

AI与视觉表达部分

jee & 散步

“有个性”的图像生成

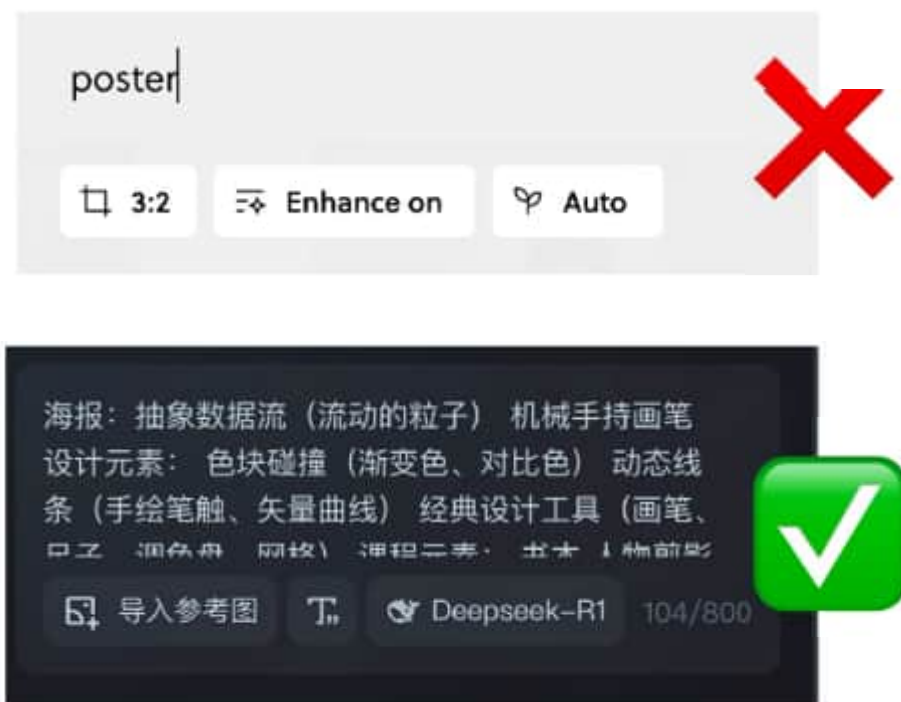
散步 sanbuphy

Datawhale

physicoadam@gmail.com

什么叫“有个性”

让内容言之有物



控制生成内容



如何用 AI 制作海报-元素生成

我现在要设计一个有关 AI+设计通识课 的海报，帮我设计一下这个海报的元素（要体现出设计、AI，并且你要记住这是一个课程），给出画面元素和里面写的字的一些参考



已深度思考 (用时 31 秒) ^

*** 可自由决定是否启用深度思考 (提示词长建议不开启)**

嗯，用户需要设计一个关于AI+设计通识课的海报，我得先理清清楚他们的需求。首先，他们想要体现出“设计”和“AI”这两个核心元素，同时还要明确这是一个课程。那我得想想怎么把这三个要素结合起来，既要有科技感，又不能太复杂，毕竟通识课可能面向的是不同背景的学生。

首先，海报的视觉元素。设计方面，可能需要一些设计工具或元素，比如几何图形、线条、网格或者画笔之类的。AI的话，可能需要用一些科技感的元素，比如数据流、电路板、像素点或者机器人图标。不过要避免过于常见的元素，比如机器人可能太直白，可以考虑用抽象一点的，比如神经网络或者流动的数据点。

然后是课程相关的元素，可能需要书本、学位帽、黑板或者学习路径的图标，但如何把这些和AI、设计结合起来呢？比如书本旁边有数据流环绕，或者黑板上有代码和设计草图。颜色方面，科技感通常会用蓝色、紫色这种冷色调，但设计可能需要一些活泼的颜色，比如橙色、黄色，来增加活力。所以颜色搭配可能需要兼顾两者，比如主色调用蓝紫渐变，点缀亮色。

如何用 AI 制作海报-元素生成

一、画面核心视觉元素

1. AI与设计的融合符号：

- 抽象大脑：用几何线条或像素点组成的大脑轮廓，内部填充设计元素（如色块、画笔笔触、网格线）。
- 数据流环绕工具：AI芯片、电路纹理与设计工具（如画笔、尺规、PS图标）交织成环形，体现跨界结合。
- 动态粒子：背景中漂浮的发光粒子，象征AI算法的流动性与设计灵感迸发。

2. 课程属性强化：

- 书本与代码结合：一本打开的书籍，左页显示设计草图，右页显示AI代码（如Python片段）。
- 学习路径箭头：从“设计基础”到“AI工具”的阶梯式路径，用渐变线条连接。

3. 科技感背景：

- 低多边形（Low Poly）风格的渐变蓝紫色背景，叠加透明网格线或微弱的二进制代码纹理。

二、文字内容与排版

主标题

- 文案示例：



如何用 AI 制作海报-元素生成

我现在要设计一个有关 AI+设计通识课 的海报，给我视觉符号的6组例子，每组例子用逗号分割每个符号，每组需要不一样，并且不少于 6 个符号



1. 大脑图标，电路板图案，调色板，画笔，齿轮，数据流箭头
2. 机器人头像，抽象几何图形，像素点，渐变色彩，网格线，无限符号
3. 眼睛图标，神经元连接图，3D立方体，光影效果，数字代码，云朵
4. 手掌轮廓，AI芯片，动态波纹，抽象线条，星云图案，时间轴
5. 地球图标，虚拟现实眼镜，算法公式，拼图块，光点，螺旋结构
6. 书本图案，AI语音波形，抽象人脸，折线图，光晕效果，未来城市剪影



如何用 AI 制作海报-元素生成

我现在要设计一个有关 AI+设计通识课 的海报，给我视觉符号的6组例子，每组例子用逗号分割每个符号，每组需要不一样，并且不少于 6 个符号；并且需要划分出主体物、辅助元素，想要表达古典和现代结合、艺术和计算机结合的思想



1. 古典与现代结合

- 主体物：古典雕塑（如大卫像）、AI芯片、毛笔、数字像素、古典建筑柱式、现代玻璃幕墙
- 辅助元素：古典花纹、电路板纹理、水墨晕染、几何网格、古典书籍、现代平板电脑

2. 艺术与计算机结合

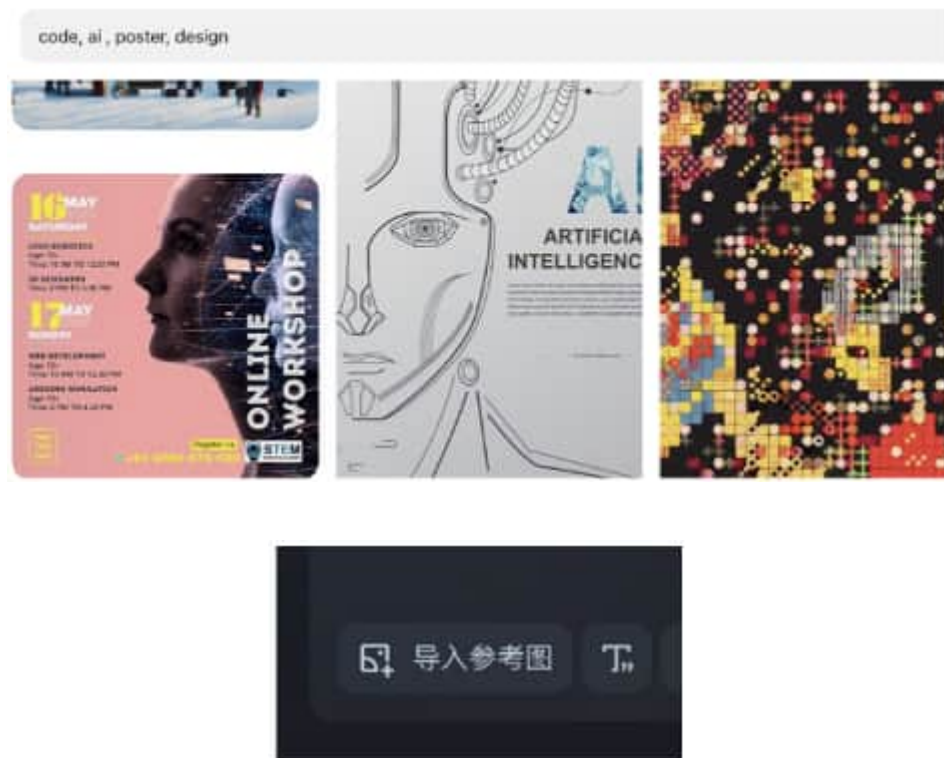
- 主体物：调色板、代码编辑器、画笔、二进制代码、画布、机器人手臂
- 辅助元素：色彩渐变、数据流、艺术画作、算法流程图、颜料滴落、键盘按键

3. 传统与科技融合

- 主体物：古典钟表、智能手表、传统算盘、量子计算机、古书、电子书
- 辅助元素：齿轮、数字时钟、算珠、量子比特、书页翻动、电子屏幕

如何用 AI 制作海报-寻找参考

获得符号灵感后，我们可选择 1、直接生成海报 2、寻找类似元素的海报作为底图

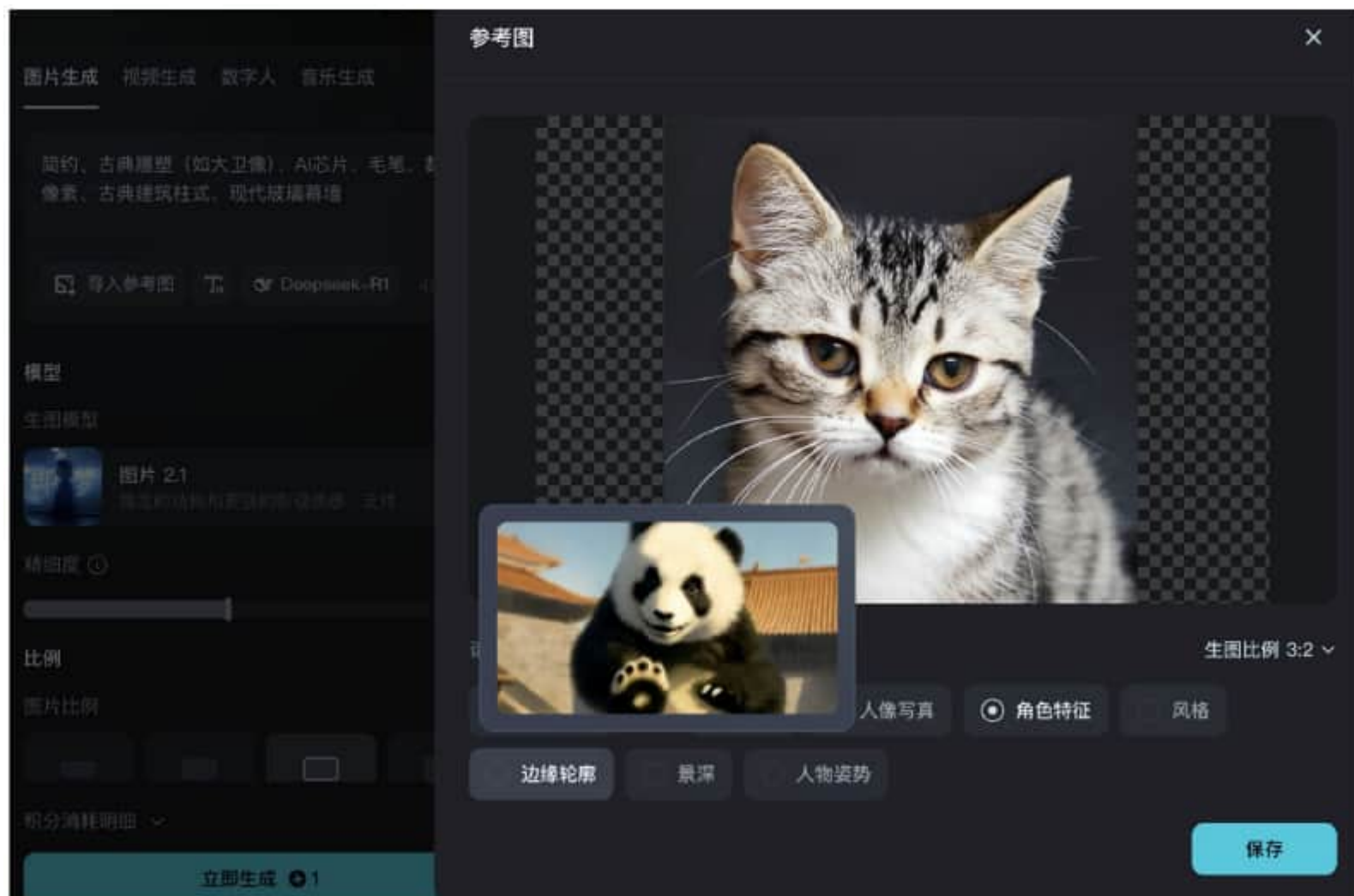


<https://jimeng.jianying.com/ai-tool/home>

如何用 AI 制作海报-直接生成



如何用 AI 制作海报-参考图



如何用 AI 制作海报-作为基底

*不写“海报”等字眼可用于生成底图，除非需要参考板式



如何用 AI 制作海报-排版参考



即梦-图片生成 03-23 19:46

古典雕塑（如大卫像）、AI芯片、毛笔、数字像素、古典建筑柱式、现代玻璃幕墙、课程海报

图片 2.1

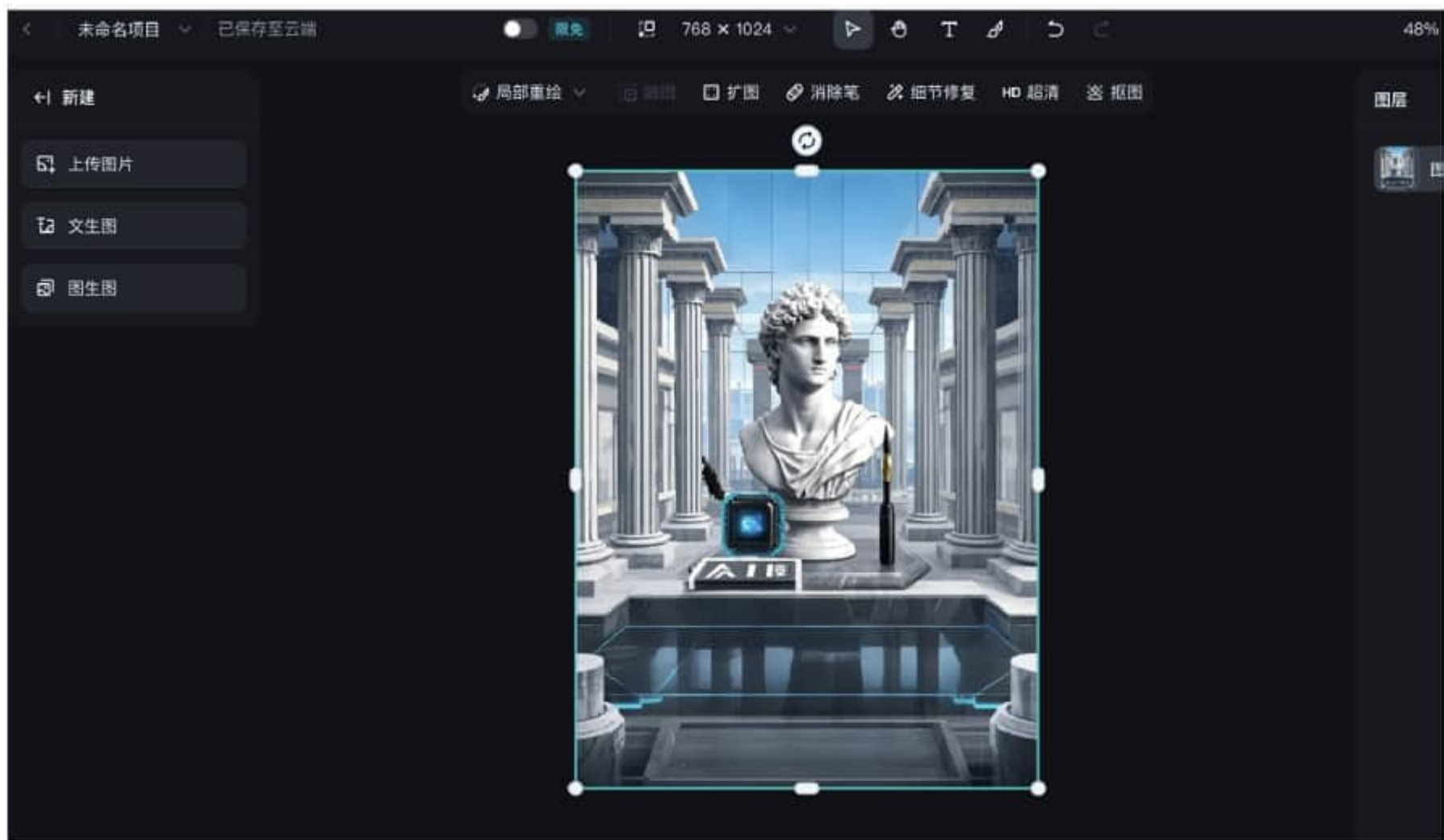
3:4



如何用 AI 制作海报-筛选结果



如何用 AI 制作海报-画布编辑



如何用 AI 制作海报-结果例



可控图像生成·为什么生成图像不稳定



回顾上节课的内容

- 模型能力不够：模型太小 / 训练数据太少（SD一开始手画不好，到 Flux 架构逐渐成熟）
- 提示词没写好，使用太多自然对话，而不是模型“喜欢的语言”，即与训练集相似的语言
 - ❌ 你这个颜色不够红，给我鲜艳点颜色的的苹果
 - ✅ 红色的苹果、深红色、亮



1、我们需要什么稳定性？ 2、“参考”图片发生了什么事？

可控图像生成·为什么生成图像不稳定

当参考一张图并输入提示词后，扩散模型在做什么？

```
# 4. Preprocess image
image = self.image_processor.preprocess(image)

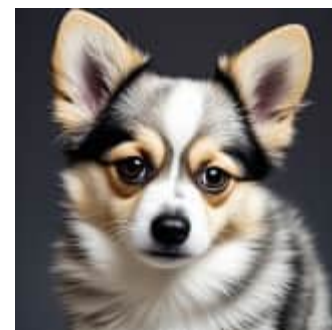
# 5. set timesteps
timesteps, num_inference_steps = retrieve_timesteps(
    self.scheduler, num_inference_steps, device, timesteps, sigmas
)
timesteps, num_inference_steps = self.get_timesteps(num_inference_steps, strength, device)
latent_timestep = timesteps[:1].repeat(batch_size * num_images_per_prompt)

# 6. Prepare latent variables
latents = self.prepare_latents(
    image,
    latent_timestep,
    batch_size,
    num_images_per_prompt,
    prompt_embeds.dtype,
    device,
    generator,
)
```

*基于原图 —— 加噪声 —— 文字条件生成



可爱的狗



可控图像生成·为什么生成图像不稳定

*基于原图 —— 加噪声 —— 文字条件生成 仍然在“文字引导”！天然具备不确定性



💡 当我们需要稳定，我们在寻找什么？

- 主体不变
- 画面元素不变
- 画面风格不变
- 画面可以完全改变但语义不变 / 部分画面需要不变
- 不变的程度：轮廓一致 / 姿势一致

明确自己在“寻找什么”

可控图像生成·图像编辑

一般而言，我们希望对部分区域“言出法随”

下命令的形式



(a) Source Image (b) Target Image in Turn 1 (c) Target Image in Turn 2 (d) Target Image in Turn 3

Multi-turn Editing



(a) Source Image (b) Target Image

Single-turn Editing

参考原图修改提示词影响部分结果



"The boulevards are crowded today."



"Photo of a cat riding on a bicycle."



"My fluffy bunny doll."



"a cake with decorations."

























jelly beans









<https://osu-nlp-group.github.io/MagicBrush/>

<https://arxiv.org/abs/2208.01626>

可控图像生成·图像编辑 如何获得“言出法随”的能力

single-turn, multi-turn, mask-provided, and mask-free editing

		screen	
		replace some bags with throw pillows	
		turn the stop sign into a no parking...	
		remove the lower sign	
		make the airplane a toy airplane	
		What if it was another car on it?	
		The background should be of a...	
		let the cellphone screen be blue	

		["change the back ground"]	Z
		["Making dark color "]	Z
		["make all the colors more vivid..."]	Z
		["Change the whole image into ..."]	Z

只有使用“编辑数据集”的模型，才能随心编辑

可控图像生成·图像编辑 那么，在哪里可以体验呢？

CosXL Model from Stability AI 没错，就是那家做了SD的公司

https://comfyanonymous.github.io/ComfyUI_examples/edit_models/

Gemini 2.0 Flash (Image Generation) Experimental



Editing light settings with a prompt



Creating quick variations of original image (left) by prompting using CosXL Edit model



Add some chocolate drizzle to the croissants.



Run settings

<> [Get code](#)

Model

Gemini 2.0 Flash (Image Generation) Experimental

Output format [NEW](#)

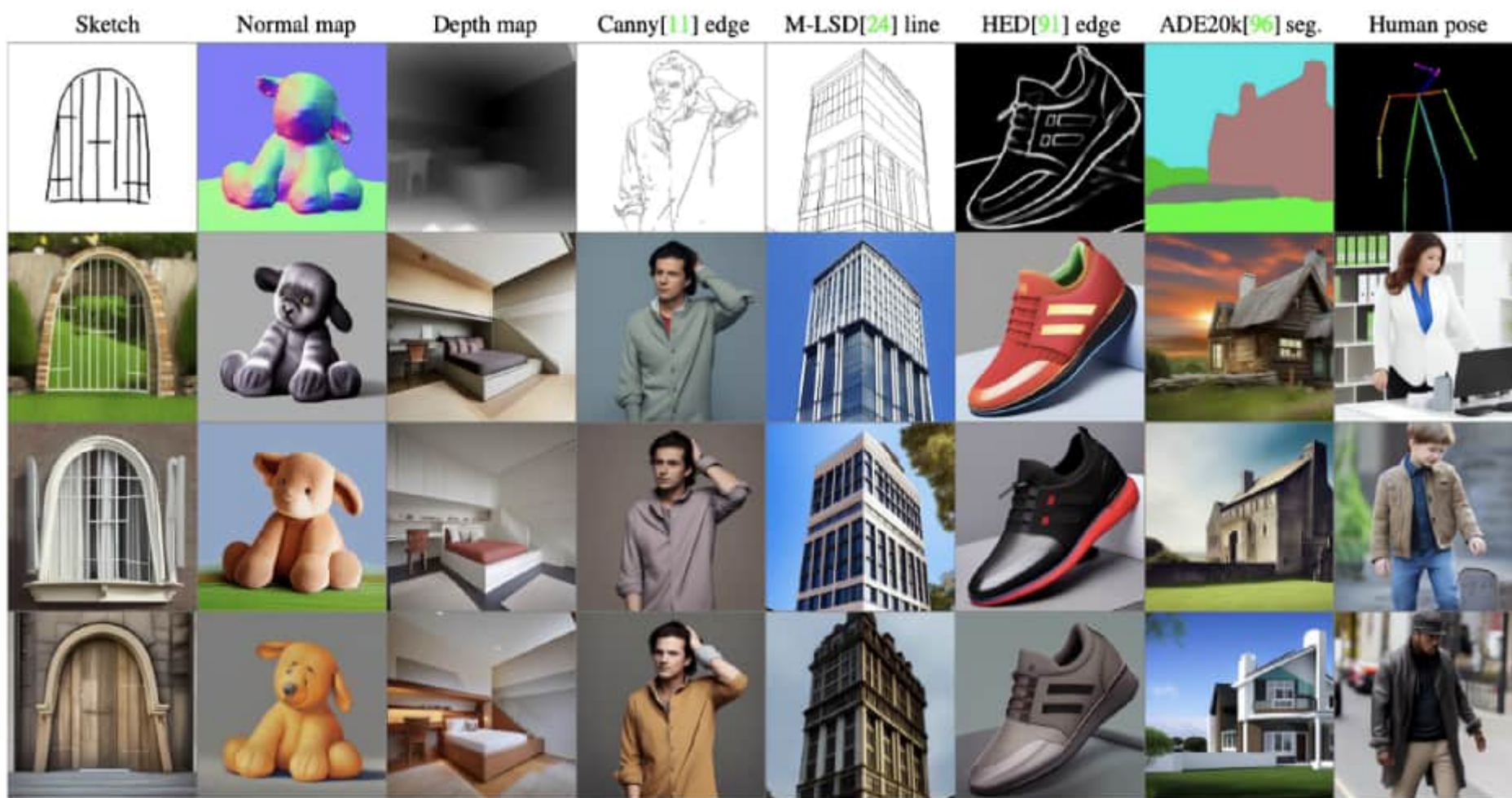
Images and text

Token count 527 / 32,768

Temperature

Advanced settings

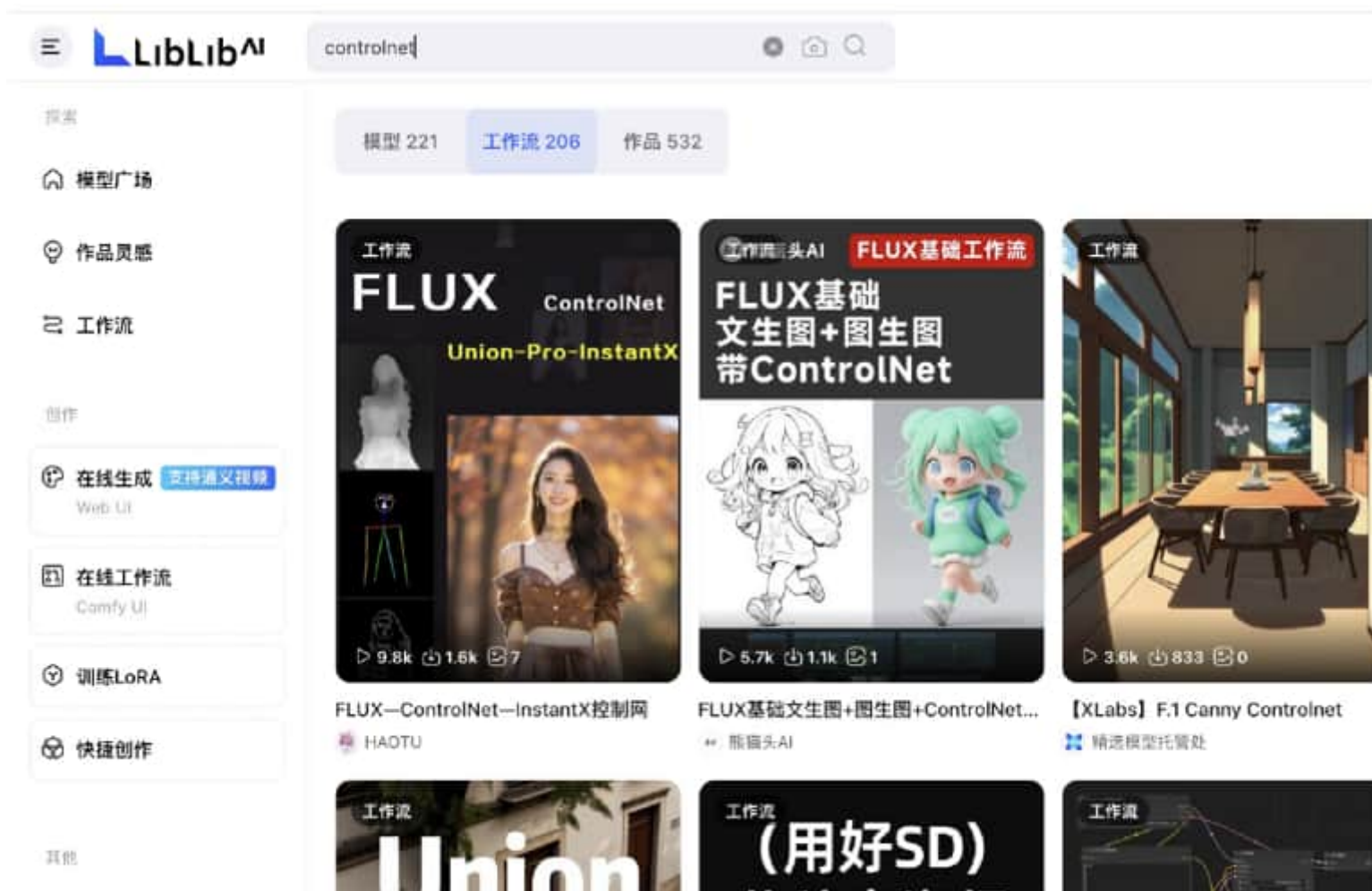
可控图像生成·controlnet <https://arxiv.org/pdf/2302.05543>



可控图像生成·controlnet

那么，在哪里可以体验呢？

<https://www.liblib.art/>



可控图像生成·controlnet

那么，在哪里可以体验呢？

<https://www.liblib.art/sd>

LibLib



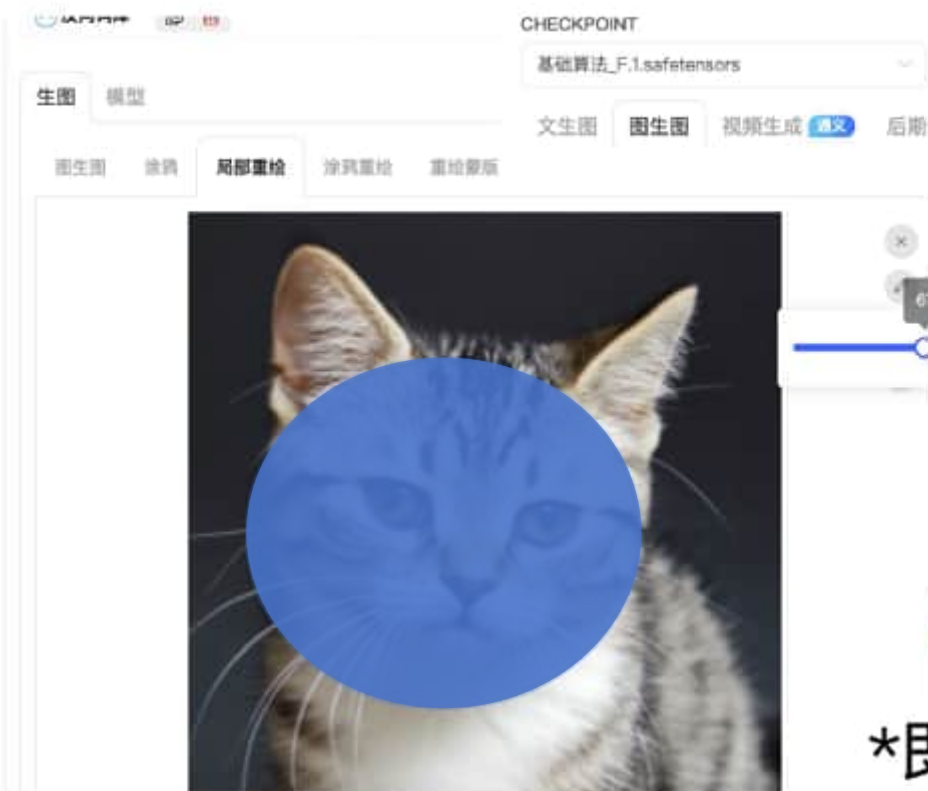
可控图像生成·inpaint（局部处理）

没办法一次关注所有事，我们可以一步步实现

<https://www.liblib.art/sd>

LibLib^{AI}

<https://www.liblib.art/search?keyword=inpaint&type=text>



* 只涂抹需要更改的区域



Flux-Inpainting超强重绘 workflow V2

大筑智障君



Kolors+Inpainting局部重绘（完美版）

耀耀AR松露糖味芋



基础工作流合集【4】-inpainting（局...

J_

*即梦——画布——上传图片——局部重绘

可控图像生成·lora 获得“风格”的强稳定性



都不是我喜欢的结果，怎么办？

- 收集“统一风格”的数据集（30~100张）



打标+训练

放心，这不难



可控图像生成·lora

那么，在哪里可以体验呢？

<https://www.modelscope.cn/aigc/modelTraining>

ModelScope AIGC专区

一站式 A I 创作平台

首页 创意广场 模型 图片生成 视频生成 模型训练 工作流 资源库

发起训练 训练管理 数据管理

模板选择

FLUX

FLUX.1-dev

麦橘超然

动漫

写实

自定义

参数设置

专业参数 团

使用底模

FLUX

MusePublic/489_ckpt_FLUX_1

2172

Checkpoint | FLUX.1

单张次数 Repeat

10

训练轮数 Train epochs

20

推荐范围: 10-12

LoRA模型名称

图片 共 40/100 张 清空所有图片 示例

点击图片

564*381

cozy kitchen scene, sunlight streaming, floral-patterned mug, woven placemat, bowl ...

736*552

twilight sky, star-shaped lights, illuminated buildings, urban setting, serene atmosphere,...

583*560

564*564

720*480

裁剪方式

裁剪尺寸

打标模型

无需裁剪

512*512

JoyCaption

打标

可控图像生成·lora

那么，在哪里可以体验呢？

<https://www.modelscope.cn/aigc/modelTraining>

发起训练

训练管理

数据管理

Checkpoint | FLUX.1

单张次数 Repeat

训练轮数 Train epochs 推荐范围: 10~12

* LoRA模型名称

模型触发词 ?

模型是否公开 ? ☒ 是

⚡ 开始训练

发起训练

训练管理

数据管理

训练进度

训练数据

当前步数/总步数: — 训练轮数/总轮数: —

LoRA

实时样图

模型预览

模型信息

模型名称: test1

触发词:

样图设置

样图分辨率: 768 * 1024

随机种子数: -1

提示词:

保存设置

每N轮保存一个LoRA:

LoRA保存精度: bf16

学习率&优化器

总学习率: 0.0001

优化器: AdamW

数据集设置

图片分辨率: 768 * 1024

是否启动art: 是

可控图像生成·lora

那么，在哪里可以体验呢？

<https://www.modelscope.cn/aigc/modelTraining>

ModelScope AIGC专区

一站式 AI 创作平台

首页创意广场模型图片生成视频生成模型训练

发起训练

训练管理

数据管理



test1 训练中

LoRA类型: FLUX.1-dev

操作人: sanbuphy

开始时间: 2025-03-23 22:44:39



dream

训练时长: 6h

LoRA类型: FI

操作人: sanb

开始时间: 2025-03-23 22:44:39



test_lora 训练完成

训练时长: 4小时36分39秒

LoRA类型: Stable Diffusion XL

操作人: sanbuphy

开始时间: 2024-11-02 09:57:48

可控图像生成·lora

那么，在哪里可以体验呢？

<https://www.modelscope.cn/aigc/modelTraining>



可控图像生成·ipadapter

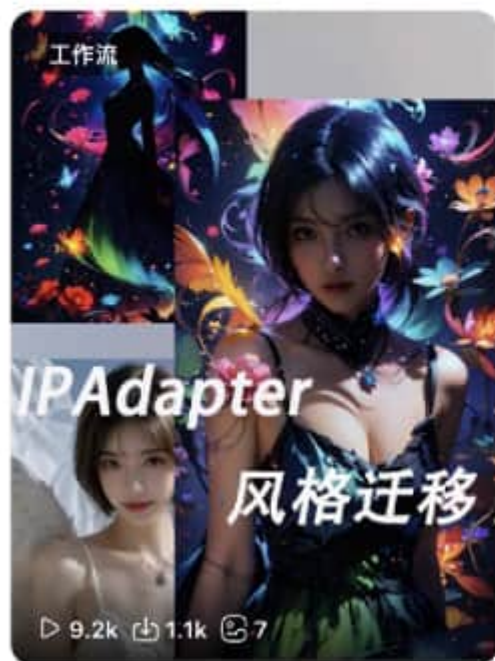
<https://www.liblib.art/search?keyword=ipadapter&type=text>

ipadapter

ipadapter 可以当做特殊的条件注入（类似提示词）

能够协助我们将参考图风格迁移至扩散过程

模型 7 workflow 93 作品 26



一键IPAdapter风格迁移



人物换背景-IPAdapter参考



FLUX最新IPAdapter

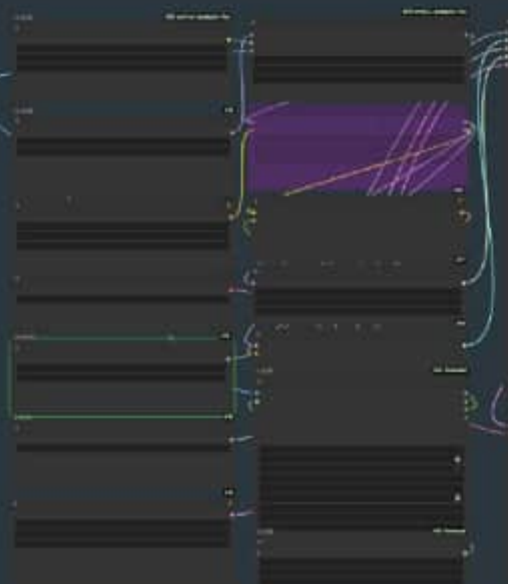
可控图像生成·ipadapter

<https://www.liblib.art/search?keyword=ipadapter&type=text>

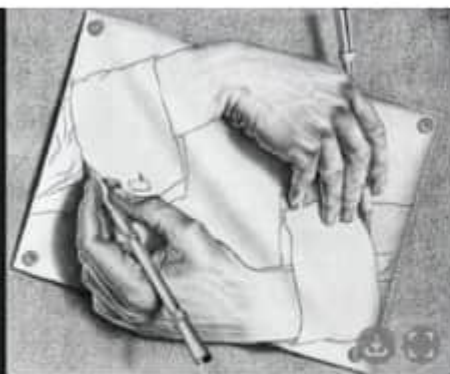
风格参考图



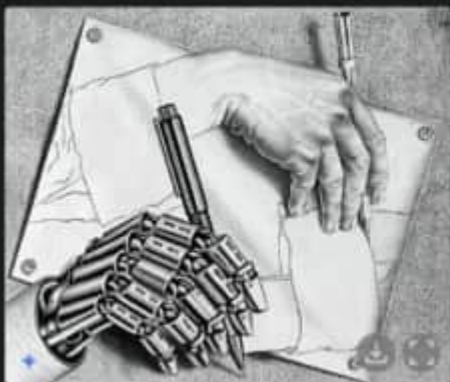
提示词反推图



AI作图的工作流思维 重新审视开头的问题



pls modify the left-bottom hand to a cool realistic robot hand holding a pen



即梦-图片生成 03-16 20:44

参考景深左下的机器手与右上的人类手在相

⚠ 将一件事情分成多步骤

“把这个图片左下角变成机械手，右上角变成油画”

多次“抽卡”，多轮迭代 ↓ 使用多种工具、插件作为节点

1、左下角区域变成机械手 2、右上角区域变成油画风格

- 使用专门的图像编辑模型（支持命令式操作，否则只用提示）
- 使用局部重绘
- 使用局部重绘 + controlnet 锁定姿势
- 使用局部重绘 + controlnet 锁定姿势 + ipadapter 参考风格

优秀实践参考·百度MEUX·controlnet



AIGC 大厂实战设计案例 《用AI提高10倍效率的方法》
<https://www.uisdc.com/zt/aigc-team>

<https://www.uisdc.com/stable-diffusion-34>

优秀实践参考·百度MEUX·controlnet

· SD随机生成

logo+lineart/canny >>>



优秀实践参考·百度MEUX·controlnet

· SD随机生成

logo+canny/lineart+depth >>>



logo



深度图



抽卡跑图

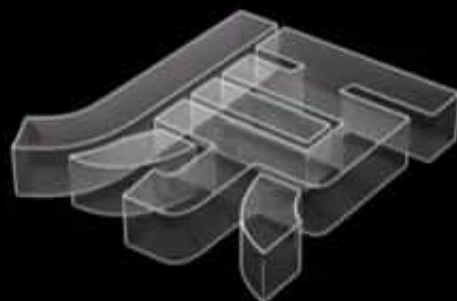


挑一张做海报

优秀实践参考·百度MEUX·controlnet

· 建模辅助制作简单3D图形

立体logo+canny/lineart+depth >>>



优秀实践参考·美团平台人物库的炼丹之旅·lora



<https://www.zcool.com.cn/work/ZNjY4Mjl0MjQ>

优秀实践参考·美团平台人物库的炼丹之旅·lora

#设计策略

对于人物库的构建策略，既要满足日常设计中的多元场景搭建、以及多种风格诉求，还要能够高效的去生产呈现。因此核心从以下几个方面展开设计：

统一风格：重塑人物的风格造型，提升不同产品触点上人物形象的一致性和辨识度。

形成体系：完善人物库成体系化的搭建，沉淀和丰富设计资产。

提升效率：提升人物库的易用性、适用性，并帮助设计师提高他们的工作效率。

角色分类

基于用户画像
设定角色划分

角色实现

角色造型 样图生成
风格设定 模型训练

角色应用与拓展

SD角色生成
多风格拓展

美团平台人物库 AIGC 训练全流程

优秀实践参考·美团平台人物库的炼丹之旅·lora

四、AI样图生成

AI生成要基于大量的样本学习。训练一个LoRA模型大约需要150张不同造型、姿态、表情和服装的高质量样图，每张样图都需要经过拍摄、生成、合成和修图等步骤进行打磨。建立单样图的标准流程可以显著提升样图生成的质量和效率。

• 单张样图——7步生成法



Step01 拍摄（参考图）：需保持动作自然和面部清晰
Step02 通过SD生成初步AI图并生成
Step03 基于AI生成图片，用AI进行修图
Step04 优化输出效果，以原图作为参考图，减少后期修图及后期处理



1. 训练集处理

1.1 图片要求



1.2 Tag要求



优秀实践参考·美团平台人物库的炼丹之旅·lora

Tips

- 训练集图片尺寸长宽数值需要满足 64px 倍数
- 训练集图片可根据景别，使用不同尺寸图片
- 图片需要整体抠图，增加白背景

512px

64px

768px

1024px

2. LoRA训练

2.1 分组渐进式训练

主要优势：效率高、阶段结果及时展示、快速降低模糊度，可根据每个单元测试反馈，进行下阶段增量及迭代。

训练集是LoRA模型训练的基石，训练集样图质量越高，LoRA模型实现效果越佳。

单个LoRA
训练集图量

150+



优秀实践参考·美团平台人物库的炼丹之旅·lora

3.2 压力测试



景别测试



Tag可控性测试



openpose动作测试

优秀实践参考·饿了么AIGC实践·元素控制



解决步骤：

1、通过人工把画面上不必要的乱码剔除

○ 图片上有乱码

○ 通过ps，去掉全部乱码



2、通过 mj 生成设计元素，丰富视觉氛围



<https://mp.weixin.qq.com/s/8xiyNGXnKI6m6srcnZGUwA>

优秀实践参考·饿了么AIGC实践·元素控制

3、人工将各层通过合成技法进行组合，达到符合预期的效果

① 人工合成前景和背景



② 人工将mj生成的设计元素进行合成



③ 人工加上氛围元素丰富画面



④ 人工加上文案信息



1、通过 mj 和人工合成为人物的五官和饰品进行整体调整优化

○ 人物的头饰不够丰富

● 通过mj和ps合成处理生成丰富美观的头饰



合成

调整头饰样式和组合形式，提升美观

绘制

绘制眼睛细节，加强深情对视的神态



合成

增加飘带头饰，丰富人物造型

绘制

调整眼睛细节，加强深情对视的神态

优秀实践参考·盒马设计AIGC探索·符号的产生

Step 1/ 提取营销策略

项目背景

圆满中秋，在盒马的食材/商品中，体会陪伴和热爱的种种（食）光

营销策略

以中秋团圆为背景做最契合氛围的内容演绎，打造最具盒马商品力的中秋尖货

创意策略

月满中秋，欢度佳节，揭开帷幕，盒马陪你一起共赏中秋一场圆满食光好戏

Step 2/ 情绪版建立

	团圆	吃月饼	月桂树	送礼	水果
中秋	赏月	唐王游月	嫦娥奔月	玉兔	饮酒对月
	吴刚	赏桂花	猜字谜	桂花糕	赏花灯

<https://www.shenmabidu.com/post/225083.html>

优秀实践参考·盒马设计AIGC探索·符号的产生

Step 3/ 设计风格确定



团圆, 卡通, 插画, 灯笼, 月亮, 自然

柔和, 温暖, 暖色, 可爱, 渐变

等距视角, 高清, 最佳质量

根据提示词对画面进行更精准的描述, 进行ai绘图以生成精准化的内容, 并且可以通过微调参数, 来按照实际的生成效果可以进一步补充和调整。

优秀实践参考·快手 AIGC 商业运营设计·业务素材细化

步骤拆解：

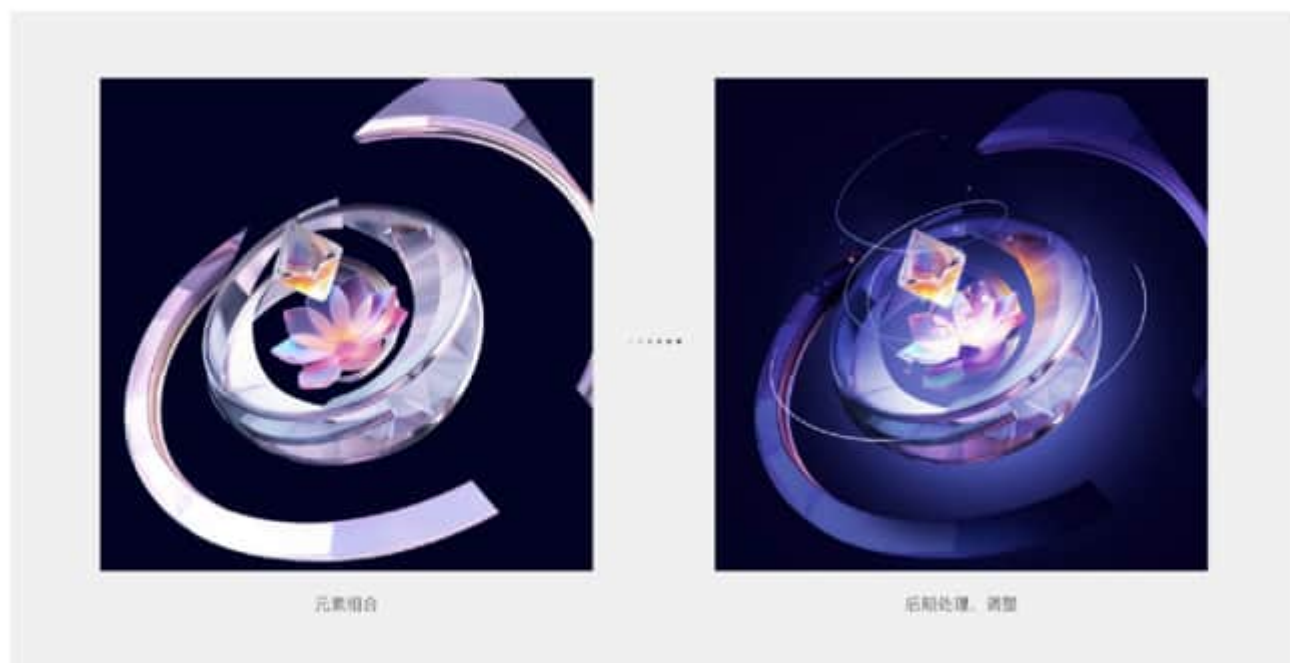
STEP1：搜集风格参考，确定画面的表现风格，构思画面需要的元素。



<https://www.zcool.com.cn/article/ZMTU2ODIyMA>

优秀实践参考·快手 AIGC 商业运营设计·业务素材细化

STEP3: 组合生成的元素, 通过PS后期处理



<https://www.zcool.com.cn/article/ZMTU2ODIyMA>



Thanks

散步 sanbuphy

Datawhale

physicoad@gmail.com

课后作业·必做

1. 尝试用 Deepseek 等大语言模型生成海报元素。
2. 在 liblibai 的 comfyui 处尝试几种控制网络的工作流以及局部重绘功能 / 尝试在 modelscope 社区训练lora
3. 尝试为自己的作品制作海报，可以直接生成或多步骤拼接生成。

课后作业·（选做）



SUBJECT :

SCENE:

STYLE:



Prompt :

毛绒小狗抱着东方明珠



SUBJECT :

SCENE:

STYLE:



Prompt :

巨大的毛绒小狐狸在黄浦江泡澡



SUBJECT :

SCENE:

STYLE:



Prompt :

毛绒小狐狸在咖啡店喝咖啡

学有余力的同学，尝试复刻：<https://labs.google/fx/zh/tools/whisk>