

神经网络 第四次作业

2017012289 李岱轩

1 实验概述

本次实验中，我根据要求补全了VAE和GAN的代码，并调整了部分其他超参数（如latent dim，learning rate与batch size）进行了测试。测试的结果与讨论如下。

2 问题回答

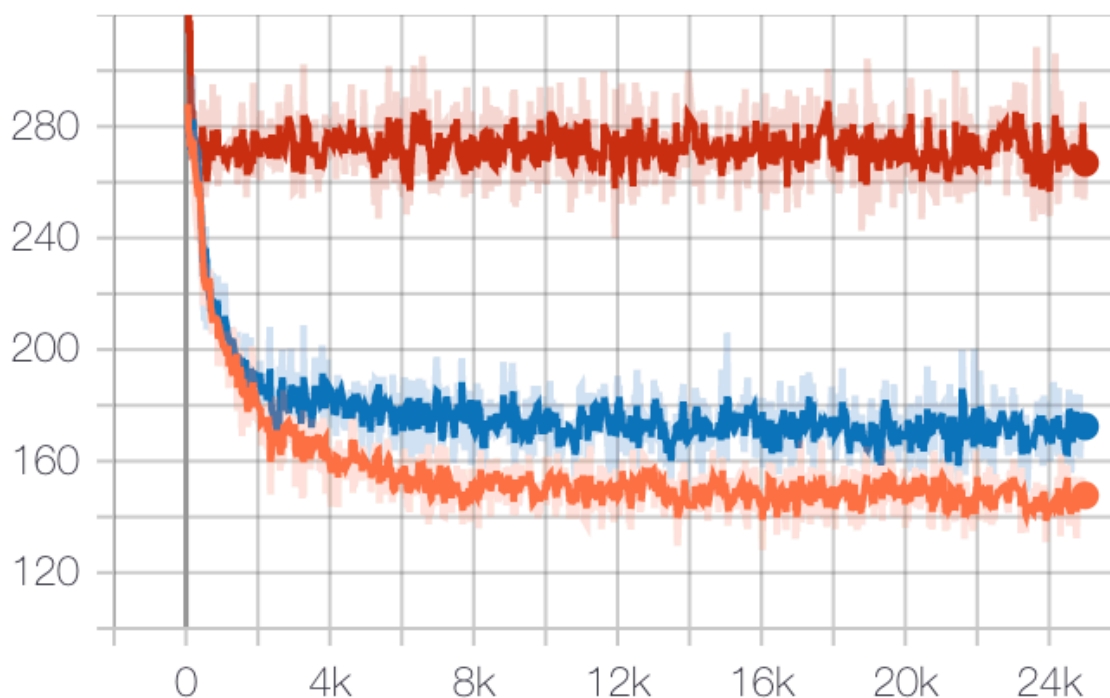
2.1 对latent dim为3，10，100的VAE网络进行训练，并给出训练集，验证集Loss

本实验中采取的超参数集合为：

batch_size	learning rate	num_training_step
32	0.001	25000

latent dim分别为3,10,100时，对应红蓝橙三种颜色的线，训练集loss为：

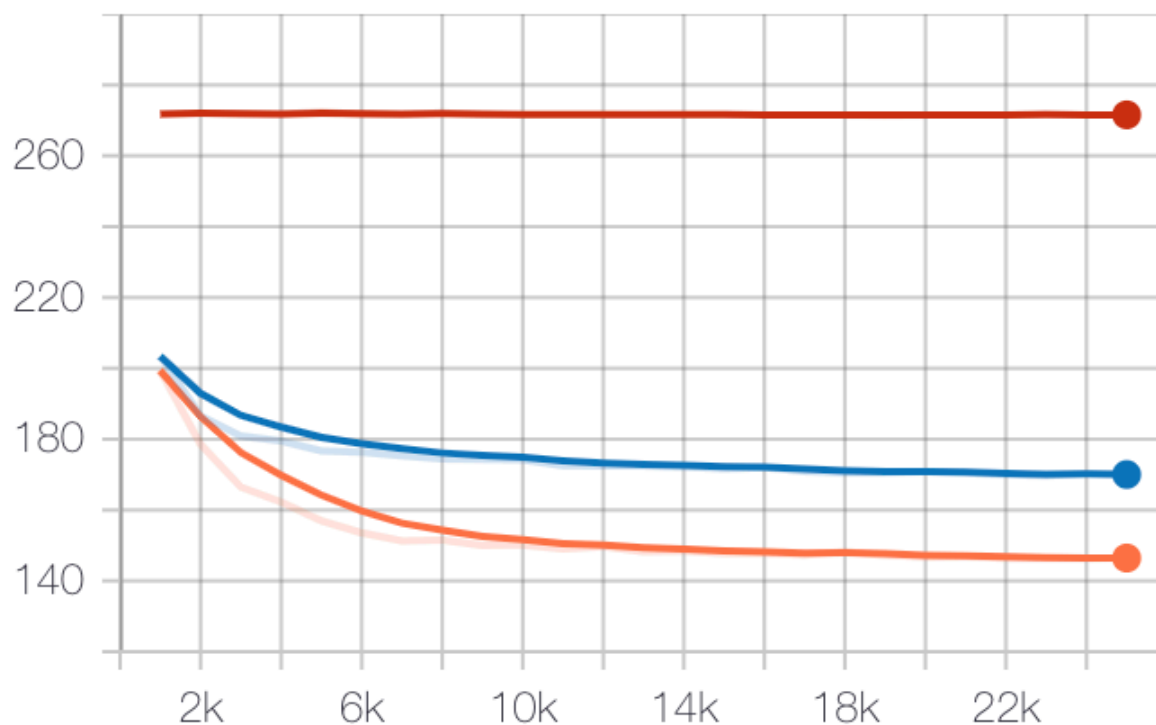
loss



验证集loss为：

loss
tag: dev/loss

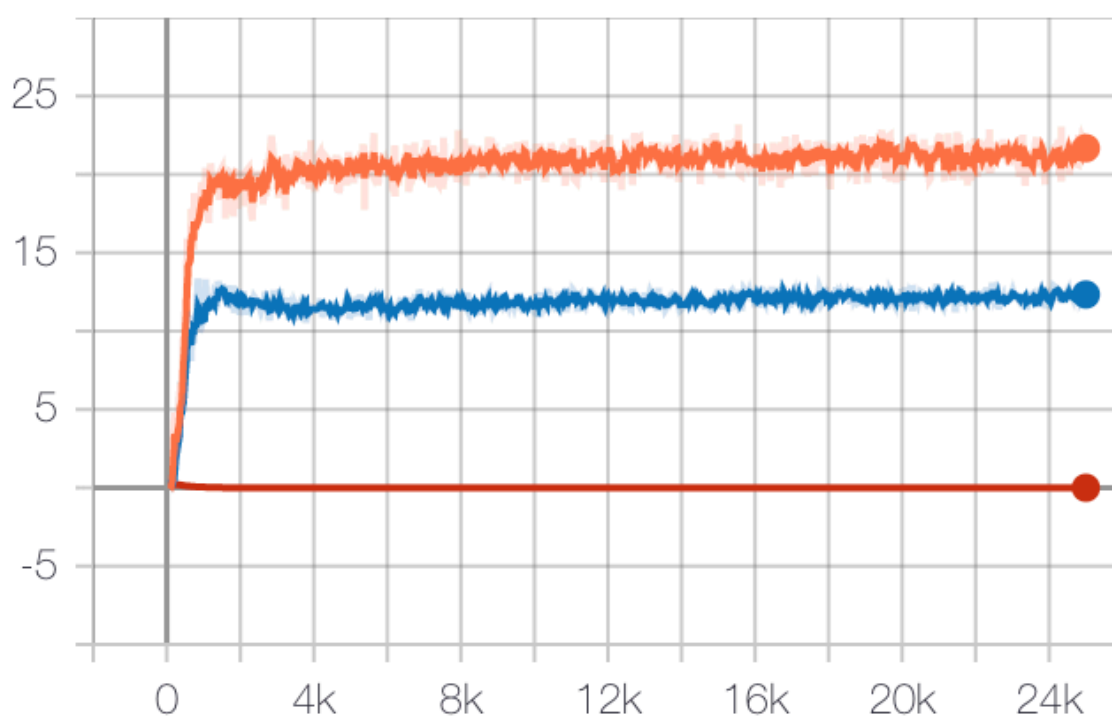
r
t



可以看到，dim=3时训练集loss仅下降很少就维持稳定，验证集loss几乎不变。dim=10,100时，训练集与验证集loss同步下降并趋于稳定。相比之下，dim=100时验证集与训练集loss的下降和最终收敛的值均更低。

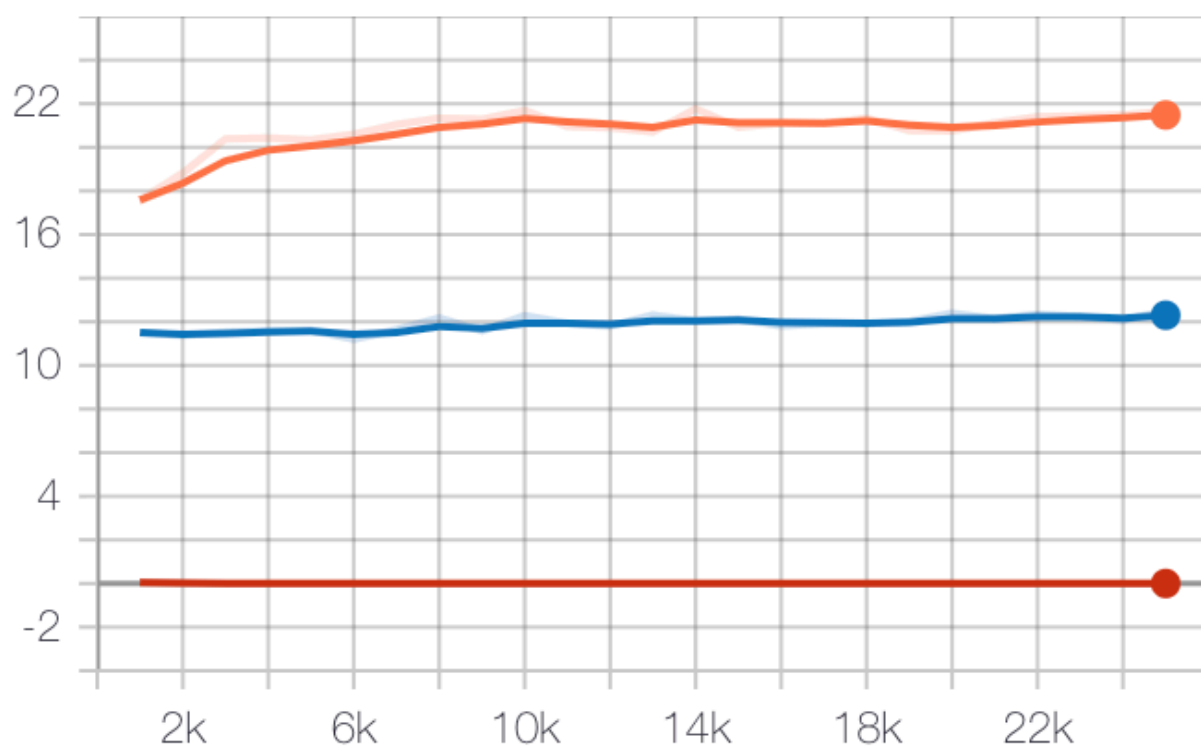
训练集kl散度为：

KL_divergence



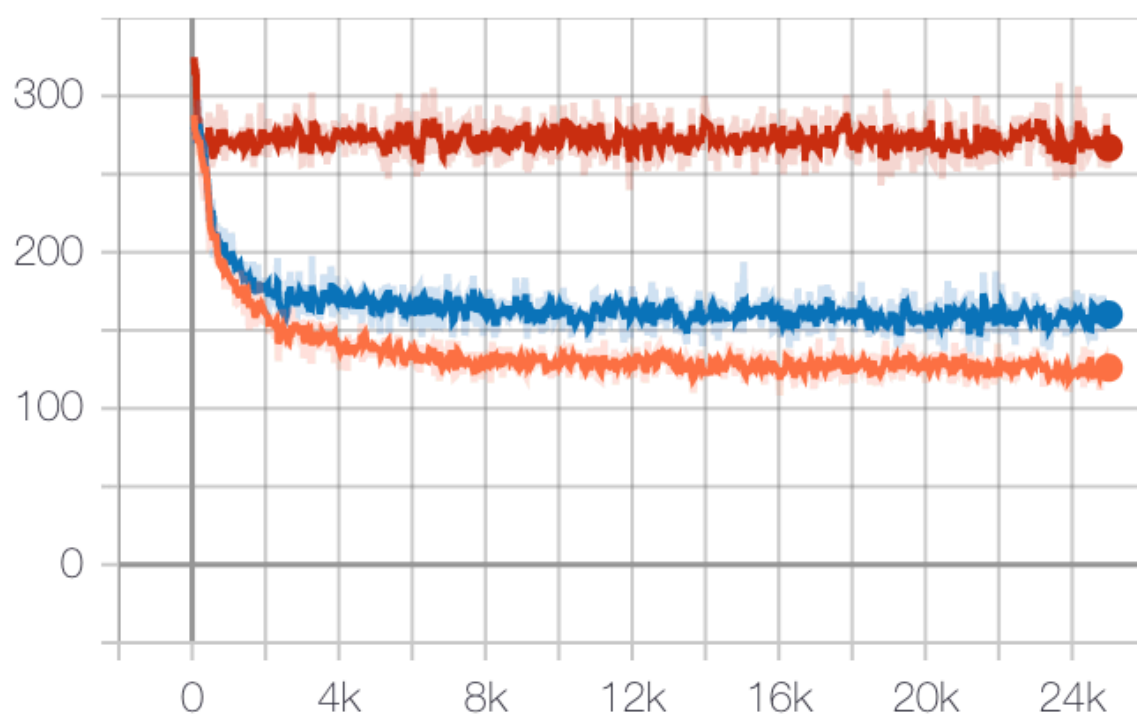
验证集kl散度为：

KL_divergence
tag: dev/KL_divergence



训练集重构误差为：

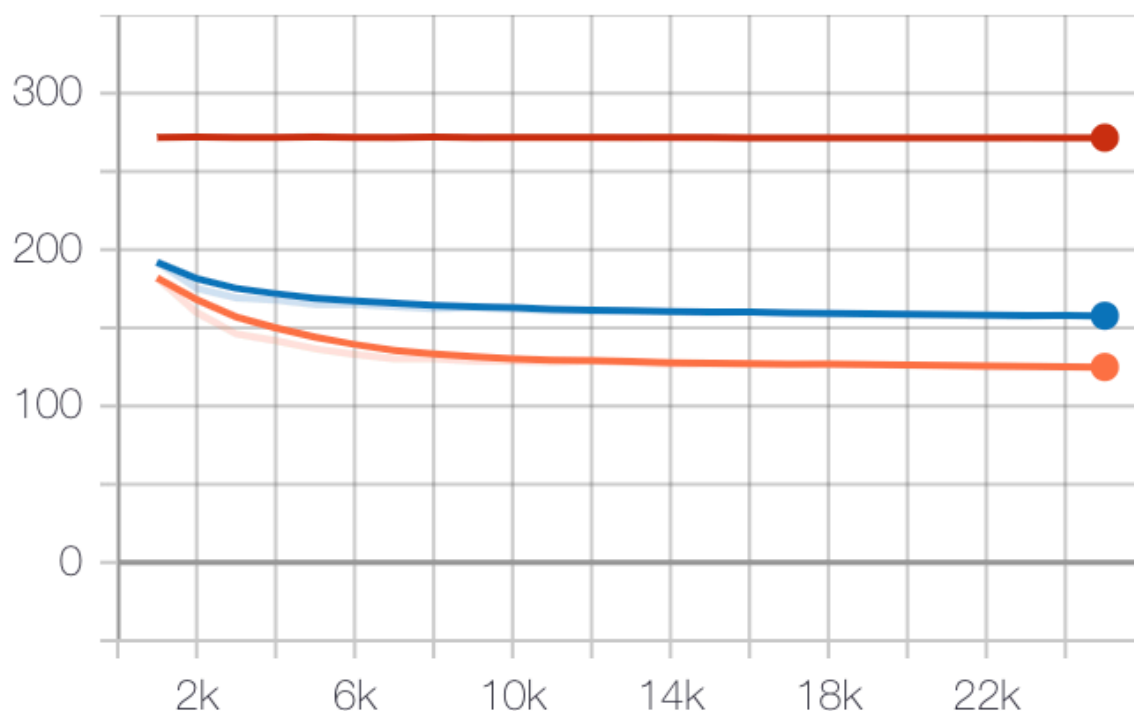
reconstruction_loss



验证集重构误差为：

reconstruction_loss

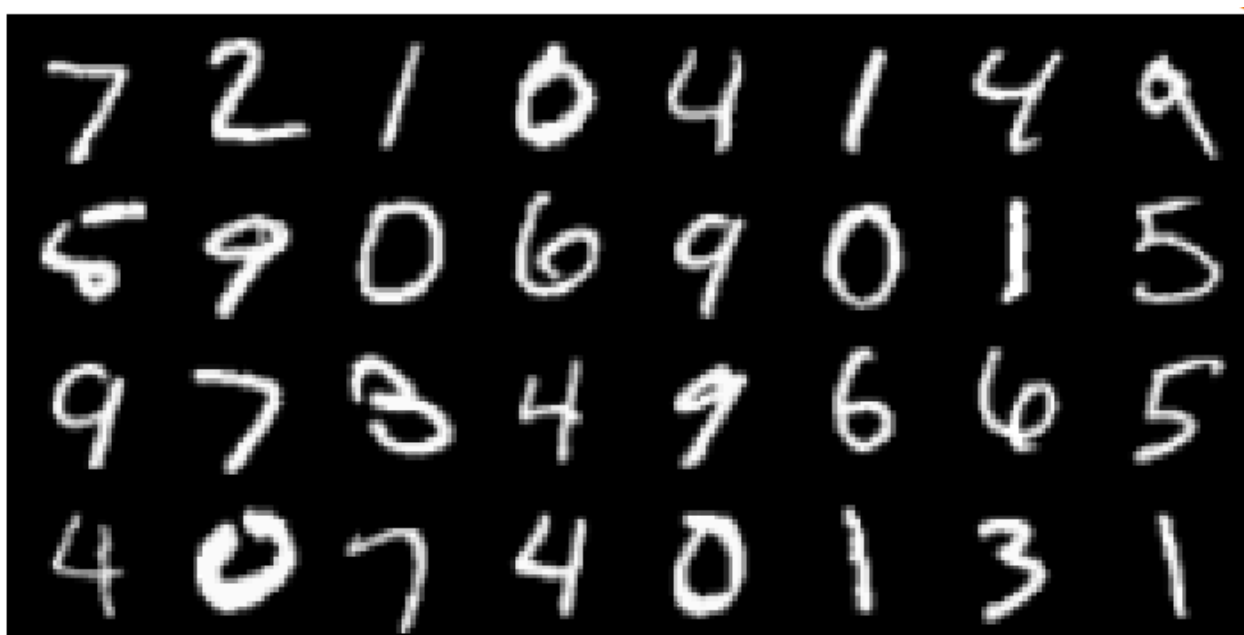
tag: dev/reconstruction_loss



VAE的训练是在重构误差的下降与kl散度的上升之间找到平衡，使得在最终的生成图像上loss最小。可以看到，dim=3时kl散度与重构误差几乎不变，说明完全没有起到训练的效果。dim=10,100时，在训练和验证集上随着kl散度上升，重构误差逐渐下降，最后趋于收敛。相比之下dim=100时kl散度更大，重构误差更小，说明算法寻找平衡点更加aggressive。然而对于本任务（图像生成），我们还需要比较真实图像来得到最终哪一个参数更优秀的结论。

2.2 对latent dim为3，10，100的VAE网络比较验证集重构与生成的图像

验证集图像均为：



dim=3,10,100时的重构图像分别如下（依次从左至右）：



dim=3,10,100时的随机生成图像如下所示（依次从左至右）：



可以看到，无论是重构还是生成，dim=3都无法给出有意义的结果。而dim=10,100都可以看出来重构和生成的数字。相比之下，dim=100重构或者生成的数字更加清晰和准确，如重构中第二行第四个的6，在dim=10中这部分不够清楚，而在dim=100中很清楚，又如生成图中dim=10的多个数字（如第四行第六个）完全看不出是什么数。而dim=100只有个别数字不清楚（如第三行第五个）。

2.3 对latent dim为3，10，100的VAE网络比较FID

三个网络的FID如下：

dim	FID
3	320.686
10	46.27
100	28.83

可以看到，dim=3,10,100，fid逐渐减小，表明效果变得更好。

2.4 总结latent_dim的影响

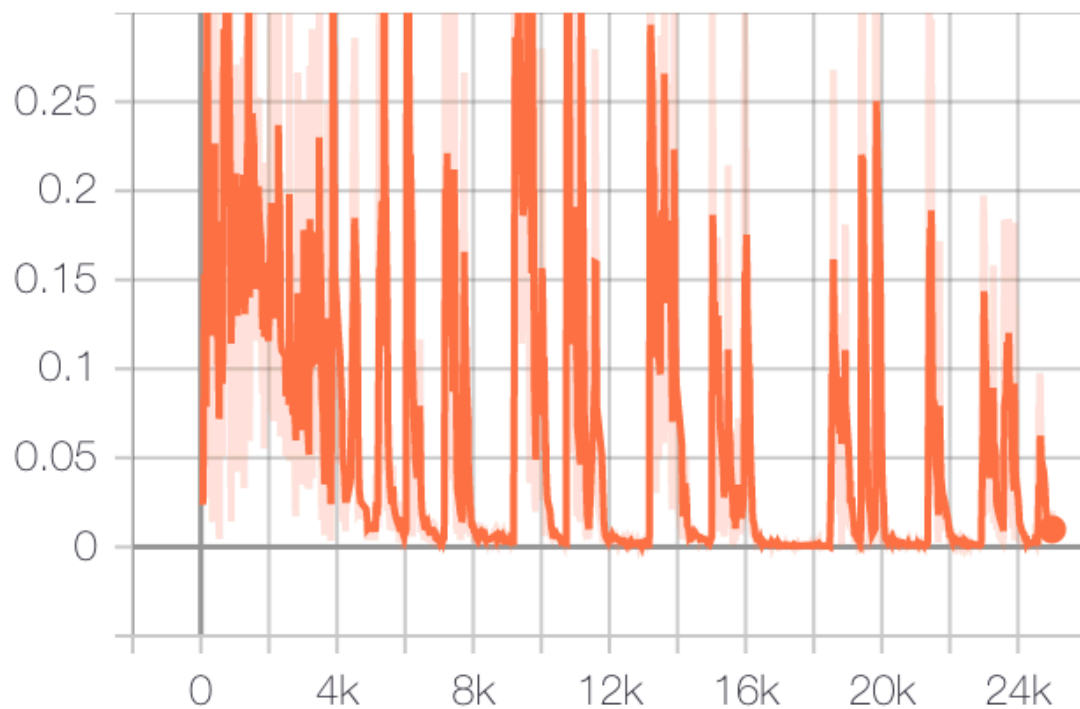
总的来说，随着latent dim增大，模型重建和生成图的效果逐渐变好，经过实验，latent dim到达一定大笑之后在增大则效果不是很明显（默认参数64的训练fid与100接近）。

分析其原因，很小的latent dim（如3）无法捕捉生成图像的信息，效果很差。随着dim增大，可以捕捉到的信息变多，生成和重建图的准确性和清晰度也变高，同时评价指数fid也变得更高。但是增大到一定程度后，可捕捉的信息饱和，再增大dim效果不是很明显。

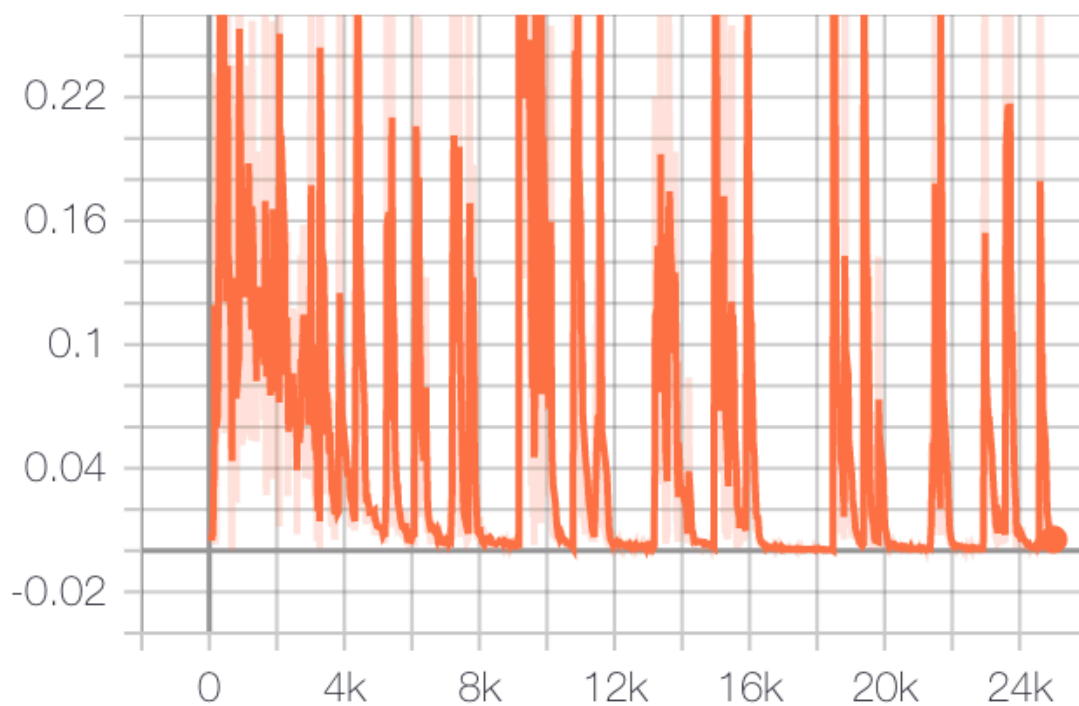
2.5 GAN训练曲线与FID

将隐层大小参数设置为100，generator隐层大小设置为64，其他参数采用默认参数，得到训练曲线如下所示：

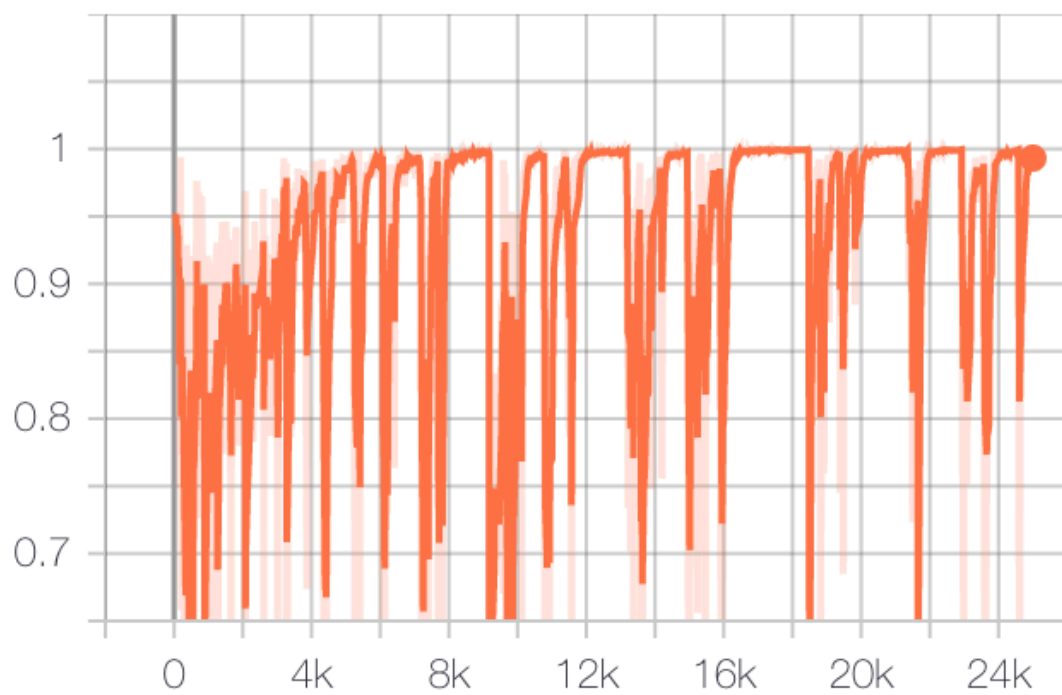
D_G_z1__



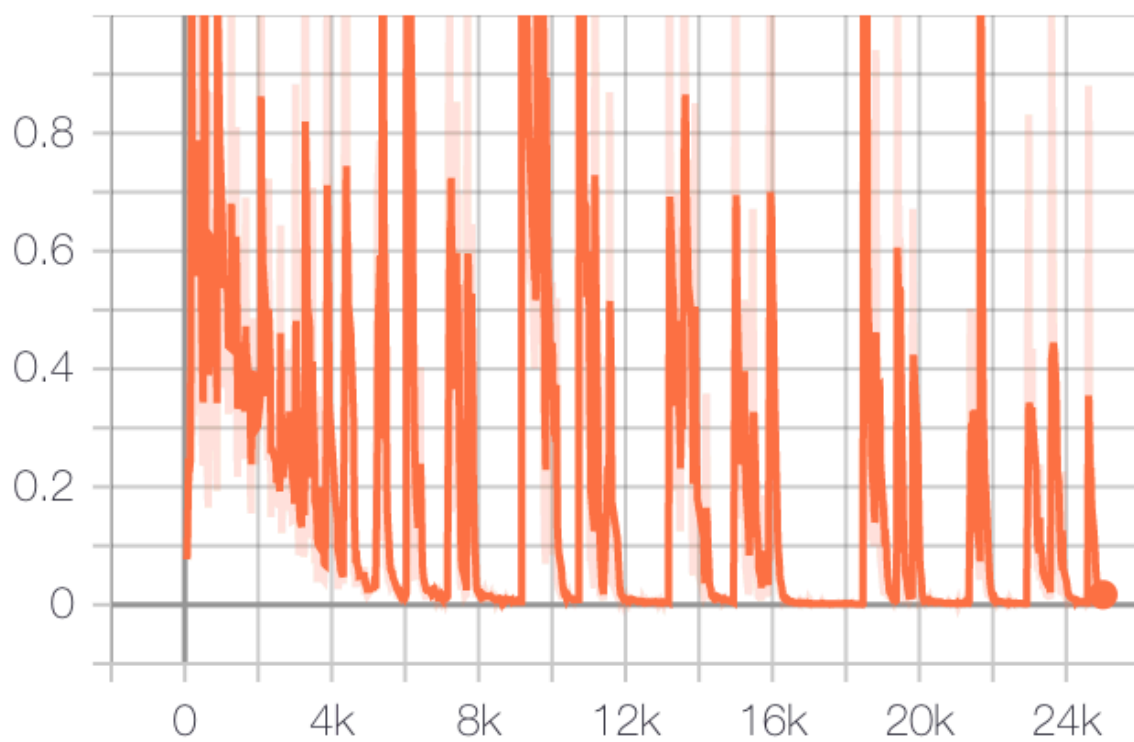
D_G_z2__



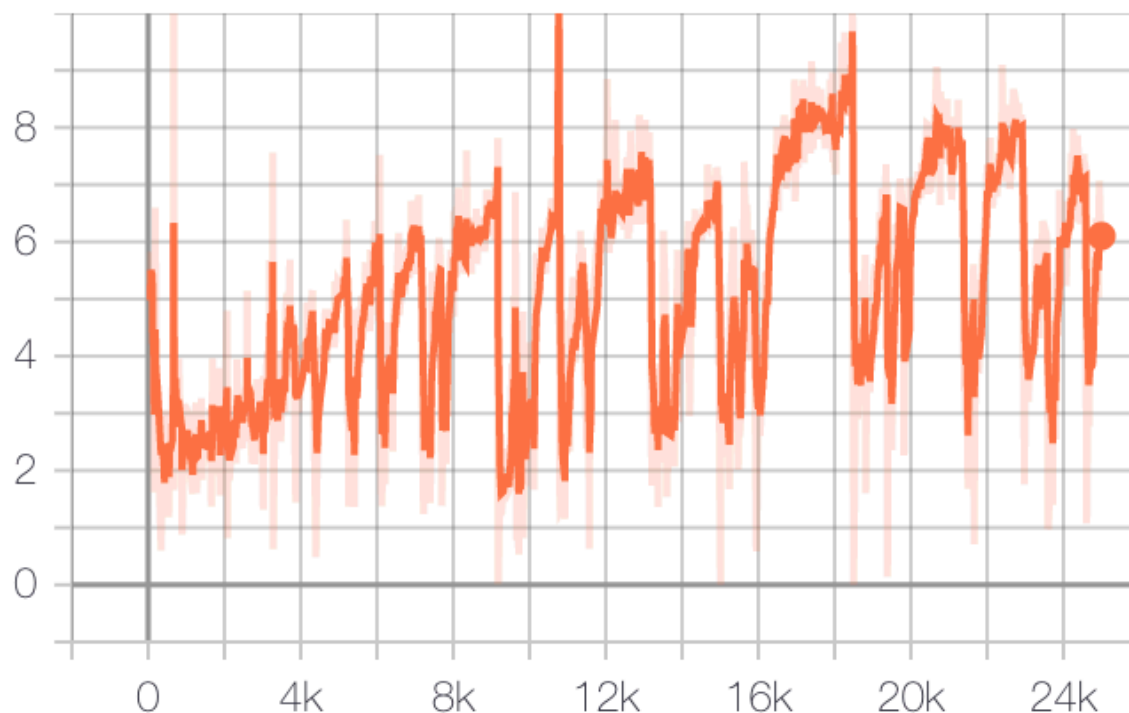
D_x_



discriminator_loss



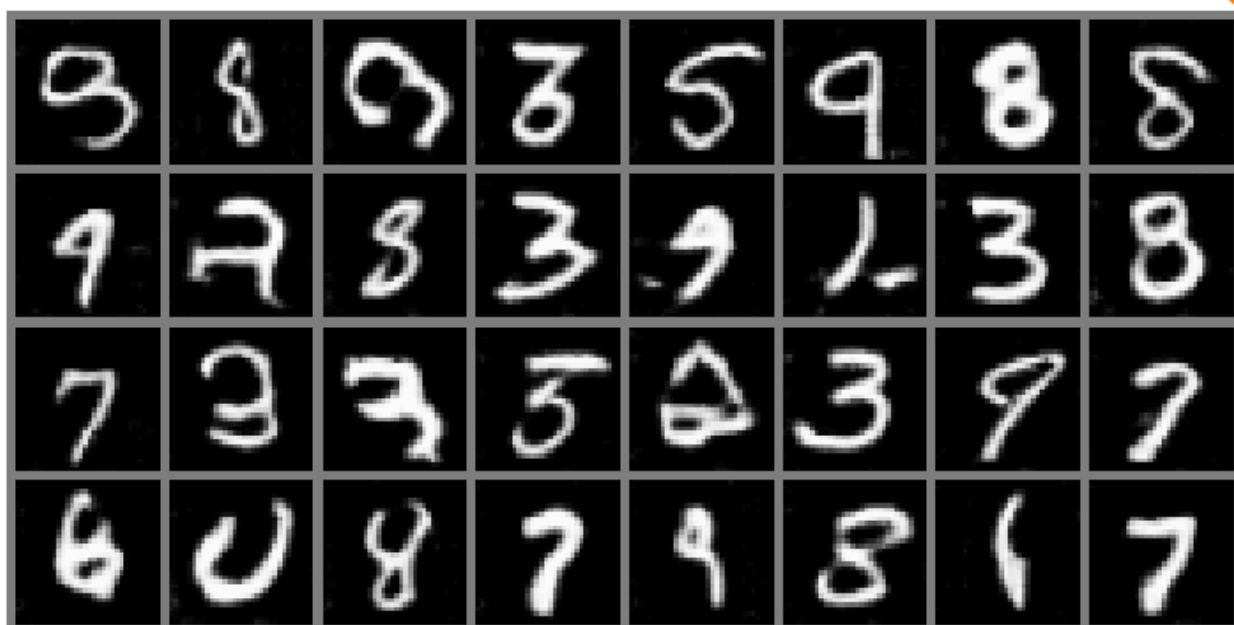
generator_loss



可以看到，discriminator和generator的损失函数波动都非常大。分析其原因，是GAN网络在不断通过平衡G和D两个网络来达到最好的生成效果，因此训练中两个网络相互竞争波动加大。同样，我们可以通过接过来判定是否GAN是有效的。

FID的值为45.02，满足题目要求。

2.6 GAN生成图像，以及它与VAE的对比



GAN的生成图像如上所示，其中有一些可以明显看出是数字，另一些则意义不明。总体情况类似dim=10的VAE。我认为原因是GAN网络需要更复杂的模型调参来得到性能最大化。GAN在对抗生成中波动过大，应该继续调整模型参数来进一步优化GAN。

