

公司代码：688343

公司简称：云天励飞



深圳云天励飞技术股份有限公司
2024 年年度报告摘要

第一节 重要提示

1、 本年度报告摘要来自年度报告全文，为全面了解本公司的经营成果、财务状况及未来发展规划，投资者应当到 www.sse.com.cn 网站仔细阅读年度报告全文。

2、 重大风险提示

公司已在本报告中详细阐述公司在生产经营过程中可能面临的各种风险，敬请查阅第三节“管理层讨论与分析”之“四、风险因素”。敬请投资者注意投资风险。

3、 本公司董事会、监事会及董事、监事、高级管理人员保证年度报告内容的真实性、准确性、完整性，不存在虚假记载、误导性陈述或重大遗漏，并承担个别和连带的法律责任。

4、 公司全体董事出席董事会会议。

5、 天职国际会计师事务所（特殊普通合伙）为本公司出具了标准无保留意见的审计报告。

6、 公司上市时未盈利且尚未实现盈利

是 否

公司采用《上海证券交易所科创板股票上市规则》第 2.1.2 条第（二）项的标准上市，上市时尚未盈利。公司 2024 年度实现营业收入为 91,737.19 万元；归属于上市公司股东的净利润为 -57,904.70 万元；归属于上市公司股东的扣除非经常性损益的净利润为 -68,345.58 万元。报告期内公司尚未实现盈利。由于公司所属的人工智能行业属于技术和人才密集型行业，具有技术门槛高、高端人才密集、研发投入大的特点，加之近年来国内人工智能行业竞争愈加激烈，在公司收入规模尚未达到一定量级的情况下，持续高强度的研发投入是导致公司尚未实现盈利的主要因素。

7、 董事会决议通过的本报告期利润分配预案或公积金转增股本预案

公司于2025年4月25日召开的第二届董事会第十四次会议审议通过了《关于公司2024年度利润分配方案的议案》，鉴于母公司当前累计未分配利润为负数，尚不满足利润分配条件，为保障和满足公司正常经营和可持续发展需要，公司拟定2024年度利润分配方案为：不派发现金红利、不送红股、不以资本公积金转增股本。

本次利润分配方案尚需经股东大会审议批准。

8、是否存在公司治理特殊安排等重要事项适用 不适用**第二节 公司基本情况****1、公司简介****1.1 公司股票简况**适用 不适用

公司股票简况				
股票种类	股票上市交易所及板块	股票简称	股票代码	变更前股票简称
A股	上海证券交易所科创板	云天励飞	688343	/

1.2 公司存托凭证简况适用 不适用**1.3 联系人和联系方式**

	董事会秘书	证券事务代表
姓名	邓浩然	和邈
联系地址	深圳市南山区粤海街道深圳湾科技生态园10栋B座14-15楼、33楼	深圳市南山区粤海街道深圳湾科技生态园10栋B座14-15楼、33楼
电话	(0755) 26406954	(0755) 26406954
传真	(0755) 86529704	(0755) 86529704
电子信箱	ir@intellif.com	ir@intellif.com

2、报告期公司主要业务简介**2.1 主要业务、主要产品或服务情况****1、公司主营业务基本情况**

公司积极布局边缘 AI 发展的技术和业务，从底层的算法、芯片到场景业务深耕边缘 AI，推动中国通用人工智能发展为目标，持续引领人工智能产业发展。

公司拥有算法和芯片两大核心技术平台，凭借“算法芯片化”的核心能力和“端云协同”的技术路线，不断推进边缘 AI 在消费级、企业级和行业级场景的产品应用，为各行业带来安全、智慧、便捷的 AI 体验。

在算法层面，经过多年的技术研发与积累，构建了人工智能算法平台，拥有大模型研发能力，并推出了“云天天书”大模型。在芯片层面，公司基于对人工智能算法技术特点及行业场景计算需求的深刻理解，通过自定义指令集、处理器架构及工具链的协同设计,实现算法技术芯片化，构建

了神经网络处理器平台。

公司报告期内实现核心技术落地、获取营业收入的主要手段包括面向消费级、企业级和行业级场景的客户销售 AI 产品和解决方案。



产品出货形态及面向的终端场景

2、公司主要产品基本情况

算法芯片化基础技术平台：该平台是公司基础技术研究平台，基于该平台重点研究算法与芯片协同设计技术，包括大模型算法技术、视觉语义类算法技术、AI 处理器设计仿真验证技术、芯片工具链技术等。公司基于算法芯片化基础技术平台，不断提升基础技术研发效率，为公司构建公司核心技术能力提供有力支撑。

基于公司基础技术，公司开发了算法服务、数据挖掘与算法训练、知识库大数据、智能调度中枢、SDC 端边服务、AIoT 物联感知汇聚、模型迭代进化、数字孪生技术系列化核心能力平台。通过上述平台，公司实现快速整合各类核心能力技术组件，面向行业场景业务高效落地应用。公司报告期内的主要的业务由面向消费者、企业和行业的产品和服务构成。

The image displays three product cards, each with a central image and surrounding feature descriptions:

- 噜哇博士 AI 拍学机 (Luwawa Doctor AI Learning Camera):** Features include '识万物' (Recognize everything), '学英语' (Learn English), '讲故事' (Tell stories), and '懂科普' (Understand science). The tagline is '把世界变成孩子的课本' (Turn the world into children's textbooks) and '到户外去, 到广阔天地中去, 在万物自然中研学' (Go outdoors, go to vast lands, study in nature).
- AI "拍拍镜" (AI "Snap" Glasses):** Features include '好看' (Looks good), '好戴' (Easy to wear), '好用' (Easy to use), and '好人文' (Good for humanities). The tagline is '国内首款量产AI眼镜' (Domestic first mass-produced AI glasses) and '外置大脑 数字永生' (External brain, digital immortality).
- 研丞 无线耳机模组 (Yancheng Wireless Earbuds Module):** Features include '加密算法' (Encryption algorithm), '超低功耗' (Ultra-low power consumption), '主动降噪' (Active noise cancellation), and '智能交互' (Smart interaction). The tagline is '为无线耳机提供产品设计服务' (Provide product design services for wireless earbuds) and '设计产品2024年总出货量超三千万副' (Designed products with total shipment of over 30 million pairs in 2024).

在消费级业务方面，基于公司核心技术，打造面向各类边缘 AI 场景的各类硬件设备并对外销售，包括可穿戴设备类产品、AI 交互类产品等。报告期内，公司成立子品牌“噜咔博士”，推出大模型原生教育硬件产品，和合作伙伴闪极科技推出联名智能眼镜产品，以及子公司研丞技术为多个知名品牌提供智能耳机和智能手表的芯片模组、操作系统及产品设计服务。

在企业级业务方面，公司打造了覆盖 NPU IP、SOC 芯片、SOC 模组、边缘计算盒子、AI 推理服务器及智算集群的系列产品，广泛应用在摄像头终端、车路云、机器人、智算中心等场景。公司报告期内的企业级服务业务包括向企业出售芯片模组，和落地大规模异构高性能算力集群为客户提供 AI 训练及推理算力服务，包括智能算力调度及 AI 大模型开发配套服务。



智算中心



天舟系列云/边服务器



边缘计算盒子



DeepEdge 系列芯片



“天舟”大模型一体机



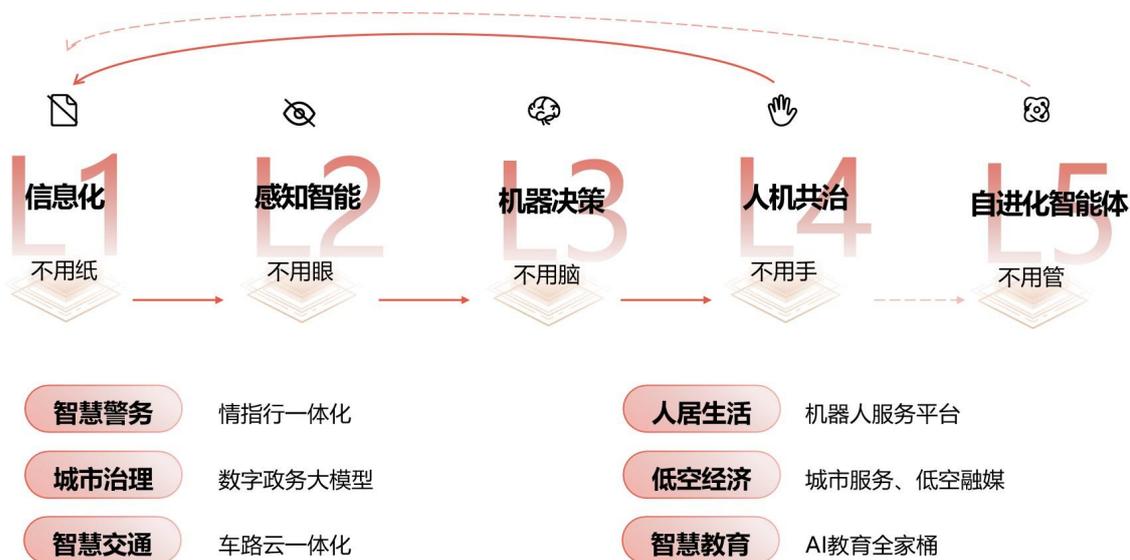
IPU-X6000 推理卡



“深目”AI 模盒

在行业级业务方面，基于公司的软件和硬件通用平台，开发面对行业的软硬一体化解决方案，

目前已形成城市治理、智慧交通、智慧应急、智慧园区、智慧泛商业、智慧警务、智慧教育等行业产品与解决方案体系，并结合多模态大模型技术特点和行业需求，持续拓展创新业务。



2.2 主要经营模式

公司的盈利模式为致力于通过人工智能技术推动边缘 AI 应用的落地，并不断拓展产品在消费级、企业级和行业级市场的规模化落地。

1、消费级：最终面向个人消费者的产品，涵盖可穿戴设备产品、智能硬件等，依托公司自研的大模型或系统、软件应用等为消费级产品提供技术基础，以硬件形式进行交付。

2、企业级：包括 IP 授权、芯片、服务器整机和智算中心等，盈利模式包括：将相关研发成果通过“IP 授权”的方式实现收入；面向设备厂商进行标准化芯片产品交付实现收入；以自研 DeepEdge 系列芯片为核心，向客户提供服务器整机；搭建异构高性能算力集群，为客户的 AI 训练、推理等场景提供高效、稳定和安全的运行环境并实现服务收入。

3、行业级：基于自研的算法软件、芯片等核心产品，搭配外购的定制化或标准化硬件产品、安装施工服务等，面向政府机构等终端客户、企事业单位等集成商客户，根据其需求交付相应软硬件产品或解决方案并对外销售实现收入。

2.3 所处行业情况

(1). 行业的发展阶段、基本特点、主要技术门槛

根据《国民经济行业分类》(GB/T4754-2017)，公司所属行业为“I65 软件和信息技术服务业”中的“I6513 应用软件开发”。同时根据《中国上市公司协会上市公司行业统计分类指引》，公司所属行业为“信息传输、软件和信息技术服务业—软件和信息技术服务业”，行业代码为“I65”。

继 2023 年“百模大战”之后，基础大模型数量逐步收拢，豆包、通义千问、DeepSeek 等逐

步成为基础大模型的主流方向，随着训练成本的下降和大模型开源，落地于具体场景和应用的垂类大模型以及承载应用的推理算力、智能硬件等逐步成为市场需求的爆发点，网络搜索、会议摘要、自动化编程、自动化办公软件等应用渗透率率先提升，过去一年的行业发展总结如下：

2024 年，我们可以看到：

(1) 大模型技术向多模态、轻量化、开源化、平民化方向发展

大模型的能力迅速地从自然语言处理快速向语音、图像、视频等领域横向扩展，在文生图、文生视频等领域出现显著进步。从 Pika 到 Sora，再到 GPT4o 的发展让我们看到大模型在对 3D 场景的理解及视频内容一致性等方面取得了巨大进步，提升了人们对多模态大模型能力上限的想象力。

小型化的技术路线逐渐清晰、明朗，通过高质量 AI 技术设施、高效训练方法与优质数据集探索十亿级参数量下的模型性能极限。小型化基于十亿级别参数量，通过架构调整训练更大体量数据，从而实现匹敌千亿级参数的模型效果。小型化意味着更低的使用成本和部署门槛，部分模型已经走向 PC、手机等终端设备使用，拓宽了大模型的赋能场景。

2024 年 12 月，中国公司 DeepSeek 推出开源大模型 DeepSeek-R1，其性能接近 OpenAI 的 o1 等闭源模型，这一开源策略迅速吸引了全球开发者，HuggingFace 平台基于 DeepSeek 的衍生模型下载量超 500 万次，推动开源社区的技术迭代与应用创新。另外，云端大模型 API 调用成本不断下降，各大厂商打出超低价格。OpenAI 在 5 月将其 API 价格调低 50%，随后国内各家大模型厂商掀起了降价潮，甚至部分模型免费。

(2) 模型能力提升及成本下降推动推理市场需求爆发

2023 年底 OpenAI 发布的 o1 模型标志着 AI 训练范式的重要转型——其训练过程中推理验证循环耗能占比超 60%，传统参数训练占比降至 40%，折射出算力需求从单纯规模扩张向效率优化的跃迁。DeepSeek R1 模型训练成本仅为 560 万美元，以指数级降低的训练成本大幅降低了行业门槛。随着模型全生命周期成本中推理环节占比持续攀升，“训练一次、多场景微调”的架构成为行业新共识，众多中小型企业也能参与大模型开发。成本的降低一方面带来应用需求的增加，另一方面也带来了开发者的增加，开发者的加入会促使更多应用投放市场，使得行业进入良性循环的上升期，而应用需求的整体增长会极大促进推理算力的需求量级攀升，这为国产半导体产业提供了历史性机遇。

根据巴克莱的分析师估计，未来两年内，“前沿 AI”（指最大且最先进的系统）在推理方面的资本支出将超过训练，从 2025 年的 1226 亿美元跃升至 2026 年的 2082 亿美元。博通预测 2027

年仅 3 家客户（谷歌/微软/Meta）的 ASIC+网络需求将达 600-900 亿美元；Marvell 预计 2023-2028 年 AI 推理芯片市场 CAGR 达 45%。同时，国内云厂商阿里、字节、腾讯等均表示，未来在云和 AI 基础设施投入将大幅提升。根据浪潮信息和 IDC 联合发布的《中国人工智能计算力发展评估报告》，2025 年中国智能算力规模将达到 1,037.3EFLOPS，预计到 2028 年将达到 2,781.9EFLOPS。2023-2028 年期间，中国智能算力规模的五年年复合增长率预计达到 46.2%。

同时，一体机作为集成化方案同时将 AI Agent 产业链上游的模型、算法、硬件和中游的平台及下游的应用集合在一起，可以给客户提供针对垂类领域便捷高效的解决方案，符合当前主要需求方——政企、金融客户众多细分场景的部署要求，是当前 AI Agent 的理想实现载体。AI 智慧政务 DeepSeek 一体机部署全国各单位正在推进。根据 IDC 测算，2025 年中国服务器市场出货量有望达到 488 万台，政府、金融、公共事业、医疗等 6 大政企行业由于涉及隐私数据，存在本地私有化部署需求，2021 年上述行业占中国服务器市场需求约 28%。根据中金公司向《通信产业报》全媒体的表示，预计在乐观情景下 2025 年上述行业约 5%的需求转向 DeepSeek 一体机，则需求达到 7 万台，市场规模有望达到 540 亿元。

（3）智能硬件生态进入重构周期

移动设备“手机+APP”的统治格局正被 AI 能力突破所改变，大模型时代会重塑硬件的生态，包括 AI 玩具、AI 可穿戴设备、自动驾驶、机器人/具身智能等硬件使边缘 AI 逐步发展成为行业的重要力量。Meta 眼镜的商业化突破、闪极眼镜等带来的现象级关注等带动行业加速布局智能硬件领域。边缘 AI 让物联网升级成为 AIoT，助力大模型落地。国泰君安证券研究所数据显示，随着 AIoT 的发展，未来 3 年边缘智能硬件设备将保持 50%以上的复合增长率。据 IDC 预测，2025 年全球 AIoT 连接设备数将达 300 亿台，形成万亿级市场规模。

根据 Wellsenn 统计数据，2024 年全球 AI 智能眼镜销量将达 200 万副，2025 年预计类似传统手机厂商、互联网厂商等企业将会积极布局探索 AI 智能眼镜领域，推动 2025 年全球出货量进一步增长至 400 万副。到 2030 年后，AI+AR 技术发展成熟阶段，到 2030 年全球 AI 智能眼镜出货有望达到 8,000 万部，渗透率约为 4.3%。2035 年，全球 AI+AR 智能眼镜销量达到 14 亿台规模，与智能手机规模相当，AI+AR 智能眼镜最终实现传统智能眼镜的替代，达到 70%的渗透率，成为下一代通用计算平台和终端。

（4）国家战略层面持续构建系统性创新生态

我国通过“十四五”规划纲要、“十四五”数字经济规划等顶层设计，明确将人工智能列为战略性新兴产业。2024 年中央经济工作会议明确提出实施“人工智能+”行动，强调通过 AI 与实体经济深度融合，提升传统产业效率并培育新质生产力，这是“人工智能+”行动首次写入国家级的

政府工作报告。2023 年 10 月，国家发改委发布《新一代人工智能基础设施高质量发展行动计划》，提出到 2025 年建成一批国家级 AI 算力枢纽节点，推动通用/专用算力并网调度。工信部等四部委印发《国家人工智能产业综合标准化体系建设指南》，通过系统化的标准布局，旨在解决当前 AI 产业面临的技术碎片化、安全风险高、应用落地难等问题，推动技术、产业与治理的协同发展。

(2). 公司所处的行业地位分析及其变化情况

公司的竞争力来源于自研的算法、芯片技术及两者间的高效适配，以及基于上述核心技术、市场需求快速推出各类高性价比 AI 产品的能力。

在 AI 算法层面，公司算法技术达到业内领先水平。2018 年，公司获得被誉为“中国智能科学技术最高奖”的“吴文俊人工智能科学技术奖”。公司的“面向智慧城市的大规模动态人像识别和实时检索系统”获得 2018 年度深圳市科技进步（技术开发类）一等奖。此外，公司还获评第二十一届中国专利奖。2023 年，公司推出自研千亿级基础大模型“云天天书”，该大模型使用自研训练框架，全栈适配国产算力，在可伸缩训练、断点续训等业界难点问题上屡次取得创新突破。采用首创 SPACE 高性能推理引擎，同等条件下推理速度提升超 80%，做到“算得快、算得多、算得精”。在通用问答、语言理解、数学推理、文本生成、角色扮演等方面均达到行业先进水平，在 2023 年 9 月上旬的 C-Eval 中文大模型榜单中，“云天天书”大模型位列第一。2024 年，公司紧跟 AI 模型的发展趋势，基于视觉和语言两大前期模型的技术积累，发布了视觉语义多模态大模型，协助边缘、端侧设备更好地理解物理世界。公司的视觉语义多模态大模型，以设计轻量化和推理高效为核心，在约十亿量级参数下，通过高效的算子设计，实现了极低的推理成本。在万物识别任务上，效果比肩 InternVL2-40B、Qwen-VL-Max 和 GPT-4o 等主流模型，在某些类别（如动植物的识别方面）实现了超越。通过图文对齐，天书大模型展现了强大的跨模态理解能力。自研图像动态分辨率和词元压缩技术，在有效提高细粒度图文理解准确性的同时，降低推理成本。2024 年，由公司参与研发的“国产 E 级高性能人工智能算力平台”项目获得广东省科技进步奖特等奖。

在 AI 芯片领域，公司是业内少数基于对人工智能算法技术特点的深度分解及对行业场景计算需求的深刻理解，通过自定义指令集、处理器架构及工具链的协同设计，自主研发芯片并已实现流片、量产及市场化销售的公司之一。公司获得 2020 年“吴文俊人工智能专项奖芯片项目一等奖”；作为工信部 2019 年新一代人工智能产业创新重点任务的揭榜单位，开展“面向智能安防及机器人视觉应用的终端神经网络芯片”项目，公司因该项目获评为工信部“第一期人工智能产业创新揭榜优胜单位”；承担国家发改委 2019 人工智能芯片专项“自主指令集的异构芯片”重大专项；承担科

技部 2019 新一代人工智能重大专项“神经网络处理器关键标准和验证芯片”重大专项并顺利通过验收。2023 年，公司新一代芯片 DeepEdge10 完成流片。目前已完成 DeepSeek R1 系列模型、国产鸿蒙操作系统、QwQ-32B 模型适配，不断为国产 AI 生态注入新动能。DeepEdge10 基于公司自研的神经网络处理器（NNP400）架构，采用自主可控的先进全国产工艺打造，采用了创新的“算力积木”架构设计，通过支持多芯粒扩展的 Chiplet 技术，以及 C2C Mesh Torus 互连技术，可以将多个 DeepEdge10 标准计算单元像搭积木一样，封装成不同算力的芯片和多芯片互连的可扩展计算系统，覆盖 8T-256T 算力需求，可实现 7B、14B、130B、671B 等不同参数量大模型的高效推理，赋能各类智算推理硬件产品。针对各类应用场景，云天励飞已开发出 DeepEdge10C、DeepEdge10 标准版、DeepEdge10Max 和 DeepEdge200 四款芯片。

在产品层面，消费级场景下，岍丞技术产品业已进入华为、荣耀、OPPO、VIVO、安克、boAt、Noise 等终端品牌供应链，2024 年全年累计出货超过三千万套；企业级场景下，公司自研芯片与海康威视、阿里巴巴平头哥等头部客户建立了业务合作关系。2024 年公司推出全新大模型推理加速卡 IPU-X6000，采用全国产工艺、D2D Chiplet 技术，专门针对大模型推理任务设计，研发了符合大模型演进趋势的统一工具链，包括分布式并行策略、基于硬件的流水线排布、先进的量化策略、多机并行的编译机制等，在同等条件下能够做到推理性能更高、消耗成本更低；行业级场景下，公司已经在多城市落地多个重点项目，参与建设了多个城市的智慧安防、巴士智能调度系统、智慧社区、万科印力智慧商业、智慧书城等解决方案，2022 年公司与中国科学技术大学联合完成的“大规模视频结构化关键技术研发及产业化”项目荣获第十一届吴文俊人工智能科学技术奖一等奖。目前各场景中 AI 渗透率仍较低，未来发展空间较大，行业及公司在各细分场景的产品应用仍有较大增长空间。

同时，公司还多次受邀参与人工智能国家级行业标准的制定，参与由国家人工智能标准化总体组、中国电子技术标准化研究院发起的《2021 人工智能标准化白皮书》编撰工作，参与起草由中国电子工业标准化技术协会发布的《信息技术人工智能机器学习模型及系统的质量要素和测试方法》《信息技术人工智能面向机器学习的系统框架和功能要求》《信息技术人工智能面向机器学习的数据标注规程》等多项标准。在产业标准方面，2022 年公司参与编写中国在人工智能领域发布的第一批国家标准，包括《信息技术计算机视觉术语》《信息技术人工智能术语》《信息技术生物特征识别人脸识别系统技术要求》等。2020 年入选全国信息技术标准化技术委员会人工智能分技术委员会（首届人工智能国家标准化组织）单位委员名单并当选全国信息技术标准化技术委员会可信赖研究组副组长。

(3). 报告期内新技术、新产业、新业态、新模式的发展情况和未来发展趋势

报告期内，国内外大模型技术都有了长足的进步。以 OpenAI 发布的 Sora、GPT-4o，Google 推出的 Gemini 1.5，字节跳动的豆包大模型为代表，大模型从单模态向多模态演进，融合文本、语音、图像、视频等多种模态的技术逐渐成熟，实现跨模态生成和交互，显著提升了大模型对复杂场景的理解能力，也提升了内容创作和人机交互能力。以 OpenAI 推出的 o3-mini、o1、o1-mini 模型，DeepSeek 发布开源 DeepSeek-R1 模型，阿里通义千问团队发布 QwQ-32B 模型为代表，推理模型通过强化学习和内化思维链，提升了模型的复杂推理思考能力，能解决 STEM 领域的难题。

模型架构技术也有较大的优化，上下文窗口长度得到拓展，主流大模型支持的上下文长度达到 128K，通过稀疏注意力机制和训练算法优化提升了长文本处理能力；混合专家架构成为趋势，通过多个专家分工协同处理，提升了模型效率。随着大模型技术的快速发展和能力边界的拓展，大模型与企业级工具和服务集成，提升办公生产效率，成为新质生产力发展的重要引擎，开源大模型生态带动开发社区活跃，降低模型应用的门槛，加速行业垂类应用落地，借助 AI 模型推理实现 AI 应用规模化落地，正成为人工智能产业发展的关键趋势。AI 应用的规模化落地也带动了智能硬件的创新，在边缘 AI 推理芯片、模型轻量化边缘部署的配合下，各类智能硬件迎来爆发式增长，涵盖大模型推理计算集群、大模型一体机、AIPC、AI 手机、AI 耳机、AI 眼镜、拍学机、自动驾驶以及人形机器人等。

我们看到以下趋势及机会：

(1) 大模型通用性、泛化能力增强的趋势，与行业结合的趋势同时存在，并行发展。一方面，模型参数量及规模在过去两年中快速增加，大模型出现“智能涌现”。如谷歌推出的多模态具身视觉语言模型 PaLM-E 参数量已超过五千亿，而 ChatGPT4 参数量达到 1.8 万亿。规模的增长使得模型在知识密集型、多步骤复杂型等任务下的表现都得到显著提升，泛化能力增强；另一方面，大模型在落地应用过程中与各行业“Know-how”相结合，亦呈现小型化的趋势，使模型学习特定领域知识，形成针对该行业、领域的“轻量版”行业大模型和领域大模型，提升适用性及准确度。如微软的 Phi-2，Meta 的 lightLLM，谷歌的 Gemma 等都展现了小模型在特定领域的优异表现。

(2) 大模型的能力迅速地从自然语言处理快速向语音、图像、视频等领域横向扩展，在文生图、文生视频等领域出现显著进步。过去一年中，从 Pika 到 Sora 的发展让我们看到大模型在对 3D 场景的理解及视频内容一致性等方面取得了巨大进步，提升了人们对多模态大模型能力上限的想象力。多模态大模型具备对图、文、音等不同规格、类型的数据的理解能力及生成能力，能更进一步提升生产效率，已成为大模型发展的重要方向之一。

(3) 大模型商业化应用加速, 边缘 AI 成为重要发展方向。随着大模型在多模态理解生成(如文本、图像、音频等多种模态信息的精准解析与融合生成)、交互能力(实现更加自然、流畅、智能的人机交互体验)以及复杂问题推理思考(具备对复杂逻辑、因果关系、不确定性问题的深度分析与推理能力)等方面取得显著进步, 大模型深耕多个行业, 加速拓展应用场景, 在场景需求驱动下大模型开始在很多领域转化为生产力, 提升生产效率, 成为办公工作、生产制造、物流运维、创意创作的高效工具, 例如网络搜索、会议摘要、自动化编程、自动化办公软件、专业领域咨询、广告设计、文本创作等。大模型应用部署从云端将逐渐向边缘和端侧延申的趋势逐渐明朗, 用户对应用提出实时性、安全性、隐私性的需求, 面向大模型推理应用的边缘 AI 成为重要发展方向。

(4) 技术革新推动单位算力成本下降, 促使需求显著增长。多头潜在注意力 (MLA)、模型-硬件协同设计、精准资源分配、动态缩放策略等技术手段使得模型训练及应用的单位成本进一步降低, 促使大模型的商业化应用进程不断加速, AI 推理算力需求开始呈现出快速增长, 例如目前大模型在搜索业务中呈现的爆发式增长趋势。这类需求增长进一步提出对 AI 推理性能、能效比等更为严苛的要求, 而依托高效 AI 推理芯片, 借助软硬件协同优化的技术手段, 构建高性价比的软硬一体 AI 基础设施已成为关键。

(5) 智能硬件将成为推动 AI 应用广泛普及的核心载体。各类形态丰富的智能硬件, 在 AI 应用落地进程中承担着多样化的关键角色: 部分智能硬件专注于边缘端, 以高效响应与数据安全为特色, 提供即时的 AI 推理算力支持; 另一部分则致力于构建人性化交互体验, 同时强化个性化数据隐私保护机制; 更有智能硬件借助端云协同技术, 成为全时、全场景大模型人机交互的便捷入口。

3、公司主要会计数据和财务指标

3.1 近 3 年的主要会计数据和财务指标

单位: 元 币种: 人民币

	2024年	2023年	本年比上年 增减(%)	2022年
总资产	4,918,496,203.98	4,853,584,032.04	1.34	1,906,471,597.77
归属于上市公司股东的净资产	3,936,669,224.62	4,380,324,880.02	-10.13	1,405,132,129.18
营业收入	917,371,920.91	506,008,643.32	81.30	546,217,591.67
扣除与主营业务无关的业务收入和不具备商业实质的收入后的营业收入	908,021,020.49	500,201,808.26	81.53	544,900,854.57

归属于上市公司股东的净利润	-579,047,039.30	-383,117,186.32	不适用	-447,109,776.69
归属于上市公司股东的扣除非经常性损益的净利润	-683,455,758.85	-499,686,723.87	不适用	-528,580,927.00
经营活动产生的现金流量净额	-294,519,855.49	-541,011,153.43	不适用	-410,749,047.10
加权平均净资产收益率(%)	-13.93	-10.55	减少3.38个百分点	-35.79
基本每股收益(元/股)	-1.63	-1.15	不适用	-1.68
稀释每股收益(元/股)	-1.63	-1.15	不适用	-1.68
研发投入占营业收入的比例(%)	43.59	58.27	减少14.68个百分点	63.44

3.2 报告期分季度的主要会计数据

单位：元 币种：人民币

	第一季度 (1-3 月份)	第二季度 (4-6 月份)	第三季度 (7-9 月份)	第四季度 (10-12 月份)
营业收入	98,448,460.88	191,014,198.89	193,510,653.25	434,398,607.89
归属于上市公司股东的净利润	-137,817,011.37	-171,929,581.12	-115,553,504.65	-153,746,942.16
归属于上市公司股东的扣除非经常性损益后的净利润	-156,741,752.89	-189,151,898.05	-154,541,731.05	-183,020,376.86
经营活动产生的现金流量净额	-166,268,010.57	-387,654,719.63	105,899,726.09	153,503,148.62

季度数据与已披露定期报告数据差异说明

适用 不适用

4、 股东情况

4.1 普通股股东总数、表决权恢复的优先股股东总数和持有特别表决权股份的股东总数及前 10 名股东情况

单位：股

截至报告期末普通股股东总数(户)	28,877
年度报告披露日前上一月末的普通股股东总数(户)	34,732
截至报告期末表决权恢复的优先股股东总数(户)	0
年度报告披露日前上一月末表决权恢复的优先股股东总数(户)	0

截至报告期末持有特别表决权股份的股东总数（户）							0
年度报告披露日前上一月末持有特别表决权股份的股东总数（户）							0
前十名股东持股情况（不含通过转融通出借股份）							
股东名称 （全称）	报告期内 增减	期末持股 数量	比例 （%）	持有有限 售条件股 份数量	质押、标记或冻 结情况		股东 性质
					股份 状态	数量	
陈宁	0	83,672,080	23.56	83,672,080	无	0	境内自 然人
东海云天	0	28,459,385	8.01		无	0	其他
珠海创享一号	-1,500,000	9,041,860	2.55		无	0	其他
投控东海	0	6,377,840	1.80		无	0	其他
深圳创享二号	-3,154,098	6,316,162	1.78		无	0	其他
明德致远	0	6,083,700	1.71	6,083,700	无	0	境内非 国有法 人
王孝宇	600	6,022,720	1.70		无	0	境内自 然人
中电信息	0	5,748,934	1.62		无	0	国有法 人
中电金控	0	5,748,934	1.62		无	0	国有法 人
珠海创享二号	-1,500,000	3,266,460	0.92		无	0	其他
上述股东关联关系或一致行动的说明	截至报告期末，上述股东中：1）公司董事长、总经理陈宁持有明德致远 100% 股权，互为一致行动人；2）中电信息、中电金控均为中国电子信息产业集团有限公司全资子公司，互为一致行动人；3）投控东海和东海云天的执行事务合伙人均为深圳市投控东海投资有限公司。						
表决权恢复的优先股股东及持股数量的说明	/						

存托凭证持有人情况

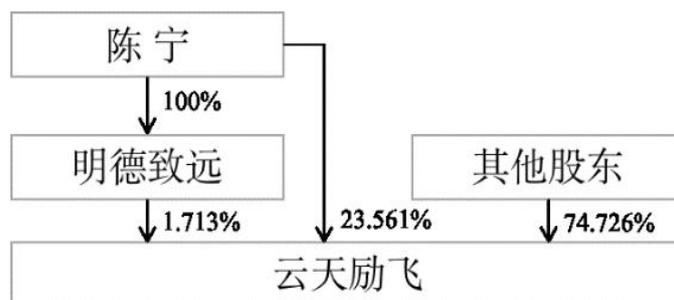
适用 不适用

截至报告期末表决权数量前十名股东情况表

适用 不适用

4.2 公司与控股股东之间的产权及控制关系的方框图

适用 不适用



4.3 公司与实际控制人之间的产权及控制关系的方框图

适用 不适用

4.4 报告期末公司优先股股东总数及前 10 名股东情况

适用 不适用

5、公司债券情况

适用 不适用

第三节 重要事项

1、公司应当根据重要性原则，披露报告期内公司经营情况的重大变化，以及报告期内发生的对公司经营情况有重大影响和预计未来会有重大影响的事项。

报告期内，公司实现营业收入 91,737.19 万元，同比增长 81.30%，归属于母公司所有者的净利润-57,904.70 万元，较去年同期亏损金额扩大 19,592.99 万元。具体经营情况分析详见本章节“一、经营情况讨论与分析”相关内容。

2、公司年度报告披露后存在退市风险警示或终止上市情形的，应当披露导致退市风险警示或终止上市情形的原因。

适用 不适用