



中国美术学院出版社
CHINA ACADEMY OF ART PRESS

艺术设计专业系列丛书

- | | |
|--------------------|-------------------|
| 设计概论 | 构成基础 |
| 文创设计 | 造型基础 |
| 品牌视觉识别设计 | 服饰品设计 |
| 卡通形象设计 | 服装立体裁剪 |
| 包装设计 | 商业摄影 |
| 居住空间设计 | Premiere影视后期编辑 |
| 酒店空间设计 | 影视后期特效制作与剪辑合成 |
| 室内设计全过程案例教程 | 3ds Max建模项目与应用 |
| 环境设计制图 | 三维模型制作实践教程（Maya版） |
| 公共环境艺术设计教程 | 三维角色建模实战教程 |
| 环境艺术模型设计与制作 | 三维动画制作 |
| ▶▶ AIGC设计与绘画创作实践教程 | 定格动画短片创作 |
| 漆艺 | 家具设计（第2版） |
| 民间美术 | Photoshop CC实战教程 |
| 立体构成 | |



扫描二维码关注
中国美术学院出版社官方订阅号



定价：52.00元

艺术设计专业系列丛书

AIGC设计与绘画创作实践教程

张鹏 宣兴磊 卢晓洁◎主编

中国美术学院出版社

安徽省高等学校省级质量工程项目立项优质教材
艺术设计专业系列丛书

AIGC SHEJI
YU HUIHUA CHUANGZUO
SHIJIAN JIAOCHENG

AIGC设计与 绘画创作 实践教程

张 鹏 宣兴磊 卢晓洁◎主编



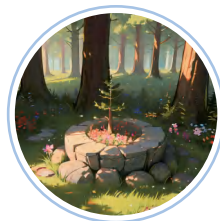
中国美术学院出版社

安徽省高等学校省级质量工程项目立项优质教材
艺术设计专业系列丛书

AIGC SHEJI
YU HUIHUA CHUANGZUO
SHIJIAN JIAOCHENG

AIGC设计 与绘画创作 实践教程

张 鹏 宣兴磊 卢晓洁◎主编



 中国美术学院出版社

责任编辑：孟海江
执行编辑：周 赟
特约编辑：张荣昌
图书制作：宏图文化
封面设计：唐韵设计
装帧设计：宏图文化
责任校对：杨轩飞
责任印制：张荣胜

图书在版编目（CIP）数据

AIGC 设计与绘画创作实践教程 / 张鹏，宣兴磊，卢晓洁主编. -- 杭州：中国美术学院出版社，2025. 8.
(艺术设计专业系列丛书). -- ISBN 978-7-5503-3617-9
I. TB21-39
中国国家版本馆 CIP 数据核字第 20254A9E29 号

艺术设计专业系列丛书

AIGC 设计与绘画创作实践教程

张 鹏 宣兴磊 卢晓洁 主编

出 品 人：祝平凡
出版发行：中国美术学院出版社
地 址：中国·杭州南山路 218 号 / 邮政编码：310002
网 址：<http://www.caapress.com>
经 销：全国新华书店
印 刷：北京荣玉印刷有限公司
版 次：2025 年 8 月第 1 版
印 次：2025 年 8 月第 1 次印刷
开 本：889 mm × 1194 mm 1/16
印 张：10
字 数：274 千
书 号：ISBN 978-7-5503-3617-9
定 价：52.00 元

著作权所有·违者必究

编写委员会

| 主 编 | 张 鹏 宣兴磊 卢晓洁

| 副主编 | 李娜娜 刘咏松 张婧婧

| 编 委 | 潘晓虎 张 翔 王婷玉

周雨成



在科技飞速发展的今天，人工智能（Artificial Intelligence，AI）与人工智能生成内容（Artificial Intelligence Generated Content，AIGC）已经成为引领未来科技革命和产业变革的重要力量。AIGC 技术在艺术设计领域的应用也越来越广泛而深入，为设计师提供了更加高效和智能的创作工具和平台。为了适应这一趋势，培养更多掌握 AIGC 技术的艺术设计人才，编者精心编写了本书。本书旨在系统地介绍 AIGC 在艺术领域的基本原理、关键技术及实践应用，为广大学者和技术爱好者提供一份权威、实用的学习指南。

党的二十大报告指出，教育、科技、人才是全面建设社会主义现代化国家的基础性、战略性支撑。必须坚持科技是第一生产力、人才是第一资源、创新是第一动力，深入实施科教兴国战略、人才强国战略、创新驱动发展战略，开辟发展新领域新赛道，不断塑造发展新动能新优势。^①目前 AIGC 技术无疑为数字文化产业注入新的活力。AIGC 技术的学习对于提升个人技能、培养创新思维、适应未来发展趋势、提升综合素质、增强个人竞争力都具有重要意义，对 AIGC 技术的学习也十分重要和迫切。然而，目前市场上关于 AIGC 技术在艺术领域应用的教材相对较少，且存在内容零散、不系统的问题。因此，编者结合行业最新的技术和实践经验，编写了本书。

本书在编写的过程中注重理论与实践相结合，既深入剖析了 AIGC 技术的原理和使用方法，又提供了丰富的案例和实践经验。同时，本书还贯彻立德树人的根本任务，在知识讲解中融入思政教育和党的二十大精神。整体上，本书主要具有以下特色。

思政融入：在任务实施过程中融入了中华优秀传统文化、乡村振兴等元素，贯彻党的二十大精神，实现思想政治教育 with 知识体系教育的有机统一。

^① 中国政府网：《习近平：高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告》，2022 年 10 月 25 日，https://www.gov.cn/xinwen/2022-10/25/content_5721685.htm。

内容全面：涵盖了 AIGC 图像生成技术在艺术领域的多个应用场景，帮助读者全面了解 AIGC 技术的全貌。

结构清晰：按照从基础到进阶的顺序编排章节，使读者能够循序渐进地掌握 AIGC 技术的知识和技能。

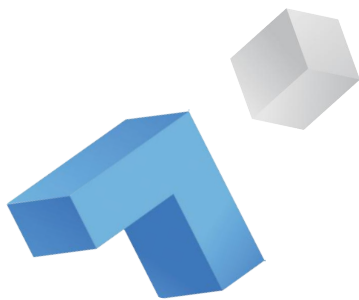
图文并茂：书中配有大量的图表、示意图和案例，帮助读者更好地理解 AIGC 技术应用。

资源丰富：编者为广大一线教师提供了服务于本书的教学资源库，有需要者可致电教学助手 13810412048 或发邮件至 2393867076@qq.com。

此外，本书为校企合作共同开发的新形态教材，成书过程中得到合肥顽皮机器教育科技有限公司、合肥太空硅基数字科技有限责任公司，以及 AI 高级工程师周雨成、孟诚的大力支持与帮助，诸多案例均来自企业实际项目。本书亦为安徽省高等学校省级质量工程项目立项优质教材。

本书以 Stable Diffusion 为媒介，通过模块式的教学，介绍了软件的安装、模型的种类与应用、关键词的编写技巧、图像控制应用等内容，并设置了涉及角色、场景、环艺、平面等方向的任务实践模块，帮助读者掌握重、难点知识，提高能力和培养素质。本书可作为计算机科学与技术、人工智能、艺术设计等相关专业的教学用书，也可作为 AIGC 技术爱好者的参考用书。

通过对本书的学习，读者将能够深入了解 AIGC 技术的原理和应用，掌握相关技能，为进一步的学习和实践奠定坚实的基础。由于 AIGC 技术发展迅速，同时受编者自身水平所限，本书若存在疏漏和不足之处，敬请广大读者提出宝贵的意见和建议，以便我们不断完善和提高教材的质量和水平。



目录

模块 1

AIGC 开篇 1

任务 1.1 初识 AIGC 2

1. AIGC 的概念 2
2. AI 的工作原理 2

任务 1.2 了解常用的 AIGC 工具 3

1. 文心一格 3
2. Midjourney 3
3. Stable Diffusion 3

任务 1.3 探究 AIGC 的社会影响及争议 5

模块 2

Stable Diffusion 通识篇 7

任务 2.1 安装 Stable Diffusion 8

1. Stable Diffusion 安装配置要求 8
2. 克隆安装方法 9
3. 整合包安装方法 14

任务 2.2 认识 Stable Diffusion 基础界面 16

1. “底模”及“VAE 模型”加载位置 16
2. “文生图”和“图生图”界面介绍 16
3. “附加功能”界面介绍 18
4. “图片信息”界面介绍 19
5. “训练”界面介绍 20
6. “设置”界面介绍 20
7. “扩展”界面介绍 22

模块 3

Stable Diffusion 中阶篇

23

任务 3.1 认识 Stable Diffusion 的模型种类

24

1. 模型下载地址 24
2. Checkpoint 模型 27
3. Lora 模型 28
4. Embedding 模型 29
5. Hypernetworks 模型 31
6. VAE 模型 32

任务 3.2 应用模型

34

1. Checkpoint 模型应用 34
2. Lora 模型应用 36
3. Embedding 模型应用 37
4. Hypernetworks 模型应用 37
5. VAE 模型应用 38

任务 3.3 使用 Stable Diffusion 进行绘制

40

1. Stable Diffusion 绘制基础 40
2. Stable Diffusion 绘制参数调整 41
3. Stable Diffusion 绘制分辨率提高 43
4. Stable Diffusion 图片信息复刻 45
5. Stable Diffusion “随机种子” 控制 46

模块 4

Stable Diffusion 进阶篇

51

任务 4.1 撰写 Stable Diffusion 提示词

52

1. Stable Diffusion 正向提示词的撰写规范 52
2. Stable Diffusion 反推功能 54
3. Stable Diffusion 反向提示词的撰写规范 56

任务 4.2 调整 Stable Diffusion 提示词的权重	57
1. Stable Diffusion 提示词权重基础	57
2. Stable Diffusion 提示词权重符号种类	59
任务 4.3 安装 Stable Diffusion 常用插件	61
1. Stable Diffusion 插件下载地址及安装路径	61
2. Stable Diffusion 常规插件的介绍	63

模块 5

Stable Diffusion 高阶篇 67

任务 5.1 脚本应用	68
1. X/Y/Z 图表应用	68
2. Ultimate SD upscale 应用	71
任务 5.2 AI 绘画模型训练	74
1. Lora 模型训练	74
2. Dreambooth 大模型训练	87
任务 5.3 ControlNet 插件应用	95
1. ControlNet 插件模型安装	95
2. ControlNet 插件使用	98

模块 6

Stable Diffusion 案例篇 107

任务 6.1 AIGC 草图方案制作——古风角色方案	108
1. 任务描述	108
2. 任务目标	108
3. 工作流程	108
任务 6.2 AIGC 线稿填色细化制作	112
1. 任务描述	112
2. 任务目标	112

3. 工作流程	112
任务 6.3 AIGC 角色三种视图制作	118
1. 任务描述	118
2. 任务目标	118
3. 工作流程	118
任务 6.4 AIGC 草图分层细化制作	122
1. 任务描述	122
2. 任务目标	123
3. 工作流程	123
任务 6.5 AIGC 角色手办 IP 制作	127
1. 任务描述	127
2. 任务目标	127
3. 工作流程	127
4. 理论指导	131
任务 6.6 AIGC 古风游戏场景气氛图设计	132
1. 任务描述	132
2. 任务目标	132
3. 工作流程	132
4. 理论指导	138
任务 6.7 AIGC 乡村旅游民宿景观方案	140
1. 任务描述	140
2. 任务目标	140
3. 工作流程	140
4. 理论指导	143
任务 6.8 传统节日祝福海报设计	144
1. 任务描述	144
2. 任务目标	144
3. 工作流程	144
4. 理论指导	148

参考文献	149
-------------	------------

模块 3

Stable Diffusion 中阶篇

当我们在网上看到一些令人心动的 AI 绘图作品时，也想使用 AI 软件创作出属于自己的 AI 作品，然而通过提取并使用这些 AI 佳作的提示词，最终生成图像的品质却达不到网络宣传所呈现的精美效果，这是为什么呢？实际上，想要提高 AI 生成图像的品质，不仅需要合适的提示词，更为重要的是需要相互匹配的模型套组，两者结合才能生成具有理想效果的作品。这就好比一件游戏套装，单独使用的效果非常普通，但完整的套装所带来的品质和效果却有翻天覆地的变化。

本模块主要介绍常用模型的获取途径，Stable Diffusion 的不同模型类型及其主要作用和使用方法。

学习目标

知识目标

- 1) 了解 Stable Diffusion 模型的种类。
- 2) 掌握 Stable Diffusion 模型的作用。
- 3) 掌握 Stable Diffusion 绘图的基本流程。
- 4) 了解 Stable Diffusion 生成参数的作用。

能力目标

- 1) 能够从网络上获取所需要的模型。
- 2) 能够区分不同大模型的风格，并在项目中选择合适的大模型。
- 3) 能够正确选择并使用不同类型的模型。

素养目标

培养审美和鉴赏能力，可从众多风格的作品中筛选所需的艺术风格。

任务 3.1

认识 Stable Diffusion 的 模型种类

1. 模型下载地址

目前，AI 模型的使用者基数很大，普及也非常广，有众多 AI 绘画爱好者乐此不疲地制作出效果各异、种类不同的模型，这也使 Stable Diffusion 的模型资源变得非常丰富。Stable Diffusion 作为一款开源软件，在国内外的影响力也极为深远，网络上有各种类型的模型网站，可供下载模型也有多个版本，不同版本的模型大小存在差异，涉及的格式后缀也较为烦琐，这些问题会在后面一一讲解与介绍。本小节主要介绍几个普及较为广泛的模型网站。

(1) C 站

作为老牌 AI 模型网站，Civitai 网站（简称 C 站，图 3-1-1）提供了高质量且数量庞大的免费 Stable Diffusion 模型，具有比较快捷的模型更替速度，附带可供学习交流的社区，属于主流模型资源网站。其 AI 生成的自定义模型和成品图都附上了关键参数及提示词标签，为广大 AI 绘画爱好者提供了非常好的学习指导。

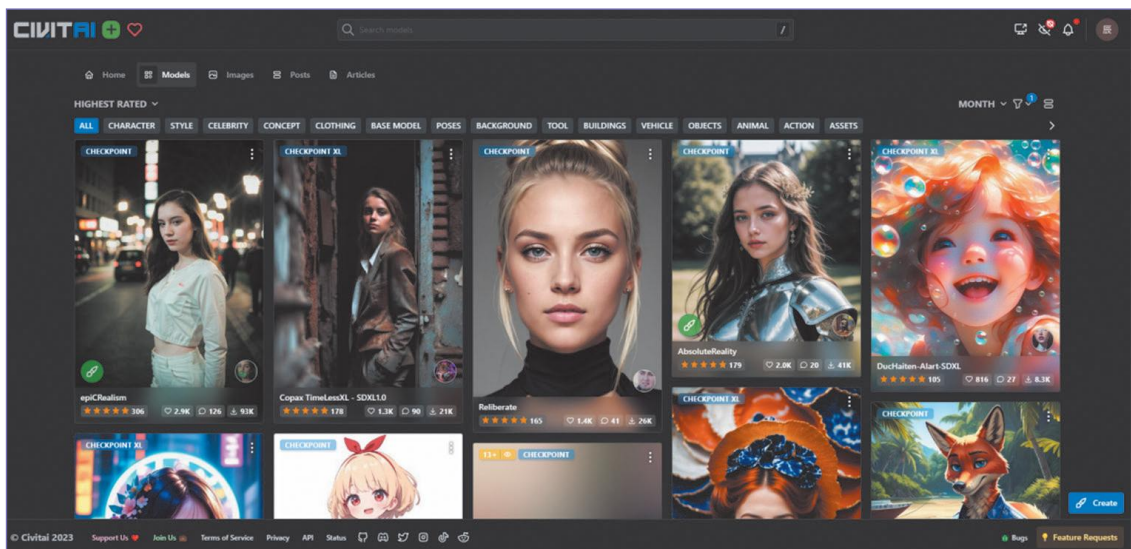


图 3-1-1 Civitai 站

(2) 吐司

吐司（图 3-1-2）是由上海的一家公司提供的 AI 模型分享社区和在线运行平台。该平台由一群认可和憧憬 AI 技术的小伙伴组成，该模型网站功能齐全并拥有手机 App 可供使用，除了整合国内外丰富的模型资源外，也镜像存储了 C 站的部分模型资源。同时，其具备的线上 AI 绘图功能也让用户

可以随时随地自由创作喜欢的 AI 作品，而不用担心因计算机配置不够而无法生成图像。该平台在资源集中也附带了目前主流的本地部署的 AI 绘图软件，一定程度上解决了 AI 绘画爱好者安装软件时找不到资源的问题。

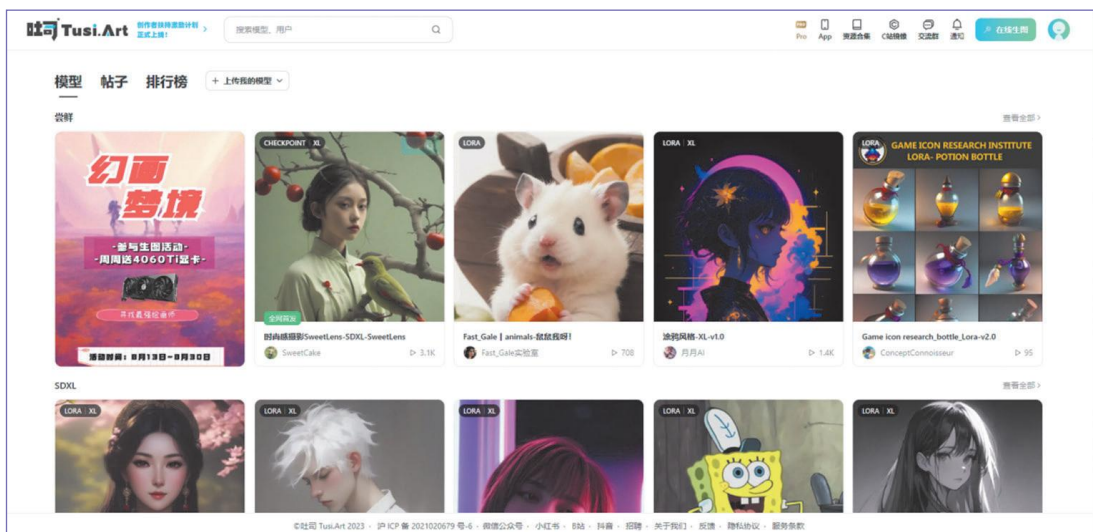


图 3-1-2 吐司

(3) Liblib AI

在 AI 模型资源网站中，Liblib（哩布哩布）AI（图 3-1-3）是专注于提供更方便的参考图片和模型资源的网站，该网站功能齐全，操作简单，其主要功能基本与 C 站相同，同样附加了在线 AI 绘图的功能，并且收藏的模型可以直接在在线 AI 绘图中使用。

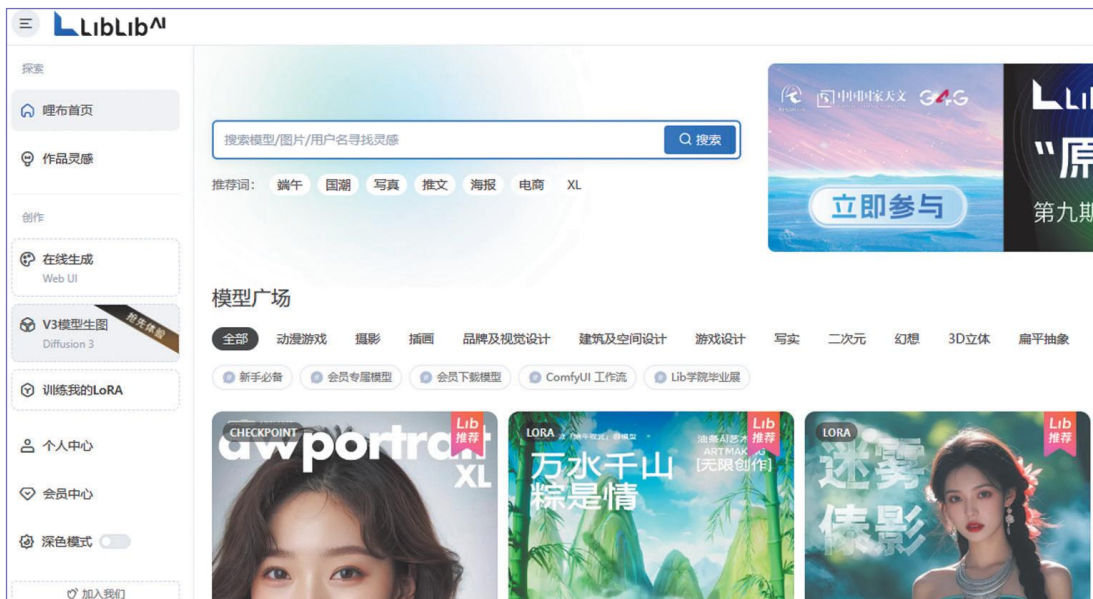


图 3-1-3 Liblib AI

(4) 搜图神器

搜图神器（图 3-1-4）是国内的图片搜索网站，除了包含丰富的 AI 绘制图片外，其附属功能也

极为全面，整合了众多国内外常用网址及小程序，感兴趣的读者可自行探索。



图 3-1-4 搜图神器

(5) PromptTool 词图 AI

PromptTool 词图 AI (图 3-1-5) 前期的主要功能是辅助填写 AI 绘画中最为复杂的提示词标签，其组词工具对于初学者来说具有非常强大的功能，该网站在后期同样增加了在线 AI 绘图的功能，并设置了共享交流社区供用户分享绘画经验和提示词填写经验。

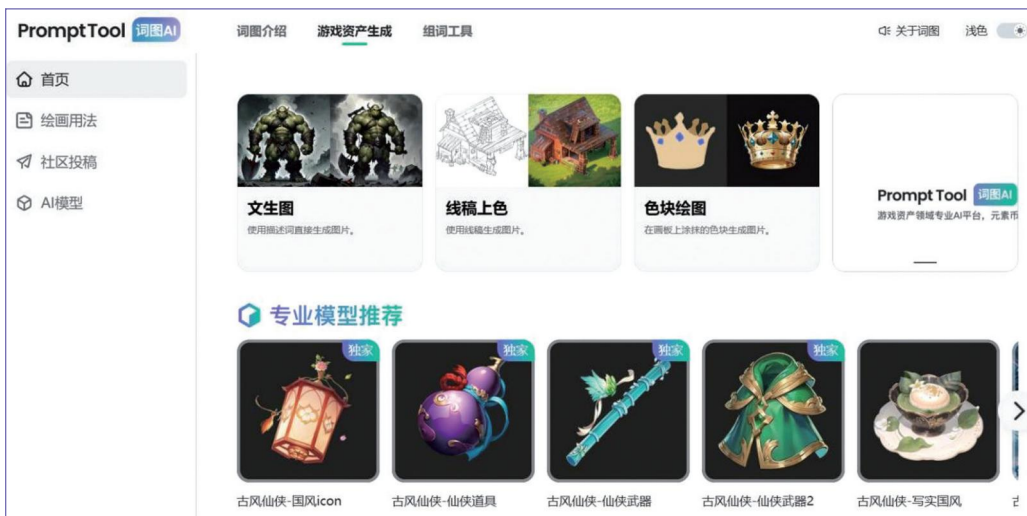


图 3-1-5 PromptTool 词图 AI

以上这些网站都有其独特的优点，对于 AI 绘画爱好者来说是学习、掌握 AI 绘画的重要途径。当然，除了这些网站之外，还有很多类型独特且质量优秀的 AI 绘图资源分享网站，读者可以从中各取所需、积极探索了。

简单了解了获取 AI 绘图模型资源的几个途径后，接下来将详细介绍在使用 AI 绘画过程中所涉及的几大模型的种类及其作用，为后期掌握 AI 绘画的操作提供一些必要的帮助。

在进行讲解的过程中，我们使用的 Stable Diffusion 本地部署程序的版本皆为 1.5.1，若出现版本

的差异，建议统一切换到此版本进行学习。

2. Checkpoint 模型

Checkpoint 模型也称为“大模型”或“底模”，可以将它理解为统一绘画风格的图片及文本的打包集合。此类特定画风指的是图像的艺术展现方式，如真实的照片、中国水墨、二次元、卡通、皮克斯等。在 Stable Diffusion 中，所有关键提示词生成的图像都是基于大模型的风格生成的，大模型中的图片风格决定了生成图像的具体艺术表现，它就像生产线的第一道门槛，通过它才能进行下一步。正因如此，如果使用 Stable Diffusion 生成图像，选择合适的大模型就显得极其重要。

打开 Stable Diffusion 的操作界面，大模型的位置一般在界面的左上角，命名为“Stable Diffusion 模型”，如图 3-1-6 所示。

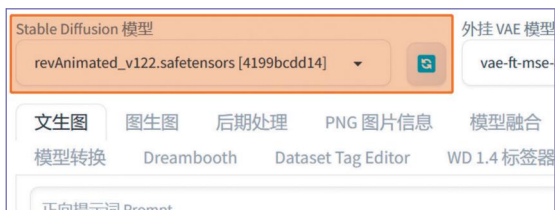


图 3-1-6 大模型的位置

在本地部署的程序中，Stable Diffusion 大模型通常存储在应用文件根目录下的“models\stable diffusion”文件夹中，存储完毕后，在操作界面中单击模型框旁的蓝色旋转按钮，便可在页面中刷新查看到大模型的名称。

“大模型”的文件后缀有“.ckpt”和“.safetensors”两种。文件体积以 GB 为单位，因为一般大模型都大于 2 GB，所以在众多模型分类中比较容易识别。

以网站 Liblib AI 为例下载、安装大模型，首先打开网页找到右侧的检索选项，单击“CHECKPOINT”类型标签，筛选出来的都是大于 2 GB 的大模型文件，在模型列表缩略图的左上角也可以看到对应的模型名称，很好辨认，如图 3-1-7 所示。当然还要注意模型的适用版本，错误版本的模型即使安装到程序中也是无法使用的。选好对应的模型后，进入文件下载页面，单击“下载”按钮即可，如图 3-1-8 所示。

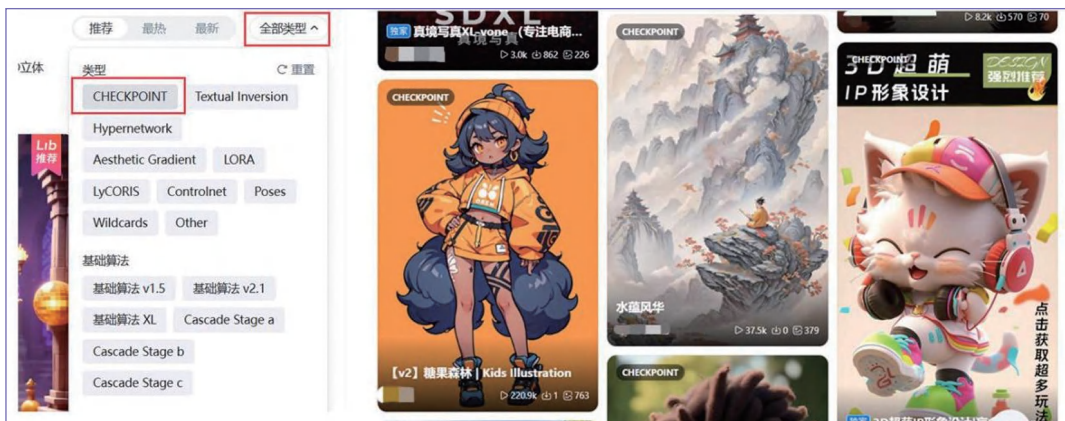


图 3-1-7 筛选大模型



图 3-1-8 下载大模型

其他的模型资源网站中，筛选的方式也基本相同，在此就不一一举例了。

3. Lora 模型

Lora 模型全称是 Low-Rank Adaptation of Large Language Models，直译为大语言模型的低价适应，这是微软的研究人员为了解决大语言模型微调问题而开发的一项技术，其作用于神经网络模型，增加运算层（神经网络模型：Stable Diffusion 的图片运算过程中在 Diffusion 模型中运行的一种计算机算法或者数学模型）。可以简单地理解为：当“大模型”的图库不足时，它通过创造独立的小型模型来填补所需的特定图像风格的空白。针对特定的图像，让 AI 通过少量的图片进行学习，以影响后来生成的图像来达到微调模型的目的，最终获得理想的效果。

该模型适用于绘制特定的人物、画风、产品等，应用范围极为广泛。Lora 模型之所以应用范围如此广泛，其主要原因在于训练 Lora 模型的过程相对于大模型来说，步骤简单且耗时较短，可以快速完成模型的训练并进行使用，以获得理想的效果。

Lora 模型放置在 Stable Diffusion 界面中的扩展模型中，通常处于隐藏状态，使用时可以单击“生成”按钮下的红色小图标，接着在提示词下方延伸出新的扩展模型界面，Lora 模型选择界面如图 3-1-9 所示。

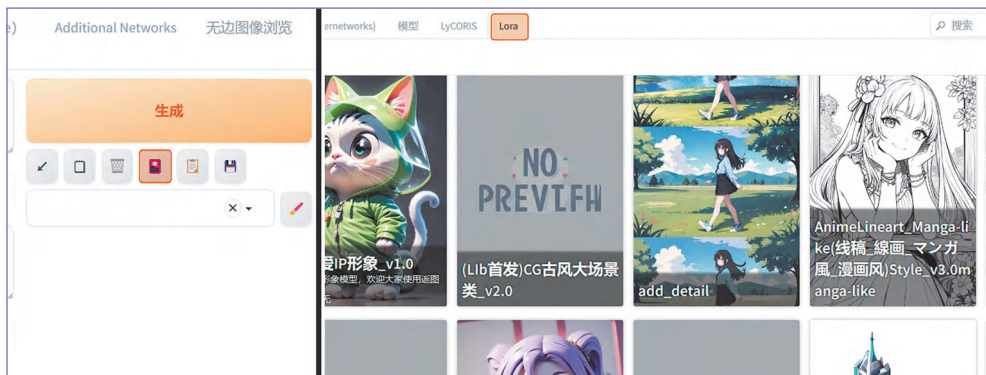


图 3-1-9 Lora 模型选择界面

在本地部署的程序中，Lora 模型通常存储在应用文件根目录下的“\models\Lora”文件夹中，存储完毕后，在操作界面单击模型旁的蓝色旋转按钮，便可在页面中刷新查看。

Lora 模型的下载方式也很简单，模型下载网站中以 Liblib AI 为例，打开网页找到右侧的筛选类型，选择“LORA”标签，就可以看到网站内收纳的所有 Lora 模型，在模型列表缩略图的左上角也可以看到对应的模型类型，很好辨认。Lora 模型筛选界面如图 3-1-10 所示。

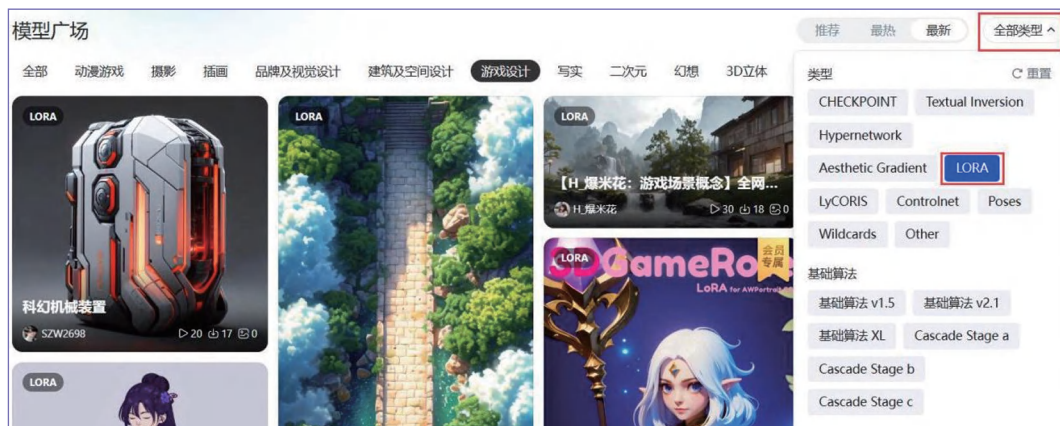


图 3-1-10 Lora 模型筛选界面

点击进入一个 Lora 模型，再单击“下载”按钮，其资源就会下载到计算机中，Lora 模型下载页面如图 3-1-11 所示。将下载后的 Lora 模型放到 Stable Diffusion 根目录的相应位置，就可以在刷新操作界面后找到该模型。Lora 文件的大小通常只有几百兆字节，虽然它占用空间小，但在图像处理方面的效果却异常优秀，因此受到广大 AI 绘画爱好者的青睐。

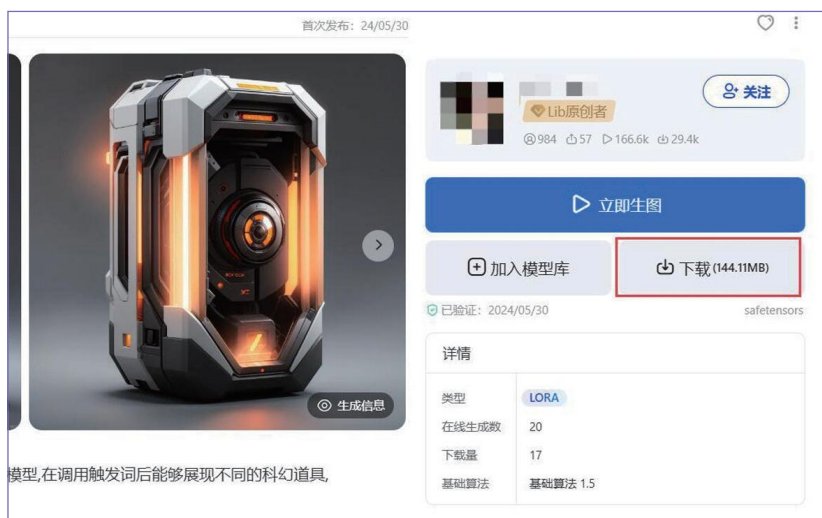


图 3-1-11 Lora 模型下载页面

4. Embedding 模型

Embedding 模型就是嵌入式（T.I.embedding）模型。T.I. 是 Textual Inversion 的缩写，即文本反转，它基于文本编码器增加新的文本向量来提升设计质量，简单理解其作用就是针对在大模型中缺少

的对于某些元素的具体描述，将其详细的文本描述进行标签化，提高对某一词汇及图像的认知，增加精准度。

与 Lora 模型不同，Embedding 模型主要是将特定的细节描述嵌入大模型中，让大模型去快速认识一个全新的概念。这种控制图像的方式虽然并不稳定，且无法进行精确的学习，但对特定图像的处理效果具有独到之处。

Embedding 模型的格式一般以“.pt”和“.safetensors”结尾，在操作界面的位置也比较醒目，通常与 Lora 模型处于同一个扩展模型栏中，如图 3-1-12 所示。



图 3-1-12 Embedding 模型位置

Embedding 模型作为早期的扩展模型，在 Stable Diffusion 的根目录下有独立的文件夹，文件放在根目录下的“embeddings”文件夹内。

想要下载 Embedding 模型资源，可以在 Liblib AI 中打开主页，找到右侧的筛选类型，选择“Textual Inversion”标签，就可以看到网站内收纳的所有 Textual Inversion 模型，在模型列表缩略图的左上角也同样可以看到对应的模型类型。Embedding 模型筛选界面如图 3-1-13 所示。

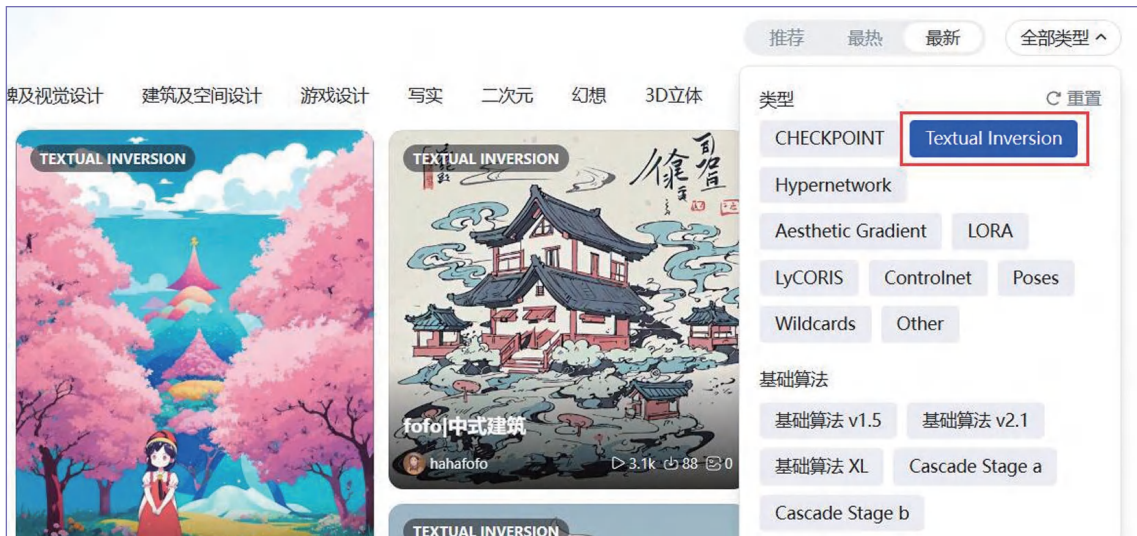


图 3-1-13 Embedding 模型筛选界面

单击其中一个 Embedding 模型后，再单击“下载”按钮，其资源就会下载到计算机中，Embedding 模型下载页面如图 3-1-14 所示。值得注意的是，对于不同的 Embedding 模型，其使用方法也有所区别。使用前需要认真阅读模型作者给予的提示和使用方法。



图 3-1-14 Embedding 模型下载页面

5. Hypernetworks 模型

Hypernetworks 模型即超网络模型，其功能与 Lora 模型和 Embedding 模型类似，是对 Stable Diffusion 生成的图片进行针对性调整以控制画面、获得理想图的一种方式。其在原理上主要作用于神经网络模型，通过修改运算权重来影响出图。Hypernetworks 模型可以理解为低配版的 Lora 模型，因其应用的领域较窄，主要以训练画风为主，训练难度大，逐渐被 Lora 模型所替代，所以简单了解即可。Hypernetworks 模型的位置如图 3-1-15 所示。

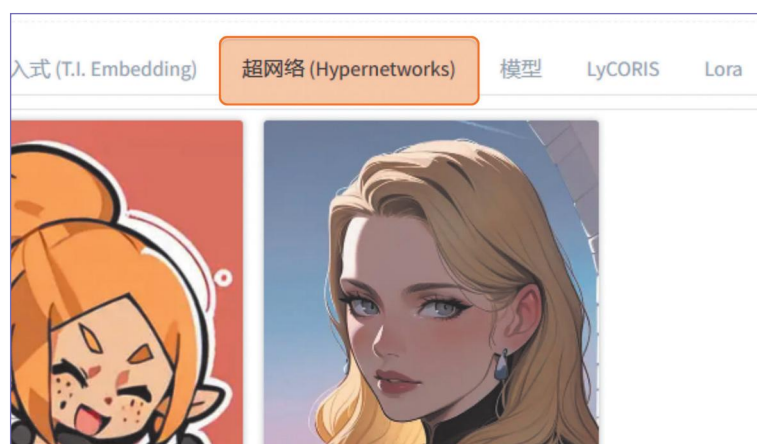


图 3-1-15 Hypernetworks 模型的位置

在本地部署的程序中，Hypernetworks 模型通常存储在应用文件根目录下的“\models\hypernetworks”文件夹中，存储完毕后，在操作界面单击模型旁的刷新按钮，便可在页面中刷新查看到。

想要下载 Hypernetworks 模型的资源，在 Liblib AI 中打开主页，找到右侧的筛选类型，选择“Hypernetwork”，就可以看到网站内收纳的所有 Hypernetworks 模型，在模型列表缩略图的左上角也同样可以看到对应的模型类型。

单击进入一个 Hypernetworks 模型后，再单击“下载”按钮，其资源就会下载到计算机中，将下载后的 Hypernetworks 模型放置到 Stable Diffusion 目录的相应位置，就可以在刷新操作界面后找到。

Hypernetworks 模型下载页面如图 3-1-16 所示。

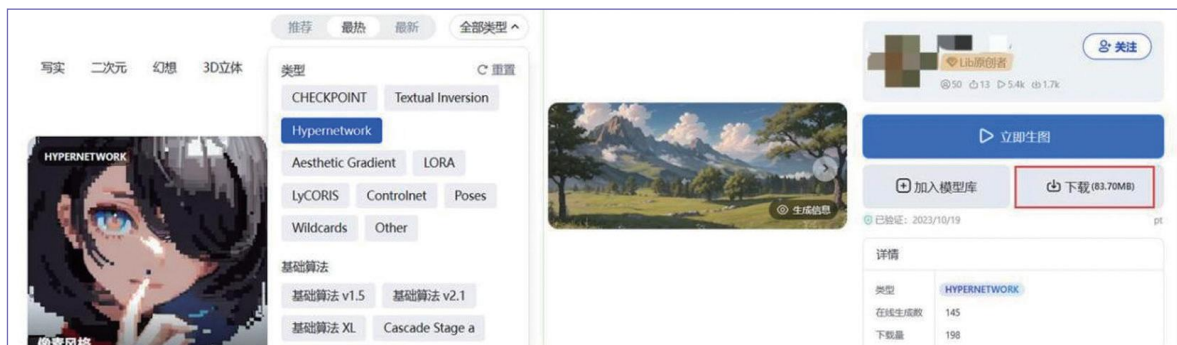


图 3-1-16 Hypernetworks 模型下载页面

6. VAE 模型

VAE 全称为 Variational Auto-Encoder，即变分自编码器，是一个编码器和解码器的集合。VAE 可以简单理解为给大模型增加滤镜和细节的工具，通常用以优化细节，修改颜色、饱和度等，作用比较强大。VAE 模型在操作界面中就放置在大模型的旁边，如图 3-1-17 所示。

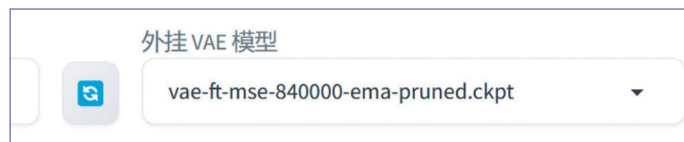


图 3-1-17 VAE 模型的位置

之所以最后介绍 VAE 模型，主要是因为它的运作原理。

Stable Diffusion 即稳定扩散模型，在运作中我们把文本正向提示词（prompt）在大模型中进行输入，通过 VAE 编码器进行处理，经过一系列的运算后，再由 VAE 解码器进行输出，生成图像。

VAE 模型在后期制作的大模型中一般处于内置选项，无特殊要求，选择基本默认选项即可。当生成图像出现明显的画质低劣、颜色灰暗时，可以适当切换 VAE 模型来提高画面的清晰度。有无 VAE 模型的对比效果如图 3-1-18 所示。



无 VAE 效果图

加 VAE 效果图

图 3-1-18 有无 VAE 模型的对比效果

VAE 模型文件的格式有三种：大模型名称加上后面的“vae.pt”后缀格式、“ckpt”格式和“.safetensors”格式，其文件存储位置位于根目录下的“\models\VAE”文件夹中。

想要生成一张理想的图片，选对大模型至关重要。在以上所提的几种模型中，Checkpoint 模型是 Stable Diffusion 程序的基础模型。而使用扩展模型需要注意的是，在使用前一定要查看它的基础模型版本，以及是否需要触发词。没有对应的版本，或者没有按照要求使用，都有可能让使用的扩展模型无效化。合理运用、正确使用是活用扩展模型及提高图像效果的重要前提。

任务 3.2

应用模型

1. Checkpoint 模型应用

大模型的作用是控制风格，接下来以三类不同风格的大模型为例进行模型的使用介绍。这三种大模型分别是写实风格模型 majicmixSombre_v20、迪士尼风格模型 disneyPixarCartoon_v10、二次元风格模型 SmoothStep_1.0。

当我们下载到所需要的模型，并将其放置到 Stable Diffusion 根目录的文件夹中时，如果在操作界面的大模型切换菜单中没有找到对应的模型，可以单击右侧的刷新按钮，等待加载完成即可找到。随着模型数量的累积，如果直接在左上角的大模型菜单中进行选择，可能无法确定模型的具体风格，这时候我们可以在扩展模型中找到模型栏，此模型栏是大模型的切换页面的另一种呈现方式，它的优点是可以将输出的图片风格附加在名称上，增加辨识度。若需要切换不同图片风格的大模型，单击图片右上角的工具符号即可，如图 3-2-1 所示。

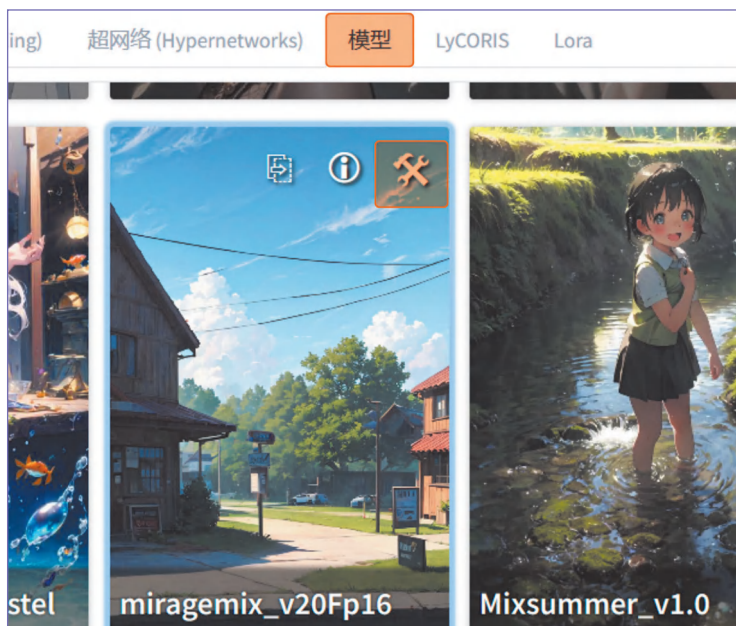


图 3-2-1 大模型图片切换

选择好大模型后，输入固定的提示词，设置统一的参数，就可以观察到这三种模型之间的差异。案例的提示词和参数设置如图 3-2-2 和图 3-2-3 所示。

依次切换三种大模型，分别利用 AI 生成图像（图 3-2-4），就可以看到风格各异的图像，这说明大模型的选择对于 AI 生成图像风格的影响是极其重要的。

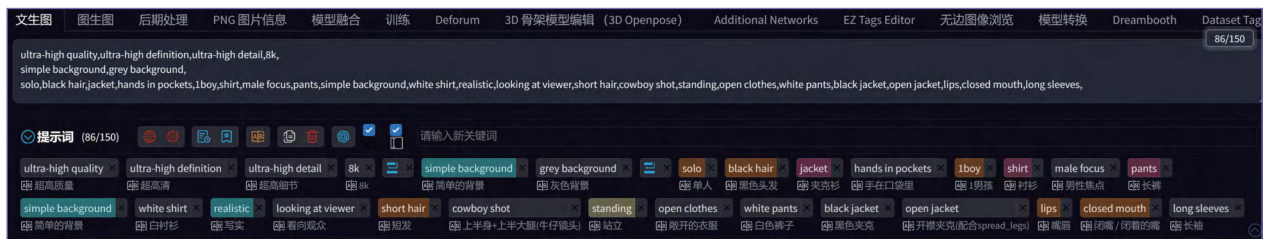


图 3-2-2 案例的提示词设置

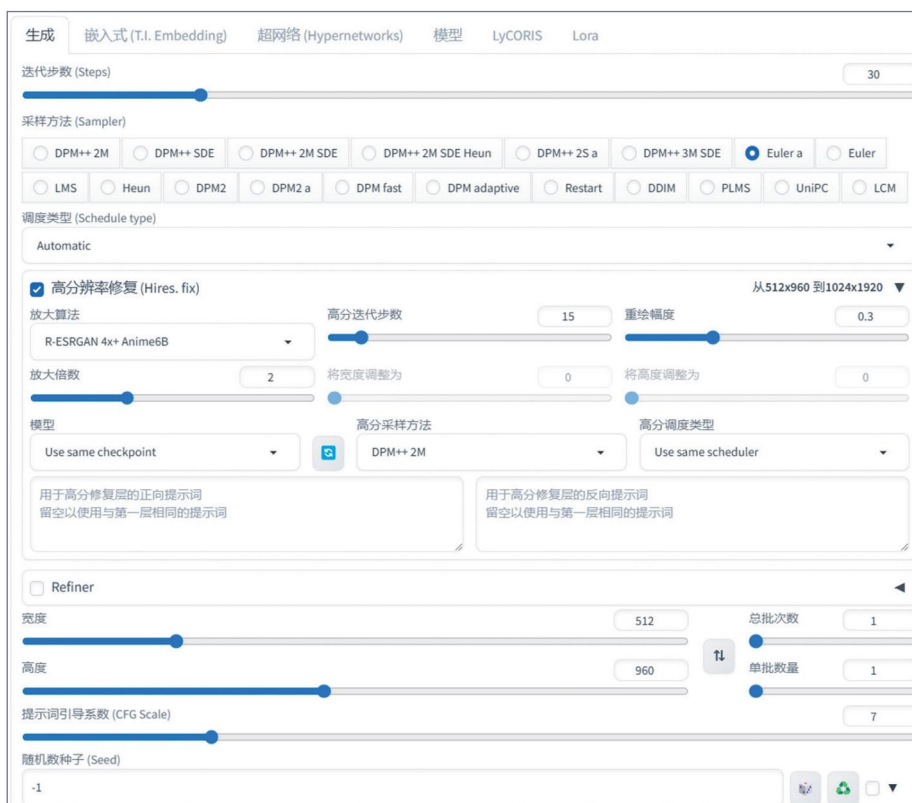


图 3-2-3 案例的参数设置



图 3-2-4 不同大模型生成的图像

2. Lora 模型应用



Lora 模型的使用效果有很多，使用方法也比较简单。它虽然被称为小模型，但是却不能单独使用，加载 Lora 模型时同样需要搭配 Checkpoint 模型来出图。当我们完成提示词的编写后，只需要进入扩展模型中，单击所要使用的 Lora 模型，相应的标签就会出现在提示词框中，如图 3-2-5 所示。



图 3-2-5 Lora 模型的应用标签

Lora 标签由三个部分组成，分别为标识项、Lora 文件名、权重。权重的范围值一般以 0 到 1 为起止位，权重代表 Lora 模型介入图像生成所占有的百分比权重，数值越大意味着 Lora 模型介入的比重就越大，其模型中参考图片的使用率就越高。由于 Lora 模型是通过少量图片进行训练获得的，因此当图像与 Lora 素材库中的元素差异过大时，就会出现严重的撕裂感和错误。通常 Lora 模型的作者会提供效果最佳的权重数值，因此在使用 Lora 模型进行图像生成时，尽量遵循指导参数进行设置、出图，而不是使用默认权重 1，这样才能获得效果更佳的 AI 图像。有的 Lora 模型在使用时，作者会提供触发词，我们需要添加触发词到提示词当中才能产生效果，因此使用 Lora 模型前就需要认真研读作者提供的使用说明书。

如图 3-2-6 所示，我们使用的是一个小型的 Lora 画风模型，可以看出去掉 Lora 模型生成的图像与添加 Lora 模型后生成的图像差别是非常大的，这也是为什么 Lora 模型在后期使用中越来越广泛。



图 3-2-6 Lora 模型使用前后对比

3. Embedding 模型应用

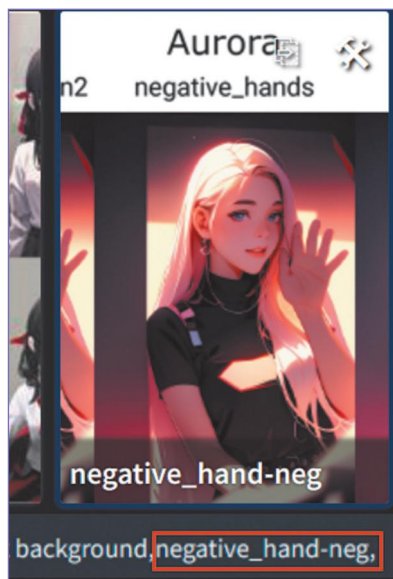


图 3-2-7 Embedding 模型使用示例

Embedding 模型的使用方法与 Lora 模型类似，在 Stable Diffusion “文生图”界面下，当我们写完提示词之后单击扩展模型，第一项就是 Embedding 模型，单击需要添加的 Embedding 模型，它的提示词标签就会与普通提示词一样出现在提示词框内。

需要注意的是，Embedding 模型通常是放置在反向提示词中来使用的，如果将反向提示词的 Embedding 模型放置在正向提示词中，生成的图像会变得千奇百怪。

如图 3-2-7 所示，选择一个手部修复的 Embedding 模型，当我们进行二次元图像生成时，经常会出现手部的结构错误，尤其是会出现手指缺少或多余的情况，该模型大概率可以修复这类问题。

与 Lora 模型不同，Embedding 模型的标签与普通提示词是一样的，所以增减权重的方式也与普通提示词的填写方式是一样的。

4. Hypernetworks 模型应用

Hypernetworks 模型作为一个注重风格演化的模型，在控制风格方面的表现也是相当出色。在扩展模型中选择相应的模型，可以发现在提示词中 Hypernetworks 模型的标签也是由三个部分组成，分别为标识项、Hypernetworks 文件名、权重。此外，Hypernetworks 模型的使用方式与注意事项也基本与 Lora 模型相同。Hypernetworks 模型使用示例如图 3-2-8 所示。



图 3-2-8 Hypernetworks 模型使用示例

Hypernetworks 模型只影响扩散模型运算中数据传输的权重参数，所以想用好 Hypernetworks 模型，需要注意以下几个方面：一是正确使用权重的范围值；二是选择与 Hypernetworks 模型相匹配的大模型，两个模型的风格差异不宜过大；三是使用作者给予的各项参数，这样通常会达到最好的出图效果，如果选择自行调整参数，出图的效果可能与理想的效果差距较大；四是如果模型作者有提供触发词，一定要使用所提供的触发词，否则无法达到理想效果。

如图 3-2-9 所示，分别是未使用 Hypernetworks 模型、使用 Hypernetworks 模型和使用 Hypernetworks 模型给予的触发词的图像效果。可以看到，越是想要得到与作者的出图相同的效果，就越需要按作者给予的参数进行调整，否则出图的差异性还是比较大的。



图 3-2-9 Hypernetworks 模型使用情况对比

5. VAE 模型应用

作为 Stable Diffusion 中的扩展模型的编码器，有的大模型自带 VAE 模型，无须加载，我们只需要在“外挂 VAE 模型”下拉菜单（图 3-2-10）中选择“无”或者“自动”即可。



图 3-2-10 “外挂 VAE 模型”下拉菜单

有的大模型不自带 VAE 模型，直接出图会出现颜色灰白、低质量的情况，此时使用官方提供的通用 VAE 模型才能达到比较好的出图效果。

判断是否内置 VAE 模型，首先可以阅读大模型作者的使用说明，也可以自己生成图片进行尝试，有的大模型会提供使用 VAE 模型的下载链接，如果想获得其他类型的 VAE 模型，可以在 Stable Diffusion 启动器的 VAE 模型列表（图 3-2-11）中搜索，也可以在网址中寻找，获取的途径有很多，可以自行探索。

模型管理

打开文件夹

Stable Diffusion 模型

嵌入式 (Embedding) 模型

超网络 (Hypernetwork)

变分自编码器 (VAE) 模型

LoRA 模型 (原生)

LoRA 模型 (插件)

计算全部新版短哈希

☒ 本地

☒ 远程

搜索模型...

文件名	模型名	作者	主页	旧短哈希	新短哈希	文件大小	创建日期
animevae.pt	NovelAI 原版 VAE	NovelAI		f458b5c6	f921fb3f29	784.69 MB	2022-11-21 11:35
final-pruned.vae.pt	NovelAI 原版 VAE	NovelAI		f458b5c6	f921fb3f29	784.69 MB	2023-07-13 14:41
vae-ft-mse-840000-em	Stable Diffusion ft-MSE	StabilityAI	https://huggingface.co/stabilityai/stable-diffusion-vae-ft-mse	223531c6	c6a580b13a	319.19 MB	2023-06-12 16:12

图 3-2-11 Stable Diffusion 启动器中的 VAE 模型列表

任务 3.3

使用 Stable Diffusion 进行绘制

在了解了 Stable Diffusion 的基础界面和模型种类及作用后，我们就可以开始利用 Stable Diffusion 绘制图像了。接下来将会使用文本生成图像、图像生成图像两种方式进行绘制，绘制过程中会用到 Lora 模型，从而获得更好的画面效果。本任务可以让大家更加快速地了解 Stable Diffusion 的奇妙之处。

1. Stable Diffusion 绘制基础

下面将使用 ReV Animated 大模型及 3D 渲染风格的 Lora 模型进行简单的绘制演示，这里大家可以根据自己的喜好去搭配合适的模型。

1) 在正向提示词框里填写如下正向提示词。

3dmm style,masterpiece, best_quality,ultra-detailed,epic,illustration,welcoming,Foxglove <lora:3DMM_V12:1.2>

2) 在反向提示词框里填写如下反向提示词。

tattooing,Neck decoration,collar,necklace,collar,badhandv4,paintings,sketches,(worst quality:2),(low quality:2),(normal quality:2),lowers,normal quality,((monochrome)),((grayscale)),skin spots,acnes,skin blemishes,age spot,(outdoor:1.6),manboobs,(backlight:1.2),double navel,muted arms,hused arms,neck lace,analog,analog effects,(sunglass:1.4),nipples,nsfw,bad architecture,watermark,(mole:1.5),EasyNegative

3) 单击右侧的“生成”按钮就可以生成效果图了，如图 3-3-1 所示。

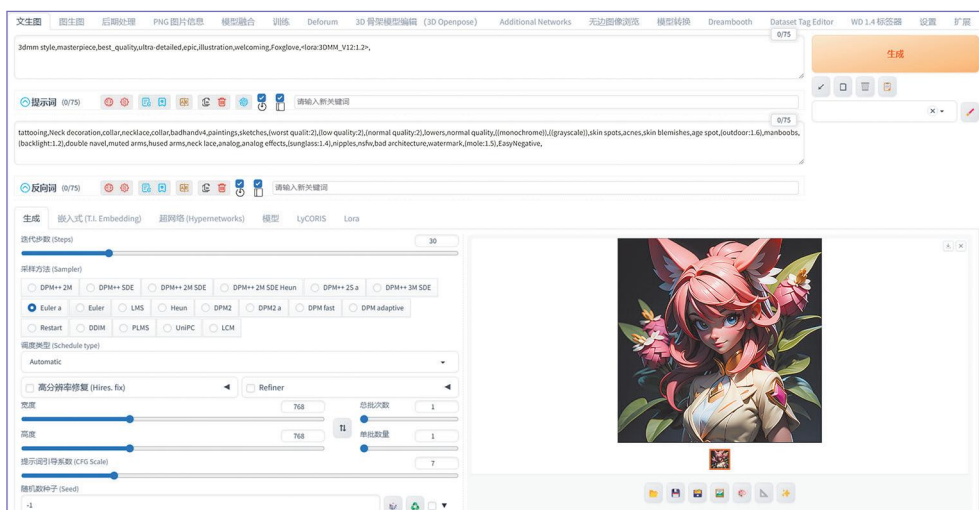


图 3-3-1 “文生图”效果图

提示

提示词全部都要用英文来填写。

2. Stable Diffusion 绘制参数调整

在利用 Stable Diffusion 进行绘制的时候，可以通过调节宽度和高度的参数来获得我们想要的尺寸比例。例如，想要一个竖屏的图（图 3-3-2），可以把参数里面的宽度和高度数值调整为“512”和“768”，就可以获得想要的竖屏画面；若想要一个横屏的图像（图 3-3-3），使用同样的方法把宽度和高度数值调换一下就可以了。

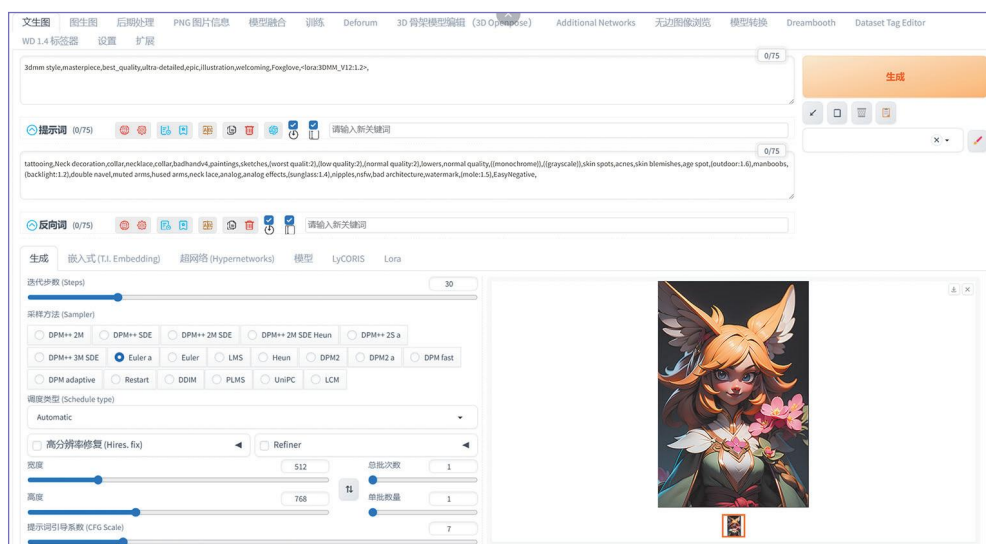


图 3-3-2 竖屏的图像

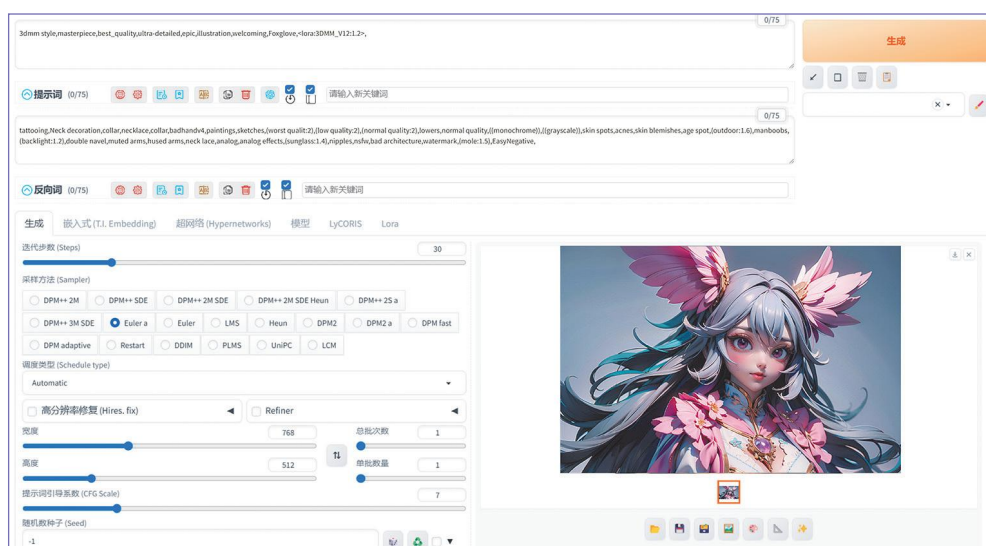


图 3-3-3 横屏的图像

下面讲解参数栏里面的各项参数的具体作用。

（1）采样方法

采样方法（Sampler）就是 Stable Diffusion 进行 AI 绘画时所使用的算法，其选择菜单如图 3-3-4 所示。常用的采样方法有三种，分别是 Euler a、DPM++ 2S a Karras、DDIM。其他的采样方法大家可以自主探索其作用，下面主要为大家介绍一下这三种常用的采样方法。

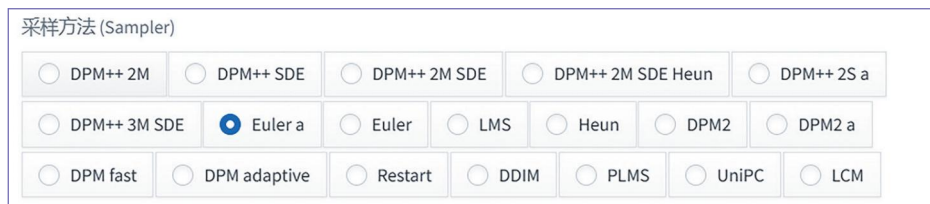


图 3-3-4 采样方法选择菜单

1) Euler a。Euler a 是一种用于控制时间步长大小的可调参数。适当的 Euler a 值能够绘制出细节和纹理，但如果值太大会导致出现过度拟合、生成的图像出现噪点等情况。在高细节图中增加采样步数可能会导致不可控突变（如人物脸部扭曲、细节扭曲等）。该采样方式比较适合用于 ICON（图标）的设计、二次元类的图像生成及小场景的绘制。

2) DPM++ 2S a Karras。DPM++ 2S a Karras 采样方法的特点是可以提高图像的品质，该采样方法在每个时间步长中执行多次操作，同等分辨率下细节会更多，比如可以在较小的尺寸下绘制角色的全身像，但代价是采样速度更慢。该采样方式比较适合用于写实类的人像生成，或者对复杂的场景、图形进行细节的刻画。

3) DDIM。DDIM 采样方法可以快速生成高质量的图像，相比其他采样方法具有更高的效率，想尝试超高步数时可以使用，随着步数增加可以叠加细节。该采样方式同样适合用于写实类的人像生成，或者对复杂的场景、图形进行细节的刻画。

（2）采样迭代步数

生成图片的采样迭代步数（Sampler）每多一次迭代就会让 Stable Diffusion 生成的图像更加接近提示词所描述的文字意思。当然，更高的步数需要花费更多的计算时间，同样，迭代步数越大，图像细节也越多。默认迭代步数是 20 步，一般情况下会将数值调整到 30 步，30 步之后生成图像就不会有太大的变化。如果计算机的显卡和 CPU 配置较好，数值可以再往上调。

！ 提示

推荐迭代步数为 30 步左右。

（3）图片分辨率

在 Stable Diffusion 设置图像尺寸的地方，长宽数值的单位都是像素。通常如果你的显卡是入门级的，建议用默认的 512 × 512 的图像尺寸即可，如果你的显卡性能比较好，可以增加尺寸，当前 Stable Diffusion 版本（Stable Diffusion v1.5）最高支持 2048 × 2048 的尺寸。

（4）生成批次和每批数量

每次单击“生成”按钮时，Stable Diffusion 会根据生成批次和每批数量（Batch&Num）的设置来

决定输出多少张图片，图片的数量 = 生成批次 × 每批数量。生成的图片张数越多，所消耗的时间越长，特别是当显卡版本较低时，消耗的时间就会更长。

！ 提示

建议每次绘制的时候调节生成批次。

(5) 提示词相关性

提示词相关性 (CFG Scale)，即图像与提示词的匹配程度。增大提示词相关性的参数 (提示词引导系数) 将使图像更接近提示词的文本含义，但它在一定程度上也降低了图像质量，建议将参数设置在 7 ~ 11。提示词引导系数设置成 7 和 1 的对比，如图 3-3-5 所示。

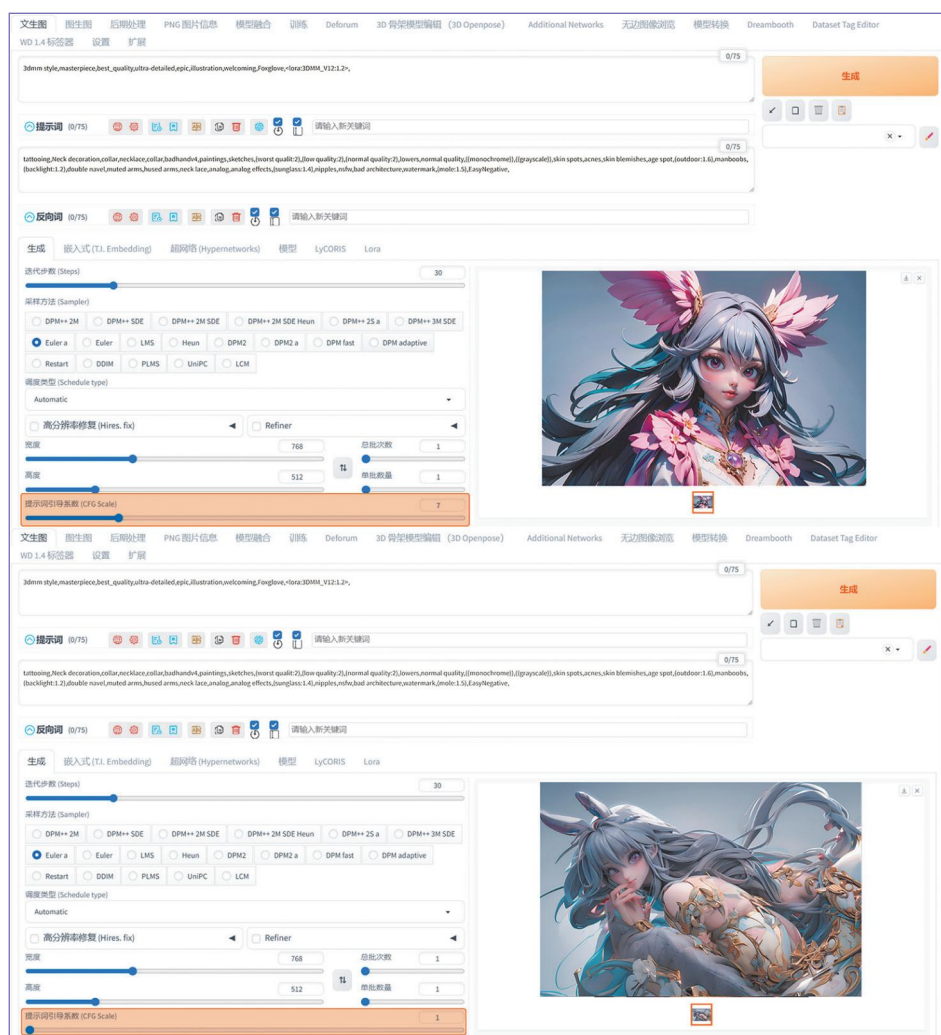


图 3-3-5 提示词引导系数设置成 7 和 1 的对比

3. Stable Diffusion 绘制分辨率提高

若在生成图片的过程中，发现图片并不是很清晰，甚至出现很模糊的情况，这个时候就可以使用

高清修复这一功能，该功能可以在生成图片的过程中，直接将图像进行放大，直至画面变得清晰，但出图的时间也会随之变慢。

在勾选“高分辨率修复（Hires. fix）”的时候会出现一些参数，如图 3-3-6 所示。在“放大算法”下拉菜单中选择“R-ESRGAN 4x+”或“R-ESRGAN 4x+ Anime6B”，这两种放大算法的图像生成效果通常会好一些。“高分迭代步数”的数值通常可以在 20 ~ 30 之间；“重绘幅度”数值范围一般在 0.25 ~ 0.5；“放大倍数”正常情况下可以设置为 2，如果生成的图像效果不好，可以把这个数值调小一些，比如 1.3、1.5 这样的数值。单击“生成”按钮，就会得到一张长、宽各放大两倍的图像，放大图像前后尺寸对比如图 3-3-7 所示。

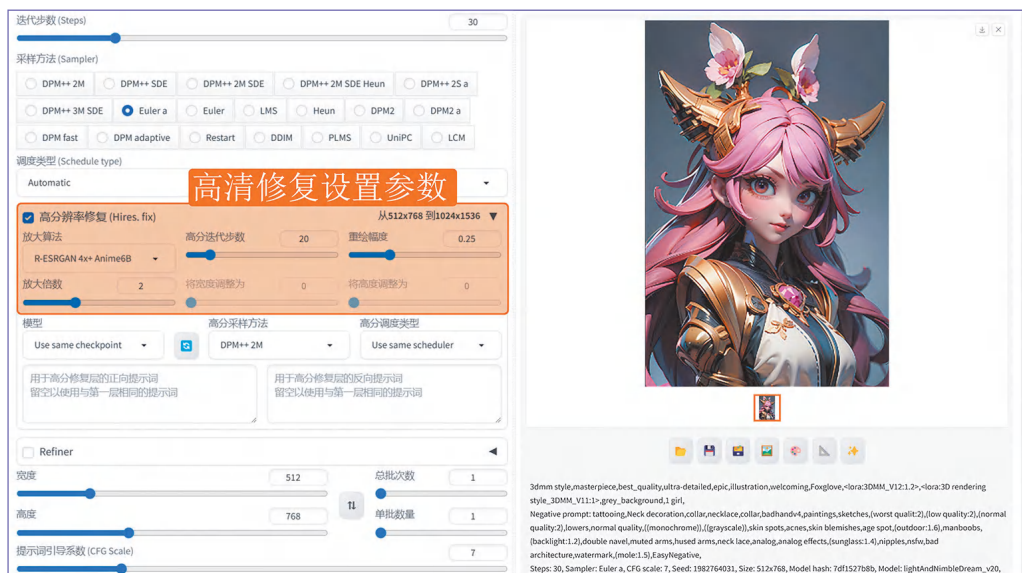


图 3-3-6 高清修复设置参数



图 3-3-7 放大图像前后尺寸对比

4. Stable Diffusion 图片信息复刻

通常大家在模型网站上会看到很多画面精美的图片，通过 Stable Diffusion 图片信息复刻功能，可以将该图片的信息完整地复刻到本地计算机上。方法步骤如下。

1) 在网站上找到 AI 模型生成的图片。先单击图像右下角的小符号，再单击“Copy Generation Data”，如图 3-3-8 所示。

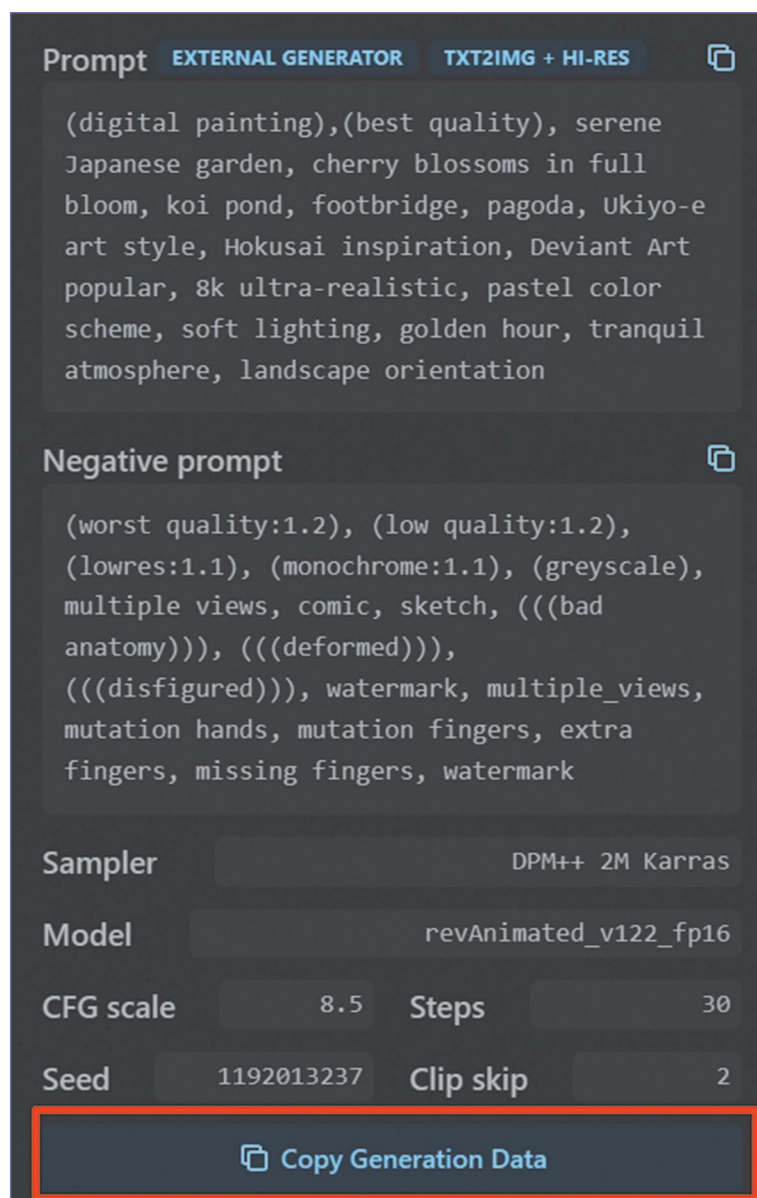


图 3-3-8 复制图片数据

2) 将图像信息复制到正向提示词栏，然后再单击“生成”按钮下面的小箭头，进行自动匹配数据，再单击“生成”按钮，就可以完成图片信息的复刻了，如图 3-3-9 所示。



图 3-3-9 复刻图像

5. Stable Diffusion “随机种子” 控制

“随机种子”（seed）是指图片在生成时的随机性，它初始化了 Diffusion 算法起点的初始值。如果想每次得到的图片都不一样，那么使用 -1 即可；如果觉得某张图片不错，可以通过查看生成图片下的文字信息，根据它的随机种子值来指定。

每生成一张图片，下面的文字信息就会有一个 seed 值，单击绿色的小按钮，这个数值就会自动粘贴过来，再去生成一次就会得到与原图一样的图片，如图 3-3-10 所示。我们可以在这张图上进行更加细致的调整。

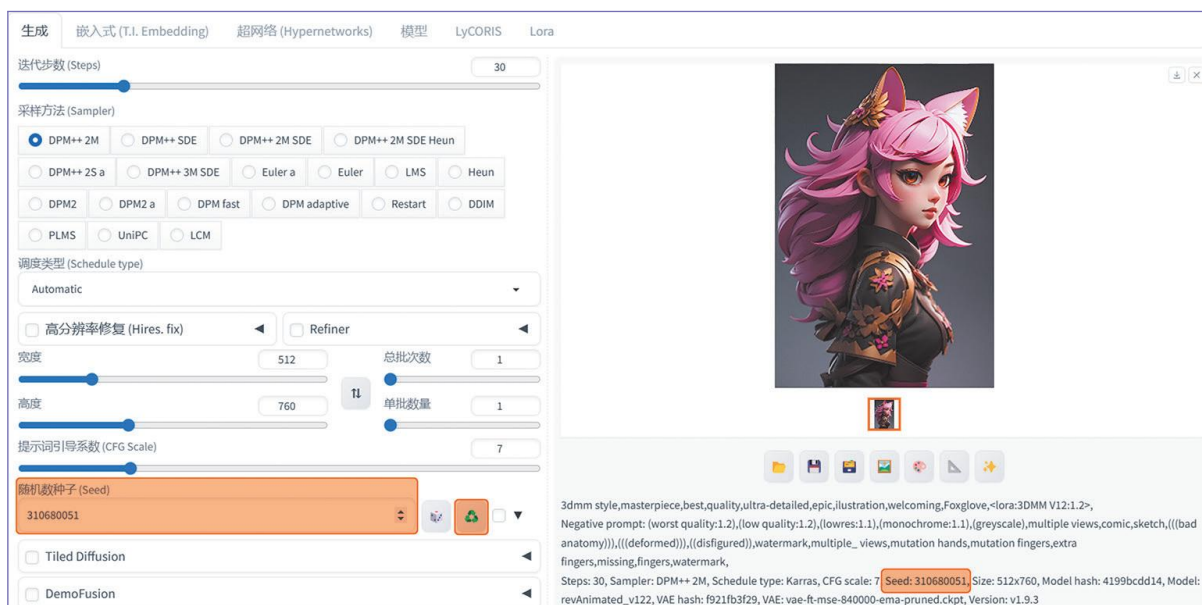


图 3-3-10 “随机种子”控制

在制作图像的时候如果这幅图像错过了，也没有进行保存，这个时候可以单击图像下面文件夹形状的小按钮（图 3-3-11），打开历史生成图像的存储文件夹（图 3-3-12），之前生成的所有图像都保存在这些文件夹里。

提示

时间久了计算机内存不够时，可以把历史生成图像清理一下，这样计算机运行会更加顺畅。

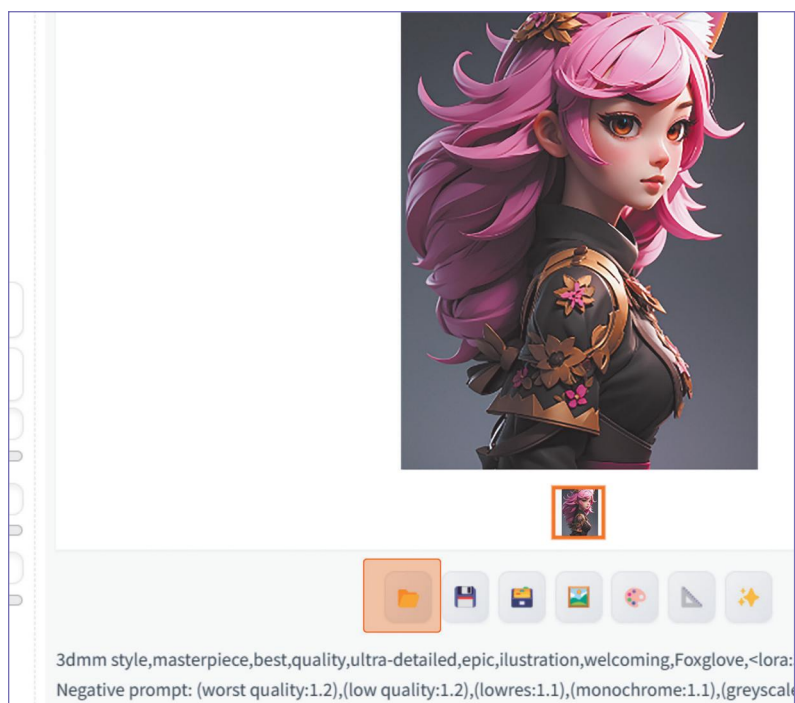


图 3-3-11 历史生成图像的存储文件夹按钮



图 3-3-12 历史生成图像的存储文件夹位置

找到“PNG 图片信息”（图 3-3-13）这一栏，把之前的图像拖曳到图像放置的位置，就可以得到想要的信息。甚至还可以直接单击“发送到 文生图”按钮，把这些信息直接发送到“文生图”页面，单击“生成”按钮就可以得到之前的图像了，如图 3-3-14 所示。

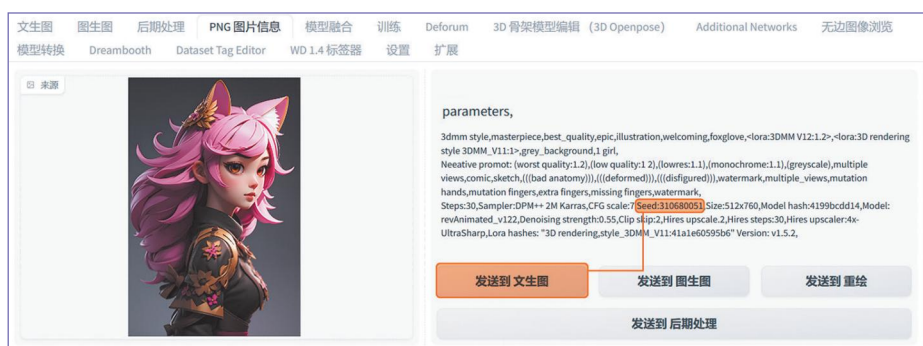


图 3-3-13 “PNG 图片信息”栏

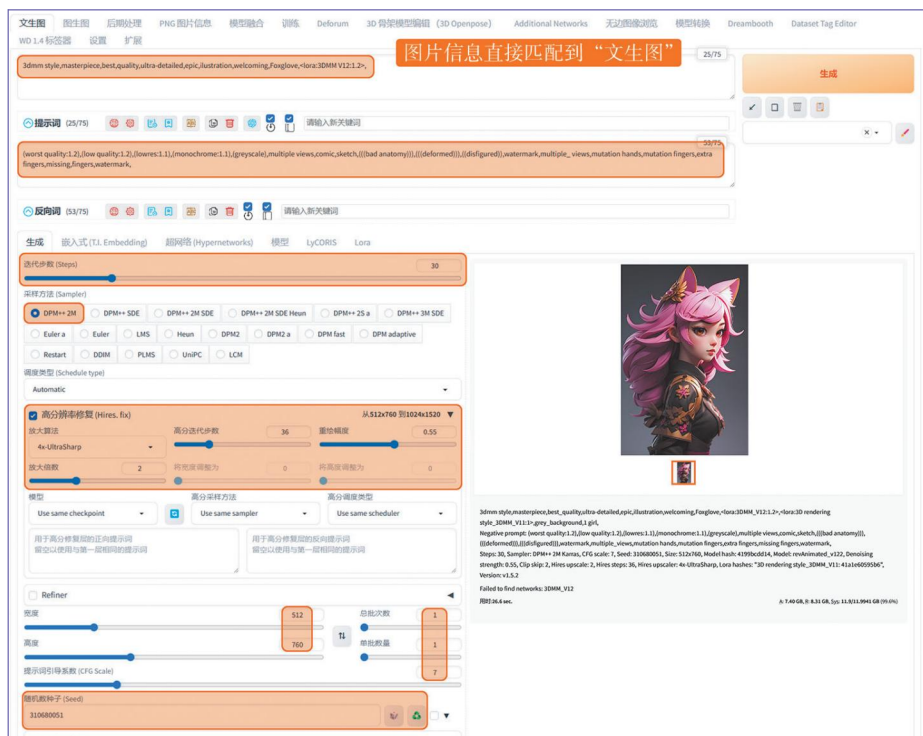


图 3-3-14 “文生图”重新生成

拓展任务

正确地选择大模型是生成符合要求的图片的基础，如果选择皮克斯风格的大模型，那么无论如何也生成不了水墨风格的图片。所以，选择合适的大模型至关重要。这也同样考验设计者是否可以正确地区分不同的艺术风格。这是一个合格的设计者应具备的基本艺术素养。

为了熟悉不同大模型的风格，巩固本模块的所学知识，请完成以下任务（也可自己编写提示词进行测试）。

使用同一套提示词，测试不同风格的大模型。

1) 提供正向提示词。

8k,intricate,elegant,highly detailed,majestic,digital photography,cowboy shot,surreal painting,broken glass,(masterpiece,sidelighting,finely detailed beautiful eyes:1.2),HDR,Black eyes,gongbihua,masterpiece,a painting,1girl,((upper body)),((background is full of flowers and leaves)),pale skin,long eyelashes,(Chinese Traditional hanfu:1.3),look at viewer,face front,(hair scrunchie:1.2)

2) 提供反向提示词。

Deep,Negative,V1.x,EasyNegative,Deep,Negative,V1.x,EasyNegative,((nsfw)),sketches,tattoo,(beard:1.3),(EasyNegative:1.3),badhandv4,(Teeth:1.3),(worst quality:2),(low quality:2),(normal quality:2),lowers,normal quality,facing away,looking away,text error,extra digit,fewer digits,cropped,jpeg artifacts,signature,watermark,username,blurry,skin spots,acnes,skin blemishes,bad anatomy,fat bad feet,cropped,poorly drawn hands,poorly drawn face,mutation,deformed tilted head,bad anatomy,bad hands,extra fingers,fewer digits,extra limbs,extra arms,extra legs,malformed limbs,fused fingers too many fingers,long neck,cross-eyed,muted hands,bad body,bad proportions,bad gross proportions,text error,missing fingers,missing arms,missing legs,extra digit,extra arms,extra leg,extra foot,missing fingers

