

2024年11月26日

# 英伟达推出全新 CPU 和 GPU AI 处理器, IDEA 研究院发布通用视觉大模型 DINO-X

—计算机行业周报

## 推荐(维持)

## 投资要点

分析师: 宝幼琛 S1050521110002  
baoyc@cfsc.com.cn

### 行业相对表现

表现	1M	3M	12M
计算机(申万)	3.8	56.5	2.1
沪深300	-2.9	15.5	8.5

### 市场表现



资料来源: Wind, 华鑫证券研究

### 相关研究

- 《计算机行业点评报告: 苹果(AAPL): Apple Intelligence 持续推进, 推动 AI 手机升级》2024-11-20
- 《计算机行业周报: 百度举办百度世界大会, 通义千问开源 Qwen2.5-Coder 全系列》2024-11-17
- 《计算机行业周报: ChatGPT 搜索功能上线, 腾讯开源最新 MoE 模型》2024-11-12

## 算力: 算力租赁价格平稳, 英伟达推出全新 CPU 和 GPU AI 处理器

Nvidia 发布了两款产品: 具有两个 Grace CPU 的四通道 B200 GPU 模块——GB200 NVL4, 以及针对风冷数据中心的 H200 NVL PCIe GPU。具体介绍如下:

1) GB200 Grace Blackwell NVL4 超级芯片是标准(非 NVL4)双 GPU 变体的更有效的变体, 具有多达四个与 NVLink 相连的 B200 Blackwell GPU 和两个基于 Grace ARM 的 CPU, 全部位于一个主板上。该解决方案针对 HPC 和 AI 混合工作负载, 具有高达 1.3TB 的一致性内存。Nvidia 宣传 GB200 NVL4 的模拟性能是 Nvidia GH200 NVL4 的 2.2 倍, 训练是 1.8 倍, 推理性能是 Nvidia GH200 NVL4 的 1.8 倍, 是其直接前身。英伟达表示, GB200 NVL4 超级芯片将于 2024 年下半年从各种供应商处提供, 例如微星、华硕、技嘉、纬创资通、和硕、华擎机架、联想、惠普企业等。

2) H200 NVL 是一款双插槽风冷 GPU, 具有 PCIe 5.0 连接 (128 GB/s)。该冷却器针对机架安装解决方案进行了优化, 采用流通式设计, 其中进气从右向左流动; 没有鼓风机式风扇。其性能比 Nvidia 即将推出的 SXM 外形尺寸的 H200 略差。H200 NVL 的额定 FP64 为 30 TFLOPS, FP32 的额定值为 60 TFLOPS。Tensor 核心性能的额定值为 FP64 的 60 TFLOPS、TF32 的 835 TFLOPS、BFLOAT16 的 1671 TFLOPS、FP16 的 1671 TFLOPS、FP8 的 3341 TFLOPS 和 INT8 的 3341 TFLOP。但 Nvidia 表示 H200 NVL 比它所取代的 H100 NVL 快得多。它具有 1.5 倍的内存容量和 1.2 倍的内存带宽, 为 HPC 工作负载提供高达 1.7 倍的推理性能和 1.3 倍的性能。Nvidia 还快速与 Ampere 进行了比较, 指出 H200 NVL 比 Ampere 的同等 GPU 快 2.5 倍。

## IDEA 研究院发布通用视觉大模型 DINO-X

IDEA 研究院发布最新成果通用视觉大模型 DINO-X。它实现视觉任务大一统, 支持各种开发世界感知和目标理解任务, 包括开放世界对象检测与分割、短语定位、视觉提示计数、姿态估计、无提示对象检测与识别、密集区域字幕等。这背后

得益于，他们构建了超过一亿高质量样本的大型数据集 Grounding-100M。与之前 DINO 家族中 Grounding DINO 1.5 类似，DINO-X 也分为 DINO-X Pro 模型和 DINO-X Edge 模型。前者可为各种场景提供更强的感知能力，后者经过优化，推理速度更快，更适合部署在边缘设备上。

实验结果表明，DINO-X Pro 模型在 COCO、LVIS-minival 和 LVIS-val 零镜头物体检测基准测试中分别获得了 56.0 AP、59.8 AP 和 52.4 AP 的成绩。尤其是在 LVIS-minival 和 LVIS-val 这两个罕见类别的基准测试中实现新 SOTA——分别获得了 63.3 AP 和 56.5 AP 的成绩，都比之前的 SOTA 性能提高了 5.8 AP。

DINO-X 的万物识别能力，让其拥有了对开放世界（Open World）的视觉感知，可以轻松应对真实世界中的众多不确定性和开放环境。IDEA 研究院创院理事长沈向洋提到了具身智能、大规模多模态数据的自动标注、视障人士服务等这几个场景。

目前，IDEA 研究院与多家企业联合开展视觉大模型及应用场景的探索研究工作，在视觉大模型的落地方面取得了实质性进展。一方面，有别于市场上的以语言为基础的多模态大模型基于全图理解的方法，通过在物体级别的理解上加入语言模块优化大模型幻觉问题。另一方面结合自研的“视觉提示优化”方法，无需更改模型结构，不用重新训练模型，实现小样本下的场景化定制。

## ■ AI 融资动向：xAI 获 50 亿美元融资，目前估值达 500 亿美元

本期 AI 初创公司的融资中，xAI 融资额排名第一，获 50 亿美元融资。xAI 成立于 2023 年的 7 月 12 日，虽然成立的时间不长，但 xAI 在人工智能领域已有不错的进展，在成立次月的 18 日，他们就宣布完成了首个旗舰模型 Grok-0 的训练，有 330 亿参数；两个多月后的 11 月 3 日，他们宣布了 Grok-1 模型；在今年的 3 月 17 日，他们宣布将 Grok-1 的权重和架构开源；3 月 28 日宣布的 Grok-1.5 模型，增加了长文本能力；4 月 12 日推出的 Grok-1.5V，则是增加了视觉能力，可以根据交错的自然语言和图像进行推理；而在今年 8 月 15 日，他们推出了推理能力更强的大语言模型 Grok-2，更小但能力不俗的 Grok-2 mini 也在当天一并推出，Grok-2 和 Grok-2 mini 在测试中的部分能力，超过了 Anthropic 旗下的 Claude 3.5 Sonnet 和 OpenAI 旗下的 GPT-4 Turbo。xAI 截至今年已打造有 10 万块英伟达 H100 的算力集群，是当前全球最大的 AI 训练系统。

## ■ 投资建议

海外应用侧最重要场景已具雏形，国内政策持续支持。海外方面，当地时间 11 月 19 日，在“Microsoft ignite 2024”

全球开发者大会上，围绕 AI Agent/AI 智能体，微软更新了多款 AI 产品。微软的目的在于建立完善的 AI 智能体生态，自推出以来，该生态联盟已有超过 10 万家组织通过其 Copilot Studio 创建或编辑过 AI 助理。国内方面，近日，人工智能辅助诊断首次被列入价格构成。国家医保局分析人工智能潜在的应用场景，在放射检查、超声检查、康复类项目中设立“人工智能辅助”扩展项，即同样的价格水平下，医院可以选择培养医务人员进行诊疗，也可以选择使用人工智能参与诊疗行为，但现阶段不重复收费。预计该政策将大幅加快人工智能辅助技术进入临床应用速度。

我们坚定持续看好 AI 软硬件机会，建议关注以 AI 为核心的龙头厂商科大讯飞（002230.SZ）、有望迎来需求爆发的 AI 应用金桥信息（603918.SH）、高速通信连接器业务或显著受益于 GB200 放量的鼎通科技（688668.SH）。

### ■ 风险提示

1) AI 底层技术迭代速度不及预期。2) 政策监管及版权风险。3) AI 应用落地效果不及预期。4) 推荐公司业绩不及预期风险。

#### 重点关注公司及盈利预测

公司代码	名称	2024-11-26 股价	EPS			PE			投资评级
			2023	2024E	2025E	2023	2024E	2025E	
002230.SZ	科大讯飞	47.04	0.28	0.40	0.56	168.00	117.60	84.00	买入
688256.SH	寒武纪-U	485.90	-2.04	-1.21	-0.50	-238.19	-401.57	-971.80	买入
688668.SH	鼎通科技	35.80	0.67	1.04	1.41	53.43	34.42	25.39	买入

资料来源：Wind，华鑫证券研究

## 正文目录

1、 算力动态：算力租赁价格平稳，英伟达推出全新 CPU 和 GPU AI 处理器 .....	5
1.1、 数据跟踪：算力租赁价格平稳 .....	5
1.2、 产业动态： 英伟达推出全新 CPU 和 GPU AI 处理器，性能大幅提升.....	5
2、 AI 应用动态：PERPLEXITY 搜索访问量环比+23.66%.....	7
2.1、 流量跟踪：Character.AI 搜索访问量环比+9.00%.....	7
2.2、 产业动态：IDEA 研究院发布通用视觉大模型 DINO-X .....	7
3、 AI 融资动向：XAI 获 50 亿美元融资，目前估值达 500 亿美元.....	12
4、 行情复盘.....	13
5、 投资建议.....	16
6、 风险提示.....	16

## 图表目录

图表 1：本周算力租赁情况 .....	5
图表 2：四路 NVLink 桥接器互联的 H200 NVL .....	6
图表 3：2024.11.18-2024.11.22 AI 相关网站流量.....	7
图表 4：DINO-X Pro 模型的成绩 .....	8
图表 5：全面检测 .....	9
图表 6：细粒度目标理解 .....	10
图表 7：多任务感知与理解 .....	10
图表 8：本周 AI 初创公司的融资动态 .....	13
图表 9：本周指数日涨跌幅 .....	13
图表 10：本周 AI 算力指数内部涨跌幅度排名 .....	14
图表 11：本周 AI 应用指数内部涨跌幅度排名 .....	14
图表 12：重点关注公司及盈利预测 .....	16

# 1、算力动态：算力租赁价格平稳，英伟达推出全新 CPU 和 GPU AI 处理器

## 1.1、数据跟踪：算力租赁价格平稳

本周恒源云 13 核+128G 价格环比上周上涨 5.88%。具体来看，显卡配置为 A100-40G 中，腾讯云 16 核+96G 价格为 28.64 元/时，阿里云 12 核+94GiB 价格为 31.58 元/时；显卡配置为 A100-80G 中，恒源云 13 核+128G 价格为 9.00 元/时；阿里云 16 核+125GiB 价格为 34.74 元/时；显卡配置为 A800-80G 中，恒源云 16+256G 价格为 9.00 元/时。

图表 1：本周算力租赁情况

显卡配置	CPU	内存	磁盘大小 (G)	平台名称	价格 (每小时)	价格环比上周
A100-40G	16	96	可自定, 额外收费	腾讯云	28.64/元	0.00%
	12 核	94G	可自定, 额外收费	阿里云	31.58/元	0.00%
A100-80G	13	128	系统盘: 20G 数据盘: 50GB	恒源云	9.00/元	5.88%
	16 核	125G	可自定, 额外收费	阿里云	34.74/元	0.00%
A800-80G	16	256	系统盘: 20G 数据盘: 50GB	恒源云	9.00/元	0.00%

资料来源：腾讯云，阿里云，恒源云，华鑫证券研究

## 1.2、产业动态：英伟达推出全新 CPU 和 GPU AI 处理器，性能大幅提升

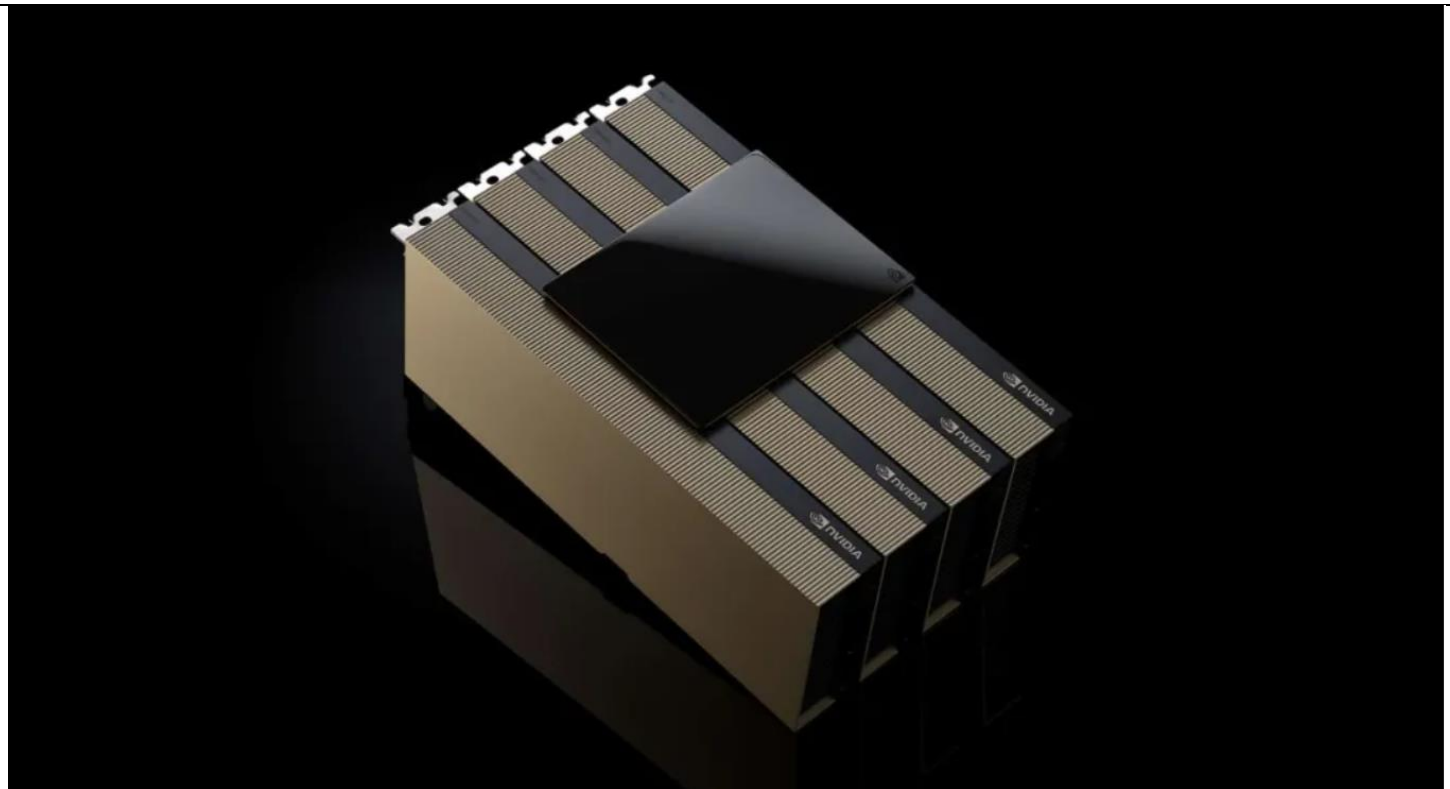
Nvidia 发布了两款产品：具有两个 Grace CPU 的四通道 B200 GPU 模块——GB200 NVL4，以及针对风冷数据中心的 H200 NVL PCIe GPU。具体介绍如下：

1) GB200 Grace Blackwell NVL4 超级芯片是标准（非 NVL4）双 GPU 变体的更有效的变体，具有多达四个与 NVLink 相连的 B200 Blackwell GPU 和两个基于 Grace ARM 的 CPU，全部位于一个主板上。该解决方案针对 HPC 和 AI 混合工作负载，具有高达 1.3TB 的一致性内存。Nvidia 宣传 GB200 NVL4 的模拟性能是 Nvidia GH200 NVL4 的 2.2 倍，训练是 1.8 倍，推理性能是 Nvidia GH200 NVL4 的 1.8 倍，是其直接前身。英伟达表示，GB200 NVL4 超级芯片将于 2024 年下半年从各种供应商处提供，例如微星、华硕、技嘉、纬创资通、和硕、华擎机架、联想、惠普企业等。

2) H200 NVL 是一款双插槽风冷 GPU，具有 PCIe 5.0 连接（128 GB/s）。该冷却器针对机架安装解决方案进行了优化，采用流通式设计，其中进气从右向左流动；没有鼓风机式风扇。其性能比 Nvidia 即将推出的 SXM 外形尺寸的 H200 略差。H200 NVL 的额定 FP64 为 30 TFLOPS，FP32 的额定值为 60 TFLOPS。Tensor 核心性能的额定值为 FP64 的 60 TFLOPS、TF32 的 835 TFLOPS、BFLOAT16 的 1671 TFLOPS、FP16 的 1671 TFLOPS、FP8 的 3341 TFLOPS 和 INT8 的 3341 TFLOP。但 Nvidia 表示 H200 NVL 比它所取代的 H100

NVL 快得多。它具有 1.5 倍的内存容量和 1.2 倍的内存带宽，为 HPC 工作负载提供高达 1.7 倍的推理性能和 1.3 倍的性能。Nvidia 还快速与 Ampere 进行了比较，指出 H200 NVL 比 Ampere 的同等 GPU 快 2.5 倍。

图表 2: 四路 NVLink 桥接器互联的 H200 NVL



资料来源：硅基 LIFE，华鑫证券研究



## 2、AI 应用动态：Perplexity 搜索访问量环比+23.66%

### 2.1、流量跟踪：Character.AI 搜索访问量环比+9.00%

本期（2024.11.18-2024.11.22）AI 相关网站流量数据：访问量前三位分别为 ChatGPT（918.4M）、Bing（421.8M）和 Canva（189.9M）；访问量环比增速前三位分别为 Gemini（9.00%）、Character.AI（3.49%）和 ChatGPT（2.86%）；平均停留时长前三位分别为 Character.AI（00:12:56）、Discord（00:12:25）和 DeepL（00:08:56）；平均停留时长环比增速前三位分别为 Kimi（3.02%）、Character.AI（0.78%）和 ChatGPT（0.54%）。

图表 3：2024.11.18-2024.11.22 AI 相关网站流量

应用	应用类型	归属公司	周平均访问量 (M)	访问量环比	平均停留时长	时长环比
ChatGPT	聊天机器人	OpenAI	918.4	2.86%	6:15	0.54%
Bing	搜索	微软	421.8	-0.07%	6:45	-0.49%
Discord	游戏社区	微软	126.2	0.88%	12:25	-0.40%
Canva	在线设计	Canva	189.9	2.10%	7:57	-0.42%
Github	代码托管	微软	117.7	0.86%	6:34	-0.76%
Gemini	聊天机器人	谷歌	68.31	9.00%	4:39	-1.76%
Character.AI	聊天机器人	Character.AI	43.86	3.49%	12:56	0.78%
NotionAI	文本/笔记	Notion	36.09	-0.25%	8:47	-0.19%
QuillBot	释义工具	QuillBot	18.55	0.82%	4:07	0.00%
Kimi	聊天机器人	Moonshot AI	7.606	1.77%	3:25	3.02%
DeepL	翻译工具	DeepL	49.79	0.50%	8:56	-0.19%
文心一言	聊天机器人	百度	5.352	2.18%	2:58	-0.56%
Perplexity	AI 搜索	Perplexity	25.10	2.41%	5:57	-1.38%

资料来源：similarweb, 华鑫证券研究

### 2.2、产业动态：IDEA 研究院发布通用视觉大模型 DINO-X

IDEA 研究院发布最新成果通用视觉大模型 DINO-X。它实现视觉任务大一统，支持各种开发世界感知和目标理解任务，包括开放世界对象检测与分割、短语定位、视觉提示计数、姿态估计、无提示对象检测与识别、密集区域字幕等。这背后得益于，他们构建了超过一

亿高质量样本的大型数据集 Grounding-100M。与之前 DINO 家族中 Grounding DINO 1.5 类似，DINO-X 也分为 DINO-X Pro 模型和 DINO-X Edge 模型。前者可为各种场景提供更强的感知能力，后者经过优化，推理速度更快，更适合部署在边缘设备上。

实验结果表明，DINO-X Pro 模型在 COCO、LVIS-minival 和 LVIS-val 零镜头物体检测基准测试中分别获得了 56.0 AP、59.8 AP 和 52.4 AP 的成绩。尤其是在 LVIS-minival 和 LVIS-val 这两个罕见类别的基准测试中实现新 SOTA——分别获得了 63.3 AP 和 56.5 AP 的成绩，都比之前的 SOTA 性能提高了 5.8 AP。

图表 4: DINO-X Pro 模型的成绩

Method	Organization	COCO-val	LVIS-minival				LVIS-val			
		$AP_{all}$	$AP_{all}$	$AP_r$	$AP_c$	$AP_f$	$AP_{all}$	$AP_r$	$AP_c$	$AP_f$
OWL-ViT	谷歌	42.2	-	-	-	-	34.6	31.2	-	-
MDETR	NYU & Meta		22.5	7.4	22.7	25	-	-	-	-
GLIP	微软	49.8	37.3	28.2	34.3	41.5	26.9	17.1	23.3	35.4
Grounding DINO	IDEA	48.4	27.4	18.1	23.3	32.7	-	-	-	-
OpenSeeD	IDEA		23	-	-	-	-	-	-	-
UniDetector	清华大学	-	-	-	-	-	19.8	18	19.2	21.2
OmDet-Turbo-B	联汇	53.4	34.7	-	-	-	-	-	-	-
OWL-ST	谷歌	-	40.9	41.5	-	-	35.2	36.2	-	-
MQ-GLIP	腾讯	-	43.4	34.5	41.2	46.9	34.7	26.9	32	41.3
MM-Grounding-DINO	上海 AI Lab & 商汤	50.4	41.4	34.2	37.4	46.2				
DetCLIP	华为	-	38.6	36	38.3	39.3	28.4	25	27	31.6
DetCLIPv2	华为	-	44.7	43.1	46.3	43.7	36.6	33.3	36.2	38.5
DetCLIPv3	华为	-	48.8	49.9	49.7	47.8	41.4	41.4	40.5	42.3
YOLO-World	腾讯	45.1	35.4	27.6	34.1	38	-	-	-	-
OV-DINO	美团 & 中大	50.2	40.1	34.5	39.5	41.5	32.9	29.1	30.4	37.4
T-Rex2 (visual)	IDEA	46.5	47.6	45.4	46	49.5	45.3	43.8	42	49.5
T-Rex2 (text)	IDEA	52.2	54.9	49.2	54.8	56.1	45.8	42.7	43.2	50.2
Grounding DINO 1.5 Pro	IDEA	54.3	55.7	56.1	57.5	54.1	47.6	44.6	47.9	48.7
Grounding DINO 1.6 Pro	IDEA	55.4	57.7	57.5	60.5	55.3	51.1	51.5	52	50.1
DINO-X	IDEA	56	59.8	63.3	61.7	57.5	52.4	56.5	51.1	51.9

资料来源：量子位，华鑫证券研究

总结来看，DINO-X 主要以下特点：

1) **全面检测**。几乎可识别所有物体。DINO-X 称得上目前业界检测最全的通用视觉模型，甚至无需用户提示。



图表 5：全面检测

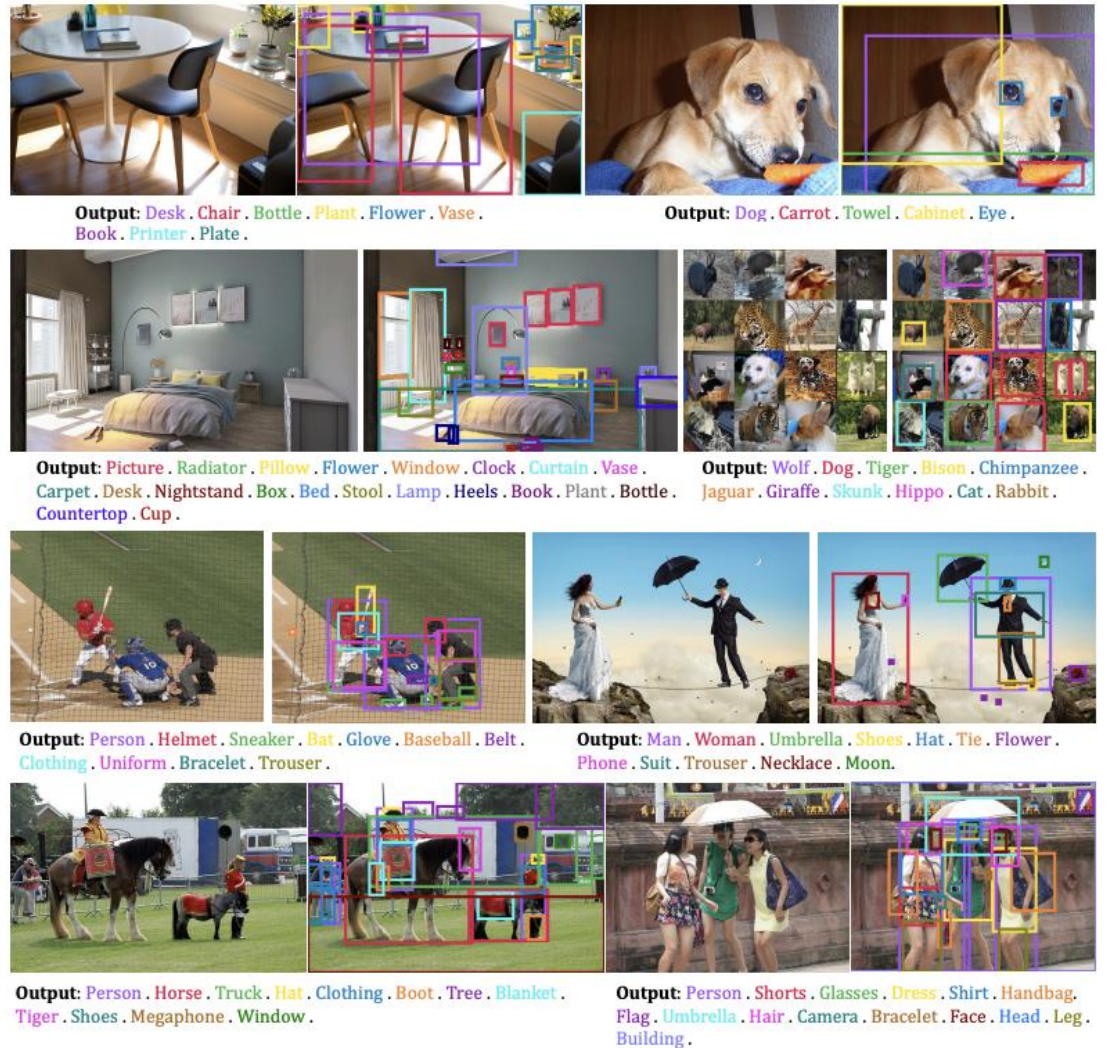


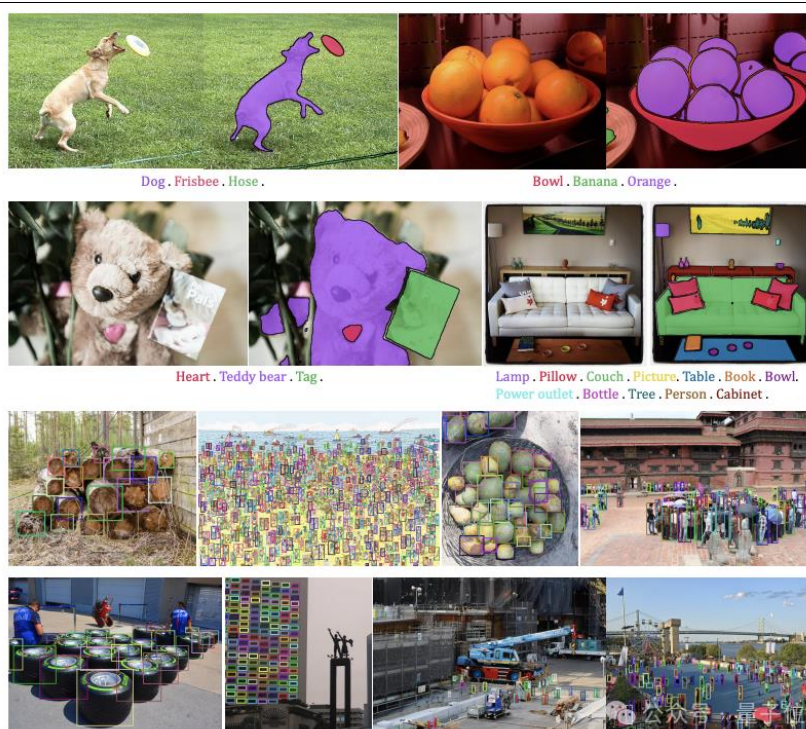
Figure 8: Prompt-free object detection and recognition with DINO-X · 量子位

资料来源：量子位，华鑫证券研究

2) 泛化和通用性。在面对未见过的物体或环境时，模型仍能保持高水平的检测性能。

3) 细粒度目标理解。DINO-X 通过统一多个视觉任务，实现了多样化输出，包括边界框、分割掩码、关键点和描述文本，提升了模型在复杂场景下的理解能力。

图表 6: 细粒度目标理解



资料来源: 量子位, 华鑫证券研究

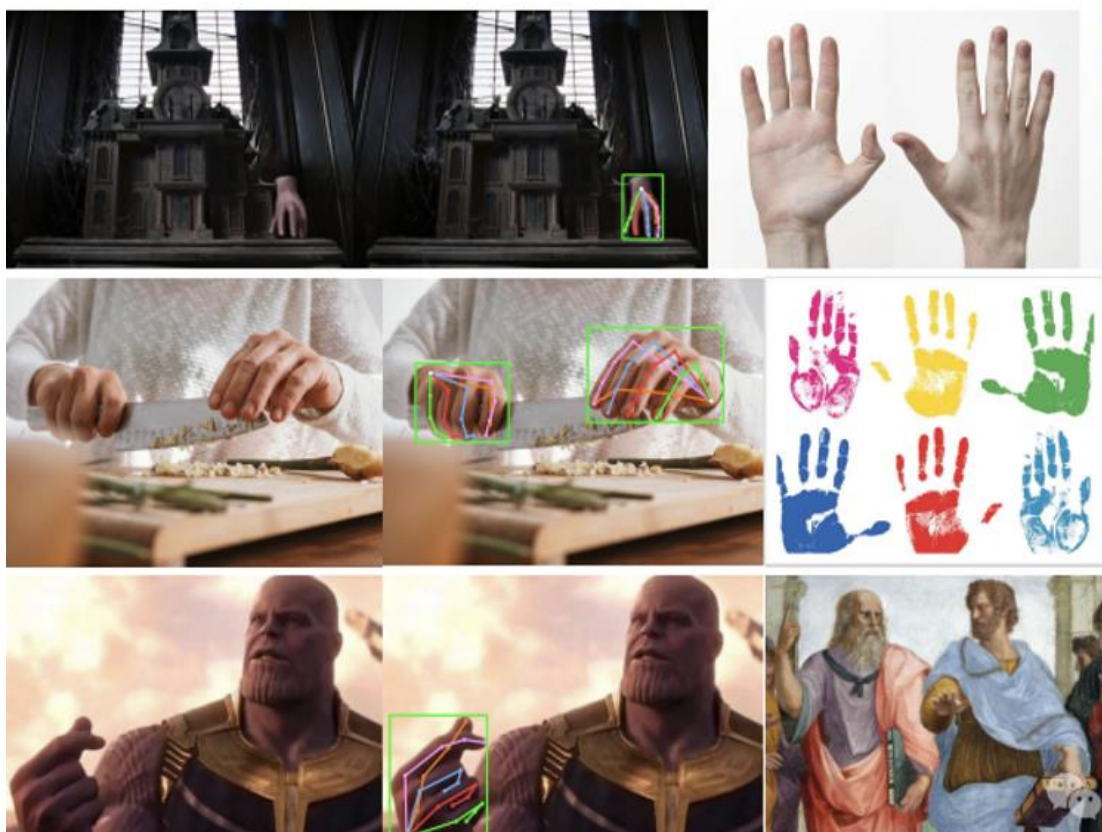
4) 多任务感知与理解。DINO-X 整合了多个感知头, 支持包括分割、姿态估计、区域描述和基于区域的问答在内的多种区域级别任务, 让感知到理解逐步成为了现实。

图表 7: 多任务感知与理解





Person.



资料来源：量子位，华鑫证券研究

**5) 长尾目标检测优化。**为了支持长尾目标的检测任务，DINO-X 不仅支持文本提示和视觉提示，还支持经过视觉提示优化的自定义提示。

跟之前的版本 GroundingDINO 1.5 Pro 和 Grounding DINO 1.6 Pro 相比，此次通用视觉大模型 DINO-X 进一步增强了语言理解能力，同时在密集物体检测场景中表现出色。

DINO-X 可接受文本提示、视觉提示和自定义提示，并能同时生成从粗略的表示（如边框）到精细的细节（包括遮罩、关键点和对象标题）等各种输出。DINO-X Pro 的核心架构，与 Grounding DINO 1.5 类似，利用预先训练好的 ViT 模型作为主要的视觉骨干，并在特征提取阶段采用了深度早期融合策略。但不同的是，他们扩大了 DINO-X Pro 在输入阶段的提示支持，除了文本，还支持视觉提示和自定义提示，以满足包括长尾物体在内的各种检测需求。而对于 DINO-X Edge 版本，他们利用 EfficientViT 作为高效特征提取的骨干，并采用了类似 Transformer 编码器-解码器架构。

此外，为了提高 DINO-X Edge 模型的性能和计算效率，他们还对模型结构和训练技术做了几个方面的改进：

1) **更强的文本提示编码器**：采用了与 pro 模型相同的 CLIP 文本编码器。

2) **知识提炼**：从 Pro 模型中提炼知识，以提 Edge 模型的性能。具体来说，团队利用基于特征的蒸馏和基于响应的蒸馏，分别调整 Edge 模型和 Pro 模型之间的特征和预测对数。

3) **改进 FP16 推理**：采用浮点乘法归一化技术，在不影响精度的情况下将模型量化为 FP16。

DINO-X 的万物识别能力，让其拥有了对开放世界（Open World）的视觉感知，可以轻松应对真实世界中的众多不确定性和开放环境。IDEA 研究院创院理事长沈向洋提到了具身智能、大规模多模态数据的自动标注、视障人士服务等这几个场景。

目前，IDEA 研究院与多家企业联合开展视觉大模型及应用场景的探索研究工作，在视觉大模型的落地方面取得了实质性进展。一方面，有别于市场上的以语言为基础的多模态大模型基于全图理解的方法，通过在物体级别的理解上加入语言模块优化大模型幻觉问题。另一方面结合自研的“视觉提示优化”方法，无需更改模型结构，不用重新训练模型，实现小样本下的场景化定制。

### 3、AI 融资动向：xAI 获 50 亿美元融资，目前估值达 500 亿美元

本期 AI 初创公司的融资中，xAI 融资额排名第一，获 50 亿美元融资。xAI 成立于 2023 年的 7 月 12 日，虽然成立的时间不长，但 xAI 在人工智能领域已有不错的进展，在成立次月的 18 日，他们就宣布完成了首个旗舰模型 Grok-0 的训练，有 330 亿参数；两个多月后的 11 月 3 日，他们宣布了 Grok-1 模型；在今年的 3 月 17 日，他们宣布将 Grok-1 的权重和架构开源；3 月 28 日宣布的 Grok-1.5 模型，增加了长文本能力；4 月 12 日推出的 Grok-1.5V，则是增加了视觉能力，可以根据交错的自然语言和图像进行推理；而在今年 8 月 15 日，他们推出了推理能力更强的大语言模型 Grok-2，更小但能力不俗的 Grok-2 mini 也在当天一并推出，Grok-2 和 Grok-2 mini 在测试中的部分能力，超过了 Anthropic 旗下的

Claude 3.5 Sonnet 和 OpenAI 旗下的 GPT-4 Turbo。xAI 截至今年已打造有 10 万块英伟达 H100 的算力集群，是当前全球最大的 AI 训练系统。

图表 8：本周 AI 初创公司的融资动态

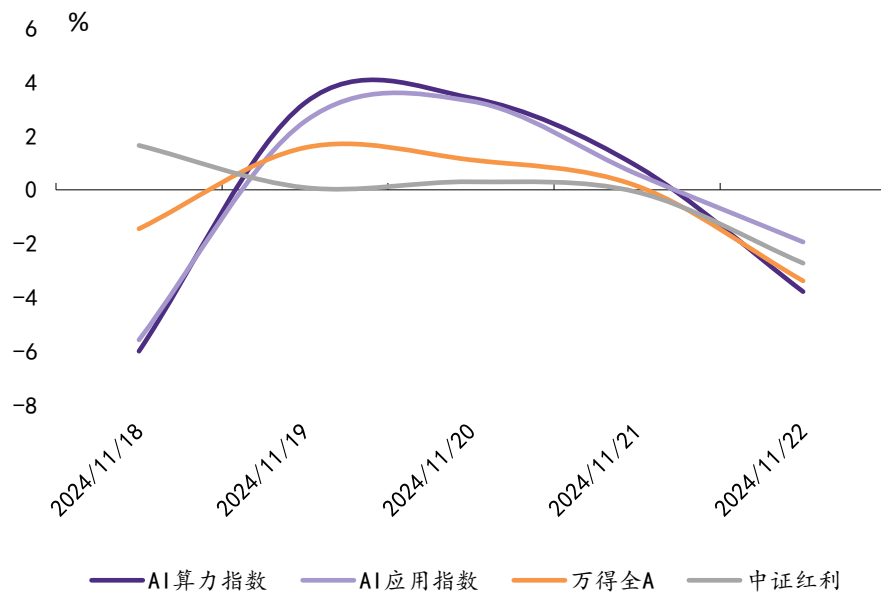
应用	应用类型	领投方	融资轮	融资额	目前累计融资额	目前估值
闪极科技	AI 硬件	光远投资	A 轮	数千万元	——	——
xAI	Grok 系列模型	卡塔尔投资局、红杉资本	——	50 亿美元	110 亿美元	500 亿美元
Gyges Labs	穿戴设备	金沙江创投	Pre-A 轮	数千万元	——	——

资料来源：Techweb，腾讯网，华鑫证券研究

## 4、行情复盘

本周，AI 算力指数/AI 应用指数/万得全 A/中证红利日涨幅最大值分别为 3.42%/3.29%/1.57%/1.65%，日跌幅最大值分别为-5.99%/-5.58%/-3.38%/-2.73%。AI 算力指数内部，依米康以+23.54%录得本周最大涨幅，天玑科技以-24.07%录得本周最大跌幅。AI 应用指数内部，国光电器以+45.79%得本周最大涨幅，因赛集团以-14.79%录得本周最大跌幅。

图表 9：本周指数日涨跌幅



资料来源: wind, 华鑫证券研究

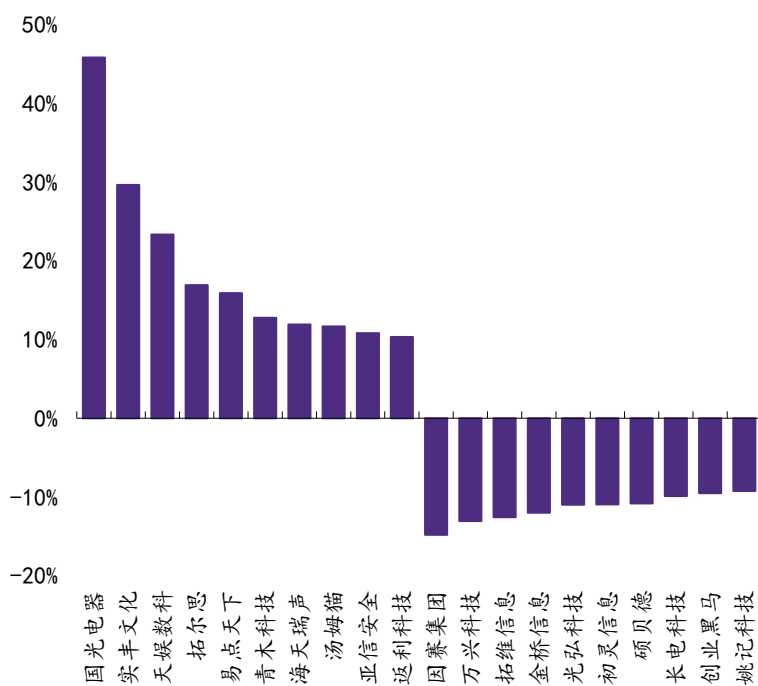
图表 10: 本周 AI 算力指数内部涨跌幅度排名



资料来源: wind, 华鑫证券研究

图表 11: 本周 AI 应用指数内部涨跌幅度排名





资料来源：wind, 华鑫证券研究

## 5、投资建议

海外应用侧最重要场景已具雏形，国内政策持续支持。海外方面，当地时间 11 月 19 日，在“Microsoft ignite 2024”全球开发者大会上，围绕 AI Agent/AI 智能体，微软更新了多款 AI 产品。微软的目的在于建立完善的 AI 智能体生态，自推出以来，该生态联盟已有超过 10 万家组织通过其 Copilot Studio 创建或编辑过 AI 助理。国内方面，近日，人工智能辅助诊断首次被列入价格构成。国家医保局分析人工智能潜在的应用场景，在放射检查、超声检查、康复类项目中设立“人工智能辅助”扩展项，即同样的价格水平下，医院可以选择培养医务人员进行诊疗，也可以选择使用人工智能参与诊疗行为，但现阶段不重复收费。预计该政策将大幅加快人工智能辅助技术进入临床应用速度。

我们坚定持续看好 AI 软硬件机会，建议关注以 AI 为核心的龙头厂商科大讯飞（002230.SZ）、有望迎来需求爆发的 AI 应用金桥信息（603918.SH）、高速通信连接器业务或显著受益于 GB200 放量的鼎通科技（688668.SH）。

图表 12：重点关注公司及盈利预测

公司代码	名称	2024-11-26 股价	EPS			PE			投资评级
			2023	2024E	2025E	2023	2024E	2025E	
002230.SZ	科大讯飞	47.04	0.28	0.40	0.56	168.00	117.60	84.00	买入
688256.SH	寒武纪-U	485.90	-2.04	-1.21	-0.50	-238.19	-401.57	-971.80	买入
688668.SH	鼎通科技	35.80	0.67	1.04	1.41	53.43	34.42	25.39	买入

资料来源：wind, 华鑫证券研究

## 6、风险提示

1) AI 底层技术迭代速度不及预期。2) 政策监管及版权风险。3) AI 应用落地效果不及预期。4) 推荐公司业绩不及预期风险。

## ■ 计算机&中小盘组介绍

**宝幼琛：**本硕毕业于上海交通大学，多次新财富、水晶球最佳分析师团队成员，7年证券从业经验，2021年11月加盟华鑫证券研究所，目前主要负责计算机与中小盘行业上市公司研究。擅长领域包括：云计算、网络安全、人工智能、区块链等。

**任春阳：**华东师范大学经济学硕士，6年证券行业经验，2021年11月加盟华鑫证券研究所，从事计算机与中小盘行业上市公司研究

**周文龙：**澳大利亚莫纳什大学金融硕士

**陶欣怡：**毕业于上海交通大学，于2023年10月加入团队。

## ■ 证券分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

## ■ 证券投资评级说明

股票投资评级说明：

	投资建议	预测个股相对同期证券市场代表性指数涨幅
1	买入	> 20%
2	增持	10% — 20%
3	中性	-10% — 10%
4	卖出	< -10%

行业投资评级说明：

	投资建议	行业指数相对同期证券市场代表性指数涨幅
1	推荐	> 10%
2	中性	-10% — 10%
3	回避	< -10%

以报告日后的12个月内，预测个股或行业指数相对于相关证券市场主要指数的涨跌幅为标准。

**相关证券市场代表性指数说明：**A股市场以沪深300指数为基准；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）为基准；香港市场以恒生指数为基准；美国市场以道琼斯指数为基准。

## ■ 免责声明

华鑫证券有限责任公司（以下简称“华鑫证券”）具有中国证监会核准的证券投资咨询业务资格。本报告由华鑫证券制作，仅供华鑫证券的客户使用。本公司不会因接收人收到本报告而视其为客户。

本报告中的信息均来源于公开资料，华鑫证券研究部门及相关研究人员力求准确可靠，但对这些信息的准确性及完整性不作任何保证。我们已力求报告内容客观、公正，但报告中的信息与所表达的观点不构成所述证券买卖的出价或询价的依据，该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。投资者应当对本报告中的信息和意见进行独立评估，并应同时结合各自的投资目的、财务状况和特定需求，必要时就财务、法律、商业、税收等方面咨询专业顾问的意见。对依据或者使用本报告所造成的一切后果，华鑫证券及/或其关联人员均不承担任何法律责任。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露。

本报告中的资料、意见、预测均只反映报告初次发布时的判断，可能会随时调整。该等意见、评估及预测无需通知即可随时更改。在不同时期，华鑫证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。华鑫证券没有将此意见及建议向报告所有接收者进行更新的义务。

本报告版权仅为华鑫证券所有，未经华鑫证券书面授权，任何机构和个人不得以任何形式刊载、翻版、复制、发布、转发或引用本报告的任何部分。若华鑫证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，华鑫证券对此等行为不承担任何责任。本报告同时不构成华鑫证券向发送本报告的机构之客户提供的投资建议。如未经华鑫证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。华鑫证券将保留随时追究其法律责任的权利。请投资者慎重使用未经授权刊载或者转发的华鑫证券研究报告。