

文化和科技融合热点观察

总第 23 期

2023 年 4 月号

本期热点：AIGC 在图像生成领域的应用

海外视野

Stability AI

Midjourney

OpenAI

国内风向

百度

商汤科技

万兴科技

西湖心辰（杭州）科技有限公司

北京计算美学科技有限公司

上海退格数字科技有限公司

中山大学人机物智能融合实验室

协会之窗

南京芯视元电子有限公司



关注“文化科技融合汇”
获取产业发展前沿资讯



扫描二维码，在线填写
读者满意度调查问卷

主办单位

中国（南京）文化和科技融合成果展览交易会组委会
紫金文化科技融合创新研究院
南京元宇宙产业协会

支持单位

南京市文化投资控股集团有限责任公司
中国人民大学创意产业技术研究院

出品单位

江苏文创文化发展有限公司

本期导读

本期我们着眼于 AIGC 在图像领域的应用，探讨国内外企业在 AI 图像领域的发展状况。自去年 8 月 Stable Diffusion 面世，AI 图像生成技术迎来了质的飞跃。秉承“开源”的精神，SD 不仅在短短两个月时间内让其背后公司 Stability AI 一跃成为市值高达十几亿美金的独角兽，更吸引了全球数上千万人积极参与其生态的建设。

Stability AI 开放了包括核心 AI 算法模型（Latent）、开放核心训练数据集（LAION）、开放 AI 生成图片的版权（CC0）。这种三位一体的开源模式，让全世界所有普通人、极客、创业者、商业团体可以随心所欲地完成对 SD 的部署、运行、改进和商业化，且大胆使用全世界人利用 SD 生成的图片，而不必受制于版权的束缚，开源生态加速了 AI 技术在视觉生成领域的发展。

早在 2022 年，AI 绘画便在社交平台上引起过一定的关注，但由于 AI 绘画的开发成本较高、生成速度较慢，没有得到大规模的市场推广。直到 Stable Diffusion 模型开源后，行业获得了关键性技术的突破，许多基于 Stable Diffusion 模型的应用才纷纷入局。以此为基础，国内文生图行业的数据训练也有了一定的发展。一方面，创新企业有了底层模型支持；另一方面这种模型可在消费级 GPU 上运行，极大降低了研发成本、提高了产品推向市场的速度。

本期热点观察，我们通过对比研究、对话访谈、实践体验等方式，深入探索国内 AI 图像的软件以及其背后公司主体的发展情况。在海外视野板块我们将简要介绍国外最热门的三款应用：Stable Diffusion、Midjourney、DALL.E2。国内风向板块，我们根据公司发展、底层技术两个维度选择了 6 家企业，包括三家上市公司和三家初创公司。主要关注公司的底层技术、应用本身的成熟度、商业模式和公司发展状况。此外，本期还关注到高校在 AIGC 图像领域的创新研究。

目 录

引 言

什么是 AIGC3

AIGC 基本工作原理3

图片生成（AI 绘画）核心模型-Diffusion Model4

海外视野

Stability AI 5

Midjourney8

OpenAI 19

DALL · E 2, Stable Diffusion 和 Midjourney 之间的比较 24

国内风向

百度 27

商汤科技30

万兴科技集团股份有限公司 33

西湖心辰杭州科技有限公司 30

北京计算美学科技有限公司 30

上海退格数字科技有限公司 43

中山大学人机物智能融合实验室 45

小 结49

协会之窗

南京芯视元电子有限公司 52

引言

一、什么是 AIGC

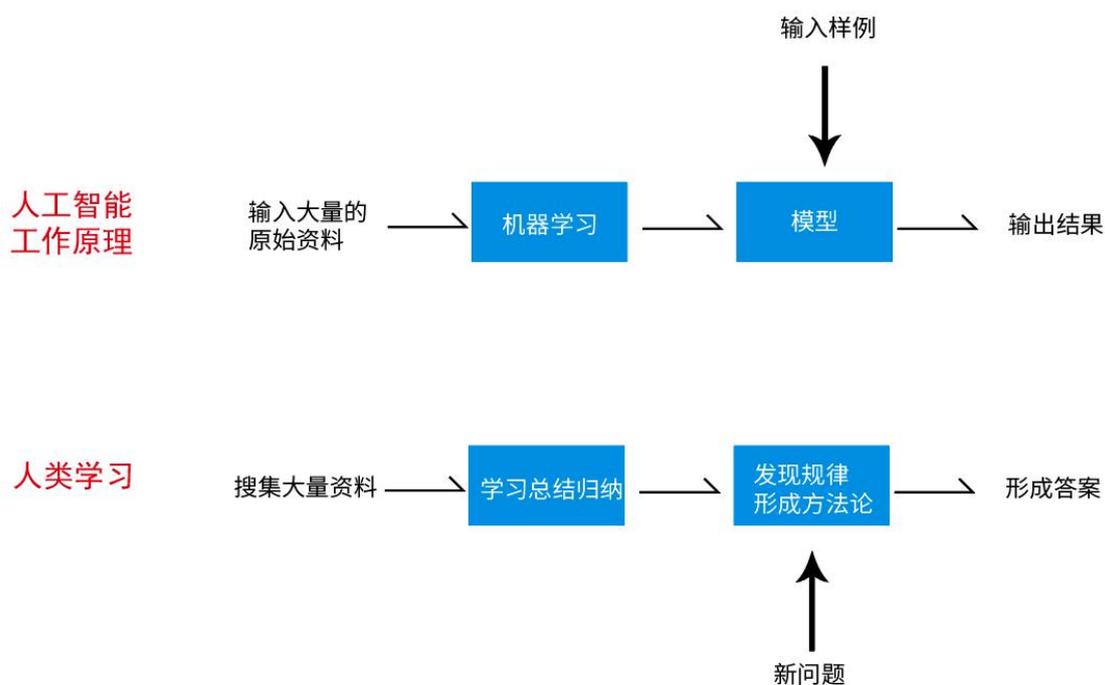
AIGC，全名“AI generated content”，又称生成式 AI，意为人工智能生成内容。AIGC 主要是人工智能(非人类)生成内容,这就区别于过去普通用户(UGC)和某一领域专业用户(PGC)生产内容。

二、AIGC 基本工作原理

AIGC 底层主要依赖的是 AI 技术，AI 技术本质主要是让机器拥有像人类一样的智能，所以需要让机器能够像人类一些学习和思考，所以目前大部分实现 AI 的底层技术叫做“机器学习”(Machine Learning)技术。机器学习是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域。

人类学习时，也是依据大量的原始资料，通过“学习总结归纳”，发现事物的各种规律形成“知识经验方法论”，当遇到新问题时就会运用这些“知识经验方法论”作出相应的反应决策(预测推理)；

人工智能是模拟人类的学习方式，给计算机输入大量的“资料”，然后通过机器学习(总结归纳抽取相似点)，最后形成了“模型”(知识经验方法论)，然后在遇到一些需要判断决策的时候，我们会把要判断决策的事物给“模型”，然后就会告诉我们输出结果(推理推测结果)；



上面这个过程我们可以简单理解为：“模型”就是一只小狗，饲养员就是那个“机器学习算法”，饲养员通过一些指令和奖惩措施反复训练小狗学习，小狗

就会学会一些技能（模型），一旦学会了，小狗就可以出去表演，表演的过程就是预测。

三、图片生成（AI 绘画）核心模型-Diffusion Model

在图片生成领域最近几年有四大主流生成模型：生成对抗模型（GAN）、变分自动编码器（VAE）、流模型（Flow based Model）、扩散模型（Diffusion Model）等，基本主要都是基于深度学习为训练方式的模型。从 2022 年开始，主要爆火的图片生成模型是 Diffusion Model（扩散模型）为主。

在计算机视觉中，生成模型是一类能够生成合成图像的模型。例如，一个被训练来生成人脸的模型，每次都会生成一张从未被该模型或任何人看到过的人脸。生成模型最著名的例子是 GAN（生成对抗网络）。它有生成器和鉴别器，它们相互对抗，然后生成图像。由于模型本身具有对抗性，因此很难进行训练。这使得很难达到一个最优的平衡。利用扩散模型可以解决这个问题。

扩散模型也是生成模型，扩散模型背后的直觉来源于物理学。在物理学中气体分子从高浓度区域扩散到低浓度区域，这与由于噪声的干扰而导致的信息丢失是相似的。所以通过引入噪声，然后尝试通过去噪来生成图像。在一段时间内通过多次迭代，模型每次在给定一些噪声输入的情况下学习生成新图像。扩散模型的工作原理是学习由于噪声引起的信息衰减，然后使用学习到的模式来生成图像。

目前国内外的图片生成技术主要都是基于 Diffusion Model（扩散模型）来进行的实现，包括不限于 Stable Diffusion、MidJourney、DALL.E、Disco Diffusion、Google Imagen 等主流产品，但是实际技术在处理方面又各有区别，也导致会有不同的表现形态。这种差异性的原因包括两点：一方面是底层模型训练的图片语料不同，另外一个方面是一些算法方面的微调区别。

Diffusion 是一种非稳态的生成模型，核心就是通过不停去除噪音期望获得好结果的生成模型。早期的扩散模型在 AI 绘画中效果不好，而且单张图生成需要 10-15 分钟，后来英国的 Stability AI 公司对模型进行了改进，把图片生成稳定性和质量都大幅提高，图片生成速度提高了 100 倍，意味着以前需要 10-15 分钟现在只需要 6-10 秒就可以生成一张图。Stability AI 把这种稳态的 Diffusion 模型叫做：“Stable Diffusion”，然后 Stable Diffusion 模型形态基本影响了包括 MidJourney 在内的众多 AI 绘画产品。

海外视野

Stability AI

Stability AI 是生成式 AI 行业中最知名的公司之一，发布了免费的开源产品 Stable Diffusion，平衡了计算能力、资金和公众之间的关系，成功解决了 AI 技术封闭的问题。Stability AI 是一项开源技术，用户可以在其代码的基础上构建与设计、电影、增强现实、视频游戏、广告甚至电子商务相关的应用程序。通过这套开源技术，Stability AI 的社区已经创造出了几乎涵盖所有媒体内容形态的 AI 模型，包括图像、语言、音频、视频以及 3D 内容。

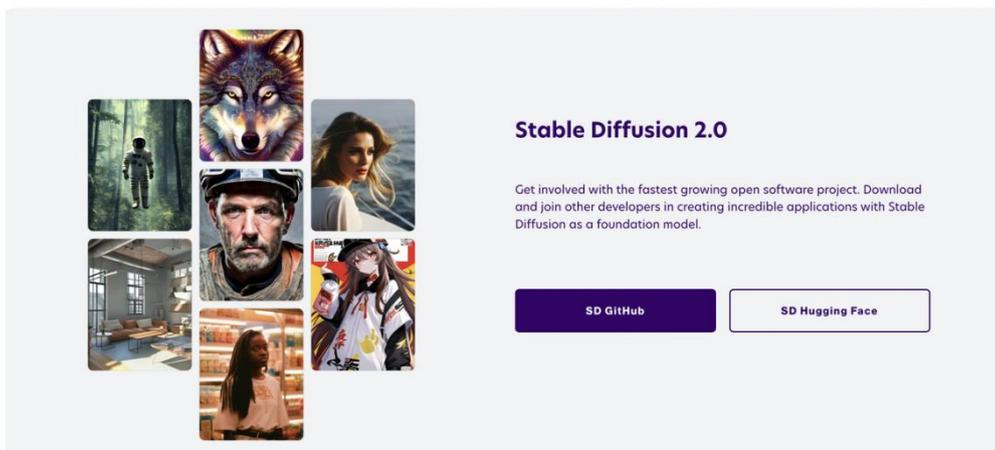
Stability AI 在训练自己的模型时，也花费了很高的成本。Stability AI 在运营和云服务方面的支出超过了 5000 万美元。公司在训练方面耗资巨大，此前微软对 OpenAI 的 10 亿美元投资，其中一半是用云计算等服务折算的。

Stability AI 并未聚焦于某个单一场景，而是通过开源来拓展场景，再去寻找具体的商业模式。当然，更大的想象空间，也需要更多的投入，所以 Stability AI 在种子轮就拿了这么多钱。

Stability 建立的社区，包括了全球各地的专家与合作伙伴，他们正在为图像、语言、音频、视频、3D 和生物学等，开发最先进的开放式人工智能模型。

Open Source Models

We put control in your hands with **open source** models



一、公司发展

Stability AI 创始人是 Emad Mostaque，他曾在多家科技公司担任工程师、分析师和投资官。2019 年，Emad Mostaque 基于 Open AI 的理念创立了 Stability AI。

其非营利模式与 Open AI 相同，但更加开放。

Stability AI 为不懂代码的用户提供无代码网站。凭借稀缺的开源特性和用户友好的体验，该项目一经公开测试就受到了大量研究人员和用户的欢迎，积累了可观的用户群体。同时，随着人脉的积累，Stability AI 在整个 AI 研究和应用领域（包括学术和工业领域）都得到了广泛的关注和认可。

最初，Stability AI 完成了 1000 万美元融资，估值 1 亿美元。2022 年 10 月 18 日，Stability AI 宣布获得 Coatue 和 Lightspeed 的 1.01 亿美元投资，估值将达到 10 亿美元，成为新晋独角兽。Stability AI 表示，公司计划利用这笔融资继续投资在研发，并将公司规模从 100 家扩大到 300 家。除了 Stable Diffusion，该公司正在开发用于生成语言、音频、视频和 3D 的 AI 生成模型，为未来的内容创作提供更多开源模型。

2023 年 3 月，据知情人士透露，Stability 已经烧掉了去年底筹集的 1 亿美元中的大量资金。两位风险投资人正在重新考虑是否参加一轮融资活动，该轮融资将把公司估值增加到 40 亿美元，但这些计划仍处于保密状态。

二、产品介绍

Stability AI 目前发布了多款产品，他们的功能各不相同（详细介绍如下表所示），而其中最热门、影响范围最广的便是 AI 技术模型 Stable Diffusion。

其核心技术来源于 AI 视频剪辑技术创业公司 Runway 的 Patrick Esser，以及慕尼黑大学机器视觉学习组的 Robin Romabach。该项目的技术基础主要来自这两位开发者之前在计算机视觉大会 CVPR22 上合作发表的潜在扩散模型（Latent Diffusion Model）研究。另外，项目也得到了一些外部开发社区，以及 stabilityAI 机构生成技术团队的支持。Stable Diffusion 模型支持多种操作，包括以文本生成图片和以图文结合的方式生成图片（文生图、图生图）。

相比较于 DALL-E 等大模型，Stable Diffusion 的特点是：

1、在消费级的显卡上迅速实现文生图——Stable Diffusion 能够在消费级显卡上实现 DALL-E 2 级别的图像生成，生成效率却提高了 30 倍。

2、Stable Diffusion 完全免费开源，所有代码均在 GitHub 上公开，任何人都可以拷贝使用。Stable Diffusion 已经有超过 20 万开发者下载和获得授权，各渠道累计日活用户超过 1000 万。

3、在生成图像方面，它在所生成图像的质量、速度和成本上都取得了巨大突破。

产品	功能	价格
Stable Diffusion	基于扩散模型的文本到图像生成器,可以根据用户输入的文本生成高分辨率、逼真和多样化的图像。	免费开源,任何人都可以下载并使用。
DreamStudio	在线生成式文本到图像套件,对生成的图像进行编辑、调整和保存。	免费试用,每月可生成 10 张图像。付费版,每月 9.99 美元,可无限制使用。
Platform API	软件开发工具包,可以在其他应用中 Stability AI 提供的模型。支持多种语言和平台。	免费试用,每月可调用 100 次 API。付费版,每月 49.99 美元,可调用 1000 次 API。
Photoshop Plugin	可以直接在 Photoshop 中输入文本,插件会生成与之匹配的图像,同时对生成的图像进行进一步编辑或合成。	免费试用,每月可生成 10 张图像。付费版,一次性支付 19.99 美元,可无限制使用。
Blender Plugin	可以在 Blender 中使用 Stability AI 提供的模型,为 3D 建模和渲染添加 AI 魔力。生成 AI 视频或 3D 模型。	免费试用,每月可生成 10 张纹理、图像或动画。付费版,一次性支付 29.99 美元,可无限制使用。

Midjourney

Midjourney 是一家专注于设计、人工智能等方向的独立研究室，团队已开发同名文生图模型及应用 Midjourney，该产品搭载于 Discord，用户通过与 Midjourney bot 进行对话式交互，提交 Prompt（文本指令）来快速获得想要的图片。

一、公司简介

Midjourney 正式团队成员十分精简，共 11 人。除了 CEO 之外，有 8 位研究人员与工程师，2 位财务与法务。除此之外，还 4 位编外顾问，以及 60 位兼职 Discord 运营。CEO David Holz 认为 AI 不应该被限制，文生图的核心是人类想象力的引擎，这样的思维方式使得 Midjourney 的风格充满了科幻色彩。

Midjourney 在 AI 公司价值链上占据了多个位置，拥有数据层、模型层、应用层整个技术栈。它参考 CLIP 及 Diffusion 构建了自己的闭源模型，抓取公开数据进行训练，并构建了 Discord 中的 Midjourney bot 应用。作为应用层公司，Midjourney 收集用户反馈数据，迭加技术的进步，不断迭代模型。

二、核心产品

Midjourney 是 Midjourney 研究实验室开发的人工智能程序，于 2022 年 7 月 12 日进入公开测试阶段，使用者可通过 Discord 的机器人指令进行操作，可以创作出很多的图像作品。

Midjourney 一直在努力改进其算法，并每隔几个月发布新的模型版本。2022 年 4 月推出第二版算法，7 月 25 日发布第三版，11 月 5 日发布第四版的 alpha 迭代版，第五版的 alpha 迭代版也已经于 2023 年 3 月 15 日发布。

（一）用户群体

Midjourney 用户群体广阔，主要包括创意设计人群、工业设计人群、Web3 & NFT 从业者以及个人爱好者。

1、创意设计人群

创意设计从业者是一个很大的群体，渗透各行各业。

小 b 端，包括产品设计师（如玩具、墙纸等）；图片设计师（网站、广告、PPT、Logo、插图等）；游戏设计师（游戏场景、角色、道具）以及自媒体创作者等等。

企业端，包括广告公司、影视公司、品牌的广告创意部门等对于艺术效果图

有大量需求的客户。



由 Midjourney 创作的食品广告



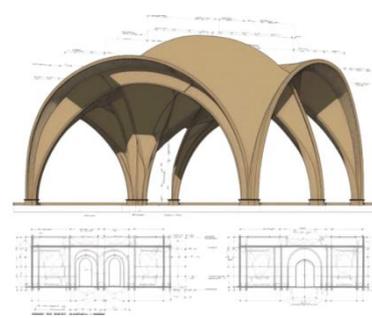
由 Midjourney 创作的角色扮演动物角色

2、工业设计行业——给设计师提供灵感

Midjourney 也进入了工业设计等精细化要求较高的领域，如建筑设计。目前 Instagram 上有许多建筑师分享文生图作品，有超过 72,000 个帖子被标记为 #midjourneyarchitecture。

在工作中，建筑师使用 Midjourney 在项目的最初阶段创建情绪板（指一系列图像、文字或样品的拼贴组合物，用来展现设计师对一个项目的想法或感觉）。目前 Midjourney 生成的图像仅能作为草图，用来激发灵感。建筑师将这些草图

芝加哥建筑师斯蒂芬·库拉斯（Stephen Coorlas）使用 Midjourney 为芝加哥北岸地区设计的露天展馆。基于这个概念图像进行简单的 3D 建模后，库拉斯将这些模型的图像反馈给 Midjourney，并令其迭代建筑图纸。然后就得到了这个展馆的建筑图纸，虽然还缺乏细节。



扎哈·哈迪德建筑伦敦工作室的计算研究小组（ZHACODE）设计师 Tim Fu，使用 Midjourney 设计的钢琴家的住宅：翻译成图纸，并建模和进行结构分析，开发出 3D 模型后，建筑师会再将图像反馈给 Midjourney，进一步迭代建筑图纸。不过近期发布的 ControlNet 将会进一步深入设计 workflow。

3、NFT 从业者

NFT 无疑也是文生图的一大应用场景。Midjourney 目前被广泛应用在 NFT 创作上，因此公司发布规定，如果在“与区块链相关的事物”中使用 Midjourney 生成图像，需要对每月超过 20,000 美元的收入部分支付 20% 的版税。



Midjourney 创建的 NFT

4、个人爱好者

Midjourney 用户群体中不乏个人爱好者，如自媒体群体。文生图大大降低了艺术创作门槛，使得普通用户也可以成为艺术家和设计师，并通过 AI 创作获取收入。

(二) 主要功能

Midjourney 通过简单的文字交互实现“文生图”及“图生图”。Midjourney 既可以输入文本描述，实现“文生图”；又可以输入图片链接并加入文字描述，实现“图生图”。具体来说一下这些功能它都可以实现：

- (1) 人像卡通化：将人物照片转换成有趣的卡通形象。
- (2) 轮廓生成：可以根据输入的文字描述生成对应的图像轮廓。
- (3) 色彩生成：利用 GAN 模型，可以生成具有艺术感和创意的彩色图像。
- (4) 视频换脸：可以实现对视频中人物的面部进行换脸操作。
- (5) 视觉问答：通过图像识别和自然语言处理技术，可以实现对图像的问题回答。
- (6) 人脸合成：可以将不同人物的面部特征进行合成，生成新的面孔。

（三）发展前景

Midjourney 未来可能会有以下几种发展前景：

（1）将目前的图库功能做到极致。

目前 Midjourney 替传统图库平台，成为新的“视觉搜索引擎”。Pinterest 拥有超 4.5 亿月活用户，若其中的 1% 能成为 Midjourney 的付费用户，则能达到约 10 亿美元的营收；乐观来看，若 5% 能成为付费用户，则能达到约 47 亿美元的营收。

（2）不再局限于 Discord 平台

Midjourney 可以作为插件接入其他设计产品，仍扮演图库功能。Canva、Adob 等等，这将进一步增加其分销渠道，触达更多的付费用户群。

（3）自主开发端到端的产品

通过不断集成最新的技术，迭代产品功能，打造产品矩阵，通过增加产品形态来为更广阔的客户群提供服务。可以和设计 workflow 紧密结合，进一步提高客单价。

（4）作为功能单一的上层应用，被技术的快速进步颠覆

Apple 计划将 Stable Diffusion 嵌入 Mac 和 Iphone 中，作为用户熟悉的平台，且使用 Apple 终端推理成本更低、速度更快，也会对 Midjourney 构成挑战。Midjourney 可能会被集成，变成多功能产品的一个环节，甚至完全被新产品取代。不过 Apple 终端也并非完美的解决方案，API 的形式将无法对垂直领域的模型进行特定化训练；将应用放在本地也面临着用户数据收集及进一步迭代模型的困难。

虽目前无法看清该行业的终局，但作为 ToC 属性很强的产品，我们认为 Midjourney 凭借着独特的图片风格及艺术效果、活跃的 Discord 社群带来的绝佳用户体验，在一定时间内能够保持用户粘性和较强的盈利能力。

三、行业案例

（一）网易严选 AIGC 小组——Midjourney 助力品牌设计

网易严选团队已经通过 ChatGPT 和 Midjourney 辅助完成项目的全流程，涉及 AIGC 在商品、品牌、营销和体验等多个领域的应用。其中 Midjourney 在其中发挥的作用包括：

1、品牌调性图

品牌传播需要符合品牌特色和气质的图像，从图片网站上查找素材很难找到与品牌完全匹配的风格感受，给品牌传播带来困难，企业不得不花更多的时间和金钱进行实景拍摄。现在，使用 Midjourney 生成符合品牌调性的图像，可以节

省较多的查询时间和采购成本。



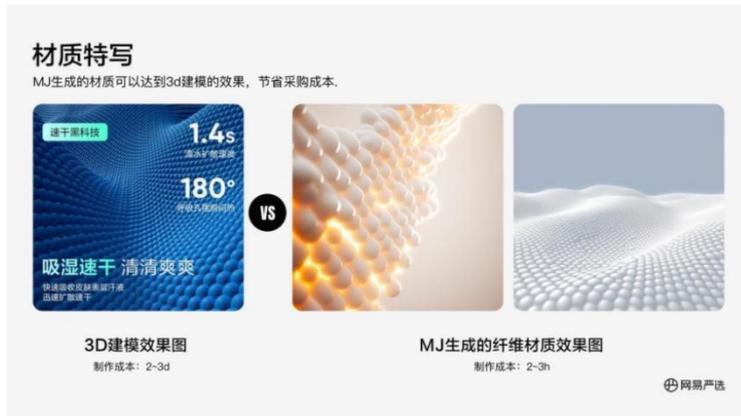
2、商品氛围图素材

在实际拍摄环节中，需要搭建空镜场景来拍摄所需的图像。现在使用 Midjourney 工具可以快速生成同样主题的空镜场景图片，并且其精细程度符合使用需求。可以节省搭建场景的时间和成本，提高生产效率。



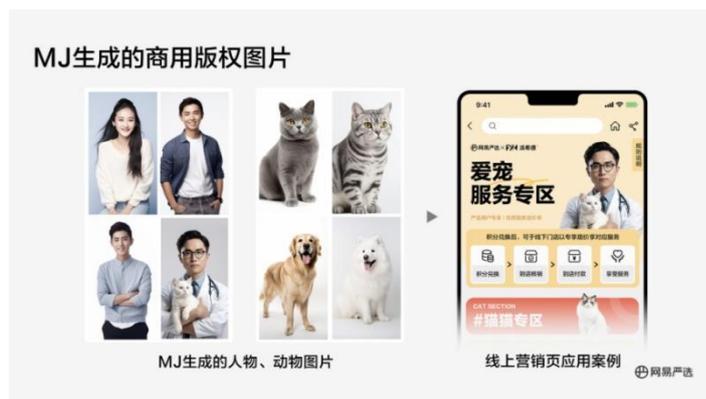
3、材质特写

在服装品类中需要展示材质细节的高清图片，通常需要进行 3D 建模和渲染制作（均价 800-1200 元/张）。现在使用 Midjourney 工具可以生成部分材质图片。节省制作时间和成本，提高生产效率。



4、商用版权图片

作为电商平台，每年需要采购大量商用版权图片用于营销素材（均价 1500 元/张）。现在使用 Midjourney 工具可以生成无限所需的多样性素材。大幅降低图片采购成本，提高生产效率。



5、服饰模特换装

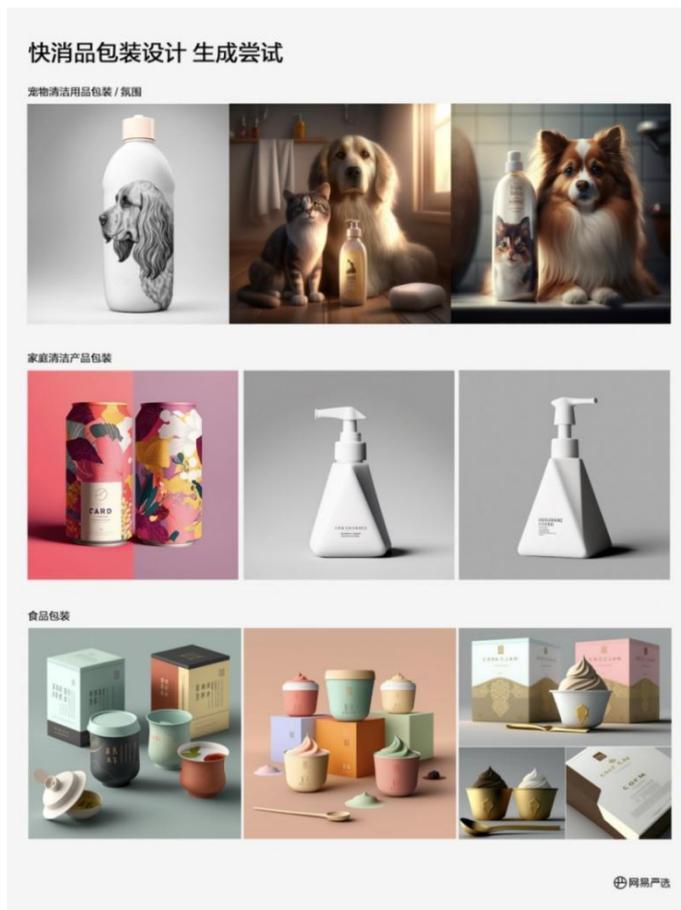
服饰品类需要大量模特拍摄素材，通过 Midjourney 也可以解决人物换装的问题。目前商品还原度并未能达到 100%，但是已具备相关可行性。“虚拟人物+服饰的换装思路”可以为降低不少的模特和拍摄成本。



除此之外，还有该小组还使用 Midjourney 辅助进行了包装设计、工业造型设计、商品图案设计以及 IP 设计。

1、包装设计

AI 帮助设计师提供容器造型，平面设计和背景等方面的创意效果。如下图，快销品包装设计师，在添加关键词后，生成了不同器型、平面及相关氛围的设计方向。



2、工业造型设计

AI 帮助设计师快速验证创意效果。以办公家具为例，以往需要 3D 软件花多日实现的模型，现在也许只需要 1 天。虽然还不能呈现落地方案，但已为初期沟通设计方向提供了有效素材。

本土新消费生活品设计师希望将郁金香的造型语言与人体工学椅相结合，得到独特的造型观感。使用 Midjourney 工具进行辅助设计，得到丰富的设计灵感和方向。此外，还可以测试不同主题风格的场景感受，以确定哪种风格最符合期望传达给用户的产品氛围。

工学椅造型设计 生成尝试



网易严选

3、商品图案设计

AI 帮助设计师生成图案并应用到商品上，以餐厨产品为例，可以快速产出图案创意方向及场景图。

餐具图案设计 AI/人工设计效果对比



AI生成效果



人工设计效果

网易严选

4、IP 设计

AI 帮助设计师提高搜索素材和建模渲染的效率。以 IP 设计为例，设计师将前期的风格调性&内容作为相应的 prompt 描述，即可在 Midjourney 中生产大量的素材。通过 Midjourney 产出的创意方案，极大地缩短建模和渲染的过程，快速聚焦理想的方向。

通过设计师的二次创作，便可直接应用于海报、包装、产品等场景丰富需求的提案内容。从设计创意探索到方案的落地过程缩减人力成本并提高设计提案的时效性。“网易天成”宠物 IP 的应用案例，如下图所示。



总的来说，使用 Midjourney 可以让素材场景更丰富多样，并且数量没有上限。而且经过设计师排版后，整体可以达到上线的要求。因此，AI 生成图代替拍摄素材的想法具有可行性。这将极大提高商品摄影内容的生产效率和创意，缩短生产周期。由此可预见在虚拟流媒体内容生产（如虚拟模特、虚拟直播、短视频）、语音对话式交互导购等方面，将会有颠覆性的改变。

（二）Midjourney 辅助游戏原画设计

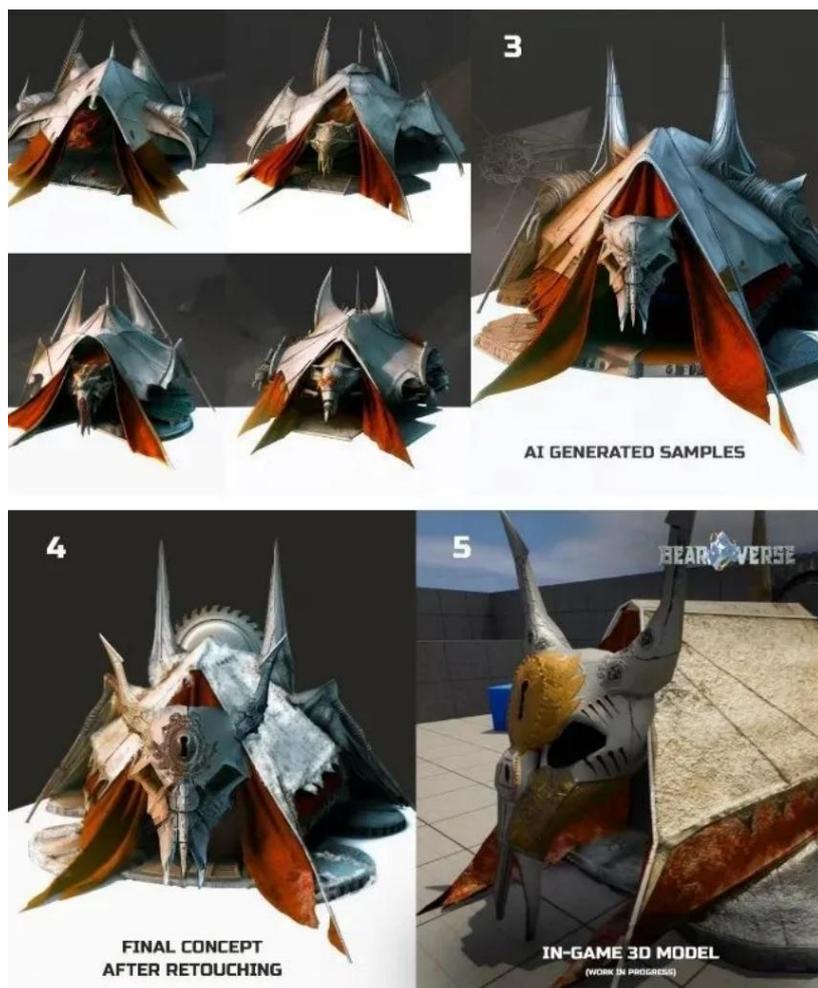
俄罗斯游戏工作室 Lost Lore 应用 Midjourney 创作角色，节省成本及时长超 80%。Lost Lore 在开发名为《Bearverse》的手机游戏时，使用 Midjourney 进行角色设计后，将开发成本从 5 万美元压缩至 1 万美元，将工时耗费从 6 个月减少到 1 个月。该游戏工作室直言：采用 AI 画图节省了 80% 美术成本。下图便是 Midjourney 根据指令所创作出的游戏角色。



团队使用 Midjourney 生成游戏角色之后进行人工调整优化。以下图为例，该团队的美术成员修改了不正确的武器、眼睛和爪子形态，纠正了阴影问题，并加入了大量的美术细节。



就像这样，以 AI 生成图为蓝本，在此基础上让人类画师进一步修改并最终拍板，最终能够大幅节省前期创意时间。此外，《Bearverse》中的部分 3D 建筑概念图也是由 AI 生成的。他们的团队向 Midjourney 输入帐篷的三视图，由系统生成一张主概念图，再对其进行人工监修，最终成为 3D 建模的参考概念图。



如今来看，作为最早接纳 AI 画图的一批公司，《Bearverse》的案例显然还仅仅是涉及到了 AI 画图的初阶使用，而即便是最“古早”的 AI 生成方式，已经足以帮助一家中小型工作室对开发进行大幅降本提效，这足见 AI 技术的潜力无穷。

OpenAI

OpenAI 是全球最著名的人工智能研究机构，发布了许多著名的人工智能技术和成果，如大语言模型 GPT 系列、文本生成图片预训练模型 DALL-E 系列、语音识别模型 Whisper 系列等。

一、公司发展

2015 年底，OpenAI 成立，直至 2018 年一直处于非营利阶段。2019 年开始向营利性转型，同年获得微软提供的 10 亿美元的投资，自此与微软构成合作伙伴关系。2023 年微软向其提供了第二笔多年投资，据报道为 100 亿美元。

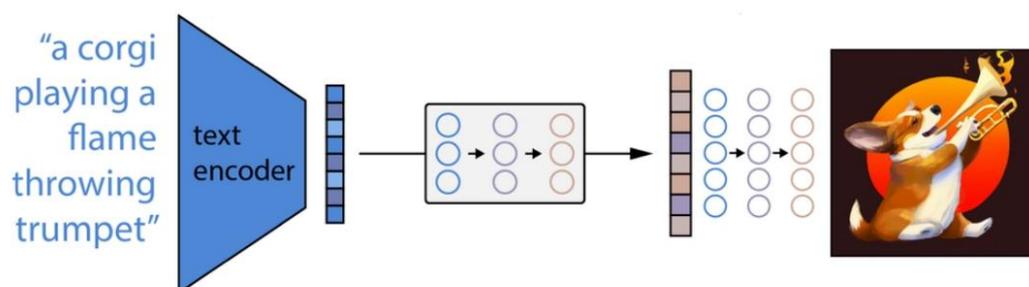
2021 年 1 月 5 日，OpenAI 发布了 DALL·E 模型，这也是一个具有很大影响力的模型，DALL-E 是一个 120 亿个参数的 GPT-3 版本，它被训练成使用文本——图像对的数据集，从文本描述中生成图像。DALL-E 可以创造动物和物体的拟人化版本，以合理的方式组合不相关的概念，渲染文本，以及对现有图像进行转换。

2022 年 4 月 6 日，DALL-E 2 发布，其效果比第一个版本更加逼真，细节更加丰富且解析度更高。

二、产品介绍—DALL-E 2

DALL-E 2 是一款人工智能图像生成器，它可以根据自然语言的文本描述创建图像和艺术形式。2021 年 1 月，OpenAI 推出了 DALL-E 模型，DALL-E 2 是其升级版。2022 年 7 月，DALL-E 2 进入测试阶段，可供白名单中的用户使用。同年 9 月 28 日，OpenAI 取消了白名单的要求，推出了任何人都可以访问并且使用的开放测试版。

1、工作原理



“一只在吹喷火喇叭的柯基” ——DALL-E 2 图片生成流程解析

针对图片生成这一功能来说，DALL-E 2 的工作原理剖析出来看似并不复杂：

(1) 首先，将文本提示输入文本编码器，该训练过的编码器便将文本提示映射到表示空间。

(2) 接下来，称为先验的模型将文本编码映射到相应的图像编码，图像编码捕获文本编码中包含的提示的语义信息。

(3) 最后，图像解码模型随机生成一幅从视觉上表现该语义信息的图像。

2、主要功能

DALL-E 精通各种艺术风格，甚至可以根据季节、光线、环境、时间的不同，绘制相应的光影。

新推出的 DALL-E 2 版本，能够以比 DALL-E 高 4 倍的分辨率，生成更真实、更准确的图像，尤其在字幕匹配和照片真实性方面优于 DALL-E。此外，DALL-E 2 还具有更快的生成速度和更低的计算成本。人们可以尽情展开想象力，DALL-E 2 可以短时间内将各种有趣的想法转化成图像。例如，当您描述“可以通向另一个世界的一碗汤”，并指定图片风格分别为：数字艺术、巴斯奎特涂鸦风格以及洞穴壁画时，DALL-E 2 将生成如下画作：



数字艺术风格



巴斯奎特涂鸦风格



洞穴壁画风格

除了画作外，DALL-E 2 还可以生成照片级的真实图像，如：宇航员闲适地待在太空中的热带度假胜地（下图所示）。



此外，DALL-E 2 还有很多强大的拓展功能：

(1) 可以基于原始图像，在原始画布之外生成新图像，创造新构图。

下图是生成的《戴珍珠耳环的少女》新构图：



(2) 可以基于自然语言描述，添加和删除元素，同时呈现出真实的阴影、反射和纹理。

(3) 可以根据局部风格来调整成像，使画面整体更具真实性。

(4) 可以根据原始图像创建不同的变体。

3、功能限制

功能上对模型进行了限制，主要有三点：

(1) 阻止生成有害内容

OpenAI 通过从训练数据中删除暴力、仇恨或少儿不宜的图像，限制了 DALL-E 2 产生这类内容的能力，最小化了 DALL-E 2 对这些概念的认知。并且还使用一些技术手段防止生成人脸的超写实照片，尤其是一些公众人物。

(2) 防止滥用

内容政策中规定，不允许用户生成暴力、成人或政治内容。如果过滤器识别到可能违反策略的文本提示词或图像上传，系统将不会返回生成图像。系统中还有自动化和人工监测系统防止滥用。

(3) 基于学习的分阶段部署

OpenAI 一直与外部专家合作，并对一定数量的可信赖用户开放预览 DALL-E 2 的权限，这些用户可以帮助开发人员了解技术的能力和限制。开发团队计划通过学习，迭代地改善安全系统，随着时间的推移邀请更多人参与预览。

4、常见应用场景

DALL-E 2 可以根据文字描述进行艺术创作，生成高质量的图像，并根据需要进行调整。因此，DALL-E 2 可以应用于广泛的领域和场景，例如广告、游戏、娱乐、教育等。

(1) 广告娱乐：生成极具吸引力的图像，以吸引更多客户。

(2) 游戏：生成游戏中的角色和场景，以增强游玩互动体验。

(3) 教育：生成教学图像，以帮助学生更好地理解课程内容。

(5) 汽车生成汽车设计图纸，以加速开发新车型。

(6) 制造和零售：生成产品图像，以优化产品销售体验。

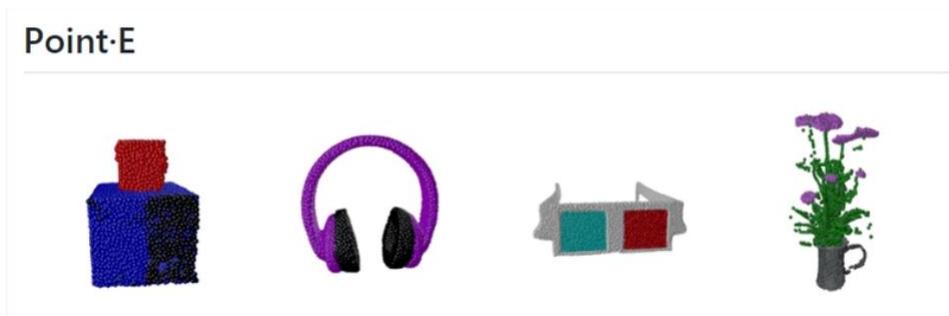
此外，DALL-E 2 十分易用，可以通过 API 与其他系统集成，使开发人员可以轻松地将其集成到他们的应用程序中，获取更多功能。

微软已经宣布，将 DALL-E 2 集成到新版必应搜索引擎和 Designer 图形应用程序中，帮助设计师尽可能轻松地创建独特的图像和设计。设计师只需输入产品或公司名称，Designer 将帮助生成符合品牌形象的独特视觉效果。当添加新素材时，它还可以自动更新，确保与整体设计完美匹配。

三、最新动态——部署 AI 3D 领域

2022 年 12 月 OpenAI 开源了 Point-E。这是一个机器学习系统，它可以基于文本提示创建 3D 模型，也被称为“3D 建模版” DALL-E。

Point-E 结合了传统的文本到 3D 合成训练算法的方法。将两个单独的模型结合在一起，Point-E 可以减少创建 3D 对象的数量。第一组算法是文本到图像模型，类似 DALL-E2，它可以创建用户给出的提示的图像。然后将此图像用作第二个模型的基础，该模型将图像转换为 3D 对象。



3D 模型合成人工智能很有可能会成为下一个巨大的行业颠覆者。三维模型广泛应用于影视、室内设计、建筑和各种科学领域。例如，建筑公司用它们来演示设计的建筑和景观，而工程师则利用模型来设计新的设备、车辆和结构。目前 DALL-E 已经彻底改变了文本到图像的生成过程，Point-E 的目标正是为 3D 领域做同样的事情。

DALL·E 2, Stable Diffusion 和 Midjourney 之间的比较

(1) DALL-E 2 使用数以百万计的图片数据进行训练，其输出结果更加成熟，非常适合企业使用。当有两个以上的人物出现时，DALL-E 2 产生的图像要比 Midjourney 或 Stable Diffusion 好得多。

(2) Midjourney 则是一个以其艺术风格闻名的工具。Midjourney 使用其 Discord 机器人来发送以及接收对 AI 服务器的请求，几乎所有的事情都发生在 Discord 上。由此产生的图像很少看起来像照片，它似乎更像一幅画。

(3) Stable Diffusion 是一个开源的模型，人人都可以使用。它对当代艺术图像有比较好的理解，可以产生充满细节的艺术作品。然而它需要对复杂的 prompt 进行解释。Stable Diffusion 比较适合生成复杂的、有创意的插图。但在创作一般的图像时就显得存在些许不足。

使用这三种不同工具，输入相同 prompt，产出的图片如下所示：



Cherry Blossom near a lake, snowing



Eerie forest, black and white, night



Dark alley at night 4k raining aesthetic



Alone astronaut on Mars, mysterious, colorful, hyper realistic



Pyramid shaped mountain above a still lake, covered with snow

功能	Stable Diffusion	Midjourney
图像自定义	高	低
上手难度	低	中
生成图片难度	低	高
修复缺陷	是	否
拓展画布	是	否
宽高比	是	是
模型变体	约 1000 个	约 10 个
负向提示	是	是
从生成中获得变化	是	是
控制构图和姿势	是	否
许可证	自由	限制（取决于付费层次）
制作自己的模型	是	否
费用	免费	每月 10-60 美元
模型	开源	专有
内容过滤器	否	是
风格	多样	真实插图，艺术风格
图像放大器	是	是

图像提示	否	是
图像转换	是	否
提示词限制	无限制	未知

由于 Stable Diffusion 对使用者的调节灵活性不同,因而适用者以及使用场景也会不同,从制图的灵活性上来讲 Dall.E 2 与 Stable diffusion 要相似一些,这两款允许使用者对于图片的部分区域进行替换调整扩展以及具体细化,更适合设计师来调教训练出更适合自己风格的 AI。而 Midjourney 更偏向艺术性,它可以通过 ChatGPT 的简单 prompt 生成细节丰富的图像,更适合设计师的灵感获取。

因为 Dall.E 2 与 Stable diffusion 在画面调节有一定程度的相似性,因而笔者将对 Stable diffusion 与 Midjourney 进行具体对比。

Midjourney 和 Stable Diffusion 都拥有庞大的用户群体,而它们都有各自的优点和缺点。以下笔者将为大家提供建议,可以依据个人需求进行选择。

选择 Stable Diffusion	选择 Midjourney
完全免费的解决方案。	无需深入学习即可生成惊人的图像。
在本地运行所有内容。	没有时间设置和学习模型。
喜欢调整设置、尝试模型组合、使用新工具。	想要 out-of-box 的 AI 图像绘制工具。
更喜欢开源工具。	不介意支付订阅费。
想要对图像做更多的控制。	可以接受使用限制条款。

国内风向

产品名称：文心·一格

公司主体：百度

融资阶段：上市公司

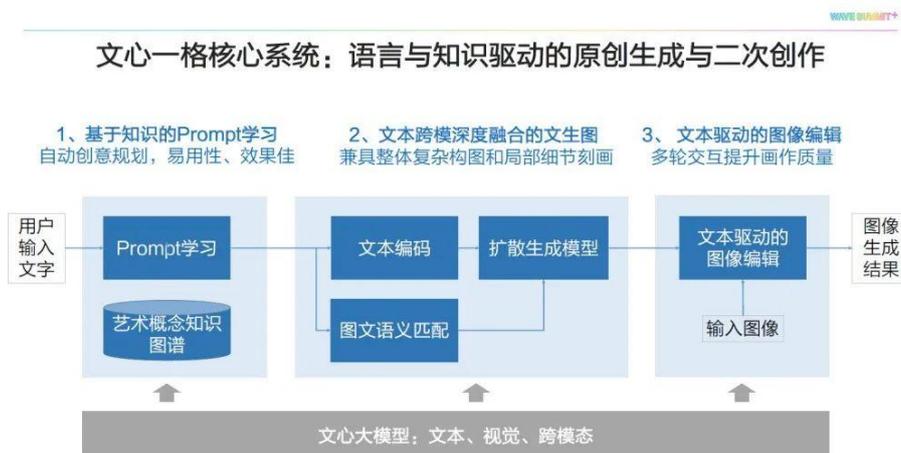
一、产品简介

百度推出了基于文心大模型的 AI 艺术创作产品——文心一格。用户在体验该产品时只需要输入一句话，AI 就能够自动生成创意画作。因此，创作者不再需要设计整个复杂过程和图像制作，而只需要输入语言，这使整个创作效率实现了大幅提升。目前该应用支持 12 种绘画风格，200 字以内的关键词输入，支持文生图和图生图。

二、核心技术

百度飞桨为文心一格研发了一个核心文生图系统，这个系统是基于语言与知识驱动的原創生成与二次创作。整个系统是基于文心大模型实现的，包括文本、视觉和跨模态三个方面：

- 1、基于知识的 Prompt 学习。对用户输入的文字，通过自动创意规划，能够大幅提升易用性。
- 2、文本跨模深度融合的文生图。在图像生成中，通过文本和跨模态的深度融合，模型兼具整体复杂构图和局部细节刻画能力。
- 3、文本驱动的图像编辑。通过文本驱动的图像编辑，允许用户进行多轮交互，提升画作质量。



百度还致力于构建生态完整的开放社区，将大模型的能力释放给普通人。更具体地讲，百度为开发者、广大科技爱好者提供飞桨开源工具和文心大模型 API 服务能力，满足他们多样化的创意探索需求。一方面，开发者可以在飞桨 PaddleHub 上直接调用相关的开源模型，享受极简易用的开发体验，并综合使用提供的 400 多 AI 开源模型，组合开发有趣的 AI 应用，满足更多艺术从业者、开发者对图片生成场景的需求；另一方面，文图生成大模型 ERNIE-ViLG 为开发者提供了 API 体验调用的入口，登录飞桨畅谷社区 API 体验专区即可体验其前沿技术能力，并能够灵活方便、高效快速将文图生成大模型能力集成到产品中。

三、产品功能

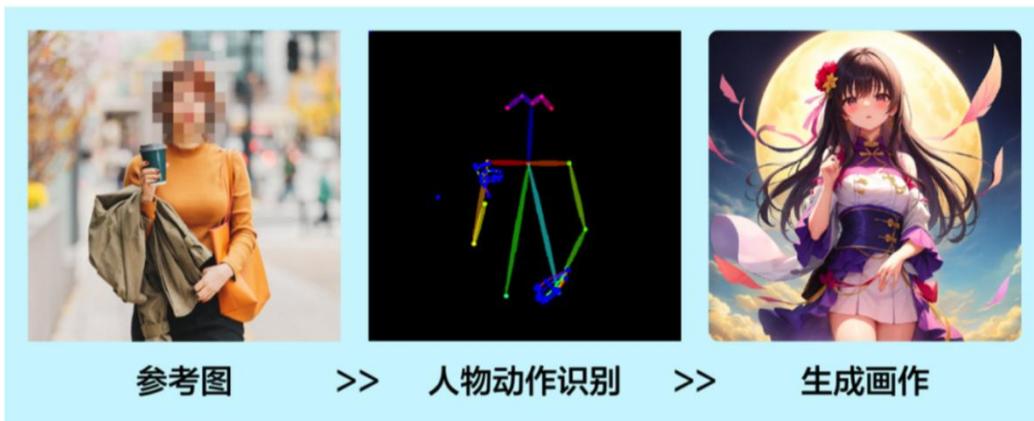
1、AI 编辑功能：图片叠加支持两张图片风格特征融合；自由涂抹精准实现涂抹区域画作内容重新生成。



2、人物动作识别再创作：AI 可识别人物图片动作，再结合输入的描述词，生成动作相近的画作；



3、线稿识别再创作：一键提取图片线稿，还可以在线稿基础上再创作。



产品名称：秒画 SenseMirage

市场主体：上海商汤智能科技有限公司

融资阶段：上市公司

一、产品简介

秒画是基于商汤自研的文生图大模型，依托商汤“日日新 SenseNova”大模型体系和 AI 大装置 SenseCore 强大算力集群的 AI 作画平台。秒画“SenseMirage”不仅融合了超 10 亿参数的商汤自研文生图生成模型，还支持一键导入多个平台的开源模型或上传用户本地模型并对其进行特异性推理加速优化，结合自研模型及训练能力，用户可免除本地化部署流程，并基于开源模型自训练模型，高效地生成更多样的内容。

商汤自研的 AIGC 文生图大模型拥有超 10 亿参数，支持二次元、三次元等多种生成风格，最高可实现 6K 高清图。商汤自研作画模型的文生图能力，每 2 秒就能生成 1 张 512*512 分辨率的图片。

二、核心技术

【自研模型高调对标 Stable Diffusion】

据商汤科技官方宣传，相比于 Stable Diffusion，商汤作画大模型基于 19 年开始研发的通用大模型设计体系，采用更先进的大模型结构设计与大 batch 训练优化算法，模型参数量大小为前者数倍。

核心技术包含了自研的分层推理专家（hierarchical inference experts）、图像质量感知的分布式训练（image quality-aware distributed training）、图像纹理指导的交叉注意力学习（texture-guided cross-attention learning）等优化算法，改进了扩散模型在生成的语义完备性、物体宏观结构、细节（比如手指、四肢等）、推理步的加速等方面的问题，使其具备更优的文本理解泛化性、图像生成风格广度以及图像高质量生成细节。不过，该应用目前并未对公众开发，只有少量 B 端客户可以申请到内测资格，因为并不能亲自感受该软件与 Stable Diffusion 之间的区别。

此外商汤科技还表示，目前市面上所有的 LoRA 模型都是基于 Stable Diffusion 或者其变体训练的，而商汤自研的基模型拥有强大的泛化能力，使用其替代 Stable Diffusion 模型可以基于更少量数据快速训练出质量更高的 LoRA 模型，实现更优的风格定制化图像生成，这样的功能在秒画平台通过 5 分钟即可实现。

同时，用户也可以一键导入 Hugging Face、Civitai、GitHub 等第三方社区

的开源模型进行体验，让用户免除本地化部署的繁琐流程，通过商汤 AI 大装置 SenseCore 强大的 GPU 算力集群、推理加速功能，即可高效生成更多样的内容，创作更便利。



商汤自研模型生成



Stable Diffusion生成

Prompt关键词：行驶中的红色保时捷，美观的建筑，概念设计，白色天光，水墨
running red Porsche car, beautiful buildings, concept design, white sunlight, (ink
and wash:1.2)



商汤自研模型生成



Stable Diffusion生成

Prompt关键词：梦幻的蝴蝶，透明材质，水晶质感，翅膀闪光，细节刻画，超写实
A dreamy butterfly, transparent material, crystal texture, wings emitting particle
light, detailed, ultra-realistic style

三、产品功能

秒画定位于 to B 的 AI 作画平台，将广泛应用于设计、营销、游戏动画、工程建筑、医学、教育科普等领域。在秒画平台上，商汤自研的文生图大模型是一个基础模型，B 端客户仅用少量数据就可以迅速微调得到一个适用于下游任务

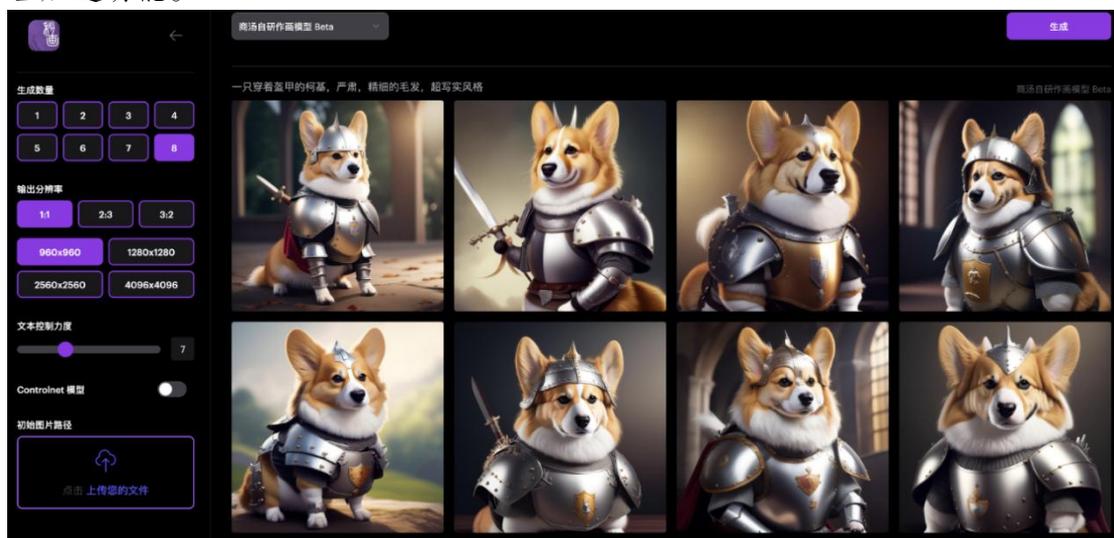
的垂类模型，大大降低了训练门槛。

除了使用商汤自研模型外，在秒画平台上用户可以一键导入 Hugging Face、Civitai、GitHub 等第三方社区的开源模型，并且添加至用户「我的模型库」列表中，方便进行体验，让用户免除本地化部署的繁琐流程。这将大幅降低 AI 作画的商用门槛。对于一些中小型企业来说，本地化部署大模型的成本是非常高的。硬件方面，训练垂类模型需要搭建一定规模的 GPU 算力集群，还需要专业的技术团队，人力成本也是比较高的。

秒画为商汤自研文生图模型和开源模型提供了一套算法框架和强大的算力基础设施——AI 大装置 SenseCore。B 端用户可以直接在秒画平台上训练定制化的 LoRA 模型，基于现有商汤自研文生图模型的高度泛化性，B 端用户只需要拖拽 20 张左右图片作为训练集，最低仅需 5 分钟就能完成一个高质量模型的训练。这让中小型企业也能轻松拥有属于自己的垂类模型。

此外，秒画还提供了特异性推理优化服务，开源模型导入后会自动采用秒画平台底层的模型编译技术进行加速。借助这项加速技术，在本地 RTX 3070 显卡上需要 10 秒生成的图像，秒画只要 2 秒的推理时间就可以生成。

总体来说，秒画是一个依托商汤 AI 大装置 SenseCore 的 to B 文生图平台，它基于商汤自研的超 10 亿参数文生图大模型，也能一键导入第三方开源模型，为 B 端客户免去了本地化部署的繁琐流程，并提供了简单快速的训练服务和推理加速功能。



产品名称：万兴爱画

公司主体：万兴科技集团股份有限公司

融资阶段：创意软件 A 股上市公司

一、公司简介

万兴科技深耕数字创意软件领域，推出万兴喵影、万兴优转、万兴录演等视频创意软件，Wondershare Filmstock、万兴喵库视频创意素材平台，万兴爱画、万兴播爆等 AIGC 应用；此外面向全球市场还推出万兴图示、万兴脑图 MindMaster、墨刀等绘图创意软件，万兴 PDF、Wondershare PDFelement 等文档创意软件，万兴恢复专家、万兴数据管家、万兴易修等实用工具软件。

万兴科技的产品涵盖视频创意、绘图创意、文档创意、实用工具四大产品线，目前已于多个产品内落地虚拟人、文生图、AI 智能抠像、AI 智能降噪、AI 音乐重组、AI 换脸等功能。

视频创意 >	AIGC应用 >	绘图创意 >	文档创意 >	实用工具 >
 万兴喵影	 万兴爱画	 万兴图示	 万兴PDF	 万兴恢复专家
 万兴喵库	 万兴播爆	 万兴脑图	 万兴PDF阅读器	 万兴数据管家
 万兴优转		 墨刀		 万兴易修
 万兴录演				 万兴手机管家

万兴科技丰富的产品线

二、公司发展

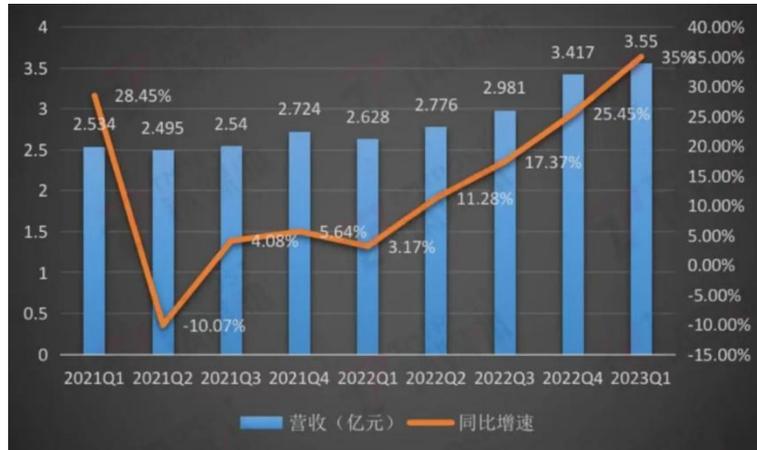
借助 AIGC 的东风，万兴科技的股价在不到两个月的时间便翻倍，Q1 收入利润双增长。

2023 年一季度，万兴科技实现营业收入约 3.55 亿元，同比增长 35%，归母净利润 2085.58 万元，同比增长 686.26%，去年同期则亏损 355.74 万元。扣非净利润 1798.50 万元，同比增长 306.80%。

目前万兴科技旗下多款产品也已集成 AIGC 功能，在技术投入上，万兴科技的研发费用也是持续增长。2020 年至 2022 年，万兴科技研发费用连续快速增长，分别为 2.19 亿元、3.28 亿元、3.48 亿元，占各期营业收入比例分别为 22.39%、31.88%、29.47%。

2023 年第一季度，其研发费用为 9430 万元，同比增长 16%，占收入比重为 26.56%。最新数据显示，截至 2022 年 12 月 31 日，公司在岗总人数为 1480 人，较上年同期增长 8.42%，研发人员数量较上年同期增长 5.96%。

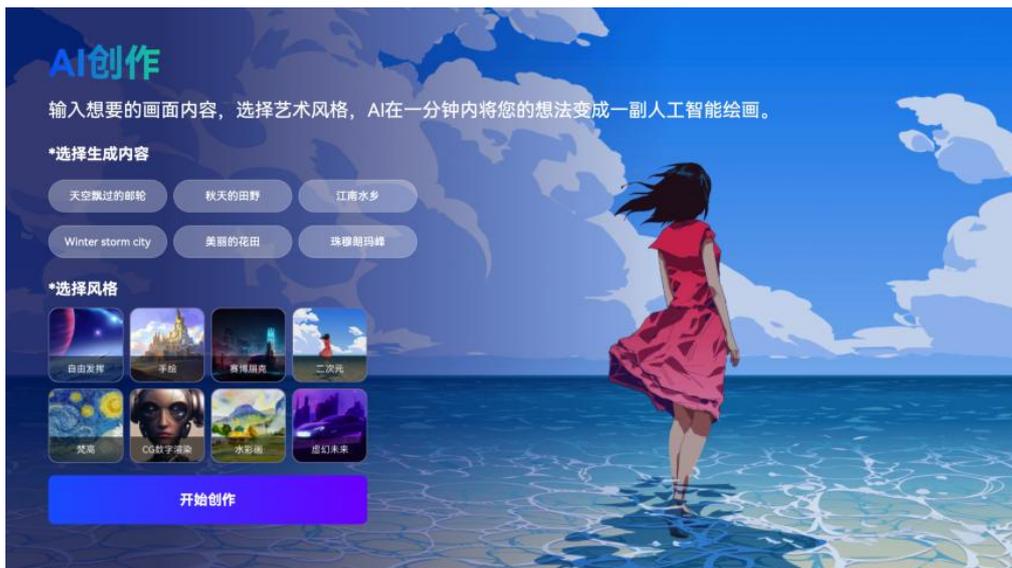
过去的三个月时间里，万兴科技专注视觉领域的基础、分割和生成类算法的研发与落地，于报告期内完成多人图像换脸、超清人像、特效人像分割、裸眼3D等算法的功能化落地及优化。同时，持续深入AIGC绘画算法，打通垂类训练技术，提升AI绘画模型效果，实现高保真人像/风景特征图片和二次元、国风等多风格图片生成。



万兴科技收入及增速 图片来源：钛媒体

三、产品介绍：万兴爱画

万兴爱画已实现网页端、iOS、安卓、微信小程序多端覆盖，并推出「AI简笔画」新功能，支持文字绘画、以图绘图、简笔画三种AI创作模式，具备交互型「图生图」AI绘画功能。用户输入想要的关键词，选择图片比例和艺术风格之后，即可获得由AI自动生成的绘画作品。此外，用户运用全新的「AI简笔画」功能，只需简单描摹几笔，即可5秒内「人机共创」绘出一幅高品质艺术画作。



为了降低用户在输入“关键词”时的学习成本，万兴爱画全新上线文本辅助输入功能「咒语生成器」，预设超500个常用Tag，并提供人物&角色、五官、

表情、头发、服装等十余个分类选项及多样化的风格效果，用户可以点选所需效果并对其权重进行调整，万兴爱画可据此自动生成一段丰富的文本内容，从而在 AI 算法的加持下创作出具有独特风格的绘画作品。该功能解决了用户在 AI 绘画过程中需要手动输入大量文本内容的繁琐问题，让 AI 绘画更加简单易用。

不同于 TIAMAT 和西湖心辰的自研模型，万兴爱画选择了接入 openAI，而专注于软件性能的研发。同时，万兴科技旗下视频创意软件万兴喵影、Wondershare Filmora 也已集成 AIGC 能力，推出文生图功能，实现了视频创作领域图片素材的在线智能生成，用户可在视频创作软件中，智能生成创作图片素材。而万兴科技旗下亿图脑图协同版已经正式开启 AIGC 功能公测，用户输入一句话，即可一键生成头脑风暴、演讲大纲、SWOT 分析、活动策划、生活计划等脑图加持下，亿图脑图大大简化思维脑图的创作流程并革新创作模式，助力用户创作提质增效。

目前万兴爱画“文生图”可供选择的风格仅有二次元经典、彩墨风、浮世绘、黑白默剧、古典油画。“图生图”功能可选择二次元经典、赛博朋克、厚涂油彩、二次元—男、日式漫画、流光厚涂。便于新手入门的功能是“咒语生成器”，提供丰富多彩的关键词选择，有效降低了应用的学习成本。但是万兴爱画目前并没有提供类似于 midjourney、TIAMAT、造梦日记等应用的“社区分享”功能。



四、最新动态

2023 年 2 月 25 日，万兴科技表示“公司已对接 Azure OpenAI 开通商用服务权限”，是国内首批获得 Azure OpenAI 商用服务权限的企业。2023 年 2 月，万兴科技旗下视频创意软件 Wondershare Filmora 新版本全新接入 ChatGPT 母公司 OpenAI 相关服务，率先在视频创作领域集成 AIGC 新技术。据介绍，新版本可自由调用 OpenAI 的 AI 绘图功能，一键即可智能生成图片素材。未来万兴科技

将逐步深化与以 OpenAI 为代表的 AI 上游技术商合作，借助 AIGC 技术持续升级和推出产品。



万兴科技AIGC落地应用明星产品

- 万兴播爆**
无需外籍演员出镜，可调用模板和国籍数字人，一键生成AIGC“真人”出海营销短视频。
- 万兴爱画**
支持文字绘画、以图绘图、简笔画三种AI创作模式，快速创作大师“同款”高品质艺术画作。
- 万兴喵影/Wondershare Filmora**
集成AIGC功能，视频创作一键智能生成图片素材。
- 亿图脑图**
一键生成头脑风暴、演讲大纲、SWOT分析、活动策划、生活计划等脑图。

更多产品正在研发中……

随着 OpenAI 的接入，Wondershare Filmora 用户在创作视频时，可自由调用 OpenAI 强大的 AI 绘图功能，一键即可智能生成图片素材，快速创作大师“同款”作品；未来，用户还将在产品内体验更多 AIGC 新功能。

产品名称：造梦日记

公司主体：西湖心辰（杭州）科技有限公司

融资阶段：Pre-A 轮融资

地区：杭州

一、公司简介

西湖心辰专注于人工智能服务研发的科技公司，旗下产品涵盖 AI 绘画、AI 智能写作、AI 心理咨询等领域，为 B 端、C 端用户提供相关优质服务。「造梦日记」AI 绘画平台拥有超强算力，最快 0.8 秒出图。平台内有 20+绘画风格和 10+画家可供选择，且支持多尺寸、多清晰度、批量出图等功能，用户输入简单的描述词或上传参考图即可生成原创图片。「Friday」AI 智能写作平台依靠自主研发的创新性 AI 算法，可以生成不同形式和语言风格的内容，拥有 50+写作模板，适配多种写作场景，25 秒即可生成一篇高质量原创文章。「聊会小天」AI 心理咨询平台能为用户提供数字化的、7x24 小时的、千人千面的心理支持与陪伴，是轻度心理倾诉及专业的心理咨询服务平台。

二、公司发展

西湖心辰成立于 2021 年，专注于大模型的研究和应用。西湖心辰创始人蓝振忠曾在谷歌担任研究科学家，也是轻量化大模型 ALBERT 的第一作者（学术引用超 4400 次），2020 年加入西湖大学担任深度学习实验室负责人、博士生导师。

西湖心辰专注于探索多模态大模型和通用人工智能技术，并已基于自研大模型衍生出多个上层应用。基于对大模型的研究成果，西湖心辰构建了模型层—中间层—应用层的产品架构。

在模型层，西湖心辰打造了 RIO 系列语言大模型和双模态图文模型，未来的目标是打造多模态大模型。按照模型参数量大小，RIO 系列现在已有十几个成员，参数量从 100 多亿到 1000 多亿不等，支持中英文双语，除了记忆外，还掌握了搜索引擎、知识图谱、外部数据等认知工具调用能力，经测试，目前 RIO 模型的能力相当于 GPT-3.5 的水平。在底层模型的智能化能力方面，西湖心辰今年也将完成两次迭代，目标是达到、甚至在部分任务上超过 GPT-4 的表现。

中间层，西湖心辰可提供深入各个领域（如营销等）的模型，大模型在处理非结构化数据的能力上有了极大提升，过往在企业内部沉淀下来却未被充分利用

的数据，将在大模型赋能下展示出更多价值。西湖心辰在去年开始接触了大量 B 端客户，也在不断和各行业交流、共同探索场景机会，目前已与多家客户在营销、社交领域开展了深度合作。

应用层，西湖心辰推出了图片生成产品「造梦日记」、智能写作产品「Friday」和心理咨询 AI 助手「聊会小天」三款产品，覆盖文字内容生成、图片内容生成、聊天对话等领域。在这一产品架构下，终端应用和底层大模型的联动十分紧密，依据应用的难度和需要推理的速度选择合适的模型，再通过该应用相关领域的数据进行调优。

总体而言，西湖心辰的特点在于能力全面，凭借团队在算法、工程、产品等方面的科研和产业经验，打造了大模型-中间层-应用层三层架构，并在多模态技术上具备较深经验。

2021 年 12 月西湖心辰完成了蓝驰创投、凯泰资本等投资的近千万美元天使轮融资。2023 年 3 月，完成了数百万美元的 Pre-A 轮融资。本轮融资领投方为百度风投，老股东凯泰资本和西湖教育基金会可持续发展平台跟投。本轮融资将主要用于团队扩张和建设，加速多模态大模型的研发，提升大模型的通用人工智能能力。

序号	披露日期	交易金额	融资轮次	估值	比例	投资方
1	2023-04-22	未披露	Pre-A+轮	-	-	汤姆猫 蓝驰创投
2	2023-03-27	数百万美元	Pre-A轮	-	-	BV百度风投 凯泰资本 西湖教育基金会
3	2021-12-23	近千万美元	天使轮	-	-	蓝驰创投 凯泰资本 西湖科创投 西湖大学产业投资基金

三、产品介绍：「造梦日记」AI 绘画平台

造梦日记 AI 绘画平台由西湖心辰和西湖大学共同研发，使用西湖心辰自研的双模态图文模型，超强算力的支撑使得该应用可以达到最快 0.8 秒的出图速度。造梦日记目前有网页端、iOS、安卓、微信小程序多端覆盖，功能上包括了 AI 画图和 AI 视频两大部分，其中 AI 视频功能还在内测中。曾达成单日新增 5 万用户的记录，如今用户数已达百万。造梦日记平台基于 cco 协议，保证用户拥有生成图片的版权，你们可以利用这些 AI 产物玩转社交平台、成为自己艺术创作的素材、抑或是创造更大的价值。

造梦日记的 AI 画图提供了包括摄影 pro、动漫风、超拟真、油画、水彩、赛博朋克、印象主义等多达 30 种绘画风格，徐悲鸿、莫奈、毕加索等 13 位艺术

家风格，在风格的多样性上较之于国内其他应用丰富很多。除了文生图外，造梦日记将“文生图”与“图生图”相结合提高作画的准确性，创作者可以结合“关键词”“参考图”以及“反向关键词”综合进行 AI 作画，从而使得生成的图片更加符合创作者的预期。面向兴趣性的 c 端用户，造梦日记推出了“头像生成器”“萌宠变变变”两个趣味功能，一键生成自己的动漫头像以及将自己的宠物拟人化。

目前「造梦日记」对关键词的学习成本较高，目前仅支持 300 字的关键词，输入“关键词”的专业性对生成画面效果的影响较大。该应用目前对动漫风格的图片处理能力比较强，而摄影人物的生成质量仍然有很多不足，生成的图片有明显的“人工智能”痕迹。这可能是由于造梦日记相较于 midjourney 仍然缺乏更多的模型训练以提高模型的作画能力。

笔者以相对简单和详细的两种关键词输入，生成动漫风和摄影图以作对比。不难看出该应用在画面效果上仍然有很大的提升空间。当关键词不专业时生成的图片，我们能看到画面中素材拼凑的痕迹，比如说图中的少女直接站在了溪水上。此外，对写实性人物的描绘也存在很大问题，摄影图中人物的面部、手部、腿部细节处理都不到位。在给到详细关键词之后，这两种情况得到了改善，但是仍然存在一些问题。

关键词	动漫风	摄影
少女 大眼睛 正面 漫步 夏日 树木 森林 小溪		
高品质 杰作 大师之作 精致的五官刻画 精致的头发刻画 精致的眼睛刻画 4k画质 绚丽的光影 美丽的少女 丁达尔效应 光晕 长发 幼态 华丽的场景 精致的衣服 羽毛		

产品名称：画宇宙

公司主体：北京计算美学科技有限公司

融资阶段：Pre-A 轮融资

地区：北京

一、公司简介

Nolibox 是聚焦于「人工智能」和「设计创意」交叉领域的科技公司，致力于“让每个人都能享受设计和创意之美”。旗下有「画宇宙」和「图宇宙」两款产品，为用户和客户提供全链路 AIGC 和人工智能设计创意解决方案。Nolibox 将美学原理、设计模型及人工智能技术进行深度融合，基于 AI 技术进行知识表示与图形生成，支持多场景之间的图形变换与融合。Nolibox 纵向层面赋能视觉设计的多个垂直领域；横向层面为企业提供千人千面的智能营销解决方案，建立面向未来的智能设计生态系统。

二、公司发展

Nolibox（计算美学）成立于 2020 年，是一家聚焦人工智能及设计创意深度融合的设计科技企业。核心团队来自清华大学，北京大学，RCA 等国内外相关领域顶级高校及实验室。nolibox 提出「人机协创」的概念，通过搭建设计中台、智能设计工坊及美学评估系统，为设计供需市场提供新的生产方式。

2021 年完成数千万元 Pre-A 轮融资，投资方为高瓴创投，融资将用于团队扩张、产品研发、素材授权和用户增长。Nolibox 在 3 个月内连续完成了两轮融资，此前已获初心资本千万级天使轮融资。2021 年 12 月 30 日，Nolibox 武汉全资子公司（智能设计创意中心）在武汉洪山区光谷新发展国际中心正式开业。Nolibox 将在武汉深度布局优质创意研发团队，为旗下智能设计产品提供大量关于设计数据化、设计模型、设计知识图谱的底层积累。在 2022 年 10 月 Nolibox 正式上线发布了「画宇宙」AI 绘画平台，并快速获得大量客户和用户。

序号	披露日期	交易金额	融资轮次	估值	比例	投资方
1	2021-12-27	数千万人民币	Pre-A轮	-	-	高瓴创投
2	2021-10-27	千万级人民币	天使轮	-	-	初心资本

三、产品介绍

计算美学的 AI 绘画产品——画宇宙是一个面向 B 端专业场景的创作平台，Nolibox 团队在产品定位上致力于将 AI 绘画技术与各行各业的需求进行结合。为了符合专业级用户的需求，在功能设计上，“画宇宙”集成了 SD、百度文心模型等各类第三方模型，将文生图、图生图、局部绘制、画面外延、相似生成等多项 AI 绘画功能进行整合，这使得用户可以在一个无限空间内实现创意的生成、创作、参考、对比、延展、整理、融合、局部修补等操作。



[定位：API 开放平台 & 企业解决方案]

“画宇宙”定位于解决企业的设计创意需求，包括内容生产力工具需求、灵感创意需求、流量和品牌运营需求等，并根据不同的需求分别研发了多个模块。

Nolibox 为企业提供“定制化”的服务：

1、对前期创意和灵感落地有需求的客户，需要结合自身行业特点和项目要求，生成大量的设计草案。在这一过程中，行业知识、行业经验的总结与积累是非常重要的，Nolibox 会与企业进行快速碰撞，探索共创出一套完整的行业知识系统；在功能层面，「画宇宙」专门为企业客户提供了「无限灵感」（批量生成）的功能，通过企业与 Nolibox 共创的知识系统，批量生成千张结果，极大满足了客户在灵感创意阶段的设计需求。Nolibox 整理众多客户的共性需求，总结出了一套较为完整的共创模式，已经应用到建筑设计、工业设计、包装设计等领域。

2、对于有流量运营需求的企业，Nolibox 提供标准化和定制化的接口服务，例如推出结合时事热点的一秒国潮、阿凡达头像、魔幻变妆、兔年变兔等 AI 绘画接口，赋能企业品牌运营和流量运营。提供按量、月租、部署等多种结算模式

产品名称：Tiamat

公司主体：上海退格数字科技有限公司

融资阶段：A 轮融资

地区：上海

一、公司简介

上海退格数字科技有限公司是一家 AI 图像生成技术服务商，成立于 2021 年。公司旗下产品“Tiamat”，是具有独特中国风的 AI 图像软件。其自研的 MorpherVLM 是国内首个基于概念融合范式提出的近百亿级跨模态生成模型。作为一家人工智能领域的创业公司，退格数字科技的创始人擅长社区运营和市场营销，因此 Tiamat 在社交媒体上的热度很高。早在今年上半年，Tiamat 就在国内 AI 绘画还未火起来的时候，率先在小红书和抖音上运营 AI 绘画的社区，让更多人上来社区使用产品，从而收集大量的用户反馈。目前其“Tiamat 人工智能艺术”标签在小红书有 1776.1 万浏览。

二、公司发展

上海退格数字科技有限公司成立于 2021 年，孵化于上海科技大学。Tiamat 团队的核心成员来自上海科技大学信息学院和伦敦艺术大学，在 AI 图片引擎、数字孪生等方面有着深厚的技术积淀，创始人青柑拥有项目策划和社区运营经验。

作为国内第一批入局 AI 生成领域的技术团队，致力于通过自研的算法、底层引擎、平台为个人用户和企业客户提供创意生产、广告等场景形式下的服务，积累了可观的社区用户和商业客户。在市场营销和 B 端用户的开发方面，Tiamat 已经尝试与科幻 IP、央视网、《时尚芭莎》《嘉人》等一线时尚大刊建立长期合作关系，并与众多头部游戏公司，特别是与持有故宫文化相关 IP 的非遗文化 IP 机构知造局等开展了战略合作。

2022 年 10 月 12 日，Tiamat 完成数百万美元天使轮融资，本轮投资方为 DCM。2023 年 2 月，Tiamat 完成近千万美元 A 轮融资，由老股东 DCM 和绿洲资本联合领投。

序号	日期	融资轮次	估值金额	融资金额	投资机构
1	2023-02-28	A 轮	未披露	约千万美元	DCM Ventures, 绿洲资本
2	2022-10-12	天使轮	未披露	数百万美元	DCM Ventures

三、产品介绍

Tiamat 生成式 AI 模型是基于扩散模型（Diffusion Models）和审美模型实现的，研发团队在扩散模型采样、数据集构建上开展了许多工作，进一步提高了图像生成质量。Tiamat 在开源模型上做了“中国风模型引擎”的优化，作为一款国产优化自研系统，Tiamat 系统支持中文和各种复杂的描述风格。Tiamat 研发团队创新性地使用梯度引导和条件生成结合的技术，对系统进行优化，使得 Tiamat 系统能够在图片生成符合文本输入的同时，让最终结果的色彩细节、构图更加完善。同时，Tiamat 面临着优化和成本算力等问题，包括如何让 AI 学会结构、构图透视等。生成式人工智能的整个生成系统需要庞大的算法支持，一方面需要极大的成本投入，另一方面也需要进一步对用户量和数据安全等问题进行测试。

目前，Tiamat 已经开放小程序版，网页版还在内测阶段。Tiamat 卓越的表现其中国风模型引擎上，Tiamat 开发一系列国风特色的绘画模型，能够生成精美的高质量国风图像。其次，Tiamat 通过国内社区的运营积累了大量的国内用户数据和反馈，能够提供更适合国内用户需求特点的图像，并将核心用户定位于专业人群和商业行为。在产品侧，Tiamat 正在努力积累用户和打造社区，已经实现了内测高活跃用户数千人、月度新增数万人，社区用户涵盖国内设计、艺术创作、多模态技术等头部博主；同时，有超过三分之一的用户有持续付费的意愿，C 端商业化能力可观。在 B 端用户层面，Tiamat 与众多头部游戏公司、时尚期刊等展开合作，并联合德高申通，将社区用户创作的作品投放至多个上海地铁沿线站点。Tiamat 还将上线专业版创作平台，品牌方在网页工具中上传任意一张产品图片素材，输入想要的场景背景，即可快速生成产品广告图。



中山大学人机物智能融合实验室

中山大学人机物智能融合实验室围绕“人工智能前沿技术与产业化”布局研究方向与课题，并深入应用场景打造产品原型，输出大量原创技术及孵化创业团队。在多模态认知计算、机器人与嵌入式系统、元宇宙与数字人、可控内容生成等领域开展体系化研究。实验室由林惊教授创办于2010年，目前由包含3名国家级人才的教授团队领衔，其中林惊教授任教育部超算工程软件工程研究中心副主任，入选 IAPR Fellow、IET Fellow；梁小丹教授曾任卡耐基梅隆大学研究科学家，获得阿里巴巴青橙奖。

实验室承担或者已完成各级科研项目60余项，包括国家2030重大项目、国家重点研发计划项目、国家自然科学基金委-广东大数据科学中心项目、国家自然科学基金委重点项目等，共获得科研经费超过5千万元。实验室同样注重成果转化，与知名IT及人工智能企业，例如中兴通讯、华为昇腾计算事业部、华为诺亚方舟实验室、广州视源科技等，开展广泛合作，完成企业委托研发项目。

2016年至2018年期间，实验室与人工智能独角兽企业商汤科技共建研发团队，大量成果应用于商汤科技的视频大数据分析、手机拍照成像、自主机器人等核心业务线。2019年至2021年实验室核心成员与暗物智能科技深度合作，致力于打造强认知人工智能技术平台。2022年实验室孵化创新企业拓元智慧，打造通用多模态认知AI技术平台，深度赋能元宇宙及内容创造各主要应用场景。

一、最新动态

中山大学人机物智能融合实验室在 CVPR 2023 (IEEE Conference on Computer Vision and Pattern Recognition) 共有15篇论文被接收，涉及了AI内容生成、面向开放词汇类别的视觉理解、数据高效的视觉分类与检测、视觉语言多模态理解等多个计算机视觉核心研究领域。

(CVPR是计算机视觉和模式识别领域的顶级国际会议之一，自1979年以来每年举办一次。它是世界三大顶级计算机视觉会议之一，也是中国计算机学会推荐的国际学术会议排名中人工智能领域的A类会议。)以下是十五篇论文：

(一) AI内容生成

- 1、零样本开放式文本到动作的生成
- 2、基于局部流和全局解析协作学习的通用虚拟试穿方法
- 3、人脸关键点和外观先验驱动的身份数字人说话视频生成
- 4、基于参数化隐式表达的音频驱动数字说话人视频生成

(二) 面向开放词汇类别的视觉理解

- 1、视觉范例提示驱动的自动驾驶统一感知任务
- 2、统一预训练的稠密描述生成和开放域检测
- 3、基于现实点云数据的语言-图像-点云对比预训练
- 4、基于词区域对齐的可扩展的开放词汇物体检测预训练

（三）数据高效的视觉分类与检测

- 1、基于改进分布匹配的数据集蒸馏方法
- 2、基于 DETR 框架的半监督物体检测
- 3、基于定制学习的主动领域自适应方法
- 4、基于图像掩码的反事实样本构造及模型微调方法

（四）视觉语言多模态理解

- 1、基于点到点的指称物体分割
- 2、基于动态图增强对比学习的医学报告生成
- 3、场景知识增强的跨模态视觉目标定位：标准数据集与方法

二、科研成果

【DreamArtist——通过对比提示调整实现可控的一次文本到图像生成】

具有指数进化的大规模文本到图像生成模型目前可以基于文本引导合成高分辨率、特征丰富、高质量的图像。然而，他们经常被总是出现的新概念、新风格或对象实体的词语淹没。尽管最近有一些尝试使用微调或即时调整方法来教导模型一个新概念，作为来自给定参考图像集的新伪词，但这些方法不仅仍然难以在没有失真和伪影的情况下合成多样化和高质量的图像，而且可控性也很低。

为了解决这些问题，中山大学人机物智能融合实验室提出了一种 DreamArtist 方法，该方法采用对比提示调整的学习策略，将正嵌入和负嵌入作为伪词引入，并对它们进行联合训练。正嵌入积极学习参考图像中的特征，以驱动模型的多样化生成，而负嵌入则以自我监督的方式进行反思，以纠正嵌入中的错误和不足。它不仅学习什么是正确的，而且学习什么是应该避免的。在图像质量和多样性分析、可控性分析、模型学习分析和任务扩展方面的大量实验表明，我们的模型不仅学习概念，还学习形式、内容和上下文。DreamArtist 的伪词与真词具有相似的财产，可以生成高质量的图像。

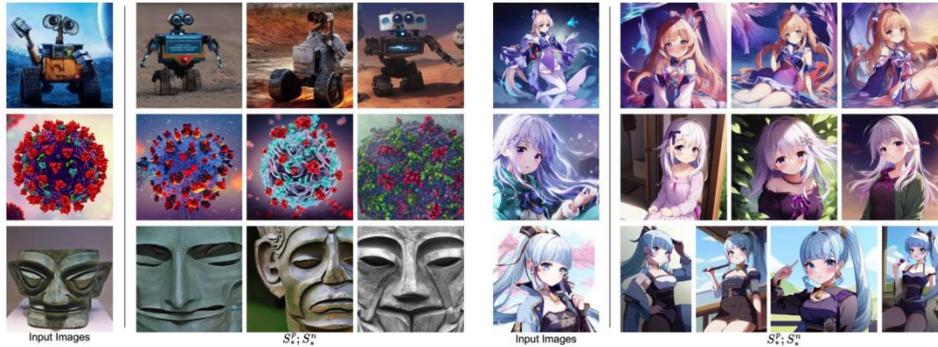
DreamArtist 可以教授文本到图像的扩散模型，以学习一个只有一个图像作为配对伪词的新概念。这些来自 DreamArtist 的伪单词可以生成高质量、多样化和高度可控的图像，就像模型的其他原始单词一样。它可以用一个给定的参考图像有效地生成高质量和多样化的图像，并在不同的上下文中包含各种新的呈现形

式。

功能：

1、Generation without Text Descriptions 图生图

使用 DreamArtist 创建绘画作品，不需要额外的文本描述便可以生成图片。DreamArtist 不仅可以生成具有显著光、影和细节的高度逼真的图像，还可以保持生成的图像的高度多样性。



2、Style Clone 样式克隆

与文森特·梵高的《星夜》这样的抽象风格相比。通常情况下，用户需要一些更实用的风格，例如电影或游戏场景的风格，赛博朋克或蒸汽朋克的风格，中国画笔或剪纸的风格，或者你最喜欢的艺术家的绘画风格。DreamArtist 可以很好地学习一些实用且高度精致的风格。生成的图像在颜色和纹理特征上与训练图像相同，显著地克隆了训练图像的风格。



3、Concept compositions 概念

DreamArtist 可以轻松地组合多个学习到的伪词，不仅限于组合对象和风格，还可以同时使用对象或风格，并生成一些合理的图像。将 DreamArtist 训练的多个伪单词组合在一起，无论是在自然场景还是动画场景中都能显示出出色的效果。伪单词的每个组成部分都可以在生成的图像中呈现，甚至可以组合两个完全不同

的对象或样式。这些对于现有的方法来说是难以实现的。



4、Prompt-Guided Image Editing 提示引导图像编辑

使用该方法学习的伪单词适用于遵循稳定扩散范式的文本引导图像编辑。修改后的区域不仅展示了学习的形式和内容，而且很好地融入了环境，看起来很和谐。使用该方法的学习特征进行图像编辑的性能与来自稳定扩散的原始特征一样有效。这些对于现有的方法来说是难以实现的。



小结

目前国内的 AI 作画软件层出不穷，有人工智能大厂依托强大的算力涉足 AI 绘画赛道，亦有初创公司开发的垂直应用。

从底层技术上来看，一类企业拥有自研文生图大模型，另一类企业选择使用开源模型进行二次开源或接入 OpenAI、百度文心等第三方模型。自研文生图大模型的企业包括百度、商汤科技等头部企业，也有西湖心辰这种专注于多模态大模型研究的创业企业。自研文生图大模型能够依据中文语言习惯、本土市场需求和中国绘画风格进行模型的开发更适应于本土市场。但是其开发难度和对算力成本的要求较高。AIGC 浪潮之下，一些设计公司、软件公司也开始涉足 AIGC 绘画，依托自身在设计领域或者软件开发领域的基础，快速接入第三方模型，开展自己的 AIGC 业务。

	产品	公司主体	技术路径
自研模型	文心一格	百度在线网络技术（北京）有限公司	文心跨模态大模型
	秒画 SenseMirage	北京市商汤科技开发有限公司	超十亿参数文生图模型
	造梦日记 Printidea	西湖心辰（杭州）科技有限公司	双模态图文模型
二次开源	万兴爱画	万兴科技集团股份有限公司	接入 OpenAI
	触手 AI	杭州水母智能科技有限公司	基于 SD 模型二次开源
	画宇宙	北京计算美学科技有限公司	基于 SD 模型优化的画宇宙模型接入文心大模型
	天工巧绘 SkyPaint	北京奇点智源科技有限公司	多语言的 Stable Diffusion 分支模型（与昆仑万维合作研发）
	Tiamat	上海退格数字科技有限公司	MorpherVLM 近百亿级跨模态生成模型

从最底层的算法模型和技术原理上来看，目前主流的文本生成图像的厂商做都算是 Diffusion（扩散）模型的各种变体。Stable Diffusion 所开源的模型也是基于其中一类 Diffusion 模型的工程化实现。SD 模型的优点在于，它采用的 Latent Diffusion 模型会先把图片进行某种“压缩”、并转换成某种潜在的表达后，再做后续计算，使得图片生成的运算量降低，运算速度大幅提升。至于其他的主流 Diffusion 模型，以 Disco Diffusion 为例，虽然他们的效果甚至可能会更好一些，但速度就会比较慢，很难满足商业化的需求。

AI 作图要想真正走向 B 端市场，满足垂直领域企业的设计需求，AI 作画的可控性至关重要。目前的 AI 绘画都是基于扩散算法的，这个算法的优势在于它的随机性，随机性可以给我们带来很多意想不到的结果，扩散模型生成的每一张图都不一样。这个特性最大的好处就是可以带来无穷无尽的想象力，很多时候这些想象力甚至会超越人类自己。但是这些随机性带来的想象力有一个很大的问题，就是它不可控。大家都戏称这种创作叫抽卡式创作，你也不知道最终的生成效果是什么样的。有时候运气好出来一张非常惊艳的图，但是不可复现。

这就是去年开始早期 AI 绘画的状况，它虽然可以生成很多漂亮的图片了，但是它本质上还是一个玩具，因为所有的设计领域都需要对结果可控，对于设计师来说，不能修改，不能控制画面中的内容，无法精准地满足甲方的要求，这种工具也只能在早期头脑风暴的阶段使用。但是后面出现的微调模型和 controlnet，把这个状态打破了。controlnet 的出现补完了 AI 绘画工业化生产的最后一块拼图，AI 绘画这个工具正式地进入设计师的工作流了。

国内的 AI 图像软件大都是基于 Stable Diffusion 模型的二次来源和优化。因此，各个应用并没有绝对的优势，技术门槛不高、图像生成的可控性也远没有达到大规模商用的水平。当前的 AI 绘画应用都是开源算法加上自己特有的一些算法和模型，所以谁都没有底层能力。从用户层面来讲，当前阶段对于普通用户来说，文生图不是刚需，平台通过重度付费用户付费来获得商业收益，很难达到收支平衡。

因此，To B 的商业模式需要重点关注，除了直接向客户提供算力服务，国内 AI 绘画公司正尝试与文娱、地产、快消等各类 IP 和企业进行合作，已有一些落地的营销案例。普遍采用 toB 的 SaaS 服务的商业模式：作为一种 AIGC 解决方案，给 B 端的客户做营销或者是创造平台，让 B 端用户用它画广告或者视频。

对于 AI 绘画软件开发的公司来讲，哪家企业的平台能够真正解决垂直领域 B 端用户的需求，帮助 B 端用户“降本增效”才是获得盈利的关键。因此，我们关注到企业在模型开发上的能力、图像语料的丰富程度、软件功能的垂直性等因素。而对于一些面向 C 端用户的娱乐性绘画产品，我们更多是考量其学习成本

的高低、功能的丰富性与趣味性、绘画风格的多样性等因素。

技术是中性的，当 AI 绘画的浪潮席卷而来，或许不同的对待方式会有不同的结果。一个有趣的类比是，汽车刚发明的时候，一部分马车夫成为打砸内燃机的卢德分子，另一部分马车夫当上了司机。

协会之窗

南京芯视元电子有限公司

一、企业简介

南京芯视元电子有限公司是一家专注于智慧显示芯片研发的高新技术企业。公司坐落于“芯片之城”南京江北新区核心地带。公司产品主要有硅基 LCoS 微显示芯片、硅基 OLED 微显示芯片、硅基 Micro LED 微显示芯片、空间光调制器。产品广泛应用于 AR/VR/MR 眼镜、车载 HUD、头戴显示器、光通讯、光计算、3D 打印等新兴领域。

芯视元凭借过硬的技术、一流的产品、全面的服务在业内具有良好的声誉。一直以来对自主创新的坚持，使得芯视元在行业中具备独有的竞争优势，获得了众多企业的认可并达成战略合作关系。

AI 为脑、AR 为眼，未来已来。芯视元欢迎更多的合作伙伴一起来推动产业发展，共创美好未来！

芯视元获得高新技术企业认定、3A 企业信用等级认证、质量管理体系认证、2018 年工信部主办的芯片大赛企业组一等奖、2019 “创业金陵”南京科技创新创业大赛成长组一等奖。



此外，芯视元还拥有 80 多件专利，集成电路设计布图 15 件、软件著作权 4 件。

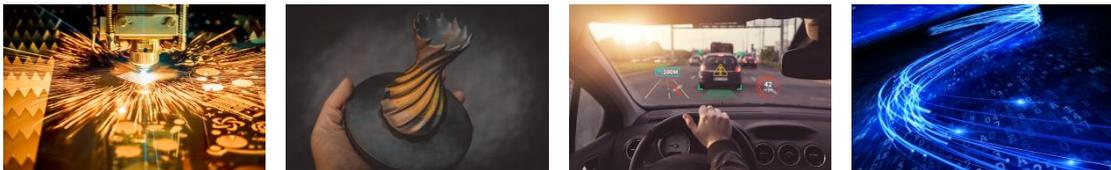


二、企业探索

纵观全球可独立设计硅基背板的厂商寥寥可数。业内企业通常选择专攻某一细分领域，如 LCoS，亦或将 OLED、Mirco LED 作为主战场，鲜有企业能做到多元化发展。这不仅意味着前期大量的投入，更是一场风险未知的持久战。但在芯视元 10 年的发展轴线里，逐步拓展产品线，持续推出产品系列服务行业上下游企业，同时更对自己提出“LCoS+OLED+Mirco LED”三方面齐头并进的高要求。从自主研发到批量提供硅基 LCoS、OLED 和 Mirco LED 等相关芯片，芯视元用技术说话，获得了业内合作伙伴的广泛认可，向业界提交了一份“不止于此”的优秀答卷。



芯光 200



其中“芯光 200、天目 III-010-A”，作为芯视元公司 2021 年度明星产品，曾亮相于 2022 第三届中国半导体投资联盟年会，参与竞选中国 IC 风云榜“年度优秀创新产品奖”。两款新品有望在不同领域大放异彩，被广泛寄予厚望。其中，芯光 200 系列下的两款产品，芯光-I-200 和芯光-II-200，均搭载了自主研发、具有完全自主知识产权的高性能硅基液晶芯片(LCoS)，在 360hz 驱动频率下可达到低于 200mW 的功耗。它的高开口率、高光利用率、高衍射效率、高相位调制精度、大相位调制和低 phase flicker 特性拓宽了应用范围。芯光-I-200 主要针对可见光波段，应用于全息成像，3D 打印等领域；芯光-II-200 主要作用于 1550nm 波段，分辨率为 1920*1080，可应用于光通讯、光计算等尖端领域；天目 III-010-A 是第一款公司自主研发的 Micro LED 芯片，因亮度高、响应快，颇受 AR/VR HMD、轻量型 AR 眼镜、车载抬头显示、智能车灯等新兴市场和领域的青睐。



参数	值	参数	值
芯片规格	0.39"	芯片规格	0.26"
有效显示尺寸	8.62 mm x 4.86 mm	有效显示尺寸	5.76 mm x 3.24 mm
灰阶	128 (7bit)	灰阶	256 (8bit)
显示分辨率	1920x1080	显示分辨率	1280x720
像素尺寸	4.5 um x 4.5 um	像素尺寸	4.5 um x 4.5 um
PPI	5645	PPI	5645

今年3月，在“洞察元宇宙，数字新江北”2023中国元宇宙产业生态大会上，芯视元展示了两款面向XR应用的硅基Micro LED显示背板SVC2K39M和SVC1K26M。目前国内外已经有十多家Micro LED模组厂采用这两款显示背板制作Micro LED显示模组，其中单绿色产品亮度最高可达上百万尼特，单红色产品和单蓝色样品亮度最高也能到几十万尼特。

三、未来展望

目前，集成电路产业和VR&AR产业处于迅猛增长时期，用户群体在不断的扩大，为满足广大VR&AR用户优质显示芯片的需求，必须提高硅基LCoS的用量和参数指标要求以满足市场的需求。随着公司不断发展，以芯视元自身技术研发优势，势必可以提高技术投入和研发力度，在深耕微显示芯片领域的技术优势日益凸显。可以期待，在未来将会有更多产品持续落地，现实世界与虚拟世界的距离将进一步缩小。

芯视元始终秉持“感知未来，智慧显示；同建生态，合作共赢”的理念。一直专注于微显示技术领域的研发，坚信技术才是产品的驱动力，芯视元创始人何军表示：我们会继续深耕于微显示芯片技术研究领域，在AR/VR/MR眼镜、车载HUD、头戴显示器、光计算、3D打印等新兴领域全面布局，致力于成为具有国际一流竞争力的微显示芯片领域龙头。我们认为的未来是一个人工智能的世界、物联网的世界。人工智能为脑，增强现实为眼，它们是表里关系，虽然感觉不到，但确实已经离不开。“AI+AR=AIR”，这是创始人何军9年前创造的等式，以它来定义企业未来。同时，何军也呼吁业界关注AR/VR，“如果说过去的15年到20年是‘互联网+’时代，那么未来10年、20年甚至更长的时间里，是‘XR+’时代。”

附录

产品名称	公司主体	发展状况	创始团队	核心技术	应用场景	商业模式	
国外公司	Midjourney	Midjourney	初创公司	创始人: David Holz, 北卡罗来纳大学教堂山分校, 应用数学系。	Midjourney V5 model	产品设计、服装设计、广告设计、艺术创作、模拟场景等领域。	采用SaaS付费。
	Stable Diffusion	Stability AI	独角兽企业	创始人: Emad Mostaque, 牛津大学, 数学和计算机专业。	Stable Diffusion	计算机视觉、数字艺术、视频游戏等领域。	收费产品: DreamStudio
	DALL-E 2	OpenAI	独角兽企业	创始人: Sam Altman, 8岁学会编程, 20岁从斯坦福辍学创业。	DALL-E系列模型	广告娱乐、游戏、教育、汽车设计图纸、制造零售等领域。	API费用以及软件许可。
国内上市公司	文心一格	百度公司	上市公司	创始团队核心成员: 王海峰, 哈尔滨工业大学, 任深度学习技术及应用国家工程研究中心主任。	飞浆、文心大模型	教育、娱乐、设计、艺术等领域。	“一定免费次数+付费点数”的模式。
	万兴爱画	万兴科技集团股份有限公司	上市公司	创始人: 吴太兵, 曾任银行系统工程师, 上海交通大学, 安泰管理学院硕士。	接入OpenAI	C端市场	“一定免费次数+付费点数”的模式。
	天工巧绘 SkyPaint	昆仑万维科技股份有限公司	上市公司	创始人: 周亚辉, 清华大学, 精密仪器系。	多语言Stable Diffusion分支模型(与奇点智源合作研发)		“一定免费次数+付费点数”的模式。
	秒画 SenseMirage	北京市商汤科技开发有限公司	上市公司	创始人: 汤晓鸥, 本科毕业于中科大, MIT博士学位, 香港中文大学信息工程系教授。	基础大模型“日日新 SenseNova”、商汤自研超十亿参数文生图模型	B端: 设计、营销、游戏动画、工程建筑、医学、教育科普等领域。	
国内初创公司	触手AI	杭州水母智能科技有限公司	A轮	创始人: 苗莢, 设计和创新管理海外双硕士。	基于Stable Diffusion二次开源	C端市场	“一定免费次数+付费点数”的模式。
	造梦日记	西湖心辰(杭州)科技有限公司	Pre-A轮	创始人: 蓝振忠, 卡耐基梅隆大学, 人工智能博士。	自研双模态图文模型, 支持中英文双语。	B端+C端市场	次数和会员两种收费方式。B端和C端同时进行。
	画宇宙	北京计算美学科技有限公司	Pre-A轮	创始人: 徐作彪, 清华大学美术学院, 艺术设计-工业设计专业硕士。	基于SD模型优化、接入百度文心大模型	B端: 建筑设计、工业设计、包装设计等领域。	提按量、月租、部署等多种API费用结算模式。
	Tiamat	上海退格数字科技有限公司	A轮	创始人: 毛钰炜, 上海科技大学, 信息科学与技术学院, 计算机科学与技术专业。	基于扩散模型和审美模型进行优化		聚焦中小型企业定制化合作需求。
	Vega AI	北京右脑科技有限公司	初创公司	创始人: 梁健, 北京大学计算机科学硕士。	基于Stable Diffusion二次开发		

产品名称	公司主体	用户量级	绘画风格	适用功能	生成速度	收费标准	
国外公司	Midjourney	Midjourney	超1300万用户。	多样	人像卡通化、轮廓生成、色彩生成、视频换脸、视觉问答、人脸合成。	60秒。	每月10-60美元。创作的图像用于商业用途，收益超过每月2万美元需支付20%作为分成。
	Stable Diffusion	Stability AI	超1000万用户。	多样	生成风景类画、艺术性作品、复现相同图片、根据中文古诗词进行绘画、动漫角色绘制等。	10-20秒。	免费
	DALL-E 2	OpenAI	超300万人。	多样	图生图、文生图、根据局部风格调整成像、根据原始图像创建变体。	10-20秒。	创作需消耗积分，每月可免费领取一定数量，15元可购买115积分。生成每张图片约0.033美元。
国内上市公司	文心一格	百度公司	累计用户500万。	27种	文生图、图生图。	2分钟内。	创作需消耗电量，任务活动可免费领取一定数量，39元可购买100积分。生成每张图片约0.78元。
	万兴爱画	万兴科技集团股份有限公司		17种	AI文字绘画、AI以图绘图、AI简笔画。	30秒，AI简笔画可在5秒内完成。	20次免费生成额度。
	天工巧绘 SkyPaint	昆仑万维科技股份有限公司		5种	二次元模式、头像生成器。	约10秒内。	每次消耗5金币，新用户免费发放200金币。
	秒画 SenseMirage	北京市商汤科技开发有限公司			文生图。	2秒。	根据B端用户的需求，为客户训练并封装成SDK交付。
国内初创公司	触手AI	杭州水母智能科技有限公司		数百种	文生图。	普通通道，1分钟左右。	创作需消耗积分，任务活动可免费领取一定数量，6元可购买600积分。生成每张图片约0.15元。
	造梦日记	西湖心辰(杭州)科技有限公司	累计C端用户超百万，B端用户十余家。	38种	手绘增强，给草图上色、图生图等。	理论出图速度0.8秒，审图速度1-2秒。	创作需消耗星星，任务活动可免费领取一定数量，12元可购买150颗。生成每张图片约0.16元。
	画宇宙	北京计算美学科技有限公司		267种	文生图、图生图、图像外延、四方连续生成、智能商品图、智能人像等。	5分钟内。	创作需消耗画力星，任务活动可免费领取一定数量。会员58元每月，不限量使用。
	Tiamat	上海退格数字科技有限公司	内测高活跃用户数千人、月度新增数万人。	多样	文生图。	60秒。	20次免费生成额度。
	Vega AI	北京右脑科技有限公司		8种	文生图、图生图、条件生图。	16秒。	免费



中国（南京）文化和科技融合 成果展览交易会

CHINA NANJING

Culture & Technology Integration Exhibition Achievement Fair

中国（南京）文化和科技融合成果展览交易会是全国首个聚焦文化和科技融合的专题性展览交易平台，每年在南京举办。展会由江苏省人民政府指导，江苏省委宣传部、江苏省科学技术厅、江苏省委网信办、江苏省文化和旅游厅、江苏省广播电视局、南京市人民政府联合主办，南京市委宣传部、南京市科学技术局、南京市委网信办、南京市文化和旅游局、南京市文化投资控股集团联合承办，围绕“新时代 新体系 新体验”主题，聚焦广播影视、动漫游戏、演艺娱乐、文化旅游等行业产业数字化转型需求，展示和交易新技术、新产品、新一代解决方案，树立文化科技融合的“风向标”，展示数字文化产业的“创新源”，打造产业数字化的“指南针”。展会获得 2020 年、2021 年中宣部《文化体制改革和发展工作简报》专报推广，评为“2019-2020”年度全国品牌影响力会展项目。

联系人：缪岑岑

联系电话：13270888656



紫金文化科技融合创新研究院

Purple Academy Of Culture And Technological Integration & Innovation

紫金文化科技融合创新研究院是在国家文化科技创新工程专家组、江苏省委宣传部指导下，由南京市文投集团、紫金文创研究院牵头发起的全国首家文化科技融合创新研究平台。研究院下设行业研究中心、创新联合体中心、产业金融中心、交流展示中心，成立由国家文化科技创新工程专家组组长张树武教授领衔、全国高校专家学者和龙头企业代表构成的专家组，以开放性眼光打造全国文化科技融合高端智库、新型研发机构孵化苗圃、新兴文化业态培育基地、龙头文化企业加速平台。

联系人：黄庭

联系电话：13912992856



南京市文化投資控股集團

Nanjing Cultural Investment Holding Group

南京市文化投資控股集團成立於 2009 年 9 月，是南京市委、市政府直屬的大型國有文化企業。十餘年來，南京市文投集團始終秉承“讓城市文化生活更美好”的文投使命，始終堅守“國有文化資產管理運營主體、重大文化項目投資建設主體、新興文化業態投資主體、美好文化產品示範供給主體、城市文化 IP 孵化培育主體”五大主體定位，聚焦建設互聯網文娛、十竹齋藝術、文化綜合體、文化金融投資、文化科技創意五大產業板塊，共築“扎根南京、輻射南京都市圈、示範江蘇、走向全國的城市文化生活示範運營商”的企業願景，致力成為全國文化龍頭企業三十強和全國國有文投集團前三強。目前集團管理資產總量超百億元，營業收入和利潤每年保持 20% 增速，實現了社會效益和經濟效益雙豐收；2019 年-2021 年，連續三年榮膺“全國文化企業 30 強”提名單位。

聯繫人：江浩

聯繫電話：13851814061





扫描二维码，在线填写读者满意度调查问卷

报：中央文化体制改革和发展工作领导小组办公室、科学技术部
办公厅、文化和旅游部办公厅

送：国家文化和科技融合示范基地所在地各省、自治区、直辖市
宣传。部办公室、科学技术厅办公室、文化和旅游厅办公室

发：国家文化和科技融合示范基地所在地各市宣传部办公室、科
学技术局办公室、文化和旅游局办公室

中国（南京）文化和科技融合成果展览交易会组委会
紫金文化科技融合创新研究院

2023年05月01日发
