

# STABLE DIFFUSION

## 上手指南

20 分钟简单上手 Stable Diffusion  
另有进阶操作，了解原理，从入门到精通

Sirius

# 说明

本文档主要内容整理自 B 站 UP 主 Nenly 同学（<https://space.bilibili.com/1814756990>）发布的 Stable Diffusion 系统课程视频。部分内容来自互联网以及个人的理解与看法。

本文档为个人的学习笔记，仅供学习参考使用，不作为任何商业用途。如若侵权，请联系删除。

# 目录

<b>一、20分钟简单上手 .....</b>	<b>1</b>
<b>第一步：确定风格.....</b>	<b>1</b>
<b>第二步：描述图片.....</b>	<b>3</b>
<b>第三步：细节调整.....</b>	<b>5</b>
<b>二、提示词 .....</b>	<b>8</b>
<b>内容提示词.....</b>	<b>9</b>
<b>标准提示词.....</b>	<b>11</b>
<b>提示词权重.....</b>	<b>13</b>
<b>反向提示词.....</b>	<b>17</b>
<b>素材网站.....</b>	<b>17</b>
<b>三、其他参数 .....</b>	<b>21</b>
<b>采样步数.....</b>	<b>21</b>
<b>采样器.....</b>	<b>22</b>
<b>分辨率与高清修复.....</b>	<b>24</b>
<b>四、模型选择 .....</b>	<b>26</b>
<b>文件基础.....</b>	<b>26</b>
<b>分类与推荐.....</b>	<b>27</b>
<b>素材网站.....</b>	<b>32</b>
<b>五、图生图 .....</b>	<b>35</b>
<b>基本步骤.....</b>	<b>36</b>

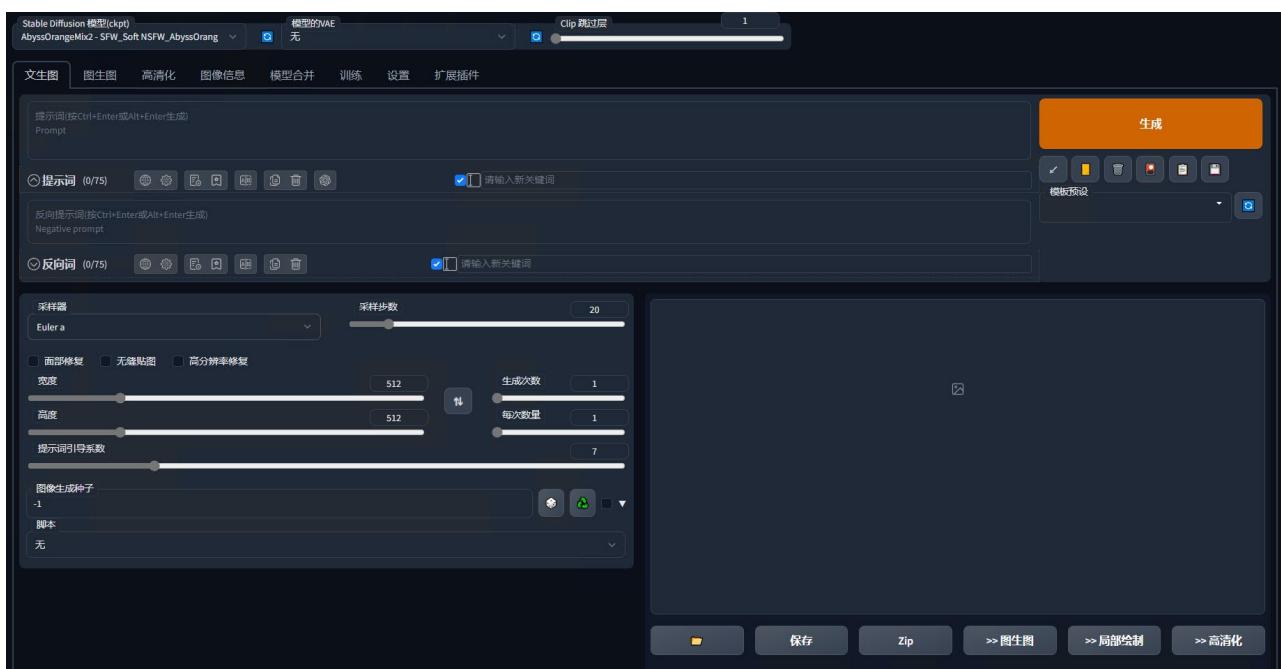
参数解析.....	37
随机种子.....	38
应用场景.....	40
<b>六、超分辨率.....</b>	<b>42</b>
高清修复.....	42
SD Upscale 脚本.....	45
附加功能高清化.....	47
<b>七、局部重绘.....</b>	<b>50</b>
基本操作.....	50
参数解析.....	52
涂鸦绘制与涂鸦蒙版.....	54
上传蒙版.....	56
<b>八、模型进阶.....</b>	<b>60</b>
文本嵌入 <i>Embeddings</i> .....	60
低秩模型 <i>LoRa</i> .....	65
超网络 <i>Hypernet</i> .....	66
<b>九、LoRa .....</b>	<b>68</b>
加载方式.....	68
人物角色类.....	70
画风类.....	70
服饰类.....	72

特定元素类.....	74
<b>十、ControlNet.....</b>	<b>76</b>
安装方式.....	77
基本步骤.....	78
参数解析.....	80
五大模型与应用.....	82
多重应用.....	88
应用实例—二维码生成.....	90
<b>附 1：Stable Diffusion 安装 .....</b>	<b>92</b>
<b>附 2：插件安装 .....</b>	<b>100</b>
<b>附 3：常见问题与解决办法.....</b>	<b>113</b>

## 一、20 分钟简单上手

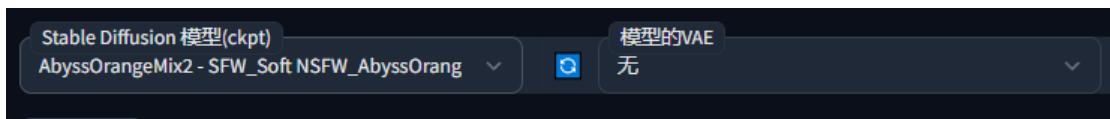
本指南不涉及 Stable Diffusion (SD) 的工作原理。快速上手 SD 不需要任何代码知识，只需要能够描述出自己想要什么样的图片，能看懂简单的英文即可。本章会带领你简单操作一遍，熟悉 AI 画图流程。

首先运行 SD 后会有一个黑色窗口（安装方法见附录 1），这是 SD 的后台。等启动完成后会自动打开浏览器界面，如果没有自动打开，复制黑色窗口中的地址到浏览器即可，之后所有的操作都是在网页中进行。但是黑色窗口一定不要关闭，否则 SD 就会掉线，有时可以在黑色窗口中查看图片渲染进度。



启动完成之后，就会看到这样一个界面。其中我已经安装了一些工具辅助完成绘画，比如汉化插件，用 ChatGPT 辅助描述图片等，后续将会介绍这些工具如何使用，并介绍简单的绘画步骤，如何安装可见附录 2。

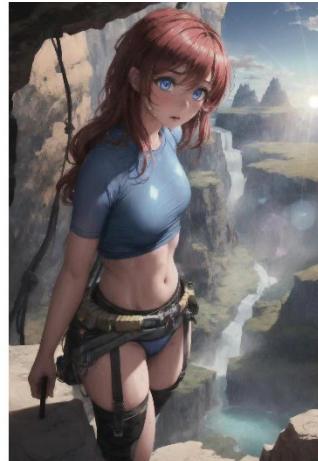
### 第一步：确定风格



图片会有很多种类型，如漫画、写实等，而**模型**就是用来确定风格，选择合适的**模型**来确定图片的**总体绘画风格**。不要看每个模型选项都很长很复杂，其实就是模型的名字而已，这里我们选用 AbyssOrangeMix2（全称：AbyssOrangeMix2 - SFW\_Soft NSFW\_AbyssOrangeMix2\_sfw），是二次元风的模型。

sd-v1-4 是 SD 自带的通用模型，可以拿来练手。模型可以从网站上下载，有各种各样的风格，并且网站上都会带有样片，来帮你了解模型会生成什么类型的图片，之后会介绍资源库网站以及如何下载。

VAE: None



VAE:  
Color101-v1.vae.pt



**VAE 选项**，就是图片的滤镜，和模型一样，也可以从网站上下载喜欢的滤镜，但是与模型不同的是，模型必须要选择一个，而 VAE 可以不选，也可以让系统自动选择合适的。这里我们选无（其实许多模型内置了 VAE）。



当确定好绘图的总体风格时，还会有不同的功能：用文字生成图片，用图片生成图片，增加图片清晰度等等。其中最常用的就是文生图，这里我们也选择文生图，之后的例子也是用文生图举例。

当确定好绘图的总体风格时（之后的例子模型选用 AbyssOrangeMix2, VAE 为无），就可以进行下一步了。

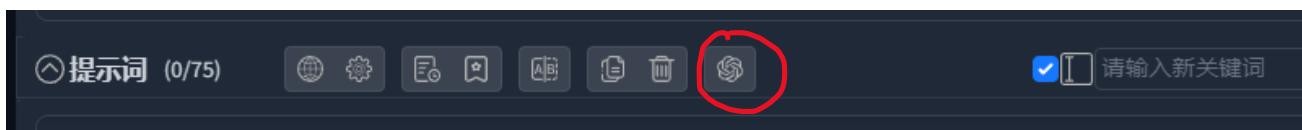
## 第二步：描述图片



上图中上面的框为**提示词 (Prompt)**，对你想要生成的图片内容进行文字描述，不同模型及参数下，同样的提示词生成的图像也会不相同。它支持自然语法、标签语法、emoji 以及颜文字等丰富的语法格式。但是建议使用标签语法，即用词或者短语描述图片，每个词或短语之间用半角逗号隔开。这里我们使用示例：  
Beautiful girl,gothic dress, clear details, Gothic architecture interior, [pink|blond]long hair

下面的框为**反向提示词 (Negative prompt)**，用文字描述你不希望在图像中出现的东西。使用的描述方式与提示词相同。这里我们使用示例：  
(worst quality:2), (low quality:2), (normal quality:2), lowres, normal quality, ((monochrome)), ((grayscale)), skin spots, acnes, skin blemishes, age spot, (outdoor:1.6), manboobs, backlight,(ugly:1.331), (duplicate:1.331), (morbid:1.21), (mutilated:1.21), (tranny:1.331), mutated hands, (poorly drawn hands:1.331), blurry, (bad anatomy:1.21), (bad proportions:1.331), extra limbs, (disfigured:1.331), (more than 2 nipples:1.331), (missing arms:1.331), (extra legs:1.331), (fused fingers:1.61051), (too many fingers:1.61051), (unclear eyes:1.331), bad hands, missing fingers, extra digit, (futa:1.1), bad body, NG\_DeepNegative\_V1\_75T, pubic hair, glans,((nsfw)),(chubby:1.8),(swimsuit:1.5),((Belly)),(exposed skin:1.8)

上述示例直接复制即可，先不要在意符号什么意思，只关注提示词本身都是负面词汇即可，后续会详细介绍。



另外两种提示词输入都只支持英文，此时可以采用翻译软件的方式，但效率较低，所以安装翻译插件 all-in-one 可以稍微提高效率。在右边输入框中输入中文提示词后直接回车就会生成对应的英文提示词。也可以点击红圈中的 ChatGPT 按钮，我已经配置好了 API，用自然语言输入你想生成的内容之后，让 ChatGPT 帮你生成提示词。反向提示词同理（有时可能会不显示输入框，多次刷新网页即可）。

提示词越详细，生成的图片越精细。理论上到此为止就可以生成图片了，可以点击“生成”体验一下。只不过质量较差（还需要第三步进一步调整参数）。在此我用以下两个例子体现提示词对成图的影响。

第一张图片的提示词为 a pretty girl，没有反向提示词。第二张图片的提示词为 absurdres, 1girl, ocean, railing, white dress, sun hat，反向提示词为(worst quality:1.2), (low quality:1.2), (lowres:1.1), (monochrome:1.1), (greyscale), multiple views, comic, sketch, animal ears, pointy ears, blurry, transparent, see-through。

可以看到第二张图细节更多，更清晰，也更有质感。



既然是较为开放的 AI 创作，这时就可以发挥甲方的特点，对图片提出各种

要求。比如第一张图片，我想让她背对我，这时就可以在提示词中加入 back to me，这样就可以得到如下符合要求的图片。



但是两张图的差别却是比较大的，这时就需要使用后面的进阶知识，精准地控制发色、服饰、画面构图，风景等。当然，也有网站可以自动生成提示词，或者将提示词分类，根据需要复制提示词，比如 lexica.art。后续会有更多网站的介绍和使用说明。

### 第三步：细节调整



这一部分就是对 AI 绘画的方法进行规范，我们示例就选用上图参数，直接照抄即可。下面是对每个参数的简单解释，之后会有更为详细的解读。

①高度宽度就是图片的大小。

②生成次数就是生成几次，调整为 3 就相当于你点了三次生成；每次数量就是一次生成几张图片，这两个选项是因为 AI 绘画不可控，为了得到心仪的图片需要多次生成。

③图像生成种子就是一个随机数，-1 就代表随机生成，这样即使其他所有参数全部一模一样，生成的图片也不一样；当为一个正整数时，就指定了随机种子，这样当其他参数都一样时生成图片也是一样的，图片丢失的时候就可以复现图片。

接下来简单介绍剩下的看起来很陌生的参数。

现实中画画可以用铅笔也可以用水彩笔，画出的图画也有很大的差别。**采样器 (Sampling method)** 就是生成符合文本描述的图片的方法，使用不同的采样器就是使用不同的笔绘图。之后会介绍一些常用的采样器，默认可以使用 **DPM++ 2M Karras**，它可以在速度和质量上进行平衡。

AI 绘画就像走路一样，一步一步地走到目的地（最终生成的图像）。**采样步数 (Sampling Steps)** 就是 AI 要走多少步才能生成最终的图像。更多的步数意味着从起点到最终图像的更小、更精确的步骤。增加采样步数直接增加了生成图像所需的时间。但是步数越多不代表生成图片质量越高，取决于采样器，一般建议值在 20 左右，根据不同的采样器，需要进行调整。

**提示词引导系数 (CFG Scale)** 用于控制图片和提示词之前的契合度，这个值越高，生成的图像也越接近你的提示词，但它也在一定程度上降低了图像质量，

限制了 AI 的想象力。**可以用更多的采样步数来抵消负面效果。**过高的 CFG Scale 体现为粗犷的线条和过锐化的图像（就是很抽象）。一般建议在 7-15。

之后我们直接点击生成，稍等片刻之后，如果你生成了和下面一样的图片，就说明你已经成功了（可能会有极其细小的差别，也是成功的）。



**注意，提示词结尾是否有逗号也会影响图片的内容！**

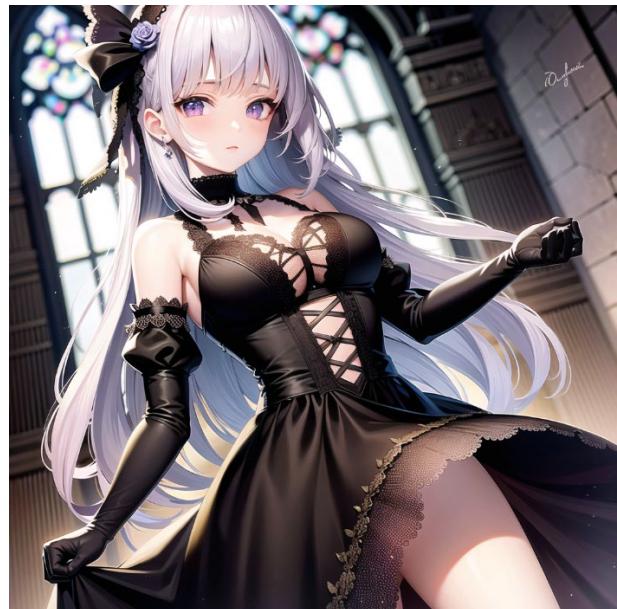
到此为止，基础的参数都已介绍完毕，已经完全可以使用 SD 生成不错的图片了，但是 Stable Diffusion 的能力远不止如此，之后的章节将会详细介绍细节，以生成出色的图片。

## 二、提示词

提示词书写较为自由，但有几点要求：①只支持英文，符号要使用英文的半角符号。②以词组为单位，使用逗号隔开，不需要语法。③可以换行，但要在行末加上逗号。④描述的细节越多，生成的图片质量越好。

```
1girl, walking, forest, path, sun, sunshine,  
shining on body,
```

其实提示词是支持自然语言的，简单描述一下想要的画面，如下左图的提示词就是用自然语言’A beautiful girl wearing a detailed Gothic dress stood inside the Gothic architecture.’生成。



自然语言缺点是生成质量不佳，难以调整细节，所以我们将自然语言描述的句子中的关键词提取出来，用特定的方式进行排列，调整权重，使画面更接近自己想要的效果。如上右图，就是使用提示词：Beautiful girl, gothic dress, clear details, Gothic architecture interior 生成（也修改了其他一些参数配合词组）。

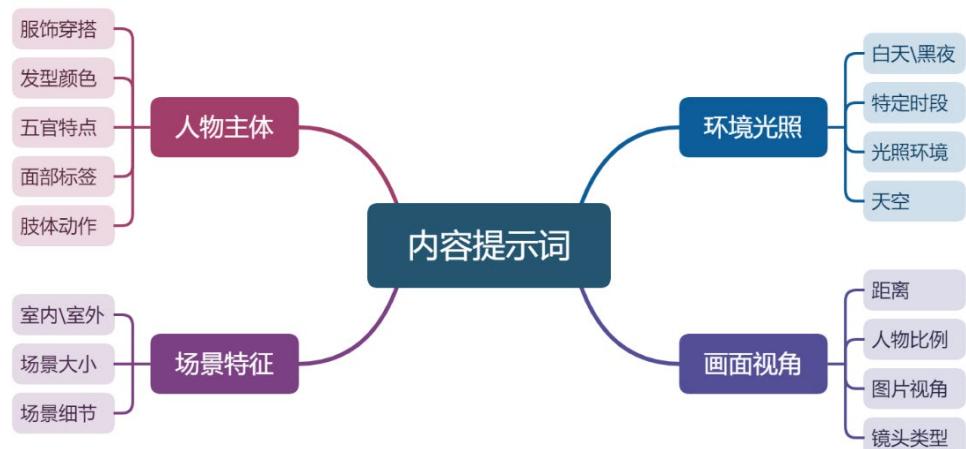
生成的图片可能会符合要求，也可能会奇奇怪怪，因为 AI 绘画就像抽卡一样，具有随机性，想要“抽”到好的图片就需要一定的运气。如何增加“中奖”的概率

呢？这就需要提示词补充更多的细节。

如“一个在丛林漫步的人”是一个非常概括的描述，是男是女？森林里面有什么？时间是几点？天气怎么样？这些细节在提示词里都没有，AI 就只能瞎蒙，“中奖”的概率自然就低了。

## 内容提示词

提示词也不是越多越好，也需要有逻辑：先描述一个雏形，再慢慢细化、补充和微调。这里将具象化的提示词称为**内容提示词**，它描述了图片中具象的部分（人物、场景等）。内容提示词简单的分为四大类，这样在写提示词的时候就可以考虑这四个方面进行补充。



当然内容提示词不止于这四种，这只是帮助你打开绘图的思路，不要限制自己的想象。

### 人物主体

还是采用上述的“一个在丛林漫步的人”的例子。首先是对于人物或者主体的描述，如下面的例子，一个什么样的人？→女孩；女孩穿的是什么衣服？→裙子；

什么样的裙子？→连衣裙；裙子什么颜色的？→白色。甚至可以对衣服的穿着方式进一步细化。越是具体 AI 的思路也越清晰。

1girl, walking, forest, path, sun, sunshine, shining on body,  
white dress, blonde hair, long|

**翻译:** 白色连衣裙，金发，长发，微笑，伸展手臂，举手

可以加入一些形容词，例如 beautiful、happy 等，尽管形容词对 AI 来说比较抽象，但是也可以在一定程度上让画面向想要的方向倾斜。

## 场景特征

对于场景首先就是要加入场景特点，一个丛林肯定是不是空无一物的（当然你也可以让它很空），这时就要加入一些点缀，如白色的花，林间小路，在对这些点缀描述的时候也遵循之前人物主体的规律，只不过只需要简短的描述，不然可能会喧宾夺主。

trees, bush, white flower, path|

**翻译:** 树木，灌木，白色的花，(林间) 小路

当然，还要描述场景的整体位置，如果是室外就加上 outdoor，室内就是加上 indoor，这个会显著影响画面氛围。

## 环境光照、画面视角

关于环境的描述其实也算是场景的一部分，比如画面发生的时间，天空有没有云，光源的位置等等。

day, sunlight, blue sky,

**翻译:** 白天，阳光，蓝天，多云的天空

这部分根据不同的绘图场景会有较大的区别，总之就是要把自己想象成一个摄影师，你要考虑如何构图如何打光才能让画面更好看，之后把你构思的细节全

部告诉 SD。

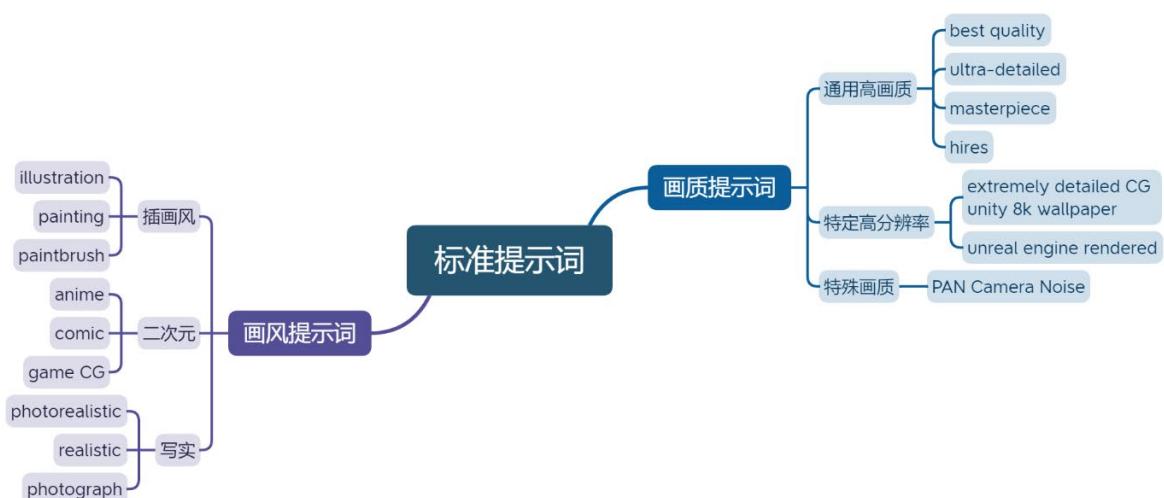
既然是构图，自然就要考虑“拍照”的视角，也就是画面视角。比如你是想要特写（close-up）还是全身照（full body），“拍照”的视角是俯视还是背影，甚至镜头的类型都可以描述出来：广角、sony A7III 等等。

```
1girl, walking, forest, path, sun, sunshine, shining on body,  
white dress, blonde hair, long hair, smiling, stretching arms, hands up, beautiful, happy,  
trees, bush, white flower, path, outdoor,  
day, sunlight, blue sky, cloudy sky,|
```

最终根据书写提示词的思路，就可以写下较为完备的提示词。但是只有内容提示词还是不够的，有时我们的画面会很模糊，细节不清晰，这时就需要另外一种提示词——标准提示词。

## 标准提示词

有时我们对于图片的画质、画风等会有特别的要求，此时就需要标准提示词来为画面的风格设定一个目标标准。标准提示词其实是相对固定的，比如高清的图片就是 8k，想要虚拟仿真就是 unreal engine 等等，因此可以总结归纳以便以后使用，以下是对标准提示词的划分以及部分总结。



当想要改变照片的画质时，就加入改变画质的词语，上述只是列举了一小部分提示词，还有一些常见的 HDR、HD、UHD 等等，生成图片时可以都加上“best quality, ultra-detailed, masterpiece, finely detail, highres, 8k wallpaper”；当想改变照片艺术风格，也就是画风时，也是改变对应的标准提示词即可，无需改动内容提示词。但是有一点，图片的画风非常依赖模型，特别是真实系的图片更依赖于基于真实照片训练出的模型，让不同的模型画擅长的画风才是正确的选择。



一个 AI 能读懂的好提示词，应该是内容充实丰富，且画面有清晰标准的。将提示词模块化，当我们书写提示词时，就可以基于下面的模板对号入座地填写每一段内容，如果想对图片部分内容进行微调时，只需要修改对应的部分即可，这样使得提示词十分清晰有条理。

通用模版	描述人物	→ (1girl:2.0), solo, nilou \{genshin impact\}, solo, long hair, jewelry, blue gemstone, earrings, horns, crown, cyan satin strapless dress, white veil, neck ring, red hair, {green eyes},
	描述场景	→ indoor, room, house, sofa, wooden floor, plant, flowers, trees, windows,
	描述环境(时间、光照)	→ day, morning, sunlight, dappled sunlight, backlight, light rays, cloudy sky
	描述画幅视角	→ full body, wide angle shot, depth of field
	其他画面要素	→ light particles, fantasy, wind blow, maple leaf, dusty, ... (其他往后增加)
高品质标准化	高品質标准化	→ {{masterpiece}}, {best quality},{highres}, original, reflection, unreal engine, body shadow, artstationextremely detailed CG unity 8k wallpaper
画风标准化	画风标准化	→ (illustration), (painting), (sketch), anime coloring, fantasy,
其他特殊要求	其他特殊要求	→ exaggerated body proportions, greasy skin, realistic and delicate facial features, SFW,

## 提示词权重

```
SFW, 1girl, walking, forest, path, sun, sunshine, shining on body,  
yellow skirt and white t-shirt, blonde hair, long hair, smiling, stretching arms, hands up, beautiful, happy,  
trees, bush, white flower, path, outdoor,  
(masterpiece:1,2), best quality, masterpiece, highres, original, extremely detailed wallpaper, perfect lighting,  
(extremely detailed CG:1,2), drawing, paintbrush,
```

我们可以看到，有些提示词后面会有括号和数字，这些内容是用来增强或者减弱某些提示词的优先级和权重。

如上图提示词中，虽然有 white flower，但是实际生成的图片（下图）中并没有太多明显的白花。原因就在于，如果输入了很多不同的元素，AI 在处理的时候不知道你最侧重什么，然后就优先画了比较靠前的内容提示词。



这时如果非常想要白花，就增加白花的权重和优先级。提示词默认权重为 1，有两种基础的调整权重分配的方式，分别都可以增加或者减弱权重：

<b>括号加数字</b>	<b>例:</b> <i>(white flower:1.5)</i> , 含义: 调节白花 (White flower) 的权重为原来的1.5倍 (增强)	<b>例:</b> <i>(white flower:0.8)</i> , 含义: 调节白花 (White flower) 的权重为原来的0.8倍 (减弱)	
<b>套括号</b>	<b>圆括号:</b> <i>((white flower))</i> , 每套一层, 额外 $\times 1.1$ 倍。 此处: 调节白花 (White flower) 的权重为原来的约 $1.1 \times 1.1 = 1.331$ 倍 (增强)	<b>大括号:</b> <i>{{{white flower}}}}</i> , 每套一层, 额外 $\times 1.05$ 倍。 此处: 调节白花 (White flower) 的权重为原来的约1.15倍 (增强)	<b>方括号:</b> <i>[[white flower]]</i> , 每套一层, 额外 $\times 0.9$ 倍。 此处: 调节白花 (White flower) 的权重为原来的约0.729倍 (减弱)

**不要调节个别提示词的权重太高, 建议范围  $1\pm0.5$ 。**当提示词权重大于 2 时就会扭曲图片效果, 如果权重很高还是达不到效果, 就需要更多同类型的词条协同增强效果, 当然也有其他办法调节。



调节权重还有快捷方式, 选中想要调节权重的词组, 按住 ctrl+上下键, 就可以快速地给选中词组加括号和权重。

## 进阶权重分配

这部分不太用的到, 对画面掌控更加地精细, 如果感觉其他办法调节的力度不是很合适时可以使用这部分的内容。

之前的章节说过, AI 绘画就像走路一样一步一步得到成片, 这样我们控制某个词组从第几步开始画, 以此来弱化这个词组的效果。如 [flower:5], 就是从第 5 步才开始画花, 这样就使得花只占画面的一小部分; [flower::10] 代表从开始就画,

但是画到第 10 步就不画了; [[flower::30]:5] 代表从第 5 步开始画, 到 30 步结束。



由于是通过控制步数来控制权重, 步数本来就不高的情况下很容易画不出来,  
比如总步数只有 10 步, 但是你要从第 8 步开始画花, 就很可能画不出来。

除了一个元素的精细控制以外, 还可以通过步数对画面比重进行控制, 比如 [girl:flower:50], 前面 50 步用来画人, 之后的步数用来画花。但是这种写法不太好用, 一般用百分百来控制占比, [girl:flower:0.5], 这样就表示前面的 50% 步数是画人的, 后面的 50% 用来画花。

除了通过步骤调整权重, 还有元素融合的方法, 可以定义一个混合特征的词组。比如金色头发过于单调, 想要粉金色的效果, 就通过[pink|golden]long hair 实现, 用中括号包含所有颜色, 颜色用 | 隔开, 渲染的时候, 是一步粉红一步金色, 最后出来的是调节过后的粉金色。中括号起到了混合的作用。使用 **AND** 连接也起到同样的效果 (**AND** 一定要大写): pink long hair AND golden long hair。如下左图。



如果想调节混合的比重，就可以使用上述的调整步数的方式：[pink long hair: golden long hair: 0.7]，这样头发会偏粉色一些。如上右图。

元素也可以随机选择，如{Crown|Corolla|Hairpin|Bowknot}就是随机选择一个发卡类型，前面介绍到“每次生成数量”这个参数，如果每次只生成一张图片就只会选择其中的一个，因此要生成多张图片才能看出差异。

### 提示词位置/数量的影响

由于 SD 代码上的特性，导致开头与结尾的词往往作用性更强，提示词数量越多，单个提示词的作用性越低。

关于提示词数量，你可能已经注意到了，会有数量限制。但是在 我们使用的 SD 中，是可以写 75 个词汇以上的提示的（虽然标着正向提示词最多 75）。SD 会自动通过对提示词进行分组，当提示词超过 75 个，会以 75 个为一组，然后将所有组进行合并，然后生成。分组会对成图造成一定影响，为了避免将短语或者同类的词组分成俩组，SD 会自动检查，也可以人为控制分组，通过输入 BREAK 来手动分组，BREAK 必须为大写。

## 反向提示词

如果不希望画面中出现什么，就把它往反向提示词中放，反向提示词可以为空，但是大多数情况我们都会放一些通用的项目，比如拒绝低质量、手指不要畸形、不要单色调等等。

110/150

NSFW, (worst quality:2), (low quality:2), (normal quality:2), lowres, normal quality, ((monochrome)), ((grayscale)), skin spots, acnes, skin blemishes, age spot, (ugly:1.331), (duplicate:1.331), (morbid:1.21), (mutilated:1.21), (tranny:1.331), mutated hands, (poorly drawn hands:1.5), blurry, (bad anatomy:1.21), (bad proportions:1.331), extra limbs, (disfigured:1.331), (missing arms:1.331), (extra legs:1.331), (fused fingers:1.61051), (too many fingers:1.61051), (unclear eyes:1.331), lowers, bad hands, missing fingers, extra digit, bad hands, missing fingers, (((extra arms and legs))),

上图反向提示词较为通用，文本如下，可直接复制使用：

NSFW,(worst quality:2),(low quality:2),(normal quality:2),lowres,normal quality,((monochrome)),((grayscale)),skin spots,acnes,skin blemishes,age spot,(ugly:1.331),(duplicate:1.331),(morbid:1.21),(mutilated:1.21),(tranny:1.331),mutated hands.(poorly drawn hands:1.5),blurry,(bad anatomy:1.21),(bad proportions:1.331),extra limbs,(disfigured:1.331),(missing arms:1.331),(extra legs:1.331),(fused fingers:1.61051),(too many fingers:1.61051),(unclear eyes:1.331),lowers,bad hands,missing fingers,extra digit,bad hands,missing fingers,(((extra arms and legs)))

另外第一章中给出的反向提示词也是较为通用的。下一节会介绍提示词素材网站，可以更加便利地获取适用于不同模型的、最新的词汇。

## 素材网站

除了借助翻译插件和 ChatGPT 辅助书写提示词之外，还可以从提示词库中挑选提示词，由于是网站，提示词素材也再不断更新，而且不受自身想法的限制。以下介绍几个易上手的网站，每个网站之前相辅相成，基本上可以覆盖所有需要的提示词。

<https://tags.novelai.dev/>

Danbooru 标签超市，打开网站第一页即是中文使用教程，像逛超市一样，网站将词汇进行分类存放，便于查找同一类型的词汇，同时也可直接中文搜索。唯一需要注意的是里面的词汇并不是适用于所有模型，特定于某些模型的词汇会特别标出。

The screenshot shows the Danbooru Tag Supermarket interface. On the left is a sidebar with a tree view of categories: 人文景观 (Human Landscapes), 位置 (Location), 建筑物 (Buildings), 城市 (City), 食物 (Food), 人物 (Person), 作品角色 (Work Roles), 构图 (Composition), 物品 (Items), 自然景观 (Natural Landscapes), 艺术风格 (Art Style), 预设 (Presets), and 嵌入模型 (Embed Models). The '室内' (Indoor) category under '位置' is selected, showing 35 items. The main area displays a grid of 12 tags, each with a name, count, and various icons:

- classroom <11k>
- bedroom <8150>
- bathroom <7501>
- bathtub <6294>
- kitchen <4098>
- library <3932>
- stage <3110>
- locker room <2479>
- office <1521>
- gym storeroom <1418>
- messy room <748>
- gym <1142>
- 健身房 <611>
- fitting room <580>
- 教室
- 卧室
- 浴室场景
- 浴缸
- 厨房
- 图书馆
- 舞台
- 更衣室
- 办公室
- 体育器材库房
- 杂乱的房间
- 健身房
- 医室

On the right, there are sections for '购物车' (Cart) with '我想买' (Want to buy) and '我不想要' (Don't want), both showing 'No Data'. At the bottom are buttons for '导入标签' (Import Tags), '清空购物车' (Empty Cart), and '结算' (Check Out).

[t.csdn.cn/lAv76](https://t.csdn.cn/lAv76)

其中有一些 AI 绘图实战项目，可以用来练手，最主要的是其后有中英文对照表，在 SD 的提示词中有许多英文缩写的使用，如 NSFW 就 not safe for work，还有一些专业词汇如 panoramic view 就是全景，便于查找。

[www.atoolbox.net/Tool.php?id=1101](http://www.atoolbox.net/Tool.php?id=1101)

一个工具箱——AI 绘画提示词生成器。通过打标签的形式，选择自己想要的标签，会自动地按照之前提到的语法规则整理在一起，直接复制即可。同时也支

持生成负面标签。

## AI绘画提示词生成器

The screenshot shows a web-based tool for generating AI drawing prompts. At the top, there's a grid of various negative labels such as 'lowres', 'bad anatomy', 'missing fingers', etc. Below this is a search bar with placeholder text: "+增加权重; “-”减少权重; ‘x’删除标签; 拖动标签改变顺序;". To the right of the search bar are buttons for '加权符号' (Weight Symbol) and '重置' (Reset). A '自定义标签' (Custom Label) button is also present. The main area is divided into two sections: '正面' (Positive) and '负面' (Negative). Under '正面', the text 'best quality, masterpiece, highly detailed' is listed. Under '负面', a long list of negative labels is shown, including 'lowres', 'bad anatomy', 'bad hands', etc. Below these sections is a large grid of clothing items with checkboxes, under the heading '衣服' (Clothing). The grid includes items like 'evening dress - 晚礼服', 'skirt - 短裙', 'long skirt - 长裙', etc.

[ai.dawnmark.cn/](http://ai.dawnmark.cn/)

可以当做翻译器使用，输入中文生成提示词，也可以直接选择希望的图像类型，每种类型都是效果图帮助选择。

This screenshot shows a user interface for generating images. At the top, there are several AI models listed: Novel AI, Nijijourney, Midjourney, Dreamstudio, and Stable Diffusion. Below them is a search bar with placeholder text '+这里可以输入更多关键词,关于生成图片的'. There are also fields for '加入参考的图片URL链接' (Add reference image URL) and '复制插画词' (Copy illustration word). The main area has a search bar with '美丽对称的面容' and '美丽的风景' selected. A checkbox '选择更多图像细节' (Select more image details) is checked. Below this is a 'Art Medium' dropdown set to 'Drawing', with other options like 'Pencil', 'Ink', 'Pastels and Chalk', etc. A 'Search' button is also present. A grid of image thumbnails shows various styles of drawings, including 'Illustration', 'Sketch', 'Drawing', 'Doodle', 'Children's Drawing', and 'Dot Art'.

*openart.ai/*

*arthub.ai/*

这两个是模型和绘画分享网站，有些创作者会主动分享作图用的提示词和模型，前者偏欧美风，后者偏亚洲风和二次元。这时就可以挑选喜欢的图片，利用之前学习的提示词分类知识，对提示词进行筛选，如果喜欢图片的画风和质感，就抄其中的标准提示词。

### 三、其他参数



在第一章中已经简单地介绍了这一部分每个参数的意思，如果说提示词是咒语，这一部分就是魔杖，控制了咒语的具体施放形式。[以下会详细介绍每个参数](#)，[首先声明的是在讲解一个参数的时候，默认其他参数不变，这样就可以对比调整一个参数对图片的具体影响。](#)

#### 采样步数



[AI 生成图像就是一个加噪再去噪的过程。加噪就是将图片扩散，图片信息也就被分散，从而让 AI 可以进行重新组织。去噪就是用像素一点点模拟最终生成](#)

的图像，每模拟一次图画就会更清晰。

采样步数就是模拟的次数，理论上步数越多最终效果越清晰，一般默认采用 20 左右，实际步数大于 20 步之后效果提升并不大且会增加很多运算时间。以下依次为 10 步（条纹和白点即缺陷），20 步和 30 步。



如果算力充足且想要更高的精细度，就把步数设置到 30-40。需要注意的是，步数不应少于 10，否则会很诡异。

### 采样器

采样器就是 AI 进行图像生成的时候用的某种特定的方法，我们使用的 SD 中附带了十几种采样器，其实常用的也就 4-5 种，这里介绍一些常用的采样器。需要注意的是，下面给出的都是笼统的评价，根据模型和提示词的不同每个采样器都会有差异，实际画图推荐带有加号的，因为是复合调整过的更为稳定。可以多试几种，每种多试几次，AI 绘图就像抽奖一样，中奖概率大并不代表一定能中奖。

**Euler a 和 Euler**：富有创造力，适合二次元图像、插画、小场景，但是出图比较朴素。左图为 Euler a，右图为 Euler。可以看到人物的头发不是很自然，并且有些重复，衣服的褶皱也不是很细腻。



DPM2 和 DPM2 a 出图速度很快，DPM++ 2S a 细节会比较丰富。另外，某些模型也会推荐采样器，这都是作者亲自测试过的。比如 AbyssOrangeMix 就推荐 DPM++ SDE Karras。在下载模型的时候可以看到相关信息，如下图就是 TMND-Mix 模型的作者推荐参数。

模型简介：

### Suggestion/建议：

Step 1/第1步: txt2img/文生图:

Hires. fix/高分辨率修复: On/开

Upscaler/放大算法: R-ESRGAN 4x+

Hires steps/高分迭代步数: 20-30

Denoising strength/重绘幅度: 0.25-0.35

Step 2/第2步: Send to img2img/>> 图生图:

Resize mode/缩放模式: Just resize/仅调整大小

Denoising strength/重绘幅度: 0.1-0.2

Script/脚本: SD upscale/使用SD放大(Upscaler/放大算法:R-ESRGAN 4x+)

此外再介绍一些采样器笼统的特点，方便选择：

**DDIM:** 可以快速生成高质量图像，适合写实人像、复杂场景。

**DPM++2S a Karras:** 画像细节会逐步增加与完善，相比其它采样器，可以在小分辨率画像中，容纳更多的细节。适合写实人像、复杂场景。

**DPM++ 2M Karras**: 在速度与质量之间进行平衡，并生成更精确的图像及其细节。适合二次元图像、三次元图像。

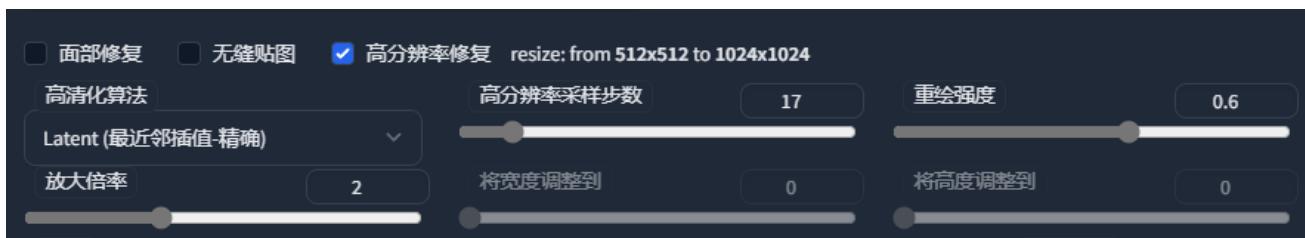
**DPM++ KDE Karras**: 整体图片细节会不断变化。适合写实人像、复杂场景。

当然不同的采样器也不是绝对地只适合特定的场景，需要多次尝试选择心仪的采样器。

## 分辨率与高清修复

宽度和高度就是最终出图的分辨率，默认是 512\*512，但是默认分辨率下哪怕细节再多也是很模糊的，设备允许的情况下会把像素提升到 1k 左右。但是分辨率不是越高越好，一方面是对设备的显存要求很高；另一方面是会出现多人、多手的情况，原因是训练 AI 的图片分辨率可能比较小，生成的图片就是多张图片拼接而成的。

当想生成更高分辨率的图片时，会先用低分辨率绘制，之后靠高清修复（Hires Fix）来放大。

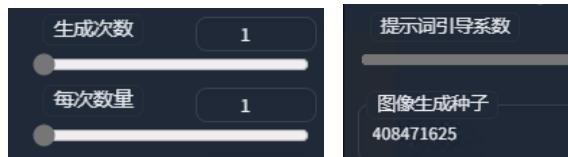


本质是进行了一次额外的图生图，放大倍率如果是 2, 512\*512 的分辨率出图就是 1024\*1024，宽和高都放大一倍。高清修复的参数调整和设备有很大关系，要根据自己的设备调整合适的参数。

SD 还有专门的页面用于高清修复，之后会系统讲解三种高清化方式。

高清修复旁边还有面部修复，一般都会勾选上，和日常使用的智能美颜差不

多（但对显存要求极高！）。无缝贴图（Tilling）用来生成可以无缝贴满整个屏幕纹理性图片的，没有需求不要勾选，会使得图片很奇怪。



生成次数就相当于点几次生成按钮，每次数量就是一次生成几张，CFG 提示词引导系数就是图片与提示词的关联程度，这些参数都很简单，第一章已经介绍过，就不再详细展开。但有一点需要注意，每次生成多张图片其实是把多张图片看成一张很大的图片进行生成的，会对显存有较高要求。

如果想让 AI 按同一组提示词和参数出图，就把生成次数提高，每次数量适当提高，生成上百张图，总有符合条件的最佳选择，然后结合后面将讲到的局部重绘功能，就可以生成百分百满意的作品。

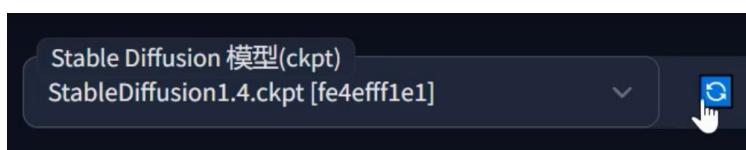
## 四、模型选择

AI之所以能生成图片，是因为它对很多其他画作进行了深度学习，我们一般把这个拿图让AI学习的过程称为喂图。AI学习的内容不只包括如何描绘具体的事物，还包括事物的呈现形式，也就是画风。如果给AI喂的都是二次元风格的，那么这个AI所画出的人、风景图片都像是二次元插画。就像一名学生学的是素描，画什么物体都是满满的线条。

我们喂给AI的图片，以及AI针对这些图片学习的过程，最后会被打包整合到一个文件中，这个文件就是模型了。每个模型就像掌握不同技术的画师，使用不同的模型就可以画出不同画风的画作。

### 文件基础

模型一般的存储位置为：“下载位置/stable-diffusion-webui/models/Stable-diffusion”，如果下载了新的模型文件，也是需要放在这个位置。其中一类模型称为检查点模型（Checkpoint），类似于打游戏一样，每玩一段时间就记录一个检查点，AI对图片的学习也是每运算到某个关键位置的时候就记录下来，方便下次回滚继续运算。这一个个检查点就构成了模型文件，足以支持AI出图。



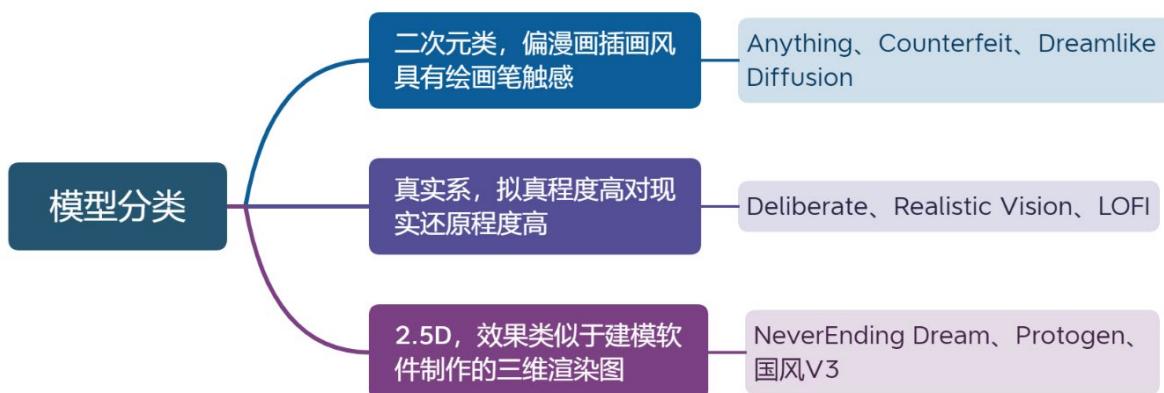
当把模型放到指定位置之后，如果已经打开了WebUI，则需要点一下蓝色的刷新按钮才能看到新加入的模型。由于之后在资源网站上下载会有其他类型的模型，为了区分，以.ckpt 和.safetensors 结尾的文件都放在上述指定文件夹内（也不是绝对的，但目前可以这样理解）。如果安装模型遇到了问题，可以查看附录3。

在模型旁边的 VAE 选项，可以简单理解为 AI 绘图的调色滤镜，最直观地就是影响画面的色彩质感。目前比较新的模型里，已经把 VAE 整合进模型文件了。少数没有的模型中，在下载的界面作者会推荐适合模型的 VAE，也有一些普适的 VAE，比如 kf-f8-anime。需要注意的是，VAE 的后缀一般为 .pt，存放的位置是在“[下载位置/stable-diffusion-webui/models/VAE](#)”中，与模型的位置并不相同。

其实模型的类型并不只有检查点这一种，还有三种具有其他作用的模型。用于画面微调的**超网络 (hypernetwork)**，用于优化画风的**嵌入式向量 (embeddings)**，和固定特定人物角色特征的**低秩适应 (LoRa)**，在之后的模型进阶中会讲到。

## 分类与推荐

市面上的所有模型，基于出图特点和风格倾向，大致都可以划分为以下三类：



其中每一类都推荐了三种模型，作为新手而言，有这 3\*3 个加上 SD 自带的一个总共 10 个模型，足够熟练运用 SD 了。在介绍如何查找和下载这些模型之前，先来了解这 9 个模型各自的特点，以便快速生成满意的作品。

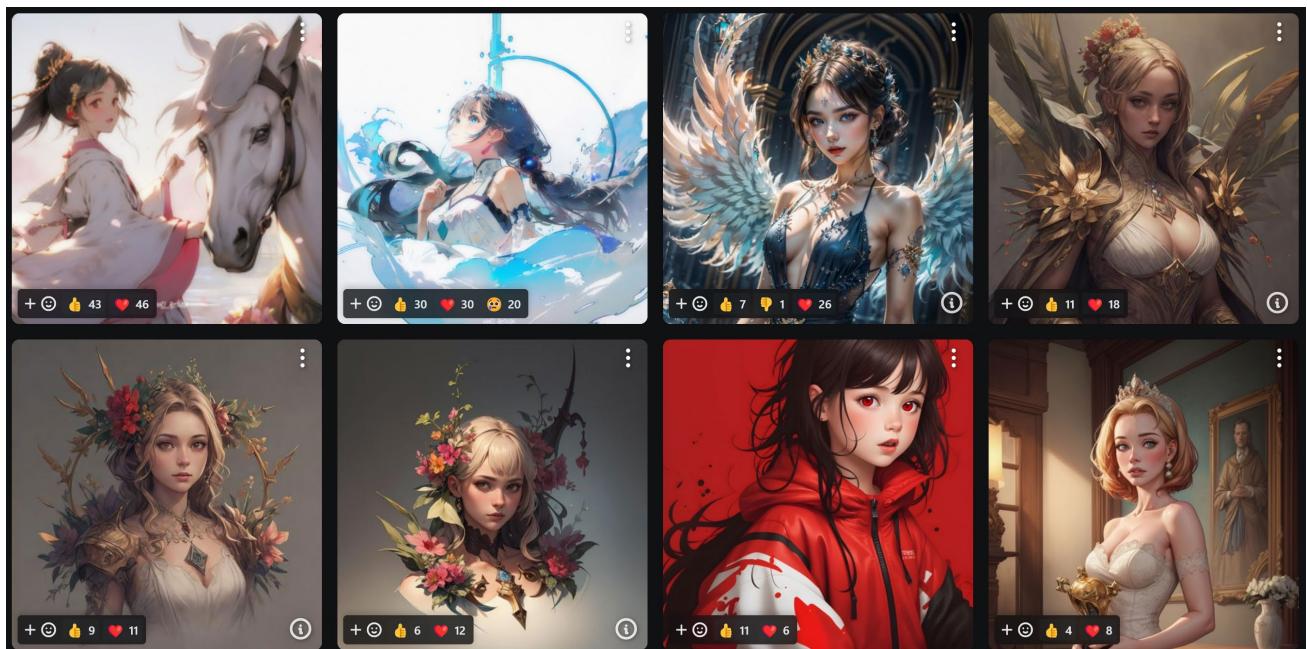
需要注意的是，每个模型都会不定期的更新，所以在查找的时候模型名字中的版本号可能会不一样，但都是同一个模型。

## 二次元类

万象熔炉 Anything 是比较经典的二次元模型，他对于提示词的要求不高，用简单的词语也可以创作不错的效果。可以用它生成接近于动漫、插画、角色立绘等的画面风格。



Counterfeit 是泛用性很广的插画风模型，可以生成出包含复杂室内外场景的绘画作品。场景的精细度、细节还原度高，看起来很精致。



Dreamlike Diffusion 是一个比较魔幻的插画风模型，画风十分独特。在合适

的提示词和参数下，会生成具有幻想色彩的作品，还可以用来生成食物、动物等画面。

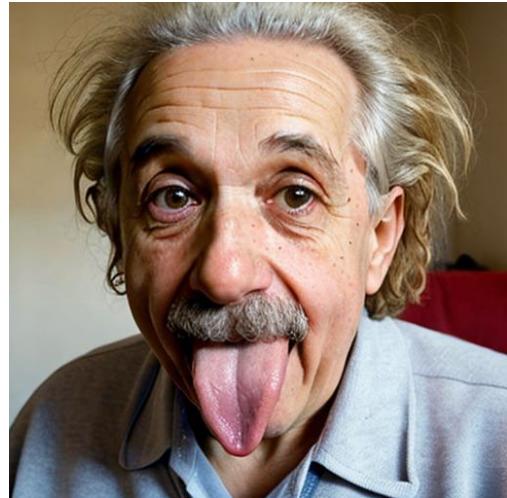


## 真实系

Deliberate 是目前最好的真实系模型之一，更像是超级升级版的 SD 官方模型。像下面这张图，无论是背景的植物，还是人物的头发和肌肉纹理，都清晰可见。模型同样自由度很高，可以创作人物、动物、风景等不同的图片。



如果想让写实的质感更为朴素一点，就可以使用 realistic vision，可以用来生成食物、动物等画面。有些创作者甚至会拿来生成具有真实感的假新闻照片。



如果想生成真实系的人像，还有一个好用的融合模型 LOFI，可以用来生成精致的照片级专精模型。这个模型对人物面部的处理比前面两个更为精致。



## 2.5D

NeverEnding Dream 在画人方面有独特的审美，可以和 LoRa 类模型结合，进行动漫游戏角色的二次创作。第二个 protogen 其实可以算是一个真实系的模型，但在表现效果上有较大的自由度，可以很真实，也可以生成超现实的画面。



第三个国风 3 整体风格适合生成古风的人物服饰、场景，产生的人物形象也符合国人的审美，整体的质感偏三维渲染。结合其他的 LoRa 模型，还能创作水墨风等特色的作品。



除了介绍的这些，还有一些比较小的模型，可以实现特定类型的图像绘制，例如现代感的建筑、平面设计等等。可以结合不同的模型生成自己所需的部分，再通过 PS 等软件将多张图片结合，甚至自己绘画补充剩余部分，都可以得到满意的作品。

## 素材网站

首先再介绍网站之前，先了解如何搜索模型。我们如果有明确目标的话，直接搜索模型名字就可以，但如果想要找一种类型的模型，根据我们之前对模型的分类，通过搜索以下关键词来搜索对应类型的模型。



可以使用这些词作为关键词搜索对应风格的模型，也可以用他们作为提示词来激活模型的风格特性。

目前 AI 绘画里主流的下载网站有两个，一个是 HugginFace，它用于用户分享 AI 学习模型和数据集，第二个就是 Civitai。

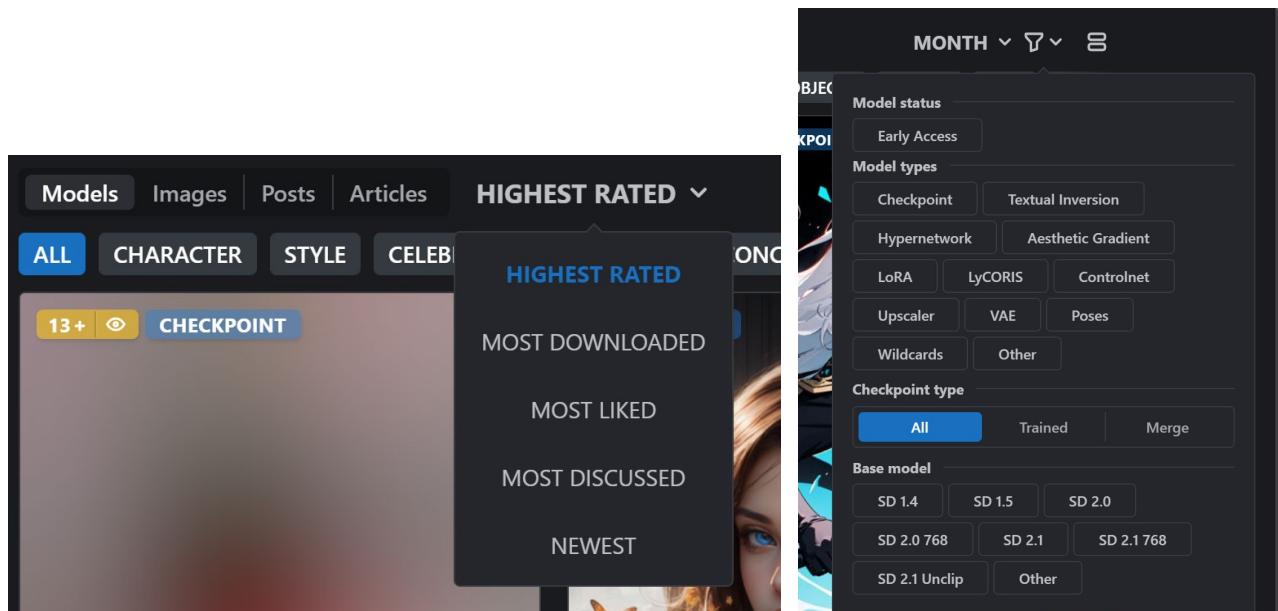
### *HugginFace*

由于它包含的内容比较广，不仅包括 AI 绘画还包含其他 AI 领域的东西，所以具有一定的专业性，第一次使用略有难度。在这里就不过多介绍了，如果需要下载官方的模型，或者寻找一些少见的模型可以来这里搜一下。

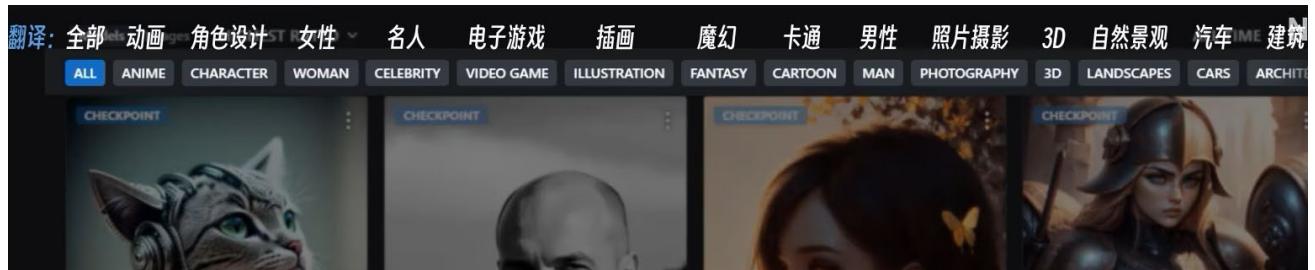
## Civitai

Civitai 则是一个完全的 AI 绘画模型分享平台。每个模型都有展示的图片，在模型详情页的下面还有其他用户生成的图片供参考，每张图片都有生成的信息，可以直接复制喜欢的图片的关键词。

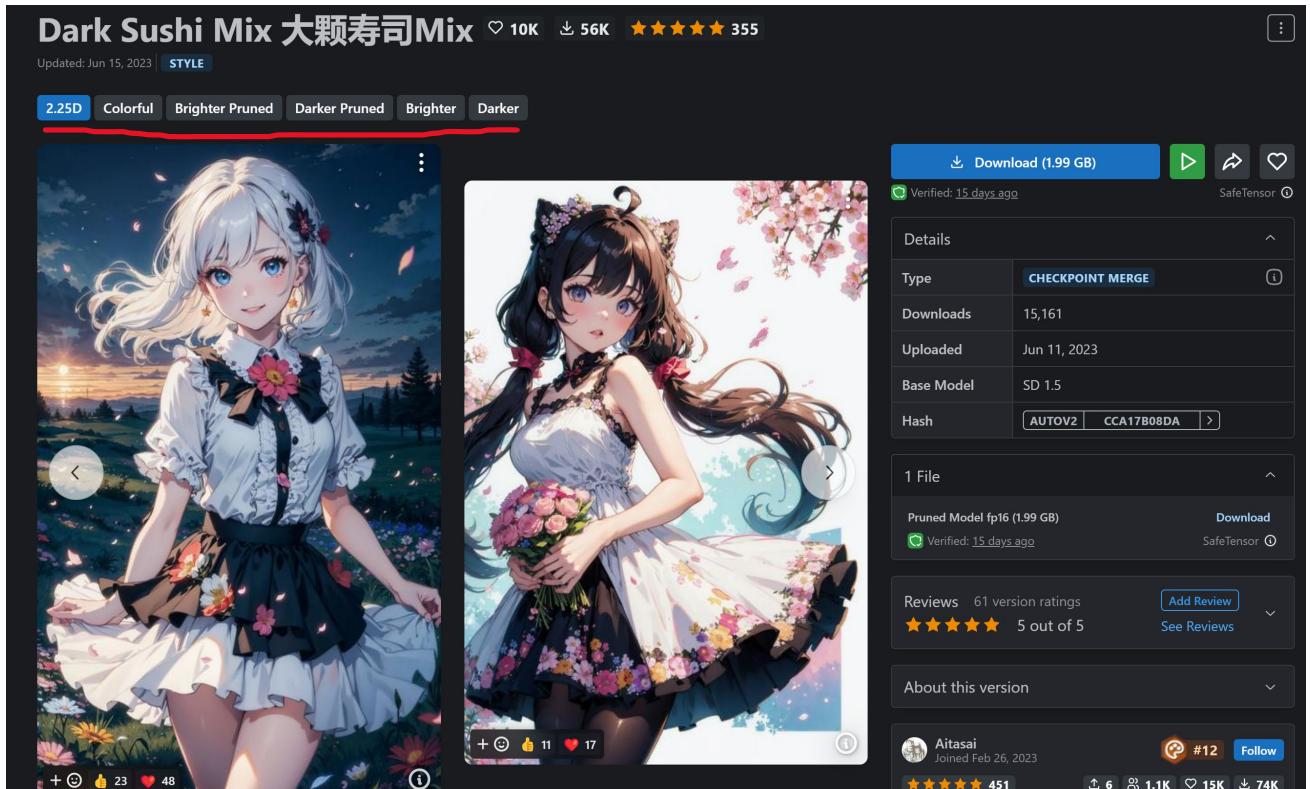
在 C 站上访问和和下载均不需要注册，使用谷歌账户登录之后可以看到一些被限制的内容。在左上角进行筛选，可以选择当下最火的模型，如左下图。



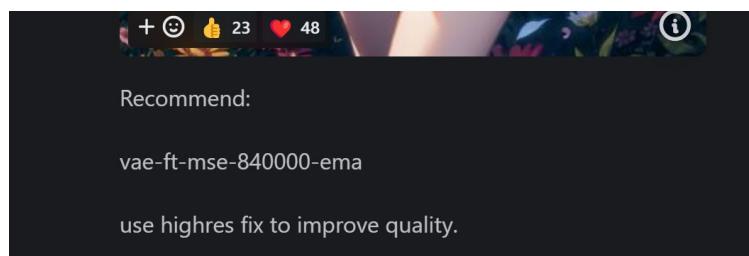
点击右上角的小漏斗，可以基于之前介绍的模型进行分类，如果要找大模型，就选择 checkpoint，也就是我们之前介绍的检查点模型。第二个检查点类型不用太在意，默认选择 all，trained 就是作者基于图像一点点创作的模型；有时为了获得更好的效果会把多个模型融合到一起就是 marge，这类模型名字一般都有 mix。第三个底层模型是训练模型需要考虑的东西，默认什么都不选即可。



我们还可以基于模型的风格进行筛选，想要真实就选 photography，都是比较灵活的推荐。



进入到模型的详情页，一个模型可能有好几种版本，可以选择不同的版本下载，如红线所示。但是一般情况下都建议下载最新的，直接点击蓝色的下载按钮即可。



在模型的下面一般都会有使用介绍，作者会把推荐的参数放在这里，如适合的风格、关键词、分辨率以及 VAE 等等，可以快速地掌握使用这个模型。

而在模型使用介绍的下面，就是其他用户使用模型生成的图片了，每张图片都有参数，可以直接复制。

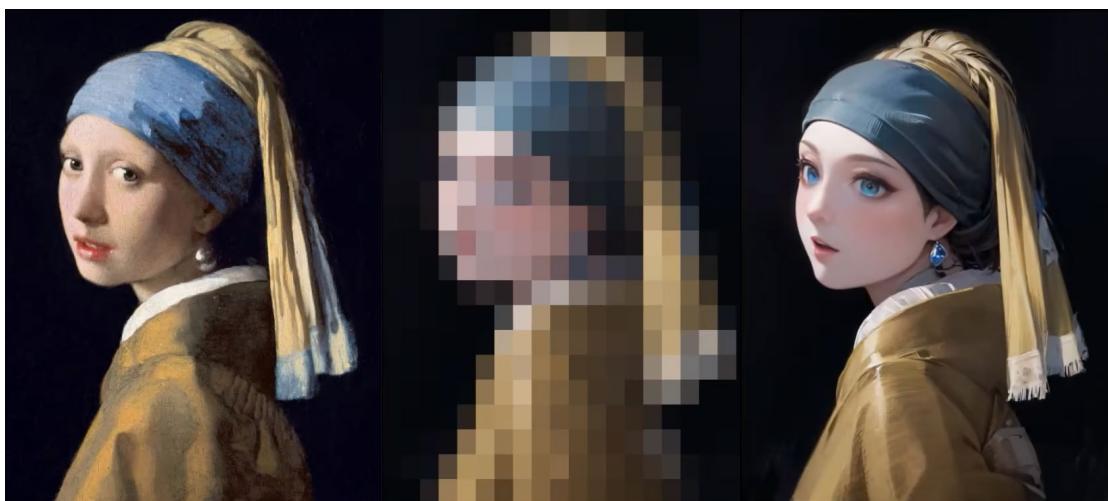
但是有一点需要注意，Civitai 的服务器在国外，而且目前屏蔽了国内（包括港澳台）的访问，所以去 Civitai 下载模型需要一点“进口魔法”。因此在这里推荐两个替代网站，这两个网站是在国内架设的服务器，从 Civitai 搬运到国内，访问和下载不需要魔法。

<https://liblib.ai/>  
<https://www.liandange.com/>

由于是搬运，难免有缺少的情况，可以先在这两个网站上寻找，没有的话再去 Civitai。但是需要声明，这种类似的搬运网站有时会不标注原作者并且作为牟利使用，不要为任何模型付费，如果可以的话使用模型尽量去原站点评论点赞原作者。

## 五、图生图

通过提示词描述图片再让 AI 生成难免出现偏差，因为 AI 绘画充满随机性，但是这时如果再丢一张图给 AI 作为参考，AI 就跟更理解你的需求。AI 会从参考图中获取颜色、衣服、姿势等原本图片上记录的信息，最终成品会具备参考图上足够多的信息。

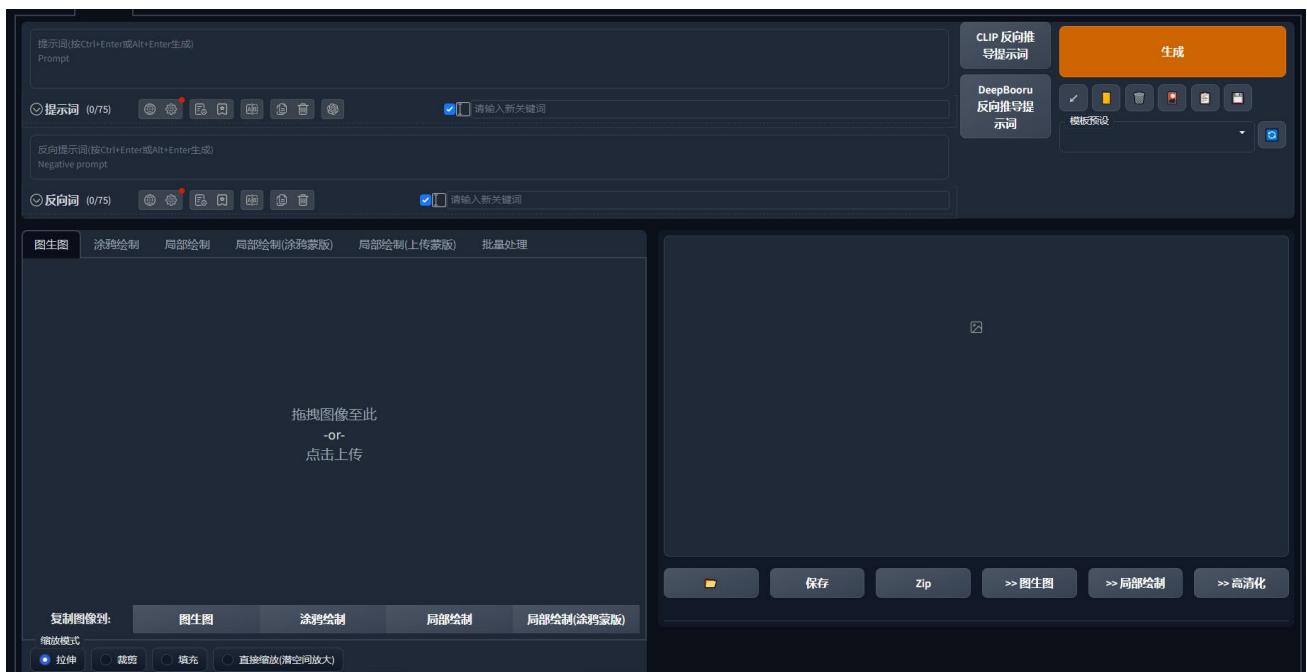


## 基本步骤

### 上传图片



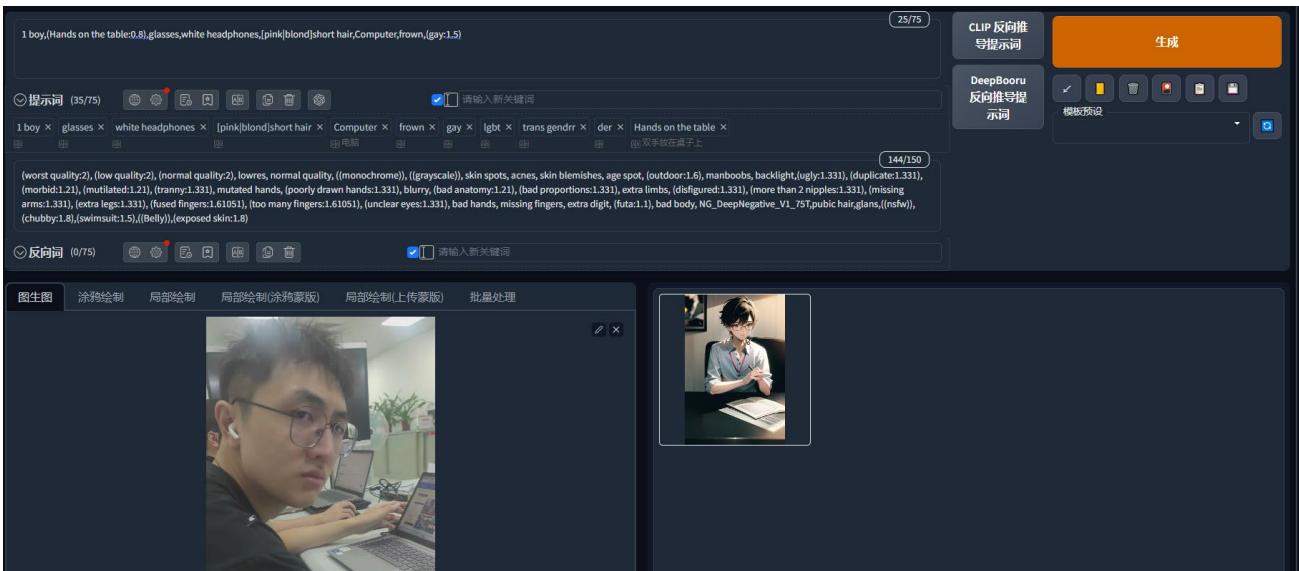
点击图生图标就进入到图生图操作界面，界面总体和文生图差距不大，也有参数调整等界面（下面的截图没截到），唯一不同的是多了图片上传模块。



首先准备一张图片，可以直接拖动照片上传，也可以选择文件上传。在上传文件的下方有局部重绘等功能，在之后的章节将会提到，可以先不用修改这部分参数。

### 书写提示词

上传完图片之后，**提示词中需要描述想要保留的图片特征**，比如下图的 1 boy, glasses 等，然后按照之前的书写提示词的格式，对成片进行描述即可。反向提示词与之前提到的相同，直接复制一些通用模板即可。



## 其他参数



绝大部分的参数和文生图是完全一致的，但也有一些独特的选项，比如重绘强度。重绘强度其实也不难理解，就是与参考图相似的程度，强度越低越相似，一般推荐在 0.6~0.8。

## 参数解析

首先是提示词，要求提示词足够精确、具体，像以下失败的案例，很明显就是 AI 只读取了图片信息，但是没有理解画面的具体内容是什么。



对人物主体个数、性别描述清楚，如果想要描述的更为精确，就加入针对性描述外貌和服饰细节的内容提示词。如下所示。之前也学习过，为了更好的质量，还需要标准提示词，这里就不再赘述。

1 man, short hair, blue eyes, beard, red plaid shirt, woolen hat,  
翻译：短发，蓝眼睛，胡子，红色格子衬衫，羊毛帽

对于成片的分辨率，建议根据自己的设备选择合适的宽度高度。建议提前裁切参考照片的长宽比例，让其与成片比例一致，否则可能会导致成片拉伸、变形。如果比例不一致，会有一下几种缩放模式，作用如字面意思很好理解，试一下就知道了。需要注意的是，放大潜变量不建议使用，对显存要求较高。



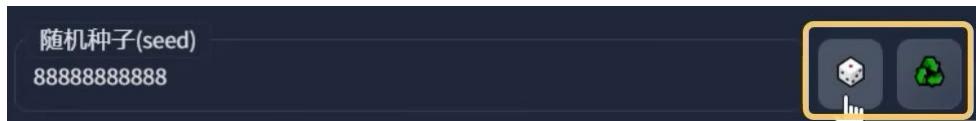
### 随机种子

举个例子，如果原图是在室外拍摄的，但是生成的图片是在室内，我们就需要在提示词中加入户外相关的提示词再生成一次，将其纠正过来，这样生成的图片就是在户外了。



但是这样做的问题在于，我们对背景改动的同时人物也发生了变化，为了不改变人物形象，这时我们就需要用到随机种子。AI 生成一幅画是随机的，每一次生成都有一套描绘方式，而这描绘方式就是随机种子。

如果使用了同一个随机种子，那生成的图像必然会有许多相似之处，有的种子效果好，有的种子效果可能就差。



随机种子还提供了两个功能，筛子是把随机参数设置为-1，就是随机。绿色循环按钮就是把种子设置为上次生成所使用的随机数，这样就把随机数固定下来。同时我们在生成完照片之后，在成图的下方中有生成图片相关的信息，seed 后就是生成这次图片所使用的种子，如下图所示。



那么如果我们只有成图的时候，该如何获取图片信息呢？这时就可以使用SD自带的图像信息工具。进入到图像信息选项卡，只要上传图片就可以获取生成图片的相关信息，但是有时如果图片被裁切或者二次编辑过，生成的信息可能会有错误或者识别不出。



## 应用场景

像把真实的人物照片变成二次元风格等过于常见，就不在赘述。自然还可以反过来，将二次元或者手绘，通过使用真实系的模型将其转变成真实的人物图片。如果想要更加具体准确地呈现某个特定的人物形象，可能还需要结合这个人物的LoRa模型，之后我们会讲到。

有时可以把物品或者场景作为参考图，提示词中尝试用描述人物的方式描述它，最终就会获得物品拟人化的图片。



还有更为进阶的使用方式，我们可以在 Photoshop 等设计软件中画出图片的框架，再借助图生图补充画面细节。有时甚至画的图片十分抽象也可以生成不错的图片，之后我们也会讲到这部分的内容。



我们随便画一个草图，如上图，红色的晚霞，落日，大海、沙滩，还有一位少女。之后我们将这张图导入到 SD，然后把我们刚才场景使用词条话的语言描述清楚，点击生成就可以获得一张精致的插画。

1girl, solo, full body, long hair, black dress, silhouette, standing besides the sea, view from back, arms up, embrace, (setting sun on the horizon:1.5), sunset, sunset glow, sunset cloud, (beach:1.5), clouds, coast, seaside, gradient sky, depth of field, big sun, sunlight, backlight, rays, fantasy, | (masterpiece:1.2), best quality, masterpiece, higres, original, extremely detailed wallpaper, perfect lighting,(extremely detailed CG:1.2), drawing, paintbrush, outdoor, forest, wild field, travel, woods in the background, depth of field

NSFW, (worst quality:2), (low quality:2), (normal quality:2), lowres, normal quality, ((monochrome)), ((grayscale)), skin spots, acnes, skin blemishes, age spot, (ugly:1.331), (duplicate:1.331), (morbid:1.21), (mutilated:1.21), (tranny:1.331), mutated hands, (poorly drawn hands:1.5), blurry, (bad anatomy:1.21), (bad proportions:1.331), extra limbs, (disfigured:1.331), (missing arms:1.331), (extra legs:1.331), (fused fingers:1.61051), (too many fingers:1.61051), (unclear eyes:1.331), lowers, bad hands, missing fingers, extra digit,bad hands, missing fingers, (((extra arms and legs))),

A detailed illustration of a girl standing on a beach at sunset, matching the silhouette in the草图. The girl has long, flowing hair and is wearing a black dress. She is standing on a sandy beach, looking out at the ocean under a vibrant sunset sky with orange, pink, and purple hues. The water reflects the warm colors of the sky.

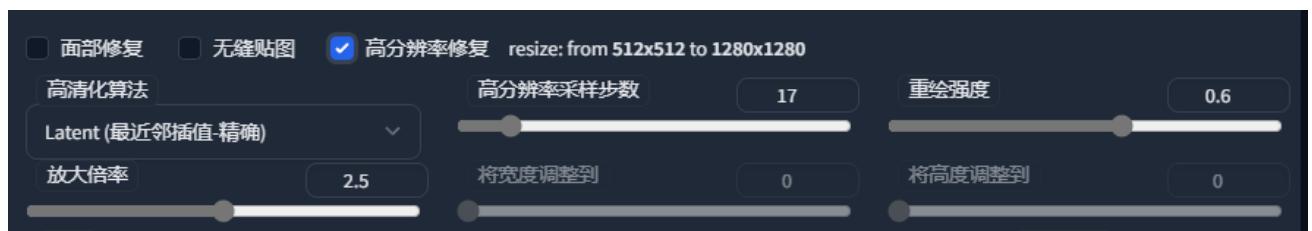
## 六、超分辨率

我们之前学习了许多让 AI 绘画更精致的方式，包括提高分辨率、添加质量标准化提示词等。在 SD 中，除了以上学过的方式，还有一系列功能专门用于优化产出图像的质量，例如高分辨率修复、Upscale 脚本以及图片放大算法。

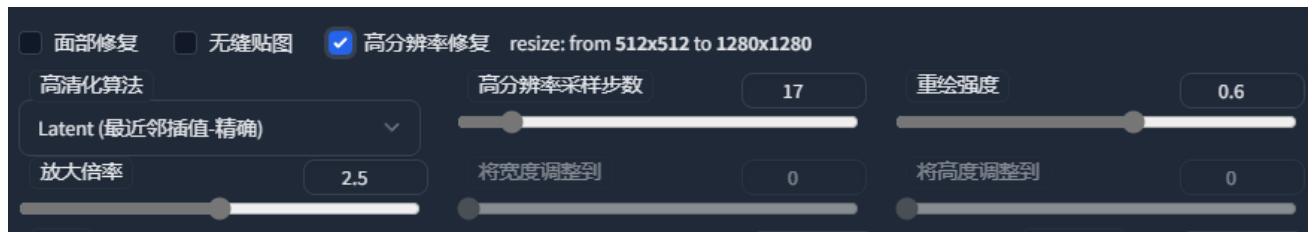
不同的生成图片的方式也会采用不同的优化方式，文生图就会采用高分辨率修复，图生图采用图片放大，对于生成之后的图像，采用附加功能（有时也翻译成高清化）中的放大操作。



### 高清修复



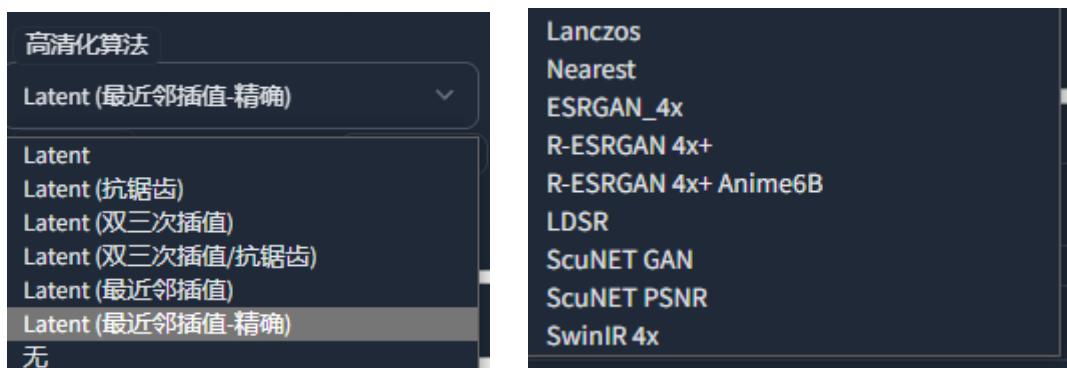
文生图中有一个高清修复 (**Hires Fix**)，之前我们在第三章中也使用过这个功能。在所有参数都一致的情况下，下左图为未开高分辨率修复，图像十分模糊。  
这是生成真实系图片很明显的问题，当分辨率过低时，AI 没有足够的操作空间去生成细节，而在二次元图片中就不会太明显。而下右图开启高分辨率修复之后就十分逼真。



在使用高清修复的时候，首先在生成之前勾选它，然后会出现两排额外选项。

第二排的参数用于设定目标图片分辨率，可以按倍率确定，也可以确定具体数值。

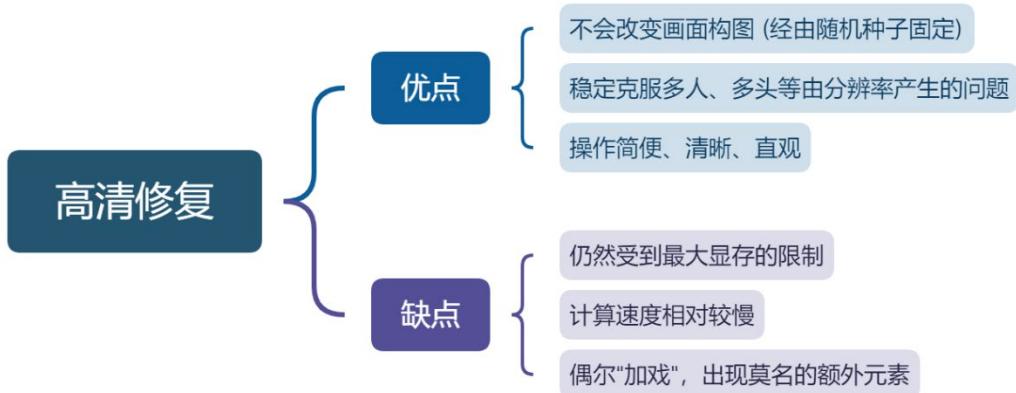
第一排的高分辨率采样步数等同于采样步数，因为高清修复本质上就是对图像重绘。重绘强度等同于图生图中的重绘强度，如果想保持画面结构不发生大变化，最好不超过 0.5。



高清化算法类似于之前采样器，每种算法在使用时都大同小异。没有绝对好的高清化算法，只有最合适的，所以需要多次测试才能得到心仪的图片。如果要

说最稳妥的，这里推荐 R-ESRGAN 4x 和 R-ESRGAN 4x Anime6B。其中第二种适合于二次元图片。

高清修复有以下几种优缺点，如下图所示。一种比较好的做法就是，先以低分辨率反复生成，确定一个合适的画面之后再固定种子进行高清修复。



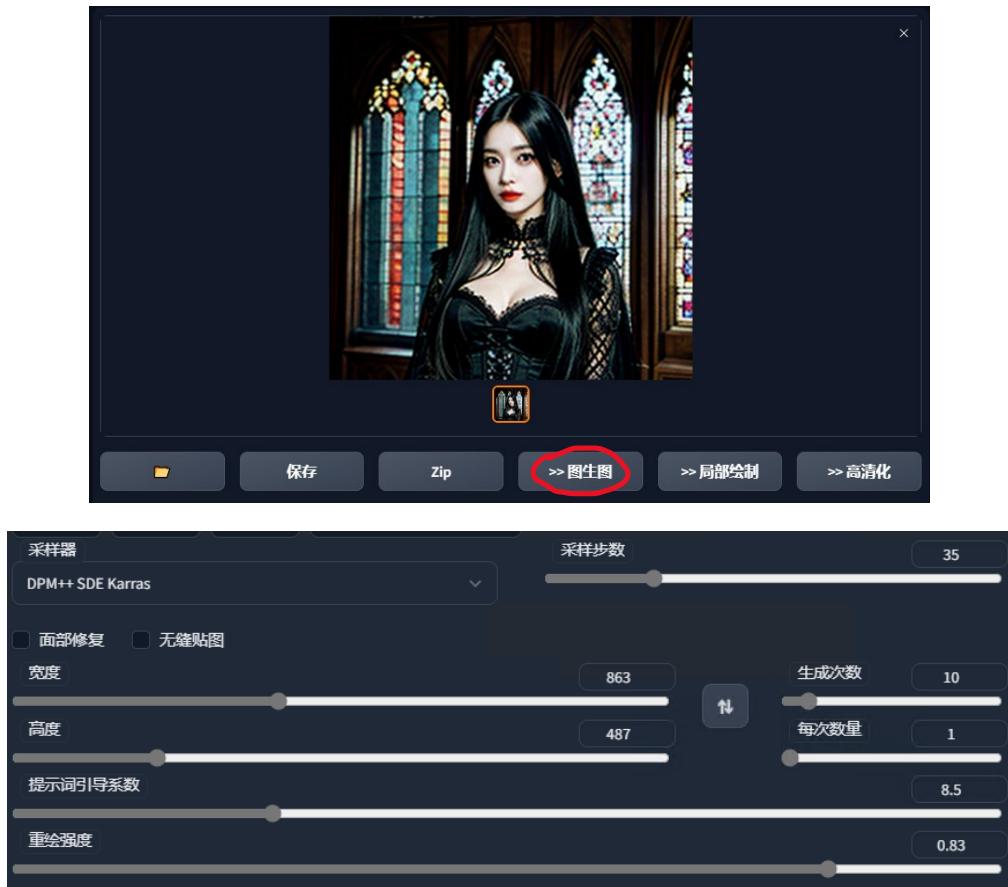
需要注意的是，高清修复有时会改变画面结构，增加一些元素等等，这些可以通过降低重绘强度来解决。当重绘强度过低时，边缘就会模糊，过高时（大约 0.7 左右）结构会变化，（大于 0.9）甚至会出现多头。



如果只是为了变得清晰，安全的重绘强度区间为 0.3~0.5；如果想让 AI 发挥更大创作空间，就设在 0.5~0.7。

## SD Upscale 脚本

图生图中是没有高清修复选项的，因为图生图本质上就算是一种高清修复，所以我们在文生图得到低分辨率的图片之后，可以点击图片下面的图生图按钮，直接导入到图生图界面中，他会自动把所有参数同步。



之后按照图生图部分的内容操作：选择采样器，调整分辨率和重绘幅度等，就可以生成高清的图片。如果不想有大的改动就调低重绘强度。

如果想让图生图的图片更大更清晰，就需要用到 Upscale 脚本，他的基本原理就是把图像划分成几张小图，分别优化，最后再合成为一张。

在随机种子下面的脚本选项中，选择 SD 模型放大，不同的汉化插件会有不同的中文翻译，注意区分。



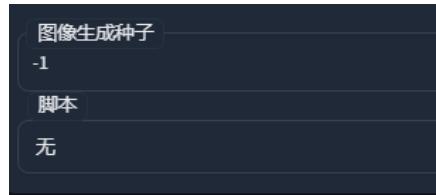
选择之后会有如上几种参数，其中缩放比例就类似于高清修复中的放大倍数，而高清化算法与刚刚看到的那几种也是差不多的。而最核心的参数就是**图块重叠范围**。前面提到过，SD 放大脚本就是把图像分成几块分别渲染，那在合成的时候如果没有缓冲区，直接拼接就会显得很不自然，如下左图。这时就需要每个小图中间有缓冲的部分，就像粘双面胶必须有重叠的部分一样，这样就很自然，如下右图。**图块重叠范围**就是控制重叠范围的大小。



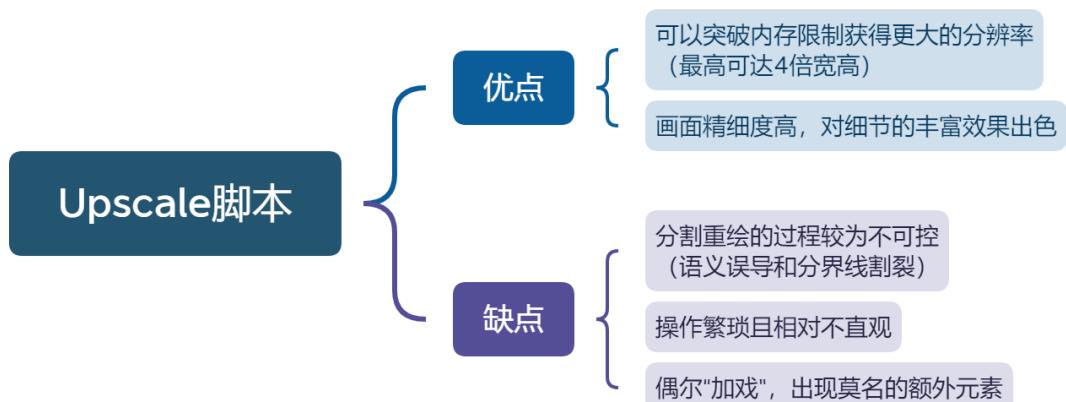
一般默认重叠范围大小是 64，如果你觉得还是过于生硬就设置为 128。需要注意的是，这里有个公式：最终图片宽高= (设置的宽高-重叠像素) × 放大倍率。所以在图块重叠范围设置成 64 时，需要在宽度和高度上加上 64。如下图的 576 就是  $512+64$ ，放大倍率为 2 时，最终成图宽高为  $1024 \times 1024$ 。



在设置→放大选项卡中，还可以自行定义图生图放大过程的算法，但是再没有完全精通 SD 的情况下建议不要修改设置。



SD Upscale 脚本也不是十全十美，因为每张图片都有缓冲区，所以每张图片的大小是偏大的，但是由于是分部分渲染，所以压力是比高清修复要小的。



放大四倍是理论上可行的操作，实际操作中，因为原图放的过大 AI 很难将提示词中每个区块的内容对号入座，就容易让画面混乱，所以还是建议 2 倍左右多次尝试。

此外，有时分割线正好会在人脸等关键元素上，可能会产生不和谐的画面，解决办法就是降低重绘幅度，和增大图块重叠范围。

## 附加功能高清化

最后是对图片进行后期处理的时候，利用附加功能（有些翻译插件也会译成

高清化、后期处理) 做放大, 可以对任意一张图片 (不限于 AI 绘图) 通过 AI 放大到一定尺寸。在原理上, 它与市面上的图片修复功能是相似的, 可以理解成重绘幅度为 0 的高清修复。操作同样也与之前的方法相似, 选择放大倍率和放大算法, 但是不同的是允许选择两个放大算法。



选择第二种算法之后, 可以再给它设置一个 0-1 的可见度, 决定了第二种算法的效果程度, 如果觉得复杂也可以不选。CFPGAN 和 CodeFormer 用于配置脸部修复比例, 保持默认值即可, CFPGAN 可以稍微提高到 0.2。如果对于脸部效果不太满意, 可以调节这两个参数进行修改。



可以看到放大之后，分辨率确实是设置的数值，但是效果不明显，只是将线条边界分清楚，将模糊色块更加清晰，效果肯定不如前两种改变绘画方式的修复。甚至有时放大还不如不放大。



## 七、局部重绘

当借助 AI 画出了一张满意的图片，但是却有一些错误的时候，就需要**局部重绘 (inpainting)**，只针对这些错误进行修改。就像写作业使用涂改液一样，针对一张大图的某个区域覆盖重画。

当然，我们可以通过加入提示词，固定种子来重新画一遍，就像之前做的一样。但是由于提示词的改变，画面的整体结构还是会改变的，如我们想让下左图闭上眼睛，生成的却是下右图。所以如果我们要对画面的 99% 都是满意的时候，就需要局部重绘。



### 基本操作

将一张图进行局部重绘的方式有很多，最简单的肯定是成片下面的局部重绘按钮（不同中文插件翻译略有不同），点击之后就会自动把参数发送到图生图界面中，因为局部重绘是图生图的一个分支。



点击局部重绘后就会自动导入到局部绘制中，在这里可以使用 AI 画的图，也可以上传其他图片。把鼠标放到图片上，鼠标这时就变成了一个画笔，用它涂抹覆盖想要 AI 重绘的区域。如果是上传蒙版就不需要手绘，控制区域更为精确。



这里有两个小按钮，一个是撤销，一个调整画笔粗细，可以让你更精确地控制重绘区域的大小。覆盖住眼睛之后，在提示词中加入 close eyes 并提高一些权重（如 1.2），在稍微提高一点重绘幅度（如 0.75）。此时我们可以先生成一次，看一下效果。



我们可以看到闭眼的效果虽然不错，但是感觉还是不太自然，可能有人已经发现了，在源图片下方比图生图还多了很多参数，想要进一步把控效果就需要调整下面的局部重绘参数。如果没能闭眼，说明重绘强度或者提示词权重太低。如果调到很高还是不行，就放开随机种子，多跑几次。如果一直不闭眼，就要考虑提示词是否输错了（逗号输成句号等）。

总结图生图的局部步骤就是：导入图片到图生图→描绘重绘区域→加入重绘提示词→提高提示词权重→提高重绘强度→修改重绘参数。

### 参数解析



在图片下方又比图生图多了一系列的选项，和高清修复一样，他决定了重绘时的实现方式。这里需要解释一下蒙版 (mask) 的意思，蒙版是图像处理领域的

专业名词，泛指用于限定处理区域的范围对象，我们画笔涂出的黑色区域就是一个蒙版。

蒙版模式中，如果选择的是绘制非蒙版内容，那么涂黑的部分就不会重绘，没有选中的部分都会重画。一般都会选择绘制蒙版内容。

蒙版蒙住的内容，可以理解成 AI 拿到的需要重绘的图像信息。选择原图 AI 拿到的就是原图。如果想让 AI 发挥更多想象，可以选择填充，它就会把原图高度模糊之后在给 AI 去生成。后两种让 AI 生成图像的过程更为复杂，理论上得到的效果也会更高，可以综合这几种方式多试几次，和选择放大算法一样，有点看运气。

重绘区域决定了 AI 绘画的逻辑，如果选择的是全图，AI 会基于新的要求（提示词、参数）把整张图重新画一遍，但最后只保留蒙版选中的区域，并把它拼回去。如果选择的是仅蒙版，AI 就只会画框出来的区域，然后再拼回去，所以会快一点。仅蒙版由于没有获取图片全貌，所以经常会出现问题，推荐使用全图。

如果选择的是仅蒙版，旁边的仅蒙版绘制参考半径（像素）作用类似于 Upscale 脚本中的图块重叠范围，发挥类似缓冲带的作用，可以保证较好的拼接效果。此选项的大小随着蒙版区域变大而变大。

蒙版模糊度决定了重绘区域边缘和其他部分是如何接触的，它与 Photoshop 中的羽化作用相同。如果蒙版模糊为 0，重绘四周与原图接触就不和谐，10 以下的数值让拼接更为顺滑，如果开太大会影响其他区域。此选项的大小也随着蒙版区域变大而变大。

## 涂鸦绘制与涂鸦蒙版

尽管局部重绘可以圈住一小块重新生成，但是还是不够强大，举个例子，如果让 AI 修复画坏了的手，即便区域被精确的选中，AI 既然能画坏一次就能画坏无数次。



在局部重绘的标签旁边，还有另外几种重绘的标签。其中**局部重绘（涂鸦蒙版）**可以帮助我们增强局部重绘的可塑性，注意，不同的中文插件可能翻译不同，位置对应即可。



使用它与使用局部重绘大致相同，不同的是在调整笔粗细的下方多了一个拾色器，这样**画笔的作用不仅仅是画出蒙版范围，而具有了一定的输入内容的能力**。如上图我想加入一个紫色的帽子，就用紫色画笔大概画一个帽子的形状。

绘画完成之后，需要在提示词中加入对应的提示词，稍微拉高权重（如 1.2 以上），然后重绘幅度稍微增大一点（如 0.8 左右）。之后就可以直接生成。你可能发现涂鸦蒙版比局部重绘还多了一个选项：**蒙版透明度**，用于控制重绘内容的透明程度，默认 0 即可。



如果增加的画面内容比较复杂（如增加一个带有蕾丝和花朵的紫色哈雷帽），就需要降低权重和重绘幅度，更稳定地传递我们想要绘画的内容。正如你想象的那样，如果你有数位板等专业的绘图工具，就可以更为精细地绘画，增加更为复杂的画面。

现在回到画手的例子，我们可以先吸取背景的颜色。然后吸取肉色把手部轮廓勾勒出来。之后就要增加提示词，如 5 fingers, detailed hand, high five 等描述手的形状，最后直接生成即可。



如果 AI 还是不能准确画出理想的样子，就可以把随机种子设为-1，多生成几次，不断调整参数。



其实涂鸦蒙版是可以单独使用的，即涂鸦绘制（不同插件翻译不同）。具体的

操作也是和涂鸦蒙版差不多的，因为没有了局部重绘，如果加入了东西，就会重新生成整张图，这肯定会导致其他部分的细节发生变化。

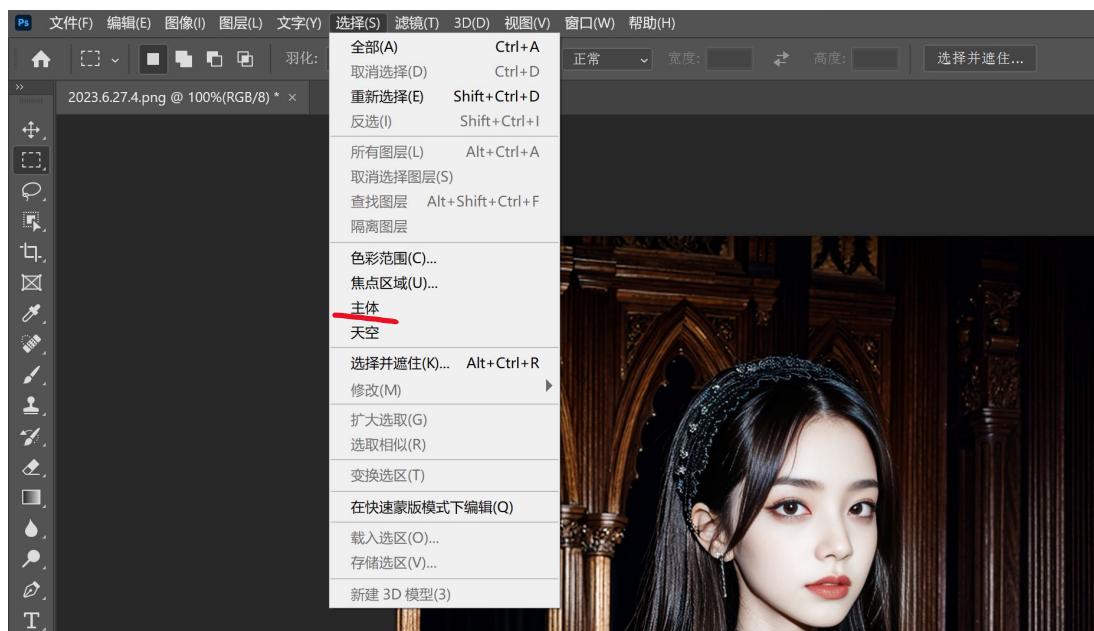
### 上传蒙版

这一功能适合有图像处理软件基础的朋友使用，可以更加精确地控制重绘区域范围，进入**上传蒙版**选项卡之后，需要上传两张照片，一张是需要修改的照片，一张就是蒙版。

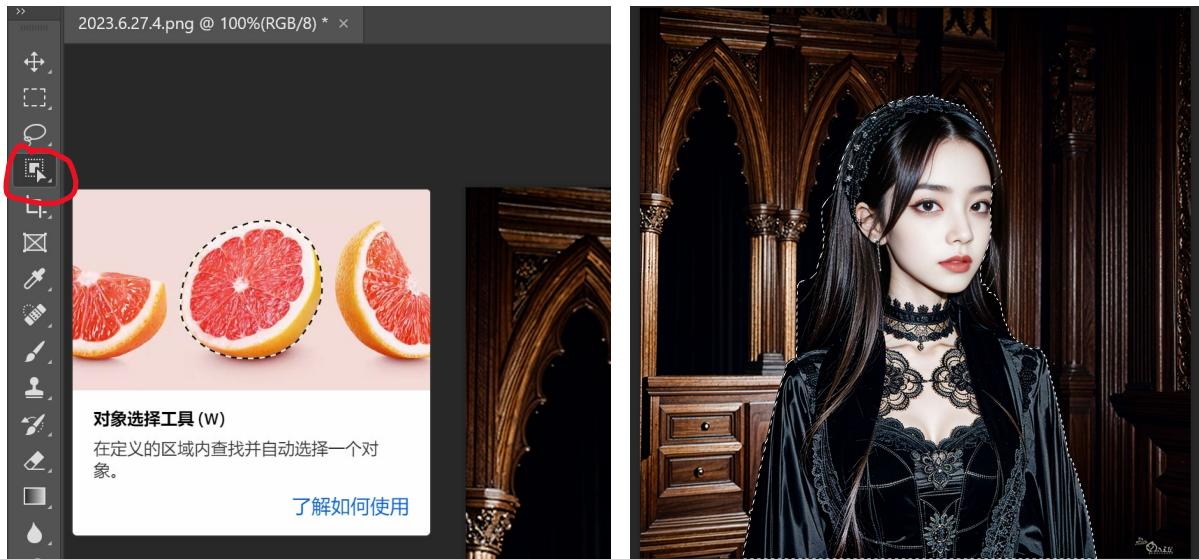


这一功能的作用就是让你可以使用 Photoshop 等软件精确选定重绘范围，在这里就不过多介绍 PS 相关的操作了，只介绍如何使用 PS 制作蒙版。

我们以重绘人物为例，使用 PS 制作蒙版。

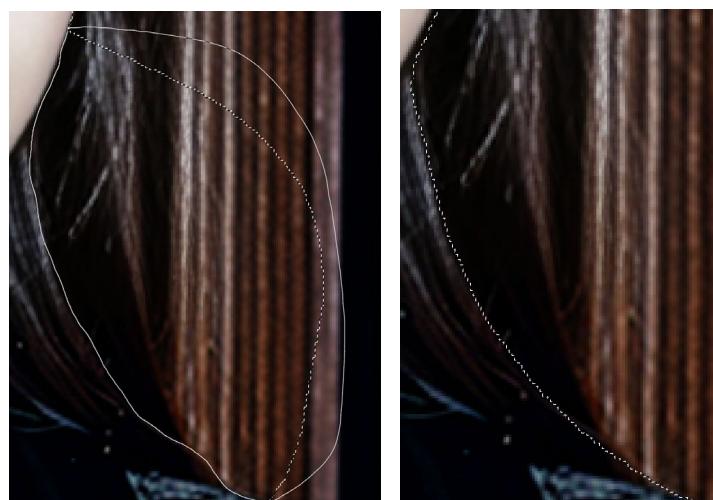


首先导入图片到 PS 中，有两种方式选择框住的主体。一种是在选择菜单中的主体选项，点击后就可自动识别主体人物。但是当画面中人物较多元素较为复杂时，更为推荐使用工具栏中的对象选择工具，下左图所示。

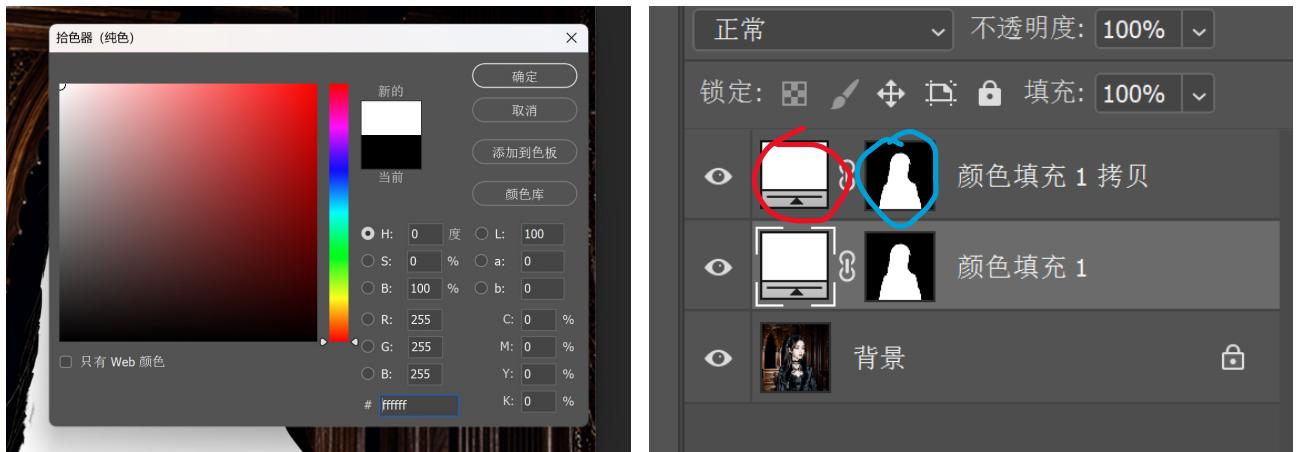


使用对象选择工具框住想要选择的主体，即可自动识别出主体，如上右图。如果生成的选区有残缺或者多余的部分，就可以切换到套索工具（在对象选择工具的上面），手动选择区域增加或者删除。

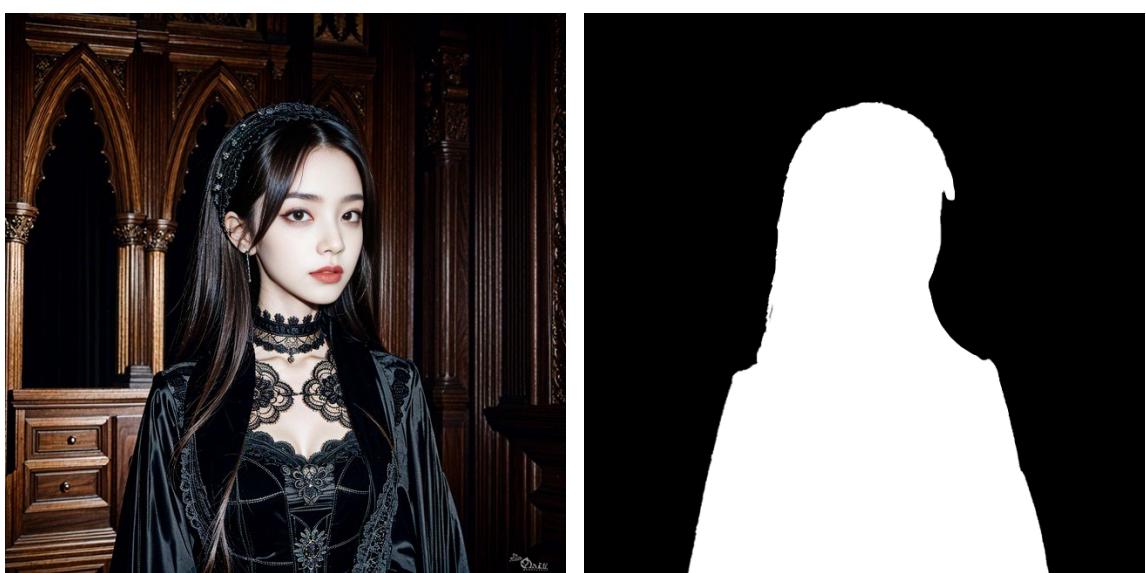
按住 alt 滚动鼠标滚轮即可放大或者缩小图片，按住 ctrl 滚动滚轮可快速左右移动图片。选择了套索工具之后，按住 shift 用鼠标选中区域即可增加选区，按住 alt 选中区域即可删除选区。



生成了一个选区之后，在最上方的菜单栏中的图层→新建填充图层→纯色，点击确定之后生成图层，之后拾色器选择白色，就会让白色区域填充选中的主体，点击确定，如下左图。



生成图层之后在右边选项卡中选择刚刚创建的纯色图层， $\text{ctrl}+\text{j}$  复制一层，再双击白色方块，如上右图红色部分，使用拾色器将填充颜色设置为黑色。之后点击蓝色圈住的地方选择图层， $\text{ctrl}+\text{I}$  交换蒙版区域，这样我们就得到了一张蒙版图像，如下右图。



之后在上方菜单中的文件→导出→快速导出为 PNG，导出蒙版。之后把蒙版上传到 SD 即可。需要注意的是，选中区域精准只是确保生成的图片保留的原图

部分不缺失，不能保证边缘过度平滑，可以尝试降低重绘幅度、使用后面会讲到的 ControlNet 中的 inpaint\_global\_harmonious 和 sd15\_inpaint 以及其他各种方法协作完成画图。

到这里局部重绘的基本操作就讲解完毕了，剩下的就是发挥创意修图了。



比如可以拿一张真实的手部图片覆盖掉想要修改的部分，然后放进图生图中生成，就可以修复手部的问题了。还可以先在 PS 中画出精细的轮廓，然后导入再生成。还可以在 PS 中单独截取手部然后重绘。

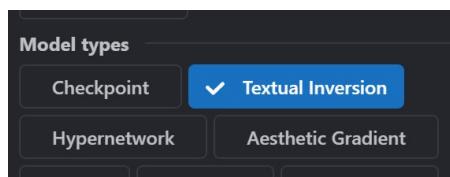
## 八、模型进阶

通过提示词告诉 AI 要画一只猫，它就会检索和猫有关的信息，AI 的所有智慧和理解都来源于人类输入给它并让它学习的知识。

在 AI 学习的过程中，Checkpoint 是最大的模型，就像是一本新华词典，记录的信息非常全。在 Checkpoint 的支持下，当你通过提示词告诉 AI 很少见的事物的时候，它也能准确描述出。就像给小学生一本新华词典，让他写出魑魅魍魉一样。

除此之外，还有其他几种专一的小模型，相对于类似于新华词典的 Checkpoint，他们更像是专业词汇词典，专用于某一领域。这些模型可以用于针对与特定人物的创作、画风自由切换，甚至角色的三维图绘制，最主要的还是解决画手的问题。其中比较容易接触到的小模型有：文本嵌入 Embeddings、低秩模型 LoRa、超网络 Hypernetwork。

### 文本嵌入 *Embeddings*



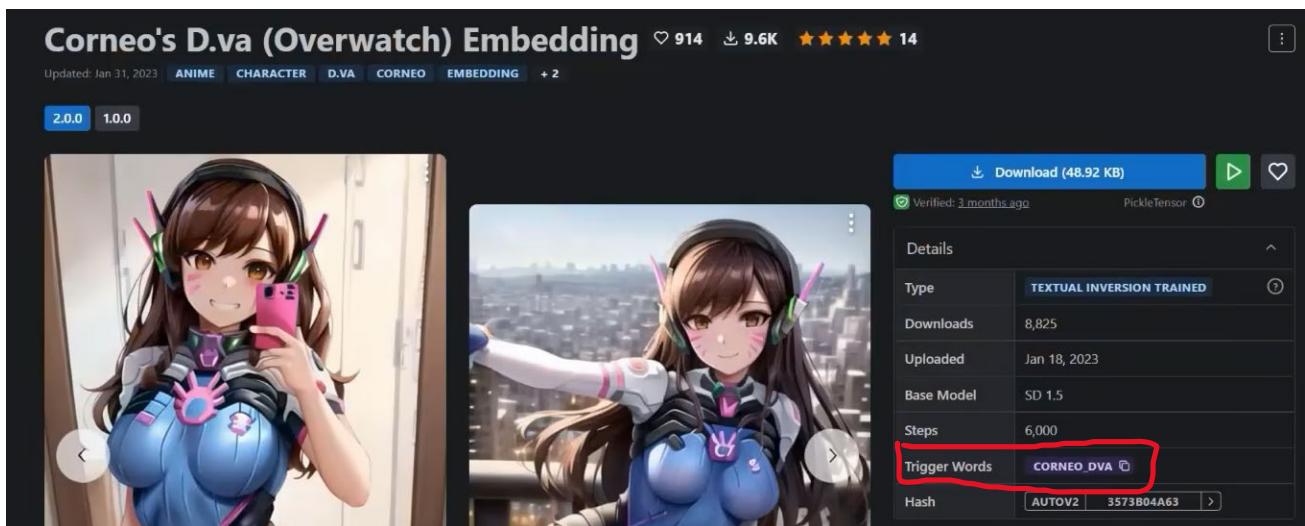
在 Civitai 等模型网站上，这类模型也被称为文本倒置 Textual Inversion。点进这些模型的下载界面，会发现它们一般都非常小，甚至只有几十 KB。Embedding就像是一个书签，能够精准得指向某个词或者字的含义，从而使得 AI 高效索引，还可以帮助我们画出特定形象。

举个例子，如果我们想让 AI 画“猫又”（半人半猫的妖怪形象），AI 不知道猫又是什么，但是它知道怎么画猫、人、妖怪。那么 Embedding 模型机会在这三

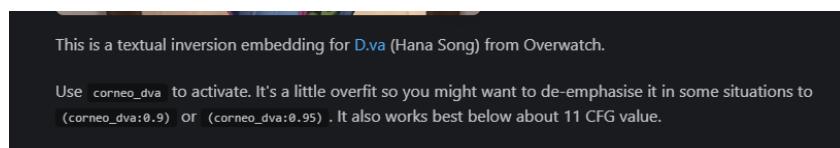
个概念上加上“书签”，当提示词中有“猫又”相关的词组时，AI就会把标有书签的信息汇总在一起，从而就理解了猫又的概念。

## 特定人物生成

首先找好自己想要训练的人物角色之后，如下图是游戏守望先锋中的游戏角色 D.VA。下载文件，文件的后缀是.pt，之后把文件放在 stable-diffusion-webui→embeddings 文件夹中。



Embedding 不需要特殊的参数，只需要使用特殊的提示词调用，而模型详情页中都会标明调用提示词，如上图模型的调用词就是 corneo\_dva。



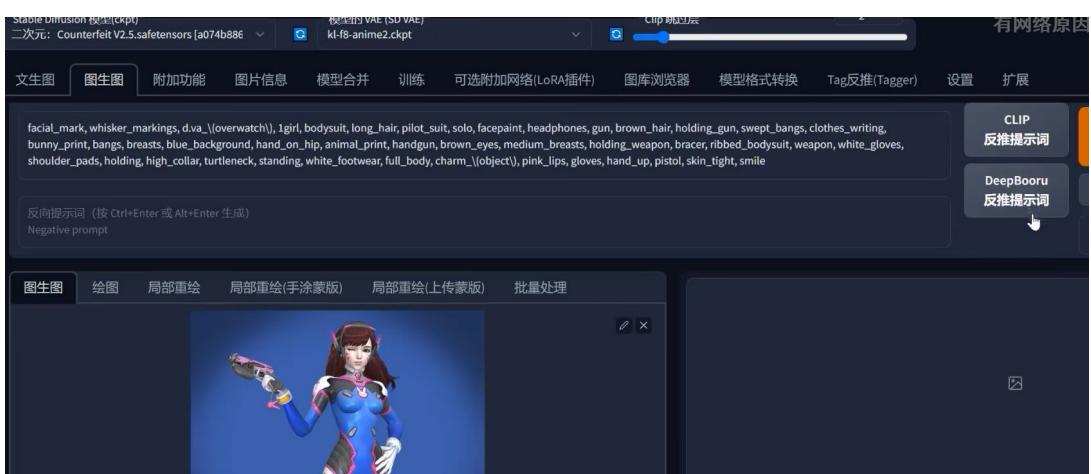
有时根据模型的不同，需要自己判断是放在正向提示词还是反向提示词中。在模型详情页中作者也会给出合适的参数建议，可以参考其进行调参。



如果你希望图片的人物更像 D.VA 一些，可以加入更多的提示词，来描述她身上本身就具有的东西。如果不知道如何描述，可以使用图生图中的反推，它可以识别图像中已有的成分，再结合模型中已有的词汇，生成描述图片的提示词。

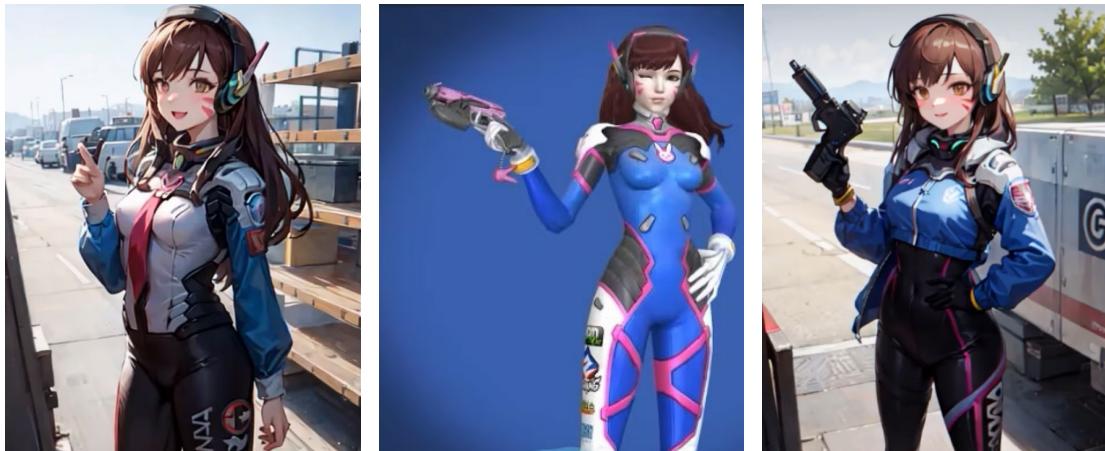


CLIP 和 DB 是两种不同的算法，其中后者在速度和精度上都有优势。导入图片后点击反推就会生成描述图片的提示词，如下图。如果报错，见附录 3 问题 9。



可以根据需要对识别的内容进行筛选，去除描述不准确的或者不需要的，符合的特征就提高权重等等。之后把修改的内容提示词，加入到正向提示词中，就

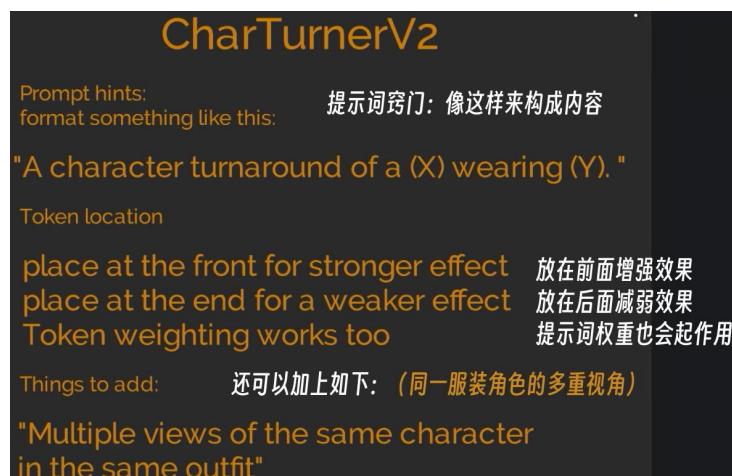
会有更多目标人物的细节。如下左图是直接生成的，下中图是用于生成描述人物提示词的图片，下右图是加入人物描述提示词后再生成的图片。



但是 Embedding 在怎么说也只是“书签”，只是让 AI 按照特定的方式搜索信息，并没有额外附加信息，所以我们针对特定人物的创作更加习惯使用 LoRa，之后会详细讲到。

### 三视图

另外一个应用是通过 CharTurner 实现的人物形象三视图设计，它其实就是把几种人物朝向并列的图片结构作为“书签”插入到了 AI 的字典里。



对于 CharTurner 的使用也是通过加入提示词的方式，但是略微复杂一些。以上是作者提供的使用 CharTurner 的句式，还有许多注意事项在原模型的网页中，

这里就不摘抄了。下面来举个例子。

```
A character turnaround of a (corneo_dva) wearing blue mecha bodysuit, (CharTurnerV2:1.2),  
(multiple views of the same character with the same clothes:1.2), ((character sheet)), (model sheet), (turnaround), (reference sheet), white background, simple background, character concept, full body, SFW, masterpiece, best quality,  
facial_mark, whisker_markings, d.va_(overwatch), 1girl, (blue:1.4) bodysuit, long_hair, pilot_suit, solo, facepaint, headphones, gun, brown_hair, holding_gun, swept_bangs, clothes_writing, bunny_print,  
breasts, blue_background, hand_on_hip, animal_print, handgun, brown_eyes, medium_breasts, holding_weapon, bracer, ribbed_bodysuit, weapon, white_gloves, shoulder_pads, holding, high_collar, turn  
standing, white_footwear, full_body, charm_(object), pink_lips, gloves, hand_up, pistol, skin_tight, smile
```

如第一行是作者提供的句式，其中 corneo\_dva 表示人物本身的特征，之后着重描写服饰，CharTurnerV2 是模型激活提示词。之后第二行就是模块化的固定图片整体风格的提示词，之后的就是普通的标准提示词和内容提示词。

需要注意的是，可以在一次生成中激活不同的 Embeddings，就像上面的例子一样，但是有可能会冲突，需要自己合理考虑。它还有很多问题，比如人物细节稳定性和具体转身动作幅度等等，之后通过 ControlNet、LoRa 等可以更为精细地操控。

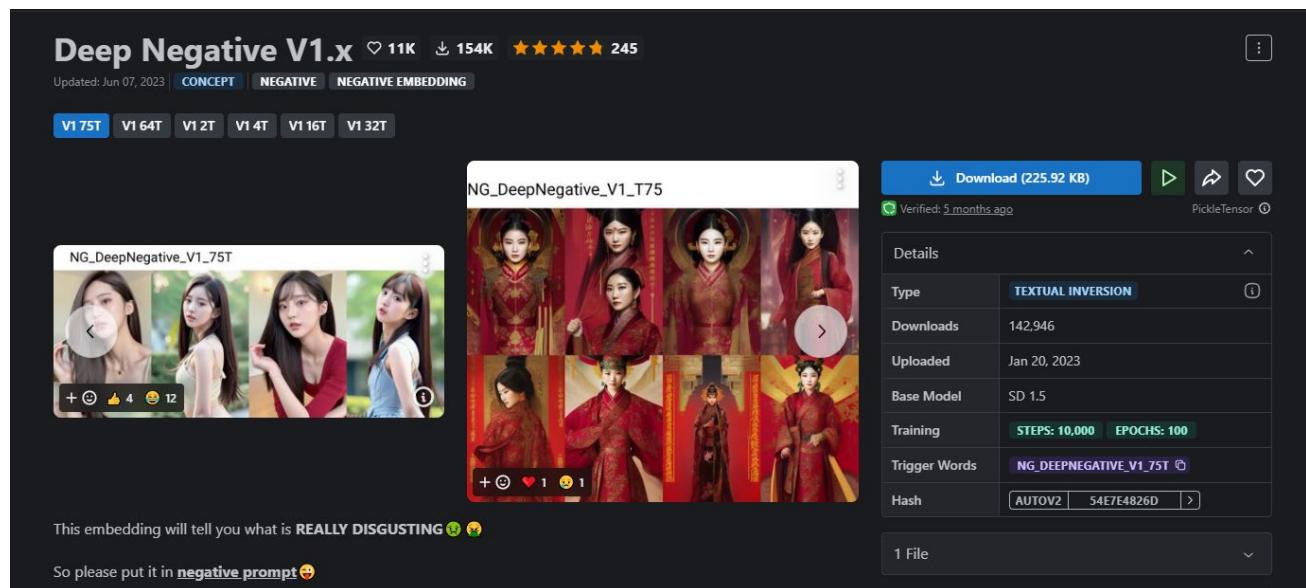
## 手部修复

想要 AI 画出正确的东西，只需要让他避开错误的案例。和之前使用 Embedding 一样，下载文件之后放到文件夹中通过提示词激活，但是有一点不同，这些告诉 AI 错误案例的 Embedding 需要把激活词放到反向提示词中（可以稍微提高权重到 1.2）。



这类加入到反向提示词中的 Embedding 也被称为**负面词嵌入**，如 Easy Negative 是综合的负面样本提炼，解决的问题可能还包含，肢体错乱、颜色混杂、噪点和灰度异常等问题。

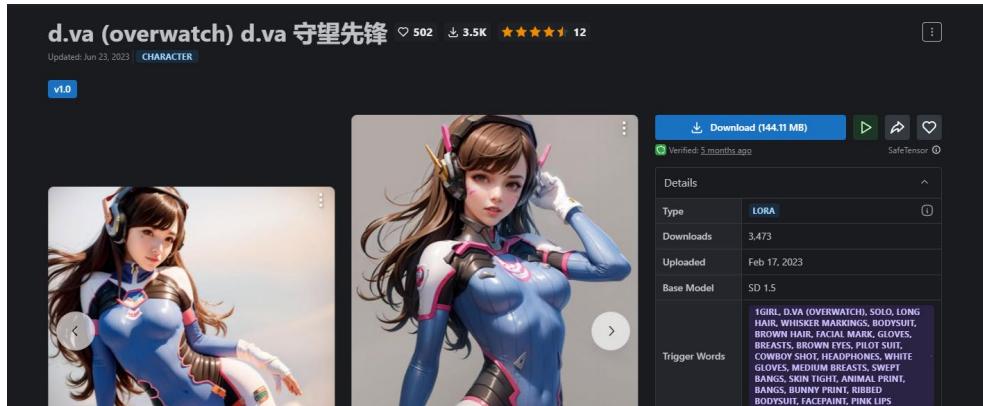
但是这类 Embedding 也不一定能解决所有的问题，不同的 Embedding 针对的类型也不同，如 Easy Negative 是针对 conterfit 模型训练的二次元图片修复嵌入，而 Deep Negative 主要是针对真实模型训练的。



### 低秩模型 LoRa

LoRa 的作用在于向 AI 传递、描述某一个特征准确、主体清晰的形象。相比于类似于书签的 Embedding 让 AI 从“字典”中东拼西凑信息，LoRa 就像是插入在字典中的附录，直接就写明了 D.VA 是谁，有什么特点，用什么方式呈现出来。

目前的 LoRa 主要是各种游戏、动漫角色的二次构建，因为训练 LoRa 需要针对一个对象的方方面面的素材，如果是人物，就需要不同的姿势、表情、动作、画风的图片作为训练素材供 AI 学习，真人的话一般素材不够全面。



找到心仪的 LoRa 模型之后，下载之后放到 stable-diffusion-webui→models→Lora 文件夹中。LoRa 的使用方式是：<lora:LoRa 文件的名字:权重>。比如上图所示的 LoRa 文件名叫 dva.safetensors，就需要在正向提示词中加入<lora:dva>激活，不加权重默认为 1。



有些 LoRa 模型也有触发提示词，说明在训练过程中基于这些提示词进行过强化训练。

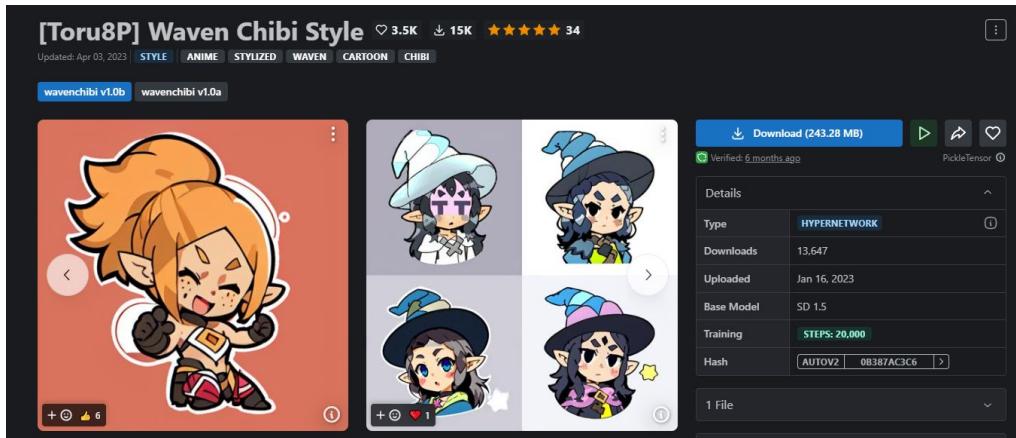
LoRa 模型在训练的时候图源十分复杂，在使用的过程中一般也会对画风构成些微影响。此时可以适当降低 LoRa 的权重，0.5~0.8 之间都可以确切地保留特征且减弱对画风风格的影响。

之后会有对 LoRa 更为细致的使用解析。

## 超网络 Hypernetwork

超网络在最后的实现效果上其实和 LoRa 是差不多的，可以通过超网络实现让 AI 学习一些原本它不知道的东西。但是与 LoRa 有一点区别，Hypernetwork —

般用于改善生成图片的整体风格，也就是画风。这里所说的画风不是像二次元和真实风那种大的差别，更像是梵高和莫奈作品之间的差别，划分十分细微。



这里我们拿 Waven Chibi Style 举例子，把文件下载到 stable-diffusion-webui →models→hypernetworks 文件夹中。使用它的方法有点绕，需要到设置→附加网络(或者叫扩展模型, 不同中文插件翻译不同)→将 Hypernetwork 添加到提示词，选择下载的模型。之后去除原本提示词中关于场景的部分，加入纯色背景(这一步是针对这一种超网络模型的特定需求)。这样就可以得到一个 Q 版插画了。



需要注意的是，现在使用 Hypernetwork 好像可以使用和 LoRa 差不多的方式，即通过在正向提示词中加入<hypernet:模型名:权重>从而调用超网络。

其实关于 Hypernetwork，它的评价大多数时候不如 LoRa 和 Embedding，它的作用其实是可以部分被 LoRa 取代，目前有很多 LoRa 已经植入了对画风的控

制。尽管如此，如果想要实现特定的艺术风格，超网络仍然是一种选择，例如雕塑风、像素风等。

## 九、LoRa

如果对于一个大模型即 Checkpoint 的出图效果不满意，最初是需要再次重新训练一个模型，而 LoRa 解决了这个问题，它是一种微调的模型方式。LoRa 的模型格式和 Checkpoint 大模型是一致的，文件后缀基本上都是 ckpt 或者 safetensor，但是大小一般只有 200M 左右，相比于大模型的 2G 左右非常小。

LoRa 要和 Checkpoint 搭配使用，从而实现对模型的微调。接下来我们介绍一下 LoRa 的使用方式和几个应用场景。

### 加载方式

第一种使用 LoRa 的方式就如第八章提到的公式一样，`<lora:LoRa 文件的名字:权重>`。另外，LoRa 是支持重命名的，可以将 LoRa 改成方便记忆的名字，方便调用。



第二种方法是图形化操作界面，更为便捷。点击生成下方的红色按钮，就会出现上右图的选项卡，安装的各种附加模型都在其中，点击想要增加的 LoRa 之后，会自动地在提示词中添加 LoRa 激活词，默认权重为 1。

在设置→扩展模型中可以修改界面的选项大小，这里就不在赘述了，维持默认肯定不会出现 bug。每个模型、LoRa 等卡片默认都是没有缩略图的，如果想为

模型增加预览，可以先用模型跑出一张喜欢的图来，然后点击下图所示的红色按钮，即可替换。



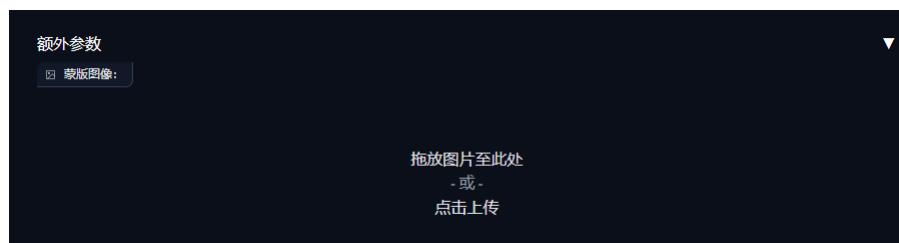
通过以上两种方式使用的 LoRa 会被记录在提示词中，但当使用的 LoRa 太多时就不方便调整参数。因此第三种方法就是借助 Additional Networks 插件（安装方法详见附录 2）。安装之后在文生图和图生图就会多出如下选项菜单。



可以选择想要增加的 LoRa 模型，权重 1 是默认效果，0 表示没有效果。大多数时候我们都会取一个 0~1 之间的值调控作用程度。



需要注意的是，这个插件默认地址和 WebUI 的 LoRa 默认地址不一样，需要到设置→Additional Networks 中添加地址，之后保存设置重启 WebUI 只有刷新模型列表即可看到。



此外这种方式还支持蒙版功能，可以让 LoRa 作用于某一特定区域。

## 人物角色类

LoRa 最广的应用方向就是对角色的把控，如何才能精确生成某个形象呢？

第一步将角色的原有形象反推，获取原有形象特征的提示词。之后在套用 LoRa 模型的时候，部分时候受限于模型，但是目前来看，现在的模型大部分都具有很高的包容度，但是如果使用风景或赛车的专用模型套用人物的 LoRa 肯定会不伦不类。使用一种模型套用 LoRa 效果不好就多试几种。



第二步自然就是启用 LoRa，在附加网络中添加 LoRa 并启用，推荐的权重在 0.6~0.7 之间，具体参数值可以多次尝试。

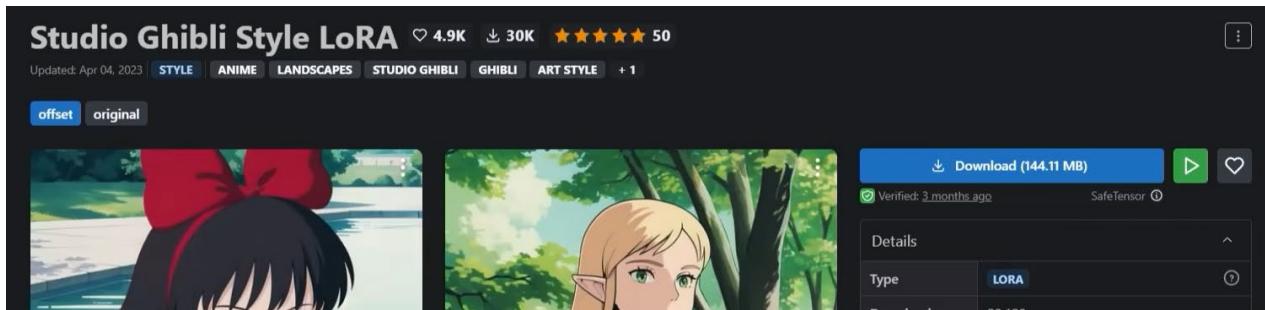
总结就是选择合适的模型，套上 LoRa，如果是文生图还可以根据角色原有的图片生成提示词，提取、整理、加权特征提示词，这样可以进一步保证图片栩栩如生。另外，如果是真实系的模型，超分辨率和其他提高细节的措施是必须的，不然会导致图片不够真实。

此外还有一些不是针对某一个特定人物的 LoRa，比如 Fashion Girl 还有 AsianMale 等等，都是符合大众审美的男性、女性面孔的 LoRa，在拟合真实系照片的时候，可以稍微加上这些 LoRa，权重调整在 0.2~0.3 之间，能稍微解决面部扭曲的问题。

## 画风类

在模型网站筛选#style 的 LoRa，就可以看到各种画风的 LoRa 模型，其中我

们使用吉卜力的 LoRa 作为示范。



之前我们也说到，可以同时使用多个 LoRa，这样我们把吉卜力画风的 LoRa，之后挑选一个喜欢人物角色的 LoRa，这两种 LoRa 是不冲突的。需要注意的是，画风 LoRa 对画面整体影响比人物 LoRa 更大，所以可以让画风的权重相对降低。



此外，画风 LoRa 还有一种特殊的 LoRa，比较概念化，它不仅包含了画风，还包含了对人物的限制，类似于画风+形象的结合体。举个例子，大部分游戏在抽卡的时候都会有一个立绘，而 Gacha splash LORA 模型就是模仿这种风格，把产生的图像也修改成类似的立绘样式（下左图是真实立绘，下右图为生成）。



这种概念 LoRa 会实现某种特定形式的输出，因此对提示词和输出参数的要求比较高，作图之前要仔细阅读模型详情页中的介绍。

### 服饰类

许多常规的服饰不需要 LoRa，因为基础模型记载了这些常规服饰的大概形象。但是对于一些特殊服装，比如十分常见的机甲少女图片，就需要特殊的服饰 LoRa 来实现。



在模型网站搜索 Mecha，就可以看到许多机甲服饰 LoRa，以机甲服饰为例，如果想生成这种类型的图片，可能还需要中加入如 cyberpunk、futuristic 等提示词作为辅助。想要 LoRa 与图片人物足够贴合，可能还需要 robotic arms/legs、

mechanical parts 等提示词，来告诉 AI 你想让人物的哪部分穿着服饰。以下是可以用于机甲的提示词参考，其他的服饰模型也可以用于思路参考。

*mecha, intricate mechanical bodysuit, mecha corset, mechanical parts, robotic arms and legs, headgear, caustics, reflection, ray tracing, demon theme, cyber effect, cyberpunk, science fiction*

翻译：机甲，复杂的机械紧身衣，机甲紧身胸衣，机械零件，机械手臂和机械腿，头饰，腐蚀性，反射，光线追踪，恶魔主题，网络效应，赛博朋克，科幻小说

对于适合使用这种科幻风的 LoRa 的大模型，二次元推荐细节较多的 Cetus Mix，真实系推荐 Experience 这种带点科幻色彩的大模型。当想强调图片某个方面的特点时，可以使用多个同类的 LoRa 叠加使用，如使用多个 Mecha 风格的 LoRa 通过 Additional Network 一并叠加使用。



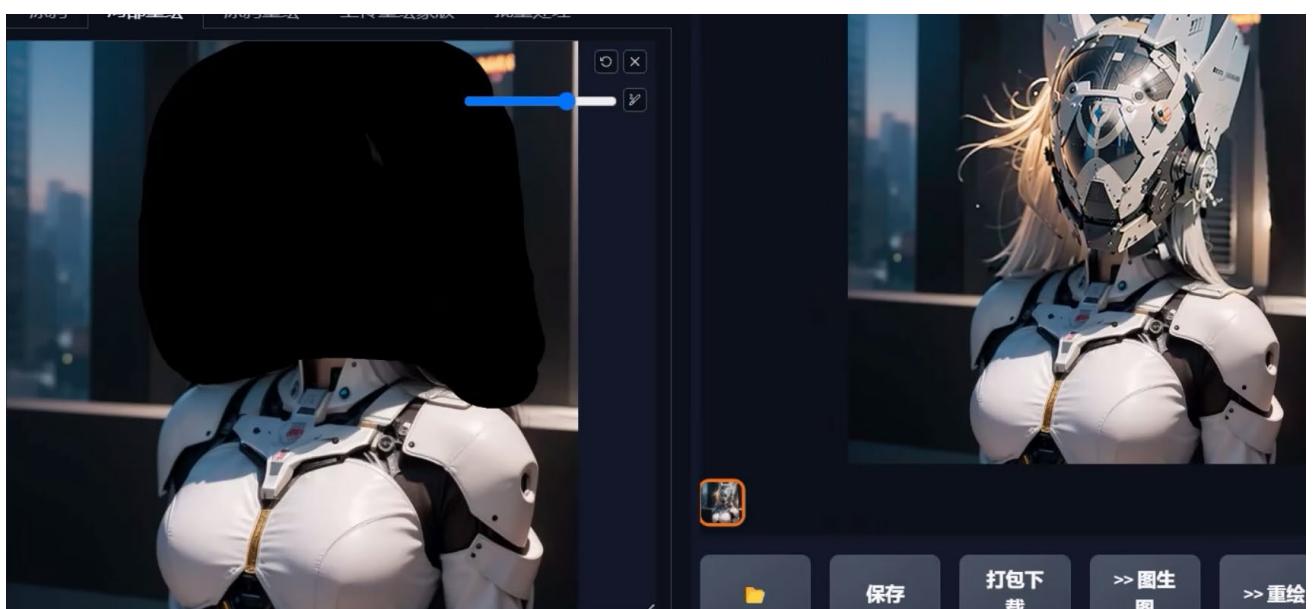
由于每个 LoRa 实现的效果略有不同，所以可以通过控制权重，来控制成片更像哪一种 LoRa。在使用服饰类 LoRa 的时候，权重一定不要太高，建议 0.3~0.4。此外还有 hanfu、holographic 等各种类型的服饰 LoRa 可供使用。

## 特定元素类

如果对同一样本的训练素材足够多，就可以实现对形象的精准传达，例如 M4 卡宾枪，美食等。甚至有些 LoRa 泛化性足够强，例如 Blue print 等这种可以用于生成产品蓝图。



还有更为进阶的用法，即使用局部重绘的方式把 LoRa 引入到图片中。我们使用 Cyberhelment 作为示范，它是生成赛博头盔的 LoRa。生成图片之后发送到图生图中，将头部选中，但注意要多选中一些，这样可以为 LoRa 增加更多的发挥空间。



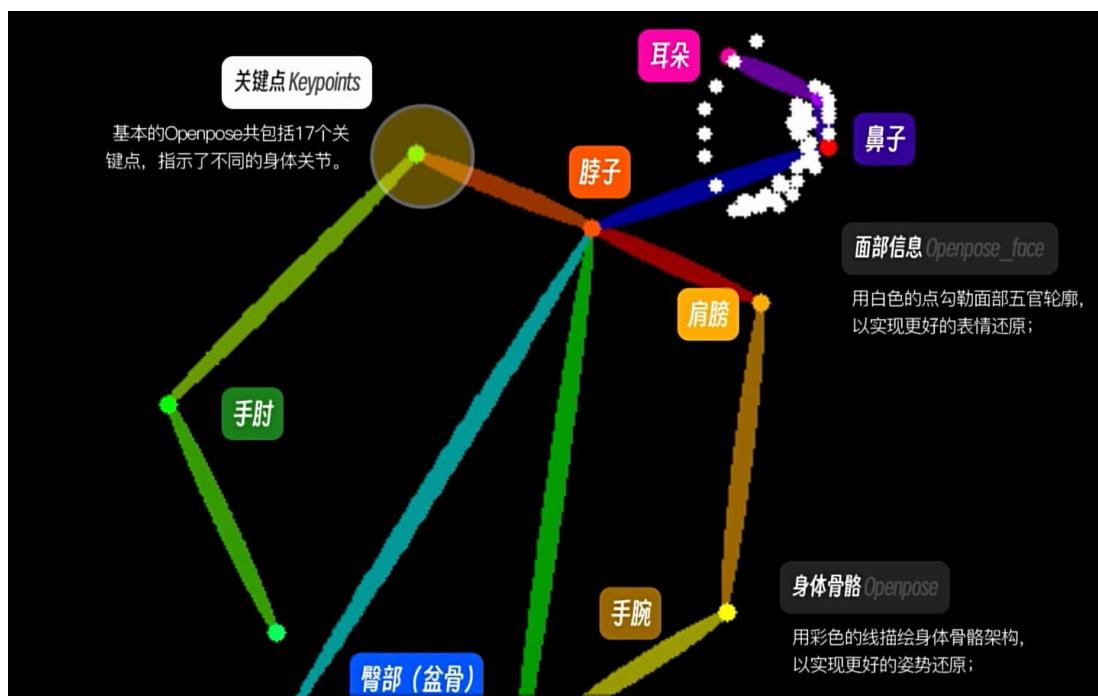
之所以使用这种方式，是因为特殊元素类的 LoRa 只是针对某个物体、部分

训练，因此如果在生成图片时就加入 LoRa 会对整个画面产生影响，局部重绘就成为更好的选择。这种方法有时也适合服饰类 LoRa。

## 十、ControlNet

ControlNet 实际上就是 WebUI 的一个插件。如果不使用 ControlNet，生成的图片充满随机性会与目标效果有一定偏差，这十分影响工作的效率，而 ControlNet 的最大作用就是控制。

ControlNet 的定位和 LoRa 类似，都是对大的图片作额外微调的插件，通过对 AI 输入一些额外的信息，来控制图片生成的方向。举个例子，如果我们通过提示词输入跳舞，画面中的人物可能有无数种姿势。



而如果使用了 ControlNet，就可以给 AI 输入一张记录了某种特定姿势信息的图片来指导 AI 作图。这张图片上不同颜色点、线，就是代表了人物的五官、四肢、关节。当然，也得让 AI 读懂类似于上图的图片，ControlNet 中搭载的许多模型就用于实现这一点。

这与图生图的形式有点类似，但是 ControlNet 的输入信息更为简洁、清晰，只是输入“姿势”这一种信息，不包含颜色、服饰等等元素，最终就不会对图片产

生额外的影响。当然，画出上图的“姿势点位图”也很麻烦，所以就有一些预处理器 Annotator从素材中智能识别提取相应的信息。

最新的 ControlNet 不仅可以控制姿势，还可以控制图片的多重信息，不仅如此，通过使用多种预处理器和多重控制网 Multi-ControlNet还能产生更强的复合控制效果。

## 安装方式

首先可以在 WebUI→拓展插件→可用选项卡中直接搜索 ControlNet 进行安装，也可以通过 GitHub 网址安装，详细操作步骤见附录 2。安装之后重启 WebUI 就可以看到 ControlNet 的选项菜单。**目前下载的是 V1.1 版本。**



但是此时还无法使用 ControlNet，因为还没有下载相应的模型。可以通过作者在 Hugging Face 的主页下载（**以下链接对应 V1.1 版本，如有更新，可以在主页查询最新模型**）：

<https://huggingface.co/llyasviel/ControlNet-v1-1/tree/main>

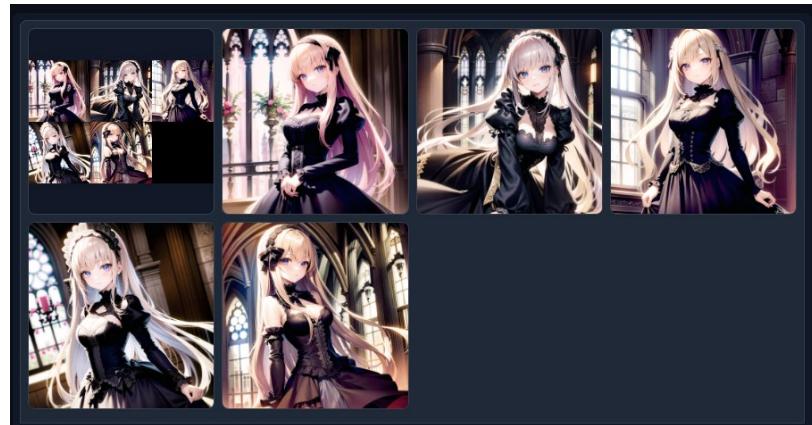
单个模型大小为 1.4GB 左右，需要“魔法”才能平稳下载。由于总共有 14 个模型，这里先推荐 5 个模型练习使用：openpose (控制姿势)、depth (深度)、canny (线稿)、softedge (柔和边缘)、scribble (涂鸦)。

**在下载时需要同时下载.pth (模型文件) 和.yaml (模型的配置文件)。**注意，需要将 .pth 和 .yaml 文件都放在 stable-diffusion-webui\extensions\sd-webui-controlnet\models 当中，文件夹当中可以有很多同名的文件，直接覆盖即可。



点击右侧的刷新按钮之后，看到有模型选项就说明已经可以使用了。

## 基本步骤



还是拿第一章中的参数举例，按照之前的步骤调高总批次数，随便生成几张图片，人物的具体姿势和位置都是随机的。

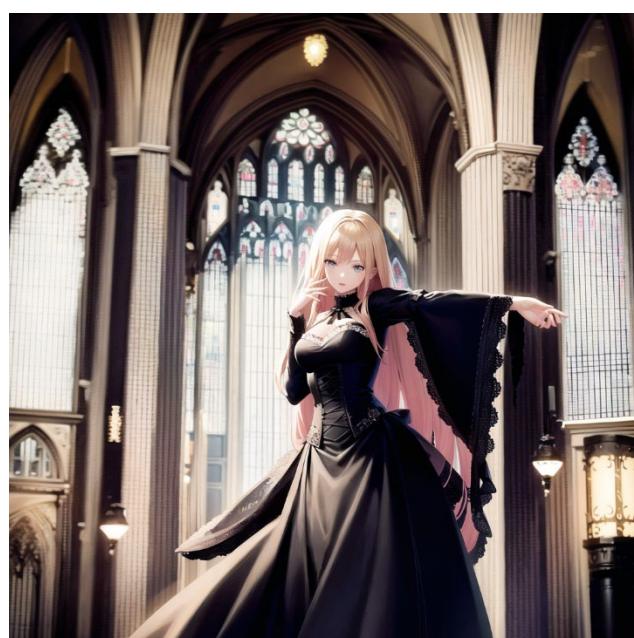


我们使用 ControlNet 对其进行控制，我们首先要启用插件。之后需要上传一张目标姿势参考图片。



之后使用预处理器提取参考图片中的姿势信息，提取姿势信息的预处理器为 openpose。共有 5 中不同类型的 openpose，有的带上手部姿势，有的带上脸部表情。我们选择没有后缀的预处理器作为举例。

之后就是选择模型，openpose 预处理器对应的就是 openpose 模型了。选择完模型之后，下方还有很多参数，之后会详细解析，先维持默认即可。点击生成，图片的姿势和参考图大部分是一模一样的（有时预处理器会识别错误，后面会讲修改方式）。



在 ControlNet 参数不变的情况下，生成多张图片人物的姿势是不会改变的，所以可以通过提示词等其他手段改变其他地方，例如套上一层 LoRa 就可以让目标角色摆出特定姿势。

在图生图中也可以使用 ControlNet，因为图生图中本身就上传了一张图片，所以如果启用了 ControlNet 并且没有上传任何图片，它就会自动通过 ControlNet 先处理一次获取姿势，再进行生成。

## 参数解析



ControlNet 可以控制图片的多种元素，所以插件内置了一些预设参数，对控制类型进行了分类，如上图所示，如果不想要对参数过多调整，或者不知道预处理器搭配的模型，就可以使用以上预设的参数。

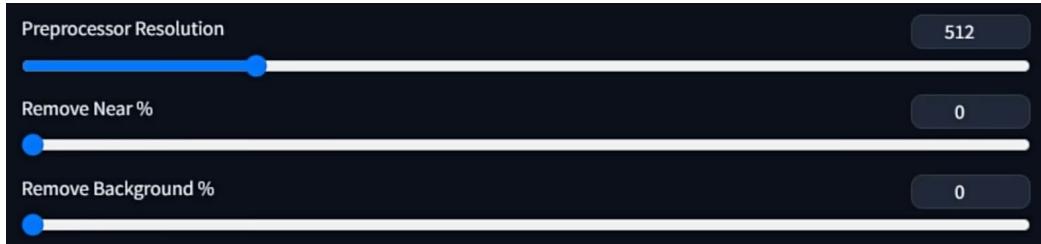


控制权重多数情况下我们维持默认 1，之后的引导介入和终止时机则是 ControlNet 的生效时间，默认的 0 和 1 代表全程生效，可以把介入时机调后一点，赋予图片更多的自由度。



控制模式可以控制图片效果和 ControlNet 的贴合程度，默认均衡即可，但是

如果觉得控制不够强，可以选择后者。**缩放模式**与之前提到的概念相同，在生成图片和信息图片像素不匹配时发挥作用。回送用于连续的图片生成，默认关闭即可。



选择好预处理器之后可能还会多出不同的控制选项，大多数情况下维持默认即可。第一个是**预处理分辨率**，它表示的是标注信息图片的分辨率，一般都会比原图小很多以节约资源。但是如果使用对精度有较高要求的模型时，如 canny、depth 等，可以将其调大或者设置成和图片一样大。

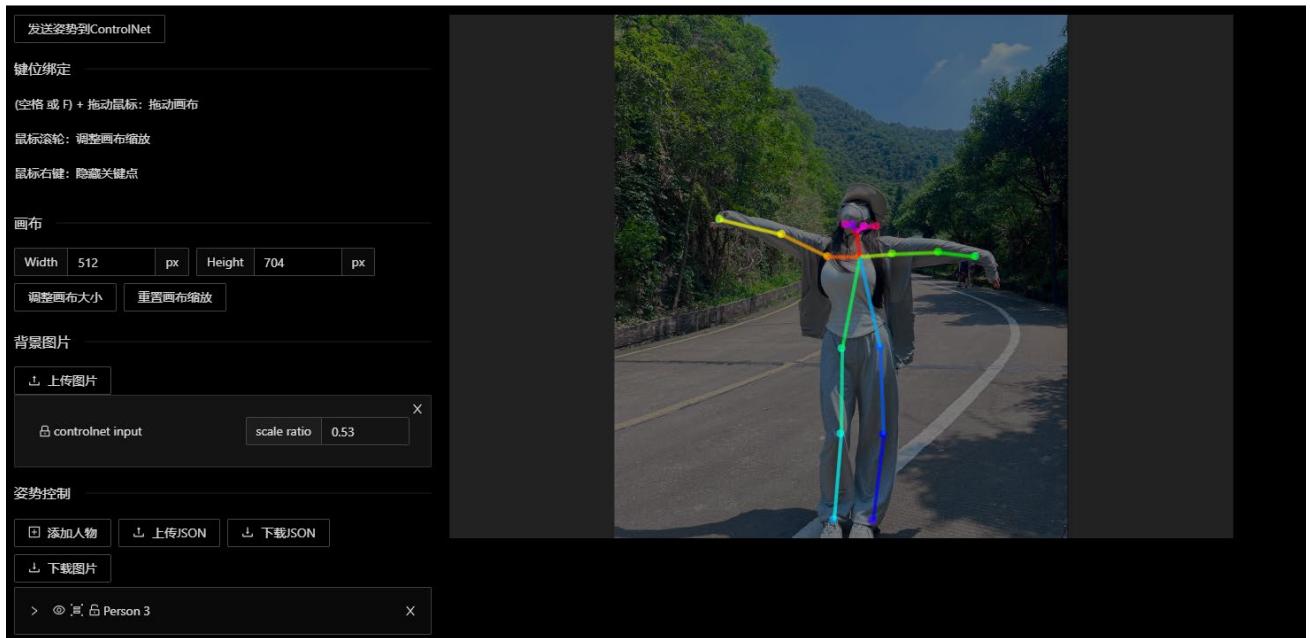


但是如果不知道如何选择可以使用**完美像素模式**，他会自动调整一个适合的像素。**低显存模式**顾名思义就是占用较少的显存，因为 ControlNet 会额外占用显存，代价就是速度会变慢。**允许预览**打开之后就可以看到 ControlNet 预处理之后“骨骼图像”，如下图。



点击预处理旁边的红色爆炸小按钮，如上右图所示，就可以生成预处理结果预览。有时结果可能不太准确，如上图所示，此时我们可以使用编辑按钮对处理

结果进行编辑。但是需要额外的 openpose editor 插件，安装方式详见附录 2。



安装完额外插件之后点击编辑就会出现如上界面，左侧有操作介绍，大多情况下参数维持默认即可，直接用鼠标点击选中关键点，然后拖动到对应位置上即可。调整完之后点击左上角的发送姿势到 ControlNet 即可。



有时我们可能想保存某一张姿势图以便下次使用，直接点击右上角的下载即可。下次使用的时候直接上传姿势图，由于已经经过预处理，注意预处理器要选择无。需要注意的是，关闭 WebUI 之后姿势图就会自动清除。

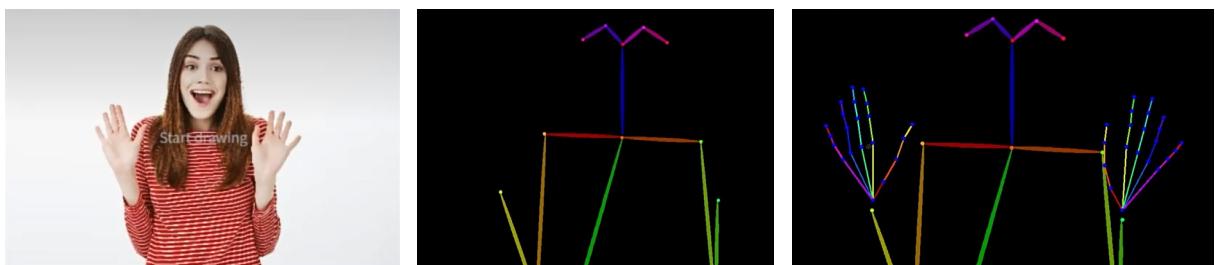
## 五大模型与应用

在 1.1 版本中，模型有 14 种，预处理器有 37 种，两者组合搭配之后数量惊人，在这里就介绍一开始提到的 5 个常用模型：openpose（控制姿势）、depth（深

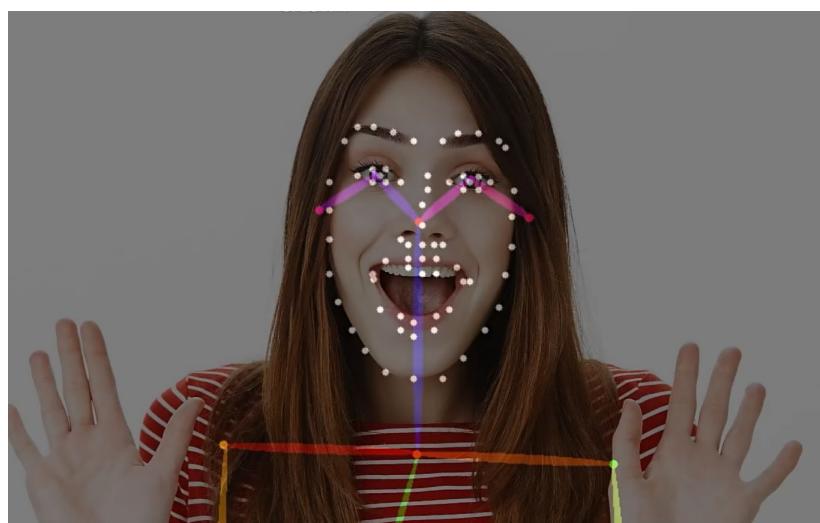
度)、canny (线稿)、softedge (柔和边缘)、scribble (涂鸦)。

### OpenPose

前面我们也介绍了其简单的操作，如果我们在预处理器中选择 openpose\_hand，就会在基础的姿势图上对手部进行精细刻画。使用这一 ControlNet 可以对手部进行精细控制，彻底解决 AI 不会画手的问题。



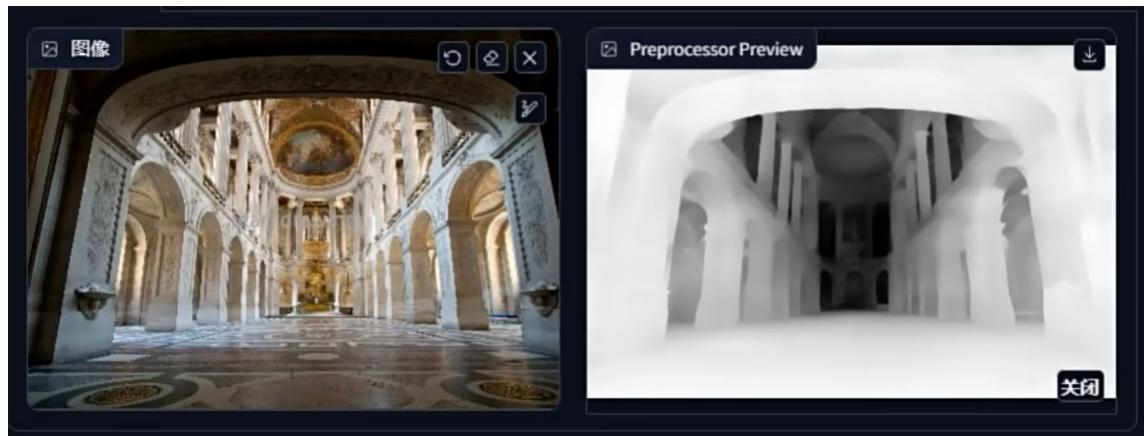
而 openpose\_face 则是用一些具体的小点，把眼睛、鼻子、嘴巴等轮廓都勾勒出来。此外还有 openpose\_faceonly 选项，它只会预处理面部信息，如果对四肢没有太高要求的话，可以使用此选项，比较适合大头照。



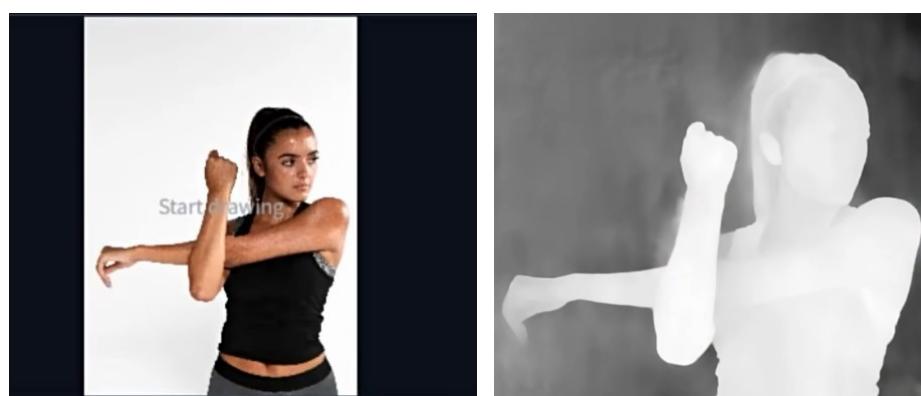
需要注意的是，如果参考图是真人，而生成图片是二次元时可能会有畸形，所以推荐参考图和生成图片是同一种类型。把所有功能都集合在一起就是 openpose\_full 了。

## Depth

它的重点是对场景的还原，由其是具有空间感的多层次背景，选择 depth\_leres++ 爆炸一下，就会识别出一张黑白图片，颜色深度就代表距离，颜色越深就代表距离越远。

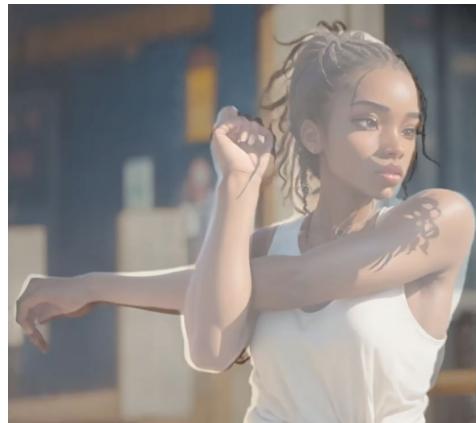


注意，depth 有四种预处理器，其中 leras++ 是 leras 的升级版，得到的效果更精细，推荐优先使用它，但对设备要求较高，如果对精细度没有太高要求可以使用另外三种，比如对人物的识别。



比如上左图肢体交叉的人像就需要用深度表示肢体关系，如果直接使用 openpose 可能就无法自动识别，需要手动调整，调整之后生成的图片也可能肢体错位，因为 AI 没有或得到深度信息，它不知道哪只胳膊在前面。而上右图就是使用 depth 生成的图像，可以很清晰的看到哪只手臂在前。在之后讲到的 Multi-

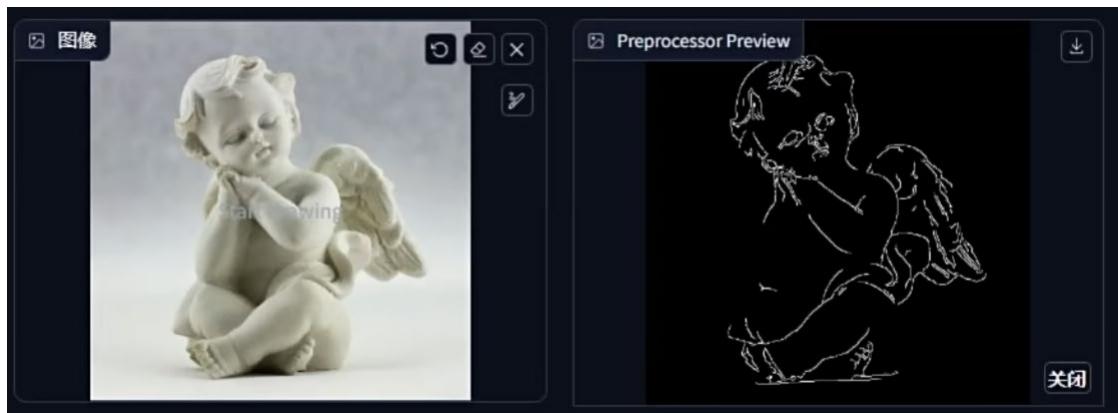
ControlNet 中可以同时用这两种预处理器，完全拿捏 AI。



需要注意的是，使用 depth 生成的图片会被限制，AI 无法自由创作人物的姿态，所以人体的形状会较为固定。如果过于受到限制，可以降低控制权重，或者控制模式中选择更偏向提示词。

### Canny

它是一种边缘检测算法，可以用于任何需要还原物体形状的场景。使用 canny 爆炸生成的预处理图是一副简单的黑白线条图，勾勒出主体的大致轮廓和细节。



如果图片中有需要精确表达的部分，如文字、标识，使用 canny 可以确保它不变形，如果有些线条不能被识别出来，可以稍微调低 threshold 的数值，将对比度不是很高的线条也识别进来，但是如果太高就会多出很多无用的细节。



有时设计师会手动画出线稿，一般都是白底黑线，但是由于在 ControlNet 中逻辑是相反的，这时预处理器可以选择 invert，就可以将颜色翻转，再用提示描述画面，就可以让 AI 自动为图片上色。



在 1.1 版本中，canny 还有一个升级版 lineart，它是专门用于动漫风格的线稿实现上色功能的预处理器和模型的组合，操作更方便，处理的线稿更为精细。

### Soft Edge

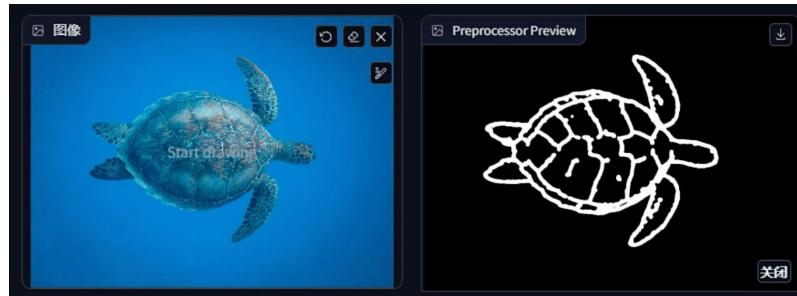


softedge 和 canny 实现的效果类似，描绘图像的边缘特征，但描绘边缘的线条比较模糊，这种模糊会让 ControlNet 的控制稍微放宽一点，让 AI 有了更大的创作空间。



上图可以很直观地看到，右边的 canny 更关注人物内部细节，而 softedge 只会把大的轮廓保留下，内部让 AI 有更大的发挥空间。softedge 也有很多种预处理器，hed 比 pidinet 生成质量相对较高，带 safe 的则是精简版。

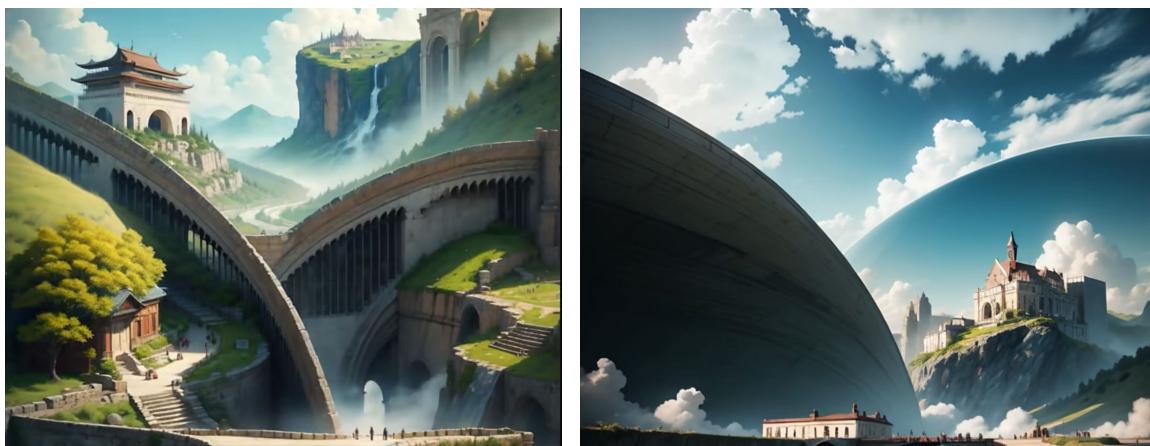
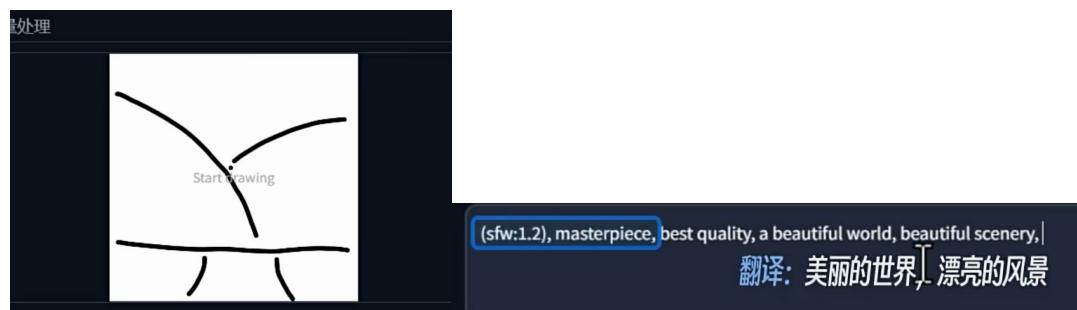
### Scribble



Scribble 是一种比 softedge 更为自由的描摹，只会为 AI 提供一个大致的轮廓，生成的具体内容更依赖于提示词，比如上图乌龟的预处理图，加入赛博朋克、科幻等提示词，就可以生成乌龟形状的宇宙飞船。



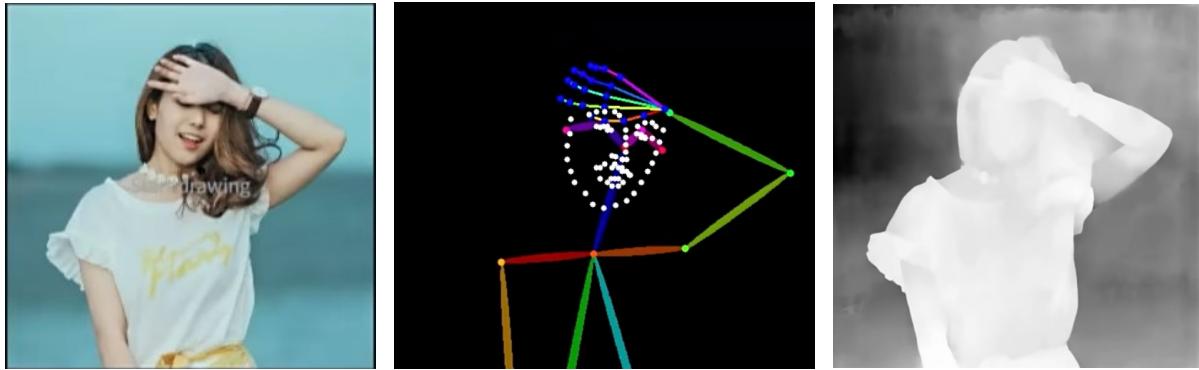
此外还可以像第五章中那样，简单描绘一个草图，注意开启 invert，之后加上少量的提示词，就可以生成多种富有 AI 的创意的图片。



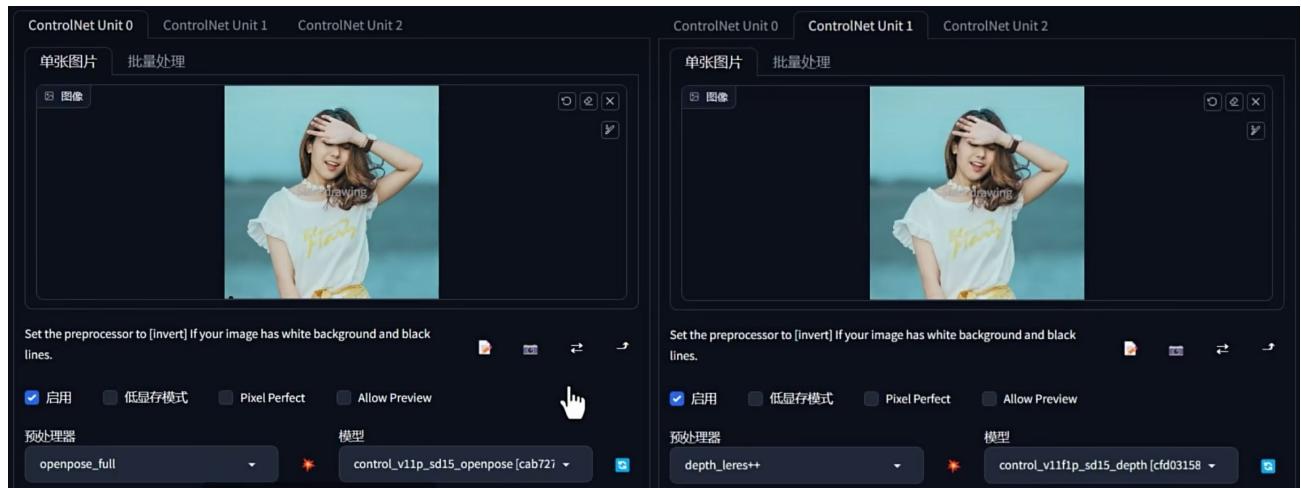
## 多重应用

ControlNet 默认只能使用一组控制，可以在设置→ControlNet→最大数量中设置最多使用的组数，重启之后界面就会如下图所示。





例如，对于处理肢体以及肢体相对位置，就可以使用 openpose+depth，openpose 可以识别精细到手指，depth 可以识别肢体的深度差异，两者组合就可以精确地控制手在放在额头前的动作。



在 unit0 中设置 openpose 参数，这里预处理器我们选择 full。在 unit1 中设置 depth 参数。**使用多重控制网最重要的就是调整权重**，对于我们的需求，最终重要的是保持姿势，所以就把 openpose 的控制权重保持 1 不变，把 depth 的权重降低到 0.6，这样不会影响识别手在头前，同时可以降低 depth 对人物边缘轮廓的控制。

需要注意的是，ControlNet 再强大也只不过是 SD 的一个插件而已，想要生成完美的图片还要综合之前的内容。比如在提示词中加入 hand on eyes 等辅助生成，就可以得到如下图片。



相比于单独使用一个控制，多种控制叠加可以获得更好的效果。除了上述的组合，还可以使用 `depth+canny`，用 `canny` 勾勒的线条补充 `depth` 生成的深度阴影中的细节；或者 `openpose+softedge`，用 `openpose` 的姿势图保持 `softedge` 大轮廓下的姿势准确性。

组合的关键是让控制网相互辅助，如 `canny` 和 `softedge` 两者叠加使用就不会有太好的效果。另外需要注意的是，多重控制网对显存的需求较高，建议开启低显存模式，或降低预处理分辨率。

### 应用实例——二维码生成

掌握了上述 ControlNet 理论之后，我们使用它来实践生成特殊二维码。使用 ControlNet 生成的二维码不同于普通二维码，它具有较高的美观性，甚至有时候看起来不像是二维码。

首先第一步就是下载 `tile.pth` 和 `brightness.safetensors` 两个模型，两个模型都要放在 `stable-diffusion-webui\extensions\sd-webui-controlnet\models` 文件夹中。  
`tile.pth` 的下载方式见：第十一章→安装方式，`brightness` 从以下网址下载（目前模型名字叫做 `control_v1p_sd15_brightness.safestentors`）：

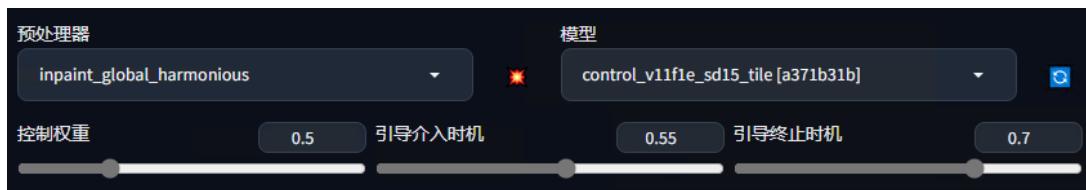
<https://huggingface.co/ioclab/ioc-controlnet/tree/main/models>



我们首先启用两个控制网络，分别都上传上我们的目标二维码图片，这里使用的是 png 格式。



之后对于 unit0 的参数，预处理器选择 inpaint\_global\_harmonious，模型选择我们之前下的 brightness。控制权重设置在 0.35 左右。引导时机先默认不变。



对于 unit1 的参数，预处理器与 unit0 相同，模型选择 tile，权重控制权重设置在 0.5 左右（**如果太低就不容易扫出结果**），引导介入时机设置在 0.55 左右，引导终止时机设置在 0.7 左右。

需要注意的是，对于选择的不同的检查点模型，以上参数会有细微的调整，根据具体生成效果判断。图片是否可以扫一扫，是否美观，也要微调参数。另外一种应用“光影文字”也是类似的操作，都是通过 brightness 固定光影部分。

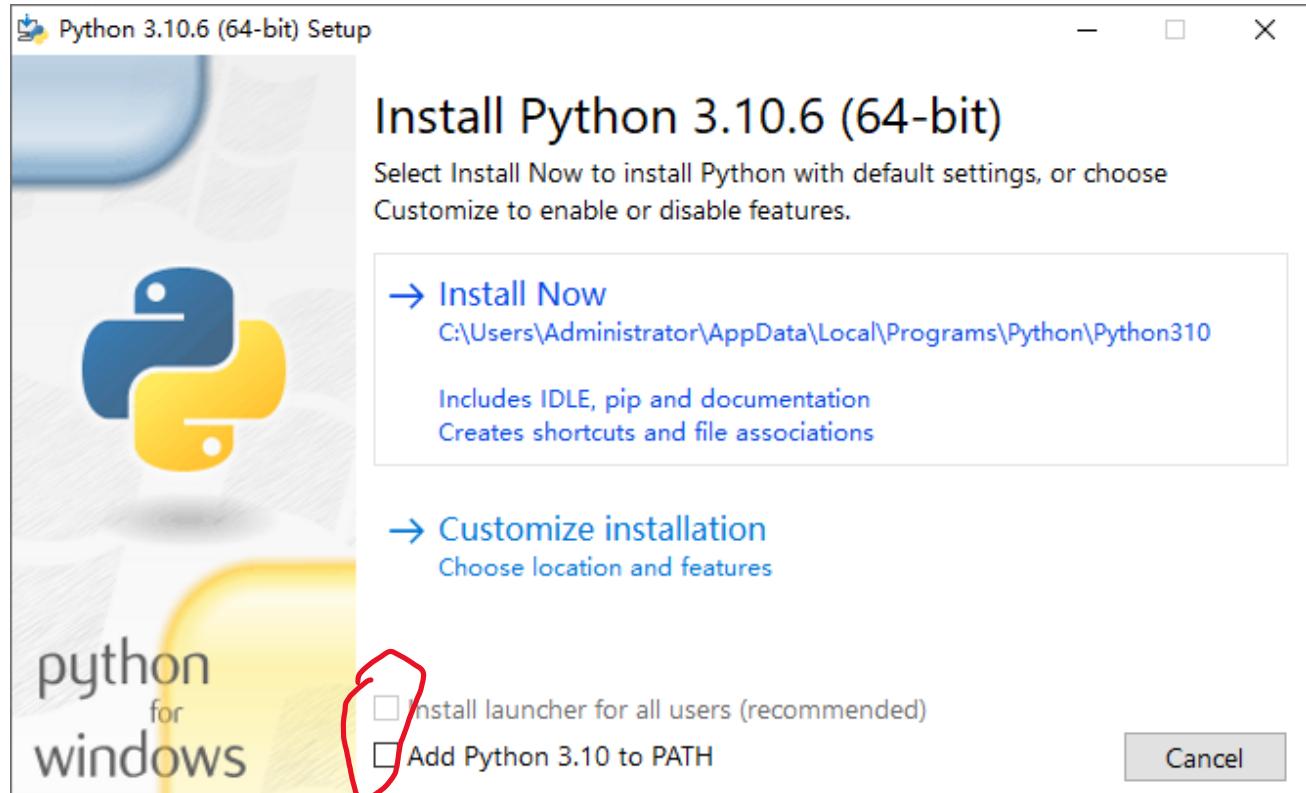
## 附 1：Stable Diffusion 安装

这里只讲解 Windows+NVIDIA 环境下，安装原生 SD（非启动器）的操作步骤。需要安装几个软件：Python、Git、CUDA（GeForce Experience）、SD 本体。其中 SD 第一次启动时会自动下载依赖文件，GeForce Experience 用于显卡驱动升级。注意，安装过程中请全程开启魔法，不然会有各种莫名其妙的错误。

第一步要先下载 python，如果精通 python 的话，可以考虑使用 anaconda 创建个虚拟环境，在这里就不介绍了。SD 是基于 3.10 版本开发的，所以下载的时候要下载 3.10 版本，推荐下载 3.10.6 版本，不要下载 3.11 版本，下载链接如下：

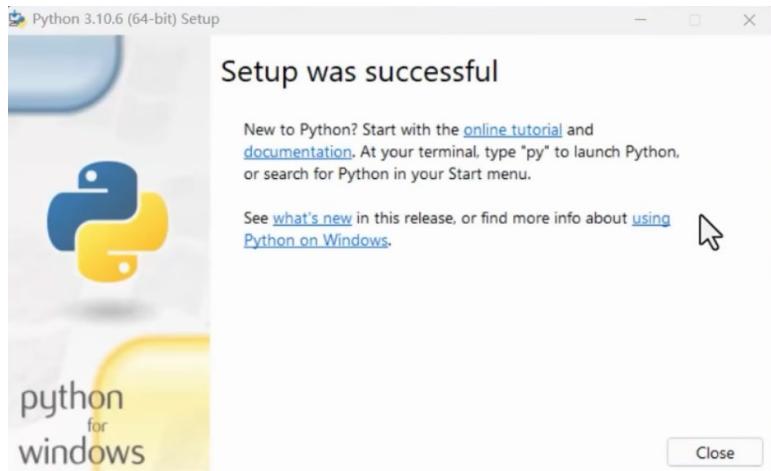
<https://www.python.org/downloads/release/python-3106/>

滚动到页面最下面，选择适合自己电脑的版本，不出意外的话应该是：Windows installer (64-bit)。下载好后直接运行。



注意要勾选下面的两个选项（安装过 python 第一个就为灰色了），如果是小

白的话直接而 install now 即可，如果想改变安装位置可以选择第二个 customize installation。



看到这个就算是安装完成了。如果想验证一下安装是否成功，可以 win+R，在弹出的界面输入 cmd。之后在 cmd 中输入： python --version，如果得到以下界面则代表安装成功（你们的版本信息应该为 3.10.6）。

A screenshot of a Windows command prompt window titled "C:\WINDOWS\system32\cmd.". The window shows the system information: "Microsoft Windows [版本 10.0.22621.1992]" and "(c) Microsoft Corporation. 保留所有权利。". In the command line, the user types "C:\Users\HYY>python --version" and presses Enter. The output is "Python 3.10.1".

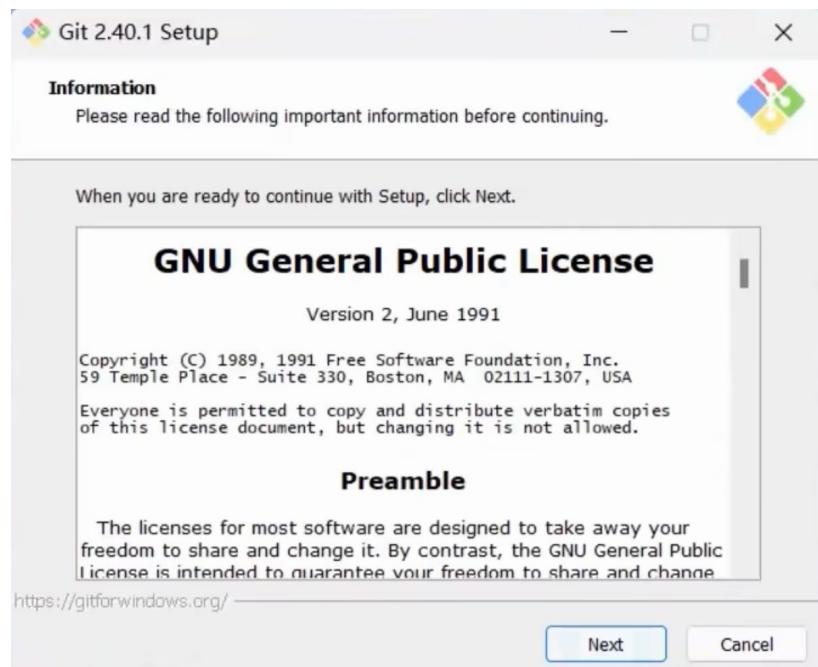
如果提示'python --version' 不是内部或外部命令，也不是可运行的程序或批处理文件。则大概率是没有将 python 添加到环境变量中，百度解决办法即可。

python 到此算是安装完成，之后我们安装 Git，Git 用于从 GitHub 上下载文件，比起自己手动下载，Git 更为安全可靠，它可以将网站的内容原封不动的下载下来。我们首先访问以下网址下载 Git：

<https://git-scm.com/download/win>

根据你的电脑选择对应的版本，不出意外的话应该是：64-bit Git for Windows

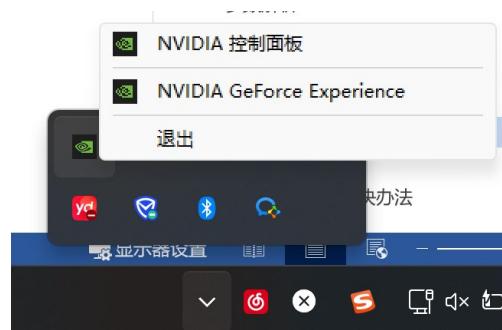
Setup。注意不要下载 portable 版本。



下载好后直接运行，一直点击 next 即可，可以自己选择安装位置，但其他的参数如果是小白的话不要改动。安装好后，还是在 cmd 中输入：git --version，如果看到版本号则代表安装成功。

A screenshot of a Windows Command Prompt window titled "管理员: C:\WINDOWS\system32\cmd.exe". The command "git --version" is entered, and the output "git version 2.41.0.windows.3" is displayed.

之后我们安装 CUDA，CUDA 是 NVIDIA 开发的让程序能够调用 GPU 的软件，所以如果是 AMD 显卡或者是 MAC 的话就要另寻他路了。



首先查看自己电脑的 CUDA 版本，电脑右下角右击 NVIDIA 图标，打开 NVIDIA 控制面板。如果没有这个图标，一种情况是你的显卡不是英伟达的，另一种情况是你的显卡驱动服务没在工作，百度解决。



打开控制面板后点击左下角的系统信息，点击组件，找到 NVCUDA64.DLL 一项，右边的即为 CUDA 版本信息。也可以在 cmd 中输入：nvidia-smi 查看。



之后我们到以下网址，找到对应的 CUDA Toolkit 版本安装，比如我的 CUDA 信息为 12.2.128，那么我就可以安装 12.2 版本系列的 Toolkit，CUDA 向下兼容，所以还可以安装更低版本的 Toolkit。

<https://developer.nvidia.com/cuda-toolkit-archive>

<b>Operating System</b>	<input checked="" type="checkbox"/> Linux	<input type="checkbox"/> Windows		
<b>Architecture</b>	<input checked="" type="checkbox"/> x86_64			
<b>Version</b>	<input checked="" type="checkbox"/> 10	<input type="checkbox"/> 11	<input type="checkbox"/> Server 2019	<input type="checkbox"/> Server 2022
<b>Installer Type</b>	<input checked="" type="checkbox"/> exe (local)	<input type="checkbox"/> exe (network)		

Download Installer for Windows 10 x86\_64

The base installer is available for download below.

<b>Base Installer</b>	<b>Download (3.0 GB) </b>
Installation Instructions:	
1. Double click cuda_12.2.0_536.25_windows.exe 2. Follow on-screen prompts	

选择了对应的 Toolkit 版本后，选择自己电脑的系统信息，注意要选择 exe (local)，点击 download 下载即可，建议不要改动安装位置。下载完后以管理员身份运行，点击 OK 和同意协议后会看到以下界面。

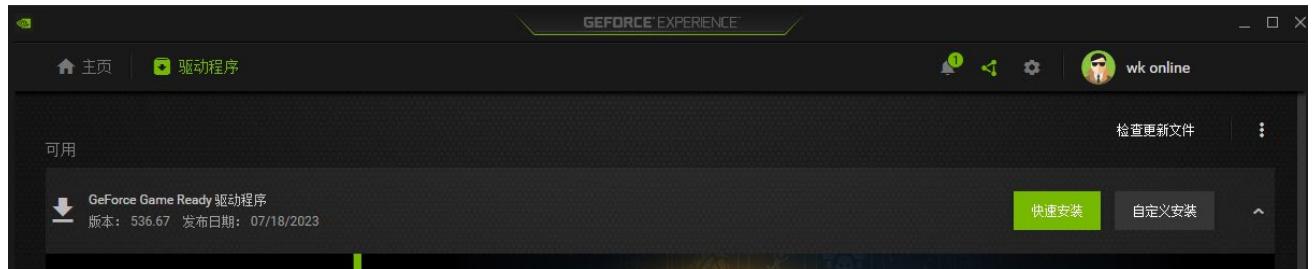


这里推荐选择精简安装不容易出错，但是不会自动 GeForce Experience。点击下一步之后就会自动安装，稍等片刻后即可安装成功。之后打开 cmd，输入：nvcc -V，输出版本信息则代表安装成功。

```
管理员: C:\WINDOWS\system32\cmd.exe
C:\Users\Administrator>nvcc -V
nvcc: NVIDIA (R) Cuda compiler driver
Copyright (c) 2005-2023 NVIDIA Corporation
Built on Fri_Jan_6_19:04:39_Pacific_Standard_Time_2023
Cuda compilation tools, release 12.0, V12.0.140
Build cuda_12.0.r12.0/compiler.32267302_0
```

如果没有安装过 GeForce Experience，需要手动到以下网址单独安装：

<https://www.nvidia.com/en-us/geforce/geforce-experience/>



安装好后登录，在驱动程序界面先点击检查更新文件，之后下载安装最新驱动即可。安装时直接点击快速安装。

安装完所有前置软件之后，建议关机重启一次，以使某些配置生效。之后打开想要安装的位置，建议硬盘剩余空间在 100GB 以上，后续 SD 大概会占用 20GB 的基础空间，添加更多的模型只会更大，目前我的 SD 已经占到了 130GB。



找到安装的目标文件夹后，右击→Open Git Bash Here，如果右键菜单中没有此选项，就在资源管理器地址栏直接输入 cmd 后回车。



打开 Git 或者 cmd 之后，输入以下命令后回车：

```
git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui.git
```

这条命令会自动创建一个名为 stable-diffusion-webui 的文件夹，并把所有项目文件拷贝到此文件夹中，提示“Resolving deltas: 100% (16893/16893), done.”则代表安装成功：

```
A85E-459B-A1CF+Administrator@A85E-459B-A1CF MINGW64 /e/ai/aiigc
$ git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui.git
Cloning into 'stable-diffusion-webui'...
remote: Enumerating objects: 24138, done.
remote: Counting objects: 100% (14/14), done.
remote: Compressing objects: 100% (10/10), done.
remote: Total 24138 (delta 5), reused 9 (delta 4), pack-reused 24124
Receiving objects: 100% (24138/24138), 31.03 MiB | 3.91 MiB/s, done.

Resolving deltas: 100% (16893/16893), done.
```

如果安装失败，大概率是网络问题，选择一个稳定的魔法并再次尝试。克隆完之后即可关闭 Git，打开 stable-diffusion-webui 文件夹，找到“webui-user.bat”文件（注意后缀不是.sh），**在双击这个文件之前，务必确保网络稳定且开启魔法，否则很多依赖无法下载。**确保网络没问题后，双击此文件运行，会弹出 cmd 运行框，之后静静等待其自动安装依赖即可。

安装过程中可能会在下载 GFPGAN 时卡住若干时间，不要着急，这个文件本身就很重。如果提示 GFPGAN 安装失败或者升级 pip 版本，参考附录 3 问题 10、问题 3。

```
Loading weights [d6548414b4] from E:\ai\aicg\AUTOMATIC1111\stable-diffusion-webui\models\Stable-diffusion\disney Pixar Car  
toon_v10.safetensors  
Creating model from config: E:\ai\aicg\AUTOMATIC1111\stable-diffusion-webui\configs\v1-inference.yaml  
LatentDiffusion: Running in eps-prediction mode  
DiffusionWrapper has 859.52 M params.  
preload_extensions_git_metadata for 19 extensions took 2.13s  
Running on local URL: http://0.0.0.0:7860
```

如果弹出了地址，则代表安装成功，复制地址到浏览器即可打开 SD。

此时最基础的配置已经完成，但是建议调整一些参数以提高使用体验。首先关闭 SD，右击“webui-user.bat”文件编辑它，在“set COMMANDLINE\_ARGS=”字段后，增加以下参数（注意每个参数之前要有空格隔开）：

```
--listen --autolaunch -xformers
```

修改后保存即可，listen 参数可以方便我们远程使用 SD，之后有类似需求可以用到。autolaunch 参数可以在每次启动完 SD 之后自动打开 Web 界面。xformers 参数可以提高显存利用率。如果在启动时提示“No module 'xformers'. Proceeding without it.”，参考附录 3 问题 11（这个问题并不会导致 SD 无法使用，但是会降低绘图效率）。

## 附 2：插件安装

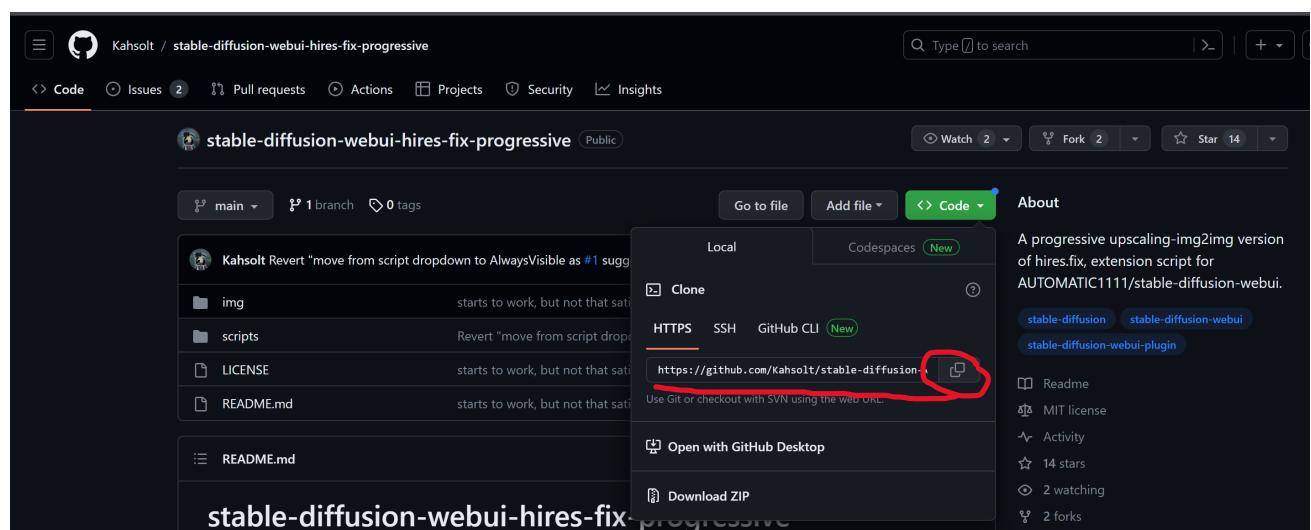
### 插件安装

插件安装有几种办法，但是以下列举的四种方式都需要 VPN 才能稳定下载，只有第五种不需要“魔法”。此外这里是以自主安装的 WebUI 为例，如果是秋叶启动器还可以在拓展管理界面通过图形化的方式安装。

第一种比较简单的方法就是通过扩展插件安装，在 WebUI 最后的选项卡中，选择从网址安装。

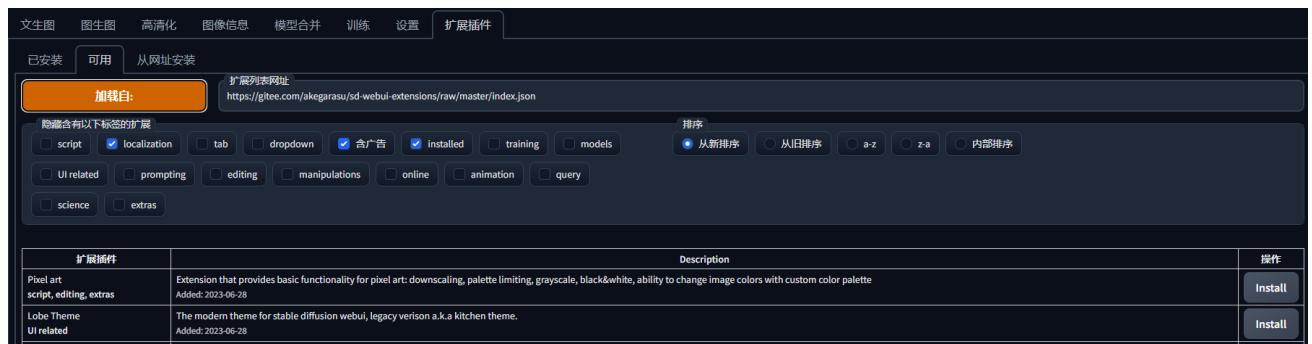


之后找到在 GitHub 上找到自己想安装的插件，如下图是一个 upscale 的插件，点击 code，点击红圈中的按钮，复制 HTTPS 的地址，粘贴到上图的“拓展插件的 git 仓库地址”中，注意不要粘贴到第二行地址栏中。之后点击安装，即可安装成功。



这种方法跟网络有很大关系，因为是从 GitHub 下载，所以需要 VPN 才能平稳下载，否则大概率是失败的。如果提示插件不安全，见附录 3 问题 12。

**第二种方法**就是点击旁边的可用，点击“加载自”，就会出现一堆可用的插件、脚本，在其中搜索直接想用的插件即可。注意不要改动“扩展列表网址”。



如果你点击加载自之后没有任何插件列出，请参照附录 3 的问题 8。

**第三种方法**就是在刚刚复制 HTTPS 地址的下方，直接下载压缩包。下载好之后把压缩包放到 stable-diffusion-webui → extensions 中，直接解压，重启 WebUI 后就有插件了。但是需要注意的是，这种方式后续不能通过 WebUI 更新，需要重新去网站下载压缩包覆盖。

**第四种方法**需要配置 git 和相关的知识。还是在第三种方法的 extensions 文件夹中，右击选择在终端中打开（或者直接选择 git bash here）。输入 git，空格后输入 clone，再空格后输入复制的 HTTPS 地址，直接回车就会自动克隆到本地。

A screenshot of a Windows PowerShell window titled 'Windows PowerShell'. The window shows the following command being run:

```
PS D:\stable-diffusion\一键安装包\novelai-webui-aki-v3\extensions> git clone https://github.com/Coyote-A/ultimate-upscale-for-automatic1111.git
```

The command 'git clone' is followed by a URL 'https://github.com/Coyote-A/ultimate-upscale-for-automatic1111.git'. The output of the command, 'Cloning into 'ultimate-upscale-for-automatic1111'...', is visible at the bottom of the terminal.

这种方法跟在 WebUI 中输入地址下载原理差不多，但是可以看到进度条，下载速度和是否安装成功。

第五种方法也是唯一一种不需要 VPN 的安装方法，就是通过各种网盘资源搜索到插件的压缩包，将压缩包放到 stable-diffusion-webui → extensions 中，直接解压，解压完成后删除压缩包，重启 WebUI 就可以了，但是这种方式也是无法通过在线的方式更新插件的，需要手动下载新版本覆盖。

## 插件更新

第一种最简单的更新办法，就是到已安装页面，点击检查更新，如果不是最新的，会有一个选项叫 behind head，此时点击“应用并重启用户界面”，即可更新插件（但是这种方法好像不适用于第三方打包安装包，如秋叶启动器）。

The screenshot shows the 'Extensions' section of the WebUI settings. At the top, there are tabs for '已安装' (Installed), '可用' (Available), and '从网址安装' (Install from URL). Below these are buttons for 'Disable all extensions' and three radio buttons for '无' (None), 'extra', and '全部' (All). A large orange button labeled '应用并重启用户界面' (Apply and restart user interface) is prominent. To its right is a 'Check for updates' button. A message at the bottom left says 'Time taken: 8.80s Torch active/reserved: 2069/2082 MiB, Sys VRAM: 3555/8192 MiB (43.4%)'. A table lists five extensions:

扩展	网址	版本信息	更新
<input checked="" type="checkbox"/> Stable-Diffusion-Webui-Civitai-Helper-main			<input checked="" type="checkbox"/> behind HEAD
<input checked="" type="checkbox"/> Tag自动补全	<a href="https://gitcode.net/ranting8323/a1111-sd-webui-tagcomplete">https://gitcode.net/ranting8323/a1111-sd-webui-tagcomplete</a>	69975587 (Sun May 21 11:54:38 2023)	<input checked="" type="checkbox"/> behind HEAD
<input checked="" type="checkbox"/> 可选附加网络(LoRA插件)	<a href="https://gitcode.net/ranting8323/sd-webui-additional-networks.git">https://gitcode.net/ranting8323/sd-webui-additional-networks.git</a>	e9f3d622 (Tue May 23 12:31:15 2023)	最新
<input checked="" type="checkbox"/> ControlNet 插件	<a href="https://gitcode.net/ranting8323/sd-webui-controlnet.git">https://gitcode.net/ranting8323/sd-webui-controlnet.git</a>	a78628d5 (Sun Apr 16 06:49:18 2023)	最新
<input checked="" type="checkbox"/> sd-webui-images-browser	<a href="https://iibulab.com/kognac/sd-webui-images-browser">https://iibulab.com/kognac/sd-webui-images-browser</a>	c615bb02 (Fri Jan 27 13:12:23 2023)	最新

第二种方法，如果不是通过在线安装的，通过压缩包或者 git 的方式安装的，这就需要去 GitHub 重新下载一次，然后到 extensions 文件夹中重新覆盖一次，重启 WebUI 后就是最新的了。



第三种方法，如果使用的第三方打包安装包，比如秋叶的启动器，在启动器界面有版本管理选项，可以直接通过图形界面更新或者删除插件。如果是非 Git 安装，同样也是不能在这里更新的，所以还是尽量能在线安装就在线安装，尽量别用压缩包。

所以推荐的安装顺序是：WebUI 在线安装→Git 安装→压缩包。

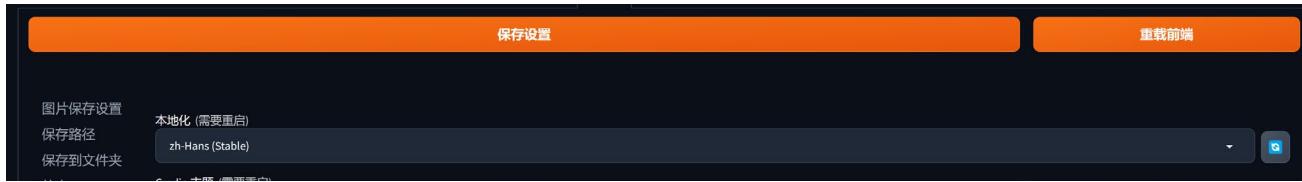
## 插件推荐

### 汉化包

扩展	描述	操作
zh_Hans Localization 本地化	Simplified Chinese localization. Added: 2023-03-13	<b>安装</b>
zh_TW Localization 本地化	Traditional Chinese localization Added: 2022-11-09	<b>安装</b>
zh_CN Localization 本地化	Simplified Chinese localization, recommend using with Bilingual Localization. Added: 2022-11-06	<b>安装</b>

首先都到此界面下载安装插件，较为方便。首先推荐的是汉化插件，先点击“加载扩展列表”，之后取消勾选 localization，搜索 zh，建议安装第一个插件

“zh\_Hans Localization”。安装完之后重启 WebUI，之后进到设置→用户界面，找到本地化选项，选择下载的汉化包（zh-Hans），之后应用设置，再重启 WebUI，界面就汉化好了。



## 图库浏览器

为了方便查找生成的图片，而不是到文件管理器中查看，同样到下载汉化包的拓展界面，搜索“image browser”，下载第二个“图库浏览器”。下载好后重启 WebUI 即可。



我们第一次进入到图库浏览器界面，不会显示任何图片，我们需要点击“首页”按钮，这时就会刷新图片出来。我们选中图片之后就会显示图片的生成信息；再点一下就会全屏预览，此时可以用键盘左右方向键切换图片。



可以更改排序顺序，也可以通过生成图片的关键词对图片进行搜索。对于满意的作品还可以收藏到收藏夹中。甚至可以对每个图片进行打分，之后就可以根据分数筛选图片，删除掉自己不喜欢的图片。

## 提示词补全

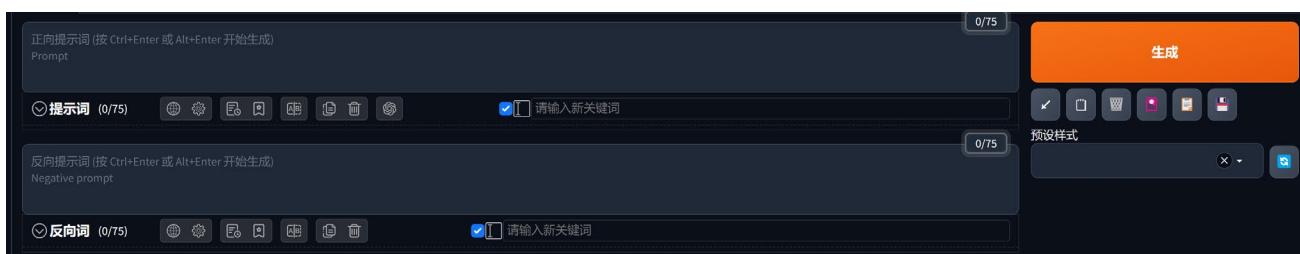
这里推荐两个提示词补全插件，一个是 all-in-one，一个是 tag autocomplete。

all-in-one 推荐使用从网址安装的方式安装，到如下界面输入此网址：

<https://github.com/Physton/sd-webui-prompt-all-in-one>



点击“安装”，显示完成之后重启 WebUI 即可，需要注意的是需要 VPN 才能顺利下载成功。



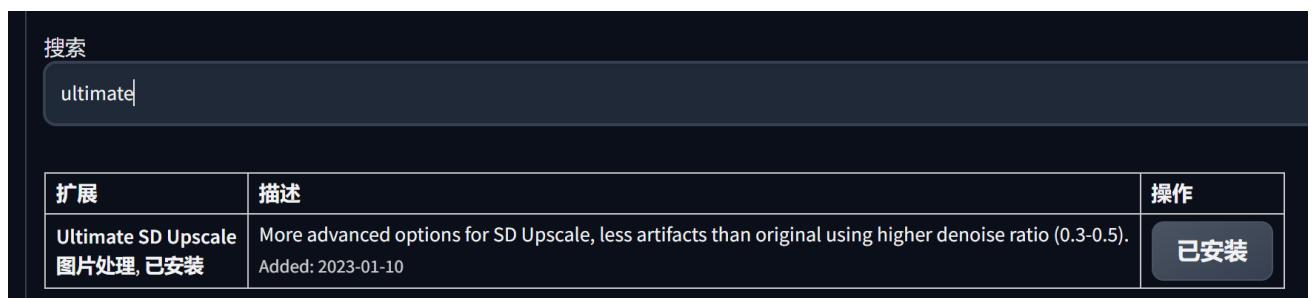
all-in-one 的界面就是上图所示，可以直接在输入框中输入中文，直接回车后稍等一会就可以翻译成英文，默认使用的是阿里翻译，可以在设置中更改。此外还可以使用 ChatGPT 翻译，但需要 ChatGPT 的 API 和 VPN 才能顺利运行。



搜索“tage auto”，第一个插件就是 tag autocomplete 了，他的作用就是把未输入完全的提示词快速补全，其次是 AI 有一套自己的语言逻辑，它就可以把词条修改成 AI 的语言逻辑表述。另外也可以快速输入 LoRa 模型和热门 IP 等。

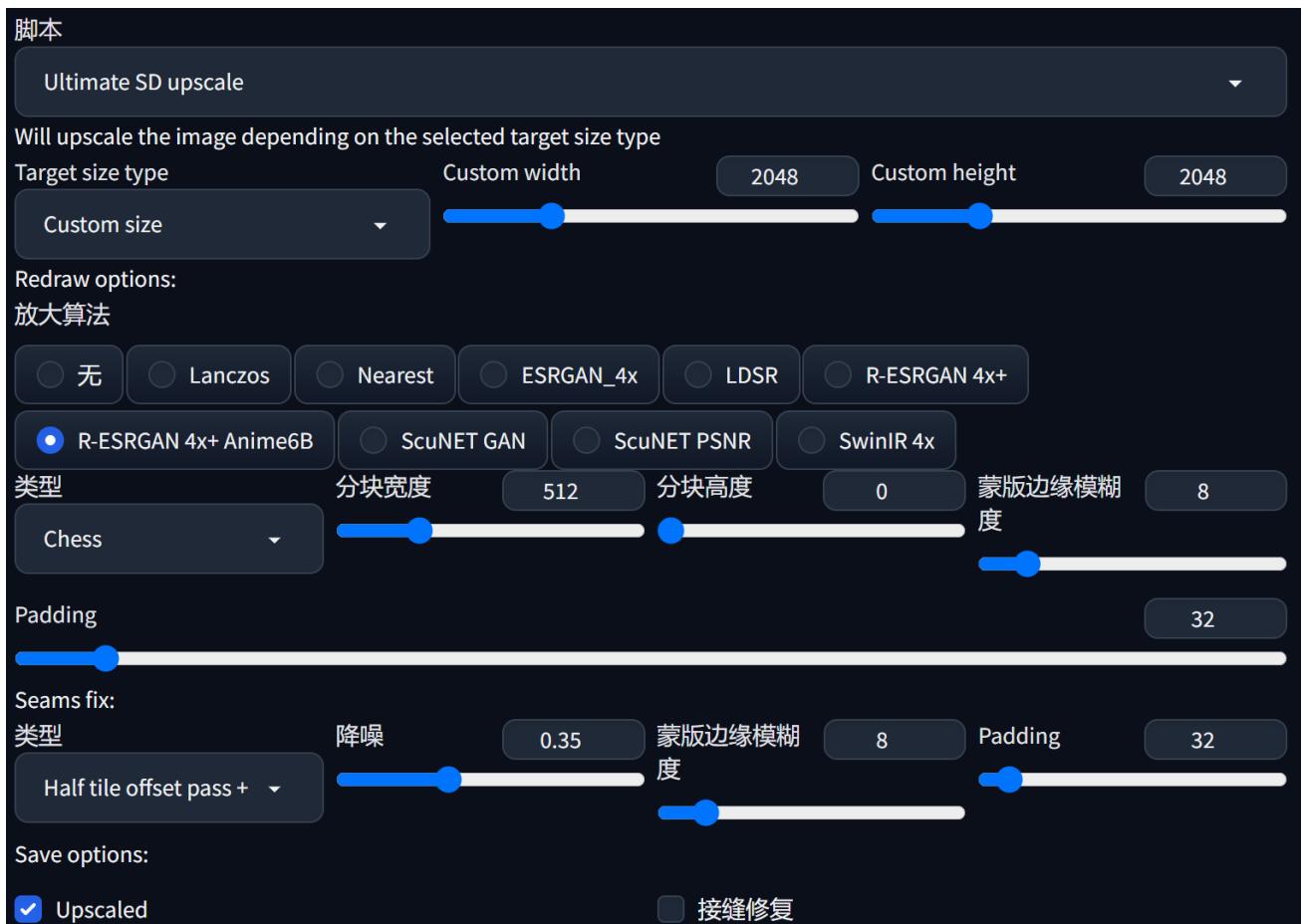
## 放大脚本

在 SD 中，除了可以安装插件以外，还可以下载脚本。在可加载界面搜索 ultimate SD upscale，会得到以下结果并安装，安装之后可以在图生图的脚本选项中看到。



我们可以认为这是之前使用的放大脚本的升级版，选择 ultimate SD upscale

之后，我们首先要设置尺寸，推荐第二种自定义尺寸。之后稍微提高重绘幅度，可以获得更多的细节。



放大算法选择图中所示的，作者指出这个算法不容易出现错误的内容。之后就是设置拆分的区块大小（之前说过图片放大脚本就是分区重绘），当分块高度和宽度有一个为 0 的时候，就是正方形（默认 512 的正方形比较稳定）。

旁边的类型有 linear（线性）和 chess（分块）两种，线性就是按顺序渲染，分块就是像国际象棋棋盘一样交叉渲染，据说分块的效果更好。padding 就是蒙版模糊度，和之前 SD Upscale 脚本部分讲到的是一个概念。

再下面的 seams fix 是接缝修复，既然是分块渲染，拼接的时候就有接缝需要处理，推荐选择第三个 Half tile offset pass + intersections。而旁边的三种选项就是

之前的内容一样的概念，就不赘述了。但有一点需要注意，这三个参数的处理对象变成了接缝附近的一小块区域。

### 局部潜空间放大

Local Latent Couple 插件不会让一张图片的尺寸变得很大，而是把一些局限于尺寸没法变得更精致的部分放大一些，在原有不改变图片结构的同时增加一些细节。



下载完成并重启之后，在文生图界面会有如上参数。再生成好了一个图片之后，就可以选择启用它。启用之后就会给整张图片增加很多细节，调整权重还有倍率可以控制细节的多少，如果想针对图片的某个区域增加细节，则需要上传蒙版，白色的区域即为增加细节的区域。四张图权重依次为：原图、0.15、0.2、0.3。

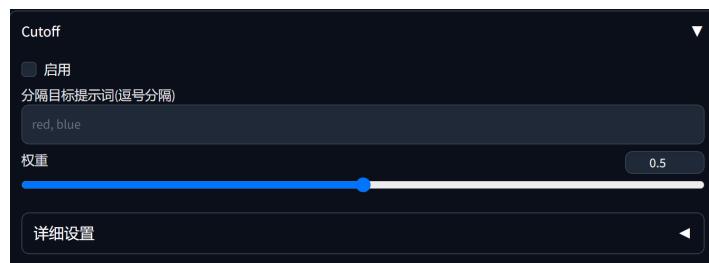


### 提示词分割控制

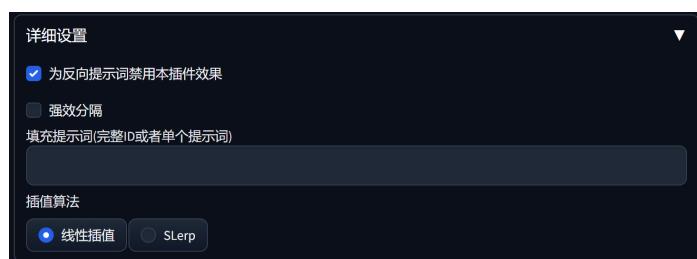
Cutoff 提示词语义分割，唯一的作用就是解决提示词之间的相互干预，比如我们生成用许多提示词生成一张图，其中一个提示词是橘色的帽子，但是在图片的其他部分还会有橘色。



安装完 Cutoff 之后，在任意一个生成界面下方就可以看到以下界面：



启用之后，将 AI 容易搞混的提示词复制进来，放到分隔目标提示词中。权重可以选择 0.3~0.5 之间的数值，代表着分隔强度。之后再生成图片就不会有混乱的现象发生。



在详细设置中，如果勾选强效分隔，并放入一个提示词，这个提示词就会拥有非常高的权重，可以把一种事物单独隔离出来处理，可以自行探索。

## 换脸

有时可能对脸部有较高要求，想换成特定人物的面部。此时就可以使用 roop 插件，相比于其他的换脸方式（提示词、图生图），这个插件直接使用目标面部图片，在文生图或者图生图的时候对面部进行修改。

安装 roop 略微有些麻烦，还需要下载 Visual Studio 构建环境。我们首先下载插件，插件地址为：

<https://github.com/s0md3v/sd-webui-roop>

插件安装方式就不再赘述，之后下载微软的 Visual Studio community，下载地址如下：

<https://visualstudio.microsoft.com/zh-hans/downloads/>

在安装的时候需要勾选 python 和 C++桌面开发两个可选安装包。

安装之后在 stable-diffusion-webui 文件夹下的 py310 文件夹中启动 cmd，之后在 cmd 中运行如下命令：

pip install insightface==0.7.3

如果没有 py310 文件夹，就在 venv/Script 中启动 cmd，需要先输入如下命令，再输入以上的命令，如果有就可以忽略。

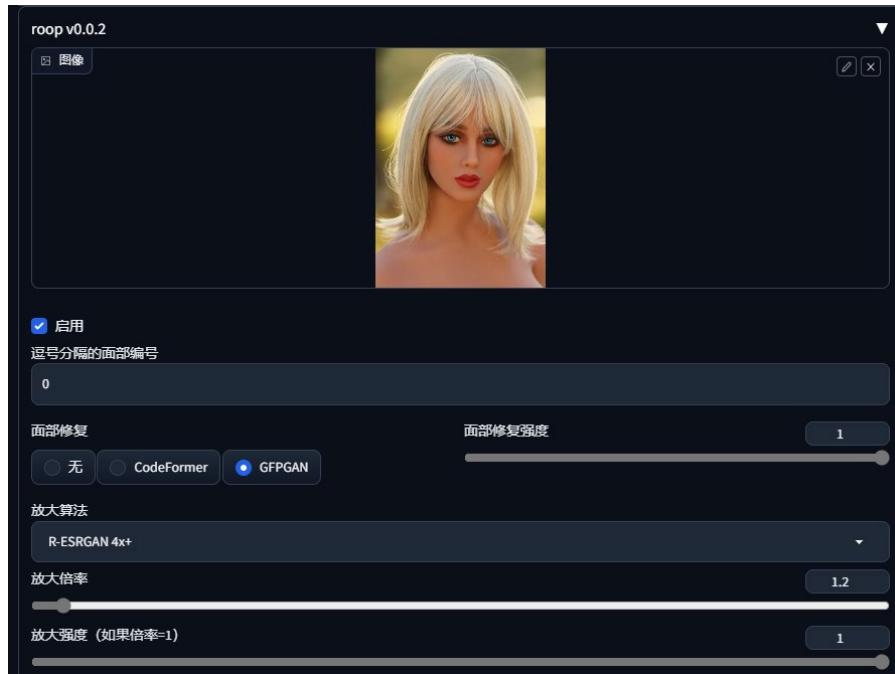
.\activate.bat

如果安装过程中有红色报错，可能是某些包安装失败，需要到 pypi.org 中搜索缺失的安装包，手动复制安装命令到 cmd 中单独安装（注意 cmd 提示符的前缀要有“(venv)”）。红色报错也有可能是 pip 版本过低，需要使用 python -m pip install --upgrade pip 升级 pip。如果还有红色报错，可以参考以下两个教程：

<https://www.youtube.com/watch?v=2gf1gdsfDU4>

[https://www.bilibili.com/video/BV1ia4y1A7KD/?share\\_source=copy\\_web&vd\\_source=d8de185b4756eeb918ddd8c6b3f8a855](https://www.bilibili.com/video/BV1ia4y1A7KD/?share_source=copy_web&vd_source=d8de185b4756eeb918ddd8c6b3f8a855)

那么如何使用 roop 呢？

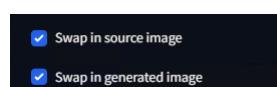


安装完成后，在文生图中会有如上界面，首先要上传一个目标人物的面部图片，干扰元素越少越好。此处我建议搭配提示词使用，先反推目标面部图片的提示词，然后加入到正向提示词中。

只有勾选了启用 roop 才会有效果，面部编号是针对多人物的，如果图片中有许多张面孔，他们的编号从左上到右下，依次为 0,1,2,3...，如果只有一个人物的话自然编号就是只有一个 0。

面部修复算法随便选择一个即可，面部修复强度代表面部与上传图片的关联程度，为 1 时代表完全贴切上传的图片。我们既然是想更换人物面部，默認為 1 即可。

放大算法是对面部细节的调整，当图片分辨率较低时没有太大区别，放大倍率可以设置在一个较低的值。



在图生图中会比文生图多出如上两个选项，分别是在原图就替换，生成之后再替换，可以两者都打开，目前实测差别不是很大。

### **Openpose Editor**

安装这款插件需要先安装 ControlNet, ControlNet 的安装方式在第十一章中。这款插件用于修改 openpose 预处理器获取到的姿势信息图，可以通过可下载安装，但要注意分辨插件名字为：sd-webui-openpose-editor。也可以通过以下网址从网址安装：

<https://github.com/huchenlei/sd-webui-openpose-editor>

安装之后概率是不可用的，通过浏览器访问以下网址检查是否安装成功：

localhost:7860/openpose\_editor\_index

如果不能访问，可能是由于地区原因无法自动下载依赖，先检查 stable-diffusion-webui\extensions\sd-webui-openpose-editor 文件夹中是否有 dist 文件夹，如果没有，从以下网址下载 dist.zip 文件：

<https://github.com/huchenlei/sd-webui-openpose-editor/releases>

下载之后将其在 stable-diffusion-webui\extensions\sd-webui-openpose-editor 文件夹中解压，并重命名文件夹为 dirt，之后重启 WebUI 再次访问上述检查网址，即可访问成功。

### 附 3：常见问题与解决办法

本节列出的只是使用 SD 常见的一些错误便于快速解决，还有很多其他错误可以自行从百度或谷歌搜索。

#### 问题 1：Can't run without a checkpoint.

安装完后运行报如下错误，说明在模型文件夹中没有任何模型，stable diffusion 启动需要至少一个模型。

```
venv "E:\ai\aicg\AUTOMATIC1111\stable-diffusion-webui\venv\Scripts\Python.exe"
Python 3.10.9 (tags/v3.10.9:1dd9be6, Dec  6 2022, 20:01:21) [MSC v.1934 64 bit (AMD64)]
Commit hash: 3715ece0adce7bf7c5e9c5ab3710b2fdc3848f39
Installing requirements for Web UI

Launching Web UI with arguments: --listen --xformers --autolaunch --enable-insecure-extension-access
sd-webui-prompt-all-in-one background API service started successfully.
No checkpoints found. When searching for checkpoints, looked at:
- file E:\ai\aicg\AUTOMATIC1111\stable-diffusion-webui\model.ckpt
- directory E:\ai\aicg\AUTOMATIC1111\stable-diffusion-webui\models\Stable-diffusion
Can't run without a checkpoint. Find and place a .ckpt or .safetensors file into any of those locations. The program will exit.
请按任意键继续. . . -
```

#### 问题 2：Failed to create model quickly; will retry using slow method.

*loading stable diffusion model: ConnectionError*

*10054, '远程主机强迫关闭了一个现有的连接。'; None, 10054, None*

本来能够运行，但是在做了某些操作之后就无法运行了。其中一种可能的原因是，下载的模型中有损坏、模型放错位置、模型类型不对等等，检查最近新增加的模型，尝试先删除他们再启动 SD，如果能重新启动则说明是模型问题。

#### 问题 3：第一次安装后无法启动

在 stable-diffusion-webui 文件夹中找到 webui-user.bat，右键点击编辑。在编辑界面中找到 set COMMANDLINE\_ARGS=

```
set PYTHON=
set GIT=
set VENV_DIR=
set COMMANDLINE_ARGS=
|
call webui.bat
```

在其后粘贴（确保其后面没有任何文字）“`--medvram --autolaunch --deepdanbooru --xformers --no-gradio-queue`”，再重启 `webui-user.bat` 即可。如果还是不行就清空 `set COMMANDLINE_ARGS=` 后的文字，再粘贴“`--listen --xformers --autolaunch --enable-insecure-extension-access`”进行尝试。

#### *问题 4: 'NoneType' object is not subscriptable*

这个报错一般是在 `stable-diffusion` 软件成功安装并运行后出现的。原因是显卡的显存太小导致的。将画面调小一点，例如调成 `512*512` 或者 `512*768` 对于 `4GB` 显存来说一般都没问题。

另外，如果勾选了面部修复（`Restore faces`），也很有可能由于显存不足而报错。不要勾选 `Restore faces` 就一般不会报错了。也有可能是因为在下载的时候没有下载完整面部修复的数据，重新下载覆盖三个文件资源即可。具体的下载链接和操作百度搜索：【`Wiki 百科`】面部修复报错怎么办？

#### *问题 5: 界面中没有 VAE 选项*

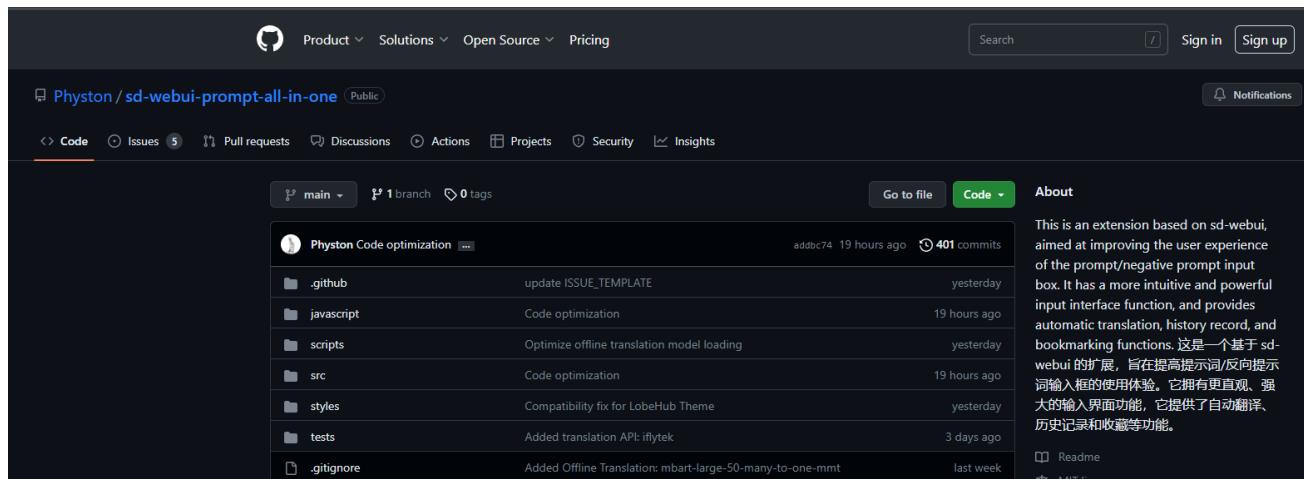
到设置→用户界面→快捷设置列表中，添加字段：`sd_model_checkpoint, sd_vae`，然后点击页面最上方的应用设置，之后点击重新加载 WebUI 即可。其他类似的问题也可以通过相似的方法解决。

#### *问题 6: ChatGPT 辅助补充不可用*

`ChatGPT` 辅助提示词填充需要正确的 API，需要现有 `ChatGPT` 账号，在账号管理中生成 API 接口，填入之后即可使用，并且要使用 VPN 才可以访问 `ChatGPT`。使用网页能访问到 `ChatGPT` 时，才能用 SD 访问到。

#### *问题 7: 插件无法更新或正常使用*

通过手动解压没通过网址下载的插件无法在 WebUI 中检测更新，要到 stable-diffusion-webui → extensions 中，删除无法更新或无法正常使用的插件，然后再 GitHub 上找到对应插件的下载地址，到如下界面。

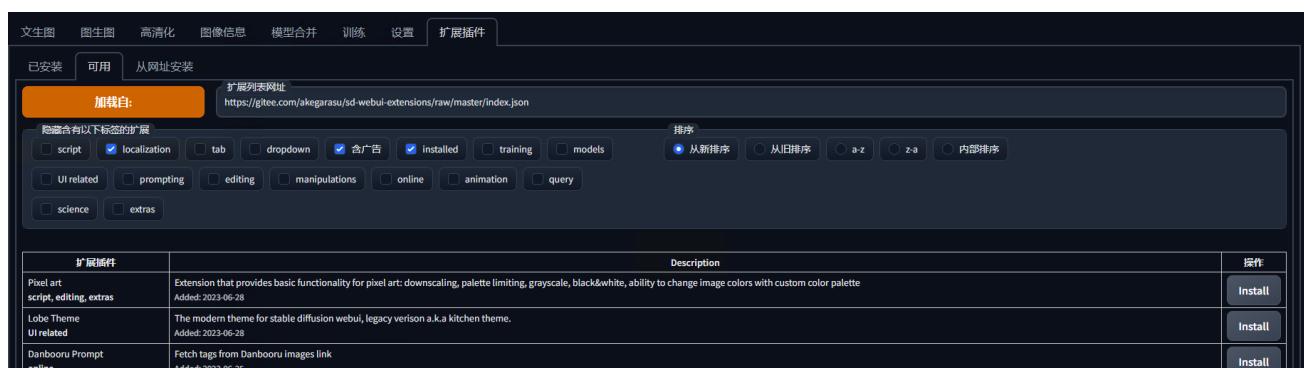


复制网页的地址，粘贴到 SD 中如下的框中，直接回车即可下载。



如果不能下载，请参照附录 2 中的插件更新内容，逐一尝试。

### 问题 8：插件加载报错



如果点击加载自之后没有任何插件列出，就需要修改下载镜像的网址。首先进到 stable-diffusion-webui→modules 文件夹中，找到 ui\_extensions.py 文件，使用

编译器或者记事本打开，然后 `ctrl+F` 搜索 `https`，注意不要修改任何东西，搜索结果如下（最新版的代码略有改动，可以找会 `python` 的朋友找到对应位置）。

```
with gr.TabItem("Available"):
    with gr.Row():
        refresh_available_extensions_button = gr.Button(value="Load from:", variant="primary")
        available_extensions_index = gr.Text(value="https://gitee.com/akegarasu/sd-webui-extensions/raw/master/index.json")
        extension_to_install = gr.Text(elem_id="extension_to_install", visible=False)
        install_extension_button = gr.Button(elem_id="install_extension_button", visible=False)
```

搜索之后会定位到 `available_extensions_index` 参数（一般情况下是搜索的第二个结果），核实和图片中参数一致之后，修改 `value` 的值，即替换原来的网址，可以用以下两个的任意一个镜像源：

<https://gitee.com/akegarasu/sd-webui-extensions/raw/master/index.json>

<https://gitgud.io/AUTOMATIC1111/stable-diffusion-webui/-/wikis/Extensions-index.md>

替换完之后保存，重启 WebUI，再次加载即有插件了。

如果在下载插件的时候报错：`extension access disabled because of command line flags`，或者报错：`JSONDecodeError: Expecting value: line 1 column 1 (char 0)`，就需要用记事本打开主目录下的“`webui-user.bat`”文件，在“`set COMMANDLINE_ARGS=`”后增加参数：`--enable-inscure-extension-access`，以避免一些不必要的检查。如果还有一些问题，可以参照问题 3。

### 问题 9：反推不可用

首先打开 `stable diffusion` 的主目录，找到 `interrogate_tmp` 文件夹（有时也叫 `interrogate`），把这个文件夹删掉，然后下载网盘中的压缩包，解压到 SD 主目录中。

还有两个模型要删除，`\models\BLIP\model_base_caption_capfilt_large.pth` 这

个是 CLIP 反推模型，删除后下载网盘中的对应文件放回来，如果是 854MB 可以不删除。第二个是\models\torch\_deepdanbooru\model-resnet\_custom\_v3.pt 这个是 DEEPbooru 反推模型，删除后下载网盘中的对应文件，放回来如果是 614MB 可以不删除。

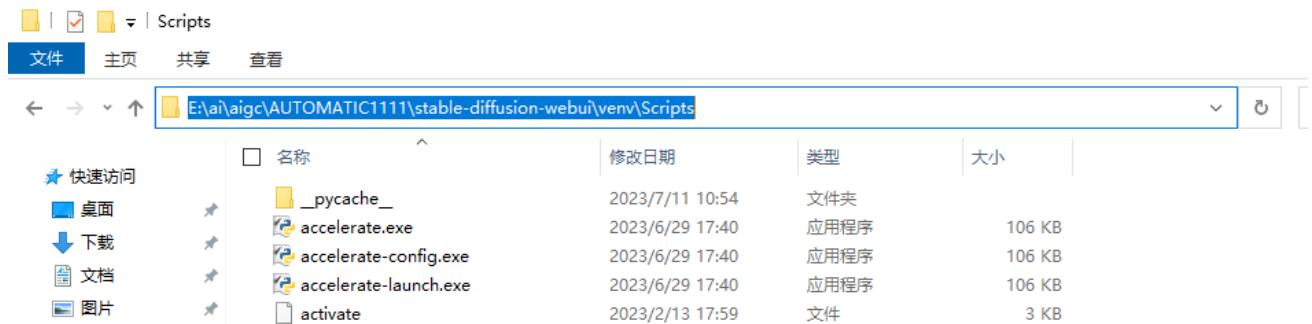
链接：<https://pan.quark.cn/s/863e95de0907> 提取码：5fgU

出处：

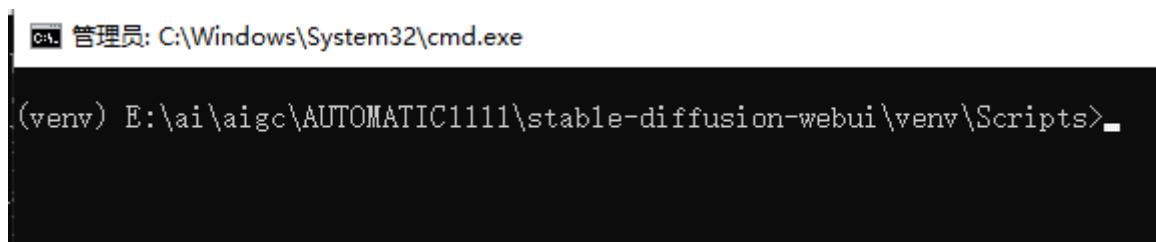
[https://www.bilibili.com/read/cv23041120?from=search&spm\\_id\\_from=333.337.0.0](https://www.bilibili.com/read/cv23041120?from=search&spm_id_from=333.337.0.0)

#### 问题 10：GFPGAN 安装失败、提示升级 pip

如果第一次启动时卡在“installing gfpgan”或者直接提示安装失败，我们首先要升级 pip 版本。关闭报错的 SD 界面，打开 stable-diffusion-webui→venv→Scripts 文件夹，在资源管理器地址栏直接输入 cmd 回车。



打开 cmd 之后输入：.\activate.bat 后回车，注意，此时提示符前为“(venv)”。



之后输入以下命令升级 pip：

```
python.exe -m pip install --upgrade pip
```

提示升级成功后再次打开“webui-user.bat”文件，它会继续安装 GFPGAN 等依

赖。如果还是提示安装不成功，则是网络的问题，需要到 python 文件中更换下载源，打开 stable-diffusion-webui→modules 文件夹中的 launch\_utils.py 文件（用记事本或者 VScode 打开都可以），ctrl+f 搜索 gfpgan\_package，定位到如下代码：

```
if not is_installed("gfpgan"):
    run_pip(f"install {gfpgan_package}", "gfpgan")
```

修改 run\_pip 一行为如下代码（就是在 install 后增加字段，以使 SD 走国内镜像源下载）。保存后再次运行“webui-user.bat”文件。

```
run_pip(f"install -i https://pypi.douban.com/simple/ {gfpgan_package}", "gfpgan")
```

此时 GFPGAN 应该可以平稳下载，之后的其他依赖出现问题也是同样的解决办法，例如 clip 下载失败，就搜索 clip\_package，定位到 run\_pip 代码行，然后在 install 之后增加以下字段后保存即可：

```
-i https://pypi.douban.com/simple/
```

如果这种方法还是卡住不动的话，就在 launch\_utils.py 文件里找到 gfpgan\_package = os.environ.get() 字段（以 GFPGAN 举例，其他的依赖方法类似），在对应的 https://github.com/前面加上 https://github.moeyy.xyz/，修改后保存再次尝试即可。

#### *问题 11: No module 'xformers'. Proceeding without it.*

首先我们要检查“webui-user.bat”文件中的“set COMMANDLINE\_ARGS=”字段后是否有--xformers 参数，如果没有添加后重新尝试。

如果还是有报错提示，则到 stable-diffusion-webui→venv→Scripts 文件夹，在资源管理器地址栏直接输入 cmd 回车。

	名称	修改日期	类型	大小
	__pycache__	2023/7/11 10:54	文件夹	
	accelerate.exe	2023/6/29 17:40	应用程序	106 KB
	accelerate-config.exe	2023/6/29 17:40	应用程序	106 KB
	accelerate-launch.exe	2023/6/29 17:40	应用程序	106 KB
	activate	2023/2/13 17:59	文件	3 KB

打开 cmd 之后输入：.\activate.bat 后回车，注意，此时提示符前为“(venv)”。

之后在 cmd 中输入以下命令安装 xformers 模型：

```
pip install xformers
```

提示安装完成后关闭 cmd，重新打开 webui-user.bat 即可。需要注意的是，一定要有--xformers 字段。

### 问题 12：从网址安装插件提示不安全

右击编辑“webui-user.bat”文件，在“set COMMANDLINE\_ARGS=”字段后增加以下参数（注意每个参数之间要有空格隔开）：

```
--enable-insecure-extension-access
```