

Report for HW3

数据集的选择

LFW (Labeled Faces in the Wild, 野外标记人脸) 是一个人脸照片数据库，旨在研究非限制人脸识别问题。该数据集包含了从网上收集的13,000多张人脸图片。

我们小组选择LFW-Funneled作为数据集经过多方面的考虑。

多样性：LFW-Funneled数据集包含了超过13,000张人脸图像，这些图像来自网上公开可用的照片。每个人可能有一张或多张照片。这些图像涵盖了各种光照、表情、姿态和年龄等变化，这为训练模型提供了丰富的多样性。

预处理：LFW-Funneled数据集中的图像已经通过商业人脸对齐软件进行了对齐。这意味着所有的人脸都已经被标准化，以便于进行人脸识别的研究。这种预处理可以减少训练模型时的复杂性，并提高模型的性能。

实用性：由于LFW-Funneled数据集中的图像来自真实世界的场景，因此在这个数据集上训练的模型更有可能在实际应用中表现良好。

挑战性：尽管LFW-Funneled数据集已经进行了预处理，但由于图像的多样性，使用这个数据集训练模型仍然具有一定的挑战性。这对于提高模型的鲁棒性和泛化能力是非常有益的。

模型训练过程

GAN模型有很多参数如batch_size ,image_size ,nc ,nz ,ngf ,ndf ,num_epochs ,lr ,beta1, ngpu。nz是生成器输入的大小，也就是潜在向量的大小。这个参数对生成图像的多样性有很大影响，因为它决定了我们可以在潜在空间中探索的维度。如果这个值太小，我们可能无法捕捉到数据的所有变化；如果这个值太大，训练可能会变得非常困难。num_epochs 是训练的迭代次数。这个参数对模型的性能有直接影响，因为它决定了模型有多少机会学习和优化参数。然而，如果设置得过高，可能会导致模型过拟合。lr是优化器的学习率。这个参数对模型的训练速度和最终性能有很大影响。如果学习率太高，模型可能无法收敛；如果学习率太低，训练可能会非常慢。经过分析之后，我认为这三个参数会对模型造成比较大的影响。

首先要确定一个比较好的num_epochs,不能太小以至于训练轮数太少无论其他参数是多少都无法有较好的输出，也不能太大以至于训练一次耗时太长无法进行调参。我先试了epoch=5000的情况，发现效果较差。改成10000后效果图已经比较清晰。

然后开始调整nz与lr两个参数的选择，我尝试训练nz=100,150,200, lr=1e-4,5e-4,1e-3。

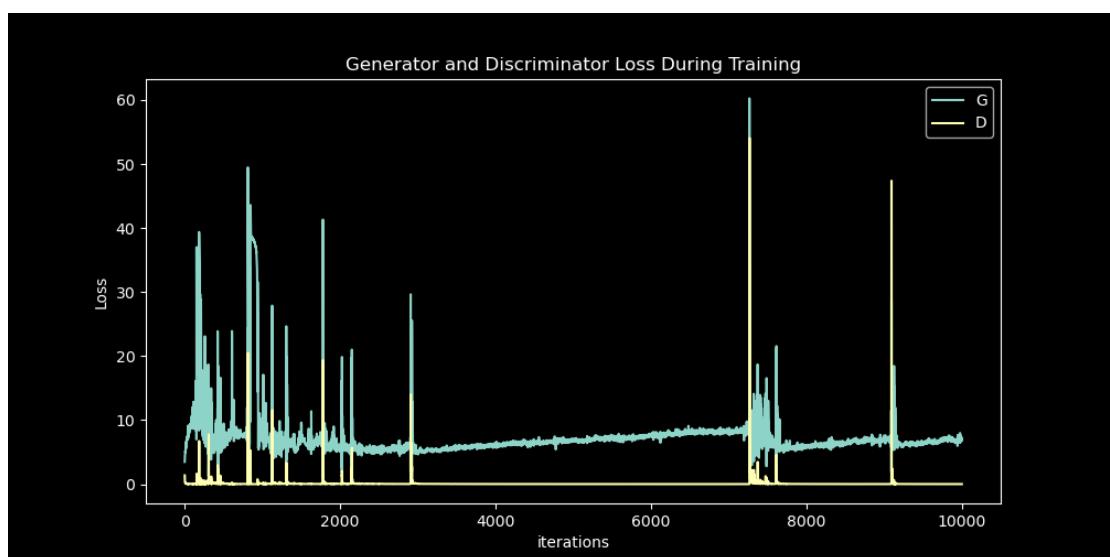
nz=100, lr=1e-3, num_epochs=10000:



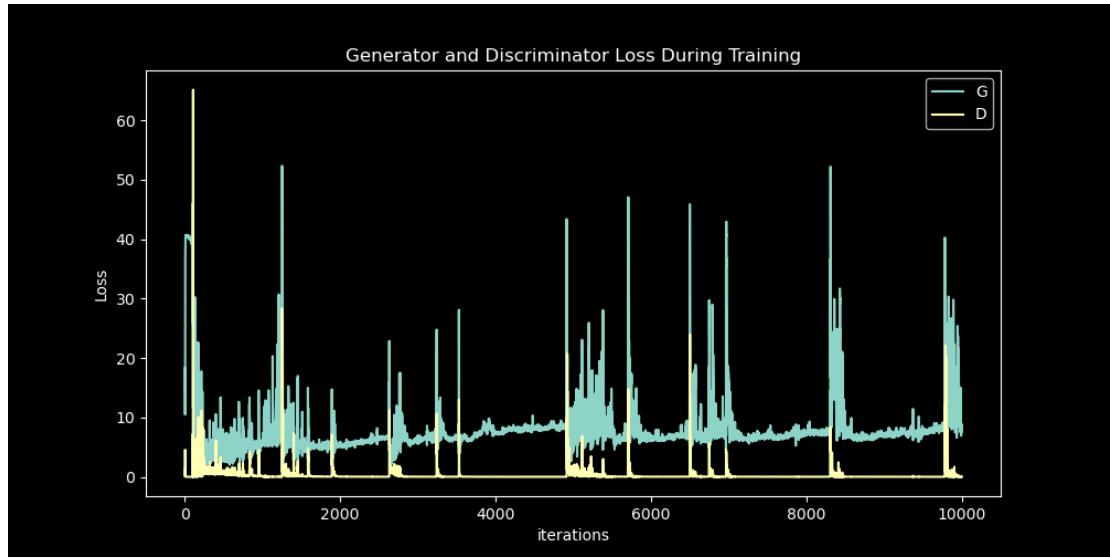
可以看出生成的图片完全无法看清，而lr更小的情况下生成的图片反而更好，说明lr过大，过拟合了。

而lr=5e-4和lr=1e-4的模型生成的图片肉眼看起来差不多，我选择从loss来进行比较。

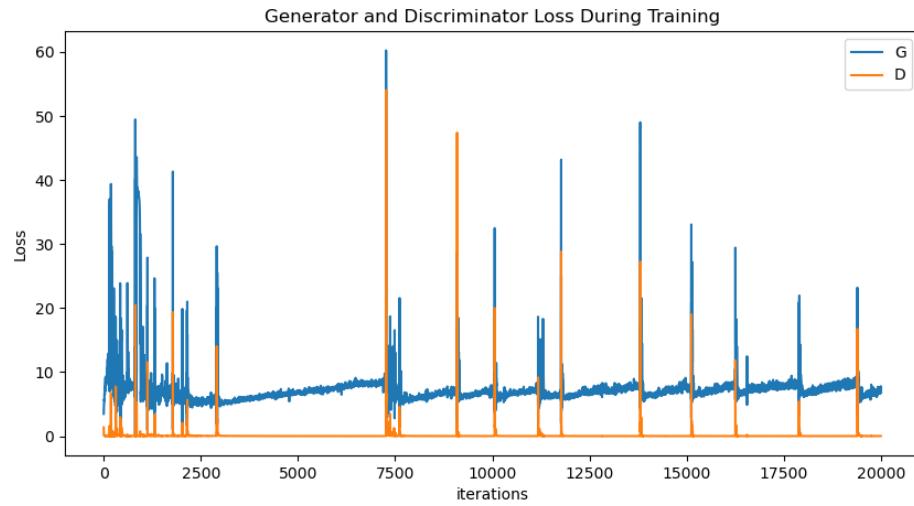
nz=100, lr=1e-4, num_epochs=10000:



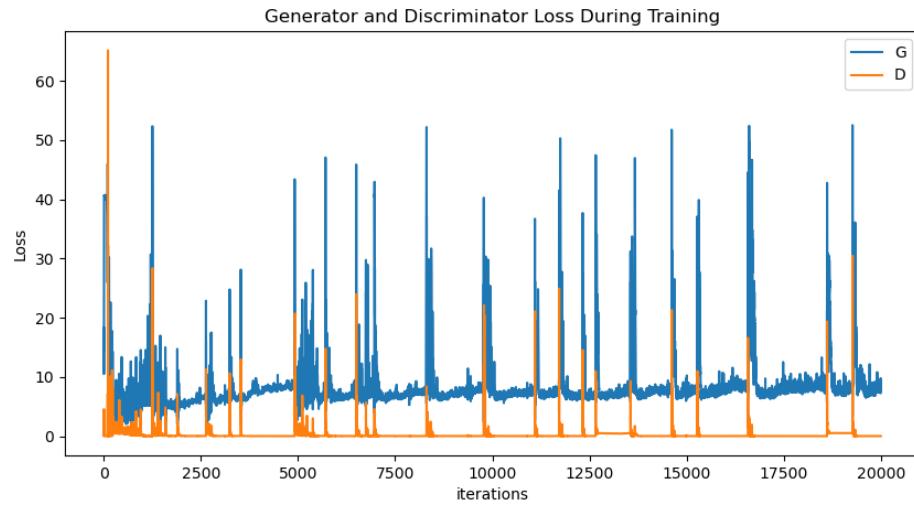
nz=100, lr=5e-4, num_epochs=10000:



$nz=100, lr=1e-4, num_epochs=20000:$



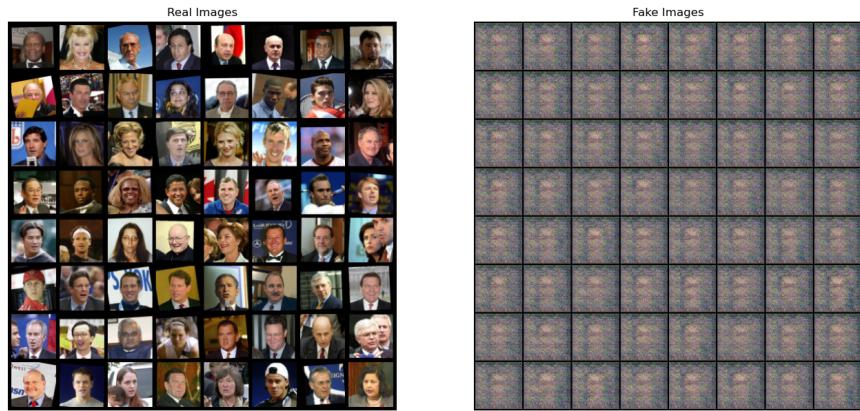
$nz=100, lr=5e-4, num_epochs=20000:$



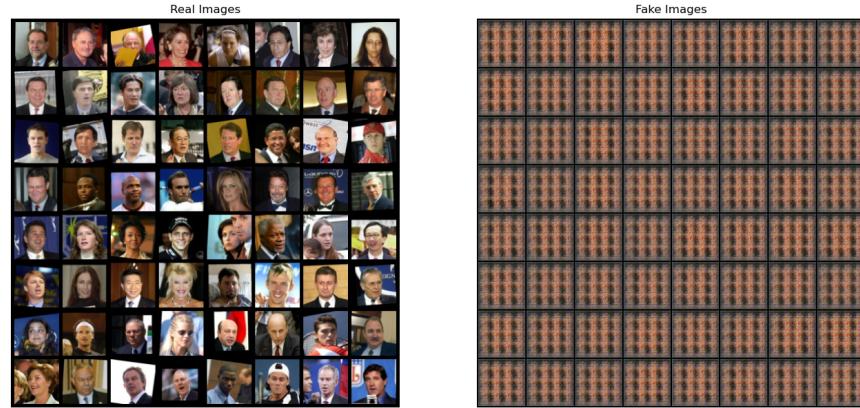
我们发现 $lr=5e-4$ 的loss曲线波动比 $lr=1e-4$ 的曲线波动更剧烈，说明 $lr=1e-4$ 的loss更稳定。

再考虑 nz 的取值

nz=150, lr=1e-4, num_epochs=10000:



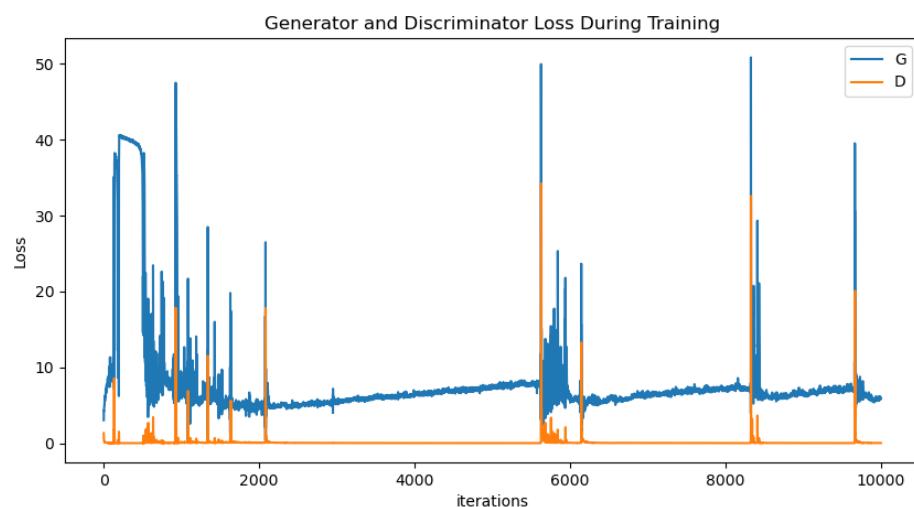
nz=200, lr=1e-4, num_epochs=10000:



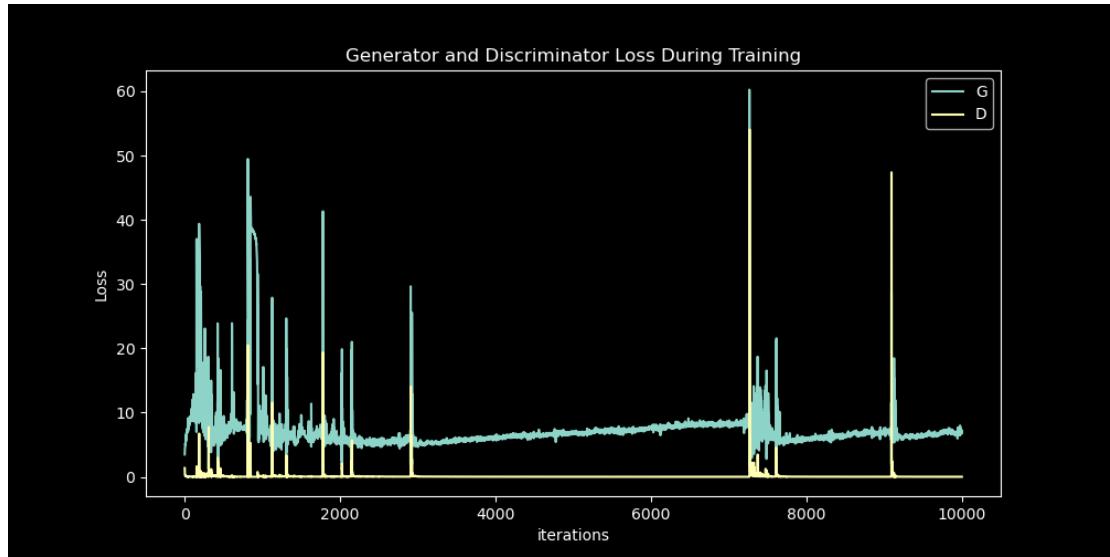
通过nz=150和nz=200生成的图片可以发现生成的图片完全看不出是一个人类，而在nz=100的情况下生产的图片较好，说明这两种情况潜在向量过大导致训练十分困难。

于是接下来我比较了nz=50与nz=100的情况。

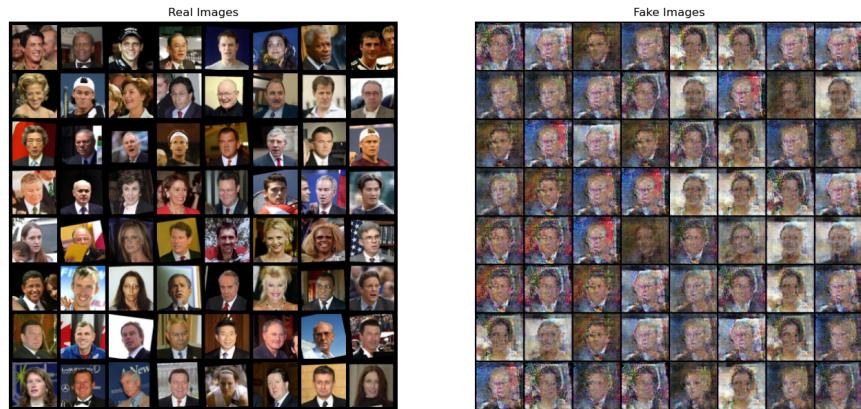
nz=50, lr=1e-4, num_epochs=10000:



nz=100, lr=1e-4, num_epochs=10000:



nz=50, lr=1e-4, num_epochs=10000:



nz=100, lr=1e-4, num_epochs=10000:

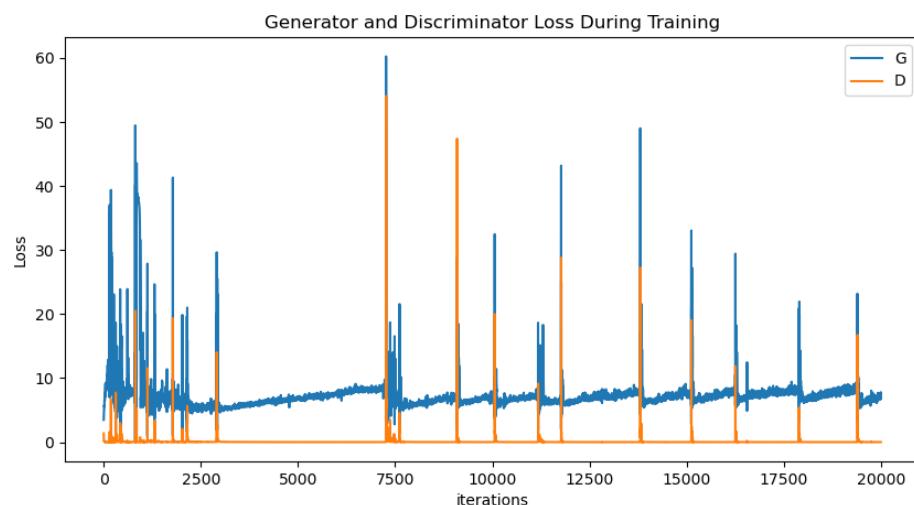


发现loss的情况相差不大，但是生成的图片来说， $nz=0.0001$ 的模型更好。

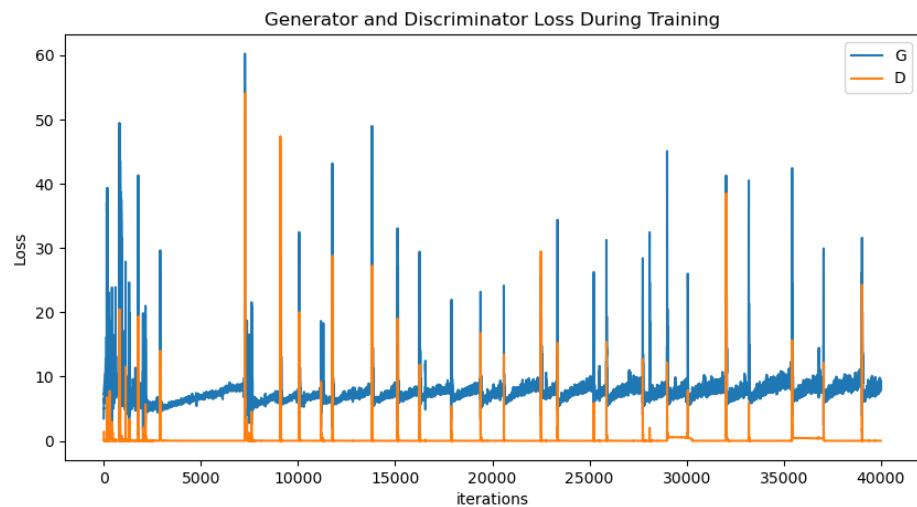
综上，我认为 $lr=1e-4$, $nz=100$ 的情况更优。

接下来我，训练了 $num_epochs=20000$, $num_epochs=40000$ 的模型。

$nz=100$, $lr=1e-4$, $num_epochs=20000$:



$nz=100$, $lr=1e-4$, $num_epochs=40000$:



从loss的角度来看并没有很大的改善。

$nz=100, lr=1e-4, num_epochs=20000$:



$nz=100, lr=1e-4, num_epochs=40000$:



从生成的图片来看20000的图片很模糊，而40000的模型中有三四张训练集的图片特征被多次训练，导致生成的图片相似度很高。

综合来看，我认为num_epochs=10000,lr=0.0001,nz=100的模型效果最好。

在任务之外，我们组还将共10000张图片一起生成了一个模型，发现图像生成还是需要更多的数据，当数据足够多的时候，模型效果相当好。

