root@autodl-container-811348b9ea-02fee580:~/autodl-tmp/PosterMaker-main# python app.py

You are using a model of type clip\_text\_model to instantiate a model of type . This is not supported for all configurations of models and can yield errors.

You are using a model of type clip\_text\_model to instantiate a model of type . This is not supported for all configurations of models and can yield errors.

You are using a model of type t5 to instantiate a model of type . This is not supported for all configurations of models and can yield errors.

Loading checkpoint shards: 100%|█████████████████████████████████████████████████████████████████████████████████████████████████████████| 2/2 [00:01<00:00, 1.54it/s]

You set `add\_prefix\_space`. The tokenizer needs to be converted from the slow tokenizers

GPU总内存: 23.55 GB

无法启用额外优化: Refer to https://github.com/facebookresearch/xformers for more information on how to install xformers

已启用内存优化技术: 注意力切片, 序列CPU卸载, 手动VAE切片, 梯度检查点, 前向分块, controlnet\_inpaint梯度检查点, controlnet\_inpaint前向分块, controlnet\_text梯度检查点, controlnet\_text前向分块

模型精度配置:

VAE数据类型: torch.float16

Transformer数据类型: torch.float16

文本编码器1数据类型: torch.float16

文本编码器2数据类型: torch.float16

文本编码器3数据类型: torch.float16

ControlNet Inpaint数据类型: torch.float16

ControlNet Text数据类型: torch.float16

适配器数据类型: torch.float16

管道数据类型: torch.float16

CUDA内存使用状态:

已分配: 10.73 GB

已预留: 11.12 GB

总容量: 23.55 GB

可用: 12.43 GB

Running on local URL: http://0.0.0.0:7861

To create a public link, set `share=True` in `launch()`.

==================================================

图像生成流程启动

==================================================

开始图像生成流程

输入参数 - 主图像类型: <class 'numpy.ndarray'>, 掩码图像类型: <class 'numpy.ndarray'>

文本字符串长度: 101, 提示词长度: 431

开始处理主图像格式...

主图像为numpy数组，形状: (1200, 800, 3)

检测到RGB格式

处理掩码图像，类型: <class 'numpy.ndarray'>

掩码图像为numpy数组，形状: (1200, 800, 3)

掩码图像处理完成，尺寸: (800, 1200)

创建随机种子生成器...

调用图像生成器...

图像生成器启动

清理GPU显存...

转换输入图像格式...

主图像形状: (1200, 800, 3), 掩码形状: (1200, 800)

解析文本布局...

解析出2个文本元素

预处理输入数据...

数据预处理完成，输入数据键: ['cond\_image\_inpaint', 'control\_mask', 'prompt', 'text\_embeds', 'target\_size', 'controlnet\_im']

推理前内存清理完成

使用一致的处理分辨率: 1024x1024

开始管道推理，步数: 28, 分辨率: 1024x1024

Token indices sequence length is longer than the specified maximum sequence length for this model (79 > 77). Running this sequence through the model will result in indexing errors

The following part of your input was truncated because CLIP can only handle sequences up to 77 tokens: ['ation.']

Token indices sequence length is longer than the specified maximum sequence length for this model (79 > 77). Running this sequence through the model will result in indexing errors

The following part of your input was truncated because CLIP can only handle sequences up to 77 tokens: ['ation.']

The following part of your input was truncated because CLIP can only handle sequences up to 77 tokens: ['visuals that blend harmonious design with sophistication.']

100%|██████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████████| 28/28 [01:10<00:00, 2.51s/it]

管道推理完成，生成了1张图像

第一张图像类型: <class 'PIL.Image.Image'>

第一张图像尺寸: (1024, 1024)

开始后处理...

后处理完成，最终图像类型: <class 'PIL.Image.Image'>

最终图像尺寸: (682, 1024)

图像已保存至: results/images/generated\_20250531\_134611.png

最终显存清理完成

图像生成成功，返回结果

生成器返回结果类型: <class 'PIL.Image.Image'>

成功: 成功生成图像，尺寸: (682, 1024)，模式: RGB

生成结果类型: <class 'PIL.Image.Image'>

图像生成流程完成

==================================================

图像生成流程结束

==================================================