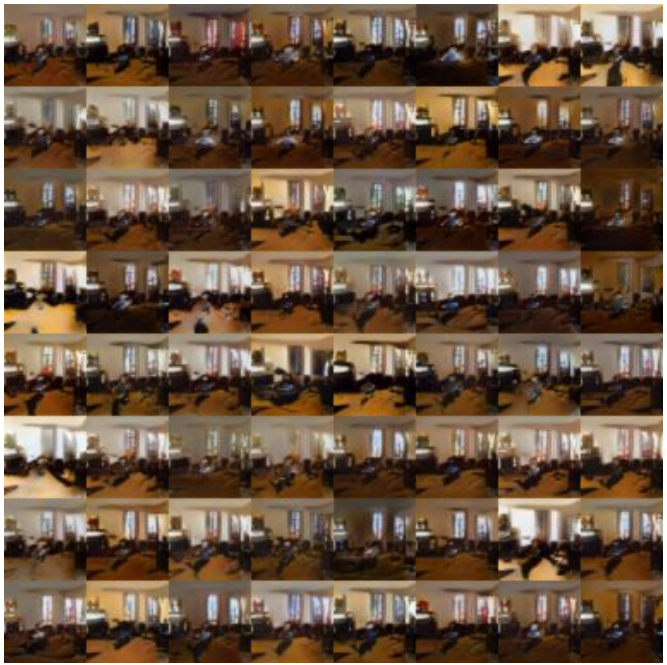
6/10 meeting

應名宥

training result

model size = 4/5
training images = 80000
batch size = 8
training step = 0.5%

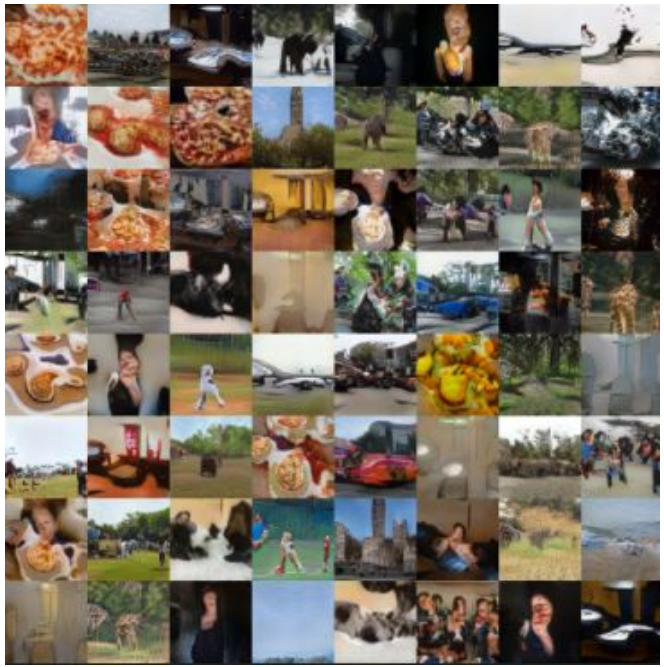
result



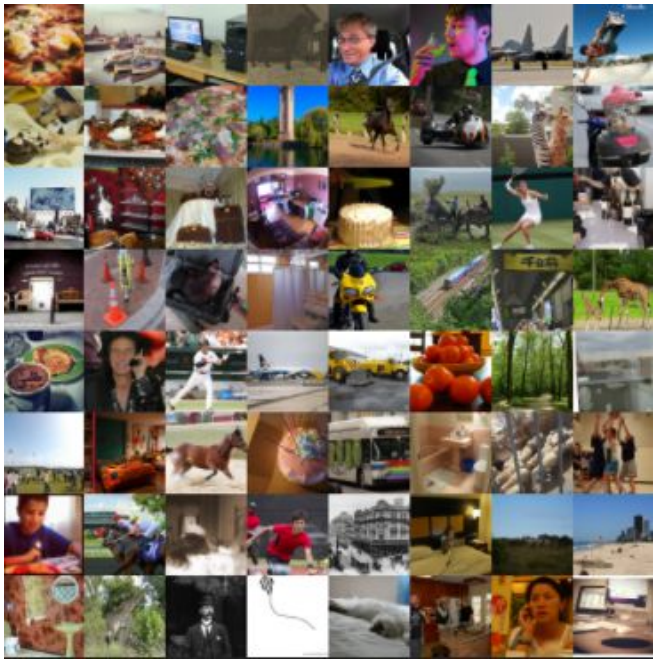
A living area with a television and a table.

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

random captions result



fake



real

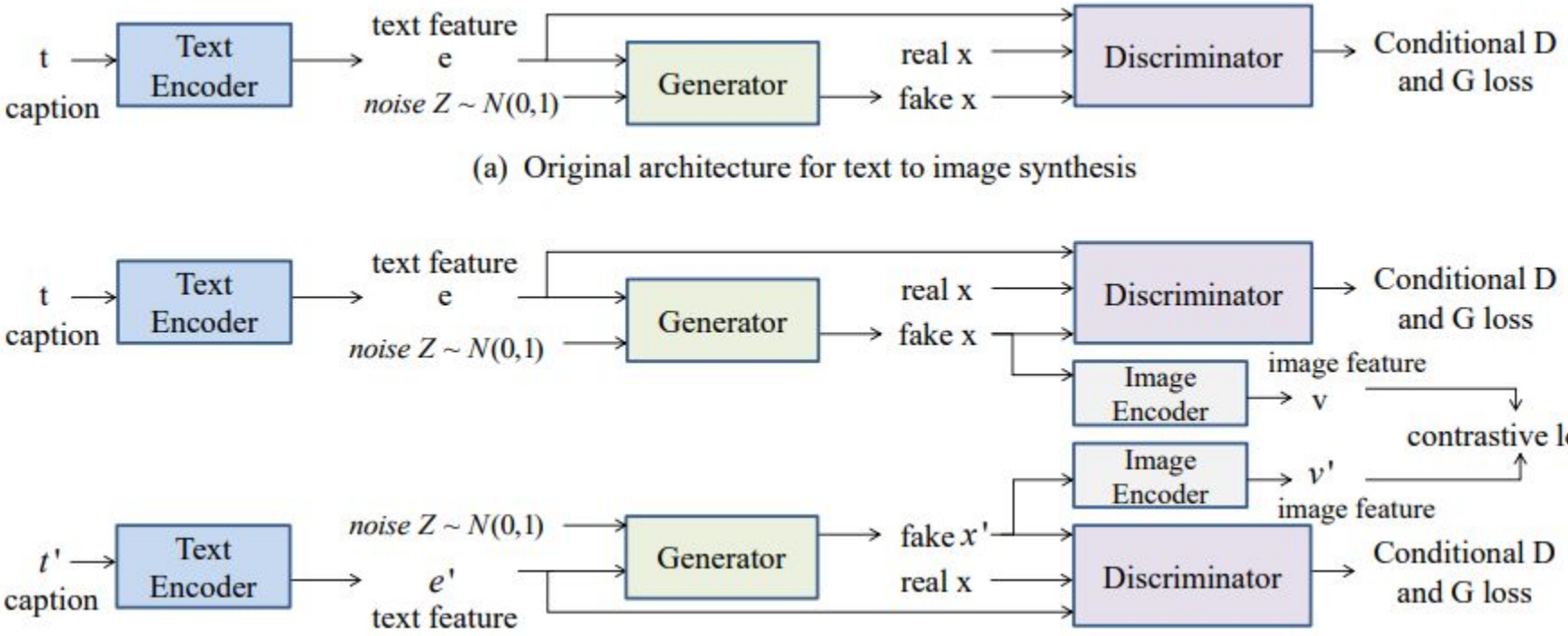
new stuff

loss function

XMC-GAN

We thoroughly evaluate the different components of XMC-GAN and analyze their impact. Table 4 summarizes our ablations[3] on the COCO-14 validation set. To study the effects of each contrastive loss component used in XMC-GAN, we experiment with four losses: (1) image-sentence, (2) region-word, (3) image-image using discriminator features, and (4) image-image using VGG features. For (3), we use the discriminator encoder projection (indicated by D in Table 4) to extract image features. For (4), we extract image features from a VGG-19 network [49] pretrained on ImageNet.

contrasitive structure



testing

- add text features mapping network
- add img-img discriminator contrastive loss
- disable img-txt discriminator contrastive loss

compare : img-txt loss

A living area with a television and a table.

A green train is coming down the tracks.

compare : img-img loss

A living area with a television and a table.

A green train is coming down the tracks.

same caption

img-txt loss

img-img loss

MOO

problem

problem

- high dimension problem
- stuck in the early part
- termination criterion
- adaptive inference
- parameter setting

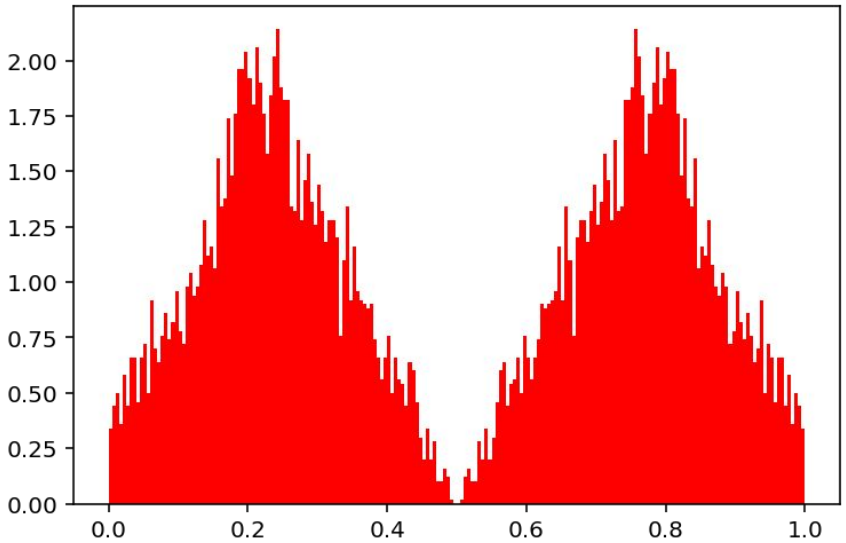
termination criterion

- Number of Evaluations
- Number of Generations
- Based on Time
- Design Space Tolerance (?)

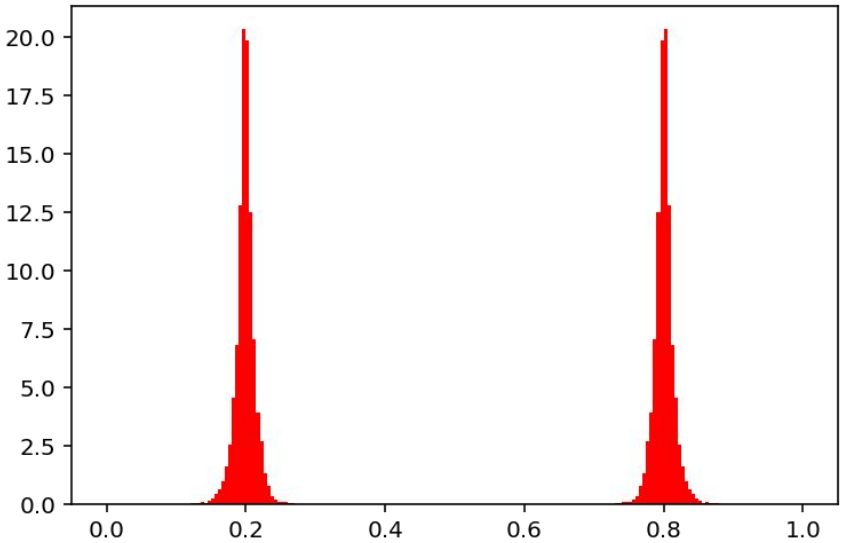
adaptive inference

- have label
  - calculate image-text score
  - calculate image-real score
  - calculate word-label score
- otherwise
  - calculate image-text score

crossover parameter

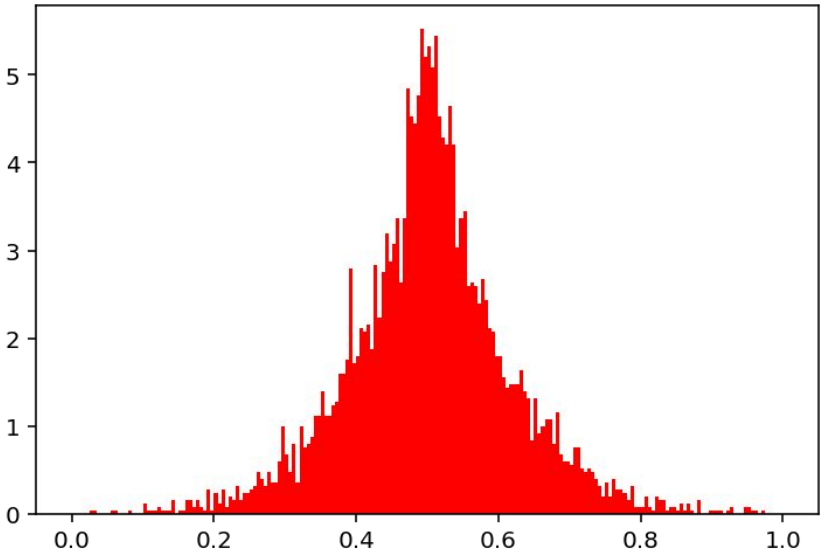


low

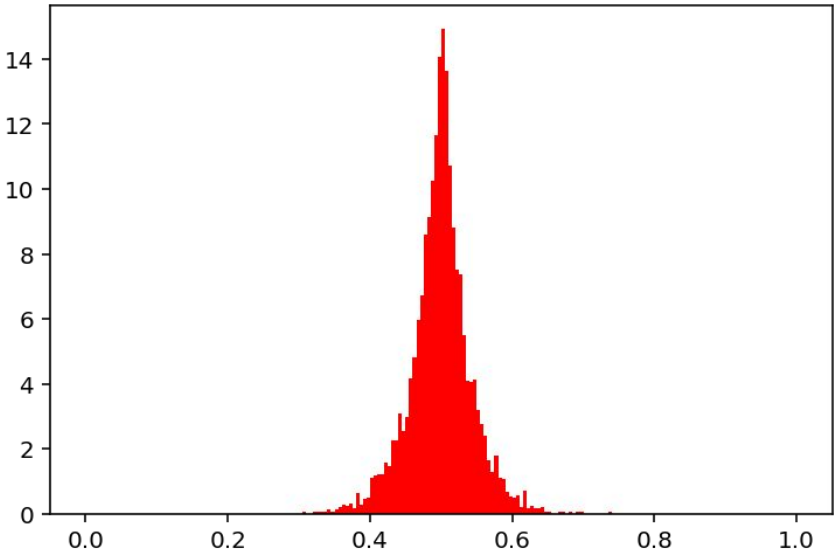


high

mutation parameter



low



high