

# AIGC之AI绘画：技术与应用双突破，生产力变革在即

## 投资要点

□ **AI绘画是AIGC重要的应用分支。近两年包含扩散模型在内的关键技术取得突破，技术可用性显著提高，技术转化为生产力的契机产生。随着Stable Diffusion等应用破圈，用户接受度和参与度持续提高，适用行业不断拓展，未来B端、AI绘画+3D的商业化潜力值得期待。**

□ **发展进入快车道，迎来转变为生产力的拐点**

2022年被称为AIGC元年。8月，凭借AI绘画作品《太空歌剧院》，参赛者没有绘画基础却获得美国科罗拉多州新兴数字艺术家竞赛一等奖，引发热议。Stable Diffusion、Midjourney等AI绘画应用推动技术进入民用领域，几月内产出效果便有直观改善，被设计师、游戏工作室等广泛用于辅助生产，生产力变革来临。

□ **关键技术取得突破，图像生成效果效率均显著提升**

AI绘画是基于深度学习技术的。生成式对抗网络GAN配合可对比语言-图像预训练算法CLIP，解决了跨模态问题，支持文本生成图像；而AI绘画的实操可分为四个环节：加噪点、去噪点、复原图片和作画。随着图像掩码建模MIM、特征处理器Transformer、扩散模型Diffusion Model和神经辐射场NeRF出现，在上述四个环节中发挥作用，AI“画技”显著提升。未来，我们预计深度学习领域将有两大主要前进趋势：大模型和人工通用智能。

□ **商业化前景广阔，B端和三维化或为突破口**

AI绘画产品不断丰富，体验持续提升，已具备较好的用户基础。但AI绘画产品目前少有营收或实现盈利，目前相关产品变现方式较为单一。对于普通C端用户，缺乏性价比加高的商业应用场景，付费意愿不强。已有的赋能收费项目集中在付费提速或者增加清晰度方面，变相弥补现有使用局限。我们认为：

- B端变现路径更为多元、成熟，如广告和营销行业均有可想见的应用情景能挖掘出较为可行的商业模式，付费的可能性和水平相对更高；
- 此外，人们对于未来元宇宙的期待是三维化、AI化及开放式的，AIGC+3D是必由之路。AIGC+3D是丰富游戏、影视、VR等数字内容，降低其制作成本的有力工具，目前已有部分2D产品具备3D迁移能力。

□ **产业链初具规模，上下游现蓝海**

产业链层面，AI绘画涉及到硬件、NLP、算法算力、应用、数据提供与处理等多环节，当前产业布局于算法和应用开发环节较为集中和领先，而在产业链上下游还有诸多可开发的蓝海领域。

具体到投资标的：

- 首先是拥有相关应用产品和场景的公司。图文类推荐关注视觉中国（素材库+数字藏品）、中文在线（AI绘画+AI文字+小说平台）、万兴科技（“万兴AI绘画”软件）、三人行（AI平面设计+营销场景）；AIGC+3D作为动态数字内容辅助创作工具，或可显著降低游戏、影视、VR/AR行业制作成本，利好相关厂商，推荐关注腾讯控股、网易、完美世界；其他AIGC模态推荐关注昆仑万维（Star X音乐平台+AI作曲）、蓝色光标（虚拟人+营销场景）；

## 行业评级：看好(维持)

分析师：谢晨

执业证书号：S1230521070004

xiechen@stocke.com.cn

研究助理：姚逸云

yaoyiyun@stocke.com.cn

## 相关报告

- 1 《外部环境致总体消费表现平缓，直播电商仍有超额增长——双11点评》 2022.11.13
- 2 《直播电商高歌猛进，头部双星持续闪耀——直播电商专题报告》 2022.10.31
- 3 《瑞幸生酪拿铁首周659万杯再创纪录，爆品策略再次验证——行业点评报告》 2022.10.27

- AIGC 和 AI 绘画依赖自然语言处理、计算机视觉和人工智能技术，我们推荐关注百度集团（文心大模型+文心一格）、拓尔思（语义智能技术）、商汤（人工智能+计算机视觉）、科大讯飞（语音识别+人工智能龙头）等。

□ **风险提示**

- 1) AI 绘画相关技术发展不及预期风险；
- 2) 商业化拓展进度和效果低于预期风险；
- 3) AI 绘画涉及的作品侵权、名人肖像侵权等风险。

## 正文目录

<b>1 发展进入快车道，迎来转变为生产力的拐点</b>	<b>5</b>
1.1 AI 绘画是 AIGC 重要的应用分支	5
1.2 今年起 AIGC 发展按下加速键	6
1.3 技术转化为生产力的契机产生	8
<b>2 关键技术取得突破，图像生成效果效率均显著提升</b>	<b>8</b>
2.1 GAN+CLIP 解决跨模态问题	8
2.1.1 生成式对抗网络 GAN—图像到图像的生成	8
2.1.2 可对比语言-图像预训练算法 CLIP—文字到图像生成	9
2.2 深度学习助力 AI 画技进步	10
2.2.1 图像掩码建模 MIM — 高效简洁的预训练方法	10
2.2.2 特征处理器 Transformer — 优化的自然语言处理模型	11
2.2.3 扩散模型 Diffusion Model — 新一代图像生成主流模型	12
2.2.4 神经辐射场 NeRF — 顺应 3D 内容消费趋势	14
2.3 大模型和人工通用智能指引发展方向	16
2.3.1 大模型催生基石模型公司崛起	16
2.3.2 人工通用智能不仅仅是想象	16
<b>3 商业化前景广阔，B 端和三维化或为突破口</b>	<b>16</b>
3.1 应用迅速丰富，用户接受度较高	16
3.2 变现仍处于尝试阶段，B 端或为切入点	18
3.3 三维化打开游戏、影视、VR 等应用空间	19
3.4 产业链初见规模，部分空缺现蓝海	21
<b>4 基建待完善，发展中风险与希望并存</b>	<b>22</b>
<b>5 建议关注</b>	<b>23</b>
<b>6 风险提示</b>	<b>24</b>

## 图表目录

图 1: 内容生产阶段.....	5
图 2: AIGC 按模态分类.....	5
图 3: AI 画作《太空歌剧院》.....	6
图 4: AIGC 历史沿革.....	7
图 5: “AI 绘画”百度搜索指数.....	7
图 6: 百家号的智能图文转视频功能.....	8
图 7: GAN 工作方式示意图.....	8
图 8: GAN 根据语义图像或素描生成现实图片.....	9
图 9: PGGAN 工作原理示意图.....	9
图 10: StyleGAN 表征控制示意图.....	9
图 11: CLIP 工作原理图示.....	10
图 12: StyleGAN 效果示意图.....	10
图 13: MAE 的非对称编码器-解码器结构.....	11
图 14: MAE 的高比例遮蔽.....	11
图 15: VIT 工作原理图示.....	12
图 16: 扩散模型工作原理图示.....	13
图 17: Imagen 架构图 (纯语言模型 T5-XXL+Diffusion).....	14
图 18: GANverse3D 图例.....	14
图 19: 人物 NeRF 模型和场景 NeRF 模型结合新场景.....	15
图 20: 按参数数量和发布日期划分的深度学习模型.....	16
图 21: AI 作画操作流程示意图.....	17
图 22: AI 绘画社区元素法典.....	18
图 23: 中国用户为 AI 产品或服务付费比例.....	19
图 24: 中国用户对 AI 绘画的认知评价.....	19
图 25: Dream Fusion 样例.....	20
图 26: Stable Diffusion 样例.....	20
图 27: Get3d 平台上的石窟模型.....	21
图 28: 2020-2025 年 AI 数字商业规模测算 (亿元).....	21
图 29: AIGC 产业链.....	22
表 1: RNN、CNN 和 Transformer 对比.....	11
表 2: 扩散模型在多领域应用实例.....	13
表 3: AI 绘画应用一览.....	17
表 4: AI 绘画应用行业示例.....	19
表 5: AI 绘画概念股一览.....	23

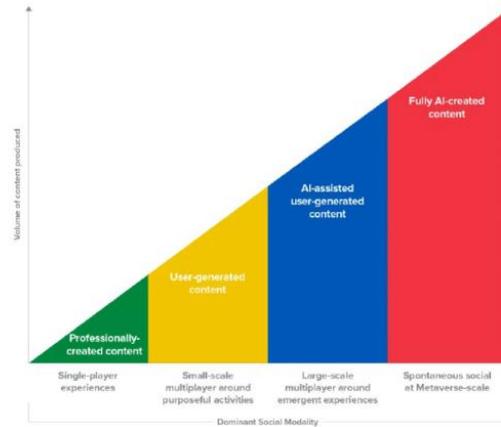
# 1 发展进入快车道，迎来转变为生产力的拐点

## 1.1 AI 绘画是 AIGC 重要的应用分支

AIGC 既是一种内容分类方式，又是一种内容生产方式，还是用于内容自动化生成的一类技术集合。AIGC 全称为 AI-Generated Content，指基于生成对抗网络 GAN、大型预训练模型等人工智能技术，通过已有数据寻找规律，并通过适当的泛化能力生成相关内容的生产方式。

作为全新的内容生产方式，AIGC 潜力无限，而我们当前处于向 AIGC 进发的过渡阶段。根据 a16z，内容生态的发展则可分为四个阶段：专家生成内容（Professionally-Generated Content, PGC）、用户生成内容（User-Generated Content, UGC）、AI 辅助生产内容（AI-assisted Generated Content）及 AI 生成内容（AI-Generated Content, AIGC）。目前我们仍处于一、二阶段为主，第三阶段为辅的境况。

图1：内容生产阶段



资料来源：a16z，浙商证券研究所

按照模态对 AIGC 进行划分最为常见。AIGC 可分为音频生成、文本生成、图像生成、视频生成及图像、视频、文本间的跨模态生成，其中跨模态生成需要重点关注。事实上，人工通用智能（Artificial General Intelligence, AGI）概念的出现反映出许多人工智能业界人士认为，在未来，基础的人工智能模型将是跨模态的，这意味着相同的模型将被用于生成不同体裁的内容，包括但不限于文本、图像、视频等等。

图2：AIGC 按模态分类



资料来源：量子位，浙商证券研究所

AI 绘画是 AIGC 重要的应用分支。AI 绘画工具中，用户通过输入不同的词汇，例如不同的艺术家风格、构图、色彩、透视方法以及修饰词，就能得到相对应的画作。目前相关生成工具已相对成熟，易用性较好，生成结果直观、易传播，部分成果在效率和质量上有不错的表现，切实触动了广大内容生产和消费者，进入主流视野。

图3: AI 画作《太空歌剧院》



资料来源：百度百科，浙商证券研究所

上图展示了 AI 绘画领域具有标志性的事件，2022 年 8 月，在美国科罗拉多州举办的新兴数字艺术家竞赛中，参赛者提交了 AIGC 绘画作品《太空歌剧院》，参赛者没有绘画基础但是却获得了此次比赛“数字艺术/数字修饰照片”类别一等奖，引发热议。

## 1.2 今年起 AIGC 发展按下加速键

PGC 和 UGC 的发展曾为我们带来内容生产和消费的空前繁荣，AIGC 的兴起是生产力革新的成果，亦是孕育自广大消费者的实际需要，具有广阔的想象空间和重要意义。

- 从供给侧角度来看，AIGC 的兴起源于深度学习技术的快速突破。在人工智能发展初期，相关算法多基于预先定义的规则或者模板，AI 在创造力层面进展缓慢，更多地在替代人类从事可重复性高的“dirty work”。近年来，深度学习算法快速迭代，神经网络的规模呈指数级增长，技术强大到可以执行非模板化的工作，更加贴近 AI 所指代的智能化概念。
- 从需求侧角度来看，日益增长的数字内容供给需求驱动着 AIGC 蓬勃发展。传统内容生产手段受限于人力有限的制造能力，在绝对产能、产能与质量的协调关系上愈发吃力。AIGC 依托其技术属性，可襄助内容生产以更加具有绝对效率和性价比的方式进行，更好地满足市场需要。

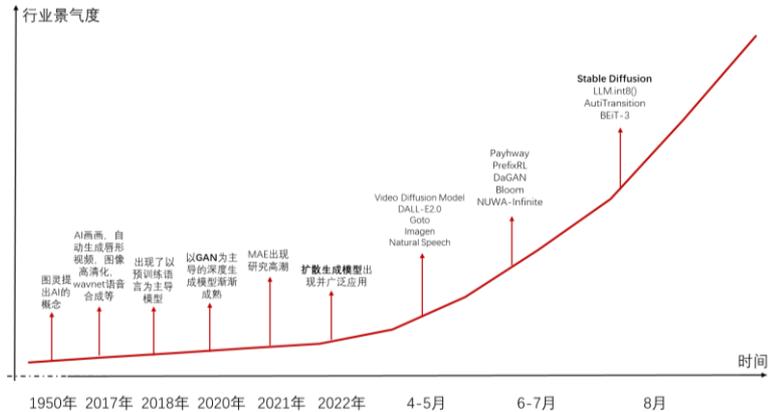
我们将 AIGC 的发展分为四个阶段：

1. **早期萌芽时期：此阶段 AIGC 仅限于小范围实验。**80 年代中期，IBM 基于隐形马尔科夫链模型（Hidden Markov Model, HMM）创造了语音控制打字“坦戈拉（Tangora）”。虽然对 AIGC 进行了一些初步尝试，但受限各种因素，实际产出效果还远远算不上是智能创作内容的程度。
2. **沉淀积累时期：AIGC 从实验性向实用性逐渐转变。**软件上，2006 年，深度学习算法取得重大突破；硬件上，图形处理器（Graphics Processing Unit, GPU）、张量处理器（Tensor Processing Unit, TPU）等算力设备性能不断提升；数据上，互联网使用数据规模快速膨胀并为各类人工智能算法提供了海量训练数据，使人工智能发展取得了显著的进步。2007 年，纽约大学人工智能研究员罗斯古德温装配

的人工智能系统通过对公路旅行中的一切所见所闻进行记录和感知，撰写出世界第一部完全由人工智能创作的小说。2012年，微软公开展示了一个全自动同声传译系统，基于深层神经网络（Deep Neural Network, DNN）可以自动将英文演讲者的内容通过语音识别、语言翻译、语音合成等技术生成中文语音。

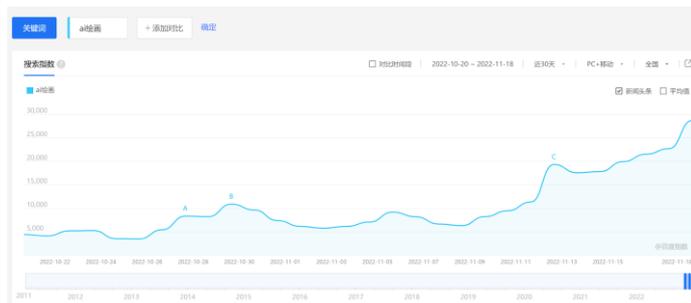
3. **快速发展时期：众多关键模型就位规模化基础已具备。**2017年，AI图像生成、自动生成唇形视频、图像高清化、图像编辑、图像风格迁移、wavnet语音合成等技术快速发展。2018年，预训练语言模型出现，降低了标注需求和成本。2019年，随着以生成式对抗网络（Generative Adversarial Network, GAN）逐渐成熟，AIGC技术研究迎来关键拐点，DeepMind发布了DVD-GAN模型用以生成连续视频，在草地、广场等明确场景下表现突出。2020年，自监督学习成为业界主流，模型体量和复杂度不断提升，其中Open AI发布的CPT3极具代表性。2021年，MAE的出现使得视觉可以用NLP自然语言同样的架构训练预训练模型，叠加多模态多任务领域发展，文本图像对齐的研究爆发。
4. **起飞破圈时期：AIGC概念伴随AI绘画应用出圈。**2022年，技术上，扩散生成模型得到广泛研究与应用，文本生成图像模型可准确把握文本信息进行创作。商业化基础已初步具备，国内外互联网巨头和独角兽纷纷下场。Open AI更新了DALL-E-2，可创作出相应极高质量的卡通、写实，抽象等风格的绘画作品。把AIGC创作最终推向平民化的是Stability.ai推出的Stable Diffusion，个人电脑即可驱动，且几个月内产出效果具有直观的改善。AI绘画迅速在微博、小红书等多平台上，内容创作者、技术研究者、投资人等各圈层里形成了声势。

图4：AIGC 历史沿革



资料来源：， Google Scholar， 浙商证券研究所

图5：“AI 绘画”百度搜索指数



资料来源：百度指数， 浙商证券研究所

### 1.3 技术转化为生产力的契机产生

2022 年被称为 AIGC 元年，这一年 AIGC 取得了里程碑式的成绩，引发了市场广泛热烈的兴趣，我们认为主要因素是：

- 包含扩散模型在内的关键技术取得突破，技术可用性显著提高，产出效率不断提高，产出效果出现分水岭，技术转化为生产力的契机产生。2021 年之前，AIGC 生成的主要还是文字，而新一代模型可以处理的模态大为丰富且可支持跨模态产出，可支持 AI 插画，文字生成配套营销视频等常见应用场景；
- 通过国内外科技公司的应用转化，技术进入民用领域，引起了广大行业从业者、文娱爱好者和投资人的关注。“智能图文转视频”和 AI 作画成功破圈，相关应用迭代速度呈现指数级爆发。以 AI 作画工具为例，水平上限有长足进步，虽发挥并不稳定，但结果输出极快，可量变引起质变，弥补其在创意、想象等方面的不足，满足一般市场需求。

图6：百家号的智能图文转视频功能



资料来源：百家号，浙商证券研究所

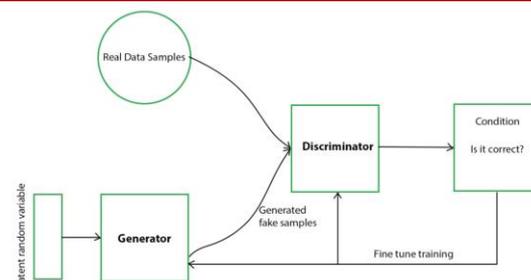
## 2 关键技术取得突破，图像生成效果效率均显著提升

### 2.1 GAN+CLIP 解决跨模态问题

#### 2.1.1 生成式对抗网络 GAN—图像到图像的生成

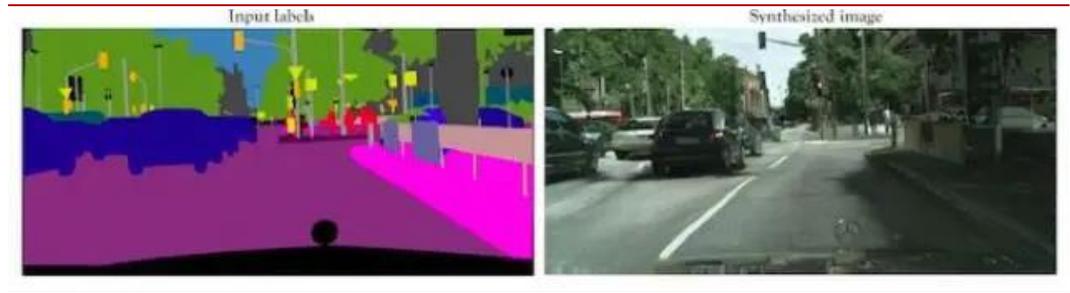
GAN (Generative Adversarial Nets, 生成式对抗网络) 在 2014 年提出后，是生成器和判别器的一代代博弈。生成器通过输入数据生成图像，并将其混入原始数据中送交判别器区分。判别器依据二分类网络，将生成器生成图像作为负样本，真实图像作为正样本。双方的训练过程交替进行，对抗的过程使得生成器生成的图像越来越逼真，判别器的分辨能力也越来越强。

图7：GAN 工作方式示意图



资料来源：Google Scholar，CSDN，浙商证券研究所

图8: GAN 根据语义图像或素描生成现实图片

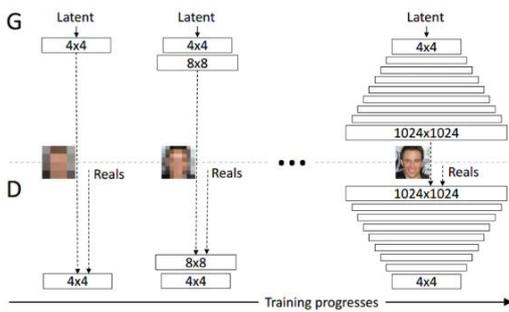


资料来源: 英伟达, 加州大学伯克利分校, 浙商证券研究所

GAN 有三个不足:

1. GAN 对输出结果的控制力较弱, 容易产生随机图像。对此, CGAN 通过把无监督的 GAN 变成半监督或者有监督的模型, 为训练加上目标, 而 DCGAN 通过缩小 CNN 在监督学习与无监督学习之间的差距使得训练过程更加稳定和可控。
2. GAN 生成的图像分辨率较低, 对此, PGGAN 逐渐的向生成器和判别器网络中添加层, 以此增加生成图片的空间分辨率, StyleGAN 则能够生成极其逼真的图像数据并且做到了高层特征可控;
3. 由于 GAN 需要用判别器来判断生产的图像是否与其他图像属于同一类别, 这就导致生成的图像是对现有作品的模仿和微调, 不能通过文字提示生成新图像, 因此 CLIP 被引入。

图9: PGGAN 工作原理示意图



资料来源: Google Scholar, CSDN, 浙商证券研究所

图10: StyleGAN 表征控制示意图

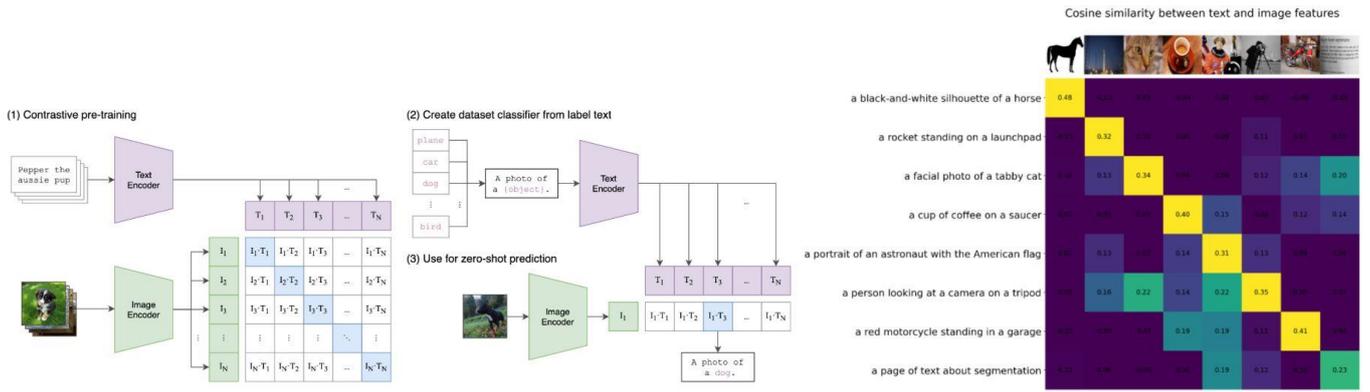


资料来源: Google Scholar, CSDN, 浙商证券研究所

### 2.1.2 可对比语言-图像预训练算法 CLIP—文字到图像生成

CLIP (Contrastive Language-Image Pre-training) 是 OpenAI 在 2021 年提出的多模态预训练的算法, 建于 NLP (Natural Language Processing, 自然语言理解) 和 CV (Computer Vision, 计算机视觉) 相结合的基础上。算法使用已经标注好的“文字-图像”数据对训练。一边对文字进行模型训练, 一边对图像进行模型训练, 不断调整参数, 使得输出的文字特征集和图像特征集相匹配。

图11: CLIP 工作原理图示

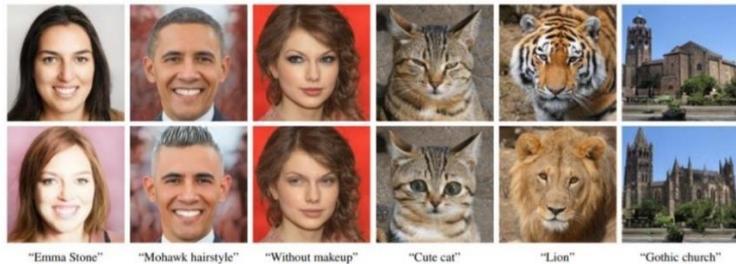


资料来源: Google Scholar, 极智视界, 知乎, 浙商证券研究所

CLIP 方法具有结构简单, 训练速度快, 效果好等诸多优良特性。CLIP 具有非常好的迁移学习能力, 预训练好的模型可以在任意一个视觉分类数据集上取得不错的效果。而且算法是 Zero-Shoot 的, 即不需要再去新数据集上做训练, 就能得到不错的结果。

现被广泛应用的 StyleCLIP 融合了 StyleGAN 和 CLIP 双方的特性。之前的 StyleGAN 的语义控制发现方法会涉及手动检查、大量带注释的数据、或者需要预训练的分类器, 且只能按照预设的语义方向操作图像, 严重限制了用户的创造力和想象力, 若需要添加一个未映射的方向, 需要大量的手工工作或大量的注释数据。StyleCLIP 支持普通用户基于文本的直观语义图像操作, 也不限于预设的操作方向。

图12: StyleGAN 效果示意图



资料来源: Google Scholar, CSDN, 浙商证券研究所

## 2.2 深度学习助力 AI 画技进步

AI 绘画的实际操作大体可以分为四个步骤: 加噪点、去噪点、复原图片和作画。其中, 加噪点即添加高斯噪声或者关键词, 通过固定公式来实现, 这方面, 快速更新迭代的 MIM 方法表现出色。去噪点即仿生物视觉神经网络在去噪过程中开展学习 (透视、颜色等), 目前 Transformer 正取代 CNN 卷积神经网络成为主流方法。而在复原图片和作画方面, AI 的“画技”主要由扩散模型 Diffusion Model 和神经辐射场模型 NeRF 决定。

### 2.2.1 图像掩码建模 MIM — 高效简洁的预训练方法

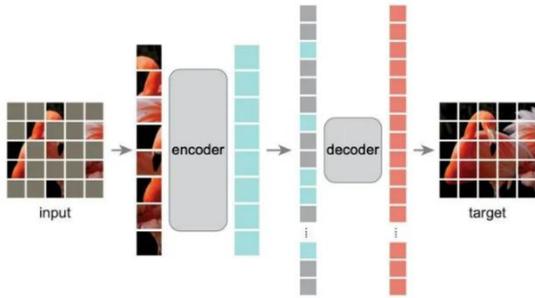
MIM (Masked Image Modeling, 图像掩码建模) 是一种自监督表征学习算法。它的主要思路是, 对输入图像进行分块和随机掩码操作, 然后对掩码区域做一些预测, 进而猜测全图。掩码信号建模在多个模型中应用发展, 例如 OpenAI 的 iGPT 模型 (通过马赛克进行信号的遮蔽和转换)、ViT 模型等。

基于 MIM 的模型在不同类型和复杂程度的广泛视觉任务上实现了非常高的微调精度，使得 AI 作画从生成不完整图像进步到可成完整图像的跨越。MIM 在语义较弱的几何 / 运动任务或细粒度分类任务中的表现明显优于有监督模型；对于有监督模型擅长的任务（语义覆盖较好的语义理解任务），MIM 模型仍然可以取得极具竞争力的迁移性能。

目前较受认可的 MAE 模型产自何恺明对 MIM 的优化。MIM 在预训练图像编码器的时候，太关注细节损失了高维抽象能力。MAE 的非对称编码器-解码器结构，使模型分工明确，编码器负责抽取高维表示，解码器则负责细粒度还原；MAE 同时对输入图像进行高比例遮蔽。将以上两种设计结合，结果用来训练大模型：训练速度提升三倍以上，同时保持高准确率，具备很好的泛化能力。

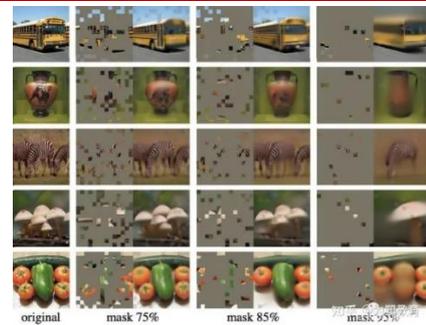
MAE 广泛应用于人脸识别等多个领域。例如，FaceMAE 作为隐私保护人脸识别范式，同时考虑了人脸隐私和识别性能，可以适配任何人脸数据集，以降低隐私泄露风险。

图13: MAE 的非对称编码器-解码器结构



资料来源: Google Scholar, CSDN, 浙商证券研究所

图14: MAE 的高比例遮蔽



资料来源: Google Scholar, CSDN, 浙商证券研究所

由北京大学、香港大学研究者在 2022 年 5 月提出的 CAE 模型、微软亚研院提出的 SimMIM 是对 MAE 方法的改进。CAE 可以更多地挖掘编码器的潜力；而 SimMIM 对 MAE 进行了化简。它们学到的表征可以区分不同类别的物体，举例来说，看到一只猫的头部可以预测出它的身体部分，看到一小片天空可以预测出它的周围大概率也是一片天空。

### 2.2.2 特征处理器 Transformer — 优化的自然语言处理模型

Transformer 是当前综合表现最优的特征提取器。模型首创于 2017 年的 Google 论文《Attention is All You Need》。它的性能优于传统的 RNN 和 CNN 特征提取器。

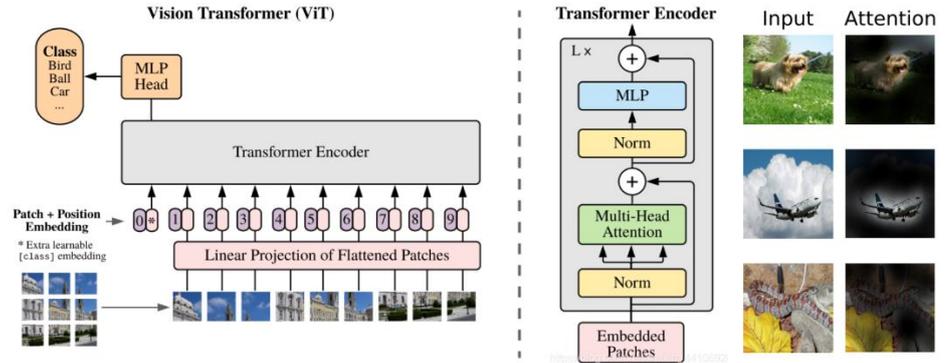
表1: RNN、CNN 和 Transformer 对比

特征处理器	应用范围	优势	劣势
RNN Recurrent Neural Network 循环神经网络	更多被用于 NLP	<ul style="list-style-type: none"> <li>具备分析位置的能力</li> <li>处理不定长度的序列</li> </ul>	<ul style="list-style-type: none"> <li>短期记忆</li> <li>不能并行计算来加速</li> <li>只能识别物体的某一个特征</li> </ul>
CNN Convolutional Neural Networks 卷积神经网络	更多被用于 CV	<ul style="list-style-type: none"> <li>有效地保留位置信息</li> <li>可以高效地进行并行计算</li> </ul>	<ul style="list-style-type: none"> <li>对远距离的特征捕获能力相对较弱</li> </ul>
Transformer	早期更多被用于 NLP, 目前 正向 CV 领域 延伸	<ul style="list-style-type: none"> <li>并行计算能力强, 效率高</li> <li>解决了短期记忆的问题</li> <li>在分类、检测、分割任务上精度超过 CNN</li> <li>与 CNN 形成互补, 学习对图像的全局理解</li> <li>通用且更强的建模能力</li> <li>大模型和大数据可扩展性强</li> <li>视觉与语言更好的连接</li> </ul>	<ul style="list-style-type: none"> <li>计算量大, 参数量大</li> <li>成本增加巨大</li> <li>带来的进步程度不够高</li> <li>实际项目中, 模型转化、量化, SDK 集成等落地链条不完善</li> </ul>

资料来源: Google Scholar, CSDN, 浙商证券研究所

Transformer 为视觉领域带来了革新性的变化，它让视觉领域中目标检测、视频分类、图像分类和图像生成等多个领域实现了长足的进步。2020 年 10 月，谷歌提出了 Vision Transformer (ViT)，它是 Transformer 用于 CV 领域的杰出例子，它在大型数据集上表现处于领先地位。2021 年 1 月，OpenAI 用的 DALL·E 和 CLIP 两个模型都利用 Transformer 达到了较好效果，前者可以基于文本直接生成图像，后者则能完成图像与文本类别的匹配。

图15: ViT 工作原理图示



资料来源: Google Scholar, CSDN, 浙商证券研究所

Transformer 的研究才刚刚起步，因此仍有很大研究和发展空间。在研究领域，CNN 研究已趋向于成熟，考虑到模型成熟度和性价比，CNN 在短期内仍不会被淘汰。

1. 现有的 Visual Transformer 参数量和计算量过大，内存占用量超过可承受范围，效率方面还需要提升，亟需开发高效 Transformer for CV。
2. 现有的 Visual Transformer 都还是将 NLP 中 Transformer 的结构套到视觉任务做了一些初步探索，未来针对 CV 的特性设计更适配视觉特性的 Transformer 将会带来更好的性能提升。
3. 现有的 Visual Transformer 一般是一个模型做单个任务，近来有一些模型可以单模型做多任务，比如 IPT，我们期待未来出现世界模型，处理全局任务。

### 2.2.3 扩散模型 Diffusion Model — 新一代图像生成主流模型

Diffusion Model 代指扩散模型，拥有比 GAN 更优的能力并快速崛起。相关研究最早可以追溯到 2015 年，奠基研究是 2020 年《Denoising Diffusion Probabilistic Models》。2022 年，借助 AI 绘画应用，扩散模型在图像生成领域展现卓越实力。

扩散模型的工作原理，是通过连续添加高斯噪声来破坏训练数据，然后通过反转这个噪声过程，来学习恢复数据。一幅画当中，衣服的纹样、树叶、云彩等带有很多细节纹理的地方，其实细节越多，越接近一个随机的噪点。对于这些细节，也许只需要几次高斯噪声的掺入（可理解为高斯模糊），就能破坏原来的纹样，接近正态分布。训练后，可以使用扩散模型将随机采样的噪声传入模型中，通过学习去噪过程来生成数据。都是给定噪声  $x_T$  生成图片  $x_0$ ，相比 GAN，Diffusion 所需数据更少，生成效果更优。

图16: 扩散模型工作原理图示

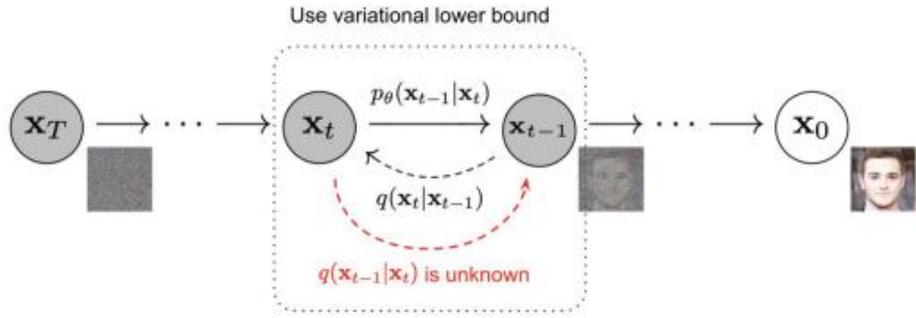
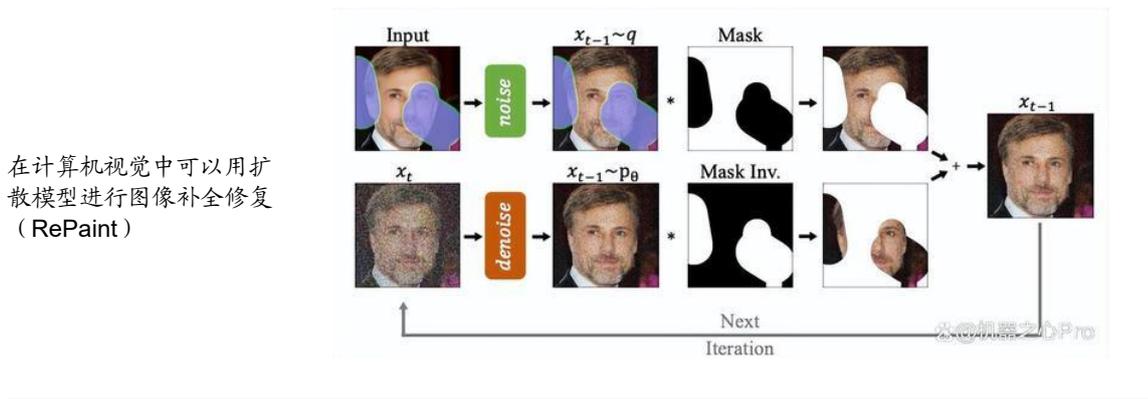


Fig. 2. The Markov chain of forward (reverse) diffusion process of generating a sample by slowly adding (removing) noise. (Image source: Ho et al. 2020 with a few additional annotations)

资料来源: Google Scholar, 浙商证券研究所

扩散模型在计算机视觉、自然语言处理、波形信号处理、多模态学习、分子图生成、时间序列以及对抗学习等七大应用方向中都有应用。

表2: 扩散模型在多领域应用实例

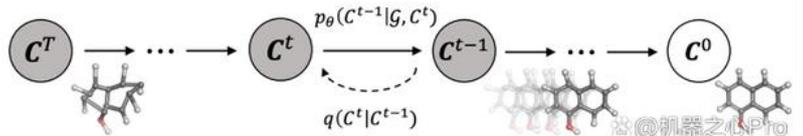


在计算机视觉中可以用扩散模型进行图像补全修复 (RePaint)



在多模态任务中可以用扩散模型进行文本到图像的生成 (GLIDE)

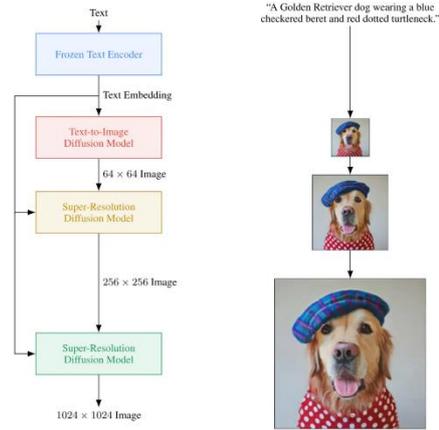
在分子图生成中用扩散模型进行药物分子和蛋白质分子的生成 (GeoDiff)



资料来源: 百家号机器之心 PRO, 浙商证券研究所

在 AI 绘画领域, 除 Disco Diffusion, 最先进的文本生成图像系统 OpenAI 的 DALL·E 2 和 Google 的 Imagen, 都是基于扩散模型来完成的。

图17: Imagen 架构图 (纯语言模型 T5-XXL+Diffusion)



资料来源: Google Scholar, CSDN, 浙商证券研究所

扩散模型还在发展中, 改进研究在采样速度提升、最大似然增强和数据泛化增强等领域持续进步。

### 2.2.4 神经辐射场 NeRF — 顺应 3D 内容消费趋势

NeRF (neural implicit representation, 神经隐式表示) 利用深度学习完成了计算机图形学中的 3D 渲染任务。这一技术从 2019 年开始兴起, 在 2020 年 NeRF 获得 ECCV best paper 之后受到了广大关注。

NeRF 在很大程度上克服了样本特征受限的问题。此前, 2D 到 3D 生成的领域也包含 GAN 方面的尝试, 比如英伟达 20-21 年推出的 GANverse3D 能够自定义对象和交换背景。但由于 GAN 在对抗训练中会受限于样本特征, 该模型当时仅适用于汽车、马匹和鸟类。

图18: GANverse3D 图例



资料来源: GANverse3D, 浙商证券研究所

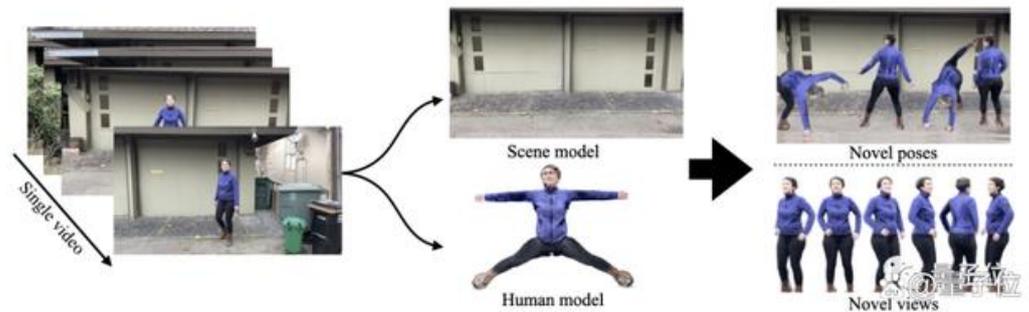
NeRF 模型的基本原理是: 将场景的体积表示优化为向量函数, 该函数由位置和视图方向组成的连续 5D 坐标定义。具体而言, 是沿相机射线采样 5D 坐标来合成图像, 将场景表示参数化为一个完全连接深度网络 (MLP), 该网络将通过 5D 坐标信息, 输出对应的颜色和体积密度值。

NeRF 对于虚拟人创建、3D 训练环境构建、增强现实、线上游戏及电影特效等都具有重要意义。自 NeRF 在 ECCV2020 提出后, NeRF 模型也持续在生成范围、生成效果、乃至所需基础数据上进行改进。例如陆续支持光影变化效果、动态 NeRF, 类实时生成, 全场景 NeRF、单张生成模型、3D 几何数据生成。

在 AI 绘画中，NeRF 通过将场景表示为隐式的神经辐射场，渲染时通过神经网络查询位置上的场景信息生成新视角图像。直观来讲，渲染就是用计算机模拟照相机拍照，它们的结果都是生成一张照片。NeRF 将场景表示为空间中任何点的容积密度和颜色值，有了以 NeRF 形式存在的场景表示后，可以对该场景进行渲染，生成新视角的模拟图片。NeRF 使用经典体积渲染（volume rendering）的原理，求解穿过场景的任何光线的颜色，从而渲染合成新的图像。

在 NeRF 之后，有人提出了 GRAF，引入了 GAN 来实现神经辐射场，并使用 Conditional GAN 实现对渲染内容的可控性。在 GRAF 之后，GIRAFFE 实现了构成。在 NeRF、GRAF 中，一个神经辐射场表示一个场景。而在 GIRAFFE 中，一个神经辐射场只表示一个物体（背景也算一个物体）。这样做可以随意组合不同场景的物体，可以改变同一场景中不同物体间的相对位置，渲染生成更多训练数据中没有的全新图像。

图19：人物 NeRF 模型和场景 NeRF 模型结合新场景



资料来源：NeuMan，量子位，浙商证券研究所

未来 NeRF 发展主要是基于 NeRF 问题的改进。NeRF 的简洁性具有优势，但也因此带来一些问题：

1. 计算量大导致耗时长：NeRF 生成图像时，每个像素都需要近 200 次 MLP 深度模型的前向预测。尽管单次计算规模不大，但完成整幅图像渲染的计算量还是很可观的，NeRF 针对每个场景进行训练的耗时较长。对此，迭代过后的 Depth-supervised NeRF 能够实现更少的视角输入和更快的训练速度。
2. 只针对静态场景：对于无法拓展到动态场景的问题，主要和单目视频做结合，从单目视频中学习场景的隐式表示。Neural Scene Flow Fields 将动态场景建模为外观、几何体和三维场景运动的时变连续函数。该方法只需要一个已知摄像机姿势的单目视频作为输入。
3. 泛化性差：NeRF 无法直接扩展到没有见过的场景，这显然与人们追求泛化性的目标相违背。因此一些文章开始对 NeRF 进行泛化性的改进。GRF 学习 2D 图像中每个像素的局部特征，然后将这些特征投影到 3D 点，从而产生通用和丰富的点表示。与之类似的还有 IBRnet、pixelNeRF 等，比较核心的想法都是卷积与 NeRF 相结合。目前这种泛化都还不够成熟，无法在复杂场景中取得理想效果。
4. 需要大量视角：尽管 NeRF 方法能够实现出色的视角合成效果，但是它需要大量的（数百张）视角来进行训练，这限制了它在现实中的应用。针对视角数量的改进，目前还局限在比较封闭的测试环境下，如合成物体或者单个物体。扩展其在实操中的可用性也是未来的一大方向。

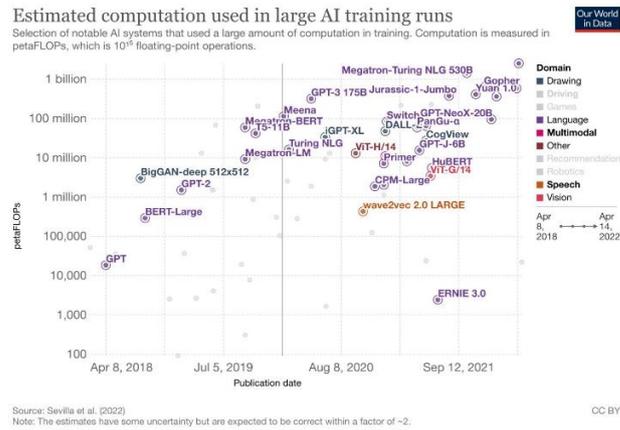
## 2.3 大模型和人工通用智能指引发展方向

我们观察到，深度学习领域有两大前进趋势：大模型和人工通用智能。

### 2.3.1 大模型催生基石模型公司崛起

深度学习领域，模型越大越好。在过去三年里，人工智能模型的规模已经增长了万倍以上。

图20：按参数数量和发布日期划分的深度学习模型



资料来源：Google Scholar，浙商证券研究所

让每家公司都进入大模型建设并不现实，我们认为市场将由少数具有先发和成本优势的供应商主导。OpenAI 作为行业领先者，开发了 GPT 语言模型和 DALL-E 图像生成模型，并不断提升其模型复杂性和规模，OpenAI 亦是大模型供应商的有力选手，而其他公司可以付费购买其更底层的 API 等服务。同时，大模型趋势也将给云计算公司如 Google、亚马逊，和 GPU 厂商如英伟达带来机会。

### 2.3.2 人工通用智能不仅仅是想象

AI 技术被区分为弱人工智能、通用人工智能、超级人工智能三种模式。弱人工智能也被称为狭义人工智能，是专攻某一领域的人工智能，例如在围棋上大放异彩的 AlphaGo 都属于弱人工智能。通用人工智能(Artificial General Intelligence, AGI)也叫强人工智能，或人类级人工智能，通用人工智能指的是一台像人类一样拥有全面智能的计算机，人类能解决的智力问题他都能解决。落到 AIGC 领域，人工智能模型将是多模态的，这意味着相同的模型将被用于文本、图像、视频等等。超级人工智能被定义为“在几乎所有领域，包括科学创造力、一般智慧和社交技能，都比最优秀的人类大脑聪明得多的智力。

通用人工智能处理复杂情况的能力无比诱人，但实现难度极高，AIGC 或为曙光。受困于技术、资源、应用方向等因素的局限，通用人工智能的发展在短期内较难突破。AIGC 的兴起，一方面可以给到一个相对特定的、具象范围的 AGI 应用空间，降低难度，再举一反三；另一方面，AIGC 提供了 AI 广泛施为的机会，无论是数据的大量生产，还是众多专业力量和资本的投入，都有利于推动技术向前演进。

## 3 商业化前景广阔，B 端和三维化或为突破口

### 3.1 应用迅速丰富，用户接受度较高

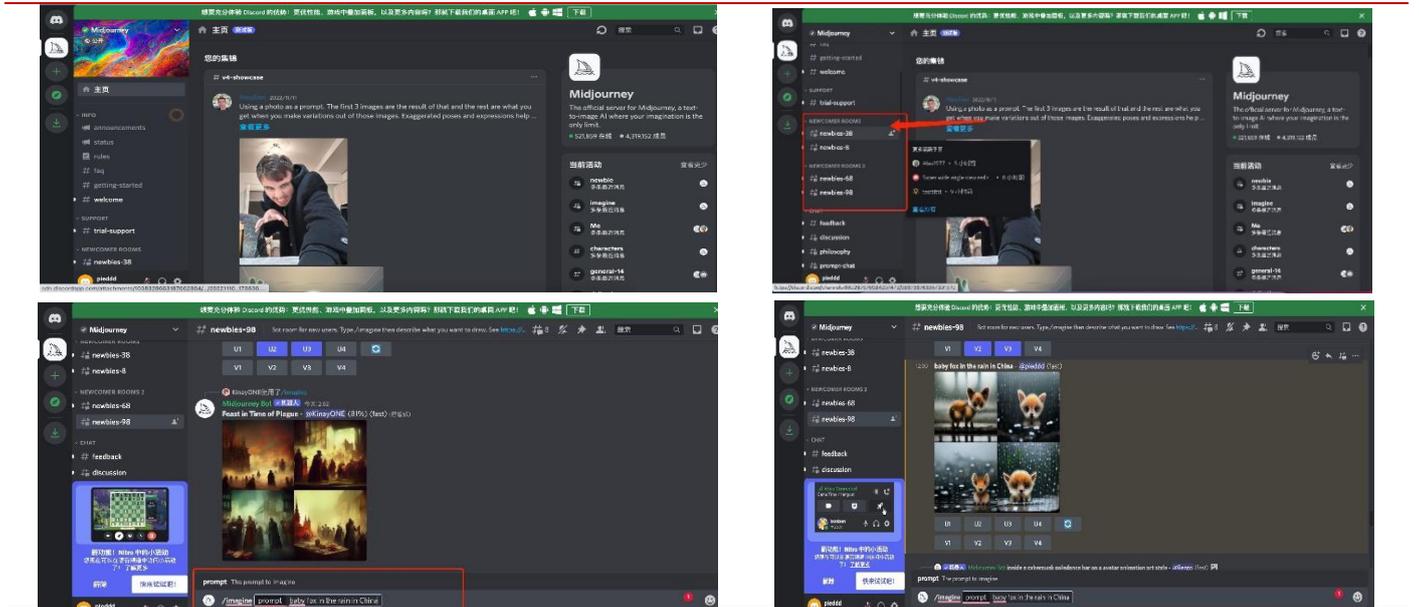
**AI 绘画产品不断丰富，体验持续提升。**近两年，海外流行借助 Disco Diffusion、MidJourney 等 AI 绘画软件来进行艺术创作，Stable Diffusion 各渠道累计日活用户超过 1000 万，面向消费者的 DreamStudio 则已获得了超过 150 万用户。在国内，2022 年是 AI 绘画产品井喷之年，诸如文心一格、TIAMAT 等产品均于今年上线，他们接受中文描述语输入、更能理解中国文化审美和用户需求，并主动利用小红书、微博等平台拓展影响力。

表3: AI 绘画应用一览

名称	Disco Diffusion	Stable Diffusion	Midjourney	DALL·E2	Novel AI	文心·一格	Parti	Make A Scene
发布时间	2021/10	2022/8	2022/8	2022/4	2022/7	2022/8	2022/7	2022/7
开发团队	艺术家 somnai +Google Colab	英国初创公司 StabilityAI	美国初创公司 Midjourney Lab	美国 AI 公司 OpenAI	美国初创公司 Noval	百度	Google	Facebook
特点	效果多元	出图速度全快 调参更多 细节丰富	出图快 上手容易 场景包容度高	画作逼真 可抠图	擅长二次元	操作简单 可理解中文 无设备要求	支持 200 亿参数, 细节逼真	草图+文字描述生成输出
局限	生成速度较慢 上手难度较高 有 GPU 使用限制	缺少东方素材 图像分辨率低	精细度上欠缺 一定申请门槛 图像分辨率高	需要内测资格 图像分辨率低	出图效果不够 稳定	图像分辨率低 出图速度快	无法处理否 定描述 物体缩放比例不稳定	可控性较差, 割裂感较重

资料来源: 各项目主页, 浙商证券研究所

图21: AI 作画操作流程示意图



资料来源: Stable Diffusion, 浙商证券研究所

AI 作画操作进过多次简化, 已经大大降低了使用门槛, 可支持文字成图、图像转化、使用文字修饰图像等。上图以 Stable Diffusion 为例, 展示了较为通用的 AI 作画流程:

- 1) 注册 discord 账号后登陆进 Midjourney 主页。
- 2) 点击进入随意一个新bies 新手社区。
- 3) 向机器人输入 “/image”命令, 在 prompt 后输入文字指令, 完成后按 enter。
- 4) 等待一分钟即可得到 AI 作画成品。

用户认知层面，已经有相当一部分用户认可 AI 绘画对人类工作的助益。根据 6pen 的调研，有 50% 以上的用户认为 AI 绘画能替代一部分，甚至完全颠覆目前的工作方式。随着更多人接受并参与进来，AI 绘画爱好者甚至已经开始形成自有生态，反哺 AI 绘画的发展。他们组建了相关社群交流技术，将作画过程戏称为“魔法吟唱”。用户在社群中分享输入参数和输出结果。

图22: AI 绘画社区元素法典



资料来源：元素法典，浙商证券研究所

### 3.2 变现仍处于尝试阶段，B 端或为切入点

AI 绘画产品目前少有营收或实现盈利。根据南方财经，视觉中国官方披露，公司拥有 AIGC 技术储备和素材资源，曾在元视觉艺术网发行过相关作品，相关作品确能产生营收，但占比极低，2022 年上半年，元视觉艺术网一共创收约 1500 万元。而 Stable Diffusion 和 Midjourney 这两家公司都还未实现盈利。究其原因，或有以下因素：

AI 绘画商业化方面仍处于摸索阶段，变现方式较为单一。用户多为生成数量或者使用时间付费，常见付费方式为订阅制或按次付费，以几大主流 AI 作画软件的商业模式为例：

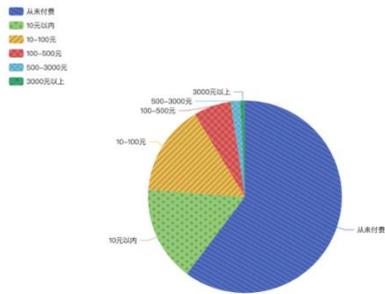
- Stable Diffusion 目前尚未形成明确的盈利模式，目前的收费方式是首次注册 DreamStudio beta 将获得价值 2 英镑的积分，大约相当于 200 次单张图免费生成的额度。试用后，可以按 10 英镑的增量购买额外的积分。但 Stability AI 的 CEO 称其未来商业模式类似红帽和 MongoDB，开源版本免费，通过商业版本盈利；
- Midjourney 采用了订阅制，新用户可免费生成 25 张，之后对于个人用户或公司年收入少于 100 万美元的企业员工用户，有两个档位的订阅套餐，分别是：1) 基本计划每月花费 10 美元，200 分钟 GPU 时间（每次生成大约 5 美分）；2) 标准计划每月花费 30 美元，15 小时 GPU 时间（每次生成大约 3 美分）。而对于大公司客户，单人一年收费约为 600 美元，生成的作品可以商用。

对于普通 C 端用户，应用场景商业化性价比较低，付费意愿有待提升。根据 6pen 的调研，60% 的用户从未在 AI 绘画产品上有过付费行为，剩下 40% 的用户中，付费超过 100 元占比仅 10%。我们认为主要原因是：

1. 商业应用场景缺失：普通大众使用 AI 绘画进行创作后，若自用（如用做头像）或者分享在社交媒体，免费软件足够尝鲜。而若用于约稿等用途，产出的作品受到素材和技术的限制，为满足客户需求，仍需大量加工以得到成品，性价比较低。《太空歌剧院》这副作品在 AI 生成之后，设计师还进行了上千次的修改，花费了近 80 个小时才完成。虽有人在抖音上做壁纸号，在闲鱼上卖描述语，或将作品卖给包装厂，但这些还未达到产业的高度，传导链条也较长。

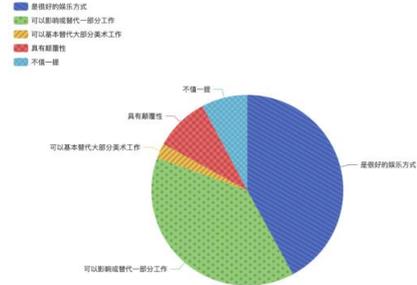
- 当前 AI 绘画平台大多为轻量级的工具应用，能够操作的玩法和赋能服务都比较有限。已有的赋能收费项目集中在付费提速或者增加清晰度方面，或类似 PromptBase 公司创立了一个 DALL-E 2、GPT-3 提示词在线交易平台，允许用户以 1.99 美元的价格买卖提示词，此几种主要在变相弥补现有产品在成像速度、质量和精确性上的局限。

图23: 中国用户为 AI 产品或服务付费比例



资料来源: 6pen, 浙商证券研究所

图24: 中国用户对 AI 绘画的认知评价



资料来源: 6pen, 浙商证券研究所

基于上述分析，我们认为，B 端变现路径更为多元、成熟，如若 B 端能挖掘出较为可行的商业模式，付费的可能性和水平相对更高。B 端可能会在以下方向产生收入：

表4: AI 绘画应用行业示例

行业	应用方式	实例
广告行业	借助 AI 制作宣传素材	《时尚芭莎》36 周年纪念 AI 海报
设计行业	AI 辅助包装设计、服装设计出图	阿里助力犀牛智造推出 AI 设计平台
营销定制行业	借助 AI 生成营销素材	百家号自动生成视频功能
漫画/动漫行业	配文自动生成漫画、插画、视频	AI 绘制《漠河舞厅》MV
游戏行业	AI 辅助角色、道具、场景设计	日本游戏开发者使用 Midjourney 生成素材开发横版射击游戏

资料来源: 钛媒体, 百家号, 时尚芭莎, 新智元, 腾讯网, 浙商证券研究所

### 3.3 三维化打开游戏、影视、VR 等应用空间

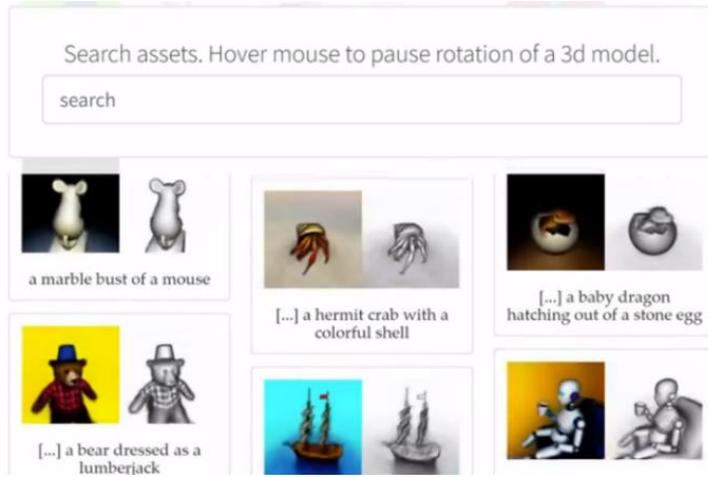
**三维化是 AIGC 视觉发展的必然趋势。**从长远趋势来看，人们对于未来元宇宙的期待是三维化、AI 化及开放式的，AIGC + 3D 内容是可见途径。中短期维度上，AIGC+3D 是丰富游戏、影视、VR 等数字内容，降低其制作成本的有力工具。

**3D 内容生产借助 AI 绘画产品快速普及。**我们看到，AIGC 进入 3D 内容领域，有效降低了参与门槛，让全民参与到 3D 内容创作和消费当中，在 3D 领域升起 UGC 的浪潮，我们认为这将大为丰富 3D 内容创作的有生力量，正如视频拍摄和剪辑工具平民化推动视频内容行业的繁荣，3D 内容创作行业也将因此迎来全新发展契机。AI 绘画产品三维化方面，代表性的产品有 Dream Fusion、Stable Diffusion、GET3D 等，并已有多种场景应用实例，覆盖多种应用场景的可能性：

Dream Fusion 是 Google 的大型 AI 图像模型 Imagen 与 NeRF 的 3D 功能相结合。Dream Fusion 训练后的模型可以在任意角度、任意光照条件、任意三维环境中基于给定的文本提示生成模型，整个过程既不需要 3D 训练数据，也无需修改图像扩散模型，完全依赖预训练扩散模型作为先验。但使用该项目的 GitHub 页面只允许从一系列预设文本提示中进行选择生成 3D 模型，暂不允许用户输入自己的文本描述，数字资产本身的分辨率也较低。

Dream Fusion 的在线画廊展示了一系列 .gltf 格式的模型，适合在 AR 项目中使用，或者作为可以手动细化以用于更高细节工作的基础网格。

图25: Dream Fusion 样例



资料来源: GitHub, 浙商证券研究所

Stable Diffusion 原本是 2D 美术生成工具，通过和 Touch Designer 可视化编程工具结合可创建 VR 场景，并且用机器学习算法为这些场景生成了文本查询对象。目前，已经成功地在虚拟现实创建了成熟的场景，这些场景以 60 fps 的速度实时渲染。

已有诸多厂商在研究将 AI 创作的 3D 场景用于游戏生产，如依赖程序生成的 Roguelike 游戏，开发者未来或许可利用 AI 制作的场景直接作为游戏关卡，大量节省游戏制作成本。

图26: Stable Diffusion 样例



资料来源: ommer-lab, 浙商证券研究所

GET3D 是英伟达推出的模型，通过 2D 图像训练后，该模型可生成具有高保真纹理和复杂几何细节的 3D 形状。它生成的是显式纹理 3D 网格，也就是说，它创建的形状是三角形网格的形式，就像纸模型一样，上面覆盖着纹理材质。因此 GET3D 不仅可以生成多种多样、高质量的模型，还可以将生成模型导入到游戏引擎、3D 建模器和电影渲染器中，对它们进行编辑，并且将 GET3D 生成的模型导出到图形应用程序后可以在模型所在的场景中移动或旋转时应用逼真的照明效果。

基于 GET3D 已建成一个用照片自动生成三维模型的平台。在实际应用过程中，文物研究人员借助摄影测量的方法实现文物三维数字化，工作人员只需用相机或手机，按照建模拍照的教程采集照片，再登录网站上传照片，云服务能够自动计算生成高精度三维模型。现产品被用于石窟寺调查，平台上已有不少石窟寺模型。

图27: Get3d 平台上的石窟模型

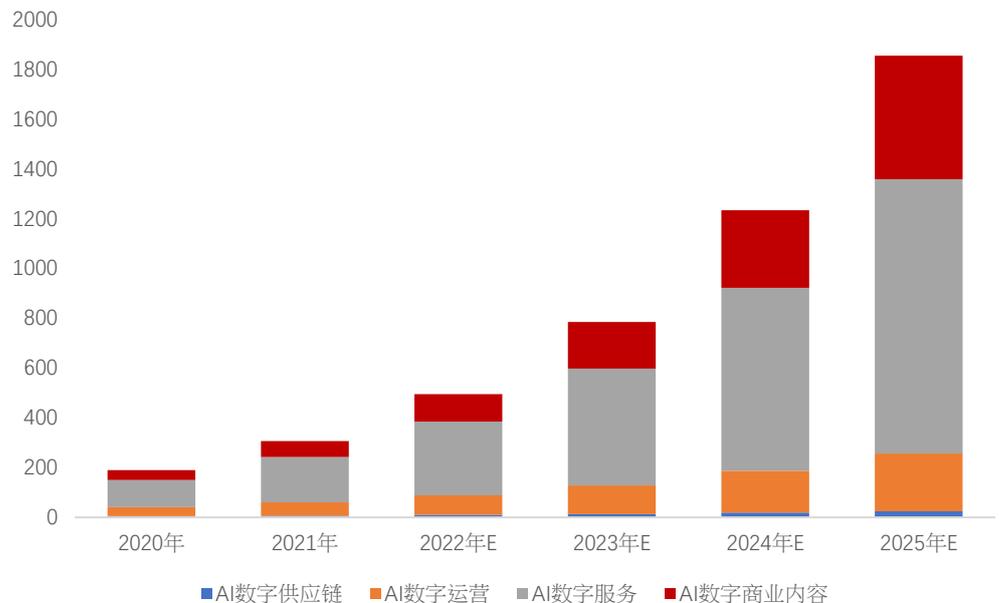


资料来源: get3d, 浙商证券研究所

### 3.4 产业链初见规模，部分空缺现蓝海

中关村大数据产业联盟发布的《中国 AI 数字商业展望 2021-2025》报告披露，至 2025 年，中国 AI 数字商业核心支柱产业链规模将达到 1853 亿元，未来五年复合增长率约 57.7%。其中，AI 数字商业内容产业规模将达到 495 亿元，AIGC 和 AI 绘画当属此类。

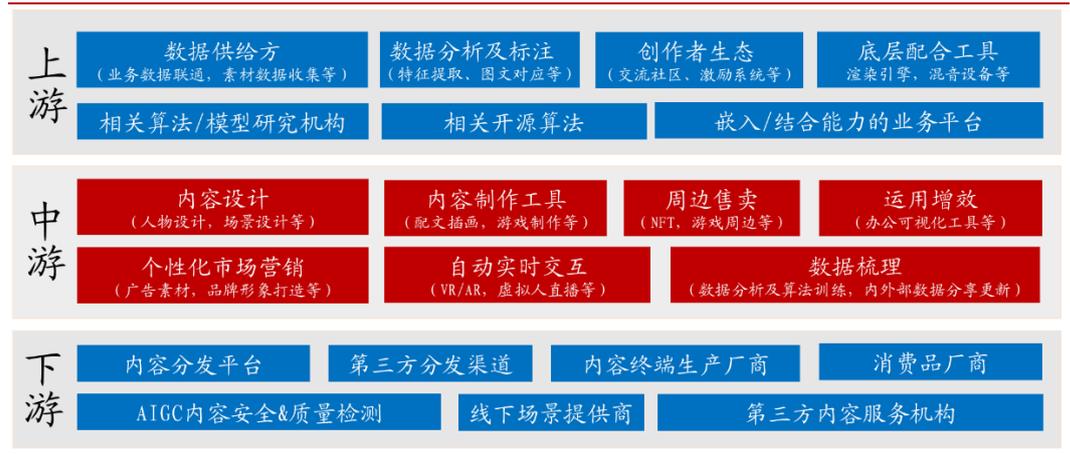
图28: 2020-2025 年 AI 数字商业规模测算 (亿元)



资料来源: 前瞻产业研究院, 中关村大数据产业联盟, 浙商证券研究所

就 AIGC 和绘画而言，其产业链涉及到硬件、NLP、算法算力、应用、数据提供与处理等多环节，当前产业布局于算法和应用开发环节较为集中和领先，而在产业链上下游还有诸多可开发的蓝海领域。

图29: AIGC 产业链



资料来源: 量子位, 浙商证券研究所

和国外类似, 参与主力分为了两类: 1) 相关行业内已有一定规模的公司, 多为龙头; 2) 初创公司。前者代表为百度, 8月发布了 AI 艺术和创意辅助平台文心一格, 它使用了百度自主研发的产业级知识增强大模型—文心大模型。另一 AI 绘画大热产品 TIAMAT 背后则为初创团队, 其成立于 2021 年, 一年内就积累了可观的社区用户和商业客户, 已获得 DCM 数百万美元天使轮融资。参见互联网发展的历史, 未来规模较大的公司很可能通过自研或者收购融合多种技术, 提供更优的体验, 打通不同场景, 连接创作者、消费者、供应商等各方。我们认为, 产业集中度将不断提升, AIGC 也将集成进入元宇宙的复杂系统。

## 4 基建待完善, 发展中风险与希望并存

我国 AIGC 行业仍处于起步阶段, 面临来自版权、伦理等各方面的困扰因素。

一方面, AI 绘画的能力之强引起了行业从业者的担忧乃至恐慌。AI 通过庞大的数据量, 无休止的深度学习飞速进步, 短期内就掌握了许多新人画手要练数年的人体、透视与光影技术, 已经有能力威胁到了底层画师的生存问题。

另一方面, 批判者认为 AI 创作没有任何情绪和灵魂, 难以和人类的艺术创作相提并论。且 AI 作画仍需学习人类创作的素材, 且创作者反复修改文本内容和调整成图才能得出相对满意的结果, 这意味着 AI 作画仍然依赖人类的智慧和劳动, 创作的主动性仍掌握在人类手上, 即 AI 作画实际是个伪命题。

此外, AI 绘画涉及的作品侵权、名人肖像侵权等都是巨大的行业风险点。日本绘画领域就曾掀起争论, 不少画师公开表示禁止 AI 学习自己的作品。根据 6pen 对原创艺术家的问卷调研, 超过 90% 的原创艺术家持相对保留态度, 约 37% 的原创艺术家呼吁向版权付费或有所标注。这也阻碍了 AI 绘画创收。AI 模型训练所使用的大量素材, 可能包含了未经授权的, 有明确版权方的图片数据, 版权归属是否应该由模型指定也未有定论。因此使用此类作品并以此盈利很有可能为使用者带来法律上的纠纷。

但 AI 绘画也同样带来了全新的希望。通过采用最新的 AI 技术来分析已故漫画大师手塚治虫生前的漫画作品, AI 在 2020 年执笔画出了“手塚治虫新作漫画”。

我们相信, 随着相关法律法规约束的健全, AIGC 行业必能在内容创作领域找到合适、合规的发展之路。

## 5 建议关注

表5: AI 绘画概念股一览

	公司代码	公司简称	市值	AIGC
下游	000681	视觉中国	109.85 亿元	高质量图像数据集, 发售数轮 AI 生成数字藏品
	002624	完美世界	273.34 亿元	已在游戏研发过程中使用 AI 绘图
	300624	万兴科技	47.59 亿元	AI 绘图软件“万兴 AI 绘画”开启公测
	300364	中文在线	77.55 亿元	AI 绘画和 AI 文字辅助创作功能 (向 17K 文学平台作者开放)
	00700	腾讯控股	2.76 万亿港币	开发“画说梦想”AI 绘画公益小程序, 写稿机器人“梦幻写手”
	S09999	网易	3523.72 亿港币	开发一站式 AI 音乐创作平台“网易天音”
	301052	果麦文化	20.47 亿元	果麦 AI 创作机器人通过采集互联网文章转化为自己的内容
	300792	壹网壹创	71.74 亿元	AI 设计智能化应用, AI 营销场景
	002354	天娱数科	72.97 亿元	推出了 SaaS 平台可用于元宇宙中的数字身份塑造等需求
	300058	蓝色光标	142.49 亿元	“创意画廊”平台支持用户生成抽象画作并用于虚拟人创作
	605168	三人行	88.92 亿元	AIGC 辅助营销素材设计和相关应用场景
	300418	昆仑万维	178.43 亿元	成立 Star X MusicXLab, 目前已经正式销售 AI 作曲
	300130	新国都	70.60 亿元	AI 深度内容生成、视频处理技术、交互式虚拟人
	SW09888	百度集团	2591.63 亿港币	文心大模型+文心一格
	300229	拓尔思	96.71 亿元	A 股语义识别龙头, 开展“大数据看世界杯”的虚拟数字人服务
	W00020	商汤	673.36 亿港币	国内领先的人工智能龙头, 深耕 CV 领域
	002230	科大讯飞	810 亿元	国内语音识别和人工智能龙头
	002657	中科金财	63.16 亿元	数字人内容制作引擎、元宇宙场景开发工具
	MVRS	Meta	2971.06 亿美元	NLP 技术领先, AIGC 应用 Meta AI
	GOOG	谷歌	1.27 万亿美元	深度学习模型 Transformer, AIGC 应用应用 Imagen
	MSFT	微软	1.80 万亿美元	NLP 技术领先, 投资 OpenAI
	000977	浪潮信息	343.83 亿元	公司发布了业内首款元宇宙服务器 MetaEngine
	NVDA	英伟达	3836.84 亿美元	视觉单元 GPU 市占绝对领先, 生成对抗网络技术领先
	688008	澜起科技	779.73 亿元	云计算和人工智能芯片供应商
	300394	天孚通信	108.36 亿元	光模块及数通厂商光学器件供应商
	300308	中际旭创	239.57 亿元	光电子领军企业
300502	新易盛	139.30 亿元	光收发器件解决方案供应商	
上游	300738	奥飞数据	66.20 亿元	IDC 服务提供商

资料来源: Wind, 浙商证券研究所

## 6 风险提示

- 1) AI 绘画相关技术发展不及预期风险；
- 2) 商业化拓展进度和效果低于预期风险；
- 3) AI 绘画涉及的作品侵权、名人肖像侵权等风险。

## 股票投资评级说明

以报告日后的6个月内，证券相对于沪深300指数的涨跌幅为标准，定义如下：

1. 买入：相对于沪深300指数表现+20%以上；
2. 增持：相对于沪深300指数表现+10%~+20%；
3. 中性：相对于沪深300指数表现-10%~+10%之间波动；
4. 减持：相对于沪深300指数表现-10%以下。

## 行业的投资评级：

以报告日后的6个月内，行业指数相对于沪深300指数的涨跌幅为标准，定义如下：

1. 看好：行业指数相对于沪深300指数表现+10%以上；
2. 中性：行业指数相对于沪深300指数表现-10%~+10%以上；
3. 看淡：行业指数相对于沪深300指数表现-10%以下。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重。

建议：投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者不应仅仅依靠投资评级来推断结论。

## 法律声明及风险提示

本报告由浙商证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格，经营许可证编号为：Z39833000）制作。本报告中的信息均来源于我们认为可靠的已公开资料，但浙商证券股份有限公司及其关联机构（以下统称“本公司”）对这些信息的真实性、准确性及完整性不作任何保证，也不保证所包含的信息和建议不发生任何变更。本公司没有将变更的信息和建议向报告所有接收者进行更新的义务。

本报告仅供本公司的客户作参考之用。本公司不会因接收人收到本报告而视其为本公司的当然客户。

本报告仅反映报告作者的出具日的观点和判断，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投资建议，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本公司的交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理公司、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告版权均归本公司所有，未经本公司事先书面授权，任何机构或个人不得以任何形式复制、发布、传播本报告的全部或部分内容。经授权刊载、转发本报告或者摘要的，应当注明本报告发布人和发布日期，并提示使用本报告的风险。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

## 浙商证券研究所

上海总部地址：杨高南路729号陆家嘴世纪金融广场1号楼25层

北京地址：北京市东城区朝阳门北大街8号富华大厦E座4层

深圳地址：广东省深圳市福田区广电金融中心33层

上海总部邮政编码：200127

上海总部电话：(8621) 80108518

上海总部传真：(8621) 80106010

浙商证券研究所：<https://www.stocke.com.cn>