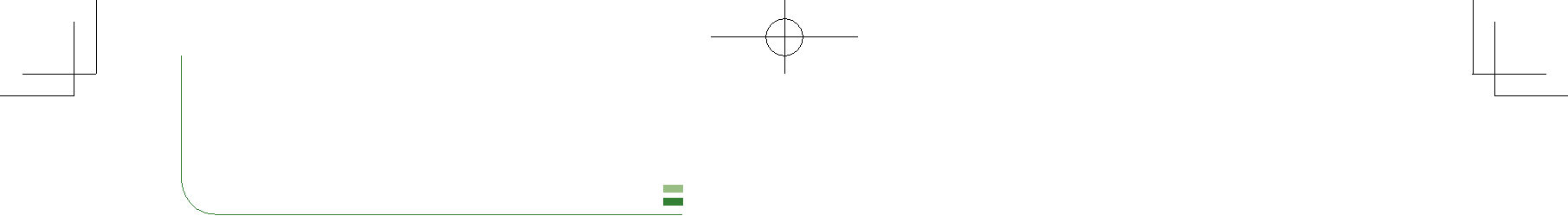
 第1章 认识AI作图

CHAPTER ONE

第 1 章

认识 AI 作图

动手学AI绘画：Midjourney从入门到实践

绘画，作为人类独创的艺术形式，其起源最早可追溯到远古时代。人类的绘画经历了漫长的发展和变迁，但从未改变的是，它体现了人类特殊的创造性和思想，这也是人类和其他生物，以及机器之间的一个关键性的区别。然而人工智能（Artificial Intelligence，AI）的出现似乎将要打破这一固有认知——计算机首次实现独立完成画

作，并在效率、复杂性、创意等方面都接近甚至突破人类水准，这让我们感到惊喜甚至恐惧。AI 作图的时代要来了吗？人类画家要被 AI 淘汰了吗？为了便于读者对上述问题构建自己的基本认知，本章将从发展背景、实用工具、商业化方式、落地实践等方面进行介绍，力争简要且全面地展现 AI 作图的全貌。



1.1 揭秘AI作图



1.1.1 AI作图爆发的前夕



毫无疑问，当前 AI 作图技术的突破是革命性的，不仅针对普罗大众，对于专业人员来说也是如此。AI 作图技术给我们带来的惊艳感，远非曾经所能比。恰如走到时代的奇点，彻底迎来了爆发。请随本书回首 AI 作图爆发的前夕，感受它发展的脉络。

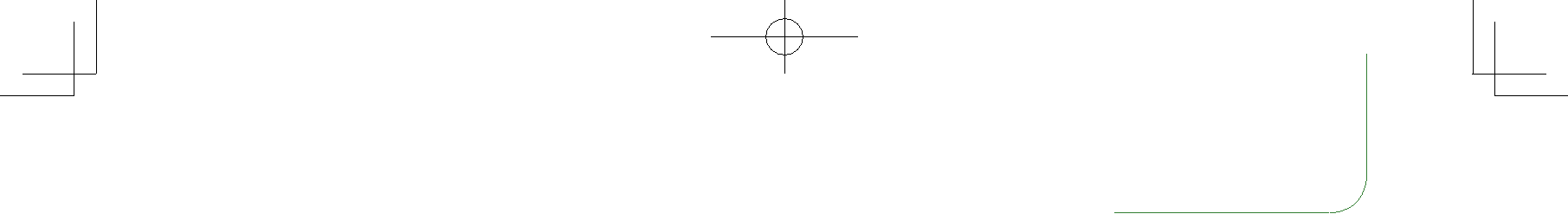
最初的 AI 作图还没有这么复杂。如果我们以计算机的视角来理解它，那么“绘画”就是以特定的规则或目标去填充一幅空白画面，并存储为一个 jpg 或其他格式的图像文件。通常意义上，图像文件本质上可以理解为 RGB 像素值组成的数字阵列，如

* 1-1 所示。RGB 数值阵列由三层组成，分别代表红（Red）、绿（Green）、蓝（Blue）三原色，如图 1-1 所示，展示了图片一小块的其中一层数字阵列。因此 AI 作图又可以被解释为对画面上一个个细小像素点的 RGB 值的计算。事实上，这正是 AI 作图，或者说计算机图像处理的基本原理。

遵循这样的模式，我们很容易想到一个最简单的 AI 作图——将全白的画面涂黑，只需要将所有 RGB 像素值从（255,255,255）改成（0,0,0）！当然，这并不是后来令人惊艳的画作，也不是本书主要讨论的对象。但我们应该理解，计算机创造绘画的方式可以有很多种，比如基于规则去修改 RGB 数值的图像生成（如改变图片的亮度、锐度等）、模拟人类绘画的笔触（如 Photoshop 软件提供的“笔触”效果）和其他绘画效果的算法，以及基于 AI 深度学习技术的算法等。尽管这些算法的技术难度并不等价于 AI 作图的效果，但更复杂的算法意味着可以创作更多样、更复杂的画作。因此，随着计算机算法和硬

002



 第1章 认识AI作图

件技术的迭代，AI 作图的水准会不断提高。计算机的发展是 AI 作图从早期雏形发展到当下惊人效果的根本动力。

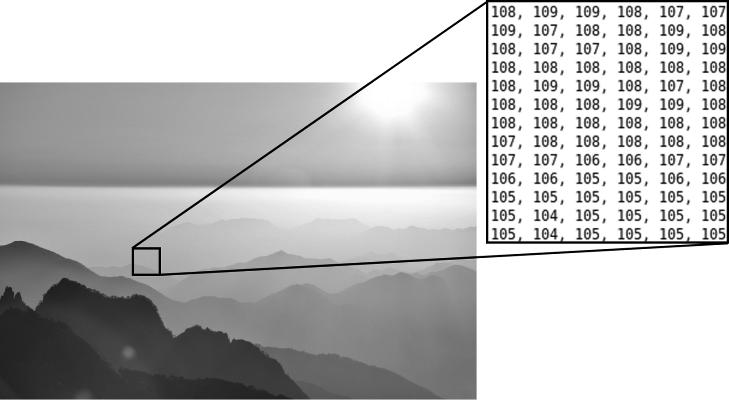


图1-1 图像中的RGB数值阵列示意图

一个公认的比较早期的 AI 作图程序是 AARON，由英国的哈罗德·科恩（Harold Cohen）开发。这位著名的抽象画家在 20 世纪 60 年代末开始开发，随后在 1973 年前往



斯坦福大学人工智能实验室工作，并在此期间完成了 AARON 的开发。该系统先创作简单的黑白图画线稿，再进行上色以完成图画。虽然 AARON 生成的画作已经具备一些人类绘画的特征，但其本质上还是基于规则的绘画，是在特定的、预先构建好的知识框架下生成的，并且它所创作的画作质量距离今的 AI 作图还有很大差距。

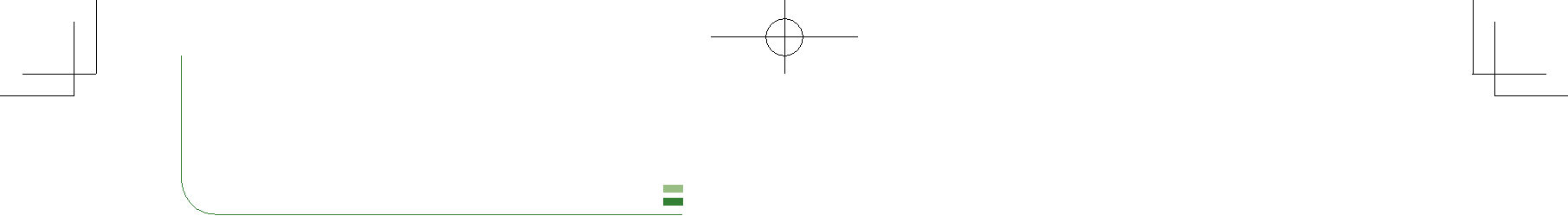
此后多年，AI 作图领域一直没有太多突破，直到 2012 年神经网络模型 AlexNet 横空出世。AlexNet 是辛顿（Hinton）和他的学生亚历克斯·克里泽夫斯基（Alex Krizhevsky）设计的，取得了 2012 年 ImageNet 图像分类竞赛冠军，这是神经网络模型

第一次突破人工设计算法的性能。两年后，第一批真正意义上的 AI 作图模型，对抗神经网络（Generative Adversarial Network，GAN）被发明。GAN 使用“生成器”来创建新图像，并使用“鉴别器”来决定哪些创建的图像被认为是成功的，通过类似“左右手博弈”的策略不断迭代学习如何创建优质的图像。至此，人类这才算是打开了 AI 作图的大门，并依托对深度学习技术的研究不断深入摸索。

基于深度学习的 AI 作图算法最早主要集中在底层视觉领域，例如超分辨率、图像恢复、去雨、去雾、去模糊、去噪、重建、风格化等。这些任务主要是针对 pixel-to-pixel （像素级）的图像编辑，在视觉上取得惊艳的效果，容易引用到以手机为代表的消费级电子产品，一度吸引了人们的目光，此类技术也受到了相关厂商的热捧。后来随着计算机

003



动手学AI绘画：Midjourney从入门到实践

视觉、多模态神经网络等技术的发展，又在视觉语义理解方面不断开拓发展。这些底层视觉和高层视觉共同孕育了更加复杂的图像生成技术。如图 1-2 所示，4 张图像展示了 AI 生成的不同风格。在 2021 年，“VQGAN（视觉生成模型）+CLIP（一种匹配图像和文本的预训练神经网络模型）”的结合使用让 AI 作图踏入新阶段，真正实现了较高质量的“输入文字生成图像”，打破了两个内容模态之间的壁垒。





（a） （b）



（c） （d）

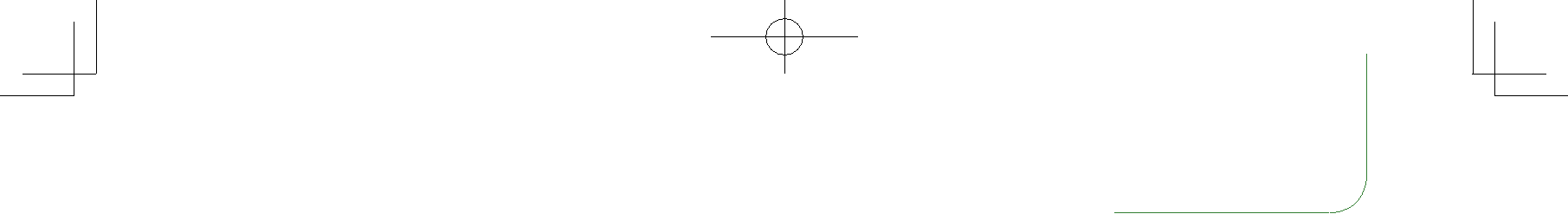
图1-2 AI生成的不同风格的图像

尽管 GAN 覆盖了大部分图像生成的领域，但科学的进步似乎不甘其独占鳌头。 2019 年，加州大学伯克利分校的研究人员发表了一篇名为 *Denoising Diffusion* *Probabilistic Models*（DDPM，去噪扩散概率模型）的论文，模型出色的性能迅速掀起了一波研究扩散模型的热潮。2021 年，德国慕尼黑工业大学发表了著名的 Latent Diffusion Model（LDM），这就是之后风靡全球的 Stable Diffusion 的前身。至此，AI

作图的两大利器——对抗神经网络和扩散模型，都已经准备就绪，而这些年中，深度学

004



 第1章 认识AI作图

习底层算法、计算框架、硬件设备也在不断迭代，为最后的临门一脚做足了准备。AI 作图爆发的前夜似乎就在此刻，然而飞速发展的科技根本不会停留片刻。

1.1.2 里程碑式的2022年

时不我待，只争朝夕！ LDM 的论文于 2021 年上线，2022 年正式发表于计算机视觉顶级学术会议 CVPR 2022，但更多的论文和基于 LDM 的产品已呈争先恐后之势！ 2021 年，DALL·E 模型震撼上线。它是知名 AI 公司 OpenAI 基于大语言模型 GPT-3（其后续产品 ChatGPT 是 2022 年自然语言处理的现象级产品）开发的一种新型神经网络，可以从文字说明直接生成图像，突破了传统图像模型和语义模型的边界。2022 年 2 月，AI作图的“文本到图像”转换产品 Disco Diffusion 正式上线。两个月后，OpenAI 宣布开放同类产品 DALL·E 2 ；当月 Midjourney 宣布开始产品内测。2022 年 7 月，Stable



Diffusion 横空出世，由于其开源特性可以接受全球网友贡献的特定模型（又称 LoRA），使其可以实现的功能更加广泛。

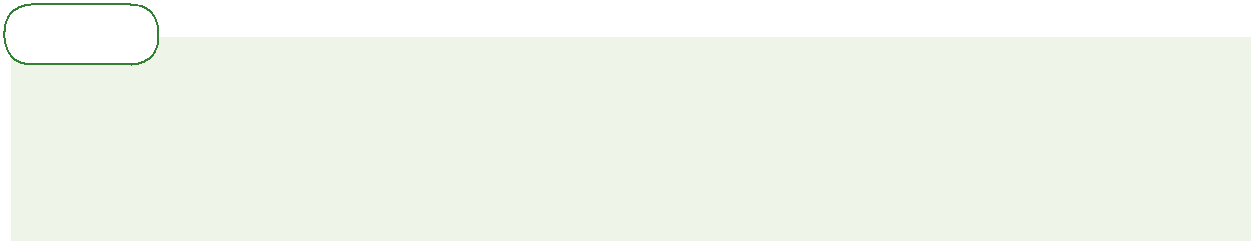


除了技术的进步和新产品发布，媒体在 AI 作图爆发的过程中也起着推波助澜的作用。2022 年 6 月 11 日，《经济学人》杂志选择用 Midjourney 生成的 AI 绘图作为当期杂志的封面，AI 作图由此进入主流媒体视野。2022 年 8 月，一幅由网名为 Sincarnate 的用户使用 Midjourney 生成的画作《太空歌剧院》，在美国科罗拉多州博览会夺得 Fine Arts Exhibition 数字艺术首奖。一位专家得出结论：“人工智能艺术现在已经无处不在，甚至连

专家都不知道这意味着什么。”有人认为 AI 会推动艺术的繁荣，有人则对 AI 冲击传统艺术领域感到担忧，还有人则对 AI 增强人类的艺术创作能力满怀憧憬。总之，AI 艺术创作开始脱离学术界的小圈子，逐渐进入我们的日常生活。它迅速占领各大社交网络榜首，以至于

AI 画作在社交媒体中渗透的速度令人“难以接受”，“一眼 AI”甚至一度成为中文网络热词。

从此，伴随着怀疑、争议和热捧，AI 作图进入寻常百姓家，势不可当地在内容创作领域攻城略地。尽管社交媒体热议、版权争论渐起，但更多人在 AI 强大的创作能力面前心服口服地选择加入。

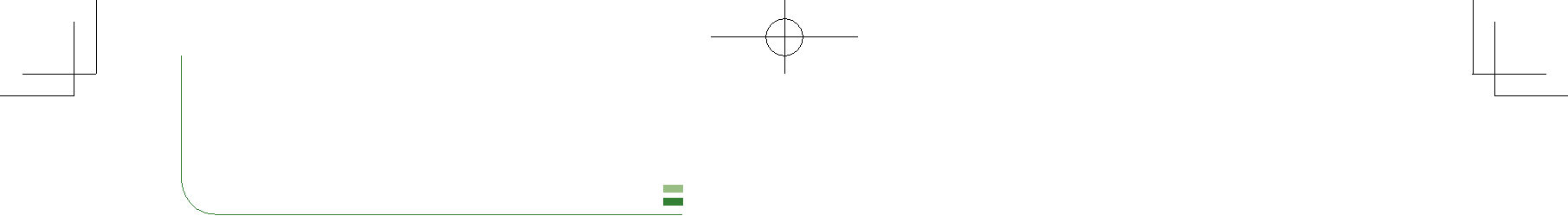


 提示

曾经关于 AI 的一大言论是 AI 无法生产创意，然而在卷进来的各行各业中，设计行业恰恰是最积极的领域之一。不仅仅是纯绘画，包括平面设计、建筑设计、服装设计、鞋类设计、室内设计师、影视动漫等一切与图像创作相关的设计行业都开始进入这个领域。

005



动手学AI绘画：Midjourney从入门到实践

由于前述的多数 AI 作图模型都具有一个共同点，基于自然语言描述（提示词，又称

Prompt）来生成对应图像，因此 AI 作图的发展也能受益于自然语言处理技术，并因此

迎来 2022 年大爆发的最后一股助力。2022 年 11 月，OpenAI 宣布上线 ChatGPT，这个现象级的产品向世人展现了大语言模型（Large Language Model，LLM）的能力。 ChatGPT 不仅登顶各大自然语言处理的性能排行榜，而且具备相当成熟的产品力，使得



普通人只需简单说明就可以上手使用，感受这一世界前沿科技的力量。在 ChatGPT 的加持下，基于提示词的 AI 作图模型得到了更加丰富和准确的文字输入，甚至可以和语言模型进行复杂的耦合，其能力得到巨大提升。

1.1.3 AI作图的成熟与落地

如果说 2022 年是 AI 作图的爆发元年，那么 2023 年就是 AI 作图的落地元年。 Midjourney 等 AI 作图产品逐步上线开放以后，短时间内就吸引了一大批用户的关注，

长期占据国内外互联网热搜。同时，人员和资本迅速进场，AI 作图相关的创业团队如雨后春笋般涌现，迅速催熟这个新生的市场。



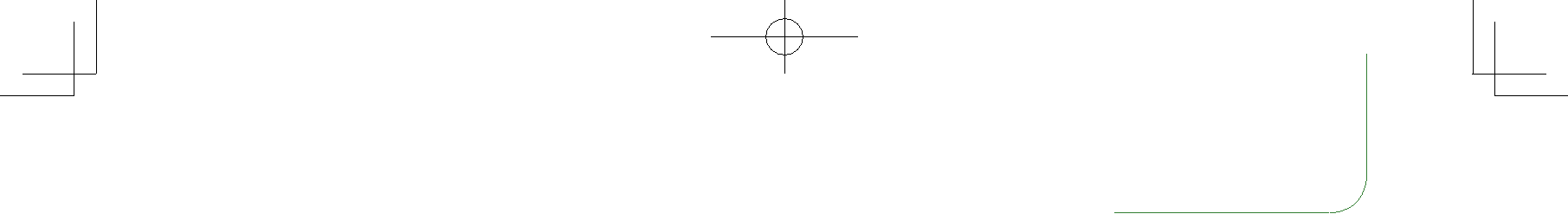
从技术的角度来说，AI 作图经过近十年的发展，在成本、效率、效果等方面都已经取得了长足的进步，在纹理细节、色彩、绘画元素融合等方面都达到了新的高度，接近甚至超越了人类的水平。早期的 AI 作图模型主要是“图像—图像”的变换，用户必须具备一定的图像和艺术基础才能控制模型的生产效，不然就只能依赖开发者或软件公司提供的预设模式，从适用性和效果来说都非常受限。而随着“文本—图像”的跨模态模型的成功，最新的 AI 作图普遍支持用户通过语言描述来引导模型的绘画。这种使用提示词作为输入的工作模式一方面解除了模型受限于原始输入图像内容的约束，实现了真正意义上的“生成图像”；另一方面大大降低了普通人参与的门槛，并显著提高了用户创作效果的上限。这些技术上的成熟使得 AI 作图成为艺术领域有力的竞争者。

从行业发展来看，随着这两年 AI 作图技术的逐步成熟，相关人员和资金也迅速到位，共同支撑行业的发展。当前，AI 作图全产业链的形态已初步形成，包括上游的数据和基础设施供应商，中游的模型和服务供应商，和下游的内容创作和商业化产业。从 2023 年起，AI 作图领域的创业和融资事件不断增多，主要包括以下三大类。

（1）以高校为主导的细分领域创业团队，如清华大学、西湖大学等，其特点是知名教师或实验室的长期科研成果转化，起点就具备一定的基础，发展迅速，但投入规模有限。

006



 第1章 认识AI作图

（2）百度、字节跳动、阿里巴巴等大公司的内部创业，其特点是可以快速聚合各类资源推动项目进展，但可能主要面向内部需求。

（3）高校或公司的技术人员出走独立创业，兼备上述两者的特点，但需要快速找到商业化的途径。

适逢互联网发展的低谷时期，各家公司团队都希望在这一新兴领域占领先机，仅在



2023 年上海举办的世界人工智能大会上就有多家公司携带超过 30 款大模型参展，激烈的

竞争让行业和市场迅速打开了局面。

当前市面上可公开使用的 AI 绘图产品在不断增多，同时主流产品的优势也十分突出。在国际上，Midjourney 和 Stable Diffusion 基本形成了分庭抗礼的“ 两强 ” 局面，前者的特点是用户上手门槛极低，完全基于网页服务，后者则因为开源而具备大量的用户自定义模型，发展势头迅猛。曾经谷歌开发的“Disco Diffusion”和 OpenAI 的“DALL·E”系列等产品也风靡一时，但由于用户体验等相对不足而未能成为主流。在国

内，Midjourney 和 Stable Diffusion 同样占据主流，但同时国内公司产品也在不断改善，不断接近二者的水平，有望夺回市场。



从落地实用来看，不同公司都在探索潜在的途径，但目前尚未找到大规模商业化的方法。受限于模型可控、精细化等方面的不足，当前模型在工业设计方面尚未大规模铺开，如建筑设计、商品设计等领域尚不能保证完全按照用户意图实现出图，但依然可以在创意、草图等阶段显著减少设计师的工作量。不过对于娱乐、艺术等确定性要求较低的领域，AI 作图则进展迅速，图像和视频的编辑软件可以直接嵌入相关插件，协同人类用户进行创作，实现了初步的商业化。美中不足的是 AI 创作的版权依然属于未规定的“灰色地带”，这可能对 AI 作图的落地造成一定的困惑。

综上，当前 AI 作图正处于飞速发展和落地的关键时期。尽管行业尚处于起步的萌芽阶段，但依然展现出技术和市场协同发展的美好愿景。随着应用场景的不断丰富，整个产业链也会变得更健壮，分工更明确，商业模式更清晰。

1.2 AI作图工具大盘点

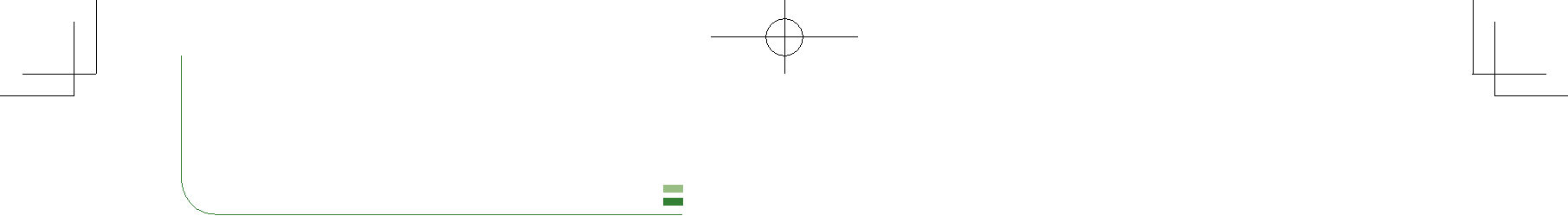


当前国内外最主流的 AI 作图工具使用的大都是 Diffusion 模型的技术，其中最火的

* Midjourney 和 Stable Diffusion。尽管创业产品大量涌现，但多数产品要么产品化程度不足，用户体验欠佳；要么仅基于 Stable Diffusion 等开源产品做简单的封装开发，生

007



动手学AI绘画：Midjourney从入门到实践

图性能较差，因此暂不列举。本文将对部分知名 AI 绘画产品进行介绍，以便读者拥有初步的了解。

1.2.1 Midjourney

Midjourney 是一个 Web 化的 AI 作图产品，确切地说是基于聊天工具 Discord 运行，于 2022 年夏天上线。它旨在为设计师、艺术家和创意工作者提供强大的视觉沟通和展示工具，用户可以基于它进行多样化的创作，包括动态场景、静态场景、角色设计、虚拟装饰等。用户可以通过简单的文本或草图，输入自己的想法和创意，然后由 Midjourney 的 AI 生图功能生成相应的虚拟图像。尽管用户受限于官方提供的接口进行操作，但由于开发团队准备了多个强大的模型和参数设置，且接受自然语言描述作为输入，因此用户依然可以创作出非常多样、高质量的图像。



相比于其他工具，Midjourney 的一大特点是完全基于远程服务运行而不依赖用户本地电脑的硬件，这大大降低了用户的上手门槛。得益于远程服务器的建设，通常在用户开始运行之后，Midjourney 在 60 秒之内就可以返回一组图像，而付费用户的生图速度还可以缩短许多。此外，由于其基于 Web 技术的特点，用户可以在任何设备上随时随地使用该功能，包括电脑、平板电脑和手机等，这一特点在当下的移动互联网时代将取得显著优势。



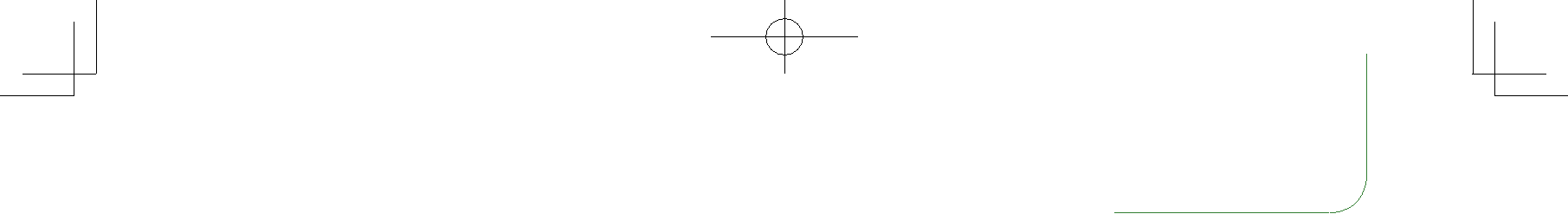
Midjourney 的社群运营也比较出色。借助 Discord 自身的讨论组功能，Midjourney得以快速构建自己的社区，极大地促进了用户的交流和增长。用户可以在社区和画廊中便捷地分享作品和创作方法，对社区整体的提升起到了促进作用。此外，用户还可以自行构建群组进行交流，这使得团队协作成为可能。

不过 Midjourney 存在一个显著的弊端，即目前还无法在保留原图主要特征的情况下生成新图片，而会产生一些相似但不完全相同的新图像。这一点使得 Midjourney 生图的多样性比较好，但在图像编辑、图像修复等特定应用场景下则不太适用。而其他软件，如 Stable Diffusion，则可能没有这点不足。

总体来说，Midjourney 是一款完成度比较高的产品。其依赖 Discord 的构建模式巧妙地规避了软件和社区的大部分构建过程，得以专注于模型和用户体验，因此这一模式也被当前许多新兴的 AI 绘图软件所模仿。不过其限制性也不容忽视，由于受官方提供的模型的限制，用户依然无法构建高度定制化的模型，这对于专有场景落地而言是明显的不足。

008



 第1章 认识AI作图



 提示

更多关于 Midjourney 的信息将在后面的章节进行详细介绍。

1.2.2 Stable Diffusion



不同于 Midjourney，更多的 AI 绘图模型提供本地化部署模式，其中最知名的便是 Stable Diffusion。Stable Diffusion 事实上是一套生图软件（或者称为代码）以及一类生图模型的总称，这有别于 Midjourney 仅限于官方提供的模型。它们的使用方式和绘图功能，相比于在线 AI 绘图软件几乎没有区别，但需要用户手动安装在本地电脑上运行。本地化部署的可能有很多，例如使用 AI 模型的早期用户通常具有一定的编程基础，从开源代码仓库（GitHub）开始到本地构建模型并不存在很多困难；AI 模型在快读迭代中，用户需要尽快使用最新版本的模型；开发团队资源有限，难以额外开发软件服务并构建背后所需的软件基础设施等。不管如何，本地化部署确实对大部分非专业用户造成了一定的门槛。然而从发展现状来看，本地化部署似乎没有阻挡 Stable Diffusion 成为最热门的产品之一，这大概得益于其令人折服的效果和 AI 作图的热度。



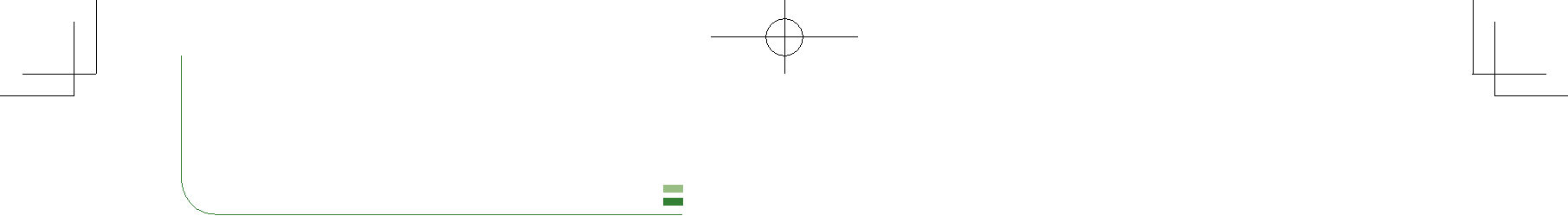
本地化部署同样为 Stable Diffusion 带来了一些独特的优势。除了前面所述的， Stable Diffusion 的开源允许用户定制自己的模型，这极大地丰富了社区中模型的丰富度。此外许多团队基于其开展研究，这对于模型技术的发展起到了良好的促进作用，特别是相对于闭源的 Midjourney 而言。此外，开源模型更便于用户反馈其中的潜在问题，有助于团队快速迭代，这符合互联网软件的敏捷开发原则。因此从软件开发的角度来说， Stable Diffusion 在将来会孕育出更加优秀的产品。

从 AI 绘图的效果而言，Stable Diffusion 也有显著的优势，其支持精细化调整，用户对生图结果的控制程度更高。它还可以和其他模型耦合构建复杂的工作流，从而变得更多样、更高效，这是作为定期发布的软件 Midjourney 所不具有的。尽管这些特性在 AI 作图的早期阶段尚未得到重视，但随着市场和落地的不断推进，其意义会逐渐展现。

此外，Stable Diffusion 团队也考虑了在线软件的优势。一方面，官方正在构建网页和软件版本的服务来提高用户使用的友好程度，如 Dream Studio ；另一方面，许多创业团队正基于此开发在线服务来作为官方模型的衍生品，从而填补空缺。总体而言，Stable Diffusion 作为一款功能强大的模型，正在逐渐获取其应有的市场。

009



动手学AI绘画：Midjourney从入门到实践

1.2.3 DALL·E 2

2021 年 1 月，OpenAI 推出了 DALL·E。一年后，其又退出了最新系统 DALL·E 2，能够以 4 倍的分辨率生成更真实、更准确的图像。OpenAI 开发了另一款革命性的大语言模型产品 ChatGPT，而 DALL·E 2 正是基于 GPT-3 模型和 Transformer 神经网络模型所构建的，具备强大的语言理解和文图转换能力。基于公开信息，DALL·E 2 首先将用户输入转化为类似图像的栅格编码数值，再通过图像生成模型（Diffusion）将栅格变成图像。



DALL·E 2 的工作模式类似 Midjourney，它运行在 OpenAI 的网页上。尽管用户

不需要下载客户端软件，但却很难在手机或平板电脑等移动端使用它，除非通过调用其

API 进行二次开发后使用。DALL·E 2 的网页界面非常简洁且操作简单，只提供了非常有限的调节和功能按钮。这大大简化了用户生图的复杂度，但也导致生图的复杂度相对有限。从运行的效果而言，DALL·E 2 的画作质量不如 Midjourney，模型丰富度不如 Stable Diffusion，但生图的速度比 Midjourney 快很多。



从发布的时间来看，DALL·E 2 是三者中最早发布的，并且一经问世便引起轰动。不过由于其综合特性不如另外二者，因此 DALL·E 2 的热度逐渐下降了。

1.2.4 对比不同AI作图工具

随着 AI 行业的日渐火热，不断有新的 AI 作图工具和平台出现。国际上比较知名的有 Adobe Firefly、Niji·journey、Novel AI 等，其中 Firefly 支持作为插件嵌入 Photoshop

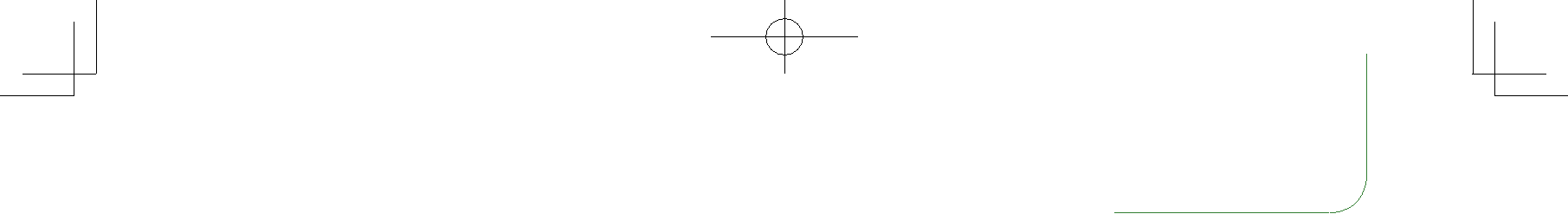
* Adobe 设计软件中，结合传统图像编辑软件使用，使其综合性能更为强大。国内的代表包括百度的“文心一格”、商汤的“秒画”、阿里的“通义”大模型等，其特点是比较类似 Stable Diffusion 的网页版，支持指定绘画风格模型或导入自定义模型，通过输入文本或图像引导模型进行创作。部分工具的简介如表 1-1 所示，注意这里的评价均为相互对比的相对评价，且随着工具的发展迭代，其表现会有所变化。
  + 1-1 部分 AI 作图工具的特点对比

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 名称 | 风格多样性 | 本地部署 | 画质高清 | 操作简单 | 自定义模型 |
|  |  |  |  |  |  |
| Midjourney | 所有风格 |  | √ | √ |  |
|  |  |  |  |  |  |
| Niji·journey | 二次元 |  | √ | √ |  |
|  |  |  |  |  |  |
| Stable Diffusion | 所有风格 | √ | √ |  | √ |
|  |  |  |  |  |  |

010



|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  | 第1章 | | 认识AI作图 | |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | 续表 | |  |
|  |  |  |  |  | | |  |  | |  |  |
| 名称 | 风格多样性 | 本地部署 | 画质高清 | 操作简单 | | |  | 自定义模型 | |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Disco Diffusion | 适合大场景 | √ | √ |  |  |  |  |  | √ | |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Firefly | 适合图像编辑 | √ | √ |  |  |  |  |  |  |  |  |
|  |  |  |  |  | | |  |  |  |  |  |
| “文心一格” | 通用场景 |  |  | √ | | |  |  |  |  |  |
|  |  |  |  |  | | |  |  |  |  |  |
| DALL·E 2 | 所有风格 |  | √ | √ | | |  |  |  |  |  |
|  |  |  |  |  | | |  |  |  |  |  |
| Novel AI | 二次元 | √ | √ | √ | | |  |  |  |  |  |
|  |  |  |  |  | | |  |  |  |  |  |
| 盗梦师 | 通用场景 |  |  | √ | | |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| 太乙 | 通用场景 | √ |  |  |  |  |  |  | √ | |  |
|  |  |  |  |  |  |  |  |  |  |  |  |



当前，AI 作图工具可以主要归为两派：基于本地部署的模型，其作图多样性、风格多样性都比较好，但操作本地部署本身具有一定门槛，不利于普通用户；基于网页或在线程序的工具，其绘画大都具有一定的局限性，如画质、风格多样性等，且无法自定义模型。在不同工具中，最热门的 Midjourney 和 Stable Diffusion 综合表现最佳，但 Adobe



* Firefly 由于可以和其他设计软件结合，且效果不错，后续可能会有较大的发展。此外，不管是企业的自研 AI 创作平台（如商汤的“日日新”、阿里的“通义”等），还是专门公司研发的 AI 作图平台（如 Tiamat、Vegaai 等），都有意延续 Stable Diffusion 的开源或有限开源模式，通过让平台支持多个细分场景的定制化模型来提升平台的生图效果和使用体验。

整体上，当前 AI 作图工具的发展局面概况如下。

（1） Midjourney 和 Stable Diffusion 依然是当下 AI 作图综合表现最佳的工具。但其他工具的发展迅速，有望追平二者。

（2） AI 作图工具的平台化趋势比较明显。鉴于 Stable Diffusion 开源社区为模型迭代带来的巨大贡献，不少新兴平台都采用了这一模式，鼓励用户为平台贡献模型，也主动接入外部模型，从而不断丰富自身的特性。

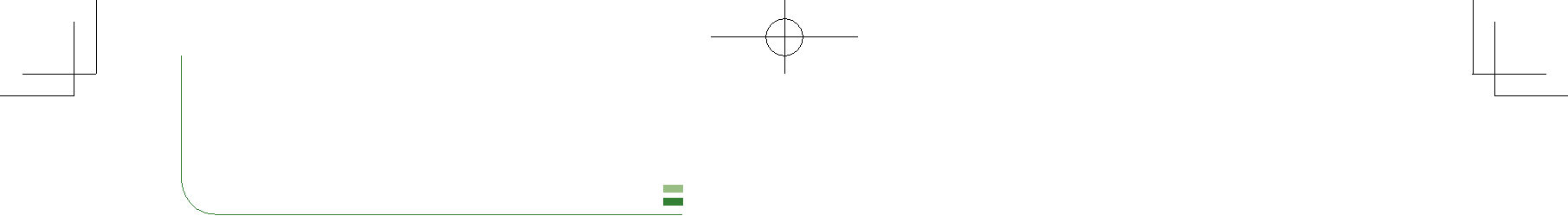
（3） AI 作图工具和其他 AI 内容生产的结合趋势比较明显。由于多模态信息可以提升

AI 大模型的效果，有条件的团队会选择将不同领域的 AI 生成模型结合开发，反哺 AI 作图的效果。

考虑到使用场景和需求，Midjourney 这类在线工具简单快捷易上手，新手即可创作出效果惊人的图画，适合希望专注 AI 创作且不排斥付费的用户；Stable Diffusion 这类

011



动手学AI绘画：Midjourney从入门到实践

离线工具定制化程度高，程序透明度和可控性较好，适合需要精细调整、具有一定专业基础的用户。但有条件的用户还可以将二者结合使用，将取得更好的效果。当前 AI 作图的发展正如火如荼，我们也期待将来会有更多成熟易用的产品出现。

1.2.5 为何选择Midjourney



以终为始，让我们从使用 AI 作图的目的出发，会更容易理解为何选 Midjourney。 “AI 作图”顾名思义，就是使用人工智能为使用者创作所需要的图片，至于作图过程中所

需要的设备条件、所需调控的模型、所需的技能都是为了达到作图目的的过程。真正的 AI 作图最终解放的是人类的双手，考验的是人类的想象力。就目前的 AI 作图工具而言，能够最大限度地实现解放双手、实现想象力的 AI 作图工具非 Midjourney 莫属。

Midjourney 作为 AIGC 现象级应用，由一个同名实验室自筹资金研发，仅有 11 名

全职员工却在一年内吸引了超过 1000 万名用户，并创造了 1 亿美元的营收。为什么众多

* AI 作图工具中，本书更推荐读者们选择 Midjourney 作为 AI 作画入门级的学习工具呢？下面介绍一下 Midjourney 的几大核心优势。

（1）使用门槛低。Midjourney 搭载在 Discord 社区中，不需要部署本地配置，不需要调试复杂的模型参数，对系统的硬件性能没有特殊要求，使用时也不需要任何代码。仅通过移动设备在线使用，输入合适的英文提示词即可生成美观的图像。这对于入门 AI



作画的普通用户来说，提供了极低的使用门槛。

（2）图片美观。不同于 Stable Diffusion 或 DALL·E 2，Midjourney 在生图反馈时模型具有适应性和对风格提示的响应能力，默认偏向于创作具有美学属性的图像，擅长运用多样的构图与和谐的色彩创作出细节清晰、完成度极高、令人满意的图像，能够在简短的提示词的基础上，反馈优质的图像。

（3）功能强大。Midjourney 不仅在作图方式上支持文生图、图生图、混合变换、图像提示等作图功能，生成图合成、图像转矢量等作图相关插件，也极大地拓展了 AI 作图的能力边界。

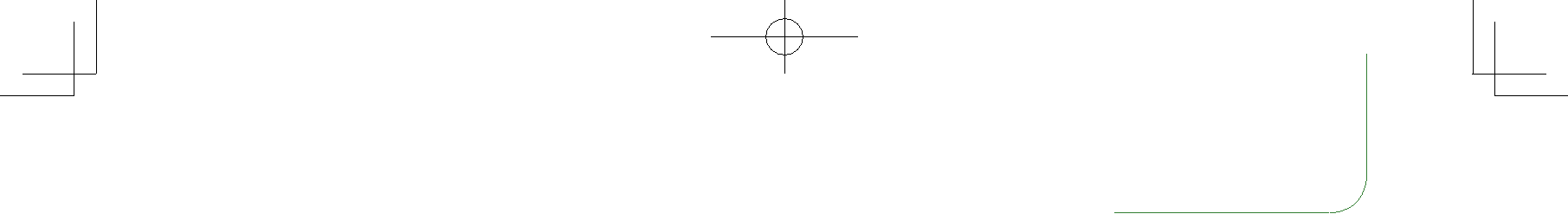
（4）潜力无限。Midjourney 自发布以来不断根据用户反馈迭代其绘图模型，至本书撰写时，已经完成发布了 5.2 版模型。在不断更新的模型中，其绘图功能也不断强大，例

* 5.2 版模型中新添加的强大的拓展图像功能，在未来更迭的模型中也将不断改善其缺点，丰富其作图功能，可谓潜力无限。

（5）无限使用。作为一款商业的付费产品，Midjourney 的用户通过订阅购买 GPU

012



 第1章 认识AI作图

的使用时间，但是对于标准计划及以上的订阅用户可以在订阅期间无限使用轻松模式进行生图。

（6）易于保存。Midjourney 通过自动存档在其官网的个人账户界面保存了用户以往生成的所有作品，用户无须担心丢失所制作的图像。此外，它还将所有缩略图保存在 2×2 网格中，用户仅需进行简单的批量操作即能将所生成的图像全部下载至本地。



1.3 AI作图的商业化道路



对于 AI 作图，与技术突破同样令人着迷的还有商业化。与其他高新科技不同的是， AI 作图在艺术、设计、娱乐等领域有明确而突出的商业价值，且这些领域具有显著的技术

壁垒。作为破局者的 AI 技术，前所未有地降低了普通人参与艺术设计的门槛，同时显著提高了设计品的生产效率，自然也拥有了一条明晰的商业化途径和巨大的发展潜力。然而，想象空间不代表毫无阻碍，不管是前期基础设施建设的成本、产业链的不完善，还是法律监管的不确定性，都需要相关从业者审慎行事。



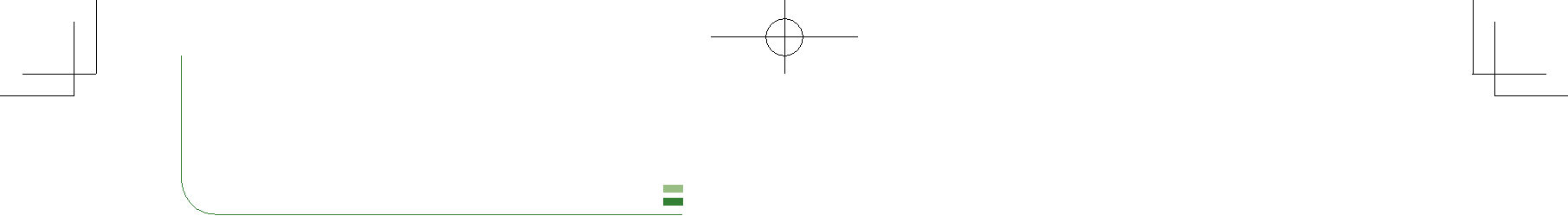
1.3.1 常见的AI作图商业化方式

尽管 AI 作图的商业化对公司和个人来说有所不同，但大体可以概括为直接商业化和间接商业化两大类。直接商业化是指借助 AI 制作可以直接销售的产品，例如漫画、表情包、绘画作品等；间接商业化是指通过 AI 参与商品设计制造的部分环节，从而降本增效并获取利益。简单来说，就是直接卖图或是借助图片售卖其他产品。对于可直接商业化的产品而言，其制作环节较少、成本和难度较低、参与门槛较低，适合新手入门，但获取的利润往往也比较有限；对于间接商业化的场景，AI 制作的图像可以用于较大的项目，其优化和提升对整体效益产生规模化的影响，例如网站优化带来的客流量可能导致营收的整体上涨，利润想象空间更大，但可能需要专业团队才能实现。

从具体的操作层面来说，笔者认为 AI 作图主要体现在两方面：对传统作图的替代，以及创新作图的应用。这里“作图”是广义的，是指一切以图像的创造和编辑为主体的工作，例如摄影、制作海报、设计效果图、设计二维 IP 形象等。在前述的两方面中，前者是指以往“人力 + 设计软件 + 采图设备”的作图模式将被 AI 作图在一定程度上取代，特别是低端、重复的工作；后者是指 AI 实现了规模化、多样化、高效率、低成本的创意制图，使得以往难以通过人力实现的工作变成了可能，例如超出从业者个人认知的超现实

013



动手学AI绘画：Midjourney从入门到实践

主义题材的设计或摄影。



 提示

不同商业化方式的利弊需要结合具体业务场景分析。



笔者根据商业化模式的不同，汇总了一些落地的具体途径，如表 1-2 所示。其中，“场景”一栏主要是指 AI 商业化的行业背景，然而可能存在同一商业化途径在多个行业均

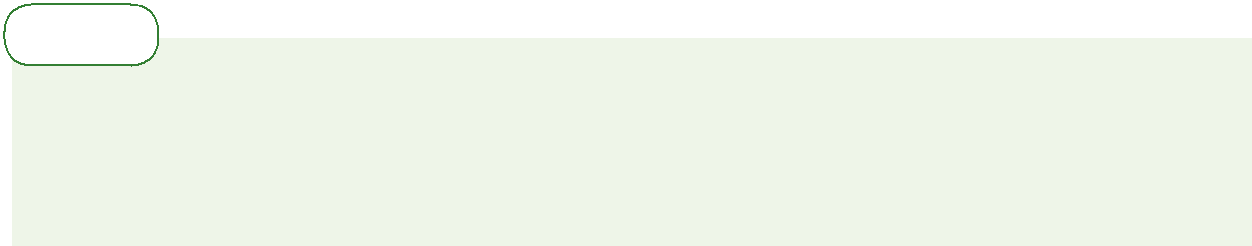
适用的情况，表 1-2 中并未一一赘述，仅列出典型作为参考说明。而“具体途径”一栏主要是指常见的具体商业化方法，其含义将在下文解释。

|  |  |  |  |
| --- | --- | --- | --- |
| 表 1-2 | | 不同场景下 AI 作图的直接和间接商业化途径汇总 | |
|  |  |  |  |
| 行业背景 |  | 商业化模式 | 具体途径 |
|  |  |  |  |
| 艺术、文化创意 |  | 直接商业化 | 美术创作、文创品设计 |
|  |  |  |  |
| 内容创作 |  | 直接商业化 | 图书插画、动漫设计、动画设计 |
|  |  |  |  |
| 新媒体创作 |  | 直接商业化 | 表情包设计、头像设计、Logo 设计 |
|  |  |  |  |
| 图像编辑 |  | 直接商业化 | 图像编辑、图像修复 |
|  |  |  |  |
| 电商、消费行业 |  | 间接商业化 | 包装设计、模特素材、商品特写、品牌 IP 设计 |
|  |  |  |  |
| 视觉广告 |  | 间接商业化 | KV 设计、海报设计 |
|  |  |  |  |
| 互联网 |  | 间接商业化 | UI 设计、网页组件设计、icon 设计 |
|  |  |  |  |
| 游戏 |  | 间接商业化 | 游戏场景、游戏角色、游戏素材 |
|  |  |  |  |
| 影视、摄影 |  | 间接商业化 | 布景设计、造型设计、服装设计 |
|  |  |  |  |



1.3.2 直接商业化：如何售卖AI作品

直接商业化的做法和传统制图、设计没有本质区别，设计师和画师依然在从事以往的工作，而仅将 AI 作图工具作为一种高效的设计制图软件来使用，正如 Photoshop 出现之前的工作者第一次接触 Photoshop 一样。

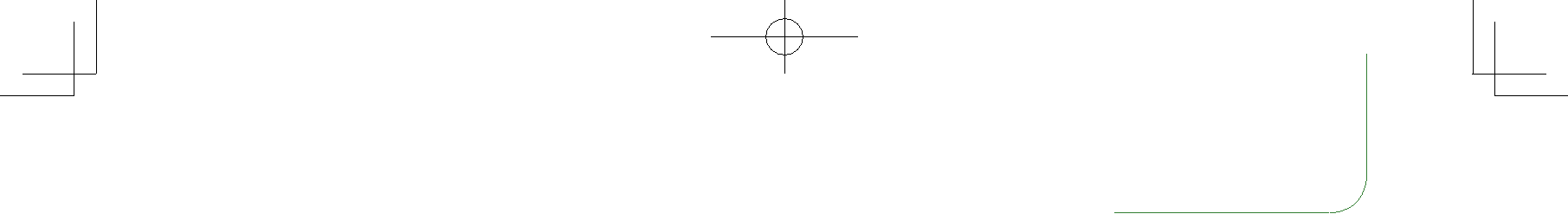


 提示

AI 作图工具的优势在于，它可以高效地生成图片，即使画面很复杂，也可以直接根据语言描述作画，大大提高了画面实现的准确度和效率。此外还可以创作出一些意想不到的效果，作为细节灵感的帮助。下面笔者会介绍在一些常见工作中 AI 制图工具的具体使用方法。

014



 第1章 认识AI作图

1. 艺术、文化创意：美术创作、文创品设计等

艺术、文化创意是最典型的 AI 作图商业化方式，即直接售卖制作的图像。不管是生成的漫画风格的图片，还是摄影风格的伪照片，只要制作精良、构思巧妙，或者正好符合买家的需求，都可以直接售卖——只要能确定这张图片的版权。目前许多软件都会宣称

AI 所制作的图像是“无版权”的，这意味着用户可以自行处理这些图片。更进一步地，用户可以将这些画作精心包装成文创商品，例如印着图画的杯子、衣服、游戏纸牌等，甚至只需将图像装裱起来就可以作为精美的装饰画来销售。通过售卖更昂贵的商品来提高



AI 图像的附加值，可以获取更高的商业价值。图 1-3 展示了 AI 图像用于文创产品的效果图。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| （a）美术创作 | | （b）定制文化衫 | （c）定制马克杯 | （d）定制枕头 |



图1-3 AI图像用于文创产品的效果图

2. 内容创作：图书插画、动漫设计、动画设计等

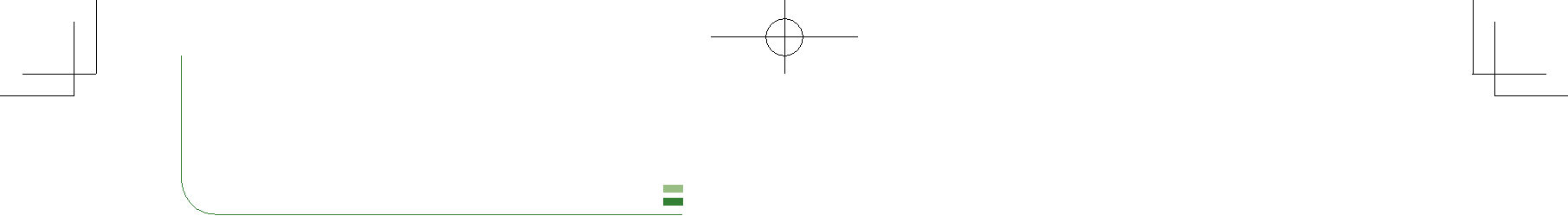
内容创作时，创作者需要绘制大量的背景图、多种姿态的角色图，并且需要一定的连续变化，从而让画面看起来更流畅。在 AI 作图工具的助力下，我们可以将图书插画、动漫或动画的生产流程解耦为独立环节，再看看怎么提高制作的效率。用户可以先拟定一个动画的故事大纲，设计基本的角色、场景和故事线等，为绘图打好基础。然后可以根据场景描述，利用 AI 作图工具制作不同的图像，特别是可以利用一些专用工具（如Niji·journey、Novel AI 等）来生成二次元画面；利用角色描述，一次性批量生产不同

姿态的角色图片，并选取合适的放置在场景中。由于 AI 作图工具可以连续多张作图，因此可以一次性生成连续的姿态动作变化的图片；或者只制作关键帧的图片，再通过 AI 插帧工具来补充中间的帧。最后，再为这些画面添加对白，即可完成一组简单的漫画或动画！如图 1-4 所示图像相关过程演示均通过 Midjourney 生成。

如果希望更加专业地制作动画，读者可以参考这套制作流水线，如图 1-5 所示，相信对你会有所帮助！

015



动手学AI绘画：Midjourney从入门到实践



（a）漫画场景 （b）漫画人物 （b）合成漫画

图1-4 AI生成漫画的效果图

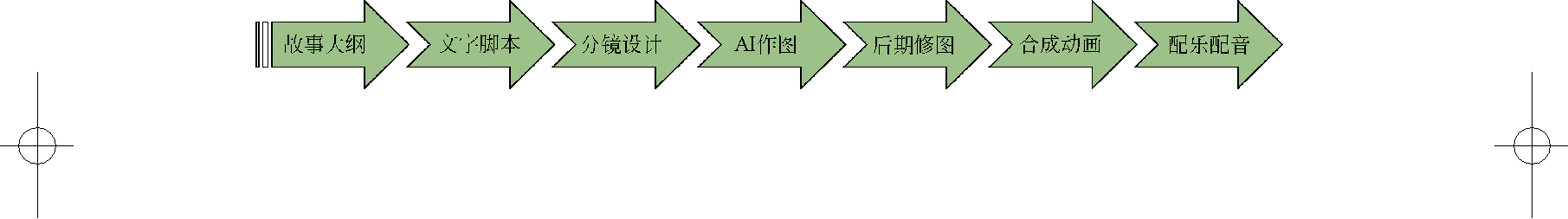


图1-5 动画制作流水线示意图

3. 新媒体创作：表情包设计、头像设计、Logo 设计等

通常我们在聊天软件中使用的头像包括二次元形象、大头照、风景、特定物品等，情侣之间还会使用情侣头像——具有暧昧意义且有强关联的两张情侣画像。对于单人所用的头像（前者），用户可以通过设定 1 : 1 的出图比例，通过提示词让 AI 生成想要的头像；对于情侣头像，用户可以设计 2 : 1 的比例制作图像，再分割为两张情侣头像。对于不满意的部分可以多次调整生成再挑选。图 1-6 展示了



一些 AI 制作的情侣头像。由于 Logo 设计的 AI 作图流程和头像类似，因此不再赘述，但用户需要注意 Logo 的商用版权问题。

|  |  |  |  |
| --- | --- | --- | --- |
| 对于表情包，可以按照头像的方式，根据预设 |  |  |  |
| 的表情含义或文字，按照九宫格的形式一次生成九 |  |  |  |
| 张图像，多次生成后挑选喜欢的几张，再用修图软 |  |  |  |
| 件添加相应的文字就可以变成一套表情包啦。最后 |  |  |  |
| 上传聊天软件之前记得转换为指定的图片格式哦！ | 图1-6 | Midjourney单次生成的一组 |  |
|  |  |
| 图 1-7 展示了一套制作好的表情包。 |  | 情侣头像 |  |

016





4. 图像编辑：图像编辑、图像修复

AI 作图技术可以用于图像编辑和修复，例如修复老照片、消除瑕疵、调整色彩等。用户可以借助 AI 提供的专业的图像编辑和修复服务，修复旧照片或改善图像质量。近年来，有人成功将 20 世纪初中国的黑白老照片修复上色，让当年的生活场景在一百年后得以再现，这正是基于 AI 修图的原理。这可以应用于个人用户、摄影师、媒体机构等，通过提供高质量的图像处理服务来获得收入。



 第1章 认识AI作图

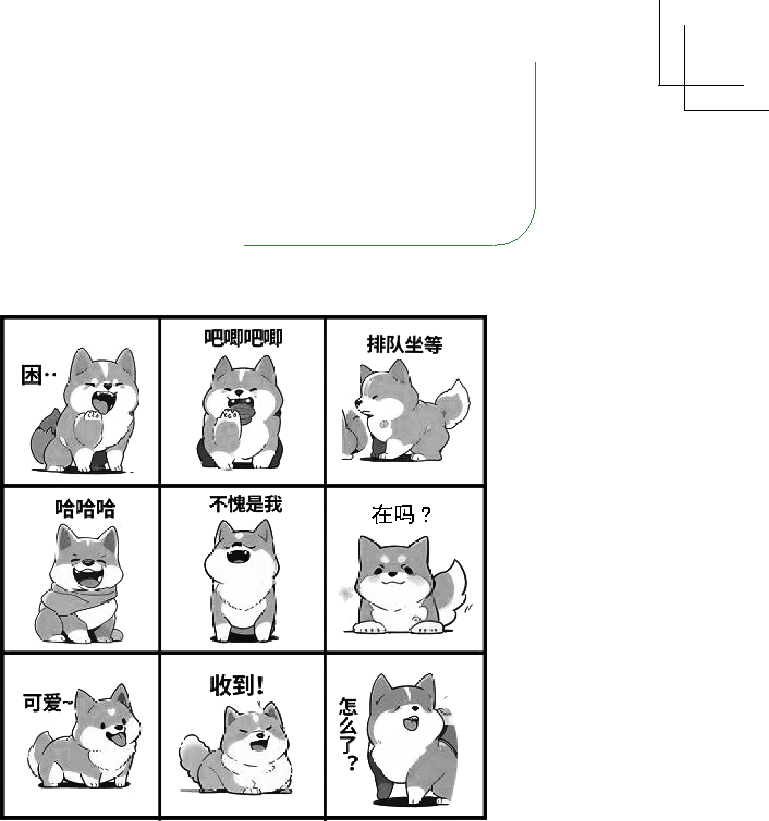
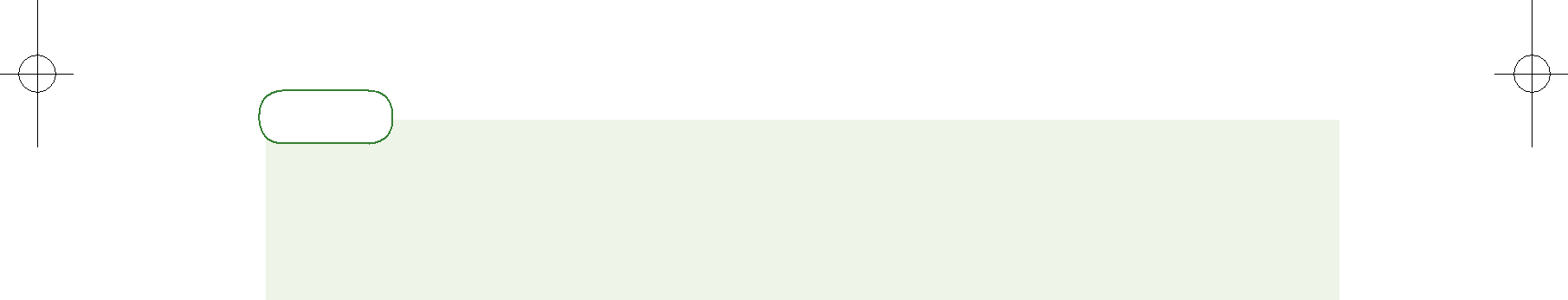


图1-7 Midjourney生成的一组表情包

1.3.3 间接商业化：借助AI实现降本增效

间接商业化依赖其他的商业模式或者工作，利用 AI 提高工作环节的效率或者质量来实现更好的销售效果，通过降本增效来获取额外利润，从而体现 AI 作图的价值。



 提示

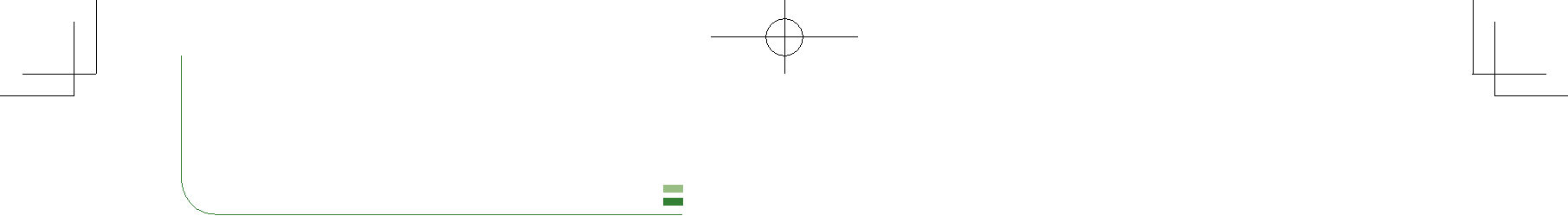
不同的行业多少都会和视觉图像打交道，例如售卖商品需要视觉广告、游戏制作需要视觉元素等，更别提影视制作这种直接输出视觉产品的领域了。因此间接商业化的形式在各行各业都有巨大潜力，等待我们去挖掘。

1. 电商、消费行业：包装设计、模特素材、商品特写、品牌 IP 设计等

在电商平台上销售产品，店家需要通过精美巧妙的商品图来展现产品的特性、吸引顾客。这些商品图包含多个方面，例如商品的包装和外观、使用效果图、材料特写图、品牌调性场景图和模特展示图等。以往为了制作这些不同的图像，店家需要分别设计、拍摄，再进行后期处理，一些环节还可能需要聘用专人模特甚至外部团队协作，如果遇到需要调整的情况，耗资将会更加巨大。如今借助 AI 作图工具，店家可以快速生成不同的包装效果图，再根据最佳的展示效果定制专门的包装材料；可以利用 AI 生成真人模特图片，再添加真实产品（如衣服、包包、玩具等）合成模特与产品的特写；利用 AI 作图工具可以对产品图合成不同场景、光影效果下的角度特写，从而节省大量专业拍摄的成本和时间。图 1-8 展示了利用 AI 生成咸鸭蛋的包装效果图，辅助产品销售。以上种种都表明了如何

017



动手学AI绘画：Midjourney从入门到实践

通过 AI 生成的优质图像来充分展示产品的特点，提高销量，获取利润。



（a） （b）

图1-8 利用Midjourney设计咸鸭蛋的包装盒

2. 视觉广告：KV 设计、海报设计等



KV 设计是一种广告设计模式，是指选取一个最具代表性的视觉元素，作为核心形象传达广告的主题和信息，本质上 KV 设计也是海报设计的一种。对于一般的海报设计，通常由背景、主体、文字三层构成，和漫画的背景、角色、对白三层结构非常相似，因此我们可以仿照漫画的 AI 作图流程。在确定好宣传方所需要的海报主旨、设计风格、尺寸比例等设计参数信息之后，就可以撰写合适的提示词，



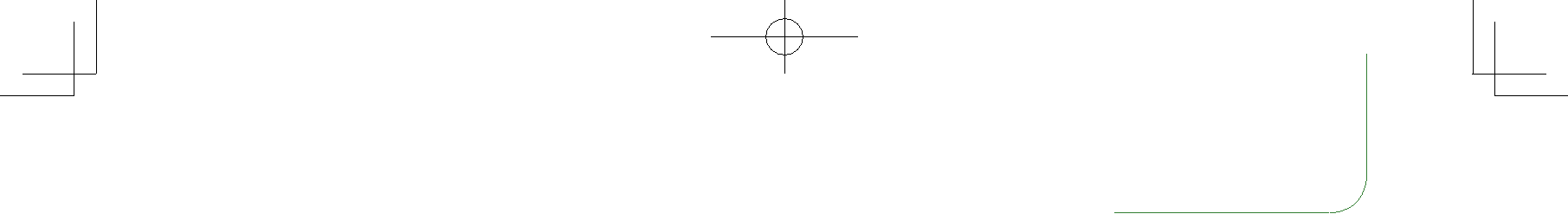
通过 AI 生成合适的图层元素。每个图层可以一次生成多张图片，再选取合适的进行调整、补充细节。最后结合海报画面的排版特点，把前景文案添加到海报上，即完成制作。图 1-9 为一幅利用 Midjourney 制作的端午节粽子海报：首先生成大粽子的背景图，其次将单独制作的两个嬉戏的小孩添加到图层上，最后添加文字即完成。AI 将原本需要将近一周时间才能完成海报工作缩短到半天，大大提高了生产效率。借助更精美的海报，宣传方可以达到更好的传媒效果，提高活动举办的成效。

1. 互联网：UI 设计、网页组件设计、icon 设计等

在互联网场景下，用户最常接触的就是网页或移图1-9 端午节粽子海报图

018



 第1章 认识AI作图

* App 的界面，这些前端页面需要高效简洁地向用户展示不同功能按钮，或者特定的信息。有些软件，例如游戏程序，还会特地将 icon 设计成某些风格以体现企业的品位或迎合用户的喜好。然而在复杂的程序中保持更新这些页面组件是一个具有挑战性的事情，大企业的 UI 设计师不得不为此奔波劳命，不然企业的软件就会被用户吐槽“万年不变”！

AI 作图工具可以帮助设计师快速生成网页组件模块的样式，并搭建出页面的雏形，进行

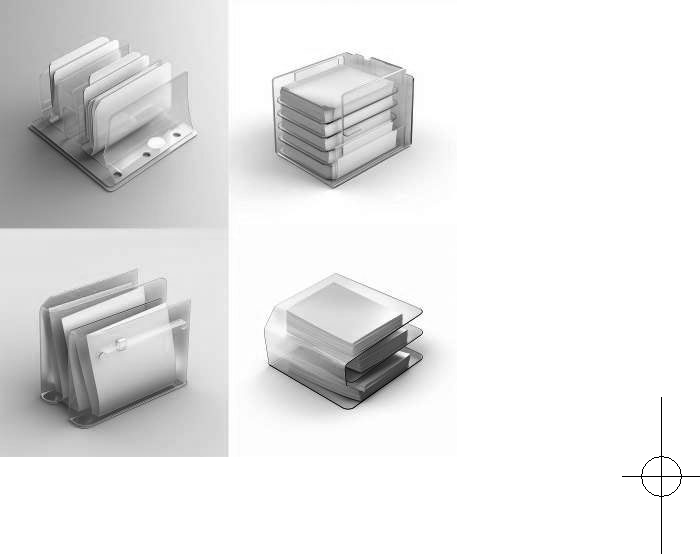


快速迭代设计，实现前端交互设计的敏捷开发！首先根据企业调性，选择合适的组件风格。以页面“按钮”为例，我们需要为一家高新技术企业设计 3D 立体风格的组件，可以通过 AI 作图工具设计出一批按钮，再选择一张合适的图片，通过“图生图”方式扩展生成其他功能的 icon。图 1-10 为一组文件夹

* icon 图片。在获得合适的模块设计后，就可以继续设计页面的背景图、动态效果等。当然也可以反过来，预先生成完整的页面效果图，再逐步完成每个组件的细节。设计师节省了烦琐的绘图时间，从而得以专注于整体的设计。
  1. 游戏：游戏场景、游戏角色、游戏素材等



图1-10 利用Midjourney生成的页面icon图像



不同于前面提过的漫画设计，游戏中所使用的素材只有一部分可以按照“背景—角色—文字”三图层的模式进行设计。游戏美工设计更加全面和复杂，涵盖地形、建筑、植物、人物、动物、怪物、道具、动画、特效、界面等，例如一整套的道具、不同角色的形象、超大场景和地图等。相关元素可以大体分为静态元素和动态元素两类，静态元素根据规模大小，所使用的 AI 作图方式会有所区别。例如道具、卡牌之类的元素可以独立生成，而诸如大地图之类的复杂图片则可能需要多次生成子图再拼合。对于动态元素，则涉及动态效果、角色模型等，这些对于现有的 AI 作图工具还存在不少挑战，可能需要结合其他建模软件才能完成。图 1-11 为一组 Midjourney 生成的小型场景的效果。

值得一提的是，AI 作图可能对中国风游戏有着特别的优势。由于国风游戏追求细腻精美的画面，由画师手工绘制这些繁复的细节是一件十分困难且耗时的工作，而 AI 工具可以高效地生产国风图像，画师只需作一定的修改即可使用，大大加快了游戏上线和更新的速度。

019



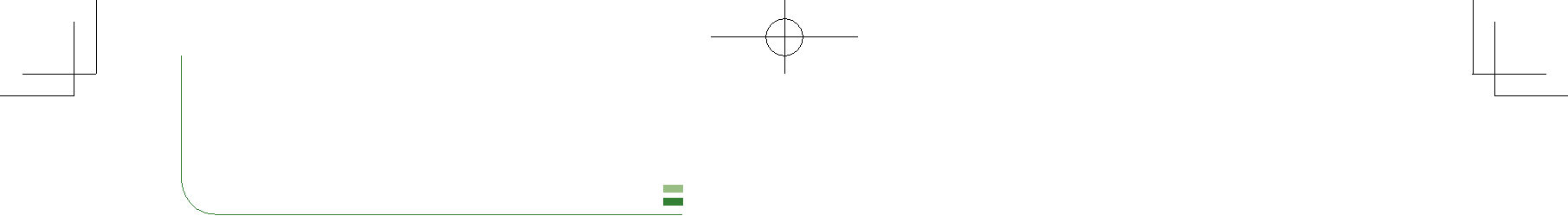
动手学AI绘画：Midjourney从入门到实践



图1-11 Midjourney生成的一组游戏街景图

|  |  |
| --- | --- |
| 5. 影视、摄影：布景设计、造型设计、服装设计 |  |
| 由于商业摄影需要选取合适的背景、机位、 |  |
| 模特、服化道、镜头、灯光等，因此提前模拟拍 |  |
| 摄效果，甚至通过 AI 生成某些场景图，对降低 |  |
| 摄制成本、提高出片效率有着重要的帮助。包 |  |
| 括 Midjourney 在内的 AI 作画工具可以直接生 |  |
| 成高清的人物、物品、风光图片，对于合成这 |  |
| 些图片有着巨大的优势。即使图像的分辨率达不 |  |
| 到要求，也可以借助图像增强放大软件来完成。 |  |
| 图 1-12 展示了一张赛博朋克主题的电影剧照， |  |
| 是通过 Midjourney 生成的，对于同题材电影的 |  |
| 编剧、摄像和导演等工作有着重要的帮助。 | 图1-12 AI生成的电影剧照 |



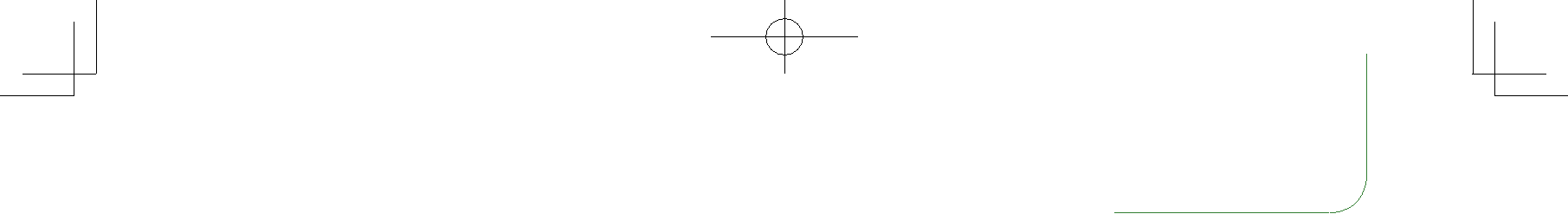
1.3.4 其他商业化方式

除了以上提到的 AI 作图商业化方式，还有其他针对 AI 作图这一概念和过程本身的商业化手段。下面列举了一些容易实现的例子，以供读者参考。

（1）开发和销售 AI 作图工具和软件。如果你具备 AI 技术的开发和编程能力，那就可以开发和销售 AI 作图工具和软件，这些工具和软件可以帮助客户利用 AI 技术进行创作、设计、处理图像等。你可以将这些工具和软件出售给个人用户、设计师、摄影师等，

020



 第1章 认识AI作图

并提供技术支持和更新服务。你也可以创建一个订阅平台或会员计划，为用户提供独家

* AI 作图内容和服务。通过提供高质量的、定期更新的作品、教程、资源等，吸引用户订阅并支付会费，正如 Midjourney 等在线平台一样。

（2） AI 图商。不管是 AI 作图模型的训练，还是广大用户利用 AI 工具生成的图像，都需要一个较大的 AI 图片素材平台来托管和交易这些图片。通过构建和运营 AI 作图平台或市场，为创作者和用户提供交流和交易的平台，让创作者展示和销售他们的 AI 作图作品，同时还可以借助平台提供购买和使用 AI 作图服务的渠道。

（3）教育和培训。利用你在 AI 作图方面的专业知识，可以提供在线教育和培训课程。可以开设教学视频、指导材料、在线课程等，帮助学习者掌握 AI 作图技术。通过出售课程或收取学费，实现收益。当前正值 AI 作图的热潮，开设相关课程将更容易吸引学员。

（4）社交媒体和内容创作。利用社交媒体平台，分享你的 AI 作图作品和技术知识，建立自己的粉丝群体。通过吸引大量的关注者和用户，可以获得品牌赞助、推广合作、付费推文等机会，从而实现商业化。

（5）研发和创新项目。你可以参与有关 AI 作图的研究和开发项目，申请相关的科研资助和项目拨款，并争取研究资助、投资或合作伙伴的支持。通过开展创新项目，推动



AI 作图技术的前沿研究，并将其转化为商业应用和收益来源。这可以帮助推动技术的进

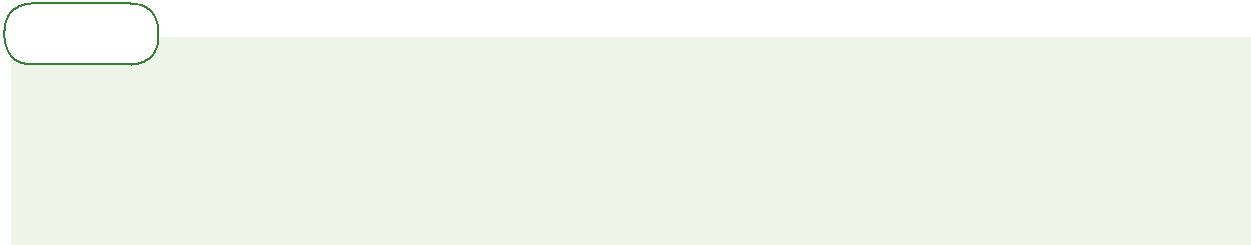
步和创新，并为公司带来声誉和商业机会。

AI 作图工具对人类日常生活的影响是普遍而深刻的，它和我们生活的交集越多，可

供商业化的渠道就越多。观察自己生活中和图像有关的场景，你将发现更多商机！

1.3.5 AI作图的法律问题

应当注意的是，在以商业目的使用 AI 制作的图像时（包括前文的直接和间接商业化），法律问题是我们始终需要关注的。



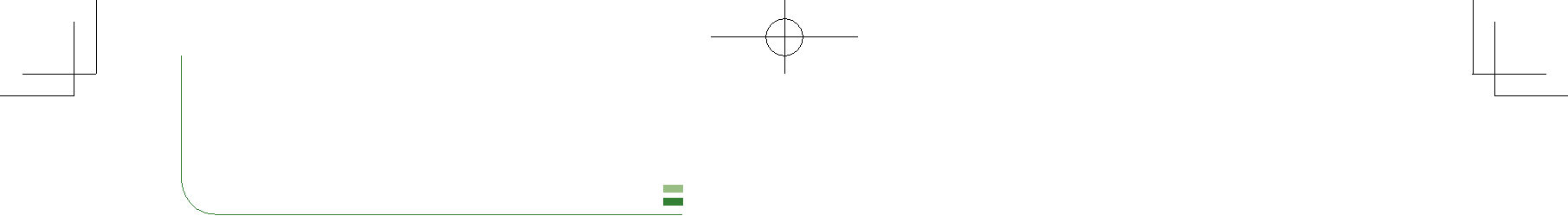
 提示

本书希望引导读者共同探讨这一问题，然而由于 AI 技术相关立法还处于“灰色地带”，现有法律条文在 AI 的冲击下面临不再适用、需要修订增补的境地，因此本书的观点仅能反映写作时的情况，读者应当结合最新情况来自行判断。

由于 AI 作图方式的特殊性，即通过计算机学习拟合海量图片后，再通过控制参数生

021



动手学AI绘画：Midjourney从入门到实践

成，因此很容易被理解为对现有图像的重新组合。在这样的观点下，不仅不能认定 AI 工具生成图像的原创性，而且面临对 AI 模型所使用的训练数据（即学习图像）的版权侵犯问题。

截至目前，世界各国对 AI 作图的版权问题依然没有定论。这意味着我们不得不相机行事。许多 AI 作图工具的开发公司发布了相关的版权限制和声明，用户可以对使用的平台自行查看。以 Midjourney 为例，在其官方的用户协议相关条款中 1，第 4 部分即为版权和商标说明，原文请查看官方网页。现将 2023 年 7 月 17 日所展示的部分相关条款翻译如下，请注意只有英文原版可以反映 Midjourney 官方的原意。



版权及商标

在本节中，付费会员指已订阅付费计划的客户。

您给予 Midjourney 的权利

通过使用本服务，您向 Midjourney 及其后继者授予一项永久的、全球性的、非排他性的、可再许可的、免费的、免版税的、不可撤销的版权许可，以按照您的指示复制、准备衍生作品、公开展示、公开表演、再许可和分发您输入本服务或本服务生产的资产中的文本和图像提示。本许可在任何一方因任何原因终止本协议后仍然有效。



您的权利

根据上述许可，您拥有使用服务创建的所有资产，前提是这些资产是根据本协议创建的。这排除了升级其他人的图像，这些图像仍然由原始资产创建者拥有。 Midjourney 不就可能适用于您的现行法律做出任何陈述或保证。如果您想了解您所在

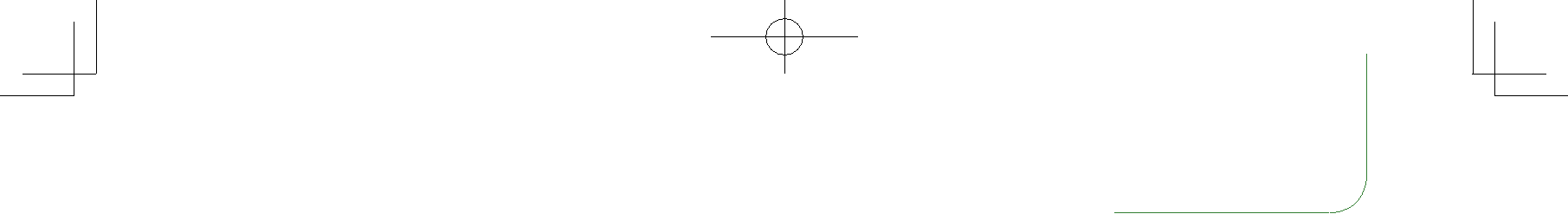
司法管辖区的现行法律状况，请咨询您自己的律师。即使在随后的几个月里您降级或取消了您的会员资格，您对您创建的资产的所有权仍然存在。但是，如果您属于下列情况，则不拥有资产。

如果您是年总收入超过 1000000 美元的公司的员工或所有者，并且您代表您的雇主使用“服务”，则您必须为代表您访问“服务”的每个个人购买“专业”会员资格，以便拥有您创建的资产。如果您不确定您的使用是否代表您的雇主，请假设它是。

1 Midjourney. Terms of Service[EB/OL]. Midjourney Documentation, 2023 年 6 月 8 日。

022



 第1章 认识AI作图

如果您不是付费会员，则您不拥有您创建的资产。相反，Midjourney 根据知识共享非商业性 4.0 国际署名许可（“资产许可”）授予您使用这些资产的许可。

请注意 Midjourney 是一个开放的社区，允许其他人使用和重新混合您在公共环境中发布的图像和提示。默认情况下，您的图像是公开可见和可重新混合的。如上所述，您授予 Midjourney 许可允许此操作。如果您购买“专业”计划，可以绕过这些公共共享默认设置。



如果您作为“专业版”订阅的一部分购买了“隐形”功能，或通过先前可用的附加组件购买了“隐形”功能，我们同意尽最大努力不发布您在服务中采用“隐形”模式的任何情况下制作的任何资产。

请注意，您在共享或开放空间（如 Discord 聊天室）中制作的任何图像，无论是否使用隐身模式，该聊天室中的任何人都可以查看。



1.4 高效的AI作图工作流



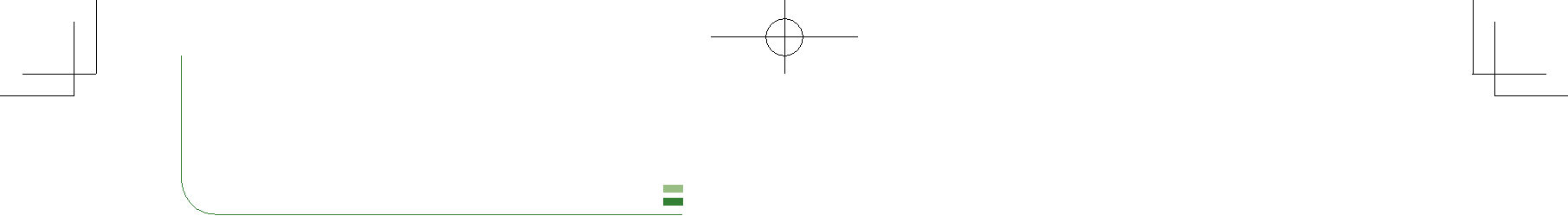
对于商业化、工业化的应用场景，存在批量化、大规模、定制化作图的需求。往常这些大量任务主要是通过多个设计师或画师分散任务或协作来完成，然而对于程序化的 AI 作图工具而言，将会有更加成熟的流水线做法，这里称为 AI 作图的工作流（Workflow）。通过构建高效的 AI 作图工作流，我们期望可以保持高水平的图像生成，且尽量提高机器生图的效率，从而充分发挥 AI 的力量！

1.4.1 什么是AI作图工作流

工作流技术起源于 20 世纪 70 年代中期办公自动化领域的研究，不同领域的“工作流”有着类似的定义。对于计算机行业而言，工作流是指业务相关的事务在计算机中的自动化，是对工作流程及其各操作步骤之间业务逻辑的抽象和概括。构建工作流的主要目的和效用，是为了实现某个业务目标，借助计算机利用自动化流程来传递信息、文件或执行任务。具体到作图领域，可以理解为将设计师的传统作图流程，分解为一个个标准化的任务来描述。不管是什么主体的作图，都应该涉及这个作图工作流的一部分环节或整体。图 1-13 给出了一个传统的作图工作流的示意图。

023



动手学AI绘画：Midjourney从入门到实践

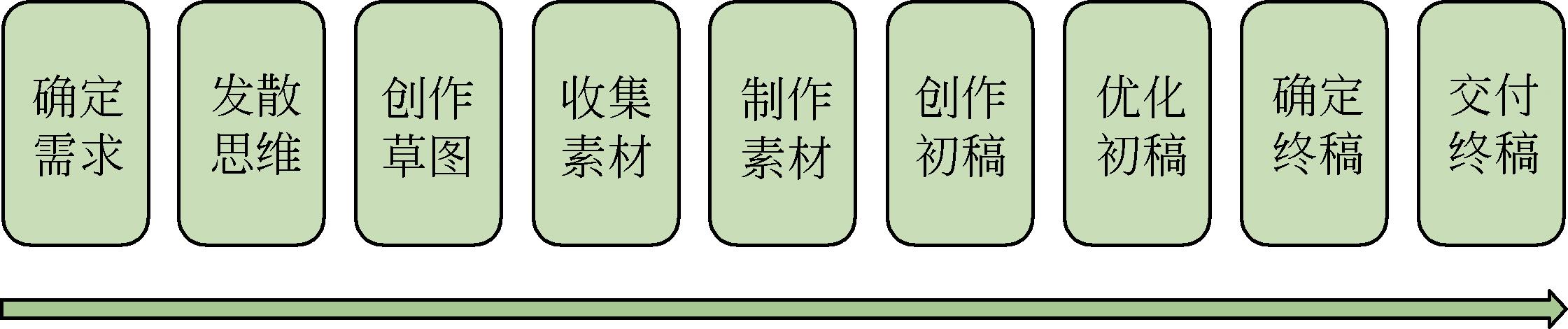


图1-13 传统作图工作流示意图，顺序为从左到右



* AI 作图出现之前，计算机或其他自动化设备对设计师作图流程的介入和辅助是有限的。这可能主要出于以下几点原因。

（1）之前的计算机图像软件局限于图像编辑的功能，用户只能利用其构建或编辑已有的图像，且编辑效果受限于已有图像和素材元素。对于 3D 建模软件而言同样如此，这导致作图效率和多样性的局限性。

（2）现有软件的操作指令完全通过操作者输入。尽管开源模板的分享使得许多设计师的工作有所减轻，并降低了新人创作的门槛，但依然依赖优秀模板设计师的努力。

（3）现有软件的操作效率和可控性都比较低，难以实现自动化运转的同时还保持一定的效果。由于 Midjourney 等基于最新技术的 AI 模型已经展现出了惊人的效果，上述问题都已经得到很大程度的缓解，甚至可能在不久的将来会被彻底解决。作为紧随前沿的从业者，我们有必要思考“设计自动化”“作图工作流”等这些全新概念对自身工作和行业的意义。

结合前文的描述，这里给出一个简单的 AI 作图工作流的示意图，如图 1-14 所示。其中发散思维和创作草图的环节变成了确定提示词，收集、制作素材的工作改为借助 AI



完成，优化初稿的工作也将大量借助 AI。此外其他环节也在不同程度上因为 AI 的使用而有所区别。根据笔者的理解，AI 带来的优化可以主要概括为以下两个方面。

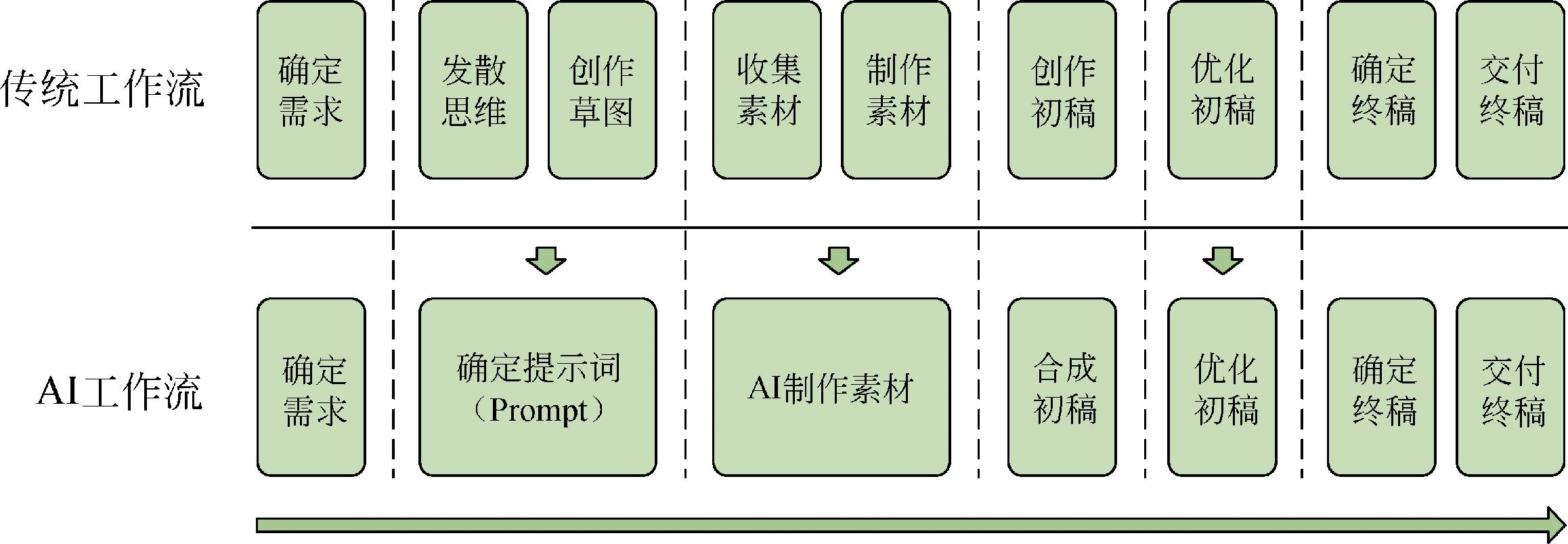
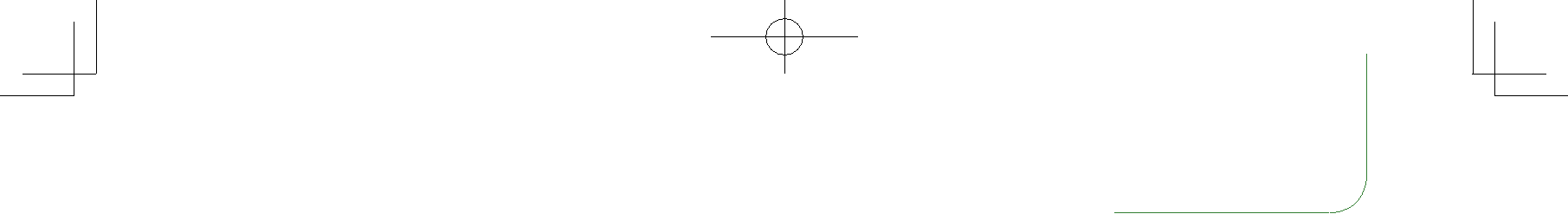


图1-14 传统作图工作流和AI工作流的对比（虚线和向下箭头指示了AI替代的环节）

024



 第1章 认识AI作图

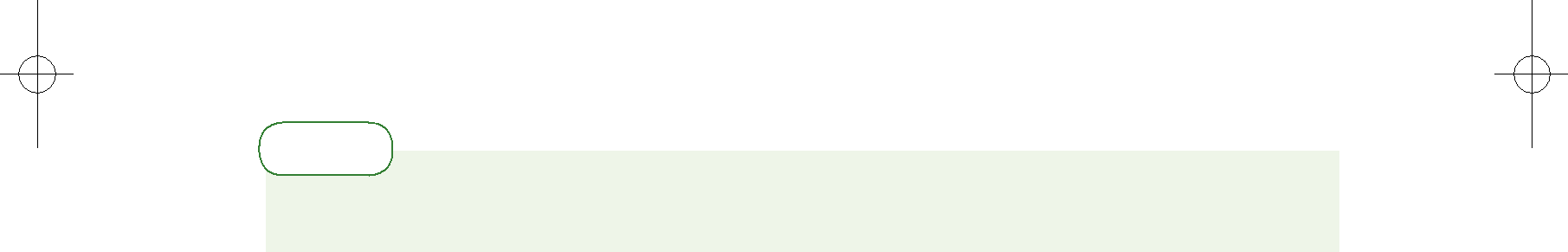
（1）加速思维创作。由于 AI 工具接入了自然语言描述作为作图的输入，因此设计师获得了“所想即所得”的创作体验，在思维风暴阶段的碎片化灵感将可以被立刻实现，即使是微弱的想法和印象也可以被 AI 强化后可视化展现出来，显著加速了前期的草稿创作。

（2）改变制图模式。首先，AI 作图将传统流程中复杂的设计过程简化为提示词的撰写，仅仅要求设计师准确地描述画面即可，在效率和准确性上都有显著的提高。其次， AI 可以制作复杂多样的图像，在多数情况下都可以替代简单的作图，因此传统的从 0 到



* 的制图模式转变成了“AI 制作基础图像 + 人类微调的辅助制图”模式，在效率和质量上有显著的提高。

根据笔者前面的描述，不难看出在 AI 作图的工作流下，设计中琐碎、简单的环节将被 AI 所取代，人类设计师只需要介入质量把控、后期优化等关键环节，包括前端需求确认在内的大部分环节甚至可以开放给需求方，让上下游协同工作，实现更好的交付效果。事实上，许多互联网大厂都已在不同程度上接入 AI 设计的工作流，例如阿里巴巴的“鹿班”AI 设计系统就一直为电商不停地生产 banner、海报等图像，极大地方便了店家的设计和运营。因此，在 AI 作图的工作流框架下，基于计算机的自动化有望成为将来商业化、工业化制图的发展趋势。



 提示

AI 对于现有工作流程和模式的影响还是一个开放性问题，你有什么样的想法呢？

1.4.2 面向商业设计的工作流

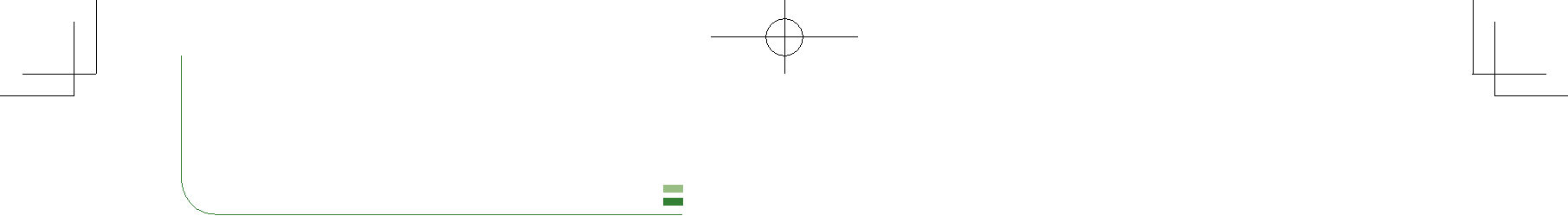
本小节我们将结合实际落地场景，展示一个面向商业设计的 AI 作图工作流的基本面貌。如图 1-14 所示，AI 作图的工作流在确定需求之后，需要根据需要撰写提示词，并根据提示词生成相应的素材，最后组合素材，进行后期编辑、迭代即可定稿。下面结合一些商业化场景给出典型案例。

（1）招聘网站视觉设计。招聘网站展示了多个不同领域不同岗位的招聘信息，此外在不同的运营宣传活动中，都需要大量的展示图片来展示信息、引导用户。参考赶集直招团队的案例 1，面向一场多岗位招聘活动的设计包括“设计资产准备”“AI 品牌语言表达”等主要部分，对应图 1-14 中的“AI 制作素材”的环节。在“设计资产准备”中，设计师需

* 58UXD. AIGC 拉升设计生产力——赶集直招 AI 设计实战案例全解析 [R/OL]. 58UXD, 2023 年 5 月 25 日。

025



动手学AI绘画：Midjourney从入门到实践

要根据具体岗位搭建符合要求的人物素材和物品道具素材，AI 生成的任务免去了模特拍摄的流程和开销，但在外貌、情绪、姿态等方面可以更加符合使用需求，并可以按相同风格和标准生成不同行业细分岗位的角色。在“AI 品牌语言表达”中，企业需要通过一套规范来保证设计调性的稳定性和独特性，而 AI 除了准确生成图像主体元素之外，还可以通过背景、装饰模块的生成更加精准地描述企业的设计调性。



（2）电商平台广告设计。电商平台在商品、品牌、营销、体验等多个领域对平面设计都存在需求，网易严选创新设计中心对 AI 作图在其中的作用进行了应用和探索 1。团队针对设计调研、设计提案、获取素材、包装设计、工业造型设计、商品图案设计、IP 设计等多个环节实践了 AI 作图的效果，结果表明 AI 作图可以普遍提高设计效率，特别是在辅助创意提案和辅助拍摄两方面。通过：①草拟内容组织 Prompt ；②修改 / 垫图控制生成结果；③二次调整修改细节；④整理并输出方案。

这样一个较短的工作流，设计师就可以在短时间内完成一个设计提案。为了更好地使用 AI 模型，还可以使用类似 ChatGPT 等自然语言 AI 模型来优化提示词的写法。围绕网易严选的调性，可以针对角色、场景、服化道等对象生成摄影素材图，多样性的素材可以大大降低制作成本，提高效率和质量。团队基于这些结论进行补充实验，表明 AI 作图可以达到一定程度降本增效的作用，并沉淀出了以下方法论：①摄影素材 Prompt ；②实拍摄物；③设计制作。在商品图的生产上可以构建稳定通用的工作流。



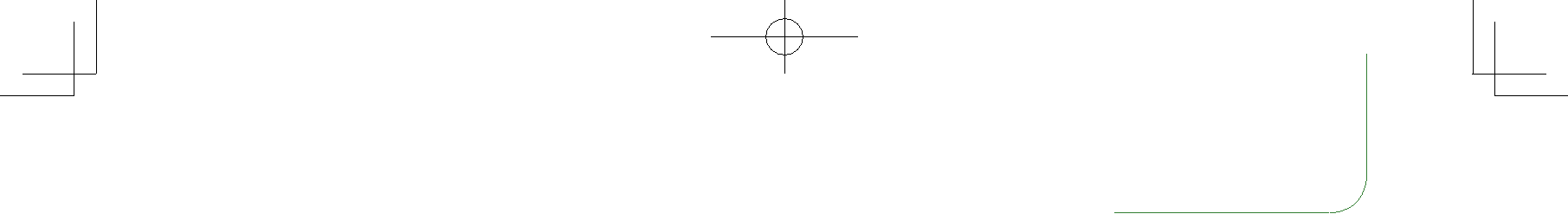
（3）游戏公司视觉设计。游戏领域同样也是一个大量需要图像制作的领域，网易互娱 ASAK 设计团队展示了其在多个方面实际落地的经验 2，包括活动弹窗、勋章图标、IP 形象、KV/ 海报、场景、动画分镜、物料等多种元素的设计。

报告提到，结合 Midjourney 和 Stable Diffusion 的协作设计可以极大地缩短概念风格、设计初稿和设计终稿的反复沟通时间，整体上可以提高 25% ～ 55% 效率。具体而言，以活动弹窗为例，游戏中新活动如果需要展现具体的奖励目标，则需要在弹窗中展现目标的形象。因此先运用 ChatGPT 获取弹窗的设计思路，尽可能提取多角度的描述词作为提示词，并结合游戏领域增加“荣耀”“宝石”“高级感”等符合玩家心理的描述词，便可以放在 AI 工具中反复生图，直到获得满意的效果为止。

* 严选创新设计中心 . AIGC——探索 AIGC 在网易严选中的应用 [R/OL]. 网易严选创新设计中心 IDC, 2023
* 4 月 23 日。
* 网易互娱 ASAK 设计团队 . 八大实战案例！ AIGC 在网易落地项目中的运用 [R/OL]. 优设网 , 2023 年 4 月 11 日。

026



 第1章 认识AI作图

对于 IP 设计，设计师对形象有一个简单而具象的预期，因此首先可以绘制线稿，线稿与提示词在生图时具有同等的提示作用。其次根据需要选择 MeinaMix、BlindBox、 ControlNet 等插件，在 Stable Diffusion 中实现预期的效果。最后设计师灵活调整参数多

生成几次，挑选出最合适的图片进行后期处理。相比于图 1-14 中展示的流程，这里在提示词阶段使用了线稿，在生成素材阶段使用了额外的软件插件，但整体思路还是一致的。



在确定好适合自身业务的工作流之后，设计师可以形成一套自己的工作习惯，也可以更进一步地借助代码或程序自动化执行软件和部分事务，真正解放自己的双手！

1.4.3 如何打造高效的工作流

在前面的小节中，我们介绍了基本的 AI 作图工作流及其在具体落地场景中的应用案例。为了进一步使工作流符合批量化、流水线生产的需求，我们还可以做进一步的改进。

（1）快速出飞机稿。在设计初期为了探索、确定需求而作的额外的草图被称为“飞机稿”，快速绘制尽可能多的飞机稿可以帮助设计师和上游需求方在需求和细节上达成一致。以往人工绘制飞机稿费时费力，沟通成本较高，而 AI 工具的高效率可以克服这些困难，可以将劣势转换为优点。设计师可以在工作流中加入多轮草稿创作的环节，将确定需求后撰写的提示词不断作图试错，快速和需求方根据大量飞机稿迭代出最贴切的需求描述和提示词，提高作图的准确性和整体效率，减少返工。



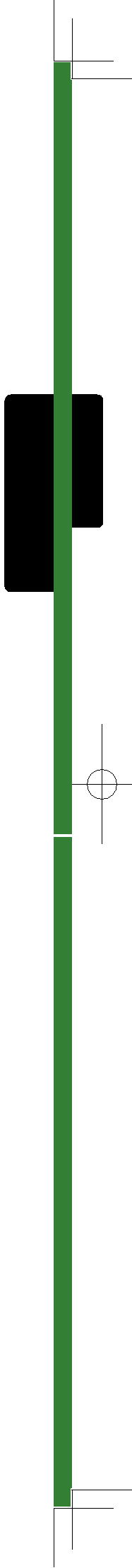
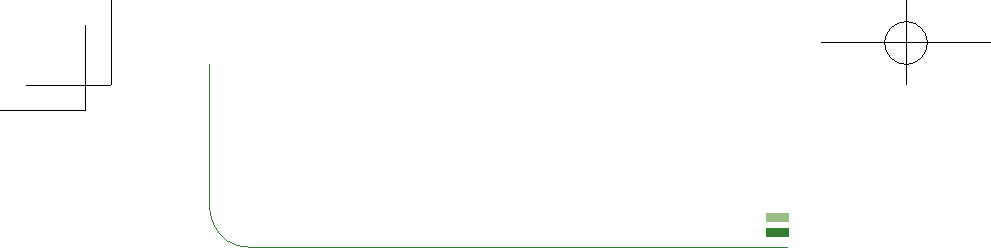
（2）分级设计。在实际运营中，不同的设计需求难度不同，复杂项目往往需要更多的设计流程和更高的交付标准，设计师所使用的方法、付出的实践精力也不一样。因此设计师可以根据实际情况将设计项目分级，分别设计不同的 AI 作图工作流来应对。根据腾讯社交用户体验设计部的实践案例 1，可以按活动类型、周期、对应设计复杂程度，将设计需求分为 S、A、B 三个级别。其中 S 级别以专题大促为主，页面以 banner 和 H5 为代表； A 级别以平台活动为主，同样是 banner 和 H5 需求并存；B 级别为日常促销，以 banner 需求居多。相对应，B 类需求可以使用轻量手段快速达成，可以直接生成画面元素再简单进行修改完成；A 类需求通常需要在探索风格、绘制画面细节上花费较长时间，需要撰写提示词并垫图后利用 AI 辅助生成，有时候还需要二次生成和额外调整；对于展示周期最长、质量要求最高的 S 级设计，其设计流程和 A 类需求基本一致，但需要设计师更加熟悉 AI 工具的参数和插件使用方法，进行更加精细的控制生成和后期调整。如果希望优化

* 类设计需求的使用体验，可以在工作流中增加批量作图的次数，设计师可以通过大量筛

1 腾讯社交用户体验设计部 . 运用 AIGC 人工智能生产内容 [R/OL]. Tencent ISUX Design, 2023 年 3 月 1 日。

027



动手学AI绘画：Midjourney从入门到实践

选来减少修改、调整图像的时间。

（3）组件化设计。为了减少 AI 作图后人工调整的负担，优化 AI 作图的质量，我们很自然地想到的一个策略就是“头痛医头、脚痛医脚”。如果 AI 作图的大部分画面都符合需求而只有一部分区域需要调整的话，我们可以考虑调整局部或重绘画面。反之，我们可以预先将制好的图像提取出合格的部分作为组件补丁，在后续的作图中可以将这些组件快速填充到新作的图中，快速定制图像。对此，网易互娱 ASAK 设计团队将设计行业内的组件化设计迁移到 AI 作图的工作流中 1。团队指出，AI 的优势之一是可以快速生成大量素材，AI 生成的可复用的素材可以统一汇总整理作为团队的组件库。例如，在动画分镜中，人物的面部表情，具有高复用性，基于这种高复用性特征，可以使用 AI 生成全局组件库，方便其他画面复用，避免每次都基于场景去重绘。

（4）构建素材库。由于 AI 设计需要大量的提示词、垫图、组件等，因此设计师可以提前收集、整理相关数据构建定制化的素材库，在正式作图的时候可以从素材库中针对性地提取相应的数据使用。此外，对于有能力训练 AI 模型的设计师而言，收集图像数据还可以帮助重新训练一个专用的生图模型，从而在业务制图的时候直接生成更好的效果。注意这一方法仅适用于支持自定义模型的 AI 作图工具。



1.5 本章小结



在本章中，我们接触了 AI 作图的发展历程；感受了 AI 作图技术在短短十年的时间里飞速发展的惊人速度；了解了不同 AI 作图工具的基本特性，为后续恰当的选择使用做好了基础；见识了多种多样的商业化方法，应当明白 AI 作图在技术和商业化的道路上都处于朝阳阶段；学习了 AI 作图的工作流方法，将其应用在实践中会更充分地发挥 AI 的实力。笔者在为读者介绍的同时，也在不断感受着最新的发展。AI 作图正如一株破土而出的春芽，洋溢着蓬勃发展的活力。更多的内容限于篇幅不能尽数落笔，愿读者在日后的学习工作中保持对这一领域的热情，继续跟进！

经过本章的学习，读者应当已认识到 AI 作图的出现是计算机科学技术发展的必然结果，同时也是艺术领域技术革命的开始。AI 并非洪水猛兽，不会野蛮地剥夺传统从业者的领地。但 AI 注定会是一叶扁舟，拒绝它的人可能会渐渐淹没在时代的洪流中，而善用的人将会乘风破浪，看到艺术设计领域“更上一层楼”的一天！

1 网易互娱 ASAK 设计团队 . 网易大厂出品！ AIGC 组件化设计方法 [R/OL]. 优设网 , 2023 年 5 月 24 日。

028

