

Supplementary Information

The source code of this study is publicly available at <https://github.com/ruijie-wang-uzh/CSegSynth>.

We primarily used two GPU servers with the following identical configurations:

```
CPU      : AMD EPYC 9124 16-Core Processor
Cores    : 64
Memory   : 755 GB
GPUs     : NVIDIA GeForce RTX 4090 (24564 MB)
          NVIDIA GeForce RTX 4090 (24564 MB)
          NVIDIA GeForce RTX 4090 (24564 MB)
          NVIDIA GeForce RTX 4090 (24564 MB)
          NVIDIA GeForce RTX 4090 (24564 MB)
          NVIDIA GeForce RTX 4090 (24564 MB)
          NVIDIA GeForce RTX 4090 (24564 MB)
          NVIDIA GeForce RTX 4090 (24564 MB)
```

For the pre-training of α -GAN, we also temporarily used NVIDIA A100 GPUs with 80GB memory on a high-performance computing cluster.

In the following, we provide snippets of our code's running log that report specific configurations of our training processes. Please adopt these configurations when reproducing the results. Also, we report the running time of each training process.

- Pre-training of VAE

Running time: 23h 27m

```
## VAE Aomic Experiments - 2024.05.15.22.25.54
* dataset: Aomic
* data_path: datasets/Aomic/output
* down_factor: 2
* split: [0.9, 0.1]
* batch_size: 7
* save_freq: 25
* num_gen_samples: 3
* num_epochs: 5000
* channel_base: 32
* pre_timestamp:
* pre_epoch: 0
* lr: 1e-05
* kl_weight: 0.0005
* random_seed: 0
* model_path: datasets/Aomic/models
* grad_clip: 1
* num_devices: 7
```

- Pre-training of GAN

Run-1:

Running time: 1d 8h 22m 36s

```
## GAN Aomic Experiments - 2024.05.22.14.49.19
* dataset: Aomic
```

```
* data_path: datasets/Aomic/output
* down_factor: 2
* batch_size: 32
* save_freq: 50
* num_gen_samples: 3
* num_epochs: 5000
* channel_base: 32
* pre_timestamp:
* pre_epoch: 0
* gp_lambda: 10
* dis_warmup: 3000
* num_g_iters: 10
* lr: 2e-05
* random_seed: 0
* model_path: datasets/Aomic/models
* d_grad_clip: 1
* g_grad_clip: 1
* num_devices: 8
```

Run-2:

Running time: 4d 9h 15m 22s

GAN Aomic Experiments - 2024.05.24.23.39.22

```
* dataset: Aomic
* data_path: datasets/Aomic/output
* down_factor: 2
* batch_size: 32
* save_freq: 50
* num_gen_samples: 3
* num_epochs: 5000
* channel_base: 32
* pre_timestamp: 2024.05.22.14.49.19
* pre_epoch: 1700
* gp_lambda: 10
* dis_warmup: 200
* num_g_iters: 4
* lr: 2e-06
* random_seed: 0
* model_path: datasets/Aomic/models
* d_grad_clip: 1
* g_grad_clip: 1
* num_devices: 8
```

Run-3:

Running time: 6d 5h 16m 9s

GAN Aomic Experiments - 2024.05.29.09.14.22

```
* dataset: Aomic
* data_path: datasets/Aomic/output
* down_factor: 2
* batch_size: 32
* save_freq: 200
* num_gen_samples: 3
* num_epochs: 5000
* channel_base: 32
* pre_timestamp: 2024.05.24.23.39.22
* pre_epoch: 6000
* gp_lambda: 10
```

```
* dis_warmup: 200
* num_g_iters: 2
* lr: 5e-07
* random_seed: 0
* model_path: datasets/Aomic/models
* d_grad_clip: 1
* g_grad_clip: 1
* num_devices: 8
```

- Pre-training of LDM

Training the autoencoder architecture:

Running time: 18h 43m 31s

```
## LDM VAE Mix - 2024.10.03.17.34.03
* dataset: Mix
* data_path: datasets/Mix/output
* batch_size: 6
* adv_weight: 0.01
* perceptual_weight: 0.001
* kl_weight: 1e-06
* autoencoder_warm_up_n_epochs: 5
* val_interval: 2
* num_gen_samples: 3
* num_epochs: 200
* lr: 0.0001
* random_seed: 888
* model_path: datasets/Mix/models
* num_devices: 6
* is_eval: False
```

Pretraining:

Running time: 5h 11m 43s

```
## LDM Aomic - 2024.10.04.13.20.04
* dataset: Aomic
* data_path: datasets/Aomic/output
* down_factor: 2
* split: [0.9, 0.1]
* batch_size: 6
* val_interval: 1
* num_gen_samples: 3
* num_epochs: 200
* lr: 0.0001
* random_seed: 888
* model_path: datasets/Aomic/models
* num_devices: 6
* vae_path: datasets/Mix/models
* vae_timestamp: 2024.10.03.17.34.03
```

- Pre-training of α -GAN

Run-1:

Running time: 4d 21h 57m 29s

```
## Alpha-GAN Experiments - 2023.09.08.16.25.00
* data_path: datasets/output/aomic
* down_factor: 2
* split: [0.8, 0.2]
```

```
* batch_size: 8
* save_freq: 5
* num_gen_samples: 3
* num_epochs: 10000
* channel_base: 32
* gp_lambda: 10.0
* recon_w: 10.0
* dis_warmup: 3000
* num_eg_iters: 10
* lr: 2e-05
* random_seed: 0
* out_path: outputs/
* num_devices: 1
```

Run-2:

Running time: 6d 17h 56m 34s

```
## Alpha-GAN Experiments - 2023.09.13.20.55.16
* data_path: datasets/output/aomic
* down_factor: 2
* split: [0.8, 0.2]
* batch_size: 8
* save_freq: 5
* num_gen_samples: 3
* num_epochs: 10000
* channel_base: 32
* pre_timestamp: 2023.09.08.16.25.00
* pre_epoch: 410
* gp_lambda: 10.0
* recon_w: 10.0
* dis_warmup: 3000
* num_eg_iters: 10
* lr: 2e-05
* random_seed: 0
* out_path: outputs/
* num_devices: 1
```

Run-3:

Running time: 6d 21h 54m 59s

```
## Alpha-GAN Experiments - 2023.09.19.09.10.18
* data_path: datasets/output/aomic
* down_factor: 2
* split: [0.8, 0.2]
* batch_size: 16
* save_freq: 5
* num_gen_samples: 3
* num_epochs: 10000
* channel_base: 32
* pre_timestamp: 2023.09.13.20.55.16
* pre_epoch: 450
* gp_lambda: 10.0
* recon_w: 10.0
* dis_warmup: 3000
* num_eg_iters: 10
* lr: 2e-05
* random_seed: 0
* out_path: outputs/
```

* num_devices: 2

Run-4:

Running time: 6d 5h 22s

Alpha-GAN Experiments - 2023.09.27.11.22.15

* data_path: datasets/output/aomic
* down_factor: 2
* split: [0.8, 0.2]
* batch_size: 12
* save_freq: 5
* num_gen_samples: 3
* num_epochs: 10000
* channel_base: 32
* pre_timestamp: 2023.09.19.09.10.18
* pre_epoch: 1105
* gp_lambda: 10.0
* recon_w: 100.0
* dis_warmup: 3000
* num_eg_iters: 4
* lr: 2e-05
* random_seed: 0
* out_path: datasets/all_outputs
* num_devices: 6

Run-5:

Running time: 18d 20h 43m 37s

Alpha-GAN Experiments - 2023.10.25.17.59.58

* data_path: datasets/output/aomic
* down_factor: 2
* split: [0.8, 0.2]
* batch_size: 12
* save_freq: 50
* num_gen_samples: 3
* num_epochs: 10000
* channel_base: 32
* pre_timestamp: 2023.09.27.11.22.15
* pre_epoch: 2150
* gp_lambda: 10.0
* recon_w: 100.0
* dis_warmup: 3000
* num_eg_iters: 4
* lr: 2e-05
* random_seed: 0
* out_path: datasets/all_outputs
* num_devices: 6

- Fine-tuning of C-VAE

Running time: 3h 26m 29s

Conditional VAE CamCan - 2024.06.05.16.21.53

* dataset: CamCan
* data_path: datasets/CamCan/output
* down_factor: 2
* split: [0.8, 0.1, 0.1]
* batch_size: 16
* kld_weight: 0.0005

```
* vol_loss_w: 100
* num_gen_samples: 3
* num_epochs: 2000
* channel_base: 32
* pre_timestamp: 2024.05.15.22.25.54
* pre_epoch: 499
* pre_model_path: datasets/Aomic/models
* dis_warmup: 200
* base_lr: 1e-06
* con_lr: 0.0001
* random_seed: 0
* model_path: datasets/CamCan/models
* num_devices: 8
* is_eval: False
* gradient_clip_val: 1.0
```

- Fine-tuning of C-GAN

Running time: 5d 8h 18m 29s

Conditional-WGAN CamCan - 2024.06.07.15.21.31

```
* dataset: CamCan
* data_path: datasets/CamCan/output
* down_factor: 2
* split: [0.8, 0.1, 0.1]
* batch_size: 8
* num_gen_samples: 3
* num_epochs: 5000
* channel_base: 32
* pre_timestamp: 2024.05.29.09.14.22
* pre_epoch: 10400
* pre_model_path: datasets/Aomic/models
* continue_timestamp:
* gp_lambda: 10
* vol_loss_w: 100.0
* triplet_w: 100.0
* dis_warmup: 2
* num_g_iters: 4
* d_lr: 1e-06
* g_lr: 1e-06
* c_con_lr: 0.0001
* random_seed: 888
* model_path: datasets/CamCan/models
* num_devices: 8
* is_eval: False
* gradient_clip_val: 1.0
* margin_weight: 0.0
```

- Fine-tuning of C-LDM

Running time: 3h 23m 39s

Conditional-LDM CamCan - 2024.10.06.15.27.01

```
* dataset: CamCan
* data_path: datasets/CamCan/output
* down_factor: 2
* split: [0.8, 0.1, 0.1]
* batch_size: 6
* val_interval: 1
```

```
* num_gen_samples: 3
* num_epochs: 200
* lr: 2e-05
* random_seed: 888
* model_path: datasets/CamCan/models
* num_devices: 8
* vae_path: datasets/Mix/models
* vae_timestamp: 2024.10.03.17.34.03
* pretrain_path: datasets/Aomic/models
* pretrain_timestamp: 2024.10.04.13.20.04
* is_eval: False
```

- Fine-tuning of CSegSynth

Running time: 6h 12m 26s

```
## CSegSynth CamCan - 2024.08.13.14.27.40
* dataset: CamCan
* data_path: datasets/CamCan/output
* down_factor: 2
* split: [0.8, 0.1, 0.1]
* batch_size: 8
* num_gen_samples: 3
* num_epochs: 200
* channel_base: 32
* pre_timestamp: 2023.10.25.17.59.58
* pre_epoch: 3000
* pre_model_path: datasets/Aomic/models
* continue_timestamp:
* gp_lambda: 1
* recon_w: 10
* vol_loss_w: 100.0
* triplet_w: 100.0
* dis_warmup: 200
* num_eg_iters: 4
* d_lr: 1e-06
* g_lr: 1e-06
* e_lr: 1e-06
* cd_lr: 1e-06
* c_con_lr: 0.0001
* random_seed: 0
* model_path: datasets/CamCan/models
* num_devices: 8
* is_eval: False
* freeze_img_dis: True
* save_freq: 20
* gradient_clip_val: 1.0
* margin_weight: 0.0
```