Machine Learning
HW8

R10921104 王怡堯

May 9, 2022

1 Make a brief introduction about variational autoencoder (VAE). List one advantage comparing with vanilla autoencoder and one problem of VAE.

One advantage of variational autoencode is that we can control the distribution of the latent representation. One problem of variational autoencode is that the reconstruct image might be blurry.

2 Train a fully connected autoencoder and adjust at least two different element of the latent representation. Show your model architecture, plot out the original image, the reconstructed images for each adjustment and describe the differences.

```
fcn_multi_encoder(
  (encoder1): Sequential(
    (0): Linear(in_features=12288, out_features=3000, bias=True)
    (1): BatchNorm1d(3000, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (2): ReLU()
    (3): Linear(in_features=3000, out_features=3000, bias=True)
    (4): BatchNorm1d(3000, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (5): ReLU()
    (6): Linear(in_features=3000, out_features=1000, bias=True)
    (7): BatchNorm1d(1000, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (8): ReLU()
    (9): Linear(in_features=1000, out_features=1, bias=True)
  )
  (encoder2): Sequential(
    (0): Linear(in_features=12288, out_features=1000, bias=True)
    (1): BatchNorm1d(1000, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (2): ReLU()
    (3): Linear(in_features=1000, out_features=100, bias=True)
    (4): BatchNorm1d(100, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (5): ReLU()
    (6): Linear(in_features=100, out_features=100, bias=True)
    (7): BatchNorm1d(100, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (8): ReLU()
    (9): Linear(in_features=100, out_features=100, bias=True)
    (10): BatchNorm1d(100, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (11): ReLU()
    (12): Linear(in_features=100, out_features=1, bias=True)
  )
  (encoder3): Sequential(
    (0): Linear(in_features=12288, out_features=2500, bias=True)
    (1): BatchNorm1d(2500, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (2): ReLU()
    (3): Linear(in_features=2500, out_features=1500, bias=True)
    (4): BatchNorm1d(1500, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (5): ReLU()
    (6): Linear(in_features=1500, out_features=500, bias=True)
    (7): BatchNorm1d(500, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (8): ReLU()
    (9): Linear(in_features=500, out_features=100, bias=True)
    (10): BatchNorm1d(100, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (11): ReLU()
    (12): Linear(in_features=100, out_features=1, bias=True)
  )
  (decoder): Sequential(
    (0): Linear(in_features=3, out_features=100, bias=True)
    (1): ReLU()
    (2): Linear(in_features=100, out_features=500, bias=True)
```
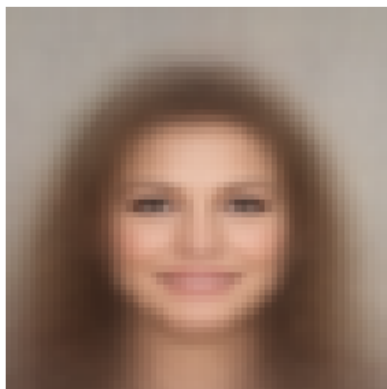
```
48      (3): ReLU()
49      (4): Linear(in_features=500, out_features=1500, bias=True)
50      (5): ReLU()
51      (6): Linear(in_features=1500, out_features=2500, bias=True)
52      (7): ReLU()
53      (8): Linear(in_features=2500, out_features=12288, bias=True)
54      (9): Tanh()
55    )
56  )
```
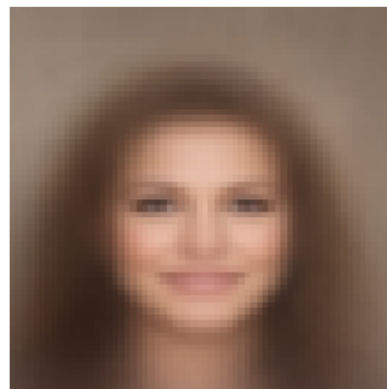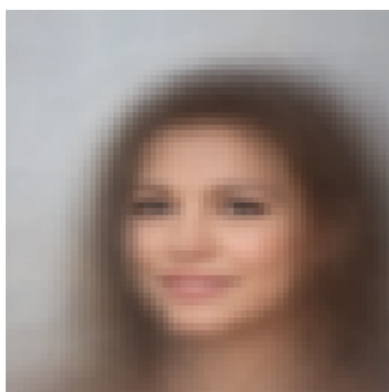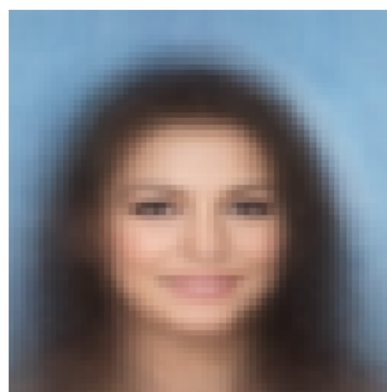

original image   econstructed images   first element * 3

second element * 3   third element * 3

We can observe that modifying the second element in the latent representation may let the human turn their head.