像设备，因为它们是数码相机的核心组件。它通过将光转换为电信号来捕捉图像。

下载日志已记录，仅供内部参考

CMOS 传感器性价比高且节能，适合电池供电设备，包括智能手机。CMOS 传感器体积较小，所以它们在医学成像领域中越来越多地被应用到内镜、显微镜和诊断成像工具中。此外，这些传感器耗电更低，可拍摄高质量图像，适合医疗应用。所以这一趋势正迅速加速，推动市场增长。

CMOS 市场由成熟的关键企业主导，如半导体零部件工业（安赛米）、索尼半导体解决方案公司、三星电子、松下公司、OMNIVISION 和佳能。这些企业专注于收购拥有创新技术或独特传感器产品的小型企业或初创企业，以获得竞争优势。这一因素预计将推动全球市场增长。

按终端用途，CMOS 市场分为航空航天与国防、汽车、消费电子、医疗保健、安全与监控和其他工业领域。消费电子细分市场在 2026 年占据了最大的市场份额，占 11.87%。由于产品创新不断增加，消费设备市场竞争激烈。相机性能是客户购买设备的首要标准。为了吸引客户关注，制造商主要专注于为相机配备创新技术，如 AI 增强的计算摄影、4K/8K 视频、HDR、慢动作和夜间模式。预计汽车行业在预测期内将以最高的复合年增长率增长。现代汽车每辆车配备了多个摄像头和传感器，尤其是在电动车和高端车型中。对支持 ADAS 的汽车需求不断增加，需要 CMOS 图像传感器用于车道和行人检测。

北美在全球 CMOS 市场保持强势地位，2025 年市场规模达到 66.6 亿美元，占比 27.37%，预计 2026 年将达到 72.5 亿美元。北美预计将以显著的速度增长。该地区在视频监控、边境监控和关键基础设施保护方面投入巨大，这需要高性能 CMOS 传感器。在安全和监控应用领域，低光性能、近红外灵敏度和夜视设备的需求很高，这对推动该地区市场增长起着至关重要的作用。预计到 2026 年，美国市场将达到 50.7 亿美元。在美国，CMOS 传感器主要用于医学成像应用，如牙科成像、内窥镜、眼科等，因其高分辨率的图像质量、体积小巧且性价比高。这一因素在推动全国市场增长中起着关键作用。亚太市场在 2025 年占 106.9 亿美元，占全球 CMOS 行业的 43.94%，预计 2026 年将达到 115.8 亿美元。并且是智能手机、笔记本电脑、可穿戴设备和平板电脑的最大制造商。亚太地区拥有三星电子、索尼、松下、佳能、小米等 OEM 厂商。此外，中国、印度、东南亚和韩国拥有快速增长的技术用户和移动优先消费者群体，这些群体在推动该地区市场增长中发挥着关键作用。此外，CMOS 市场高度依赖亚洲供应商提供晶圆和组装。中国是配备 CMOS 图像传感器的智能家居设备、智能手机、平板电脑和笔记本电脑的主要生产和消费国。据业内专家称，2024 年中国国内出货智能手机超过 3.5 亿部。此外，该国拥有显著的 CMOS 传感器封装和相机模块组装网络，这些公司在降低封装成本和加快上市时间方面发挥着关键作用。这些因素在推动中国市场增长中起着重要作用。日本市场预计到 2026 年将达到 22.7 亿美元，中国市场预计到 2026 年达到 23.3 亿美元，印度市场预计到 2026 年达到 16.5 亿美元。

机器视觉摄像头：工业视觉中的核心采集设备

在工业视觉领域，机器视觉摄像头，也就是工业相机，是工业场景里的图像采集设备，主要用于工厂检测、缺陷识别、尺寸测量、分拣、机器人引导等场景。它通常由 CMOS/CCD 图像传感器、镜头、光源、接口和图像处理模块构成。根据 Future Market Insights 报告，机器视觉摄像头市场预计将从 2025 年的 42 亿美元增长到 2036 年的 101 亿美元，复合年增长率为 8.3%。机器视觉相机市场范围包括基于 CMOS 和 CCD 传感器的摄像头，应用于汽车、制药、食品、电子和包装检测领域的区域扫描、线扫描和三维功能。机器视觉相机市场正经历强劲增长，得益于制造、汽车、电子和物流行业自动化的兴起。此外，对实时监控和预测性维护的需求加速了机器视觉解决方案在终端行业的采用。技术进步，包括传感器性能提升、更快的处理能力以及可扩展的部署选项，正在塑造该市场的未来。自动导引车辆、基于 AI 的缺陷检测和装配线监控的广泛应用，持续创造高精度成像解决方案的机遇。随着制造商越来越重视运营效率和质量保证，机器视觉摄像头市场有望在发达和新兴地区持续增长。中国以最高的复合年增长率（11.2%）领先。

机器视觉相机市场的关键参与企业有美国的康奈克斯（Cognex）公司和 Teledyne 科技公司、日本 KEYENCE 公司、德国的巴斯勒股份（Basler AG）公司。其中，当前

下载日志已记录，仅供内部参考

全球高端机器视觉市场仍由美日德企业主导 Cognex 与 Keyence 合计占据全球 50%以上份额，在高端工业相机、智能视觉软件领域对中国市场的控制力达 55%。Cognex 公司主导视觉引导和深度学习系统，用于缺陷检测、可追溯性、条码读取、机器人引导以及制造和物流中的良率优化，以卓越的速度、稳健的准确性、自动化学习、无缝的可追溯性和高运营投资回报率取得市场优势；KEYENCE 公司主要运用于拣货、物体识别、生产质量控制、动态工厂检查和自动化的视觉系统，以灵活校准、广泛的市场支持、多样化的检测模式等优势提升应用场景的生产力；Basler AG 公司主要生产先进的摄像头（2D、3D、SWIR、线扫描、嵌入式）以及基于 AI 的视觉软件，凭借高图像精度、快速缺陷检测、模块化系统、可靠的 AI 分析、易于集成等优势广泛应用于工业检查、机器人、物流、制药、汽车和智能基础设施。

AI 摄像头：从单一采集设备向智能感知终端升级

在智能终端层面，AI 摄像头代表了摄像头从图像采集工具向智能感知终端的升级方向。全球摄像头行业已告别爆发式增长阶段，进入“存量替换+细分需求驱动”的稳健发展期。行业增长核心动力从新增需求转向技术升级与场景深耕，整体保持平稳增长态势。2024 年全球消费智能摄像头出货量达 1.37 亿台，同比增长 7.7%；2025 年出货量 1.43 亿台，同比增长 2.2%，增速虽较过往有所放缓，但长期增长趋势未改，2028 年有望攀升至 1.65 亿台。

中国作为全球核心市场，同样呈现“稳中有升”特征。2024 年中国消费智能摄像头出货量 5139.8 万台，同比增长 4.2%；2025 年出货 5310 万台，同比增长 3.3%，预计 2028 年突破 6187 万台大关。驱动因素方面，一方面是传统摄像头设备的更新换代需求，老旧设备在分辨率、智能化水平上难以满足当前需求；另一方面是新兴场景的增量需求，如智慧家居、工业监控、户外安防等领域对专业化摄像头的需求持续释放，为行业增长提供新支撑。

而 2025 年全球 AI 摄像头市场规模约 177 亿美元，同比增长 27%。AI 摄像头是指搭载端侧/云端 AI 算法（含人形识别、人脸抓拍、车牌识别、行为分析、区域警戒等智能功能），具备自主感知、识别与决策能力的摄像设备，核心区别于仅提供基础录像、移动侦测功能的普通摄像头。目前行业正处于“结构升级+存量替换+场景扩容”的高景气阶段，价格端消费级 AI 与普通机型价差收窄至 50-150 元，商用/工业级价差 30%-50%，驱动渗透率加速提升，未来 3-5 年将以端侧 AI、云边协同、多模态感知与合规为核心主线。

2.4.2. 云端算力硬件：GPU 与 AI 芯片支撑大模型训练

在算力端，GPU 和 AI 服务器是视觉 AI 模型训练与推理的核心基础。随着 CNN、Transformer 和 VLM 等模型的不断发展，视觉 AI 需要处理的数据规模和模型复杂度持续提升，对并行计算和高性能算力的需求不断增强。在海外市场中，NVIDIA、Microsoft、Google、Amazon、Intel、Teledyne 等列为主要参与者。

GPU 是当前视觉 AI 训练和大模型推理的核心算力硬件。相比 CPU，GPU 更适合并行计算，能够高效处理图像、视频和深度学习模型中的矩阵运算，因此广泛运用于 CNN、Transformer、VLM 等模型训练。随着视觉 AI 模型从单一图像识别向是视频理解、多模态推理升级，GPU 和 AI 服务器在产业链中的战略地位进一步提升。

全球 AIGPU 基础设施市场规模在 2025 年为 456 亿美元。预计该市场将从 2026 年的 531 亿美元增长到 2034 年的 1478 亿美元，预测期内复合年增长率为 14.2%。AI GPU 基础设施市场的快速扩展，得益于生成式 AI 应用需求的激增、云端 AI 服务的广泛采用以及需要强大计算能力的深度学习算法的进步。此外，从医疗、金融到汽车和零售等多个行业，也在将 AI 驱动的分析整合进运营，从而提升了对能够高效处理复杂工作负载的可扩展 GPU 基础设施的需求。

AI GPU 基础设施市场正经历快速增长，主要受 AI 和机器学习应用日益增长的需求推动。多家主要参与者竞争市场份额，提供多种 GPU 硬件、云服务和软件平台。这

下载日志已记录，仅供内部参考

种竞争环境促进创新，并为终端用户提供多样化的选择以满足特定的计算需求。市场的扩展凸显了高性能计算在推动 AI 研究和部署方面的关键作用。
成熟的科技巨头和专业半导体制造商在该市场中占据重要地位。这些公司正在大量投入研发，打造更强大、节能的 GPU 以适应 AI 工作负载。此外，云服务提供商正在扩展其 GPU 产品，为各种规模的企业提供可访问且可扩展的 AI 基础设施。竞争格局以 GPU 架构、软件优化和集成解决方案的持续进步为特征。从全球地域格局来看，北美 AI GPU 基础设施市场以其快速发展和大量投资为特征。领先科技公司和研究机构的汇聚营造了一个充满活力的生态系统，不断寻求更强大、更高效的 GPU 解决方案来应对复杂的 AI 挑战。随着下一代 GPU 架构和专为 AI 工作负载设计的专用硬件的出现，市场正持续演变。玩家之间的激烈竞争激发了创新，降低了成本，使 AI GPU 基础设施更加易获得。重点仍在于可扩展性、性能和能效，以满足 AI 应用日益增长的需求。亚太地区是 AI GPU 基础设施增长最快的市场，主要由快速扩张的经济体和蓬勃发展的 AI 生态系统推动，尤其是在中国和印度。AI 在电子商务、金融和智慧城市等领域的日益普及，正在创造对强大计算能力的巨大需求。政府对 AI 开发的支持以及技能人才的可用性进一步加速了市场增长。该地区既有机遇，也面临挑战，包括基础设施建设和监管复杂性。

2025 年，全球 AI 服务器市场规模为 1946.2 亿美元。预计该市场将从 2026 年的 2622.2 亿美元增长到 2034 年的 28473.2 亿美元，预测期内复合年增长率为 34.73%。
AI 服务器市场代表了现代 AI 基础设施的关键支柱，可实现数据密集型 AI 工作负载所需的高性能计算。 AI 服务器是专门构建的系统，针对云数据中心、企业和研究机构的机器学习、深度学习和数据分析应用程序进行了优化。AI 服务器市场分析强调了自然语言处理、计算机视觉和大规模数据建模等基于 AI 的工作负载的日益普及所推动的快速部署。这些服务器旨在处理大规模并行处理、高内存带宽和加速计算。训练高级 AI 模型需要处理大量数据集，而传统服务器无法有效处理这些数据集。AI 服务器通过 GPU、ASIC 和其他加速器提供加速计算能力。对可扩展和节能计算环境不断增长的需求继续塑造 AI 服务器市场前景，将 AI 服务器定位为数字化转型战略的基本要素。金融、医疗保健、零售和制造等行业部署 AI 服务器以提高运营效率和竞争优势。AI 工作负载日益复杂，直接推动了对先进服务器基础设施的持续需求。AI 服务器市场前景的一个关键限制是部署 AI 服务器基础设施所需的高资本投资。AI 服务器涉及昂贵的组件，例如高端加速器、先进的冷却系统和高速互连。企业还必须投资支持基础设施，包括电力输送、冷却和专用网络。部署复杂性增加了与现有 IT 环境的集成挑战。较小的组织可能会面临预算限制，从而延迟采用。此外，快速的硬件发展可以缩短升级周期，增加长期拥有成本。尽管具有强大的性能优势，但这些因素限制了成本敏感型企业的采用。
美国在 AI 服务器市场中占据主导地位，这得益于先进的数据中心基础设施和企业对 AI 技术的大力采用。AI 服务器广泛部署在云服务提供商、技术公司和研究组织中，以支持大规模 AI 模型的开发和部署。美国市场受益于金融、医疗保健、自主系统和零售等行业对 AI 训练和推理工作负载的高需求。企业对高性能计算环境的投资推动了服务器的持续升级。美国 AI 服务器市场洞察表明 GPU 加速和定制 AI 服务器架构得到广泛采用。在快速数字化和大规模 AI 采用的推动下，亚太地区占据全球 AI 服务器市场约 30%的份额。该地区的企业大力投资 AI 基础设施，以支持自动化和分析。制造、电子和技术行业是主要采用者。各国政府通过国家战略积极推动 AI 发展。亚太地区仍然是全球增长最快的 AI 服务器市场之一。
AI 服务器市场趋势表明，市场明显转向针对特定 AI 工作负载优化的专用服务器架构。企业越来越多地部署专为并行处理、高密度计算和加速数据吞吐量而设计的 AI 服务器。最显著的趋势之一是越来越多地采用能够处理大型语言模型和复杂神经网络的 AI 训练服务器。这些系统优先考虑内存带宽、互连速度和可扩展性。AI 服务器行业分析的另一个关键趋势是对能源效率和热优化的日益关注。数据中心部署液体冷却和先进的气流设计来管理与 AI 工作负载相关的高功耗。随着组织将 AI 模型从开发环境转移到生产环境，AI 推理服务器也越来越受到关注。此外，将本地 AI

下载日志已记录，仅供内部参考

服务器与基于云的基础设施相结合的混合部署模型变得越来越普遍。这些趋势共同重塑了整个 AI 服务器市场的采购策略和系统设计。

从全球市场看，NVIDIA 是 AI 算力层最具代表性的企业之一。截至 2025-2026 年，英伟达在全球独立显卡市场占有率约94%，在高端AI GPU市场占据绝对主导地位。根据 Jon Peddie Research 的报告，英伟达 2025 年第二季度全球独立显卡（AIB）出货量达到 1160 万块，英伟达市场份额升至 94%，AMD 下降至 6%，英特尔保持在1%以下 。在数据中心高端训练 GPU 市场，英伟达凭借 Blackwell 架构和 CUDA 生态系统占据主导地位，2026 年 Q1 全球 AI GPU 服务器市场份额达到 68%，AMD 维持 5%-6%，国产 GPU 整体不足 4%。英伟达数据中心业务收入快速增长，反映出全球大模型训练、AI 推理和数据中心建设对高性能 GPU 的旺盛需求。
在中国本土市场，英伟达与华为形成“双核”格局，2025 年两者市场份额均约 40%，国产厂商如寒武纪、海光、百度昆仑等各占据 3%-4%。2026 年 Q1，国产 GPU 整体市场份额突破 41%，主要集中在政企、边缘算力和中端推理场景，显示国产化替代进程加速 。

2.4.3. 端侧 AI 芯片与边缘计算硬件：推动实时化、本地化部署

AI 芯片与推理加速：支持视觉 AI 规模化部署
与 GPU 主要用于云端训练和大规模推理不同，端侧 AI 芯片和边缘计算硬件主要服务视觉 AI 的实时推理和本地化部署。智慧安防、工业质检、自动驾驶和机器人等场景通常要求毫秒级响应，如果全部依赖云端传输和计算，会面临带宽、延迟和数据安全问题。因此，NPU、ASIC、SoC 等专用 AI 芯片开始被集成到摄像头、车端计算平台、机器人和工业终端中，推动视觉 AI 从云端集中部署向端云协同部署发展。

端侧 AI 芯片主要用于摄像头、机器人、车载设备和工业终端中完成本地推理计算。相比于云端 GPU，端侧 AI 芯片更强调低功耗、低延迟、小体积和稳定性，适合部署在 AI 摄像机、边缘盒子、车载计算平台和嵌入式设备中。通过在前端设备中集成 NPU、SoC、ASIC 等芯片，视觉 AI 可以在本地完成目标识别、异常检测、行为分析和环境感知等任务。从市场趋势看，AI 芯片市场保持较快增长。Markets and Markets 预计，全球 AI 芯片市场将从 2025 年的 2032.4 亿美元增长至 2032 年的 5648.7 亿美元。AI 芯片市场的快速扩张得益于高性能 GPU、加速器和专用处理器的需求激增，这些处理器支持大规模训练、推理和边缘智能工作负载。对先进封装、HBM 产能和下一代代工厂技术的大规模投资进一步推动创新，强化市场长期增长轨迹。AI 芯片市场正在重塑整个技术生态系统的价值创造。随着云服务提供商、数据中心运营商、OEM 厂商和基础设施供应商将先进的 AI 处理器集成到其平台，下游利益相关者从 AI 开发者、工业运营商到医疗、汽车和电信团队都将获得显著增强的计算能力。这种多层次的影响最终带来了可衡量的结果，如更快的推理、更完善的自动化、更低的运营成本以及加速的企业和消费市场数字化转型的新型智能服务。
AI 芯片生态系统由芯片设计师（模拟器件、德州仪器、SK 海力士、英特尔、三星）、半导体制造商（台积电、英特尔、ASML）、芯片供应商（英伟达、AMD、谷歌）以及终端用户（西门子、谷歌、AWS、Microsoft）组成。设计师设计先进的处理器架构，由制造合作伙伴使用尖端半导体设备制造。芯片供应商提供高性能的 AI 加速器，支持云、企业和边缘应用。终端用户推动了对更快计算、能效和可扩展 AI 工作负载的需求，而跨价值链的紧密协作则促进了持续创新和市场扩展。OVH SAS 将第四代 AMD EPYC 处理器集成到其裸机服务器系列中，以提升性能和可靠性，以应对高负载的 AI 推理工作负载，这带来了 15-20%的性能提升、更高的弹性和提升的核心密度，使 OVH SAS 能够提供更具成本效益、高性能的云解决方案；腾讯采用了具备先进加速功能的第三代英特尔至强可扩展处理器，驱动其小微智能语音视频服务平台，实现高质量的神经 TTS 处理。

下载日志已记录，仅供内部参考

边缘计算是端侧 AI 芯片等应用载体，主要包括 AI 摄像头、边缘计算盒子、工业边缘服务器、车载计算平台和机器人控制器等设备。其作用是在靠近数据产生的位置完成计算，减少数据传回云端的需求，从而提升响应速度并降低带宽压力。在智慧安防中，边缘 AI 摄像头可以直接完成目标识别和异常检测；在工业检测中，边缘设备能够在生产线上实时判断缺陷；在自动驾驶中，车端计算平台需要快速处理道路、行人和交通标志等视觉信息。

2025 年，全球边缘 AI 市场规模为 358.1 亿美元。该市场预计将从 2026 年的 475.9 亿美元增长到 2034 年的 3858.9 亿美元，预测期内复合年增长率为 29.9%。边缘 AI 硬件市场受到多种因素驱动，包括对物联网边缘计算解决方案需求的增长、整合 IT 与电信的 5G 网络日益普及，以及对专用 AI 处理器用于设备端图像分析的需求增加。然而，市场增长潜力受到有限的设备内训练能力和短缺的 AI 专业人员的严重阻碍。自动驾驶汽车和机器人需求的上升也推动边缘 AI 市场增长。自动驾驶车辆需要实时向云端传输数据以高效运行，先进的 AI 和机器学习技术能帮助这些自动驾驶系统做出决策，因此自动驾驶车辆连接边缘，以提升效率、减少事故、提升安全性并减少交通拥堵。而且，机器人领域因边缘 AI 能帮助降低延迟和带宽需求而广泛使用。但边缘设备应用因被保护和存储的数据具有网络敏感性而面临安全威胁，而且边缘计算高度依赖硬件，制造原型所需的时间很长，投入生产也很耗时。此外，资源、培训和研发的成本也非常高，许多因素如使用场景、功耗、内存需求和处理器等，使得该技术的实现具有挑战性。

北美地区在 2025 年占据了全球市场的 34.80%，创造了 124.5 亿美元的收入，预计 2026 年将达到 161.2 亿美元。北美市场的增长得益于该地区企业，尤其是美国，对先进技术的高度重视，包括 AI、深度学习和机器学习，这在帮助该地区保持竞争领先方面发挥着关键作用。亚太地区预计会成为全球边缘 AI 硬件增长最快的地区。2025 年亚太地区创造了 98.3 亿美元收入，占全球市场收入的 27.40%，预计 2026 年将增长至 134.8 亿美元。亚太市场预计将创下最高的复合年增长率。中国市场规模在 2025 年达到 33.1 亿美元，中国的主要参与者通过向发展中国家提供先进技术解决方案，扩大其地理布局。

边缘 AI 硬件市场的生态系统分析涵盖了多方利益相关者，包括 OEM、硬件供应商以及边缘 AI 软件/平台提供商。NVIDIA、英特尔、苹果、联发科和华为等关键 OEM 推动硬件创新，而三星、美光、Microsoft、高通和 Hailo 等领先硬件供应商则提供关键零部件和技术。生态系统还包含了 Synaptics、TIBCO、Octonion Group 和 Tact.ai 等边缘 AI 软件和平台供应商，凸显了跨行业提供全面边缘 AI 解决方案所需的协作环境。其中，高通为无人机基础设施空中检查场景提供骁龙 Edge AI 解决方案，实现设备端实时图像分析与缺陷识别，这种低延迟处理保障飞行安全，增强巡检过程的实时监控能力，同时降低整体运营成本；Hailo 的 Hailo-8 AI 处理器集成于智能安防摄像头，实现实时异常事件检测（如入侵、越界、物品遗留等），替代传统云端分析模式，从而实现本地实时监控，大幅降低带宽消耗，同时保障数据处理的高效性与低延迟；英特尔通过 Movidius Myriad X 处理器赋能零售场景智能摄像头，实现客户行为分析（如动线追踪、停留时长统计），帮助消费者提供个性化购物体验，助力商户减少盗窃事件，优化店铺布局与商品陈列策略。

2.5. 软件层与算法平台发展现状

2.5.1. 算法模型：从识别检测走向多模态理解

从算法能力看，视觉 AI 软件与算法层承担视觉理解和能力输出的功能，是连接上游硬件基础设施和下游行业应用的核心环节。上游摄像头、传感器和 GPU 提供数据和算力基础，而软件层则通过图像识别、目标检测、视频分析多模态理解等算法能力，将原始图像和视频数据转化为可用于决策的结构化信息。随着行业发展，视觉 AI 软件不再局限于单点算法，而是逐步向平台化、工程化和行业解决方案方向演进。

算法是软件层的基础

下载日志已记录，仅供内部参考

视觉 AI 算法是软件层的核心基础，其作用是将图像和视频中的非结构化信息转化为可识别、可分析、可决策的结构化信息。当前视觉 AI 算法主要包括图像分类、目标检测、图像分割、OCR、人脸识别、行为识别、姿态估计等任务，广泛应用于安防、工业检测、自动驾驶、医疗影像等领域。

目标检测算法是一种计算机算法，可以通过在图像或视频周围绘制边界框来识别和定位特定感兴趣的物体。它可分为两级探测器和单级探测器，每种探测器在精度和速度上各有优势。目标检测分为两大系列，RCNN 系列和 YOLO 系列，另外还有著名的 SSD 是基于前两个系列的改进。这些算法在多种应用中发挥着关键作用，包括探测森林火灾。

R-CNN(Regions with CNN features) ，是 R-CNN 系列的第一代算法，其实没有过多的使用深度学习思想，而是将深度学习和传统的计算机视觉的知识相结合。R-CNN 算法在目标检测领域取得了显著的成果，被广泛应用于各种实际场景中。例如，在安防监控领域，R-CNN 算法可以用于检测视频中的行人、车辆等目标；在自动驾驶领域，R-CNN 算法可以帮助车辆识别道路上的交通标志、行人等；在医疗影像分析领域，R-CNN 算法可以用于辅助医生检测病变区域等。

YOLO（You Only Look Once）是继 RCNN，fast-RCNN 和 faster-RCNN 之后，Ross Girshick 针对 DL 目标检测速度问题提出的另一种框架。YOLO 创造性的将物体检测任务直接当作回归问题来处理，将候选区和检测两个阶段合二为一。只需一眼就能知道每张图像中有哪些物体以及物体的位置。通过一次前向传播即可同时预测多个目标的类别与位置，兼具高速度与高精度，在安防监控、自动驾驶、工业质检、医疗影像等领域广泛应用。其最新版本 YOLOv8 结合 CNN 与 Transformer 架构，在精度与实时性之间取得平衡。目标检测、深度学习、轻量化、缺陷检测、机器视觉、特征融合和损失函数是 YOLO 算法中的高频关键词。随着 YOLO 算法的研究在日趋精细化，结合深度学习和神经网络算法技术的日益更新壮大，YOLO 算法的应用将越来越广泛。

在深度学习阶段，CNN、YOLO、Faster R-CNN、Mask R-CNN 等模型推动图像识别、目标检测和分割能力快速提升，使视觉 AI 从实验室研究逐步进入商业化应用。当前，目标检测、人脸识别、车牌识别、缺陷检测等算法能力已经较为成熟，成为智慧安防、工业质检和交通管理等场景的基础能力。

VLM 和多模态推动升级

随着 Transformer、多模态模型和 VLM 发展，视觉 AI 算法能力正从“识别目标”向“理解场景”升级。传统算法更强调识别图像中有什么，而多模态算法进一步强调图像、文本和视频之间的语义关联，使系统能够进行图文问答、视频理解、事件推理和复杂场景分析。未来，算法竞争重点将从单一准确率提升，转向泛化能力、场景理解能力和工程化部署能力。

VLM（Vision-Language Model）是一种能够同时处理和理解多种模态信息的 AI 模型。它结合了计算机视觉（CV）和自然语言处理（NLP）的技术。VLM 弥合了“看”与“解释”之间的差距，使得需要结合视觉和文本信息进行联合推理的应用成为可能。常见的用例包括图像字幕生成、视觉问答和内容审核。这些模型在结合视觉上下文与语言理解能够提高准确性或用户体验的场景中特别有用。VLM 为辅助和交互式系统提供动力。在医疗保健领域，它们可以分析医学图像以及患者病史，以辅助诊断或生成报告。对于机器人技术，VLM 使机器人能够根据视觉输入执行自然语言命令，例如“捡起架子上的蓝色工具”。开发者可以使用 Hugging Face 的 Transformers 或 PyTorch 等框架构建这些系统的原型，并在特定领域的数据上对预训练模型进行微调。另一个新兴用途是教育工具，例如从教科书图表中生成测验题目。通过结合视觉和文本推理，VLM 在各个行业中解锁了实用的解决方案，同时通过开源库和云 API 仍对开发者友好。

多模态模型是指将不同类型的数据结合起来进行分析处理的模型。其通过不同数据类型的相互关联和结合，可以大幅提高模型的准确性和鲁棒性，应用场景进一步拓展。其次，多模态更接近人类学习的模式，在物理世界中，人类认知一个事物并不

下载日志已记录，仅供内部参考

是通过单一模态，所以多模态全面立体认知是未来 AI 的发展方向。多模态可以结合图像与问题文本，多模态模型可理解复杂语义关系。例如，针对“图中穿红衣服的人在做什么？”的问题，模型需同时定位人物、识别动作并关联颜色描述；多模态也可以进行视觉-语言导航，机器人根据语言指令在环境中导航，需融合视觉场景理解与语言语义解析。

多模态 AI Agent 技术已在多个行业展现出显著的实用价值。在保险行业，多模态 Agent 能够大幅提升理赔流程的效率和准确性，当用户提交理赔申请时，Agent 可以同时处理事故现场照片、语音描述的文字记录、以及结构化表单数据，自动完成损失评估和报告生成。教育领域是多模态 Agent 的另一个重要应用场景，这类 Agent 可以理解学生手写的解题过程、识别语音提问中的困惑点，并提供个性化的辅导。在工业 4.0 场景中，多模态 Agent 能够整合设备振动传感器数据、红外热成像和维修记录文本，预测潜在故障并建议维护措施。医疗诊断 Agent 结合医学影像(CT、MRI)、电子病历和最新医学文献，为医生提供第二意见和诊断参考。

2.5.2. AI 视频分析软件：软件层商业化最成熟的方向之一

AI 视频分析是视觉 AI 软件层商业化程度较高的方向。传统视频监控主要是承担录像和回放功能，而 AI 视频分析可以实现目标识别、异常行为检测、事件识别、交通流量分析和场景管理等能力。随着智慧城市、交通管理、零售分析和园区安防需求提升，视频分析软件逐渐成为视觉 AI 中游的重要产品形态，行业价值也从单一硬件销售向软件订阅、平台服务和行业解决方案延伸。

2025 年，全球视频分析市场规模为 122.9 亿美元，预计将从 2026 年的 148.1 亿美元增长到 2034 年的 650.8 亿美元，预测期内复合年增长率为 20.30%。视频分析市场正在受到自动化、实时决策和预测分析的影响。由于主要国家的智慧城市举措不断增多，以及部署基于分析的投资不断增加，对这些解决方案的需求预计将获得关注。视频监控系统。这些系统可以在执法、政府、交通、银行和金融部门得到广泛使用。此外，分析平台的不断普及，通过及时发出有关可疑或异常活动的警报来最大限度地降低犯罪率，预计将推动预测期内的市场增长。Avigilon Corporation、Honeywell International Inc.、Robert Bosch GmbH 和 IBM 等市场参与者不断致力于创新更先进的视频分析产品/服务，以保持在这个竞争激烈的市场中的份额。

从全球区域看，北美在全球视频分析市场中保持着强劲的地位，2025 年达到 39 亿美元，占 31.70%的份额，预计 2026 年将达到 45.6 亿美元。新技术的引入以及 IBM 公司、霍尼韦尔国际公司、思科系统公司、摩托罗拉解决方案（Avigilon）等主要公司的出现，预计将增加美国和加拿大对视频内容分析解决方案的需求。美国和加拿大的最终用户正在积极采取行动，采用与新兴技术相结合的视频内容分析。预计到 2026 年美国市场将达到 32.2 亿美元。2025 年，亚太地区市场规模为 26.2 亿美元，占全球行业的 21.30%，预计到 2026 年将达到 32.7 亿美元。政府资助的智慧城市项目和国家安全先进监控项目预计将推动该地区视频分析解决方案的使用。零售业和信息通信技术领域的重大发展正在推动中国和印度等国家的市场扩张，中国市场预计到 2026 年将达到 8.6 亿美元。

视频分析市场趋于整合，思科系统公司（Cisco）、Avigilon Corporation (Motorola Solutions, Inc.)、IBM Corporation、NTT Data Group Corporation、Honeywell International Inc 和 Robert Bosch GmbH 等顶级市场参与者占据了 41%的市场份额。主要参与者专注于新产品创新、升级、将产品组合扩展到新的地理区域，以及与其他参与者合作以扩大其在不同地区的业务。其中，Cisco 公司凭借集成网络驱动的分析解决方案占据主导地位。视频分析生态系统的特点是适度分散但创新密集的结构，大型企业和专业供应商在各个应用层上展开竞争。虽然 Cisco 公司和安讯士通信公司等公司凭借其全球基础设施和品牌资产保持了稳固的立足点，但新兴企业正在利用 AI 驱动的能力颠覆价值链。由于需要先进的 AI 专业知识、高资本投资以及与数据隐私相关的监管合规性，进入壁垒相对较高。然而，基于云的部署模型和开放平台正在降低敏捷初创公司的壁垒。随着供应商在性能准确性、可扩展性和集成能力方面展开竞争，竞争强度正在不断升级。定价动态正在从以硬件为中心的模型转向基于订

下载日志已记录，仅供内部参考

阅的软件和分析即服务产品。创新仍然是关键的差异化因素，各公司在深度学习、面部识别、行为分析和预测智能方面进行了大量投资。与云提供商和物联网生态系统参与者的合作关系正在进一步重塑竞争地位。

2.5.3. 软件私有化：中国市场的重要竞争领域

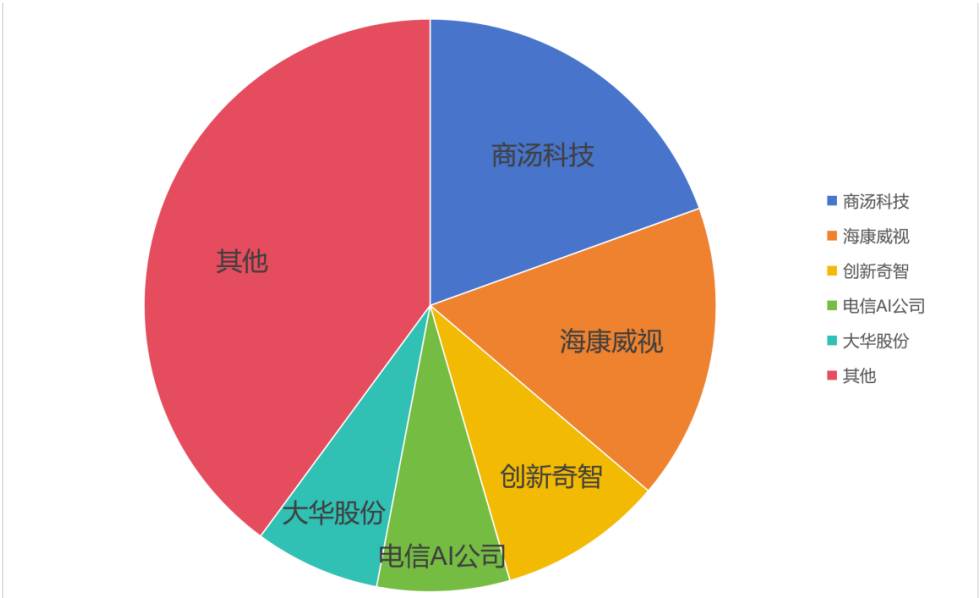
私有化部署是将视觉 AI 软件、算法模型、数据处理系统和管理平台部署在客户本地服务器、专有云或内部数据中心，而不是完全依赖公有云 SaaS 服务。与公有云 SaaS 模式相比，私有化部署更强调数据本地存储、系统自主可控和定制化交付，通常适用于政企、公安、交通、能源、金融、工业制造等对数据安全和系统稳定性要求较高的场景。

在中国市场，视觉 AI 软件呈现明显的私有化部署特征。政企、安防、工业和城市治理等客户通常更重视数据安全、系统可控和本地化交付，因此私有化部署成为视觉 AI 软件商业化的重要模式。视觉 AI 软件私有化市场的竞争主体并不局限于纯算法公司，还包括安防硬件龙头、垂直行业方案商和运营商生态企业，说明视觉 AI 软件层竞争正在从单点算法能力转向“算法+场景+交付+客户资源”的综合竞争。

根据 IDC 数据，2025 年中国视觉 AI 软件私有化市场规模达到 92.5 亿元，创新奇智凭借工业场景深耕能力蝉联 TOP3，持续稳居中国视觉 AI 第一梯队。IDC 指出，中国私有化 AI 市场已经走出技术演示阶段，进入以场景深耕、工程化交付、多模融合为核心的深水区竞争。传统的计算机视觉以感知为核心，依赖针对特定场景训练的专用模型，一个场景一套算法，部署成本高、泛化能力有限。而随着大模型的崛起，CV 2.0 正在重新定义这一市场，从多模型到统一大模型解决多场景问题，从单模态感知到图文多模态理解，从闭集识别到开集推理，从单纯的“看”到“看懂、会搜、能生成”。

视觉 AI 软件私有化市场格局上，商汤科技以 19.5%的份额居首。作为国内计算机视觉领域的开创者，商汤凭借深厚的算法积累与超大规模训练能力，在城市安全、零售、汽车等场景保持领先。海康威视以 16.7%的份额位列第二，依靠硬件与软件深度整合的产品矩阵以及遍布全国的渠道网络，在安防与工业视觉领域具备不可忽视的规模优势。创新奇智主攻工业制造场景的视觉质检与缺陷检测，已在汽车、电子制造等头部客户中实现批量复制；电信 AI 公司与大华股份则分别凭借运营商生态和安防硬件深度，在公共安全与智慧园区场景持续渗透。

图 7：视觉 AI 软件私有化市场份额



数据来源：IDC、东北证券

下载日志已记录，仅供内部参考

但私有化部署也存在一定局限。由于客户端需求差异较大，视觉 AI 软件往往需要定制化开发和本地视频，导致项目交付周期较长、标准化复制难度较高。同时，私有化部署需要持续运维和版本升级，对厂商工程化能力和服务能力提出更高要求。未来，随着 VLM、多模态模型和边缘推理发展，视觉 AI 私有化部署将从单点算法交付，逐步向平台化、模块化和端云协同方向演进。

3. 行业应用与商业化现状

下游应用层是视觉 AI 实现商业化变现的关键环节。与上游硬件和中游算法相比，下游更强调场景需求、客户付费能力和解决方案落地效果。目前视觉 AI 已在智慧安防、城市治理、工业检测、自动驾驶、医疗影像、零售分析等领域实现应用。其中，智慧安防和工业视觉商业化较高，自动化驾驶和机器人具备长期成长空间，医疗影像、零售等场景则处于持续渗透阶段。整体来看，视觉 AI 下游应用正由单一识别检测向行业决策辅助和流程化智能化升级。

3.1. 质检与缺陷检测：当前最成熟的工业视觉 AI 应用

质检与缺陷检测是工业视觉 AI 最成熟、商业价值最明确的应用场景。由于缺陷多种多样，传统的计算机视觉算法很难做到对缺陷特征完整的建模和迁移，复用性不大，要求区分工况，这会浪费大量的人力成本。深度学习在特征提取和定位上取得了非常好的效果，越来越多的工程人员开始将深度学习算法引入到缺陷检测领域中。视觉 AI 通过高分辨率工业相机采集产品表面、形状、尺寸、装配状态和外观缺陷，可以实现在线化、实时化和全量化检测，检测精度可达到微米量级，漏检率压缩至 0.1%以下。以锂电池制造为例，电池极片的涂布均匀性、极耳焊接质量等关键指标均可由视觉 AI 实现在线实时监测，替代原有的抽检模式，实现全覆盖。视觉 AI 还能通过计算机视觉技术对巡检图片及视频进行实时识别，且支持多种设备常见缺陷的精准识别，比如渗漏油、绝缘子破损、锈蚀等缺陷的识别。

从下游需求结构来看，新能源和 3C 电子是当前机器视觉质检需求最旺盛的两大领域。新能源电池领域受益于动力电池和光伏产能的大规模扩张，质检要求随产线建设同步放量；3C 电子领域受益于消费电子迭代加速，产品精密度要求持续提升，可以识别屏幕划痕、外壳瑕疵、摄像头模组缺陷和装配偏差；半导体领域在国产化进程加速的背景下，晶圆表面缺陷、封装缺陷和异物检测等环节的国产视觉 AI 空间同样值得关注；在汽车制造领域，可以用于零部件尺寸检测、焊接质量检测和装配一致性检测；在食品医药包装领域，可以用于标签识别、包装完整性检测、异物识别和批号 OCR 检测。

这一场景的关键不只是发现缺陷。而是视觉 AI 开始进入生产闭环。系统识别出缺陷后，可以联动产线控制系统，触发自动剔除、报警、返工、停机或工艺参数调整。由此，视觉 AI 从传统的事后质检工具升级为生产过程中的实时判断系统，参与“感知-判断-执行-反馈”的生产流程。

2026 年，视觉 AI 质检正向大模型与零样本检测方向发展。通过在海量工业图像上预训练，具备强泛化能力，仅需少量合格样本即可完成异常检测，无需标注缺陷数据。TVA 视觉智能体发展迅速，结合多模态感知、工业大模型认知和自主决策，实现小样本快速上线、自动适应光照和工件偏移，兼容多品类混线生产。边缘端本地化部署趋势明显，因为工厂数据安全要求高，视觉 AI 系统逐步转向私有化、离线部署，配合模型轻量化、量化和 Docker 容器化，实现快速部署和维护。

中国 AI 工业质检市场增长迅速，是工业 AI 中增速最快的细分领域。随着算法优化、深度学习和多模态检测技术突破，AI 视觉质检将进一步智能化、精准化，为企业带来更高生产效率和更低成本。

3.2. 安全生产监控：从人员监管到现场风险识别

下载日志已记录，仅供内部参考

安全生产监控是工业视觉 AI 的另一类重要应用。传统工厂安全管理主要在依赖制度约束、人工巡检和事后追责，但在大型厂区、能源基地、化工园区、矿山、仓储中心等场景中，人工巡查覆盖范围有限，风险识别效率较低，尤其是化工、钢铁、建筑工地等高危行业长期面临安全事故风险，传统录像的被动录像无法实现实时预警和主动干预，安全管理高度依赖人工巡检，存在覆盖盲区多、响应滞后等问题。视觉 AI 可以通过摄像头和视频分析系统，对生产现场人员、设备和环境进行实时监控，实现风险时间的提前识别和预警。

目前，视觉 AI 在安全生产的应用体现在边缘智能中枢协同多场景矩阵实现闭环决策。边缘智能分析聚焦于工业现场数据的实时处理与分析。通过视频流接入模块，接收来自生产一线的视频数据，并利用模型推理模块对这些数据进行实时分析，精准识别潜在的安全风险。为了确保预警的准确性，边缘智能还配备了相似去重和误报处理模块，有效过滤重复告警，降低误报率。此外，触发策略模块能够根据分析结果迅速启动相应的安全预警，而权重文件管理模块则负责维护模型推理所需的权重文件，保障分析的精确性。视觉分析模型库承担系统管理、资源协调与智能决策支持职能。通过远程升级模块实现对边缘节点功能的敏捷迭代与维护；系统配置与资源状态模块实时监控全局资源，动态优化系统性能与稳定性；告警数据合成模块对多源预警信息进行汇聚、去冗与结构化处理，形成标准工单与决策看板，真正实现“边缘感知-中心管控-闭环处置”的一体化协同，全面提升系统的整体管理能力和响应效率。

在具体应用中，视觉 AI 可以进行人员行为识别，主要识别员工是否佩戴安全帽、防护服和工牌，也可以检测人员是否进入危险区域，还能识别跌倒、聚集、违规操作等异常行为。在能源、电力、化工和矿山等高风险行业中，视觉 AI 能够将传统人工巡检升级为实时风险识别系统，减少安全事故发生概率。在环境异常检测领域，可以识别烟雾、明火、液体泄露等危险信号的早期识别。视觉 AI 也能对设备状态实施检测，通过图像分析识别设备异常运行状态，辅助预测性维护。

安全生产监控与传统安防不同，其核心不是简单监控人员，而是参与企业生产安全管理流程。视觉 AI 识别出的风险事件可以直接进入企业安全管理平台，并触发现场报警、调度处置和事后追踪。因此，视觉 AI 在安全场景中的价值体现在风险前置识别和安全管理闭环上。

目前，TSINGSEE 青犀视频平台提供基于边缘 AI 硬件与 EasyCVR 视频融合云服务平台的企业安全生产智能监管场景解决方案。方案可覆盖工地、采矿、石油化工、材料生产等行业常见的各类安全规范及行为监测功能，如穿戴规范、在岗状态、危险行为、周界异常、作业区域环境异常等。

政策层面，近年来国家安监部门持续出台工业安全数字化相关要求，部分高危行业已经将智能视频监控纳入强制合规范畴，政策催化效应明显。以矿山行业为例，应急管理部要求煤矿企业逐步推进井下智能监控系统建设，为视觉 AI 企业提供了确定性需求增量。

从商业模式来看，安全监控场景通常以软硬件一体化安防交付为主，单项目金额较大，同时叠加后续运维服务收入，客户粘性较强，收入质量优于纯硬件销售模式。基于 AI+云边端一体化管理，实现对企业生产经营的全面智能感知，解决传统手段下信息获取不及时、预警不到位、人力成本过高等痛点。基于 AI 智能识别，自动实时监测企业生产过程中的人、车、物、行为等，能准确检测识别出异常事件并及时预警，极大提升企业生产的风险防范水平，保障安全生产，实现安全生产少人、无人化、智能化监管。

未来 AI+安全生产不再是单一的文字、图片、视频或传感器分析，而是音频、视频、物联网数据、BIM/CAD 模型、业务数据的多模态融合。通过与数字孪生技术结合，在虚拟空间中 1:1 还原整个工厂、工地或场所，实现安全状态的实时映射、模拟和预测，真正做到全域感知、全局洞察。

3.3. 仓储物流：参与物品识别、分拣和调度

下载日志已记录，仅供内部参考

电商经济的高速发展驱动仓储物流行业需求持续扩张，传统人工作业模式的效率瓶颈日益凸显。一方面，订单量爆发式增长导致人工分拣错发、漏发问题频出，退货质检周期长、时效慢，直接影响客户体验与运营成本；另一方面，医药、美妆、食品等对温湿度敏感的品类占比持续提升，人工管理难以实现全程精细化监控，货损风险较高。

据艾瑞咨询《2025 年中国仓储物流行业发展报告》显示，中国仓储物流市场规模已突破 4 万亿元，年复合增长率保持在 8.5%以上，其中智能仓储、恒温仓储等细分领域增速超过 15%。随着电商、食品、美妆等行业的快速发展，企业对仓储物流的需求从“基本存储”转向“精准、智能、一体化”，空间利用需更高效、温湿度控制需更精准、信息传递需更透明、逆向物流需更便捷，传统仓储模式的局限性愈发凸显，行业正迎来从劳动密集型向技术密集型的关键转型期。而且在当今全球供应链高度复杂、客户对交付速度要求日益提高的背景下，传统仓储管理方式已难以满足企业高效运营的需求。

工业生产不仅包括生产线加工，也包括原材料入库、半成品转运、成品出库和物流配送等环节。仓储物流场景中，视觉 AI 主要参与物品识别、包裹识别、条码识别、智能分拣、库存盘点和仓储机器人导航等流程。视觉 AI 通过构建智能仓储管理系统，融合计算机视觉和 AIoT 技术，从多个维度重构仓储作业流程。视觉仓库管理系统（Visual Warehouse Management System, VWMS）应运而生，它融合了计算机视觉、AI、物联网和大数据分析等先进技术，为仓储作业提供了前所未有的自动化、可视化和智能化能力。视觉仓库管理系统是一种基于摄像头、传感器和 AI 算法的数字化平台，能够实时捕捉、分析和处理仓库内的人、货、场信息。与传统的条码扫描或 RFID 系统相比，VWMS 不再依赖人工扫码或标签识别，而是通过图像识别技术自动感知货物状态、位置、数量和移动轨迹，实现对整个仓储流程的“无感”监控与智能决策。

视觉 AI 仓管系统可以自动采集货架图像，进行商品类别、规格及数量识别，实时更新库存数据，替代传统人工盘点模式，大幅降低人为误差，提升库存数据准确率。视觉 AI 仓管系统通过热力图分析、路径追踪等功能，视觉识别分析货架空间占用状态，自动生成最优货位分配方案，避免空间浪费，提升仓库整体坪效，帮助管理者科学规划货位布局。视觉 AI 还能帮助提供全流程可视化追踪，确保操作规范，通过对仓库内人员行为和设备运行进行实时监测，识别人员是否佩戴安全帽、是否进入禁区、是否走错通道，同时识别叉车、AGV 小车是否按指定路线行驶，有无碰撞风险，一旦发现违规操作（如未按标准流程上架），立即触发警报并记录日志，实现作业合规性的智能化管理。此外，视觉 AI 还能结合红外摄像头和 AI 分析技术，提前识别潜在火源及异常热点，同时实现陌生人入侵实时预警，提升仓储安全防护水平。

视觉 AI 在仓储物流领域的商业模式以软硬件一体化方案交付为主，通常包含前期硬件设备（工业相机、边缘计算设备）、系统集成部署，以及后续软件订阅和运维服务三部分收入。相较于纯硬件销售，SaaS 化运维服务收入占比逐步提升，客户粘性强，收入可预期性好。

视觉 AI 在仓储领域的商业化逻辑主要是替代逻辑清晰、从单点工具到生产要素的演进和政策与资本双轮驱动。替代逻辑清晰主要体现在 ROI 可量化，仓储企业引入视觉 AI 系统后，可直接测算人力替代数量、错误率下降幅度、库存准确率提升等指标，采购决策逻辑明确，客户接受度高于其他 AI 应用场景。在从单点工具到生产要素的演进方面来看，早期视觉 AI 以解决单一痛点为主（如条码识别、缺陷检测），当前头部企业已逐步将视觉 AI 嵌入仓储全流程，与 WMS、ERP 深度集成，成为仓储运营不可或缺的基础设施，系统替换成本极高，形成较强护城河。政策与资本双轮驱动方面，国家“十四五”期间持续推进制造业与物流业数字化转型，叠加头部电商平台（京东、菜鸟等）大规模自建智能仓储，带动整个行业的智能化改造需求加速释放。

从渗透率角度看，目前国内智能仓储视觉 AI 整体渗透率仍处于早期阶段，中小仓储企业改造需求尚未大规模启动，随着系统成本持续下降和标准化程度提升，下沉市场的增量空间不容忽视。

下载日志已记录，仅供内部参考

3.4. 智慧工厂监控：从单点检测走向生产过程可视化

智慧工厂是视觉 AI 在工业领域应用的最高形态，与质检、安全监控等单点应用不同，智慧工厂场景下视觉 AI 承担的角色已从局部工具升级为生产要素，核心价值在于将生产现场大量非结构化的视觉信息转化为可分析、可追踪、可优化的数据资源，使视觉 AI 真正参与工艺优化与生产决策。过去的视觉 AI 更多应用于某一个检测点，例如检测某个产品表面是否存在缺陷；而智慧工厂场景要求视觉 AI 覆盖更完整的生产流程，对产线运行状态、设备状态、人员行为、产品质量和异常行为事件进行综合感知。

智慧工厂视觉 AI 的底层逻辑是将生产现场的非结构化视觉信息系统性转化为可用数据。整体流程涵盖数据采集（工业摄像机、监控设备）、数据存储与预处理、特征提取、分析建模及可视化呈现等环节，最终将原本碎片化的现场画面转化为结构化的生产数据资产，为后续决策优化提供基础支撑。在智慧工厂中，视觉 AI 可以通过多点位摄像头、工业相机、边缘计算设备和生产管理系统协同工作，对生产现场进行持续监测。例如，系统可以识别产线是否拥堵、设备是否异常、工人操作是否规范、物料是否堆放错误、产品是否出现批量缺陷。视觉 AI 产生的图像数据和缺陷数据还可以与 MES、ERP、质量管理系统打通，形成生产质量追溯和良率分析体系。

视觉 AI 帮助设备运行状态在线监测与智能预警。以冲压车间为例，视觉 AI 对冲压设备运行状态进行持续分析与判断，实现设备健康状态的在线监测、异常智能预警及故障原因自动诊断，并根据分析结果联动生成维修工单或保养计划，将设备管理从被动响应转变为主动预防，有效降低非计划停机损失。

视觉 AI 还能应用于工艺质量闭环管控。依托工业互联网平台，视觉 AI 可实现生产过程数据与质量检测的深度融合。以汽车焊接场景为例，系统实时采集焊接过程数据与车辆属性信息，结合焊接质量检测模型，对全数焊点实现 100%在线检查，自动判断焊接正常或异常状态，并为每台车辆建立可追溯的焊接质量档案，实现质量管控从抽检到全检、从事后追溯到过程控制的根本性转变。

视觉 AI 通过通过对生产现场设备数据的持续采集与建模分析，系统可实时准确反映各类生产加工状态，形成可视化看板，指导业务分析与管理改进，管理层对产线运行状况的掌握从依赖人工汇报转变为数据实时驱动，实现生产状态实时可视化与管理优化。

数字孪生的发展也将放大视觉 AI 的作用，通过三维建模技术，将整条产线完整映射至虚拟的数字系统中。而视觉 AI 正是获取真实工厂数据的重要入口，通过摄像头、工业相机等硬件设备采集每台设备的运行状态、每个工位的作业进度、每件产品的流转路径数据，并将其传输到数字孪生系统中，实现工厂运行状态的实时可视化。数字孪生将生产优化从“拍脑袋”变为科学决策。系统持续采集数千个数据点，利用 AI 算法进行深度分析，发现人眼难以察觉的优化空间。

数字孪生赋能的智慧工厂能帮助实现能源精准管控。数字孪生通过视觉 AI 采集能源消耗和生产成本，实时监测原材料价格波动、能源成本变化、设备折旧、人工费用等成本要素，立即反映在数字孪生系统的成本看板上。当某批次产品成本显著高于平均水平时，系统自动预警并推送给相关人员分析原因。

数字孪生正在重新定义生产线管理的范式，依靠视觉 AI 技术，将工业生产从不可见到透明化，从经验驱动到数据驱动，从被动响应到主动优化。市场竞争日益激烈，客户需求快速变化，只有拥抱数字孪生等先进技术，才能在智能制造时代保持竞争力。据中国信息通信研究院数据，2025 年中国数字孪生工业应用市场规模预计将达到 480 亿元。IDC 预测，到 2026 年全球数字孪生市场规模将达 482 亿美元，中国市场份额将占 28%。未来的工厂将是物理工厂与数字工厂的深度融合，生产线将具备自感知、自学习、自优化、自适应的能力。

视觉 AI 正在从单点算法工具升级工业数据入口。它不仅识别图像，还将生产现场的非结构化视觉信息转化为可分析、可追踪、可优化的数据资产。随着工业大模型、

下载日志已记录，仅供内部参考

数字孪生和生产管理系统进一步结合，视觉 AI 从看见生产市场进而深入理解生产过程，并进一步参与工艺优化和生产决策。

智慧工厂是视觉 AI 商业价值密度最高的场景，单项目交付金额远高于单点应用，通常涵盖硬件、系统集成、数据平台及长期运维服务，客户一旦完成部署替换成本极高，形成强绑定关系。当前国内智慧工厂建设仍处于早期阶段，大型制造企业率先导入，中小企业改造需求尚未大规模释放，随着系统成本持续下降与标准化程度提升，市场扩容空间显著。政策层面，工信部持续推进"智能制造示范工厂"建设，为头部视觉 AI 企业提供了明确的需求催化。

3.5. 自动驾驶：驱动交通出行的智能化变革

随着 AI、物联网、5G 移动通信等新一代数字技术的发展与成熟，自动驾驶技术得到了显著提升。汽车企业不断推进整车的自动化和智能化水平，市场对自动驾驶的接受程度也在不断提高。在当今汽车行业，自动驾驶技术无疑是最具变革性的发展方向之一，从最初简单的辅助驾驶功能，到如今越来越接近实现的完全自动驾驶，这一领域的技术迭代日新月异。

在众多关键技术中，环境感知方案的选择成为了各大车企和科技公司争论的焦点，其中以摄像头为核心的纯视觉方案与依赖激光雷达的方案之间的博弈尤为激烈。早在 2021 年，特斯拉就彻底放弃了“摄像头+雷达”的组合，全面转向纯视觉路线，并将其命名为“特斯拉视觉（Tesla Vision）”，特斯拉旗下所有车型均仅依靠摄像头实现环境感知。在马斯克看来，激光雷达和雷达会因“传感器冲突”降低安全性。当激光雷达、雷达与摄像头的感知结果不一致时，系统在决策过程中就会面临困惑，究竟该以哪个传感器的数据为准。这种传感器数据模糊的问题，不但无法降低风险，反而会增加事故发生的可能性。马斯克还特别指出，这正是 Waymo 的车无法在高速公路上自如行驶的原因之一。而特斯拉关闭车型上的雷达，目的就是为了提升安全性，让系统决策更加聚焦于摄像头所提供的信息。特斯拉所推崇的摄像头方案是通过在车辆上集成多个摄像头，模拟人类视觉系统来感知周围环境。这些摄像头从不同角度采集图像信息，然后由车载电脑运用强大的算法模型对这些图像进行分析和处理，从而识别出道路、车辆、行人以及交通标志等各种目标物体。随着计算机视觉技术和深度学习算法的飞速发展，摄像头方案的感知能力得到了极大提升。特斯拉的神经网络系统可以不断学习和进化，通过大量的实际道路数据训练，不断优化对各种复杂场景的识别和应对能力。摄像头方案的一大显著优势在于成本。相比动辄数千美元甚至更高成本的激光雷达，摄像头的成本要低得多，这使得整车的制造成本能够得到有效控制。摄像头方案在运算效率方面也具有一定优势，由于减少了对复杂雷达数据的处理，系统可以更快速地对摄像头采集的图像数据进行分析和决策，提升了运算效率。目前，在自动驾驶技术的应用上，虽然市场呈现出多种方案并存的局面，但是特斯拉作为自动驾驶技术的龙头企业，凭借在纯视觉方案上的持续投入和技术积累，在辅助驾驶领域占据了重要地位，其 FSD（全自动驾驶）功能不断升级，为用户带来了较为先进的驾驶体验。因此，摄像头所承载的视觉 AI 成为自动驾驶技术落地的重要环节。

环境感知是视觉 AI 在自动驾驶中的基础功能。通过分析道路标志、交通信号灯、行人和车辆等信息，车辆能够实时了解周围环境，为安全驾驶提供支持。基于深度学习的目标检测算法（如 YOLO 和 Faster R-CNN）使得车辆能够精准识别和跟踪目标，为路径规划和行为预测提供数据。车道线检测是实现车辆稳定行驶的关键步骤。通过摄像头捕获的道路图像，视觉 AI 技术可以准确识别车道边界，确保车辆保持在正确的车道内，提高驾驶的安全性和舒适性。此外，视觉 AI 还支持高精度地图构建，通过特征提取和匹配，增强车辆的定位和导航能力。

在自动驾驶的复杂动态环境中，精确检测各类目标，包括车辆、行人、自行车以及各种交通标识，是保障行车安全的重要任务。先进的深度学习目标检测算法，如基于区域的卷积神经网络（R-CNN）系列，特别是 Faster R-CNN 和 Mask R-CNN，通过引入区域建议网（RPN）和特征金字塔网络（FPN）等创新架构，能够在不同尺度

下载日志已记录，仅供内部参考

和视角下高效、准确地识别目标，为自动驾驶系统提供关键的先验信息，显著提升了对潜在危险的预警能力。
视觉 AI 也能帮助自动驾驶构建高精度地图。通过从车载摄像头获取的图像数据进行特征提取和匹配，可以实现对道路、交通标志等地图元素的识别和定位。这种高精度地图能够提供给车辆更准确的位置信息，增强自动驾驶系统的定位能力和导航精度。
视觉 AI 技术也可以对驾驶场景进行分析和理解。通过车载摄像头或者多个传感器的数据融合，可以对交通流程、道路拓扑结构、车辆行为等进行分析和建模。这有助于自动驾驶系统更好地理解周围环境，做出更加准确和安全的决策。

自动驾驶是视觉 AI 从识别工具走向系统参与的重要应用场景。随着智能驾驶功能从 L2 辅助驾驶向 L3 有条件自动驾驶推进，视觉 AI 在车辆系统中的作用也逐步加深。
国际汽车工程师学会（SAE）制定的 0-5 级标准是全球公认的权威框架，它的核心评判依据是在执行动态驾驶任务（DDT）中，“车辆系统”与“人类驾驶员”之间的角色与责任分配。中国采用 GB/T 40429-2021《汽车驾驶自动化分级》国家标准，与国际通用的 SAE J3016 高度一致，L0-L2 为辅助驾驶，L3-L5 为自动驾驶。2025 年底至 2026 年初，中国在 L2 普及、L3 试点破冰、L4 Robotaxi 商业化规模化方面全球领先，L5 仍处于研发阶段。政策端与产业链共同驱动，L2/L2+渗透率已超 60%，L3/L4 进入商业验证新阶段。
L2 级别（组合辅助驾驶）核心特征是同时控制速度与方向。目前中国 L2 已大规模普及并演化出城市 NOA，渗透率全球最高，2025 年前三季度新车 L2 渗透率达 64%，年底或升至 66%；城市 NOA 渗透 15%+，已下沉至低价车型。Momenta+华为占第三方供应商 80%。通过 OTA 持续迭代，体验接近 L3 但驾驶员仍需监控。已成为主流卖点，2026 年将继续渗透。典型的 L2 级别代表有比亚迪天神之眼（DiPilot）、华为 ADS（问界、阿维塔、极狐等）+ Momenta（第三方主导，中端车广覆盖）和小鹏 XNGP、理想 NOA、蔚来 NOP、小米 SU7 等。
L3 级别（有条件自动化）是特定条件下系统完成全部驾驶任务。：2025 年 12 月工信部首批产品准入+专用号牌落地，标志从测试迈入有限商业试点，是全球最早进入 L3 商用试点的国家，而国际 L3 多停留在测试阶段。中国的 L3 限于特定路段或条件，B 端运营为主，C 端预计 2028 年后逐步开放。责任主体转为车企。2026 年将成为规模商用关键年，更多车型和城市跟进，高速场景先行。典型的 L3 代表有长安深蓝 SL03、北汽极狐阿尔法 S 等。
L4 级别（高度自动驾驶）是由系统完全负责的自动驾驶。中国规模化商业化领先全球。多城全无人付费运营，车队数百至千辆级，成本已大幅下降。部分城市和路线单车盈利，2026 年将扩张至数千辆+国际布局。港口和矿区等封闭场景已成熟，干线物流 2026-2030 预计可以实现规模化使用。典型代表的 L4 级别有百度 Apollo Go（萝卜快跑）、小马智行（Pony.ai）、文远知行（WeRide）等。

相比乘用车开放道路场景，矿山、港口等封闭场景是自动驾驶更快形成商业化闭环的方向。因为这种封闭场景作业路线相对固定，交通参与者较少，环境复杂度低于开放道路。而且客户通常是矿山、港口、能源、物流企业，付费主体明确。场景中存在高危、高强度、招工难等痛点，无人化替代需求更强。车辆运行效率、安全事故率、人力成本和能耗节约均可量化，因此投资回报更清晰。
在这些场景中，视觉 AI 不只是车辆的感知模块，而是生产运输系统的一部分。它通过摄像头、激光雷达、毫米波雷达等多传感器融合，参与识别道路、障碍物、人员、车辆和作业边界，并与调度系统、车队管理系统、装卸设备协同，支撑无人车辆完成运输任务。由此，视觉 AI 从“辅助驾驶体验”转向“生产效率工具”，成为产业运行中的感知型生产要素。

矿山是无人驾驶商业化落地较快的场景之一。露天矿区通常环境恶劣、作业强度高、运输距离长，传统人工驾驶面临安全风险、招工难和运营成本高等问题。矿卡运输

下载日志已记录，仅供内部参考

又是矿山生产中的核心环节，直接关系到采装、运输、排土和生产效率。因此，无人驾驶矿卡一旦能够稳定运行，就不只是替代司机，而是直接参与矿山运输生产流程。在矿山无人驾驶中，视觉 AI 与激光雷达、毫米波雷达、定位系统和调度平台共同构成无人运输系统。车辆需要在装载区、运输道路、会车区和卸载区之间连续运行，视觉 AI 参与道路边界识别、障碍物检测、人员车辆识别和复杂作业场景判断。感知结果进一步输入车辆控制系统和矿山调度平台，使无人矿卡能够按照生产计划完成运输任务。由此，视觉 AI 嵌入的是矿山“采装-运输-卸载-调度”的生产流程，而非单点辅助功能。从落地进展看，矿山无人驾驶已经开始从单车试验进入车队规模化运营阶段。2025 年 5 月，华能伊敏露天矿投运百台级 5G-A 无人驾驶纯电矿卡车队，标志着矿山无人驾驶进入更大规模编组运行阶段。与此同时，国内头部方案商易控智驾截至 2025 年底在运行无人矿卡规模已超过 2300 台，覆盖煤矿、金属矿、非金属矿等多类矿种。矿山无人驾驶已经不再只是技术试验，而是开始参与能源和矿产生产体系中的运输环节。矿山无人驾驶商业化之所以快于开放道路自动驾驶，核心在于 ROI 更容易测算。矿山运输通常具有路线固定、作业重复、车辆数量大、连续作业要求高等特点，当无人矿卡规模达到数十台甚至百台级别后，系统可以通过减少人工驾驶、提升车辆利用率、降低安全风险和优化调度效率形成经济效益。因此，矿山无人驾驶体现了视觉 AI 由技术能力向生产要素转化的典型路径。

在港口场景中，无人集卡并非孤立运行，而是需要与岸桥、场桥、堆场管理系统和码头调度系统协同。视觉 AI 和多传感器感知系统负责识别道路、人员、车辆、集装箱和作业区域边界，调度平台则根据船舶装卸节奏安排车辆路径。由此，视觉 AI 参与的是港口“装卸-水平运输-堆场调度”的作业链条。港口无人驾驶也已进入商业化复制阶段。以天津港为例，其 5G+智慧港口项目通过 5G 网络、高精地图和云端自动驾驶系统，协调岸桥、场桥和无人驾驶平板车等设备，实现码头水平运输智能化。国内港口智能化改造也在加速推进，2025 年我国已投运自动化集装箱码头和自动化干散货码头各 30 座，天津港集装箱与干散货码头实现智能化改造全覆盖。与此同时，港口无人集卡开始出现百台级车队订单，说明该场景正从示范运营走向规模化交付。 港口无人驾驶的商业化逻辑在于，它不是单纯替代一辆车的司机，而是嵌入码头整体自动化系统。无人集卡与岸桥、堆场和调度系统协同后，可以提升水平运输效率，减少人工调度误差，并推动港口从“设备自动化”向“全流程智能化”升级。整体来看，自动驾驶领域的视觉 AI 落地呈现出“乘用车规模化装车、封闭场景率先商业化”的特征。乘用车 L2/L2+推动视觉 AI 进入量产车体系，但其核心仍是消费端智能化配置；矿山、港口等场景则不同，视觉 AI 参与的是矿石运输、集装箱转运、车辆调度和安全管理等生产流程，能够直接影响企业效率、成本和安全。因此，封闭场景中的无人驾驶更能体现视觉 AI 从技术能力向生产要素转化的逻辑。未来，随着车端算力、多传感器融合、5G 专网和调度系统进一步完善，矿山、港口、园区物流等场景有望成为自动驾驶商业化落地的先行方向。

3.6. 机器人：赋能机器感知与自主决策

机器人视觉是通过视觉 AI 技术赋予机器人“视觉能力”，使机器人能够感知、理解并与周围环境交互。传统工业机器人本质上是在预设程序和固定环境下重复动作，一旦外部环境发生变化，如工件位置偏移/光线变化或出现未知障碍物，系统便会报错停机。这种刚性作业模式决定了传统机器人只能适用于高度标准化的生产场景，难以应对真实工业环境中的复杂性与不确定性。
视觉 AI 的引入从根本上改变了这一局面。通过工业相机与深度学习算法的结合，机器人获得了对三维空间的实时感知能力，能够识别物体类别、判断位置姿态、理解场景语义，从而在非结构化环境中做出自主响应。视觉 AI 不仅是机器人的“眼睛”，更是其理解世界、做出决策的核心输入来源，是机器人从“自动化工具”迈向“自主智能体”的关键技术支撑。视觉 AI 赋予机器人感知能力、决策能力和执行能力。通过传感器和摄像头等设备，视觉 AI 赋予机器人感知环境的能力，使其能够获取和处理外界信息。例如，通过计算机视觉技术，机器人可以识别图像中的物体和

下载日志已记录，仅供内部参考

场景。视觉 AI 通过机器学习算法，帮助机器人进行决策和规划，提高其自主性和智能化水平。例如，机器人可以通过强化学习算法，优化路径规划和任务执行策略。机器人通过机械臂、移动底盘等硬件设备，执行各种任务和操作，实现与物理世界的互动。例如，机器人可以完成搬运、组装、巡逻等任务。

工业机械臂：视觉引导下的精准作业

工业机械臂是视觉 AI 落地最成熟的机器人场景。传统机械臂依赖固定夹具和预编程轨迹完成抓取、装配、焊接等动作，对工件位置的容忍误差极小。引入视觉 AI 后，机械臂可通过 3D 视觉系统实时识别工件的位置与姿态，动态调整运动轨迹，实现对无序堆放工件的随机抓取（Bin Picking），大幅降低对夹具和定位工装的依赖，显著提升产线柔性。

视觉系统在工业机械臂中承担的核心感知任务可拆解为三个层次，一是定位感知，通过结构光或 ToF（飞行时间）技术获取工件的三维点云数据，精确计算工件在空间中的位置坐标与姿态角度，为机械臂运动规划提供输入；二是状态判断，在抓取或装配过程中持续监测工件与末端执行器的相对位置，实时纠偏，确保动作精度；三是结果验证，动作完成后由视觉系统对装配质量或抓取结果进行检验，形成完整的感知-执行-反馈闭环。

在焊接场景中，视觉引导系统可实时检测焊缝位置与坡口形状，识别因工件变形或装夹误差导致的焊缝偏移，引导焊枪自动纠偏，解决传统示教编程静态轨迹无法适应动态误差的根本难题。同时视觉系统还可对焊后熔池形态和焊道表面质量进行在线检测，将质量控制嵌入焊接过程本身。在 3C 电子装配场景中，视觉 AI 通过亚像素级图像处理算法支撑微米级精度的零部件对位，满足柔性线路板、摄像头模组等高精密元件的装配需求。

人形机器人：具身智能的感知基础

人形机器人是视觉 AI 与机器人技术融合的前沿方向。与工业机械臂的固定底座不同，人形机器人需要在复杂、动态的真实环境中自主移动与作业，对视觉感知能力的要求大幅提升。

视觉 AI 需要为人形机器人提供感知能力。通过多目摄像头和深度传感器实时构建周围环境的三维地图（SLAM），使机器人知道自己在哪、周围有什么，这让机器人对环境有理解；视觉 AI 还能给机器人提供目标识别与定位，对操作对象的类别、位置、形状进行识别，为抓取和操作提供精确坐标；视觉 AI 也为机器人提供动态障碍物感知，实时追踪移动中的人员和物体轨迹，预测运动趋势，规划安全路径；机器人通过视觉 AI 对人机交互进行理解，通过视觉识别人体姿态、手势和表情，理解人的意图并作出响应。

这四个层次的感知能力相互协同，构成人形机器人在非结构化环境中自主作业的基础。以特斯拉 Optimus 为例，其视觉系统直接复用 FSD 自动驾驶技术积累，通过多目摄像头实现对周围环境的 360 度感知，结合神经网络模型完成对复杂场景的实时理解。国内宇树科技、傅利叶智能等企业亦在人形机器人视觉感知方向持续投入。值得关注的是，当前具身智能的核心突破方向是将视觉 AI 与大语言模型结合，使机器人能够理解自然语言指令并将其转化为具体操作动作，从固定程序执行迈向任务自主理解，具身智能因此成为当前机器人领域最受资本关注的方向之一。

协作机器人：与人共融的安全感知

协作机器人（Cobot）设计用于与人在同一空间协同作业，安全性是首要前提，这对视觉 AI 的实时感知能力提出了极高要求。

视觉系统在协作机器人中承担三项核心功能。一是人员安全监测，通过人体姿态识别与距离检测，构建以机器人为中心的动态安全区域，当人员进入预警范围时自动降速，进入危险区域时立即停机，实现比传统力矩传感器更快速、更主动的安全响应；二是意图预判，通过对人体运动轨迹的持续分析，预测操作员的下一步动作，提前调整机器人运动规划，避免被动响应带来的碰撞风险；三是视觉示教，操作员

下载日志已记录，仅供内部参考

无需编程，直接通过手势或肢体动作演示操作流程，视觉系统记录并学习动作轨迹，自动生成可执行程序，将机器人部署时间从数天压缩至数小时。

视觉示教能力的成熟是协作机器人渗透中小制造企业的关键驱动力。传统工业机器人编程门槛高，中小企业缺乏专业技术人员，导致自动化改造意愿低。视觉 AI 降低了使用门槛，使协作机器人真正具备在中小企业大规模普及的条件，市场空间有望随之显著扩大。

柔性制造：制造业未来重要变革方向

机器人是视觉 AI 参与工业生产的重要载体。传统工业机器人主要依赖固定程序和固定轨迹，适合标准化、大批量、重复性较高的生产任务。但随着制造业向多品种、小批量、快速交付方向发展，企业对生产线柔性化要求不断提升，传统固定式自动化设备难以适应频繁换型、复杂工件识别和非标准化作业环境。因此，柔性制造成为工业机器人发展的重要方向。

柔性制造强调生产系统能够根据产品类型、订单需求和工艺变化快速调整生产方式。在这一过程中，视觉 AI 为机器人提供了关键感知能力，使机器人能够从单纯执行预设动作，升级为能够识别环境、判断目标状态并调整作业路径的智能生产单元。具体来看，视觉 AI 可以帮助机器人识别工件类别、判断工件位置和姿态、引导机械臂完成抓取、装配、焊接和搬运，并在作业完成后进行质量检测和结果反馈。

视觉 AI 赋能机器人的核心技术路径是构建“感知-判断-执行-反馈”的闭环。在散乱零部件抓取场景（Bin Picking）中，3D 视觉系统可识别无序堆放工件的空间位置与姿态，引导机械臂完成精准抓取，替代原有依赖固定夹具和人工摆放的作业方式，显著降低产线准备成本。在装配场景中，视觉 AI 实时检测零部件装配偏差，辅助机器人动态调整路径，使装配精度达到微米级，满足 3C 电子、半导体封装等高精密制造需求。在焊接场景中，视觉系统实时判断焊点位置与表面质量，将检测结果直接反馈至控制系统，实现从事后抽检到过程实时控制的转变。

在具体行业落地方面，汽车制造是视觉 AI 与机器人结合最深入的领域。整车厂冲压、焊接、涂装、总装四大工艺均已引入视觉引导机器人。新能源电池领域，电芯装配、极耳焊接、模组组装等高精度环节对视觉引导机器人需求快速增长，随国内动力电池产能持续扩张，该领域成为视觉 AI 机器人应用增速最快的细分市场之一。3C 电子领域，手机、平板等消费电子产品迭代周期短、换型频繁，视觉 AI 赋予机器人快速适应新产品的能力，换线调试时间大幅压缩。

具身智能是视觉 AI 与机器人融合的前沿演进方向。以特斯拉 Optimus 为代表的人形机器人，直接复用自动驾驶视觉感知技术积累，通过多目摄像头构建对复杂环境的三维理解能力。国内宇树科技、傅利叶智能等企业亦在人形机器人视觉感知方向持续投入。具身智能的核心突破在于让机器人不仅能“看见”，更能“理解”，通过视觉 AI 与大模型的结合，机器人可以理解自然语言指令并将其转化为具体操作动作，真正实现从固定程序执行到任务自主理解的跨越。

从产业化角度看，视觉 AI 与机器人结合的商业价值主要体现在三个层面，一是提升作业精度，使机器人能够完成更复杂的抓取、装配和检测任务；二是提升产线柔性，减少换型调试成本，一条具备视觉 AI 能力的产线换型时间可从数小时压缩至数十分钟；三是推动生产数据闭环，视觉识别结果实时回传至 MES、ERP 等生产管理系统，为工厂整体运营优化提供数据支撑。

未来，随着 3D 视觉、多模态感知与具身智能技术持续演进，视觉 AI 有望推动机器人从固定自动化设备全面升级为具备环境理解与任务自主执行能力的柔性生产单元，成为制造业智能化转型的核心基础设施。

4. 视觉 AI 潜在空间分析

4.1. 汽车：视觉 AI 高强度应用的代表性生产场景

汽车制造长期以来被认为是对“标准化生产”和“柔性制造”要求最高的产业之一。近年来，随着电动化、智能化趋势的加速，汽车零部件复杂度和品类数量持续攀升，

下载日志已记录，仅供内部参考

叠加多车型共线、小批量多变型的产线挑战，使得传统质检模式面临巨大压力。从动力电池、电机控制器，到各类传感器和轻量化结构件，检测需求不断增长；消费者对整车外观、间隙面差、涂装质量、选配一致性愈发敏感；人工检测效率低、稳定性差、主观性强，难以支撑当下高节拍、快节奏的产线环境。

随着深度学习、图像处理、边缘计算等技术的成熟，视觉 AI 技术正在快速落地于汽车行业各大工艺车间。从冲压件裂纹检测、焊点质量识别，到整车出厂前的外观智能检验，AI 视觉正在逐步取代人工或规则算法视觉系统，解决复杂环境下的高鲁棒性、强泛化检测需求。随着汽车行业对质量控制的重视不断提升，越来越多主机厂和一级供应商正将 AI 视觉检测系统部署到关键生产工位，特别是在智能电动汽车、出口车型和高端制造基地中，机器视觉的普及速度显著加快。在这些环节中，AI 视觉技术以更高的检测精度、更强的适应能力替代人工抽检，实现从“抽检”到“全检”、从“规则检测”到“智能识别”的跨越。

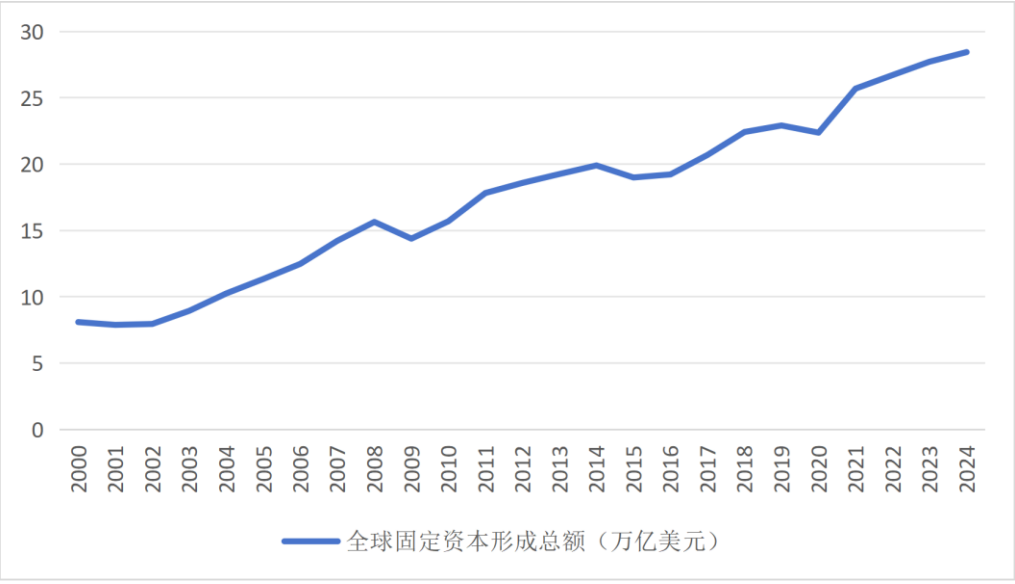
4.2. 全球固定资本形成提供较大的潜在空间

由于视觉 AI 通常以自动化检测设备、工业软件及系统集成等形式嵌入企业生产设施，所以视觉 AI 在经济核算中主要表现为固定资本形成，因此引入全球固定资本形成总额（Gross Fixed Capital Formation，GFCF）。

固定资本形成总额指的是一定时期内各经济主体取得的固定资产扣除固定资产处置后的净增加额，涵盖机器设备、厂房建筑、基础设施及知识产权产品等使用期限较长的生产性资产。与工业设备市场销售额相比，固定资本形成总额更能反映企业及其他经济主体当期实际发生的资本性投入，更符合本报告从代表性项目投资结构向宏观角度投资空间外推的逻辑测算。

世界银行数据显示，全球固定资本形成总额最新披露至 2024 年。2024 年，全球固定资本形成总额达到 28.44 万亿美元。即使不考虑未来全球固定资产投资规模的进一步增长，仅基于 2024 年全球固定资本形成规模，视觉 AI 相关投资所对应的潜在空间已较为可观。这说明视觉 AI 在工业生产领域的增长空间不仅来自固定资产投资总量的扩大，更重要的事来自现有生产设备更新、自动化改造及视觉检测渗透率的提升。

图 8：全球固定资本形成总额



数据来源：世界银行、东北证券

5. 视觉 AI 未来趋势：从识别工具走向生产要素

下载日志已记录，仅供内部参考

5.1. 技术演进：从“看见”到“理解”再到“决策”

视觉 AI 的技术演进路径，核心是感知能力、理解能力和决策能力的持续深化。早期视觉 AI 以规则驱动和人工特征为主，更多依赖模板匹配、边缘检测、阈值分割等方式，只能在固定场景下完成较为简单的识别任务。这一阶段的系统稳定性较强，但对光照变化、产品角度变化、复杂缺陷和非标准化场景的适应能力有限，因此主要用于标准化程度较高的检测场景。

深度学习普及后，视觉 AI 从规则驱动转向数据驱动，识别精度和泛化能力明显提升。以 CNN、目标检测、图像分割等技术为代表的深度学习模型，使视觉 AI 能够识别划痕、裂纹、脏污、焊点异常、装配偏差等复杂缺陷，并逐步进入工业质检、安防监控、交通管理、医疗影像等场景。在工业生产中，深度学习推动视觉 AI 从“能不能看见缺陷”升级为“能不能稳定替代人工质检”，这是视觉 AI 实现产业化落地的重要阶段。

当前阶段，视觉 AI 正在与大模型、多模态模型和行业数据体系结合，推动机器从“看见”进一步升级为“理解”。传统视觉 AI 主要回答“图像中有什么”“缺陷在哪里”“目标是否存在”等问题，而视觉大模型和多模态模型则进一步强调对场景语义、上下文关系和事件逻辑的理解。在工业场景中，这意味着视觉 AI 不再只是识别某个产品是否存在缺陷，而是可以结合工艺参数、设备状态、历史缺陷数据和生产批次信息，判断缺陷是否具有共性、是否来自某一设备异常、是否需要调整工艺参数。

多模态感知融合是视觉 AI 从“识别工具”走向“生产决策系统”的关键方向。单一视觉数据主要反映产品外观和现场图像信息，但真实工业生产环境是复杂的、多变量的。一个产品缺陷的产生，往往不仅与表面图像有关，还可能与温度、湿度、压力、振动、声音、设备状态、工艺参数、原材料批次等因素相关。因此，未来视觉 AI 需要与多类传感器数据和生产系统数据融合，构建对生产环境更加完整的感知体系。

具体来看，多模态工业视觉可以分为三个层次。第一是视觉模态内部的融合，例如 2D 视觉、3D 视觉、红外热成像、线扫相机、面阵相机等不同视觉信息的融合。2D 视觉适合外观检测和字符识别，3D 视觉能够获取深度、尺寸和空间位置，热成像则可以识别温度异常。多类视觉信息结合后，系统能够更准确地判断产品状态和设备状态。第二是视觉与非视觉传感器融合，例如将图像数据与温度、振动、声音、电流、电压、压力等数据结合，用于判断设备运行状态和异常原因。第三是视觉数据与业务数据融合，例如将缺陷图像与 MES 系统、ERP 系统、质量管理系统、生产批次、工艺参数和历史维修记录打通，使视觉 AI 能够参与质量追溯、良率分析和生产优化。

多模态融合的价值在于，它让视觉 AI 从“看到结果”走向“理解原因”。例如，在锂电池生产中，视觉 AI 可以识别极片涂布缺陷，但如果进一步结合涂布速度、烘干温度、设备压力和历史缺陷分布，就可以分析缺陷是否与某一工艺参数波动有关；在汽车制造中，视觉 AI 可以识别焊缝异常，但结合机器人运动轨迹、焊接电流、温度和工位数据后，可以进一步判断异常是否由设备偏移、焊接参数变化或工件定位误差导致；在设备巡检中，视觉 AI 可以发现设备外观异常，而结合声音、振动和温度数据后，可以更早识别设备磨损、过热或故障风险。

因此，多模态视觉 AI 的核心意义不是简单增加数据来源，而是提升系统对生产过程的综合理解能力。过去，视觉 AI 更多输出“是否异常”的判断。未来，多模态视觉 AI 将进一步输出为什么异常、异常是否会扩大、应该如何处理的分析结果。由此，视觉 AI 开始从感知层进入认知层和决策层，成为生产管理和工艺优化的重要技术基础。

与此同时，3D 视觉的普及也将推动视觉 AI 从平面识别向空间理解升级。传统 2D 视觉主要依赖平面图像，适合表面缺陷检测、标签识别和外观检查。而工业生产和机器人作业往往需要识别物体的空间位置、姿态、尺寸和深度信息。3D 视觉可以应

下载日志已记录，仅供内部参考

用于机器人抓取、装配定位、尺寸测量、焊接引导、物流分拣和复杂工业检测等场景。随着 3D 相机、结构光、ToF、双目视觉等方案成本下降，3D 视觉有望从高端制造向更多工业产线渗透，成为视觉 AI 参与柔性制造和机器人执行的重要基础。

端侧推理则是视觉 AI 从云端分析走向现场决策的重要支撑。工业质检、机器人、自动驾驶、矿山无人运输和港口无人集卡等场景对实时性要求极高，如果所有视觉数据都上传云端处理，会面临带宽成本高、响应延迟和数据安全等问题。未来，随着边缘 AI 芯片、AI 摄像头、工业边缘服务器和轻量化模型发展，视觉 AI 将更多部署在设备端和现场端，实现本地化实时推理。端侧推理使视觉 AI 能够在产线发现缺陷后立即触发剔除、报警或停机，在机器人抓取过程中实时调整路径，在无人驾驶车辆运行中快速识别障碍物并作出响应。

整体来看，视觉 AI 技术演进的方向，是从“看见目标”升级为“理解场景”，再进一步参与判断和决策。大模型提升了视觉 AI 的语义理解能力，多模态融合增强了系统对复杂生产环境的综合感知能力，3D 视觉提高了空间理解能力，端侧推理则使视觉 AI 能够实时进入生产现场。未来，视觉 AI 的竞争重点将不再只是单一识别准确率，而是多模态融合能力、场景理解能力、边缘部署能力以及与生产系统深度协同的能力。

5.2. 场景深化：从单点替代到全流程嵌入

当前视觉 AI 在工业场景的渗透仍以单点应用为主，即在某个工序解决某个具体问题。未来趋势是从单点应用向全流程嵌入演进。视觉 AI 将逐步覆盖原材料入库、生产加工、质量检测、仓储物流和生产管理等完整链条。在生产前环节，视觉 AI 可以用于来料检测、物料识别、条码/OCR 识别和批次追踪，承担生产准入判断功能；在生产中环节，视觉 AI 可以嵌入在线质检、设备监控和工艺过程监测，实时识别缺陷、尺寸偏差、装配错误和工艺异常；在生产后环节，视觉 AI 可以与质量管理系统结合，进行缺陷追溯、良率分析和工艺优化；在仓储物流环节，视觉 AI 可以参与货物识别、自动分拣、库存盘点和 AGV/AMR 导航。
这种全流程嵌入意味着视觉 AI 不再只是某个检测点的算法模块，而是成为工厂数据采集和智能决策的核心基础设施。过去，工厂中的大量现场信息以非结构化形式存在，例如产品表面图像、人员行为、设备运行状态、货物流转情况等，难以被系统直接利用。视觉 AI 的作用正是将这些非结构化视觉信息转化为可识别、可分析、可追溯的数据资产。当视觉数据进一步与 MES、ERP、质量管理系统、工业互联网平台和设备数据打通后，视觉 AI 将成为连接现实生产现场与数字化管理系统的重要入口。

具身智能与机器人发展将进一步加速这一过程。传统工业机器人主要依赖固定程序和固定轨迹，适合标准化、大批量生产。但在多品种、小批量、快速换线的柔性制造环境下，产品型号、工件姿态、摆放位置和工艺要求经常发生变化，机器人必须具备识别、定位、判断和反馈能力。视觉 AI 为机器人提供了环境感知和目标理解能力，使其能够识别工件类别、判断空间位置、引导机械臂完成抓取、装配、焊接、搬运和检测等任务。
因此，视觉 AI 在机器人场景中的意义，不只是让机器人“看见”工件，而是让机器人具备参与生产执行的能力。在视觉 AI 支持下，机器人可以根据工件位置变化调整抓取路径，可以根据装配偏差调整动作轨迹，可以在作业完成后进行质量检测并反馈结果。这使机器人从固定执行设备升级为可感知、可调整、可协同的柔性生产单元。换言之，视觉 AI 正在成为工厂“神经系统”的核心组成部分，连接感知、判断、执行和反馈。

数字孪生的发展也将放大视觉 AI 的作用。数字孪生需要将现实生产现场映射到虚拟系统中，而视觉 AI 正是获取真实世界数据的重要入口。通过摄像头、工业相机、

下载日志已记录，仅供内部参考

传感器和边缘设备，视觉 AI 可以持续采集产线状态、设备状态、产品质量和人员行为数据，并将其传输到数字孪生系统中，实现工厂运行状态的实时可视化。未来，当视觉 AI 与数字孪生、工业大模型和生产管理系统结合后，企业不仅能够发现现场异常，还可以在虚拟环境中模拟工艺调整、产线调度和设备维护方案，从而推动工厂从可视化管理走向预测性优化。

场景深化的核心在于，视觉 AI 将从局部替代人工，逐步走向全流程参与生产。其价值不再局限于某个工序节省多少人工，而是体现在对生产效率、产品良率、设备协同和管理决策的系统性改善。

5.3. 产业渗透率提升：从头部制造向更广泛产业扩散

从产业渗透角度看，工业视觉 AI 仍处于持续扩散阶段。当前，视觉 AI 在工业领域的应用主要集中在新能源电池、半导体、汽车制造、3C 电子、光伏、医药食品包装等高价值制造场景。这些行业通常具备产品价值较高、质量控制要求严格、缺陷成本较大、自动化基础较好等特征，因此更容易率先引入视觉 AI 系统。
未来，随着技术成熟、硬件成本下降和标准化方案增加，视觉 AI 有望从头部制造企业向更广泛制造业扩散。一方面，在新能源、半导体、汽车电子等高端制造领域，视觉 AI 将从单点质检继续向整线检测、良率优化和生产闭环控制升级；另一方面，在食品饮料、纺织、机械加工、物流仓储、能源巡检等传统行业，视觉 AI 也将通过低成本、轻量化方案逐步渗透。这意味着未来工业视觉 AI 的增长不仅来自高端制造的深度应用，也来自传统制造业和中小企业的广泛改造需求。

中小制造企业将是长期增量市场。过去，中小企业对视觉 AI 的接受度较低，主要原因在于系统成本较高、技术门槛较高、定制化开发周期长、内部缺乏专业 AI 团队，并且投资回报不够清晰。许多中小工厂即使存在人工质检效率低、漏检率高等问题，也难以承担复杂 AI 系统的部署成本。
未来这种情况有望逐步改善。工业相机、边缘计算设备、AI 摄像头和推理芯片的成本下降，将降低视觉 AI 系统的初始投入。而且，低代码或无代码视觉平台、智能相机和标准化行业模型的发展，将减少客户对专业算法团队的依赖。此外，头部制造企业智能化改造会逐步向供应链传导，链主企业可能要求上游供应商提升质量检测和数据追溯能力，从而推动中小企业被动或主动进行视觉 AI 改造。
相比头部企业，中小企业对成本更敏感，对系统稳定性和 ROI 要求更高。因此，未来真正能够打开中小企业市场的，可能不是重定制项目，而是标准化程度更高、部署更轻、价格更低、维护更简单的视觉 AI 产品。例如，面向特定行业的智能检测设备、预训练缺陷模型、边缘 AI 盒子、私有化轻量部署方案和按需订阅的软件服务，都可能成为中小企业市场的重要切入点。

国产替代也是视觉 AI 产业渗透的重要趋势。工业视觉系统涉及图像传感器、工业相机、镜头、光源、图像采集卡、边缘计算芯片、视觉软件平台等多个环节，其中部分高端产品过去长期由海外企业主导。随着国内工业视觉厂商、AI 芯片企业和传感器企业逐步成长，国产化进程有望降低系统成本，提高供应链安全，并增强本土厂商对中国制造业场景的适配能力。
对于投资研究而言，国产替代不仅意味着硬件环节的市场空间，也意味着软硬件一体化解决方案厂商的成长机会。具备工业相机、视觉算法、边缘计算、软件平台和行业交付能力的企业，更容易在国产替代过程中获得产业链价值量提升。未来，视觉 AI 的产业渗透将由成本下降、方案标准化、国产替代和制造业数字化需求共同驱动。

5.4. 商业模式演进：从卖硬件到卖服务

下载日志已记录，仅供内部参考

当前工业视觉 AI 的主流商业模式仍以硬件销售、项目集成和一次性解决方案交付为主。企业通常向客户销售工业相机、镜头、光源、检测设备、视觉软件和系统集成服务。这种模式能够满足早期客户的定制化需求，但也存在收入一次性较强、交付周期较长、项目波动较大、标准化复制困难等问题。

未来，视觉 AI 商业模式将从单一硬件销售向“硬件+软件+服务”转型。随着客户需求从单点检测转向生产流程优化，单纯硬件产品已难以满足复杂生产场景需求。制造企业需要的不只是相机和算法，而是一套能够持续运行、持续优化、持续产生价值的视觉智能系统。因此，未来厂商将更多通过工业视觉平台、算法授权、私有化部署、模型迭代、运维服务和数据分析服务获取持续收入。
工业视觉 AI 的商业模式不会简单等同于互联网软件的公有云 SaaS 化。工业客户往往涉及生产数据、工艺参数、质量记录和商业机密，对数据安全和系统可控要求较高。因此，未来更可能出现 SaaS 化、私有化部署和混合云并存的模式。对于标准化程度较高、数据敏感度较低的场景，SaaS 或订阅制可以降低客户初始投入；而对于大型制造企业、核心产线和高价值工艺场景，私有化部署、专有云和本地运维仍将是重要交付方式。
从收入结构看，视觉 AI 企业未来有望从一次性项目收入向持续性收入演进。例如，厂商可以向客户收取软件授权费、算法订阅费、运维服务费、模型升级费，甚至探索按检测量、按产线、按设备数量或按效果付费的模式。对于供应商而言，这种转型有助于提高收入可预期性和客户粘性；对于客户而言，也能够降低初始投入，并根据实际使用效果持续优化系统。

数据飞轮将成为视觉 AI 企业长期壁垒的重要来源。工业视觉 AI 在部署过程中会积累大量缺陷图像、生产场景、设备状态、工艺异常和质量追溯数据。这些数据可以用于模型迭代和算法优化，提升识别准确率和场景适应能力。部署规模越大、覆盖行业越深，厂商越容易沉淀缺陷样本库、行业模型经验和工程化交付能力，从而提升后续项目复制效率。
但企业中的工业数据飞轮需要谨慎引入。工业生产数据通常涉及客户的产品工艺和商业机密，并不能像互联网用户数据一样简单跨客户共享。因此，数据飞轮的本质并不是直接复用所有客户数据，而是通过行业 Know-how、模型训练方法、场景适配经验、算法模块和工程化能力形成壁垒。能够在多个行业沉淀高频应用模板、标准化检测模块和模型训练工具的企业，将在未来竞争中具备更强优势。

从项目制向产品化复制转变，是商业模式升级的关键。工业视觉 AI 早期项目往往高度定制，每个客户都需要现场调光、模型训练、产线适配和系统联调，导致交付成本高、复制效率低。未来，随着行业模型库、低代码训练平台、智能相机和标准化检测模块普及，厂商有望将部分高频场景产品化，提高交付效率和盈利稳定性。能够从项目经验中沉淀标准化产品的企业，将更容易实现规模化扩张。

综合来看，视觉 AI 在制造业中的角色演进路径正在逐步清晰，从早期辅助人工识别，到嵌入生产环节，再到参与工厂智能决策，最终成为制造业核心生产要素之一。技术层面，大模型、多模态感知、3D 视觉和端侧推理推动视觉 AI 从“看见”走向“理解”和“决策”；场景层面，视觉 AI 从单点质检、安全监控向生产全流程嵌入，并与机器人、数字孪生和工业管理系统形成协同；产业层面，成本下降、中小企业改造需求和国产替代将推动视觉 AI 向更广泛制造业扩散；商业模式层面，企业将从硬件销售和项目交付逐步转向软件平台、运维服务和数据驱动的持续收入。

因此，未来视觉 AI 的产业价值不只来自识别准确率提升，而更来自其能否进入生产流程、形成闭环反馈，并参与生产决策。随着工业制造向智能化、柔性化和数据化方向升级，视觉 AI 有望成为连接物理生产现场与数字决策系统的关键基础设施，推动制造业从自动化进一步走向智能化。这一演进过程不仅将重构制造业的生产组

下载日志已记录，仅供内部参考

织方式，也将为具备工业视觉产品能力、行业数据积累和系统集成能力的企业带来持续成长机会。

6. 典型公司分析

6.1. 海康威视：行业步入上行周期、公司业绩加速兑现

1、公司业绩持续加速兑现。2025 年公司经营表现明显改善。2025 年归母净利润 141.95 亿元，同比增幅 18.52%。从单季度归母净利润看，2025 一季度归母净利润 20.39 亿元，同比增长 6.41%；单二季度归母净利润同比增长 14.94%，单三季度同比增速进一步扩大至 20.31%；2025 年单四季度的归母净利润同比增速进一步扩大至 26%；2026 年一季度的归母净利润同比增速跳升至 36.42%。

2、安防业务步入上行周期。安防摄像头更换周期一般在 5 年左右，上次安防摄像头的安装高峰出现在 2020 年和 2021 年，安防摄像头开始步入更换周期。安防 CIS 主流分辨率从 2M~3M 逐渐过渡升级至 4M~6M 高分辨率已经成为安防应用图像传感器产品的发展趋势。2025 年全球安防 CIS 出货数量在 5.05 亿颗左右，同比增长 3.1%。从安防 CIS 各项性能提升在产品价格上也有所反映，安防设备与产品价格指数 2026 年 4 月 17 日为 139.33，较 5 年前的价格增加近 40%。安防摄像头性能提升更换市场的规模。

3、公司创新业务蓬勃发展。机器人业务 2025 年实现 64.52 亿元，同比增长 8.82%。汽车电子业务受益于 NOA 领航辅助渗透率攀升，聚焦“摄像头+雷达+控制器”产品线，2025 年实现收入 52.89，增速近 35%。智能家居业务 2025 年实现收入 56.84 亿元，同比增长 9.51%。热成像业务 2025 年实现收入 44.43 亿元，同比增长 6.01%。存储业务 2025 年实现收入 23.8 亿元，同比增长 2.85%。

盈利预测与投资评级：我们预计 2026-2028 年公司营业收入分别为 1020.86/1126.21/1244.76 亿元，归母净利润分别为 170.30/202.38/240.28 亿元。

6.2. 大华股份：海外收入持续扩张、视觉 AI 逐步扩散

1、海外业务持续扩张。在全球安防与智慧物联行业格局持续重构、地缘贸易环境趋于复杂，新兴市场智慧城市、公共安全、工业智能化需求稳步扩容的行业大背景下，大华股份持续推进全球化深耕布局，不断优化海外区域市场结构与产品业务结构，加速从传统硬件出口向 AI 物联解决方案、创新业务综合服务商转型。大华股份的海外业务已经从早期单纯的安防设备出口，逐渐发展为覆盖全球的智慧物联解决方案体系。公司产品服务覆盖拉美、中东、东南亚、欧洲等全球 180 多个国家和地区，海外业务布局分散。其中来自发展中国家和地区的收入占公司海外收入的比例较高。大华通过设立海外分支机构、本地化销售团队、区域服务中心和合作渠道，形成了较强的全球运营能力。公司核心竞争力不仅来自摄像机、录像机等硬件产品的成本优势，更来自“产品+平台+解决方案”的整体输出能力。

2、安防业务步入上行周期。安防摄像头更换周期一般在 5 年左右，上次安防摄像头的安装高峰出现在 2018 年和 2021 年，安防摄像头开始步入更换周期。安防 CIS 主流分辨率从 2M~3M 逐渐过渡升级至 4M~6M 高分辨率已经成为安防应用图像传感器产品的发展趋势。2024 年全球安防 CIS 出货数量在 5.28 亿颗左右，实现同比增长约 9%。2025 年全球安防 CIS 出货数量预计在 6 亿颗左右，同比增长 14%。从安防 CIS 各项性能提升在产品价格上也有所反映，安防设备与产品价格指数 2026 年 5 月 29 日为 139.15，较 5 年前的价格增加近 16%。

3、视觉 AI 逐步向各行业渗透。基于视频感知、AI 算法、多模态模型与行业软件平台形成的智能视觉解决方案发展起来的视觉 AI，其应用场景已经由传统安防业务逐步拓展至工业制造、能源巡检、智慧医疗与物流运营等生产领域。相比于传统安防行业，视觉 AI 领域对于研发、数据与软件生态的投入要求显著提升。

4、我们预计 2026-2028 年公司营业收入分别为 352.28/ 380.47/414.71 亿元，归母净利润分别为 39.87/45.44/53.83 亿元。

下载日志已记录，仅供内部参考

7. 风险提示

视觉 AI 技术应用程度低于预期；
基础大模型的迭代对原有视觉 AI 大模型的冲击导致的估值偏差；
业绩预测和估值不达预期。

下载日志已记录，仅供内部参考

研究团队简介：

赵丽明：北京科技大学材料学博士，现任东北证券钢铁行业首席分析师，有多年钢铁生产、市场和设备实业经验。曾在新时代证券、宏源证券、四川信托投资部、中航基金和华夏久盈先后担任研究员和投资经理，2008 年以来具有 18 年证券研究从业经历。

分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师。本报告遵循合规、客观、专业、审慎的制作原则，所采用数据、资料的来源合法合规，文字阐述反映了作者的真实观点，报告结论未受任何第三方的授意或影响，特此声明。

投资评级说明

股票投资评级说明	买入	未来 6 个月内，股价涨幅超越市场基准 15%以上。	投资评级中所涉及的市场基准： A 股市场以沪深 300 指数为市场基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为市场基准；香港市场以摩根士丹利中国指数为市场基准；美国市场以纳斯达克综合指数或标普 500 指数为市场基准。
	增持	未来 6 个月内，股价涨幅超越市场基准 5%至 15%之间。	
	中性	未来 6 个月内，股价涨幅介于市场基准-5%至 5%之间。	
	减持	未来 6 个月内，股价涨幅落后市场基准 5%至 15%之间。	
	卖出	未来 6 个月内，股价涨幅落后市场基准 15%以上。	
行业投资评级说明	优于大势	未来 6 个月内，行业指数的收益超越市场基准。	
	同步大势	未来 6 个月内，行业指数的收益与市场基准持平。	
	落后大势	未来 6 个月内，行业指数的收益落后于市场基准。	



下载日志已记录，仅供内部参考

重要声明

本报告由东北证券股份有限公司（以下称“本公司”）制作并仅向本公司客户发布，本公司不会因任何机构或个人接收到本报告而视其为本公司的当然客户。

本公司具有中国证监会核准的证券投资咨询业务资格。

本报告中的信息均来源于公开资料，本公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅反映本公司于发布本报告当日的判断，不保证所包含的内容和意见不发生变化。

本报告仅供参考，并不构成对所述证券买卖的出价或征价。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的证券买卖建议。本公司及其雇员不承诺投资者一定获利，不与投资者分享投资收益，在任何情况下，我公司及其雇员对任何人使用本报告及其内容所引发的任何直接或间接损失概不负责。

本公司或其关联机构可能会持有本报告中涉及到的公司所发行的证券头寸并进行交易，并在法律许可的情况下不进行披露；可能为这些公司提供或争取提供投资银行业务、财务顾问等相关服务。

本报告版权归本公司所有。未经本公司书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，须在本公司允许的范围内使用，并注明本报告的发布人和发布日期，提示使用本报告的风险。

若本公司客户（以下称“该客户”）向第三方发送本报告，则由该客户独自为此发送行为负责。提醒通过此途径获得本报告的投资者注意，本公司不对通过此种途径获得本报告所引起的任何损失承担任何责任。

东北证券股份有限公司

网址：http://www.nesc.cn 电话：95360,400-600-0686 研究所公众号：dbzqyanjiusuo

地址	邮编
中国吉林省长春市生态大街 6666 号	130119
中国北京市西城区锦什坊街 28 号恒奥中心 D 座	100033
中国上海市浦东新区杨高南路 799 号	200127
中国深圳市福田区福中三路 1006 号诺德中心 34D	518038
中国广东省广州市天河区冼村街道黄埔大道西 122 号之二星辉中心 15 楼	510630