

Stable Diffusion 的界面

本章学习要点：

- 掌握 Stable Diffusion WebUI 基本功能和操作流程。
- 掌握种子值在图像生成中的作用及其设置方法。
- 掌握采样方法、迭代步数以及生成批次与单批数量对生成图像的影响。
- 掌握输出分辨率的设置及其对图像质量的影响。
- 掌握高清修复功能的相关操作。
- 掌握脚本功能的应用。



3.1 功能选项

Stable Diffusion 的每种选项卡都具备独特功能，可以根据需求灵活选用。下面介绍各选项卡的功能，并举例说明一些常用功能的作用与使用方法。因为版本与系统的原因，选项名称可能会翻译不一致，但其功能都是相同的。

1. 文生图功能。此功能能够通过文字提示词生成相应的图像，其界面如图 3.1 所示。



图 3.1 文生图界面

文生图案例如下。

步骤 1：选择一个写实风格的大模型，如图 3.2 所示。



图 3.2 选择大模型

步骤 2: 在正向提示词对话框输入相关描述“一只猫”,对应的英文提示词为“1 cat”,如图 3.3 所示。

步骤 3: 单击生成按钮,生成图像,如图 3.4 所示。通过提示词引导,系统生成了一只猫的图像。



图 3.3 设置提示词

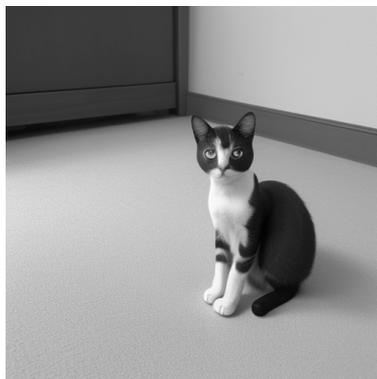


图 3.4 生成一只猫的图像

2. 图生图功能。此功能用于结合图像和输入的提示词派生出新的图像,其界面如图 3.5 所示。



图 3.5 图生图界面

图生图案例如下。

步骤 1: 选择“麦橘写实”大模型,如图 3.6 所示。

步骤 2: 输入相关描述“一个女孩”,对应的英文提示词为“a girl”,如图 3.7 所示。



图 3.6 选择大模型



图 3.7 设置提示词

步骤 3: 导入一张女孩图像到图生图界面,以这张图像为生成图的参考,如图 3.8 所示。

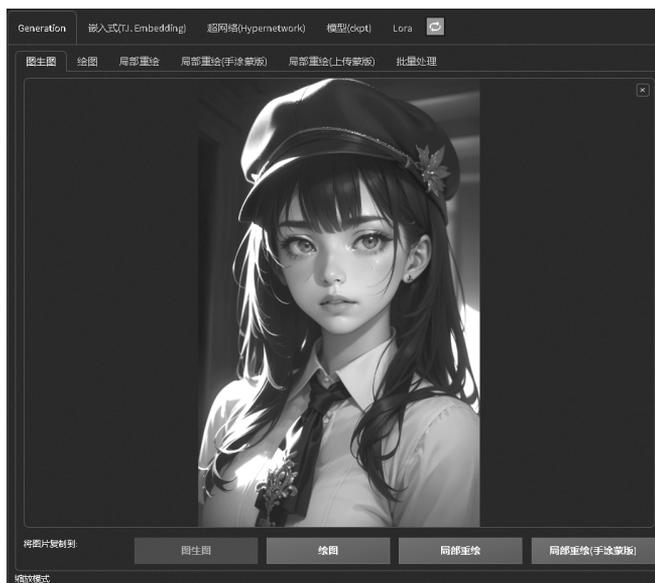


图 3.8 导入女孩图像到图生图界面

步骤 4: 选择重绘幅度选项,调整重绘幅度数值至“0.4”,如图 3.9 所示。



图 3.9 重绘“重绘幅度”

步骤 5: 单击生成按钮,生成图像,图生图功能可以利用参考图像和提示词结合生成新的图像,如图 3.10 所示。

3. 后期处理功能。此功能用于提升图像分辨率,即图像放大功能。该功能可以批量处理放大图像,其界面如图 3.11 所示。



图 3.10 生成相同内容但不同风格的女孩图像



图 3.11 后期处理界面

后期处理案例如下。

步骤 1：导入一张人物图像到后期处理界面，如图 3.12 所示。



图 3.12 导入女孩图像到附加功能界面

步骤 2：设置缩放比例为 2（数值越大，图像放大的尺寸越大），选择放大算法为“R-ESRGAN 4x+”模式，如图 3.13 所示。R-ESRGAN 4x+ 是增强型超分辨率生成对抗网络（enhanced super-resolution generative adversarial network, ESRGAN）的一个改进版本，旨在实现实时的高分辨率图像生成。其中“R”代表“实时（Real-Time）”，4x 指图像放大的倍数。

步骤 3：生成图像，得到一张画面尺寸更大的图像。图像放大完成，如图 3.14 所示。



图 3.13 放大界面



图 3.14 画面尺寸放大后的女孩图像

4. 图片信息功能。此功能可以导入由 Stable Diffusion 生成的图像，图片信息界面会反映图像生成时的具体参数，如使用的大模型、LoRA 模型以及提示词等信息，其界面如图 3.15 所示。



图 3.15 图像信息界面

图片信息案例如下。

步骤 1: 导入图像到图片信息界面,弹出了该图像的生成信息参数,如图 3.16 所示。

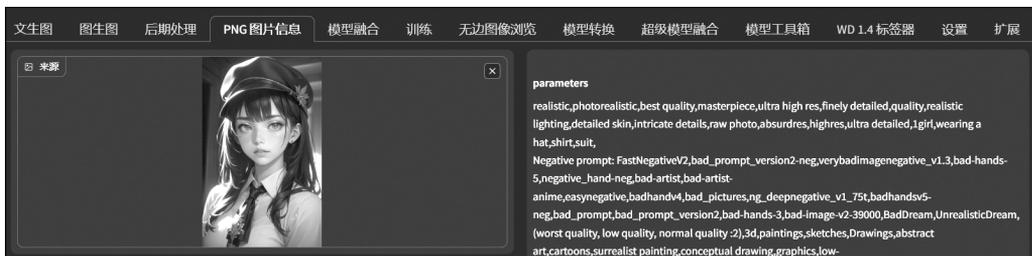


图 3.16 导入女孩图像到图像信息界面

步骤 2: 检查信息,弹出的信息参数中记录了正向提示词信息、反向提示词信息、迭代步数、采样方法、种子值、生成尺寸、模型以及插件的使用信息,如图 3.17 所示。



图 3.17 女孩图片信息内容

5. 模型融合功能。此功能用于模型的整合融合操作,其界面如图 3.18 所示。模型融合案例如下。

步骤 1: 进入模型融合界面,在模型 A 下拉菜单栏中选择“Counterfeit”大模型,在模型



图 3.18 模型融合界面

B 下拉菜单栏中选择“麦橘写实”大模型，在模型 C 下拉菜单栏不选择任何模型，在输出模型文件名中输入“一个新的模型”，这个名称代表融合后的新的模型名称。设置倍率为 0.3 (倍率(M)为模型 B 所占比例)，融合算法在一般情况下选择加权和，模型格式可以选择 ckpt。在“从...复制配置文件”中选择 A、B 或 C，如图 3.19 所示。



图 3.19 设置模型融合界面

步骤 2：单击融合按钮，等待合并结果，完成模型的合并后，单击大模型  刷新按钮，会在大模型下拉列表中找到新融合的模型，如图 3.20 所示。

一个新的模型.ckpt [1ee5920fe1]

图 3.20 新融合的模型

6. 训练功能。此功能是一个允许进行模型训练的功能选项卡，其界面如图 3.21 所示。
7. 图像浏览功能。此功能用于图像查阅和浏览，其界面如图 3.22 所示。
8. 模型转换功能。此功能可以将模型转换为 Checkpoint 格式，其界面如图 3.23 所示。



图 3.21 训练界面



图 3.22 图像浏览界面



图 3.23 模型转换界面

9. WD1.4 标签器功能。此功能是从图像反推出一些关键提示词。因为版本与翻译的不同,该功能在其他版本也被翻译为 Tagger(反推),其界面如图 3.24 所示。

WD1.4 标签器案例如下。

步骤 1: 导入图片到 WD1.4 标签器界面,单击“反推”按钮开始反推,AI 会分析画面,并根据画面生成与之匹配的提示词,如图 3.25 所示。

步骤 2: 检查挑选提示词,反推功能推导出的提示词是检索得到的结果,不可以直接用于生成创作,但是这一方法提高了组织与撰写提示词的效率。

10. 设置功能。此功能用于配置 Stable Diffusion 的各项设定。比如设定文生图的输出目录、图生图的输出目录等,其界面如图 3.26 所示。



图 3.24 WD1.4 标签器界面



图 3.25 WD1.4 标签器根据图片生成的提示词



图 3.26 设置界面

11. 扩展功能。此功能主要处理插件的安装及更新,其界面如图 3.27 所示。



图 3.27 扩展界面

3.2 采样方法

Stable Diffusion 采样通过重复执行去噪过程逐步生成图像,这些图像会与文本提示进行比较。根据这一比较,算法逐渐调整添加到图像中的噪声,并不断重复此过程,直到生成的图像与文本提示的描述相符为止。

根据速度、提示解读的准确度以及最终图像质量等因素,当前推荐的采样方法主要有以下几种: Euler a、DPM++2M Karras、DPM++2S a Karras 和 DPM++SDE Karras。其中,Euler a 是默认采样器,表现出极佳的平衡性,它能使图像展现流畅的颜色过渡和边缘效果。DPM++2M Karras 多被应用于卡通渲染,其运行速度很快。DPM++2S a Karras 和 DPM++SDE Karras 则更适用于写实风格的渲染。常用的采样方法如图 3.28 所示。

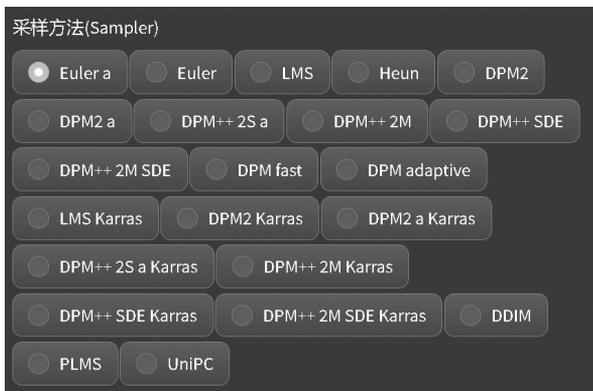


图 3.28 采样方法界面

在控制面板中,所有采样方法都会显示出来。为了界面的简洁,可以将部分不使用的采样方法隐藏起来。方法是先选择设置采样方法参数选项卡,然后在用户界面中勾选需要隐藏的采样方法。完成上述步骤后,重新启动 Stable Diffusion,就可以看到修改后的界面。

3.3 迭代步数

Stable Diffusion 是一种独特的图像生成技术,它从一个充满噪点的画布开始,采用逐步的去噪过程,以实现最优的图像效果。去噪过程的次数,也就是采样迭代步数(Steps)参数,是控制这一过程的关键。通常,去噪步骤越多,产生的图像质量越高。然而,在大多数情况下,Steps 的默认值是 20,根据实践经验,这个步数已经足够生成各种类型的图像,无论是风景、肖像还是抽象图案,都能得到满意的结果。图 3.29 为采样迭代步数界面。



图 3.29 采样迭代步数(Steps)参数界面

3.4 生成批次与每批数量

在图像生成过程中,生成批次和每批数量是两个关键的概念。这两个参数可以设置一次性生成多少张图像,并决定每次单击生成按钮时系统将要制作的图像总数量。通过将总批次数与每批数量相乘可以得知这个总数。

如果想在相同的总图像数量下让生成速度更快些,可以考虑增加每批数量。这样可以更有效地利用显卡的计算资源,节省时间。但过高的每批数量可能使显存超出负荷,导致图像生成失败。

相反,如果调高生成批次的次数,虽然可能使生成速度减慢,却能避免显存超负荷的风险。毕竟,只要时间充足,增加生成批次就能不断生成图像,直到所有的输出都完成。因此,调整这两个参数的策略,其实就是在空间(显存)和时间之间作选择,是典型的以空间换时间或以时间换空间的例证。

综上所述,可以通过调整生成批次和每批数量在速度与显存之间找到最佳平衡,在确保图像生成成功的同时提高图像生成效率。图 3.30 为生成批次与每批数量界面。



图 3.30 生成批次与每批数量界面

3.5 输出分辨率

Stable Diffusion 的图像分辨率关乎图像内容的构成和细节的展现。具体来说,画幅大小决定了画面的信息量。大的画面尺寸有足够空间表现构图中的各种细节,比如脸部、饰品、复杂的纹样等。而如果画幅太小,就无法充分展示这些细节。

随着画面尺寸的扩大,AI 模型也更倾向于在图像中塞入更多的内容。一般而言,大部分的 Stable Diffusion 模型在 512×512 像素下进行训练,未来,在 768×768 像素下进行训练将变得常规化。因此,当输出尺寸较大时,比如 1024×1024 像素,AI 会试图在图像中嵌入 2~3 幅图像的内容,这可能会导致出现人物肢体拼接、多角度等问题。因此,如果需要较大的画面尺寸,就需要提供更多的提示词,以给 AI 一个清晰的指示。而如果提示词较少,可以选择先生成较小的图像,再通过附加功能将其放大。

画幅或分辨率的设置,牵涉图像的细节表现和内容呈现,应根据实际需要和模型的训练情况综合决定。最后的输出分辨率与计算机性能息息相关,如果超过计算机运算能力上限的图像尺寸,会导致生成失败。图 3.31 为生成画面尺寸设置界面。



图 3.31 生成画面尺寸界面

3.6 种子的概念

种子分为“随机种子”和“固定种子值”。“随机种子”代表初始随机噪声,不同的随机种子将生成不同的图像。同时,Stable Diffusion 提供了一个选项设置种子值为“-1”,这表示 Stable Diffusion 环境中可以选择任何一个随机值作为初始种子。这就好比在照片上施加不同的滤镜,每个滤镜都会产生独特的视觉效果。改变随机种子,实质上是在探索一种无限的艺术可能性。“固定种子值”的原理是使随机数生成器的输出在多次运行中保持不变,这可以使每次生成的图像在一定程度上具有相似性。图 3.32 为种子界面。



图 3.32 种子界面

提示: 按钮  代表固定种子值。

3.7 高清修复功能

高清修复则用于增加图像分辨率,以显著提高图像的清晰度。然而,由于显存和显卡型号的限制,直接通过高清修复生成的高分辨率图像(如 2048×2048 像素)可能导致显卡崩溃。因此,使用此功能时应根据显存配置调整。如果需要放大图像的分辨率,建议根据实际配置谨慎操作,以确保系统的稳定性和图像的高质量输出。

勾选高清修复选项后,系统会弹出一个新的面板。这里需要根据需求选择放大算法。如果生成的图像偏向写实效果,建议选择“R-ESRGAN 4X+”模式,如果生成的图像偏向二次元效果,建议选择“R-ESRGAN4x+ Anime6B”模式的放大算法。其他参数通常可以使用默认设置。图 3.33 为高清修复界面。



图 3.33 高清修复界面

3.8 生成图像的保存、下载、转绘以及后期处理

输出图像框下方有多种图标按钮,它们具有重要的实用价值,能够对图像进行有效的管理和操作。如图 3.34 所示,它们包括“打开图像输出目录”“保存图像到指定目录”“保存包含图像的 ZIP 文件到指定目录”“发送图像和生成参数到图生图”“发送图像和生成参数到图生图局部重绘”以及“发送图像和生成参数到后期处理”。

“打开图像输出目录”按钮(图 3.34 中①)用于打开存放图像的目录,方便用户快速访问生成的图像文件。“保存图像到指定目录”按钮(图 3.34 中②)用于保存当前的工作成果,即不仅将图像写入预定的目录,还将生成的参数数据保存至 CSV 文件中,以便后续的分析 and 记录。“保存包含图像的 ZIP 文件到指定目录”按钮(图 3.34 中③)不仅可以压缩并保存文件,还进行文件的下载和传输,提升了文件管理的便利性。

此外,“发送图像和生成参数到图生图”按钮(图 3.34 中④)可以立即将图像和提示词发送到“图生图”界面。而“发送图像和生成参数到图生图局部重绘”按钮(图 3.34 中⑤)则用于将图像和提示词发送到“局部重绘”界面进行修改,使用户能够对图像局部进行更精细的调整。最后,“发送图像和生成参数到后期处理”按钮(图 3.34 中⑥)则是将图像和提示词发送到“后期处理”选项卡,以进一步放大图像分辨率,提升图像的清晰度和细节表现。这些功能按钮提供了全面且高效的图像管理和处理手段,确保各个操作步骤的顺畅和高效。

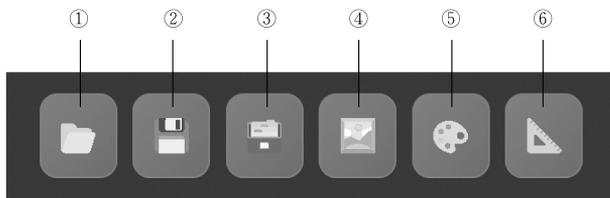


图 3.34 生成图像的保存、下载、转绘以及后期处理按钮

3.9 提示词控制区

在提示词控制区内,位于生成按钮下方的按钮控制选项如图 3.35 所示。它们提供了多种功能,便于用户对提示词进行管理和操作。以下是详细说明。

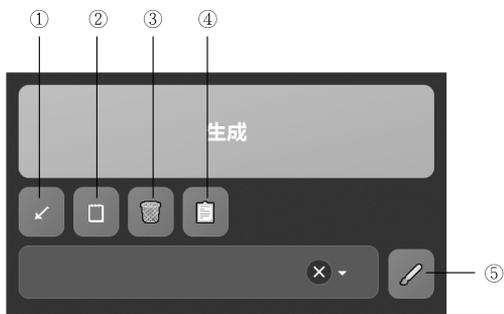


图 3.35 提示词控制区

“从提示词或上次生成的图片中读取生成参数”按钮(图 3.35 中①)的功能是从提示词中自动提取生成参数。如果当前提示词为空,它将从上次的生成信息中读取参数。该按钮允许用户将包含正面、负面、采样器、步数、模型等信息的整段生成信息全部粘贴到提示词区域,单击后会将对对应信息填入相应位置,并删除多余内容。

“从提示词或上次生成的图片中读取生成参数(对话框)”按钮(图 3.35 中②)的功能类似前者,但会弹出对话框,让用户设置更多的选项。

“清空提示词内容”按钮(图 3.35 中③)用于迅速清空当前提示词区域的所有内容,方便

用户重新输入新的提示词或进行其他操作。

“将所有当前选择的预设样式添加到提示词中”按钮(图 3.35 中④)则用于将用户当前选择的所有预设样式快速添加到提示词中,避免手动添加的烦琐操作,提高工作效率。

“编辑预设样式”按钮(图 3.35 中⑤)提供了一个界面,允许用户对已有的预设样式进行编辑和管理。通过该功能,用户可以根据不同的需求创建、修改或删除预设样式,更好地控制生成图像的风格和细节。这些功能按钮提供了便捷和高效的提示词管理方法,确保生成过程的顺畅和高效。

从提示词或上次生成的图片中读取生成参数的案例如下。

步骤 1: 导入一张 Stable Diffusion 生成的图像进入图片信息面板中,如图 3.36 所示。

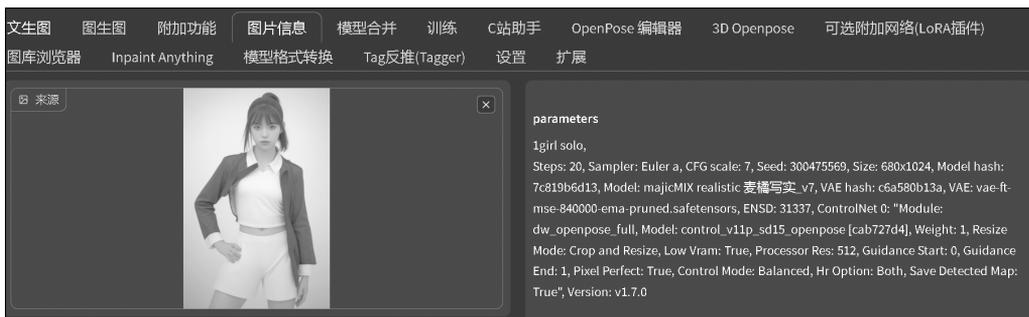


图 3.36 图片信息面板

步骤 2: 复制并粘贴生成数据到文生图界面正向提示词对话框,如图 3.37 所示。



图 3.37 复制图像全部信息到提示词对话框

步骤 3: 单击  按钮,自动提取生成参数,如图 3.38 所示。

步骤 4: 检查生成参数,在实际使用过程中,该功能可以还原生成参数,但并不能百分之百还原之前的生成过程。例如,生成数据中使用了 ControlNet 插件,该功能可以还原插件的参数设置,但不能导入参考图片,ADetailer 插件也不会自动开启,这些都需要手动再次调整。

编辑预设样式案例如下。

步骤 1: 输入要保存的提示词,并给提示词组合起一个名字,如图 3.39 所示。

步骤 2: 单击“保存”按钮。完成提示词预设存储,切换到文生图界面,新的提示词预设就出现在下拉列表中,如图 3.40 所示。通过这种方式可以预先保存常用的质量提示词组



图 3.38 自动提取生成参数



图 3.39 提示词预设面板

合、反向提示词组合,提高撰写提示词的效率。



图 3.40 预设提示词

3.10 脚本

文生图界面与图生图界面都有脚本选项,不同的是图生图界面的脚本选项功能更多一些。下面主要讲解脚本中的提示词矩阵、从文本框或文件载入提示词、X/Y/Z 图表以及 Ultimate SD upscale 图像放大功能。



3.10.1 提示词矩阵功能

提示词矩阵是一个在 Stable Diffusion 中生成图片效果检测的工具,它根据提供的提示词(即描述图片内容的提示词)来创建对比图像,其界面如图 3.41 所示。



图 3.41 提示词矩阵界面

提示词矩阵由基础提示词和变化提示词两部分组成。基础提示词是每张图片都包含的固定内容,它们位于第一个竖线(|)之前。变化提示词则位于第一个竖线(|)之后,并且它们通过竖线(|)分隔。基础提示词和变化提示词共同构成了提示词矩阵,既包含恒定的元素,又具有一定的灵活性。这种结构可以精准地描述图片内容,同时保留一定的多样性和变化性。

提示词矩阵案例如下。

步骤 1: 输入提示词矩阵: a man | blonde hair | black clothes。

步骤 2: 生成图像,提示词矩阵一次生成了 4 张图像。通过观察这些生成图像发现,图 3.42 只包含基础提示词(一个男人)。图 3.43 包含基础提示词和第一个变化提示词(一个男人,金色头发)。图 3.44 包含基础提示词和第二个变化提示词(一个男人,黑色衣服)。图 3.45 则包含基础提示词及所有变化提示词(一个男人,金色头发,黑色衣服)。生成结果如图 3.42、图 3.43、图 3.44 和图 3.45 所示。



图 3.42 一个男人图像



图 3.43 金发男人图像

通过该案例,可以清晰地看到基础提示词和变化提示词是如何组合,通过不同的变化提示词来生成不同的图像描述的。这种矩阵形式使得描述既能保持一定的固定内容,又能灵活地添加变化元素,从而生成多样化的图像内容。



图 3.44 黑衣服的男人图像

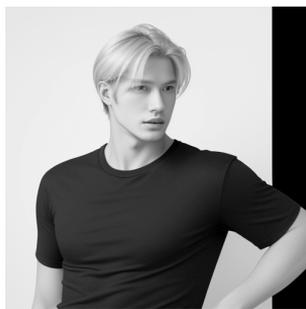


图 3.45 黑衣服的金发男人图像

提示词矩阵的调节选项包含多方面内容,首先是可变部分的位置与权重的调节选项。如果选择将可变部分放在提示词文本的开头,这些词会在生成的图片中更加突出。因为越靠前的提示词在生成过程中权重越高,能够显著影响图片的主要特征。其次是随机种子的应用选项。如果勾选此选项,提示词矩阵中的每张图片都会使用一个不同的随机种子。这保证了每张生成图片的独特性,避免了图片的雷同。第三项是选择提示词选项。用户可以决定是对正面提示词文本框还是负面提示词文本框使用提示词矩阵脚本。这一选项允许用户根据需求调整生成图片的情感基调。第四项是选择分隔符选项。用户可以选择用逗号还是空格来连接可变提示词。这会影响到生成图片时各个提示词之间的关系,从而影响图像的具体表现形式。最后是宫格图边框选项。它决定了生成的宫格对比图中各个图片之间的间隔大小。数值越大,图片之间的间隔越宽,从而影响整体视觉效果。通过这些调节选项,用户可以灵活地调整提示词矩阵的各个属性,以生成符合需求的图像内容。

3.10.2 从文本框或文件载入提示词功能



在 Stable Diffusion 中,用户可以通过特定的脚本从文本框或文件批量导入参数,以生成图片。这些参数不仅描述了图像的内容,还囊括了生成图像所需的各种详细设置。可以通过双线(--)引导每个参数名称,紧跟其后的是相应的数值。例如,用--steps 20 可以设定生成步骤的数量为 20 步;而--restore_faces true 则用于开启面部恢复功能,若未明确设置此参数,默认状态为不开启该功能。对于字符串类型的数值,应当用英文双引号括起来,以确保正确解析。参数之间必须用英文空格分隔,这样脚本才能正确解析每个设置。在这个过程中,每一行的参数组合代表一组独立的设置。

从文本框或文件载入提示词的案例如下。

步骤 1: 输入两组提示词进入文本框中,如图 3.46 所示。第一组参数:--prompt "a young girl, holding a sunflower, green dress" --steps 20。该参数请求以迭代步数为“20”步生成一张描述年轻女孩持有向日葵、穿着绿色连衣裙的图像。第二组参数:--prompt "a boy with a skateboard, red cap, denim jacket" --steps 20。此参数旨在以迭代步数为“20”步生成一张拿着滑板、戴红色帽子、穿牛仔夹克的男孩的图像。

步骤 2: 根据上述输入的参数,Stable Diffusion 将生成两张图像。第一张图像将呈现一个手持向日葵的穿绿色连衣裙的年轻女孩,如图 3.47 所示。第二张图像将呈现一个拿着滑板、戴红色帽子、穿牛仔夹克的男孩,如图 3.48 所示。采用这种方式,可以方便地批量生

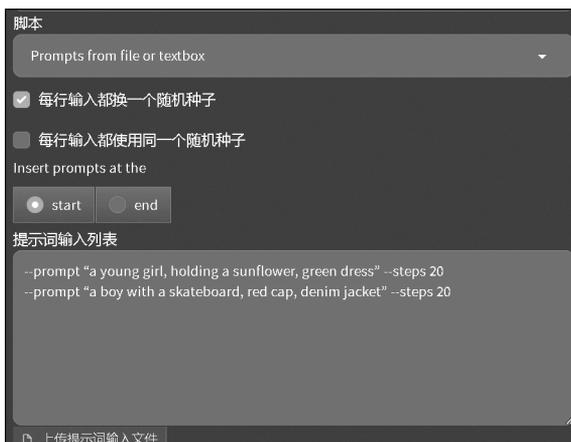


图 3.46 从文本框或文件载入提示词界面

成图像,而不需要手动更改每一张图片像的参数。



图 3.47 手持向日葵的女孩



图 3.48 戴红色帽子的男孩



3.10.3 X/Y/Z 图表功能

X/Y/Z 图表是一种用于展示数据对比的可视化工具,其界面如图 3.49 所示。



图 3.49 X/Y/Z 图表界面

在 X/Y/Z 图表中, X 轴用于展示水平方向上的数据, Y 轴用于展示垂直方向上的数据, 而 Z 轴则通过结合 X 轴和 Y 轴的数据生成按照分组展示的图表。图表会默认显示每个轴的类型和值, 以便用户能够明确地进行数据对比。通过这种方式, 用户可以更直观地理解数据的分布和关系。

在图表生成时, 保持随机种子为“-1”意味着每次生成的图像都会使用不同的随机种子, 以确保图像的多样性。如果指定了一个具体的随机种子, 那么每次生成的图像一致, 则上述的随机性设置将不再生效。用户还可以选择是否勾选“预览子图像”选项。如果不勾选, 则只生成一张包含所有对比的总图。勾选后, 除了生成总图外, 还会单独展示每个用于对比的子图像, 便于详细查看各个子图像的细节。

此外, 宫格图边框参数用于调整宫格图中各个图像间的间隔。数值越大, 图像之间的间隔就越宽。调整此参数, 可以根据需要设置图像之间的间隔, 以获得最佳的视觉效果。通过以上内容可以全面了解 X/Y/Z 图表的各项功能及其设置方法。这不仅提高了图表的可读性和可操作性, 还提供了更灵活的图表生成和展示方式。

X/Y/Z 图表案例如下。

步骤 1: 进入文生图界面下的脚本, 选择 X/Y/Z 图表进行设置。分别设置 X 轴、Y 轴和 Z 轴的参数。其中, X 轴用于测试迭代步数, 分别设置为“20”步、“30”步和“40”步; Y 轴用于测试采样方法, 分别选择“DPM++ SDE Karras”“Euler a”和“DPM++ 2M Karras”采样方法; Z 轴用于测试“Clip skip”。“Clip skip”指语言与图像的对比预训练, 可以概括理解为提示词与图像的关联程度。分别设置数值为“2”“5”和“10”, 如图 3.50 所示。



图 3.50 X/Y/Z 图表设置

步骤 2: 选择 majicMIX realistic 麦橘写实大模型, 并设置提示词内容与生成尺寸。生成参数分析图像, 如图 3.51 所示。

经过分析 X/Y/Z 图表, 能够确定语言与图像的对比预训练 (clip skip) 参数的最优取值以及最有效的采样方法和最合适的迭代步数。通过 X/Y/Z 图表功能, 可以更直观地比较不同数据集的关系和差异。

3.10.4 Ultimate SD upscale 图像放大功能

Ultimate SD upscale 是一个用于图像放大的插件, 安装步骤如下。首先, 打开 WebUI (Web 用户界面) 的扩展选项卡。单击“可用”按钮, 然后单击加载按钮, 以显示所有可用的扩展插件。接着, 在搜索栏中输入 Ultimate, 以找到 Ultimate upscale 插件。单击“安装”按

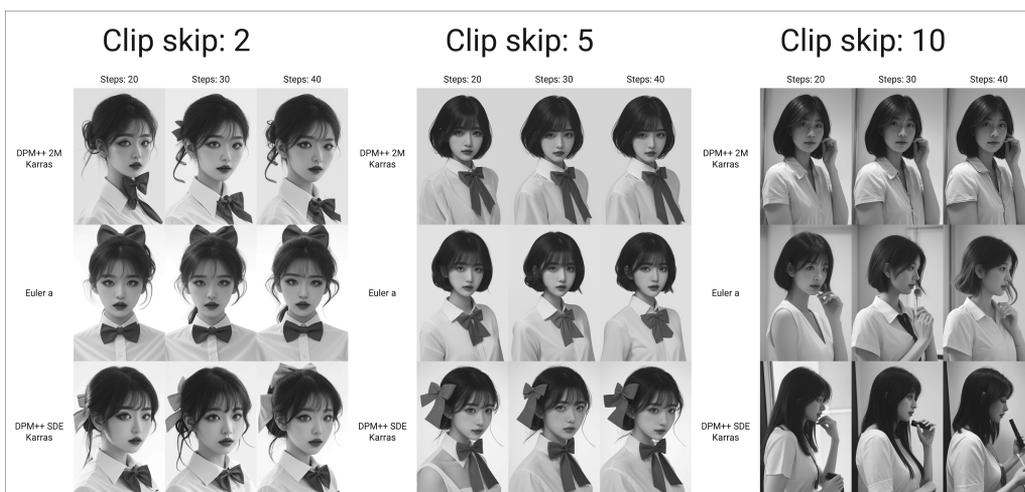


图 3.51 X/Y/Z 图表生成的分析图

钮进行安装。安装完成后,为确保插件正确加载,需重启 WebUI。这样就顺利安装了 Ultimate SD upscale 插件,实现对图像的高质量放大。

在图生图界面的脚本选项中选择 Ultimate SD upscale,可以看到插件的界面,如图 3.52 所示。

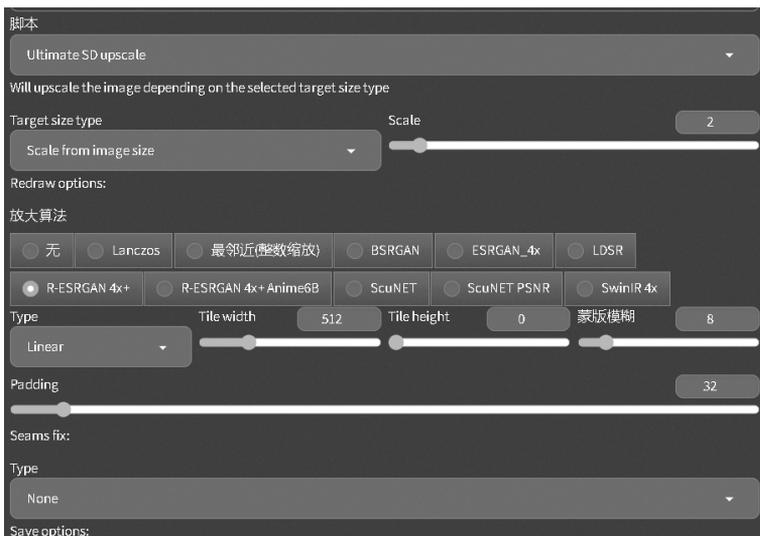


图 3.52 Ultimate SD upscale 界面

Ultimate SD upscale 提供了多种调节选项,以优化图像放大效果。目标尺寸类型参数用于调整输出图像的尺寸,有 3 个选项。首先是“从图到图设置”(from image to image setting),此选项使用宽度和长度的滑块设置,默认最大尺寸为 2048 像素。其次是“自定义尺寸”(custom size),此选项可以设置具体的宽度和高度,最大可设为 8192 像素。最后是“根据图像尺寸缩放”(scale for image size),该选项保持原图像的宽高比,通过缩放系数调整。

“放大算法”选项用于选择不同的放大算法,以优化图像放大效果。用户可以根据具体

需求选择合适的算法,以获得最佳图像质量。

“重绘类型”设置重绘图像的方式,有“线性”“分块(棋盘格)”和“禁用”3个选项。线性模式按顺序处理每一个分块,而分块(棋盘格)模式则按照棋盘格图案处理每一个分块,减少伪影。禁用模式则不进行重绘,可能会在接缝处看到不连贯的效果。

“分块宽度和高度”选项用于设置处理图像的分块大小。分块越大,处理速度越快,伪影越少。填充选项在处理时会考虑相邻分块的像素数量,从而影响接缝处的平滑度。

“蒙版模糊度”用于设置分块重绘时使用的蒙版模糊度,这有助于平滑边缘。接缝修复选项决定是否启用接缝修复功能,以消除可见的网格状伪影。有4个选项可供选择:“带状通道”(band pass)、“只在接缝处处理”、“覆盖周围小区域”、“半分块偏移通道”(half tile offset pass)。在行和列上使用蒙版,覆盖面积更大,效果也更好,但耗时更长。半分块通道加交点通道(half tile offset + intersection pass)对交点进行额外修复。无(None)为禁用接缝修复,为默认选项。

“保存选项”包括“放大”(upscaled)和“接缝修复”(seams fix)。“放大”选项用于保存放大后的图像,默认启用。如果开启“接缝修复”功能,需要选中“接缝修复”选项,这样系统会返回两张图片,一张是未修复的,另一张是修复后的。

在通常情况下,推荐不启用接缝修复,并使用分块模式进行分块。分块大小建议设置为放大图像短边像素的一半左右。通过以上内容,用户可以详细了解 Ultimate SD Upscale 插件的各项调节选项,并根据需要设置,以获得最佳的图像放大效果。

3.11 综合实践：人物场景案例

Stable Diffusion 文生图工作流程案例如下。

步骤 1: 打开 WebUI,加载大模型,如图 3.53 所示。



图 3.53 加载大模型以及 VAE 模型

步骤 2: 输入正向提示词,指导模型生成与提示词相关的图像。提示词应尽可能准确地描述想要生成的图像内容。提示词为“一个女孩,全身,长头发,走在街道上,高领毛衣,商业街,商店,食品,植物,全景,最好的质量”。这段提示词对应的英文提示词为“a girl, whole body, long hair, walk on the street, turtleneck sweater, commercial street, shop, food, plant, panorama, the best quality”。

步骤 3: 设置采样方法为“Euler a”,设置采样迭代步数为“20”步,如图 3.54 所示。

步骤 4: 设置宽高尺寸(画面生成的大小),图像的生成尺寸设置为 512×768 像素。设置生成批次为“1”批,每批数量为“1”张,如图 3.55 所示。

步骤 5: 单击“生成”按钮,生成一张女孩的图像,如图 3.56 所示。

步骤 6: 设置高清修复,选择放大算法为 R-ESRGAN 4x+ 模式,选择重绘幅度数值为