

4. 古诗转水墨画（大模型）

4.1. 实验内容

古诗和水墨画都是中华文化的重要组成部分，它们源远流长、博大精深，是中国文化的瑰宝。古诗是中国文学的珍品，它以简洁、含蓄、意境深远的语言表达着人们对自然、人生、情感等方面的情感和体验，具有较高的艺术价值和文化内涵。水墨画则是中国绘画的典型代表，它以黑白墨线和水墨渲染为主要表现手法，通过形似神似、运笔精妙的艺术手法，来充分表现山水、花鸟、人物等各种形态的艺术形象，具有极高的传承价值。

本实验内容旨在利用 Stable Diffusion 等深度学习技术，实现根据古诗自动生成水墨画的功能。实验参与者将学习 Python 编程语言在深度学习中的应用，并进一步掌握基于 Python 语言的深度学习框架（Pytorch）。通过对古诗进行自然语言处理和图像风格转换的技术实现，参与者将了解可控生成深度学习大模型的基本原理和关键技术，并实际体验模型效果。

4.2. 实验要求

- 1) 已提供 Stable Diffusion 大模型的代码与预训练参数
- 2) 已提供训练和测试数据
- 3) 要求参与者自行部署 Stable Diffusion 大模型，并自行训练和调参

4.3. 选做内容

- 1) 训练模型生成写实画作
- 2) 分析模型的实际效果，指出如何改进模型，进一步提升效果
- 3) 自行换用其他可控生成模型，尝试效果

4.4. 实验环境

要求使用 Pytorch 深度学习框架完成本次实验，允许使用基于 Python 的 Pandas、Numpy、Sklearn 等算法库进行数据处理。

4.5. 注意事项

不同 Python 库需要版本匹配才能正常运行。

4.6. 参考资料

Python:

<https://www.runoob.com/python3/python3-tutorial.html>

Numpy:

<https://www.numpy.org>

Sklearn:

<https://scikit-learn.org>

Pytorch:

<https://pytorch.org/>

参考文献：

High-Resolution Image Synthesis with Latent Diffusion Models

附：第四次课程设计指导

硬件需求：一张Nvidia 显卡，显存>22G，大概是一张rtx 3090或4090左右。训练任务：输入提示词后，可以生成水墨画风格的图像。

数据集

提供两种格式的数据集，可以根据需要使用（任选其一即可）。

- 图片+标签

标签信息位于csv文件中，格式为：图片名 标签 链接：

<https://yunpan.tongji.edu.cn/link/AACE739FD0079847BC82BE6BB2FE453160>

文件名：Paint4Poem-Web-famous-subset-20250402T083354Z-001.zip

有效期限：永久有效

- HuggingFace的parquet数据集

图片通过二进制封装为字典，使用时需要通过 bytes 访问，最后还要通过io库读取（已经封装好，保存时使用了pil库，本身就是 PIL.Image 类型）。parquet数据集列名为：image 和 text。

```
images=[]
for image in examples[image_column]:
    images.append(io_rgb(image['bytes']))
```

链接：

<https://yunpan.tongji.edu.cn/link/AA8DCF7C32B2924A81B83144503AE5639D>

文件名：dataset.parquet

有效期限：永久有效

预训练模型

链接：

<https://yunpan.tongji.edu.cn/link/AADD470591CCEA4E77A952DCCBA222CCAC>

文件名：stable-diffusion-v1-5.zip

有效期限：永久有效

一些参考资料

HuggingFace: [stable-diffusion-v1-5/stable-diffusion-v1-5 · Hugging Face](#)

Diffusers: [huggingface/diffusers: 😊 Diffusers: State-of-the-art diffusion models for image, video, and audio generation in PyTorch and FLAX.](#)