**Tutorial：从头训练及使用VisualDiffuser**

**NOTE:**

由于上传所有依赖项需要几个G的空间，这里只提供每个步骤的关键代码和操作指导，剩余依赖项需要到相关github仓库下载：

Soft-IntroVAE (https://github.com/taldatech/soft-intro-vae-pytorch),

T2I-Adapter (https://github.com/TencentARC/T2I-Adapter),

MindDiffuser (https://github.com/ReedOnePeck/MindDiffuser)

1. **安装环境**

进入项目VisualDiffuser，根据environment.yml进行安装

cd ./VisualDiffuser/

conda env create -f environment.yml

conda activate VisualDiffuser

1. **训练针对三种Conditions的Soft-Intro VAE**

进入项目soft-intro-vae-pytorch下的soft\_intro\_vae文件，执行 main.py

cd ./VisualDiffuser/soft-intro-vae-pytorch/soft\_intro\_vae/

------------------- Sketch --------------------

python main.py --dataset sketch --device 0 --lr 2e-4 --num\_epochs 100 --beta\_kl 0.5 --beta\_rec 1.0 --beta\_neg 1024 --z\_dim 256 --batch\_size 32

------------------- Color --------------------

python main.py --dataset color --device 0 --lr 2e-4 --num\_epochs 100 --beta\_kl 0.5 --beta\_rec 1.0 --beta\_neg 1024 --z\_dim 256 --batch\_size 16

------------------- Depth --------------------

python main.py --dataset depth --device 0 --lr 2e-4 --num\_epochs 100 --beta\_kl 0.5 --beta\_rec 1.0 --beta\_neg 1024 --z\_dim 256 --batch\_size 32

1. **训练fMRI2Conditions模型**

**3.1提取condition (i.e., sketch, depth, or color) features：**  
进入项目soft-intro-vae-pytorch下的soft\_intro\_vae文件，执行 encode\_and\_save.py

cd ./VisualDiffuser/soft-intro-vae-pytorch/soft\_intro\_vae/

python encode\_and\_save.py

**3.2训练condition feature decoder：**

进入项目MyProjects下的vscode文件，执行fMRI2latent\_ori2ori.py或fMRI2latent\_z-score2ori.py或fMRI2latent\_z-score2z-score.py（ori2ori表示在训练岭回归模型时fMRI (input) 与latent (label) 都未经过正则化，z-score2ori表示只有fMRI经过z-score正则化，z-score2z-score表示两者都经过了正则化。）

cd ./VisualDiffuser/MyProjects/vscode/

python fMRI2latent\_ori2ori.py

四**、训练fMRI2Text模型**

**4.1提取semantic feature：**

进入项目MindDiffuser下的Feature\_extractor文件，执行Semantic\_feature\_extraction.py

cd ./VisualDiffuser/MindDiffuser/Feature\_extractor/

python Semantic\_feature\_extraction.py

**4.2训练semantic feature decoder：**

进入项目MindDiffuser下的Feature\_decoding文件，执行Semantic\_feature\_decoding\_gs.py

cd ./VisualDiffuser/MindDiffuser/Feature\_decoding/

python Semantic\_feature\_decoding\_gs.py

**五、整合并启动图形化界面**

进入项目T2I-Adapter，执行pipeline\_demo\_online.py。执行之后可以点击生成的链接（http://127.0.0.1:7860/）访问基于Gradio的在线demo。其中，访问账号为：demo，密码为：casia-bjut。

cd ./VisualDiffuser/T2I-Adapter/

python pipeline\_demo\_online.py