

1.1 Stable Diffusion 简介

Stable Diffusion 于 2022 年 8 月 22 日由 Stability AI 正式发布，它基于慕尼黑路德维希马克西米利安大学 CompVis 小组的研究成果，结合了 Transformer 架构与扩散模型，是一种基于深度学习技术的模型，能够依据文本描述生成详细图像，即人们常说的“文生图”。当用户在文本框中输入一些文本提示词时，该模型能够生成与文本内容相匹配的图像作为输出。例如，打开 Stable Diffusion，在“提示词”文本框中仅输入“a dog, (一只狗,)”这样简单的提示词，单击“生成”按钮后，软件即可生成图 1-1 所示的小狗图像。



图 1-1 基本的文生图功能示例

Stable Diffusion 自推出至今短短两三年内，已先后发布了 13 个版本。目前最新的版本是 2024 年 10 月 22 日发布的 Stable Diffusion 3.5，该版本在图像质量、字体处理、复杂提示理解及资源利用效率等方面均有显著提升，还具有可定制性强、支持多样化输出和多种风格等特点，并且在 Stability AI 社区许可证下可免费用于商业和非商业用途。

Stable Diffusion 3.5 包括三个主要版本：Large 版本、Large Turbo 版本和 Medium 版本。

(1) Large 版本：参数量为 80 亿个，支持生成 100 万像素及以上的高清图像，适合专业设计师或有高级绘画需求的用户。

(2) Large Turbo 版本：参数量与 Large 版本相近但做了精简，分辨率几乎和 Large 版本相同，生成速度更快，适合对生成速度要求高的用户。

(3) Medium 版本：该版本于 2024 年 10 月 29 日发布，参数量为 25 亿个，支持生成 25 万像素到 200 万像素的图像，适合计算机配置较低的用户。

用户可根据自己的计算机配置情况，选择适合的版本。需要注意的是，随着硬件技术的快速发展，软件版本更新的速度必将越来越快，作为终端用户，只要把握住软件的核心要素，即可做到“以不变应万变”。

1.2 Stable Diffusion 的应用领域

Stable Diffusion 由于其免费开源、拥有强大的模型和插件等优点，面世至今仅两三年，即在艺术设计等多个领域拥有了广大用户，并得到了广泛应用。目前，其主要应用领域如下。

1.2.1 艺术与设计领域

1. 绘画与插画

对于画家和插画师来说，Stable Diffusion 是一个强大的创意工具。它可以根据艺术家输入的文本提示词快速生成草图，帮助其拓展思路。例如，在创作一幅奇幻主题的插画时，艺术家可以输入“精灵森林中的神秘城堡，有发光的藤蔓缠绕”，模型会生成一个大致场景草图，包括城堡的风格、森林的布局等细节，如图 1-2 所示。艺术家可以在这个基础上进行二次创作，加入自己的绘画风格和细节，如独特的色彩搭配、人物形象等。



图 1-2 奇幻主题插画

2. 平面设计

平面设计主要包括广告设计、海报设计、展示设计等细分领域。以产品展示设计为

例，Stable Diffusion 可以生成产品在不同场景下的展示效果。例如，为一款化妆品生成在精致的化妆台上、被模特拿在手中等不同场景的图像，从而为展示设计提供灵感，并且可以通过不断调整文本提示词来获取满意的设计效果，如图 1-3 所示。



图 1-3 产品展示设计

1.2.2 广告与营销领域

1. 广告创意生成

广告公司可以利用 Stable Diffusion 快速生成广告创意图像。例如，在推广一款无人机时，可输入提示词“一架无人机正在空中飞行，蓝天，白云”，模型即可生成具有吸引力的广告画面，展示无人机的外观和工作状态，如图 1-4 所示。这些图像可以作为初步的创意提案，为广告团队和客户沟通提供可视化的素材，帮助其更快地确定广告的主题和风格。



图 1-4 无人机广告图

2. 社交媒体营销

品牌在社交媒体平台上进行营销时，需要大量吸引人的内容。Stable Diffusion 可以生成适合社交媒体的图像，如产品特写、有趣的场景搭配产品等，如图 1-5 所示的美食特写。



图 1-5 美食特写

1.2.3 游戏开发领域

1. 游戏场景设计

游戏开发者可以使用 Stable Diffusion 生成游戏中的各种场景。对于一款开放世界角色扮演游戏，它可以生成不同地貌的场景，如茂密的森林、荒芜的沙漠、寒冷的冰原等。例如，输入“一个神秘的古代遗迹，位于热带雨林深处，周围环绕着巨大的藤蔓和古老的雕像”，则可得到图 1-6 所示的符合游戏风格的场景图像。



图 1-6 游戏场景设计

2. 角色与道具设计

在角色设计方面，Stable Diffusion 可以生成不同种族、职业的游戏角色与道具形象。例如，为一款魔幻游戏生成精灵族的弓箭手形象，包括其服饰、武器、发型等细节，如图 1-7 所示，美术团队能够在此基础上进行优化和细化，从而加快游戏开发进程。



图 1-7 游戏角色设计

1.2.4 影视与动画领域

1. 影视前期概念设计

在电影和电视剧制作前期，概念设计师可以使用 Stable Diffusion 生成故事场景、角色形象、外星生物等概念图。例如，在一部科幻电影的制作中，对于外星文明的都市、外星生物的外貌等设计，通过输入文本描述，能够快速得到多种概念方案，帮助导演和制片人更快、更好地确定影视作品的视觉风格，如图 1-8 所示。



图 1-8 影视概念图设计

2. 动画制作辅助

在动画制作中，Stable Diffusion 可以生成背景场景。对于长篇动画系列，如需要大量不同的城市街道、自然风光等背景，Stable Diffusion 可以帮助生成初步的背景图像，如图 1-9 所示，然后动画师可以根据动画风格进行调整和绘制。同时，它还可以生成一些非

关键角色的形象，以减轻角色设计的工作量。

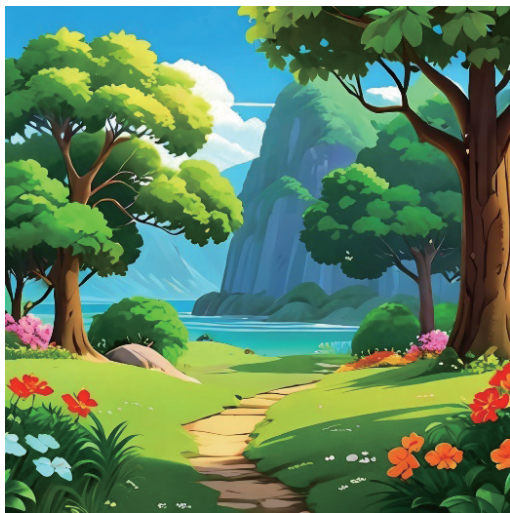


图 1-9 动画场景设计

1.2.5 建筑设计领域

Stable Diffusion 在建筑设计领域的应用非常广泛，在概念设计阶段，建筑师可利用 Stable Diffusion 来获取灵感，帮助探索不同的建筑风格，如图 1-10 所示；在方案设计阶段，在考虑场地布局与环境融合、功能空间布局可视化等方面，Stable Diffusion 可起到很大的作用；在建筑可视化方面，Stable Diffusion 可以生成高质量的建筑外观效果图及室内设计效果图，帮助设计师和客户沟通设计方案，如图 1-11 所示；在城市规划与建筑遗产保护方面，Stable Diffusion 可以生成城市规划模拟图及建筑遗产改造后的概念图等，为相关人员决策提供参考。



图 1-10 地中海风格建筑设计



图 1-11 儿童房室内设计

除此之外，Stable Diffusion 还可应用于教育培训、医学、科研、电商、图像修复与处理等诸多领域，用户可自行尝试。

1.3 Stable Diffusion 的安装

1.3.1 本地安装

用户如果想在自己的计算机上安装部署 Stable Diffusion，则首先需要检查计算机配置是否符合安装条件，因为 Stable Diffusion 对计算机配置的要求比较高。若想流畅地运行 Stable Diffusion，推荐的计算机配置如下。

(1) 显卡 (GPU): 由于 Stable Diffusion 主要依赖显卡的 GPU 进行计算，因此显卡是最为关键的部分。推荐使用英伟达的 RTX 系列显卡，如 RTX 3060 12GB、RTX 4060 Ti 16GB、RTX 4070 Ti Super 16GB、RTX 4080 Super 16GB、RTX 4090 24GB 等。这些显卡支持 CUDA 和 Tensor Core 技术，能够显著提升 AI 计算性能。

(2) 中央处理器 (CPU): 建议选择至少拥有 6 核心 12 线程的处理器，如 Intel Core i5 或 AMD Ryzen 5 系列。对于更复杂的任务或同时运行多个 AI 模型，可以考虑更高端的 CPU，如 Intel Core i7 或 AMD Ryzen 7。

(3) 内存 (RAM): 内存容量直接影响 Stable Diffusion 的运行效率。建议至少配置 16GB 的内存，如果经常处理大型图像或同时运行多个应用程序，32GB 或更高的内存配置更为理想。

(4) 硬盘 (Storage): 硬盘的选择应考虑速度和容量。建议使用至少 500GB 的固态硬盘 (SSD)，以确保快速的数据读写速度。如果需要存储大量的图像文件，可以考虑 1TB 或更大容量的 SSD，或者添加额外的机械硬盘 (HDD)。

(5) 显示器 (Monitor): 对于 AI 绘画来说，显示器的色彩准确性和分辨率同样重要。建议选择分辨率至少为 1920 × 1080 像素的显示器，色彩覆盖至少达到 100% sRGB。如果预算允许，可以考虑更高分辨率和色彩覆盖的显示器，如 4K 显示器，以获得更细腻的图像细节。

(6) 电源 (PSU): 确保电源单位能够提供足够的功率来支持所有硬件。建议选择 80 Plus Gold 或更高级别的电源，以确保电源的稳定性和效率。

(7) 散热系统 (Cooling): 由于 Stable Diffusion 在运行时会大量占用 GPU 资源，因此散热系统也非常重要。为确保机箱有足够的空间进行空气流通，必要时可考虑使用高性能的散热系统。

1. 安装流程

以下以 Windows 10 操作系统为例，介绍 Stable Diffusion 的安装流程。

(1) 下载 Stable Diffusion 程序包。首先需要从 Stable Diffusion 官方网站或其他可信的来源下载该软件的程序包，下载完成后，将压缩文件解压到用户想要安装的目录下。

(2) 安装 Python 环境。由于 Stable Diffusion 是使用 Python 语言开发的，因此用户需要在本地安装 Python 环境。需要注意的是，Stable Diffusion 要求使用 Python 3.6 以上的版本。

(3) 安装依赖项。依赖项指为了使 Stable Diffusion 能够正常运行，需要安装和配置的其他相关的软件库或组件，这些依赖项可以是编程语言、框架、库文件或其他软件包。在

安装 Stable Diffusion 之前，用户需要确保下列依赖项已正确安装。

① PyTorch: PyTorch 是一个开源的 Python 机器学习库，提供了易于使用的张量 (tensor) 和自动微分 (automatic differentiation) 等技术，这使其特别适合于深度学习和大规模的机器学习等项目。

② NumPy: NumPy 是 Python 的一个数值计算扩展，它提供了快速、节省内存的数据，以及用于数学和科学编程的常用函数。

③ Pillow: Pillow 是 Python 的一个图像处理库，可以用来打开、操作和保存不同格式的图像文件。

④ SciPy: SciPy 是一个用于数学、科学、工程领域的数学计算库，可以处理插值、积分、优化、图像处理、常微分方程数值解的求解、信号处理等问题。

⑤ tqdm: tqdm 是一个快速、可扩展的 Python 进度条库，可以在长循环中添加一个进度提示，让用户知道程序的运行进度。

2. 一键安装

用户可以使用由 B 站 (bilibili, 哔哩哔哩) “大咖” 秋葉 aaaki 分享的“秋葉整合包”，如图 1-12 所示，一键实现 Stable Diffusion 的本地部署。只需运行“启动器运行依赖 -dotnet-6.0.00.exe” 安装程序，然后单击“安装”按钮即可。执行操作后，等待出现“控制台”窗口，不必在意“控制台”窗口中的内容，保持其打开状态即可。稍待片刻，将会出现一个浏览器窗口，表示 Stable Diffusion 的基本软件已经安装完毕。



图 1-12 秋葉整合包使用说明视频封面^①

1.3.2 云服务器部署

如果用户的计算机配置达不到 Stable Diffusion 的安装条件，可以通过云服务器部署 Stable Diffusion。用户可以在百度飞桨、阿里云、腾讯云、Google Colab 等常用的云服务

^① 参考来源: https://www.bilibili.com/video/BV1iM4y1y7oA/?spm_id_from=333.337.search-card.all.click&vd_source=6855f4822a9b7c132d1b998d4571b410.2023-04-16。

器平台上部署 Stable Diffusion，这样只需在云端输入自己的文本描述，即可得到 AI 生成的图像。具体的部署操作依平台不同而有所差异，用户可根据实际情况在相关平台进行学习和部署。

1.4 Stable Diffusion WebUI 界面简介

Stable Diffusion 软件有多种界面，如 Stable Diffusion WebUI、Stable Diffusion ComfyUI 和 StableSwarmUI，其中最为常用的是 Stable Diffusion WebUI，因此本书均以 WebUI 界面为例进行介绍。

Stable Diffusion WebUI 的界面如图 1-13 所示，主要包括以下几大区域。

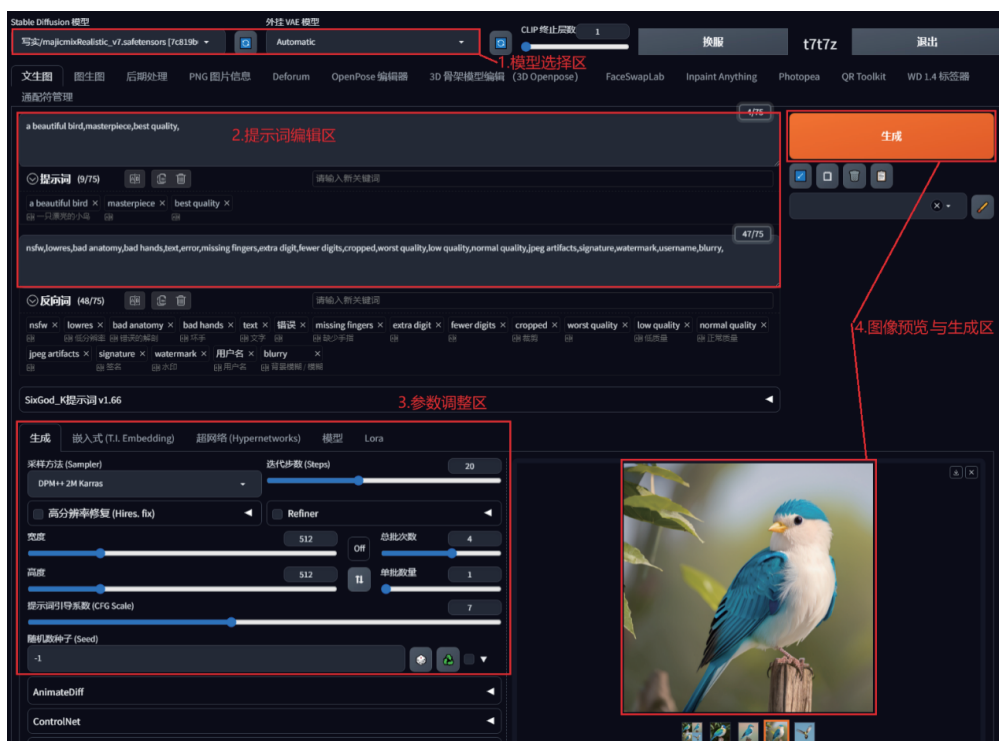


图 1-13 Stable Diffusion WebUI 界面

(1) 模型选择区。用户可在模型选择区选择不同的 Stable Diffusion 大模型及外挂 VAE 模型。

(2) 提示词编辑区。提示词编辑区主要有“文生图”“图生图”等选项卡共十多个，以及提示词及反向词输入文本框等，用户可选择相应的选项卡进行适当的操作。

(3) 参数调整区。参数调整区主要包含图像生成过程中所需调整的各项参数，如采样方法、迭代步数、图像尺寸、提示词引导系数等。

(4) 图像预览与生成区。图像预览与生成区主要包括“生成”按钮和图像预览窗口，用户在执行完前述操作后，可通过单击“生成”按钮得到最终图像，并在图像预览区域实时预览所生成的图像。

1.5 使用 Stable Diffusion 生成一张图像

使用 Stable Diffusion 生成图像的操作流程非常简单，共分为 4 个步骤：①选择大模型；②输入提示词（或上传原图）；③设置生成参数；④单击“生成”按钮。最终的图像效果是由大模型、提示词（或原图）和生成参数三者共同决定的。其中，大模型主要决定图像的画风；提示词（或原图）主要决定画面内容；生成参数则主要用于设置图像的预设属性。通过这个流程，用户可以轻松地使用 Stable Diffusion 生成符合自己要求的各种图像。

下面通过一个简单的实例介绍如何使用 Stable Diffusion 快速生成一张图像，效果如图 1-14 所示。




图 1-14 小鸟图像

1. 选择大模型

在“Stable Diffusion 模型”下拉菜单中选择一个大模型，此例中选择的是“写实/majicmixRealistic_v7.safetensors [7c819b6d13]”大模型，作为 AI 绘画的底模。

2. 输入提示词

在“文生图”页面的“提示词”文本框中输入想要绘制的内容，用英文单词和句子输入，此例输入的是“a beautiful bird（一只漂亮的小鸟）”。

在“反向词”文本框中，用户可以输入一些反向词，即不希望在画面中出现的要素。如果不想自己输入或者没有经验，系统自带了许多类型的反向词，用户只需选择即可。具体操作：在界面右侧“预设样式”下拉列表中选择“基础起手式”，单击列表上方的“将所有当前选择的预设样式添加到提示词中”按钮 ，此时“反向词”文本框中即出现了一系列提示词，如“nsfw, lowres, bad anatomy, bad hands, text, error, missing fingers, extra digit, fewer digits, cropped, worst quality, low quality, normal quality, jpeg artifacts, signature, watermark, username, blurry, (不适宜工作场合, 低分辨率, 解剖结构糟糕, 手部画得不好, 文字, 错误, 手指缺失, 多指, 少指, 画面裁切, 质量极差, 质量低, 质量一般, JPEG 图像伪影, 签名, 水印, 用户名, 模糊,)”，同时在“提示词”文本框中自动出现提示词“a beautiful bird, masterpiece, best quality, (一只漂亮的小鸟, 杰作, 最佳质量,)”，

如图 1-15 所示。



图 1-15 确定大模型和提示词

3. 参数设置

在界面下方的参数面板中，设置“采样方法 (Sampler)”为 DPM++ 2M Karras，“迭代步数 (Steps)”为 30，“宽度”为 512，“高度”为 512，其他参数采用默认值，如图 1-16 所示。



图 1-16 参数设置

4. 生成图像

单击界面右上角的“生成”按钮，稍等片刻，即可生成图 1-14 所示的小鸟图像。

1.6 本章小结

本章主要介绍了 Stable Diffusion 的基础知识、主要应用领域、两种常用的安装方法、主要操作界面，最后通过一个实例，简要介绍了利用 Stable Diffusion 进行 AI 绘画的基本流程。