231229 이윤성

알고리즘 : StableVideo
https://github.com/rese1f/StableVideo

StableVideo : Text-driven Consistency-aware Diffusion Video Editing
- 텍스트 기반 비디오 확산 알고리즘
- VRAM Requirement : ~ 29GB
- Download Pretained Model : https://huggingface.co/lllyasviel/ControlNet
→ HuggingFace Model
- control_sd15_canny.pth (5.7GB)
- control_sd15_depth.pth (5.7GB)
- etc...

사용 포트와 torch, cuda version
aiuser08 1234
docker run -it -gpus all -name ys1 pytorch/pytorch:2.0.1-cuda11.7-cudnn8-runtime

docker start and attach

```
aiuser08@aiart:~$ docker start ys1
ys1
aiuser08@aiart:~$ docker attach ys1
root@7235caa9cd00:/workspace#
```
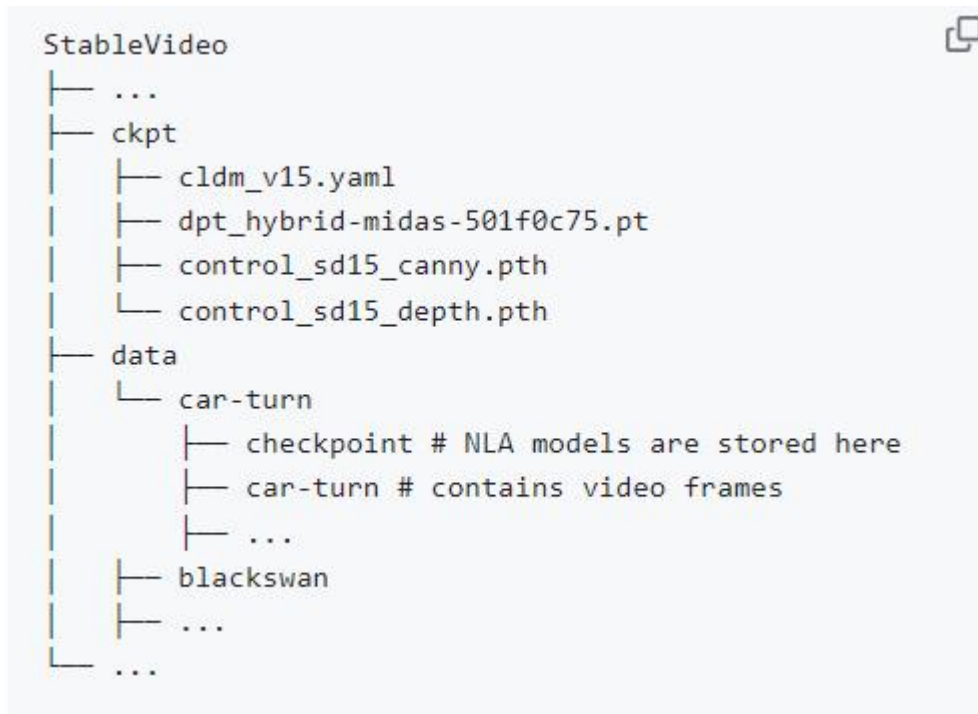
git clone and download files

```
root@7235caa9cd00:/workspace# ls
StableVideo
root@7235caa9cd00:/workspace# cd StableVideo/
root@7235caa9cd00:/workspace/StableVideo# ls
LICENSE      annotator   ckpt   ldm   requirements.txt
README.md    app.py      cldm   log   stablevideo
root@7235caa9cd00:/workspace/StableVideo#
```

현재 ckpt 폴더 안 내용

```
root@7235caa9cd00:/workspace/StableVideo# cd ckpt/
root@7235caa9cd00:/workspace/StableVideo/ckpt# ls
cldm_v15.yaml
root@7235caa9cd00:/workspace/StableVideo/ckpt#
```

필요 ckpt

```
StableVideo
├── ...
├── ckpt
│   ├── cldm_v15.yaml
│   ├── dpt_hybrid-midas-501f0c75.pt
│   ├── control_sd15_canny.pth
│   └── control_sd15_depth.pth
├── data
│   └── car-turn
│       ├── checkpoint # NLA models are stored here
│       ├── car-turn # contains video frames
│       ├── ...
│   ├── blackswan
│   ├── ...
└── ...
```

실행 결과

```
root@7235caa9cd00:/workspace/StableVideo# python app.py
No module 'xformers'. Proceeding without it.
ControlLDM: Running in eps-prediction mode
DiffusionWrapper has 859.52 M params.
making attention of type 'vanilla' with 512 in_channels
Working with z of shape (1, 4, 32, 32) = 4096 dimensions.
making attention of type 'vanilla' with 512 in_channels
Loaded model config from [ckpt/cldm_v15.yaml]
Traceback (most recent call last):
  File "/workspace/StableVideo/app.py", line 358, in <module>
    stablevideo.load_canny_model()
  File "/workspace/StableVideo/app.py", line 51, in load_canny_model
    canny_model.load_state_dict(load_state_dict(canny_model_cfg, location='cuda'
), strict=False)
  File "/workspace/StableVideo/cldm/model.py", line 18, in load_state_dict
    state_dict = get_state_dict(torch.load(ckpt_path, map_location=torch.device(
location)))
  File "/opt/conda/lib/python3.10/site-packages/torch/serialization.py", line 79
1, in load
    with _open_file_like(f, 'rb') as opened_file:
  File "/opt/conda/lib/python3.10/site-packages/torch/serialization.py", line 27
1, in _open_file_like
    return _open_file(name_or_buffer, mode)
  File "/opt/conda/lib/python3.10/site-packages/torch/serialization.py", line 25
2, in __init__
    super().__init__(open(name, mode))
FileNotFoundError: [Errno 2] No such file or directory: 'ckpt/control_sd15_canny
.pth'
root@7235caa9cd00:/workspace/StableVideo#
```

- (optional) need xformers
- Error : FileNotFoundError: no ckpt

향후 계획
- ckpt 다운로드 (서버에 완전히 다운로드, clone x) 후 다시 돌림
- ckpt 결과 보고 fine-tuning (LoRA etc...)