ADL HW4

生醫電資所 碩一 r07945029 王思敏

Explain the structure of your networks and loss terms in detail:

Auxiliary Classifier GAN (ACGAN) is used to train the Cartoon set. Inside generator and discriminator are two Resnet with 2 layers and 4 layers separately. Each layer of generator consists of 2 blocks with sequential of convolutional layer, batch normalization layer, leaky ReLU layer, convolutional layer, batch normalization layer, and leaky ReLU layer in each block. On the other hand, each layer of discriminator consists of 2 blocks with sequential of convolutional layer, batch normalization layer, leaky ReLU layer, and dropout layer in each block.

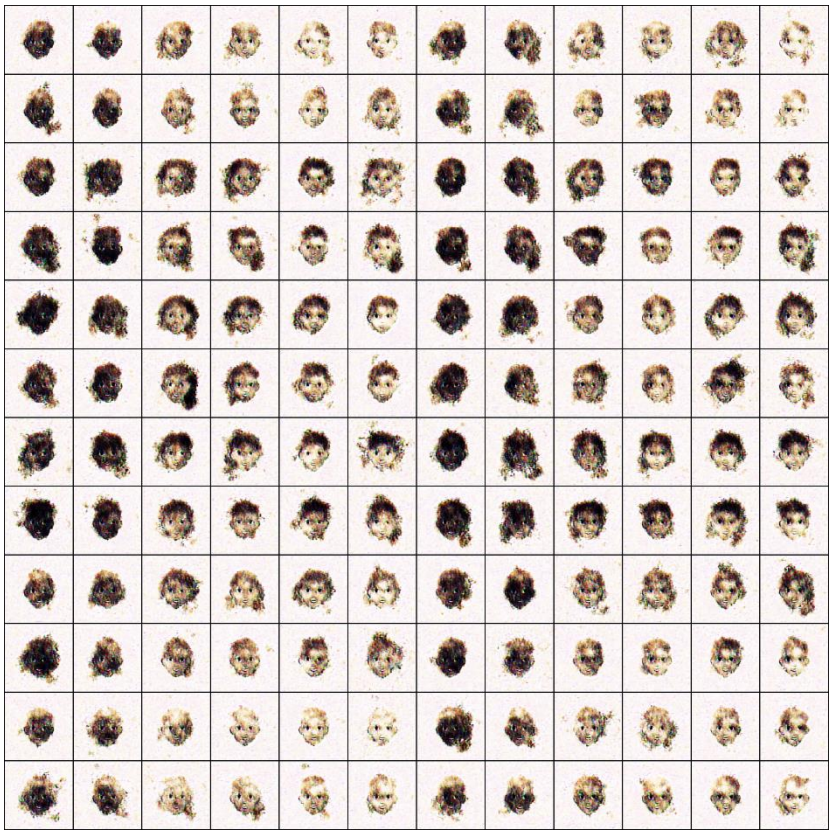
In generator, input noise will be fed into a linear layer first. Afterward, the output of the linear layer will be concatenated with the input label and fed into the Resnet of generator. The output will eventually be fed to a convolutional layer and activated by Tanh().

In discriminator, input image will be fed into the Resnet of generator. The output will then be fed into adversarial layer and auxiliary layer for validity and label.

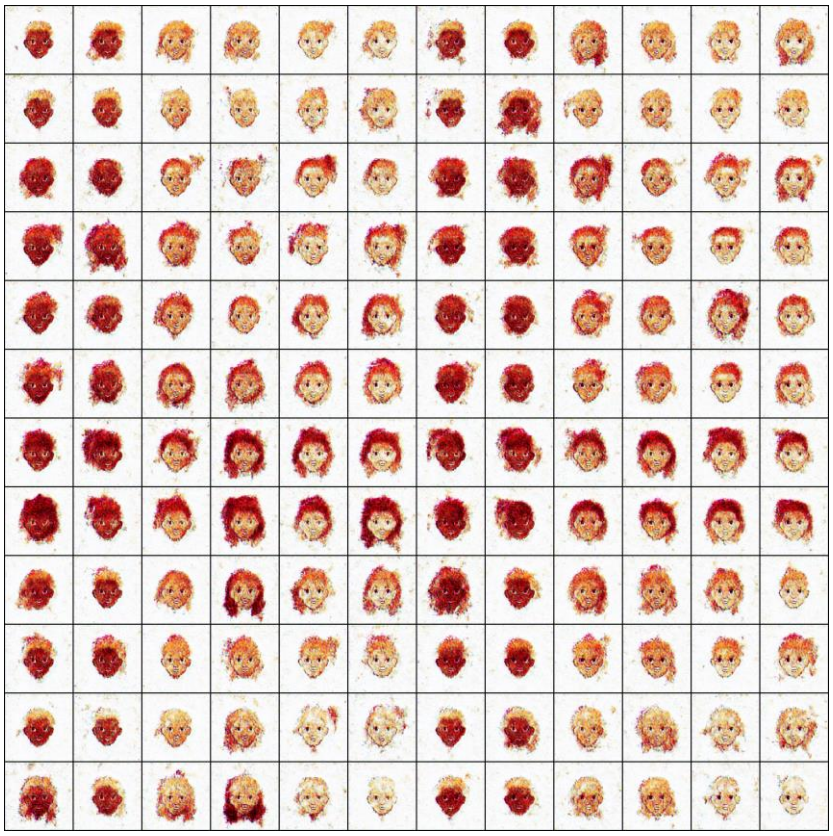
Loss items of adversarial and auxiliary are both BCELoss(). Loss of Generator is adversarial loss plus auxiliary loss over 2, and loss of Discriminator is loss of fake images plus loss of real images over 2. Loss of real and fake images are also computed by (adversarial loss + auxiliary loss) / 2.

Plot your training progress (10 pics):

Epoch0:



Epoch2:



Epoch5:

Epoch10:

Epoch20:

Epoch50:

Epoch100:

Epoch200:

Epoch500:

Epoch750:

Design at least 3 different experiments. Describe your settings, making comparisons and report your observations:

Experiments of 5 different setting are done to compare with the original ACGAN model which hit FID score of 222.26 after 1000 epochs of training (result of the original setting is not shown in the figure).

The first setting is changing the random normal samples of noise to random binomial. The outcome is better than the original setting in FID while causing serious mode collapse.

The second setting is feeding the labels into a linear layer before concatenating with the output noise. Also, the outcome is better than the original setting in FID while causing serious mode collapse.

The third setting is changing the generator to Resnet, and the forth setting is changing the generator and discriminator to Resnet. The detail of the structures of the generator and discriminator are described in the first question. The outcomes show the significant improvement comparing with the original setting. Also, changing both of generator and discriminator to Resnet shows better result than only generator.

The fifth setting is changing the input label from 0 and 1 of 15 classes to index of 1 in four main classes, such as hair, face, etc. The four main classes contain 6, 4, 3, and 2 sub-classes separately. The index of 1 of the four classes will be concatenated together and fed into the Resnet generator. The auxiliary loss is changed to cross entropy loss. The outcome show no significant different from the forth setting.

The sixth setting is the same structure and input label as the fifth setting while changing the BCE loss to Wasserstein loss, i.e. WGAN. The outcome implies that BCE loss for adversarial layer is still better than Wasserstein loss in this model structure.