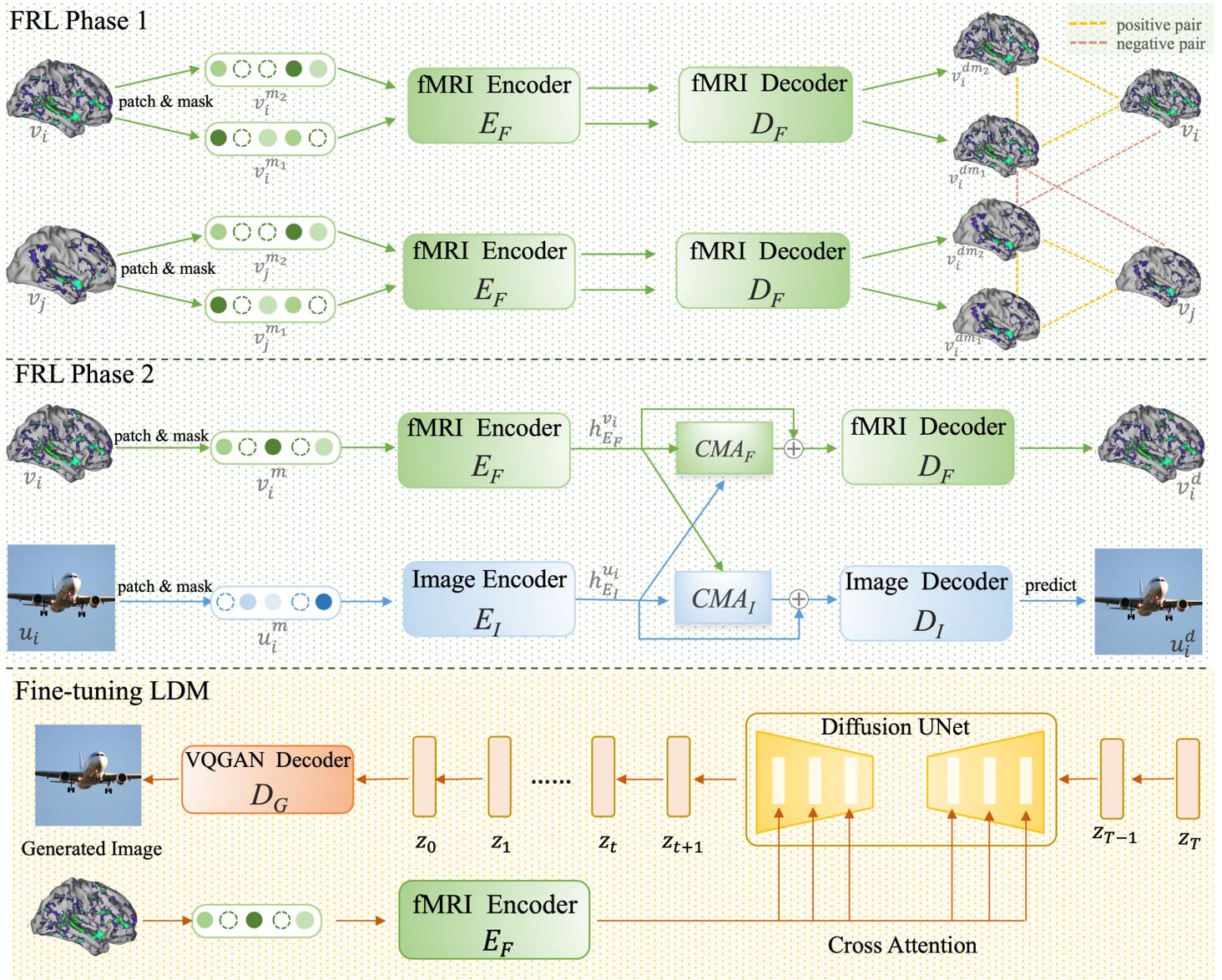
Contrast, Attend and Diffuse to Decode High-Resolution Images from Brain Activities

1. Overview of Proposed Method

We propose a double-phase fMRI representation learning framework. In Phase 1, we pre-train an MAE with a contrastive loss to learn fMRI representations from unlabeled data. After pre-training in Phase 1, we tune the fMRI auto-encoder with an image auto-encoder. When FRL Phase 1 and Phase 2 are done, we apply the representation learned by the fMRI auto-encoder as conditions to tune the LDM and generate the image from the brain activities.

2. Code Usage

Please note that, the full running of the codes needs the HCP, GOD and BOLD5000 dataset. They are very large in sizes and require downloading from additional urls. But adding additional url is not allowed due to the requirement of anominity. So this code only *serves as a demo to help understand the methods*. We will add links of the required data to gurantee reproduction once the paper is accepted.

FMRI Representaiton Learning Phase 1

```
python -m torch.distributed.launch --nproc_per_node=1 \
code/stageA1_mbm_pretrain_contrast.py \
--output_path .  \
--contrast_loss_weight 1 \
--batch_size 250 \
--do_self_contrast True \
--do_cross_contrast True \
--self_contrast_loss_weight 1 \
--cross_contrast_loss_weight 0.5 \
--mask_ratio 0.75 \
--num_epoch 140 \
```

- *do_self_contrast* and *do_contrast_contrast* control whether or not self_contrast and contrast_contrast loss are used.
- *self_contrast_loss_weight* and *cross_contrast_loss_weight* denote the weight of self-contrast and cross-contrast loss in the joint loss.

FMRI Representaiton Learning Phase 2

```
python -m torch.distributed.launch --nproc_per_node=0 \
code/stageA2_mbm_finetune_cross.py \
--dataset GOD --pretrain_mbm_path [Phase 1 checkpoint path] \
--batch_size 16 \
--num_epoch 60 \
--fmri_decoder_layers 6 \
--img_decoder_layers 6 \
--fmri_recon_weight 0.25 \
--img_recon_weight 1.5 \
--img_mask_ratio 0.5 \
--mask_ratio 0.75
```

- *fmri_recon_weight* and *img_recon_weight* denote the weight of fMRI and image reconstruction loss.
- *img_mask_ratio* and *mask_ratio* denote the masking ratio on image and fMRI input.
  Running this code reproduces the fMRI autoencoder that leads to our best result on GOD

subject 3 as reported in the paper.

Fine-tuning LDM

```
python code/stageB_ldm_finetune.py \
--group_name [wandb group] \
--exp_name [wandb run name] \
--phaseA_name CrossMAE \
--pretrain_mbm_path [FRL Phase 2 checkpoint] \
--num_epoch 1000 \
--batch_size 8 \
--is_cross_mae \
--dataset GOD \
--kam_subs "sbj_3" \
--target_sub_train_proportion 1. \
--lr 5.3e-5 \
```