

6/16 meeting

應名宥

training result

different structure, different loss

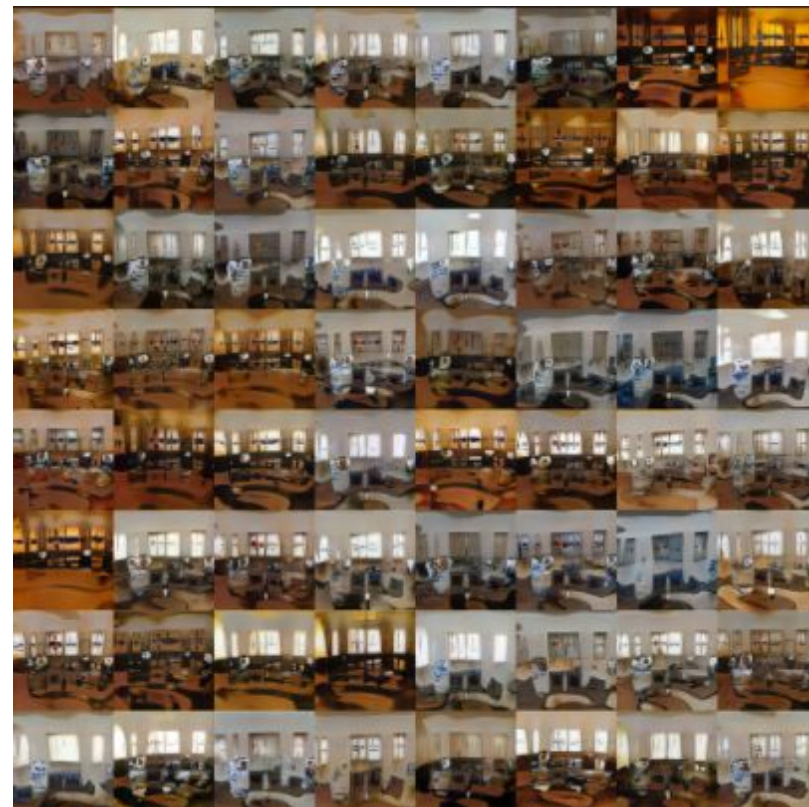
stylegan2

training step : 320 kimg (~1%)

training set : COCO 2014

model size : 0.8 (1)

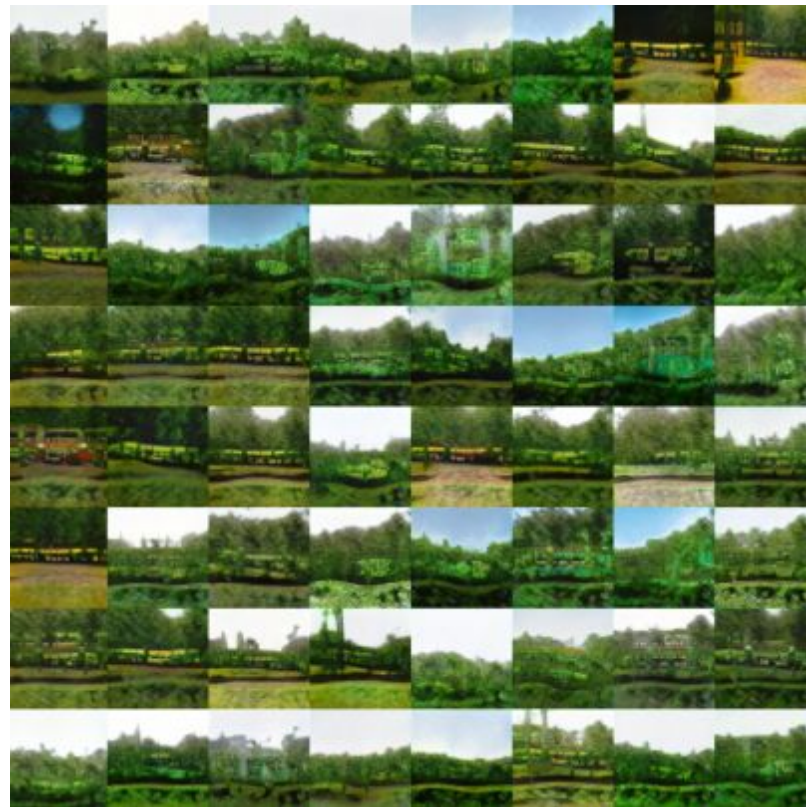
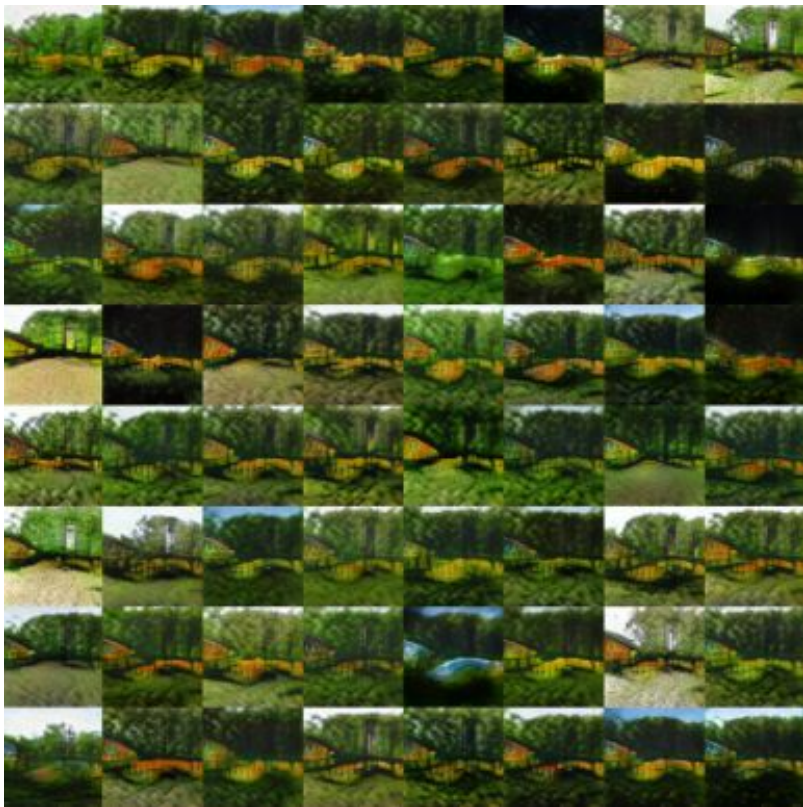
batch size : 8 (16)



A living area with a television and a table

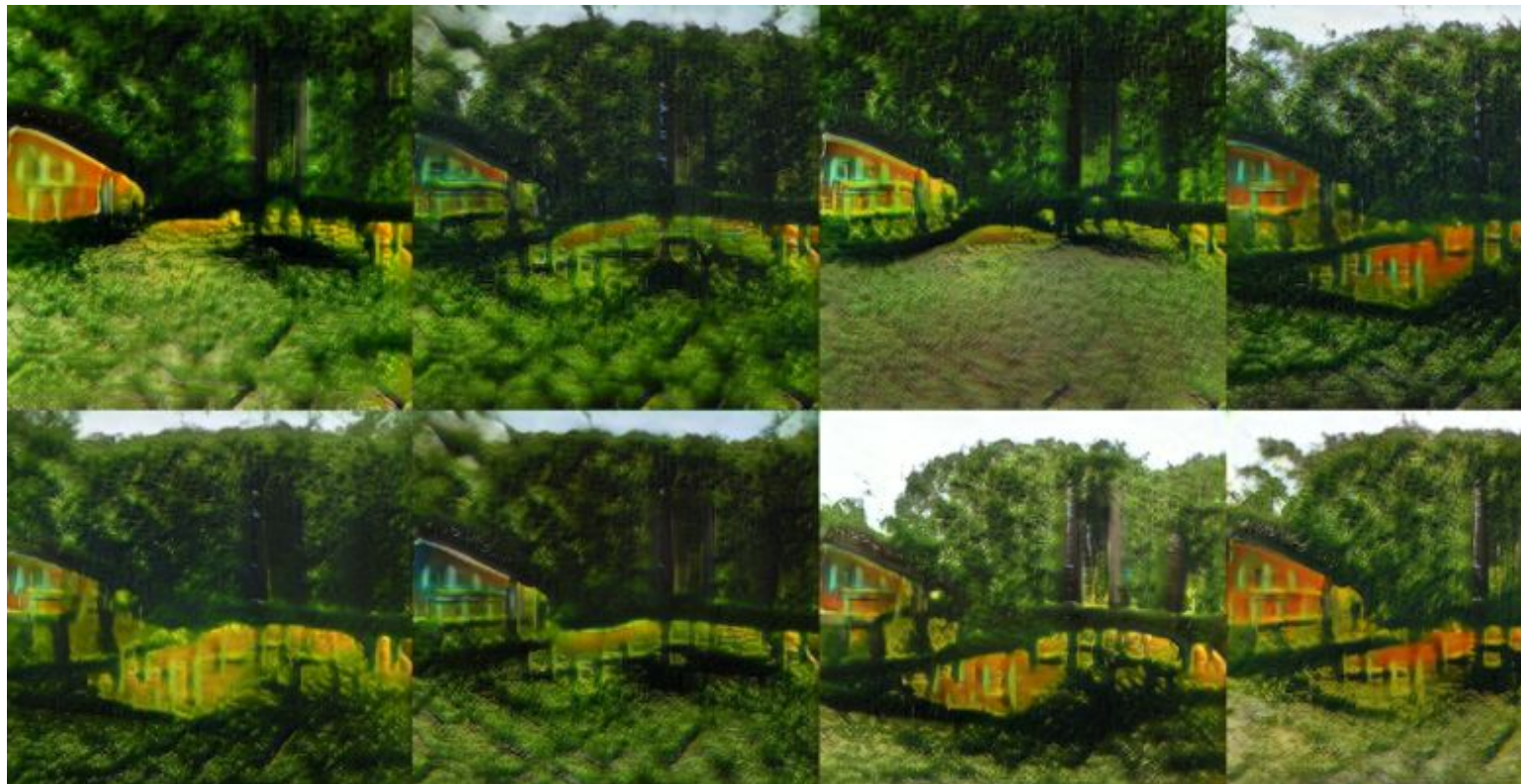


A **group of skiers** are preparing to ski down a **mountain**



A **green train** is coming down the tracks

base (320 kimg)



modified 1 (320 kimg)



modified 2 (260 kimg)





real image



base



modified 1



modified 2

change

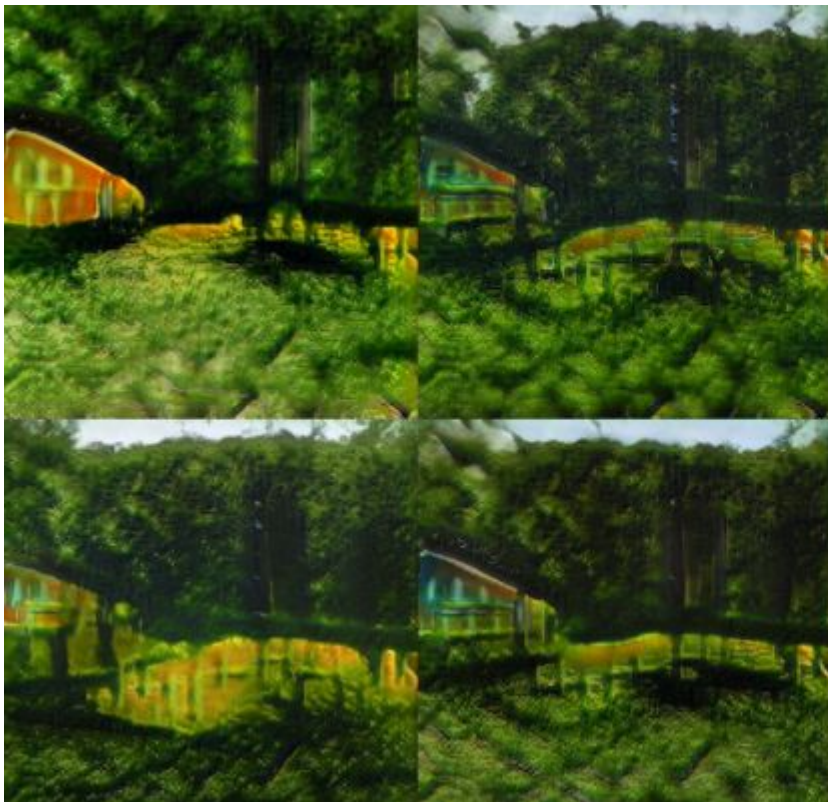
- enable new contrastive learning structure
- add discriminator img-img loss
- add vgg img-img loss
- add clip img-img loss

computation cost

- base
 - 4.25 minutes per kimg
- modified
 - 5.5 minutes per kimg

$$5.5 / 4.25 \sim 1.3$$

diversity



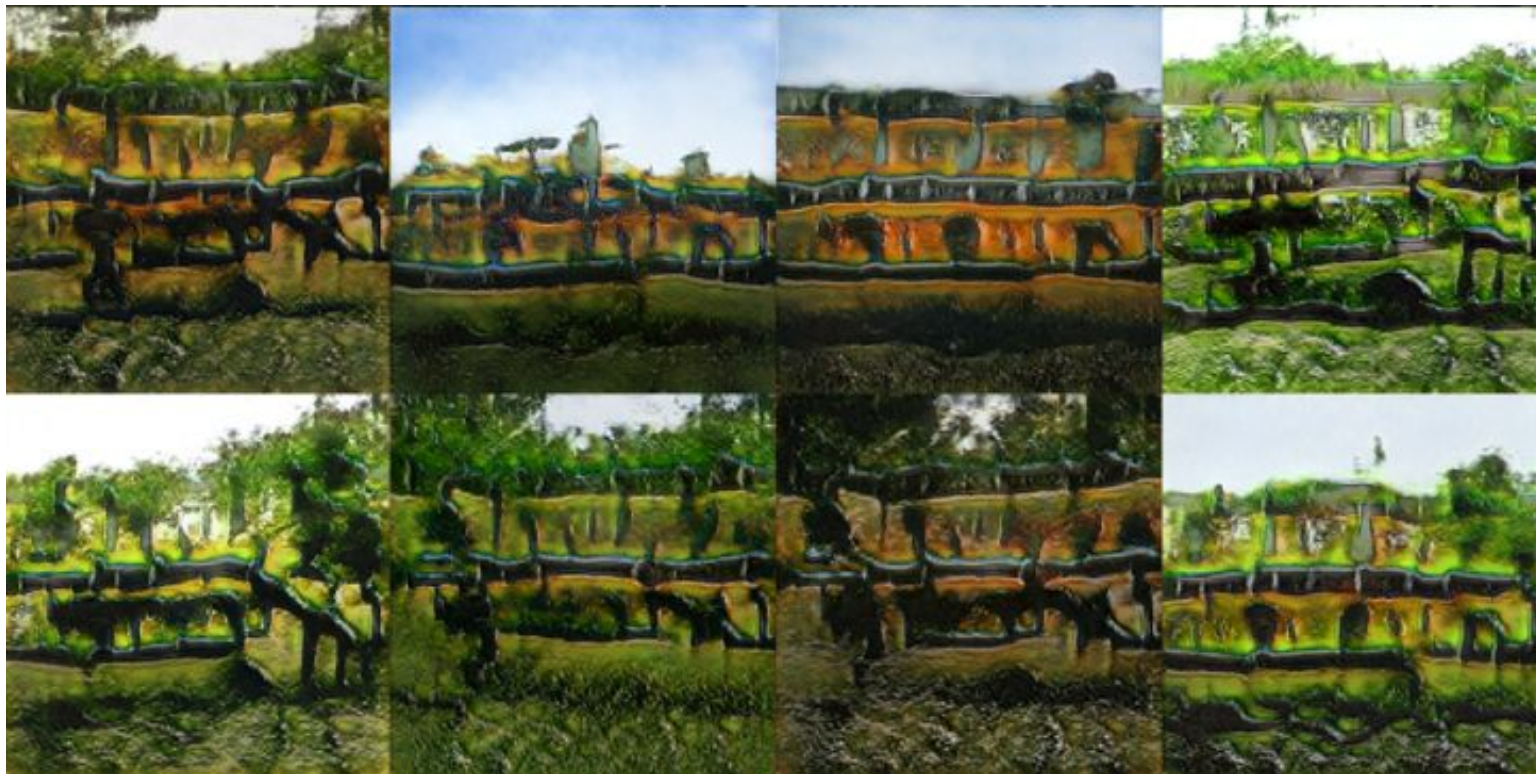
stylegan3

training step : 140 kimg (~0.4%)
training set : COCO 2014
model size : 0.5 (1)
batch size : 8 (16)

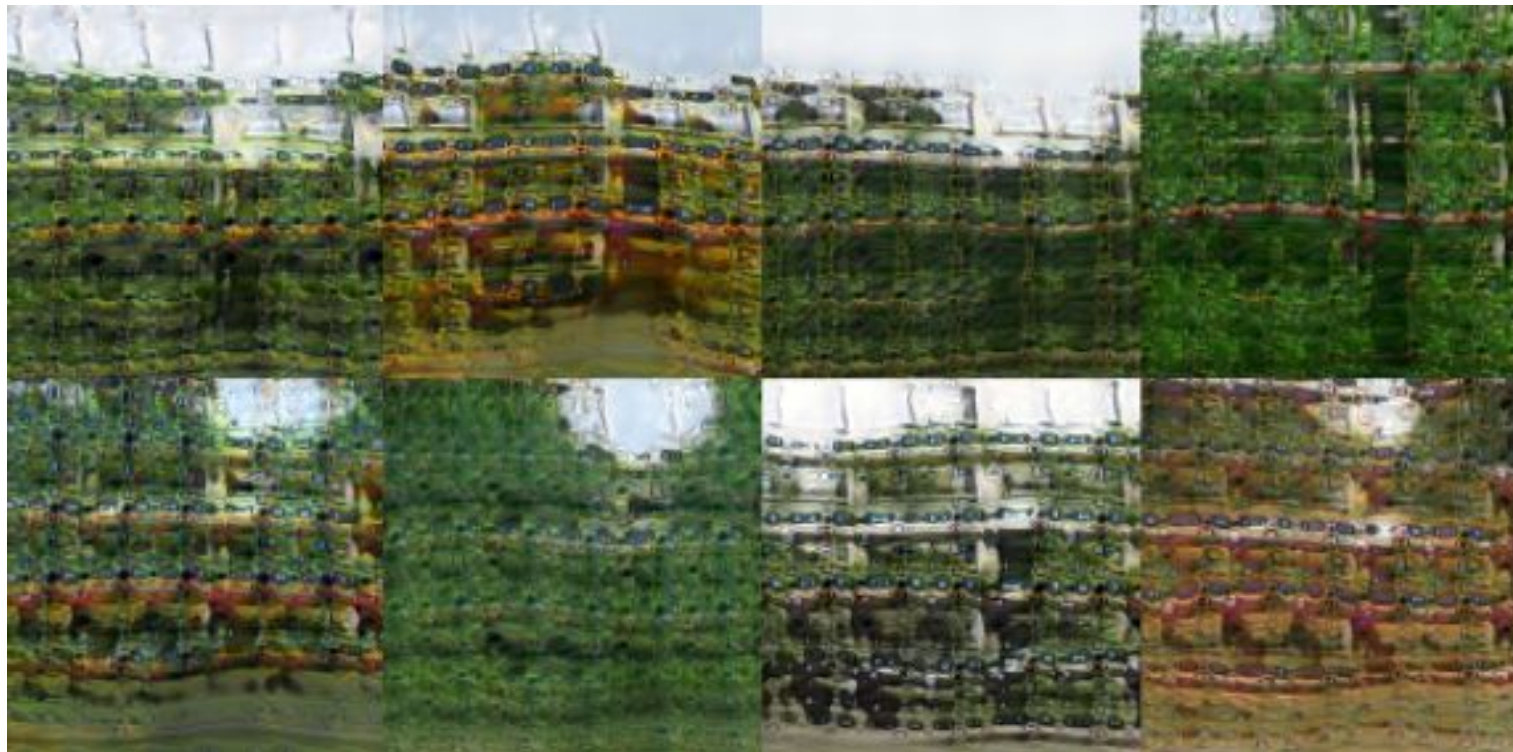


A **green train** is coming down the tracks

stylegan2



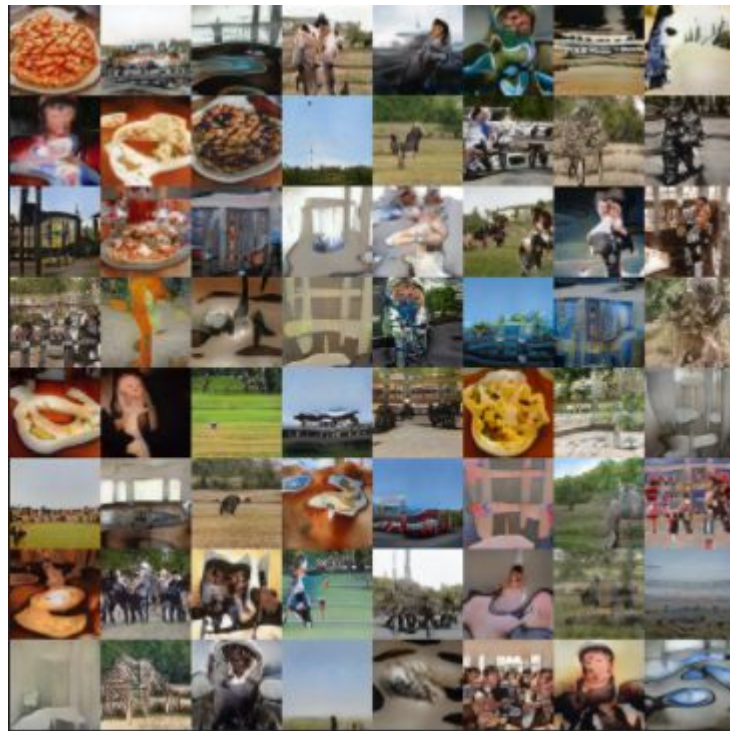
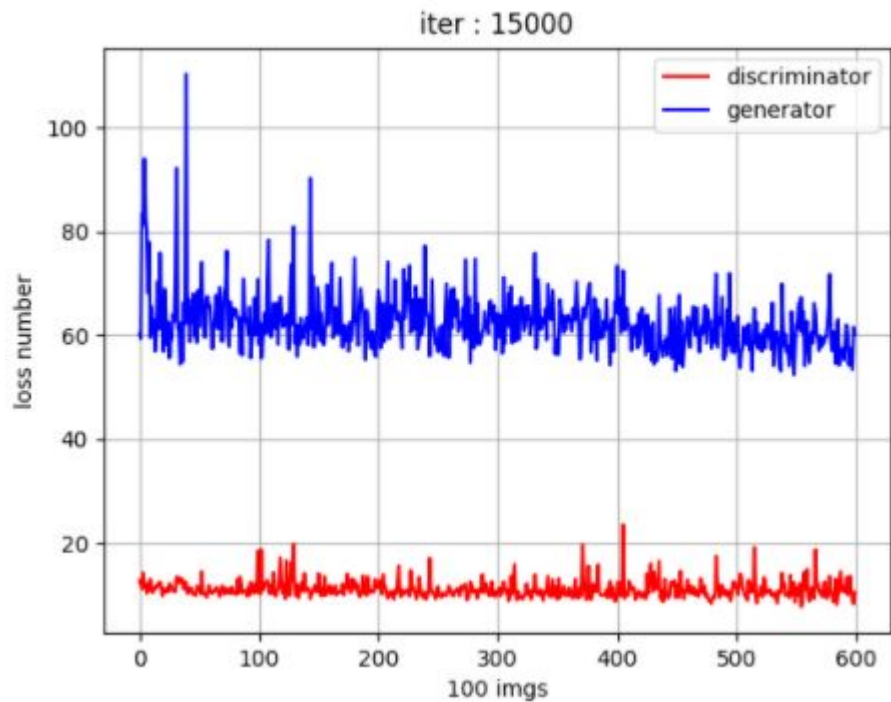
stylegan3



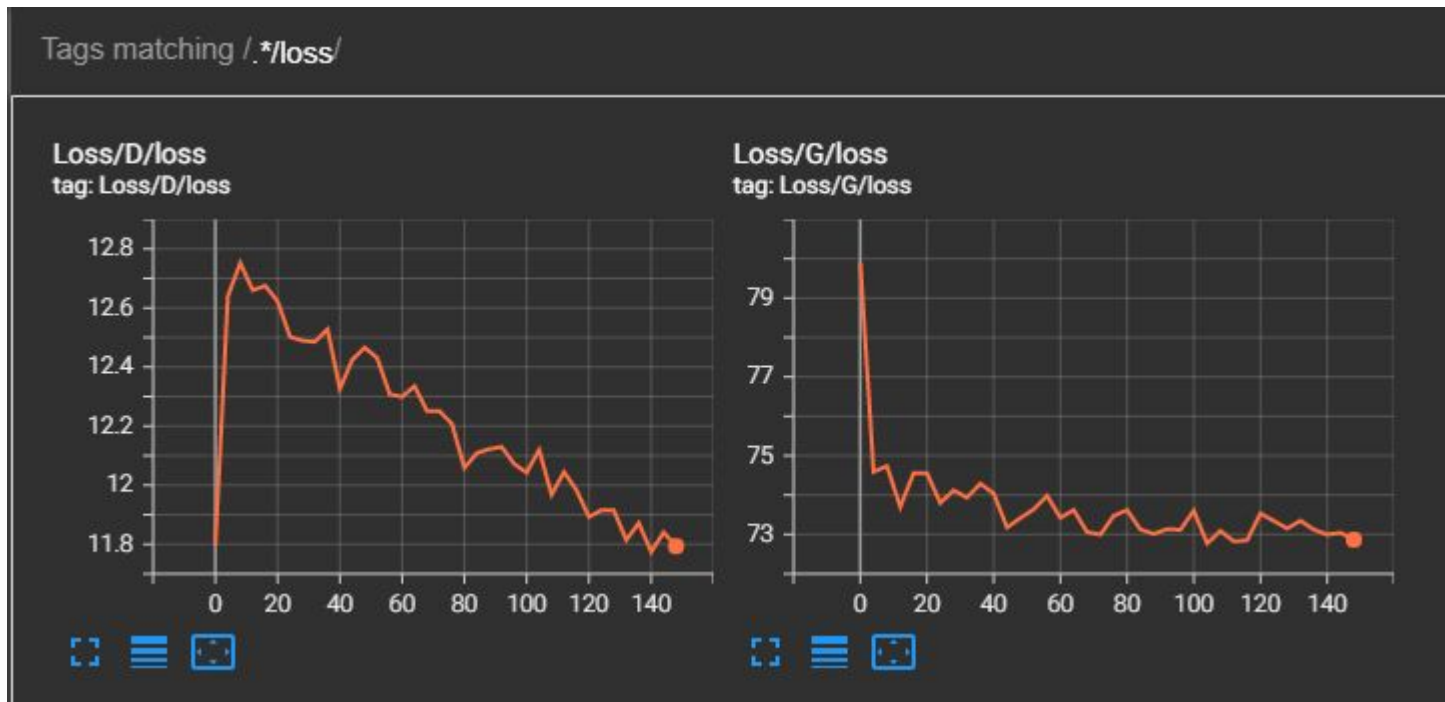
problem

- more gpu memory consumption (influence discriminator)
- hyper-parameter
- implementation problem

training visualization



visualization tool : TensorBoard



training experiment

- model size , batch size
- config ADA
- different loss hyper-parameter combination

in progress

- data augmentation in text or image
- training experiment
- adaptive inference
- inference setting (population, generation)
 - 100 1000
 - 200 500
 - 500 200
 - 1000 100

study

- Style-GAN-2 (✓)
 - <https://arxiv.org/abs/1912.04958>
- Style-GAN-2 ada (✓)
 - <https://arxiv.org/abs/2006.06676>
- Improving Text-to-Image Synthesis Using Contrastive Learning (✓)
 - <https://arxiv.org/abs/2107.02423>
- Style-GAN-3 (✓)
 - <https://arxiv.org/abs/2106.12423>