auto encoder

geo

2022-05-03

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

* autoencoder的改進版
* encoder會產生和m然後generate一組normal distribution random , 使code滿足 式子作爲添加noise后的code,這組code使得圖片在給定範圍内會產出跟encoder一樣的圖片 , 不會有產出奇特的圖片。
* 好處是根據noise的產出的code具有一定的模仿能力可以產生介於滿月和半月之間的圖片 , 不會有奇異的圖片。
* 壞處是只會針對data base裏面的image 做linear combination 沒有辦法學習一個新圖片的能力 ，只有模仿能力沒有學習能力。

1. 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.

Model architecture

# show original

    img = data.float().cuda()

    img = img.view(img.shape[0],-1)

    img = img.reshape(200,3,64,64)

    grid\_img = torchvision.utils.make\_grid(img.cpu(), nrow=5)

    x , y , w , h = 0 , 0 ,64 ,64

    crop = grid\_img[y:y+h,x:x+w]

    plt.figure(figsize=(10,10))

    plt.imshow(crop.permute(1, 2, 0).cpu())

    # element 1

    img = data.float().cuda()

    img = img.view(img.shape[0],-1)

    output = model.encoder(img)

    output[:,0] = output[:,0] \* 2

    output = model.decoder(output)

    output = output.reshape(200,3,64,64)

    output = (output+1)/2

    grid\_img = torchvision.utils.make\_grid(output.cpu(), nrow=5)

    plt.figure(figsize=(10,10))

    plt.imshow(grid\_img.permute(1, 2, 0).cpu())

    x , y , w , h = 0 , 0 ,64 ,64

    crop = grid\_img[y:y+h,x:x+w]

    plt.imshow(crop.permute(1, 2, 0).cpu())

    plt.show()

    # element 2

    img = data.float().cuda()

    img = img.view(img.shape[0],-1)

    output = model.encoder(img)

    output = output \* 4

    output = model.decoder(output)

    output = output.reshape(200,3,64,64)

    output = (output+1)/2

    grid\_img = torchvision.utils.make\_grid(output.cpu(), nrow=5)

    plt.figure(figsize=(10,10))

    plt.imshow(grid\_img.permute(1, 2, 0).cpu())

    x , y , w , h = 0 , 0 ,64 ,64

    crop = grid\_img[y:y+h,x:x+w]

    plt.imshow(crop.permute(1, 2, 0).cpu())

    plt.show()

A collage of a person

Description automatically generated with low confidence

第一row 是原始圖片 沒有經過 Encoder 的原始圖 , 第2 row 和 第3 row 分別是針對tensor第一column做2倍處理以及對整個tensor 乘于4倍 。

他們的特徵都有被捕獲到,但是並沒有做到很完善 。

第 2 row特徵比第3 row特徵更明確 。

他們主要差別在於第三row的第二張和第四張特徵的捕獲比第二row來的好 ，其他第2 row 較好 。