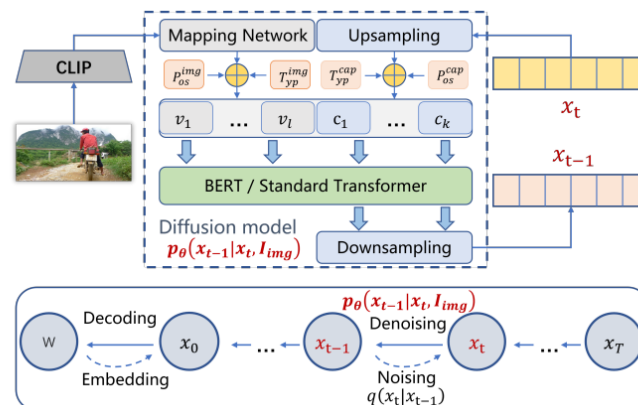


任务：图像描述，给定一张图片，生成对应的描述。



1. Setup:

cd diffusion

- `conda install mpi4py`
- `conda install pytorch torchvision torchaudio cudatoolkit=11.3 -c pytorch`
- `pip install -e improved-diffusion/`
- `pip install -e transformers/`
- `pip install spacy==3.2.4`
- `pip install datasets==1.8.0`
- `pip install huggingface_hub==0.4.0`
- `pip install wandb`

2. experiment:

cd improved-diffusion

Train:

```
bash script/train_diff.sh
```

evaluate:

```
bash script/generation.sh
```

3. others

Train process:

- ◆ Scripts/train.py→improved_diffusion/train_util.py(训练过程)
- ◆ improved_diffusion/Transformer_model2.py 中 TransformerNetModel2 类(line 706)是要训练的 diffusion 模型，本质是一个 bert 和上采样，下采样。
- ◆ improved_diffusion/gaussian_diffusion.py 包含 diffusion 的过程。
- ◆ Scripts/infill_eval_coco.py 是测试代码。