

# 项目5 AI 图像生成工具的使用



## 项目描述

AI 图像生成是 AIGC 领域的重要分支，它大幅降低了创意思像创作的门槛。即便没有美术基础或专业软件操作经验，用户也能快速生成专业、生动且富有创意的视觉作品。同时，该技术正日益成为设计师和艺术家进行创意表达的首选工具。

本项目将借助 AI 图像生成工具“豆包”，完成一个真实的设计任务：受河北省保定市古莲花池景区委托，为其设计一系列融合历史文化与自然风貌的创意旅游产品。该项目旨在有效传播古莲花池的景观魅力与文化底蕴，提升游客兴趣，助力当地旅游发展。

通过实际操作，带领读者深入体验 AI 图像生成在精准性、效率及创意辅助方面的显著优势，掌握运用 AI 工具高效产出优质设计作品的能力。



## 项目目标

- 了解 Midjourney、Stable Diffusion、豆包这三款 AI 图像生成工具。
- 理解指令（prompt）设计对图像生成效果的影响机制。
- 掌握使用 Midjourney 和豆包生成图像的方法。

## 技能目标

- 能够熟练运用豆包，进行图像生成、编辑与优化。
- 掌握图像生成过程中的指令编写与优化技巧。
- 能够根据应用场景分析需求，生成用户满意度更高且成系列的作品。

## 素养目标

- 培养创新思维，探索 AI 图像生成技术的新应用，促进艺术与科技的融合。
- 增强技术伦理意识，确保在创作过程中尊重原创、保护数据隐私。

- 激发用户导向思维，鼓励从用户需求出发，将技术创新与用户价值紧密结合，培养以用户为中心的设计思维。



## 项目准备

### 5.1 AI 图像生成工具介绍

AI 图像生成工具是利用人工智能技术，特别是深度学习算法，来自动生成或修改图像的软件。这些工具通过训练大量的图像数据集，学习到了如何根据输入的指令或数据生成逼真的、具有艺术感或实用价值的图像。以下将介绍几种具有代表性的 AI 图像生成工具及其特点。

#### 5.1.1 Midjourney

Midjourney 是于 2022 年面世的一款强大的人工智能图像生成工具，专注于通过自然语言描述生成高度逼真的图像。它使用深度学习技术来解析用户的文本输入，并根据描述生成相应的视觉内容。它是付费软件，其快速且高质量的图像生成服务依赖于租用或维护强大的服务器和云计算资源。用户可以根据自身需求选择不同服务级别的订阅计划，以满足个性化的使用要求。

与其他图像生成工具相比，Midjourney 生成的图像效果极具美感，自然流畅，且易于调整，使其成为图像生成工具领域的佼佼者。Midjourney 具备以下优点。

(1) 高效性：它能够在短时间内生成大量高质量的图像。无论是在商业项目中还是个人创作中，Midjourney 都能够快速交付所需的视觉内容。它的速度与质量并重，确保用户能够在紧迫的时间要求下，依然获得令人满意的图像成果。

(2) 灵活性：它不仅支持多种风格的创作，还能够处理复杂的描述，这使得用户可以根据自己的需求，创作出风格各异、内容丰富的图像。不论是需要简洁的设计风格，还是追求复杂的视觉效果，Midjourney 都能轻松应对。

(3) 操作简便：与许多需要专业技能的软件不同，Midjourney 的用户界面友好，操作流程简单直观，即使是没有专业设计背景的用户也能轻松上手。这种便捷性降低了图像生成的门槛，使得任何人都可以通过简单的操作，创造出专业水准的视觉作品。

图 5-1 是 Midjourney 在 Discord (一款聊天软件) 中的操作界面。Midjourney 主要通过 Discord 来提供其图像生成服务。Discord 是一个流行的聊天平台，最初主要用于游戏社区，但由于其强大的功能和易用性，逐渐吸引了其他类型的社区。Midjourney 团队利用 Discord 的特性，创建了一个方便用户访问和使用的图像生成服务。在使用 Midjourney 生成图像时，用户需要通过 Discord 平台与 Midjourney 的机器人进行交互。首先，用户需要加入 Midjourney 的官方 Discord 服务器。加入后，用户才可以在指定的频道中与 Midjourney 的机器人互动。机器人会通过解析用户输入的生成指令和文本描述来生成图像。生成的图像会直接在 Discord 频道中返回给用户，用户还可以对生成图像做进一步修改或调整。

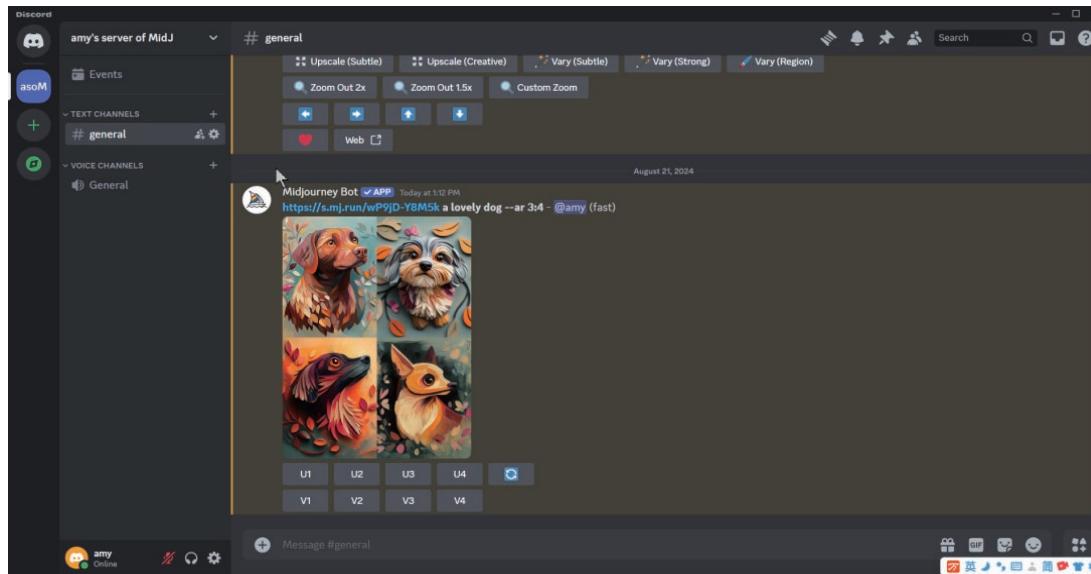


图 5-1 Midjourney 在 Discord 中的操作界面

直到 2023 年 7 月, Midjourney 才在其官网上推出直接生成图像的功能(见图 5-2)。此项改进显著提升了用户访问服务的便捷性,尤其对那些 Discord 平台了解较少的用户而言。相较于 Discord 上的多个频道和命令行操作,其官网的操作界面更加直观与简洁,功能也更加集中和便捷。然而,值得注意的是,新注册用户仍需首先通过 Discord 进行初步的使用学习与基本操作训练,待掌握必要技能后,才能使用官方网站的图像生成功能。

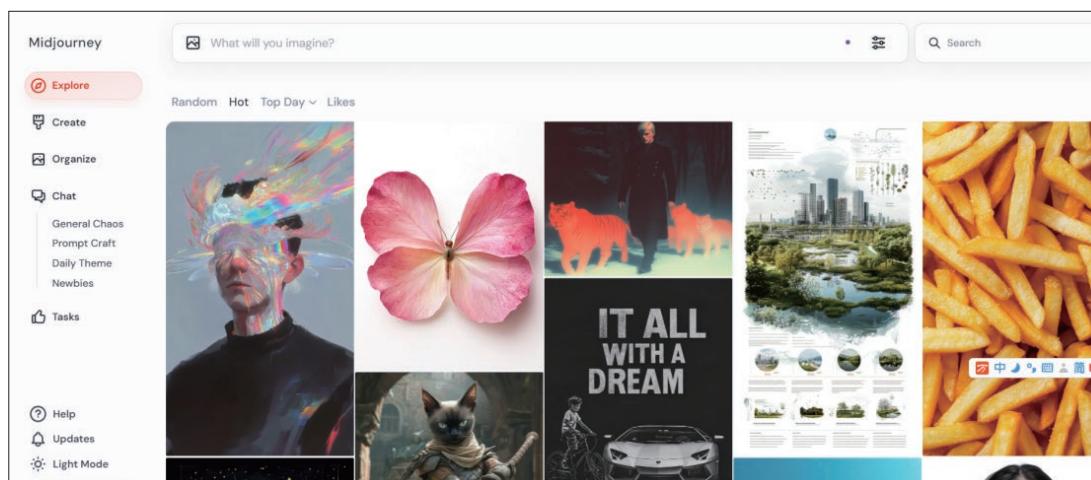


图 5-2 Midjourney 官方网页操作界面

### 5.1.2 Stable Diffusion

Stable Diffusion 是一款开源的图像生成模型,由慕尼黑大学的 CompVis 研究团体开发,并由初创公司 Stability AI、CompVis 与 Runway 合作推出,同时得到了 EleutherAI 和 LAION 的支持。自 2022 年发布以来,Stable Diffusion 因其高效性和广泛的应用潜力而迅速流行,成为 AI 绘画领域的一个热门话题。

与 MidJourney 的不同之处在于, Stable Diffusion 更适合那些需要高度定制化、注重细节和自由控制生成过程的用户, 尤其在科研、商业应用等领域表现优秀。Stable Diffusion 具备以下特性。

(1) 开源: Stable Diffusion 是一个开源项目, 用户可以自由下载、修改和部署。与许多闭源的图像生成工具不同, 它为开发者和研究人员提供了极高的自由度。

(2) 可定制化: 用户可以修改模型的参数、训练自己的数据集, 甚至开发定制版本来满足不同的需求。它适合于开发者和研究人员用来构建新应用或针对特定领域进行优化。

(3) 图像修复技术: Stable Diffusion 能够通过去噪、图像补全等技术来修复或改善图像的质量。例如, 用户可以使用它修复低分辨率、损坏或不完整的图像, 生成高质量的恢复版本。

(4) 用户控制度高: 用户可以通过调整生成过程中的不同参数 (例如扩散步数、噪声强度等) 来更细致地控制生成的图像。这使得用户不仅可以生成多样的图像, 还能精细调整每个细节, 特别适合那些需要精确控制的应用场景。

(5) 计算过程高效: 虽然扩散模型本质上计算量较大, 但 Stable Diffusion 的实现非常高效, 特别是在现代 GPU 上运行时。与其他生成模型相比, 它能在相对较短的时间内生成复杂图像。

(6) 数据更加安全: 由于 Stable Diffusion 可以在本地运行, 它能够避免将数据发送到外部服务器, 从而保护用户隐私。这对那些关注数据安全、需要在内部网络中处理敏感数据的用户特别重要。

(7) 生态化发展: Stable Diffusion 的开源社区非常活跃, 开发者不断发布新工具、插件和优化模型。它的生态系统包括了许多扩展功能, 如: 可以集成到 Photoshop、Blender 等设计工具中。

使用者可以利用绘世启动器在本地部署 Stable Diffusion WebUI (见图 5-3), 也可以通过 Stable Diffusion Online 网站在线使用 (见图 5-4)。本地部署提供了更高的自由度和灵活性, 允许用户根据自己的需求定制和扩展模型, 不过需要用户具备一定的技术能力和计算资源。而在线使用则更加便捷和快速, 无须复杂的安装和配置过程, 通过浏览器访问即可体验 Stable Diffusion 的强大功能。用户可以根据自己的实际情况和需求选择适合的使用方式。

Stable Diffusion WebUI 更适合需要频繁进行图像生成、模型训练或参数调整的专业用户或开发者。而 Stable Diffusion Online 则更适合偶尔使用、不需要复杂配置或计算资源的普通用户。

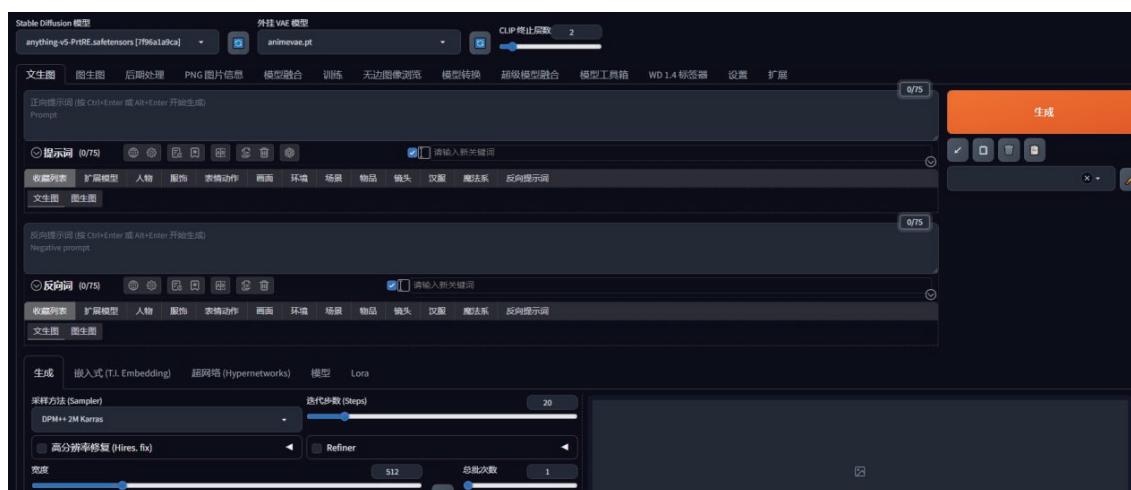


图 5-3 本地部署的 Stable Diffusion WebUI 操作界面



图 5-4 Stable Diffusion Online 网站操作界面

### 5.1.3 豆包

豆包是字节跳动公司于 2024 年 5 月推出的一款集多功能于一体的 AI 工具。当前，该工具已具备文本生成、图像生成及文生音乐等多项功能。除网页版本（见图 5-5）外，豆包还提供了移动 App 版（见图 5-6）。在豆包 App 中，采用了与虚拟人物语音对话式的方式，进行内容的输入和输出，相较于网页版仅支持文字指令输入的方式，此举显著提升了用户体验，使操作更为便捷。



图 5-5 豆包网页版操作界面

尽管豆包在生成效果的精细度方面与当前行业内的顶尖产品相比，仍存在一定差距，但其也具有以下优势。

(1) 免费使用：豆包向所有用户免费提供全部功能，并且不设任何使用次数限制，这一政策极大地提升了用户体验并降低了使用成本。

(2) 支持中文：豆包在功能设计上也充分考虑了中国用户的需求。在豆包的文生图功能中，用

户可以使用中文作为指令或提示词，这与仅支持英文输入的 Midjourney 和 Stable Diffusion 形成了鲜明对比。豆包对中文的全面支持，无疑为中文用户提供了更加友好、便捷的使用体验。

免费使用与对中文的全面支持，共同构成了豆包在中国市场中的重要竞争优势，为其在激烈的市场竞争中脱颖而出奠定了坚实的基础。

除了上述的三款工具，市面上还有很多其他 AI 图像生成工具，如 OpenAI 的 DALL-E 3 和 NightCafe Creator 等。这些图像生成工具在操作方法上大体相同，大多采用收费模式，但同时也提供一定的免费使用次数供用户试用，感兴趣的可自行体验。

## 5.2 使用 Midjourney 生成图像

### 5.2.1 Midjourney 的指令

与文心一言相似，Midjourney 也是基于用户输入的指令来生成与描述相匹配的图像。如图 5-7 所示，为 Midjourney 生成图像的主要指令 “/imagine” 的一个示例。该指令以 “/imagine” 开头，表示动作 “想象”，“prompt” 后跟想象生成的图像信息，包括：参考图像链接、文本描述信息以及参数。



图 5-6 豆包手机 App 操作界面



图 5-7 Midjourney 指令示例

第一部分是参考图像的超链接，也可以称之为图像提示。它表示用户希望生成的图像能与该链接所指向的图像风格相似。需要注意的是，此超链接必须是 Midjourney 的地址。用户可以通过上传本地图像至 Midjourney 来获取对应图像的 Midjourney 地址。

第二部分是文本描述信息，是指令的核心部分，也可称之为文本提示、提示词。目前，该部分仅支持英文输入，不支持中文。因此，在使用时，用户可能需要借助翻译软件将中文描述翻译成英文后再输入。同时，建议输入的文本描述尽量使用简单、简短的短语，避免冗长的请求和说明列表，以提高指令的执行效率。

第三部分是参数，用于更具体地控制图像的生成，或对生成结果进行微调，使其更加符合预期。需要注意的是，一条有效的指令可以仅包含文本提示，或者是包含一个文本提示和一个图像提

示，或者是包含多个图像提示，参数为非必选项。

除了“/imagine”，Midjourney 还有许多其他功能的指令，其中几个最为常用的指令如表 5-1 所示。

表 5-1 Midjourney 常用的指令

指令 / 参数	作 用	示 例
/imagine	生成图像的主要指令，后跟描述性文本	/imagine a beautiful sunset over the mountains
/blend	将多张图片融合生成一张新图	/blend image1.jpg image2.jpg
/describe	根据提供的图像生成描述性文本	/describe uploaded_image.jpg
/settings	显示或更改当前会话的设置	/settings

Midjourney 提供了丰富的参数选项，用于调整图像的生成方式。这些参数可以调整图像的宽高比、风格、质量、随机性等多个方面。“/imagine” 指令常用的参数如表 5-2 所示。

表 5-2 “/imagine” 指令常用的参数

指令 / 参数	作 用	示 例
--ar or --aspect	指定图像的宽高比（宽 : 高）	/imagine a futuristic city --ar 16:9
--v	指定版本，选择 Midjourney 的版本号	/imagine a fantasy landscape --v 5
--q	指定图像质量，范围为 0.25、0.5、1、2。较高质量会花费更多 GPU 时间	/imagine a cute puppy --q 2
--stylize or --s	控制图像的风格化程度，范围为 0 到 1 000。	/imagine a dragon --stylize 700
--seed	指定随机数种子以获得可重复的结果	/imagine a magical forest --seed 12345
--no	排除特定元素	/imagine a car --no wheels
--chaos	增加图像生成的随机性，范围为 0 到 100	/imagine a forest --chaos 50
--tile	创建可无缝拼接的平铺图像	/imagine a floral pattern --tile

## 5.2.2 Midjourney 的应用实例

### 1. 根据文本描述信息生成图像

Midjourney 拥有通过单一简洁的文字描述即能生成图像的能力。举例来说，若想让 Midjourney 创作一幅描绘“一个穿着白色长裙的女孩”的画面，由于 Midjourney 仅支持英文输入，因此需要先利用翻译工具将这句中文描述转换为英文，即“A girl wearing a white long dress”，如图 5-8 所示。

接下来，在 Midjourney 页面下方的输入框中输入“/imagine”，过程中会自动提示匹配的指令，如图 5-9 所示，单击第一条“/imagine prompt”，指令就会在输入框中自动补全。

然后在“/imagine prompt”后面输入之前翻译好的英文描述信息“A girl wearing a white long dress”，如图 5-10 所示。



图 5-8 翻译“一个穿着白色长裙的女孩”成英文

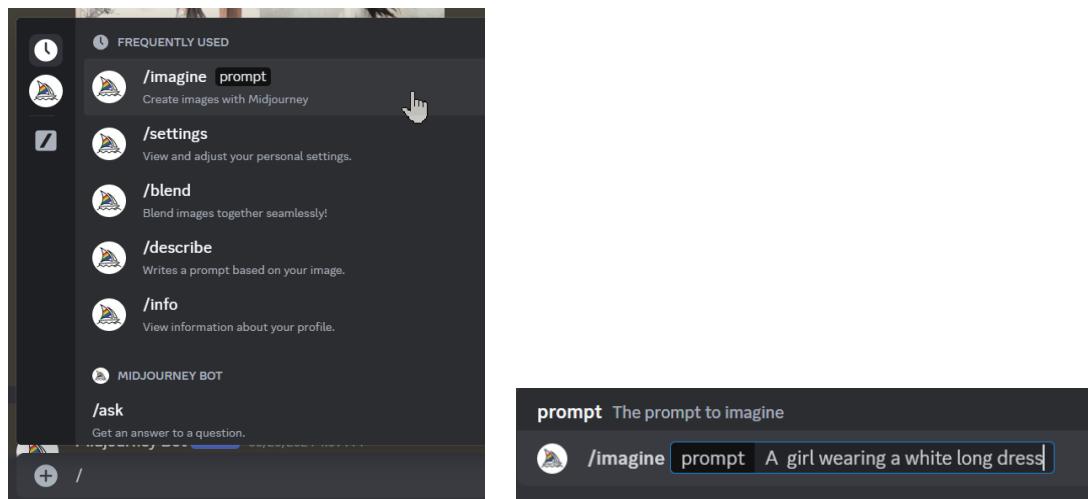


图 5-9 Midjourney 自动提示匹配的指令

图 5-10 根据文本描述信息生成图像指令

按下 Enter 键后, 等待约 1 min, 系统便会生成四张既符合描述又各具特色的图像, 如图 5-11 所示。若在这组图片中, 未发现满意的图片, 用户可单击图片右下方的刷新图标 , 以重新生成一组新的图片。

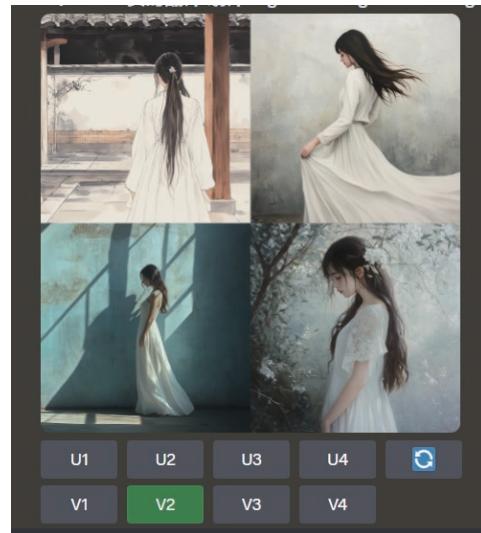


图 5-11 “A girl wearing a white long dress” 生成的图像

## 练一练

请用 Midjourney 的 “/imagine” 指令生成一幅 “穿着格子衬衫的小男孩” 的图像。

## 2. 在生成图像的基础上进行变化

通常一次生成并不能得到完全令人满意的图片，这时就可以通过 Midjourney 的变化功能对图片进行调整。

图 5-11 中，在生成图片的下方，除了刷新按钮，还有另外 8 个按钮，分别是 U1、U2、U3、U4 和 V1、V2、V3、V4。数字后缀指示图像的位置，1、2、3、4 依次对应左上、右上、左下、右下四个方位的图像。此外，U 和 V 两个前缀分别代表了升级和变化两种不同的图像操作功能。

(1) U 功能。

①含义：U 代表 “Upscale”，意为升级。

②作用：单击 U1 ~ U4 按钮，可以对图像进行升级处理。具体来说，它会将图片的大小加倍至 2 0482 048 像素，可显著提升图像的分辨率。同时，该过程还会为图像增添更多的细节，进行更为精细的刻画，使得图像在视觉上更加丰富和细腻。

③使用场景：通常在当前图像非常符合预期，希望获得更高分辨率和更多细节的图像时，会选择使用 U 功能。

(2) V 功能。

①含义：V 代表 “Vary”，意为变化。

②作用：单击 V1 ~ V4 按钮，系统会在保持当前图像风格和主要对象一致的前提下，对图像中的元素进行一定程度的变化。这种变化可能是颜色、纹理、构图等方面的微调，旨在为用户提供更多样化的图像选择。

③使用场景：认为当前图像基本符合预期，但希望在此基础上尝试一些不同的变化，以获得更多灵感或选择时，会选择使用 V 功能。

U 和 V 两个功能分别满足了用户对图像升级和变化的需求，使得 Midjourney 在生成图像时更加灵活和多样化。用户可以根据自己的喜好和需求，选择适合的功能来优化图像效果。

在图 5-11 中生成的四张以 “一个穿着白色长裙的女孩” 为主题的图像中，若发现第 2 张图像（即位于右上方的图像）较为符合预期，并希望在此基础上进一步变化出新的图像，可以直接单击对应的 “V2” 按钮。单击后，Midjourney 会基于第 2 张图像的内容，运用其独特的算法进行随机变化，从而生成全新的一组图像供用户选择。

如果不想要随机地变化，而是希望图像向着自己的预期的方向变化，则需要开启 “Remix Mode”（混合模式）。通过指令 “/settings”（见图 5-12）进入设置界面，单击选中 “Remix Mode” 按钮（见图 5-13）。若按钮显示为绿色，即为混合开启状态。

在启用 “Remix Mode”（混合模式）后，当再次按下 “V2” 按钮时，系统会弹出一个 “Remix Prompt”（混合提示）对话框。在这个对话框中，输入希望图像朝哪个方向变化的文字描述，当

然也必须是英文的。例如，如果希望生成一张展现“白色连衣裙很大，欧式，正面照片”的图像，则应在窗口中输入对应的英文描述：“The white dress is very large, European style, front view photo”，如图 5-14 所示。单击“Submit”（提交）按钮后，即可在原图的基础上，生成一组符合新增描述的图像，如图 5-15 所示。“Remix Mode”模式能够更精确地控制图像的变化，从而满足特定的需求和期望。

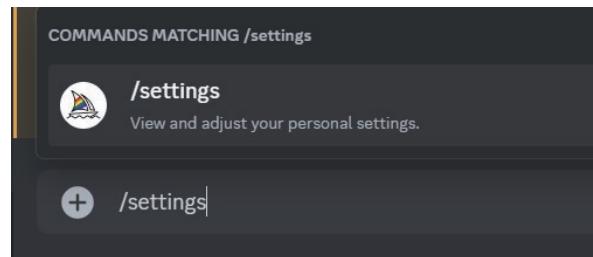


图 5-12 “/settings” 指令

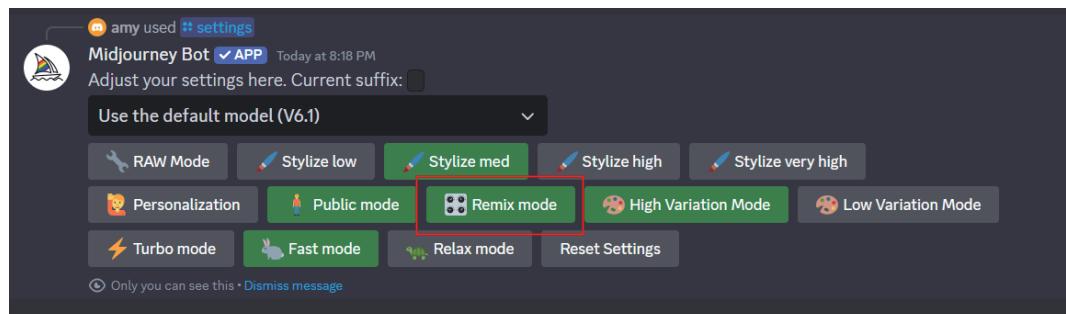


图 5-13 Midjourney 设置界面

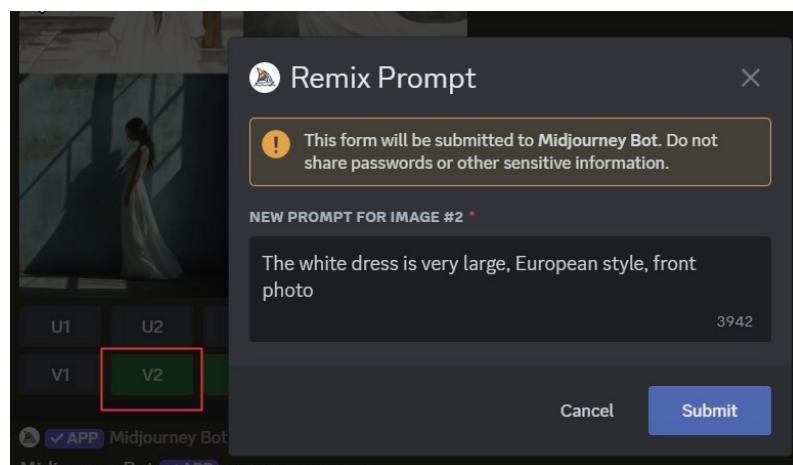


图 5-14 输入“白色连衣裙很大，欧式，正面照片”的英文提示词



图 5-15 根据“白色连衣裙很大，欧式，正面照片”英文提示词生成的一组图像

### 练一练

请你在上一小节“穿着格子衬衫的小男孩”的图像基础上进行变化，让小男孩手中拿着一个飞机模型。

## 5.3 使用豆包生成图像

### 5.3.1 豆包的指令

豆包所采用的指令体系，相较于 Midjourney 那种蕴含多重参数的复杂指令，更为贴近自然语言，并且支持直接使用中文进行描述。如图 5-16 所示，为豆包的一条指令示例。具体而言，一条完整的指令是由描述信息、比例设定、风格选择以及参考图四部分组成。

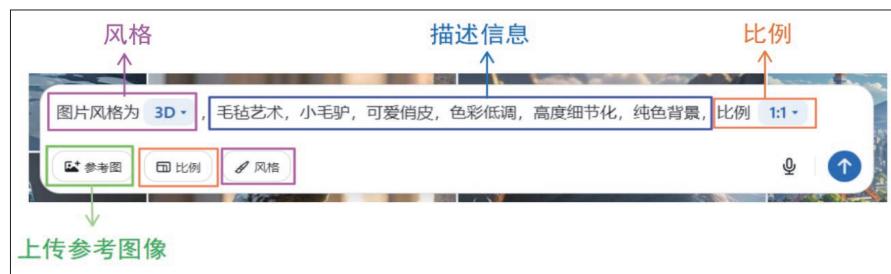


图 5-16 豆包指令示例

指令中最核心的部分是描述信息。无须再借助英文翻译软件，这里直接输入对象的中文描述短语即可。在输入描述信息时，应避免使用冗长的句子，尽量使用更多简洁明了且包含必要元素的短

语。需注意的是，描述信息的详尽程度与生成图像的预期符合度成正比：信息越丰富，图像越贴近预期；反之，信息越少，图像的随机性和不确定性则相应增大。

比例、风格和参考图均是可选项。借助输入框下方的功能按钮，可以进行快捷设定。如图 5-17 和图 5-18 所示，这里默认提供了几种常见的比例和风格选项。选定比例或风格后，系统会将其自动补充到指令中。此外，也可以不通过按钮，而在指令的描述信息中直接输入所期望的风格和比例。若期望所生成的图像能与另一幅特定图片在风格上保持相似，可利用“参考图”按钮的功能，上传被用作参考的图片。

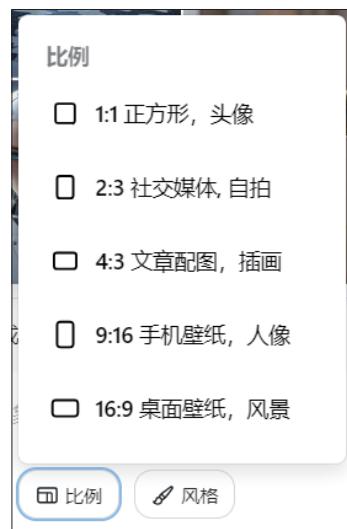


图 5-17 比例选项



图 5-18 风格选项

与 Midjourney 和 Stable Diffusion 相比，豆包的用户界面设计显得更为简洁明了，并且在操作便捷性方面展现出了一定优势。其指令系统在全面支持中文的基础上，指令本身亦被大大简化。此外，豆包所生成的图像效果正日益接近 Midjourney 的水平，因此，它是快速掌握图像生成技术的优选工具。

### 5.3.2 豆包的应用实例

#### 1. 根据文本描述信息生成图像

随着科技的飞速发展，未来机器人帮助人类完成各项工作已经逐渐成为可能。例如，这里让豆包创作一幅“一个送快递的机器人”的图像。

与 Midjourney 类似，豆包也是通过一段简单的文字描述就能生成图像，而且豆包支持中文，使用起来更加便捷。在豆包主界面的左侧菜单中，单击“图像生成”选项，或者单击输入框上方的“图像生成”按钮（见图 5-19），即可进入图像生成主界面。

在下方的输入框中直接输入中文描述“一个送快递的机器人”，等待不到 1 min 的时间，系统即生成了四幅图像，如图 5-20 所示。生成的每张图都包含了机器人和快递的元素，且呈现出较高的图像质量，展现出强烈的科技感与逼真的视觉效果。经观察，位于图像右侧中间位置的一幅图较为贴近预期设定，故单击选中此图，进入大图浏览模式以进行细致查看与后续操作。



图 5-19 豆包图像生成入口



图 5-20 根据“一个送快递的机器人”的生成图像

### 练一练

请你使用豆包生成三只可爱的小猫图像，风格为 3D 渲染。

## 2. 区域重绘

豆包区域重绘功能类似于 Midjourney 的“Vary(region)”图像局部变化功能，可以在当前已经生成的图像中，进行局部的变化。假如，这里希望机器人背后的快递箱更多些，可以单击大图上方的“区域重绘”按钮（见图 5-21），进入图像重绘界面，用画笔选中期望变更的区域，并在弹出的提示框内输入新的指令“增加更多的快递箱”，如图 5-22 所示。回车后，系统生成了一幅新图像，如图 5-23 所示。可以明显地观察到，快递箱的数量从原先的四个增加至九个，而未被选中的区域仍然保持原样，未发生任何变化。

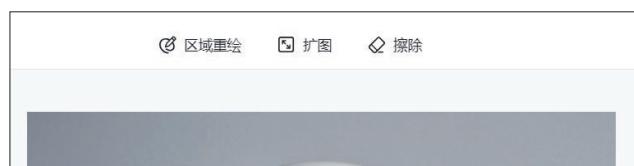


图 5-21 大图浏览模式的按钮

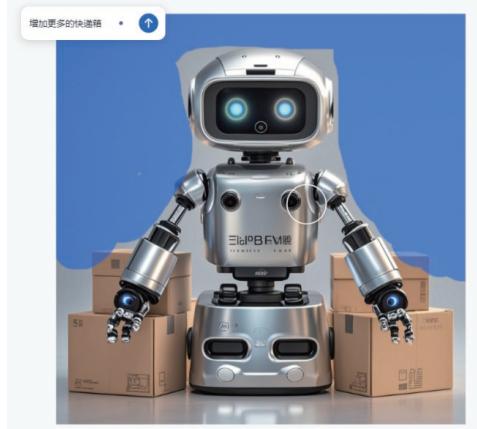


图 5-22 区域重绘界面



图 5-23 区域重绘生成的图像

### 练一练

请你将上节“练一练”生成的“三只可爱的小猫”图像中的一只小猫，通过区域重绘功能，重绘成熊猫。

### 3. 图像擦除

对于生成图像中不满意或认为多余的部分，可利用图像擦除功能进行去除。该擦除功能实质上可视为一种特殊的区域重绘，其工作原理是系统自动对所选的擦除区域进行重新绘制，以使该区域与周围背景保持一致，从而达到擦除的目的。

上一步重绘的图像 5-23 中，框线标出的区域快递箱呈现悬浮状态，这显然违背了物理常识。此时，可以采用擦除工具对相应部分进行处理。

单击图像上方的“擦除”按钮，进入图像擦除界面。用画笔选中期望擦除的区域（见图 5-24），并单击图像下方的“擦除所选区域”按钮，稍作等待，系统将生成一幅新的图像，如图 5-25 所示。在新生成的图像中，原先悬浮的快递箱部分已被合理擦除，并由其他快递箱填补，使得机器人背景的设置更符合现实状态。



图 5-24 选择需要擦除的区域



图 5-25 擦除后的效果

## 练一练

请你将上一节中“三只可爱的小猫”图像中的一只小猫，通过图像擦除功能，擦除其中一只小猫身上的部分条纹或其他特征。

## 4. 扩图

当生成的图像较为满意，但同时又感觉图像的篇幅过小，希望在原图的基础上向四周扩展，以获取远视图时，可以使用豆包的“扩图”功能实现。

举例来说，若需创作一幅描绘中式古镇、小桥流水景象的图像，可在系统下方的输入框输入“古镇，小桥流水”的描述，并选定图片风格为“水彩画”，比例为“1:1”，如图 5-26 所示。



图 5-26 输入“古镇，小桥流水”等指令

按 Enter 键后，生成的图像如图 5-27 所示，该图像贴切地反映了指令所蕴含的意境。画面中，错落有致的黑瓦房构成了背景，而一座典型的拱桥则成了视觉焦点，其在水中的倒影巧妙地与拱桥本身形成了一个圆形构图。不过，美中不足的是，此图像仅呈现了古镇的一隅。



图 5-27 古镇水彩画生成图

这时，选中图 5-27 所示的图像，进入大图模式，单击图 5-21 中的“扩图”按钮，进入扩图操作界面。在此界面中，通过图像四角附带的黑色调整框，可以灵活地调整预期扩展后的图像比例。同时，也提供了快速调整比例的选项，即图片上方显示的常用比例按钮。此外，图像本身也是可移动的，用户可根据需要将图像拖动至合适的位置，如图 5-28 所示。单击“按新尺寸生成图片”按钮，即可生成新的图像，如图 5-29 所示。

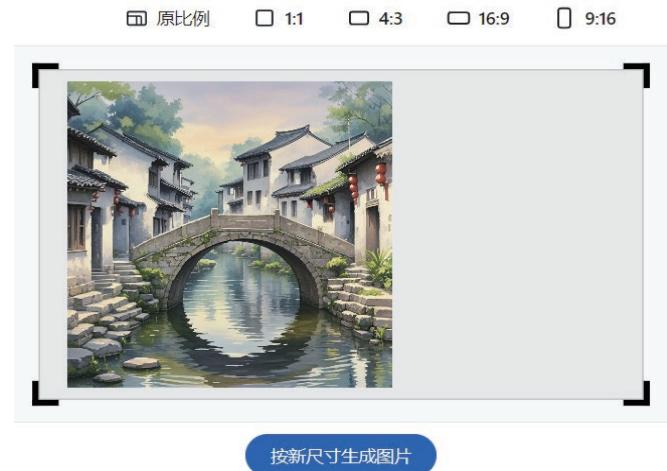


图 5-28 扩图操作界面



图 5-29 扩图后的效果图

在新生成的图像中，可以清晰地观察到：在原图的右侧区域，成功扩展出了一条小船以及一座黑瓦房，有效丰富了画面的内容与视野。

### 练一练

请你先用豆包生成一幅油画风格的山水画。然后在此图的基础上，进行扩图，扩大为原来图像的约 2 倍大小，查看扩图后的效果。

## 5. 参考其他图像的风格生成新图像

在输入框的下方，第一个按钮就是参考图，说明豆包也具有类似 Midjourney 的参考已有图像生成类似风格图像的功能。需要注意的是，豆包当前仅支持上传一张参考图像。

以图 5-30 所示的剪纸风格的小鸟图像作为参考图，单击输入框下方的“参考图”按钮，浏览选择存放在本地磁盘上的小鸟图片，随后单击“确认”按钮。小鸟图像的缩略图就出现在输入框上方，如图 5-31 所示。本例中，目标是生成具有类似风格的兔子图像。因此，在缩略图下方输入文字“小兔子”，然后按 Enter 键确认。



图 5-30 剪纸风格的小鸟



小兔子

图 5-31 上传参考图

系统会生成图 5-32 所示的兔子图像。通过观察可以发现，所生成的兔子的图像在色彩搭配及画面细节丰富度上，与上传的小鸟参考图保持一致。

为了对比参考图的影响，进行另一次尝试：此次不上传任何参考图，仅输入“小兔子”作为文本描述，所生成的图像如图 5-33 所示。该图像呈现为写实风格，以白色和灰色为主调，生动展现了小兔子的形象，与有参考图时生成的效果存在明显差异。



图 5-32 有参考图生成的图像



图 5-33 没有参考图生成的图像

图 5-32 和图 5-33 这两张图说明了豆包参考图功能确实能够影响生成的图像风格和效果。然而，同时也观察到豆包生成的效果与预期存在一定的偏差。尽管图 5-32 在配色上与参考图一致，但生成的兔子图像为卡通风格，而非预期的剪纸风格。

为减少此类偏差，可通过增加描述性提示词来实现。例如，在上传参考图后，将文本描述调整为“剪纸风格的小兔子”，而非简单的“小兔子”。通过在提示词中明确“剪纸风格”，可进一步引导图像生成的方向，可以避免因豆包识别不准确而导致的偏差。采用此方法后，生成的图像如图 5-34 所示，在风格和配色上都更贴近参考图。

这里再做一个对比：不上传任何参考图，仅输入“剪纸风格的小兔子”，生成的效果如图 5-35 所示，虽然确实呈现出剪纸风格，但在配色上与先前有参考图时存在明显差异。

综上所述，豆包的参考图功能确实能对图像生成的效果产生显著影响。在某些情况下，为确保获得更佳的预期效果，建议配合上精确的描述性提示词使用。



图 5-34 有参考图“剪纸风格的小兔子”生成的图像



图 5-35 没有参考图“剪纸风格的小兔子”生成的图像

### 练一练

请你参考本节“剪纸风格的小鸟”图像，生同一风格的狐狸图像。

## 项目分析

河北省保定市古莲花池，初名雪香园。元代时，由张柔创建，并引水入城，形成了最初的园林景观。此后，古莲花池历经明清两代的多次整修和扩建，形成了集园林、行宫、书院为一体的古典园林。清代乾隆年间，乾隆、嘉庆、光绪三朝帝后均来此驻跸巡幸，进一步提升了古莲花池的地位和影响力。

时至今日，古莲花池不仅承载着丰富的历史文化底蕴，还以其独特的自然景观吸引着众多游客。假如，为深入挖掘与传承这一宝贵的历史文化资源，同时进一步提升古莲花池的知名度与吸引力，促进当地旅游业的持续繁荣与发展，景区管理局诚邀作为设计师的你为其设计一系列兼具创意与文化内涵的旅游产品，共同推动这一历史名胜焕发新的生机与活力。

### 思考 →

如何开始一款文创产品的设计？

你计划使用豆包这款 AI 图像生成工具完成此项目。在着手开展设计工作之前，首要任务是进行详尽的项目分析，包括明确设计思路、产品类型、设计风格和产品系列等关键要素。

### 1. 设计思路

以古莲花池的历史文化、自然景观为灵感来源，结合现代审美和流行趋势，利用 AI 图像生成工具创作出既传统又时尚的文创设计。

### 2. 产品类型

明信片文创产品作为一种融合了传统与现代、文化与创意的商品，近年来在市场上受到了广泛的关注和喜爱。此类产品可以将传统文化、地域特色、艺术风格等融入设计之中，通过精美的图案、文字，传递出丰富的文化内涵和情感价值。因此，本次项目选定明信片作为设计对象，无疑是一个

极具洞察力且恰当的选择。

### 3. 设计风格

为了充分展现古莲花池深厚的历史文化底蕴，以及莲花所蕴含的高洁古朴的气质，明信片的设计选择水墨画与工笔重彩画这两种艺术形式，是最贴切不过了。

水墨画以其简约而深远的意境，巧妙地勾勒出古莲花池的清幽雅致，而工笔重彩画则以其细腻入微的笔触和丰富华丽的色彩，生动再现了莲花的绰约风姿与池畔景致的细腻韵味。这两种风格，不仅准确捕捉了古莲花池的独特魅力，也向观者传递了其背后深厚的文化意蕴与审美价值，使明信片成为承载与传播这一文化遗产的精美载体。

### 4. 产品系列

古莲花池以其每年夏季满池盛放的荷花而著称。鉴于此，拟设计两组明信片，每组包含四枚。其中一组将运用水墨画的技法，着重表现荷花的高洁雅致；另一组则采用工笔重彩的艺术手法，旨在展现园林四季的精致美景。

以上就完成了对整个项目的分析，接下来正式进入文创设计生成环节。

## 项目实施

### “古莲花池”文创设计生成

#### 1. 古莲花池之莲花系列明信片设计的生成

本系列明信片文化创意产品，拟采用水墨画的风格作为设计核心，具体涵盖四幅作品，每幅作品描绘不同形态和寓意的莲花图像。

选择水墨画作为主要风格，是因为它既是中国传统文化的代表与象征，承载民族自豪感，又善于以笔墨表现山川草木之美，体现自然意境。水墨画追求“以形写神”，以简洁线条和墨色层次营造深远意境，能恰到好处地传达莲花的精神内涵。

莲花又称荷花、芙蕖等，因古莲花池以莲花闻名，且其在传统文化中象征高尚气质，极具代表性，因此被确立为主题最为合适。

根据中国古诗词中对莲花的描写，以及莲花在中国文化中承载的吉祥象征意义，为这四幅明信片设定的标题依次为：《出淤泥而不染，濯清涟而不妖》《接天莲叶无穷碧，映日荷花别样红》《鱼戏莲叶间》及《莲生童子》。

##### 1) 出淤泥而不染，濯清涟而不妖

“出淤泥而不染，濯清涟而不妖。”这句诗出自北宋哲学家周敦颐的《爱莲说》。含义是：莲花从淤泥中长出却不被污染，经过清水的洗涤却不妖艳。此句描绘了莲花的气度、风节，赞美了莲花美丽而不显妖媚、挺拔清丽的芳姿，寄托了诗人高洁正直、超凡脱俗的人格理想。

根据上述对诗句的理解，可以在头脑中构思出这样一幅画面：“莲花为焦点，近景，莲花高洁挺拔清丽，荷叶，有大量的留白，墨色层次丰富，超凡脱俗。”

将上面的描述输入豆包的图像生成指令输入框中，并将风格设为“水墨画”，比例设为“2:3”，如图 5-36 所示。

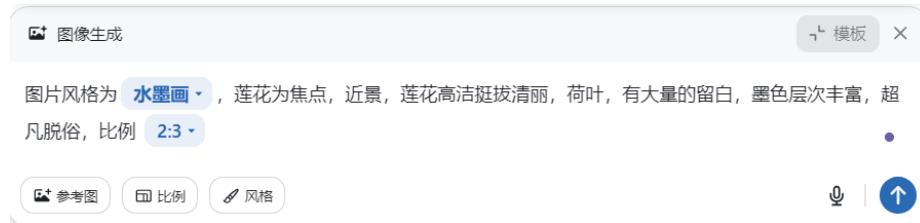


图 5-36 荷花系列“出淤泥而不染，濯清涟而不妖”图的生成指令

提交指令后，豆包生成了一组图像，如图 5-37 所示。经观察，所生成的莲花图像均高度契合指令中的描述要求。其中，位于右上方的图像（见图 5-38）质量更好，其构图最为得当，且有效地凸显了莲花高洁的气质特征。基于上述分析，该图像被选定为莲花系列明信片的首张图。



图 5-37 生成的莲花组图



图 5-38 右上莲花 图放大图

图像生成后，提交后的指令在系统中显示为：“帮我生成图片：图片风格为「水墨画」，莲花为焦点，近景，莲花高洁挺拔清丽，荷叶，有大量的留白，墨色层次丰富，超凡脱俗，比例「2:3」”，如图 5-39 所示。观察发现，刚刚通过按钮选择的风格和比例均用“「」”特殊符号括了起来。由于豆包在指令输入框可供选择的比例极为有限，若直接输入“比例为 3:2”，这样的描述无法被系统识别出来，而该特殊符号又无法直接通过键盘输入。因此，当用户需要指定特定比例时，建议采取复制现有格式“比例「2:3」”，并仅修改其中数字的方式，如改成“比例「3:2」”，以确保指令的准确执行。

帮我生成图片：图片风格为「水墨画」，莲花为焦点，近景，  
莲花高洁挺拔清丽，荷叶，有大量的留白，墨色层次丰富，超  
凡脱俗，比例「2:3」

图 5-39 提交后的指令

2) 接天莲叶无穷碧，映日荷花别样红

“接天莲叶无穷碧，映日荷花别样红。”这句诗出自宋代诗人杨万里的《晓出净慈寺送林子方》。

含义是：荷叶密密层层地铺展开去，无边无际，仿佛与天空相接。阳光下的荷花显得格外鲜艳。

根据上述对诗句的理解，可以在头脑中构思出这样一幅画面：“盛夏，碧绿茂盛的荷叶，延伸到远处，几株荷花点缀其中，荷花格外娇艳，远景，墨色层次丰富”。

将上面的描述输入到豆包的图像生成指令输入框中，同样将风格设为“水墨画”，比例设为“2:3”。

指令提交后，生成了一组图像，如图 5-40 所示。经观察，与前一组图不同的是，这次生成的图像中莲叶数量明显更多了，且几乎铺满了整个画面，符合“接天莲叶无穷碧”的情境。其中，位于左边的大图（见图 5-41）最为贴近水墨画的效果，故该图像被选定为本系列明信片的第二张。至于其余三幅图像，它们呈现出较为明显的写实风格特征，不符合要求。



图 5-40 生成的“接天莲叶”组图



图 5-41 左侧“接天莲叶”图放大图

### 3) 鱼戏莲叶间

“鱼戏莲叶间”这句诗出自乐府诗《江南》，全诗为：“江南可采莲，莲叶何田田。鱼戏莲叶间。鱼戏莲叶东，鱼戏莲叶西，鱼戏莲叶南，鱼戏莲叶北。”整首诗描绘了一幅生动的画面：莲叶层层叠叠，繁茂葱翠，鱼儿在莲叶间自由自在地穿梭、嬉戏。这样的画面不仅给人以美的享受，也让人感受到诗人内心的宁静与恬淡。

根据上述对诗句的理解，可以在头脑中构思出这样一幅画面：“盛夏，中式池塘中，有荷叶和几株荷花，荷叶下有彩色的锦鲤游来游去，近景，墨色层次丰富”。

将上面的描述输入到豆包的图像生成指令输入框中，同样将风格设为“水墨画”，比例设为“2:3”。

指令提交后，生成了一组图像，如图 5-42 所示。经观察，除了莲花和莲叶，图像中还展现了锦鲤在水中悠游的场景。其中，位于左边的大图（见图 5-43）的锦鲤最为灵动，且锦鲤的位置巧妙地构成了视觉的焦点，从而使得整个画面布局显得最为和谐。故该图像被选定为本系列明信片的第三张。至于其余三幅图像，荷花与锦鲤的分布相对分散，导致视觉中心不够明确，给人以略显杂乱的印象。



图 5-42 生成的“鱼戏莲叶间”组图



图 5-43 左大“鱼戏莲叶间”图放大图

#### 4) 执莲童子

执莲童子的寓意是“连生贵子”。在宋代，莲花与童子组合，从字面上看“莲”与“连”同音，因此莲花童子就被赋予了“连生贵子”的美好寓意。此外，执莲童子也常被看作是一种吉祥的象征。宋代人认为执莲童子能给人带来欢乐和吉祥，因此在各种生活器物上广泛采用这一形象，寄托了对美好生活的向往和追求。

根据上述的分析与理解，可以在头脑中构思出这样一幅画面：“童子坐在巨大的荷叶上，手持一枝莲花，神情自若”。

将上面的描述输入到豆包的图像生成指令输入框中，同样将风格设为“水墨画”，比例设为“2:3”。

提交指令后，豆包生成了一组图像，如图 5-44 所示。经观察，位于右下方的图像（见图 5-45）中，童子手持莲花，莲花花蕾饱满，花径弯曲的角度自然，且孩童的表情也最为讨喜，故将该图像被选定为本系列明信片的第四张。至于其余三幅图像的童子形象，有的未持莲花，有的所持莲花不够饱满。



图 5-44 生成的“执莲童子”组图



图 5-45 右下“执莲童子”图放大图

通过将上述精心挑选的四幅图像组合起来,如图 5-46 所示,最终完成了古莲花池之莲花系列明信片的设计。整套设计莲主题突出,统一采用 2:3 的比例和水墨画的风格,寓意美好,而且在视觉上和谐与统一,既展示了中国传统艺术的独特魅力,又体现了中国文化的广博与精深。



图 5-46 古莲花池之荷花系列明信片设计

## 2. 古莲花池之园林系列明信片设计的生成

本系列明信片文化创意产品,拟采用精细入微的工笔重彩风格进行设计,共计包含四幅展现不同园林景观的图像。

工笔重彩画作为中国传统绘画瑰宝,注重工整线条与层层渲染,色彩厚重沉着,兼具强烈装饰性与艺术意境,能细腻展现多样题材。本项目与人工智能结合,为现代艺术注入新的活力。

为了让明信片这一载体能够传递出更多的历史文化内涵,计划将古莲花池的历史元素巧妙融入设计之中,使得每张明信片都能独立叙述一个故事,分别代表古莲花池的一段历史。同时,为了全面展现园林随季节更迭的景致变化,本系列特设定每张明信片对应一个季节。最终呈现的效果是,每张明信片不仅是一段历史的缩影,一个季节的象征,还寄托了一种特定的情感。

为这四幅明信片设定的标题依次为:《春之书声琅琅》《夏之乾隆巡幸》《秋之叶落焚尽》及《冬之花旦映雪》。

### 1) 春之书声琅琅

古莲花池在其早期历史阶段,曾作为一处书院存在。雍正十一年(1733年)直隶总督李卫奉诏而建。莲池书院是清代直隶最高学府,名师荟萃,人才辈出。第一张明信片想要再现的是古莲花池早期为书院时,莘莘学子在莲池书院读书时的场景:琅琅的书声从学堂中传出,回荡在春意盎然的湖面上,周围中式的园林为学子们营造了一个宁静且雅致的学习环境。整个画面以春天为背景,体现欣欣向荣、生机勃勃生机之气。

根据上述的分析与理解,可以在头脑中构思出这样一幅画面:“图片风格为「工笔重彩」,近处有古典中式学堂,学堂中有清代的儿童在读书,学堂外有池塘,池塘中有新生的荷叶和荷花的花蕾,池边有抽芽的柳树,安静祥和,比例「3:2」。”

将上面的描述输入到豆包的图像生成指令输入框中,将风格和比例用“『』”特殊符号括了起来。这个特殊符号建议直接从之前提交过的指令中复制过来,如图 5-47 所示。

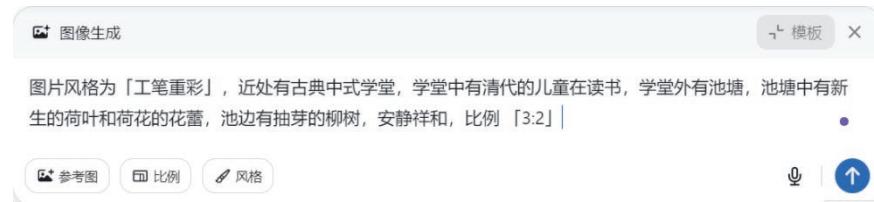


图 5-47 园林系列“春之书声琅琅”图的生成指令

指令提交后，生成了一组图像，如图 5-48 所示。经观察，位于左边的大图（见图 5-49）中的孩童表情最为自然，构成了画面的视觉焦点。同时，该图中的屋檐、柳树以及荷花元素，恰到好处地分布于画面的三个边缘位置，使得整体布局显得更为合理且和谐。故将此图像选定为本系列明信片的首张图。至于其余三幅图像，有的孩童表情模糊，有的孩童过于分散，有的则布局不和谐，均未入选。



图 5-48 古莲花池之园林系列明信片设计生成的“春之书声琅琅”组图



图 5-49 左大“春之书声琅琅”图放大图

## 2) 夏之乾隆巡幸

乾隆十年（1745 年），为了满足皇帝西巡驻跸的需要，位于古莲花池园内的保定使馆被改建为行

宫，时称莲池行宫。当时正值康乾盛世之际，是中国古代封建王朝的最后一个盛世，也是中国历史上最辉煌的时期之一。社会稳定，经济繁荣，文化科技方面也有很大的发展。其间乾隆皇帝曾七次驾幸保定均来莲池行宫驻跸或游赏。第二张明信片描绘的是盛夏时节，乾隆皇帝巡幸至莲花池行宫，置身于繁茂的荷叶与艳丽的荷花之中，尽情欣赏这满池的盛夏繁华。整个场景以盛夏为背景，不仅展现了乾隆皇帝巡游时的闲适与愉悦，更从侧面反映了康乾盛世时期，中国在经济、文化及社会生活等多个方面所达到的繁荣与昌盛。

根据上述的分析与理解，可以在头脑中构思出这样一幅画面：“图片风格为「工笔重彩」，远处有白色的单拱桥和红色的中式二层亭子，拱桥在左侧，亭子在右侧，乾隆皇帝站在拱桥上赏荷，满池塘的荷叶和荷花从远处延伸到近处，荷花盛开，盛夏，远景，比例「3:2」。”

将上面的描述输入豆包的图像生成指令输入框，提交后，豆包即可生成一组图像，如图 5-50 所示。经观察，图 5-50 右上图（见图 5-51）中的乾隆皇帝是站在地面上的，且该图中的园景与古莲花池的实际园景最为接近。故将这张图选为本系列的第二张明信片。至于其余三幅图像，乾隆皇帝都是站在荷叶上面的，这不符合逻辑，因此未予采用。



图 5-50 古莲花池之园林系列明信片设计生成的“夏之乾隆巡幸”组图



图 5-51 “夏之乾隆巡幸”图放大图

## 3) 秋之叶落焚尽

光绪二十六年（1900年），英、法、德、意四国联军侵入保定城，四处烧杀淫掠。古莲花池的建筑被焚烧殆尽，珍宝、文物被洗劫一空。第三张明信片重现的是古莲花池在遭受洗劫之后的景象：昔日乾隆年间享誉一时的“莲池十二景”在火灾中悉数损毁，仅剩残垣断壁，在秋风中萧瑟，一片凄凉景象。画面以秋天为背景，寓意着清朝正步入衰败之境。八国联军侵华战争是中国近代史上一段屈辱的历史，它深刻揭示了弱国无外交、落后就要挨打的道理。这段历史提醒我们要铭记过去、珍惜和平、奋发向前。

根据上述的分析与理解，可以在头脑中构思出这样一幅画面：“图片风格为「工笔重彩」，近处有古典中式学堂，学堂火光冲天，正在熊熊燃烧，远处有池塘，池塘中有枯萎发黑的荷花、荷叶，池塘边上有被烧黑的、光秃秃的树干，秋天，黄昏，枯黄干朽的树叶纷纷飞落，破败，凋零，比例「3:2」。”

将上面的描述输入到豆包的图像生成指令输入框中，提交后，生成了一组图像，如图 5-52 所示。经观察，右上图（见图 5-53）的亮度最低，最具凋零破败之感，且此图中的学堂与古莲花池实际的建筑风格最为接近。故这张图像被选定为本系列的第三张明信片。其余三幅图像的建筑与古莲花池实际的建筑差异太大，不适合。



图 5-52 古莲花池之园林系列明信片设计生成的“秋之叶落焚尽”组图



图 5-53 “秋之叶落焚尽” 图放大图

## 4) 冬之花旦映雪

进入21世纪之后，古莲花池复原了历史上著名的“莲池行宫十二景”，凭借其优雅的风貌与丰富的文化底蕴，吸引了来自各地的游客。第四张明信片所展现的是，经过全面修复后的古莲花池，焕然一新，名园胜景重现。在一场瑞雪之后，园区显得晶莹剔透，静谧而美丽。在中国文化中，瑞雪常被视作丰收与吉祥的象征。画面中还有一位身着华丽服饰的花旦，矗立于雪景之中，格外引人注目，为这宁静的美景增添了一抹亮色。整幅画面不仅反映了新中国成立后人民生活的安定与和谐，还展现了包括中国京剧在内的传统文化的复兴。这张图不仅向世界展示了中国园林的自然之美，也彰显了国家的强盛与中华民族深厚文化的自信感。

根据上述的分析与理解，可以在头脑中构思出这样一幅画面：“图片风格为「工笔重彩」，冬天，正在下雪，雪花漫天飞舞，远处有白色的单拱桥和红色的中式二层亭子，拱桥在左侧，亭子在右侧，京剧花旦在拱桥上翩翩起舞，池塘中的水结冰了，拱桥、亭子、冰面上都有雪，远景，高细节、高精度，比例「3:2」。”

将上面的描述输入到豆包的图像生成指令输入框中，提交后，生成了一组图像，如图5-54所示。经观察，右上图（见图5-55）的花旦扮相唯美，神态自然，气质突出，且与背景最为和谐。故这张图像被选定为本系列的第四张明信片。其余三幅图像花旦的神态和动作均不够自然，故未能入选。



图5-54 古莲花池之园林系列明信片设计生成的“冬之花旦映雪”组图



图 5-55 “冬之花旦映雪” 图放大图

通过将上述精心挑选的四幅图像组合起来,如图 5-56 所示,最终完成了古莲花池之园林系列明信片的设计。这一系列的明信片以春、夏、秋、冬四季为线索,统一采用工笔重彩的风格和 3:2 的比例,按照古莲花池历史事件的发展顺序徐徐展开。通过运用现代前沿的 AI 图像生成技术,明信片生动再现了古莲花池所经历的书院时期、行宫岁月、战争洗礼及复原重生四大重要历史阶段,实现了科技、历史、美学、爱国主义精神与民族自信心的和谐统一与精彩展现。

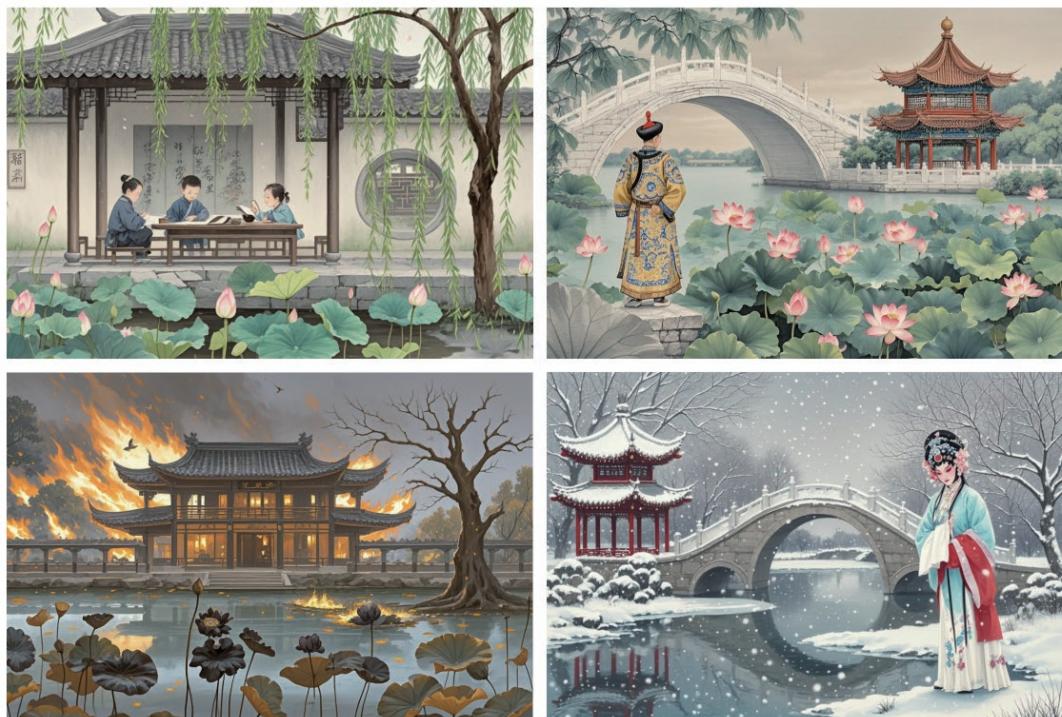


图 5-56 古莲花池之园林系列明信片设计

至此,利用豆包 AI 图像生成技术,成功地完成了以古莲花池的历史底蕴与自然景观为灵感来源,既秉承了传统元素,又彰显出时尚气息的文创设计作品。



## 项目总结

本项目围绕利用 AI 图像生成工具豆包生成保定“古莲花池”文创设计展开。

在项目准备阶段，介绍了三款主流 AI 图像生成工具：Midjourney、Stable Diffusion 和豆包，帮助学习者全面了解并选择适合自己的工具。随后，深入剖析了 Midjourney 与豆包的操作指令及实际应用案例，确保学习者能迅速掌握这些工具的核心使用方法。

在项目实施阶段，紧密围绕保定“古莲花池”文创设计主题，先期进行了详尽的项目分析，明确了文创产品的设计思路、产品类型、设计风格和产品系列。在此基础上，逐一利用豆包的图像生成功能，精心创作并筛选出了两个系列共计 8 张图像，作为最终的明信片设计成果。整个实践过程充分展示了 AI 图像生成工具在实际应用中所具备的卓越创造力和便捷性。

通过本项目的系统学习与实践，学习者不仅掌握了 AI 图像生成工具的基本使用方法，还具备了独立完成实际图像生成任务的能力。这不仅显著提升了学习者的图像创作能力和艺术修养，还为其未来的职业生涯拓展了更多的可能性，提供了更为广阔的发展空间。



## 项目拓展

假如你是一名设计师，近期接到了来自你家乡文化和旅游管理部门的委托，诚邀你为家乡设计一些富有创意的文化旅游产品。请你用豆包这一图像生成工具生成相应的产品设计。

思考你的家乡蕴藏的历史故事、名胜古迹，自然风光、美食特产以及民俗戏剧等。可以从中精选一项或多项作为设计灵感来源，精心打造一款或一系列文化创意产品。旨在让造访您家乡的游客，通过您的设计作品，深刻感受到当地的历史文化底蕴，铭记你家乡的独特风貌，并欣然将这些文化创意产品作为珍贵的纪念品带回家。