



南開大學
Nankai University

南 開 大 學

计 算 机 学 院

深度学习及应用实验作业

作业三 生成对抗网络实践

姓名：徐宇昂

学号：2011742

年级：2020 级

专业：计算机科学与技术

指导教师：侯淇彬

2023 年 6 月

摘要

本次实验基于 FashionMNIST 数据集进行实践。

关键字：pytorch, GAN

目录

一、 实验要求	1
二、 GAN 网络结构	1
三、 loss 曲线	1
四、 随机数生成图	2
(一) 修改一	2
(二) 修改二	3
(三) 修改三	3

一、 实验要求

- 掌握 GAN 原理
- 学会使用 PyTorch 搭建 GAN 网络来训练 FashionMNIST 数据集

二、 GAN 网络结构

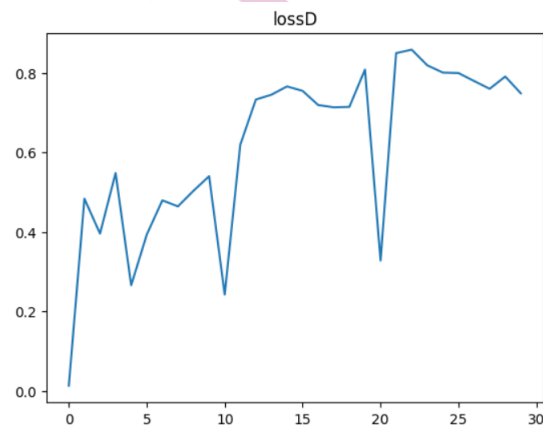
如图6所示

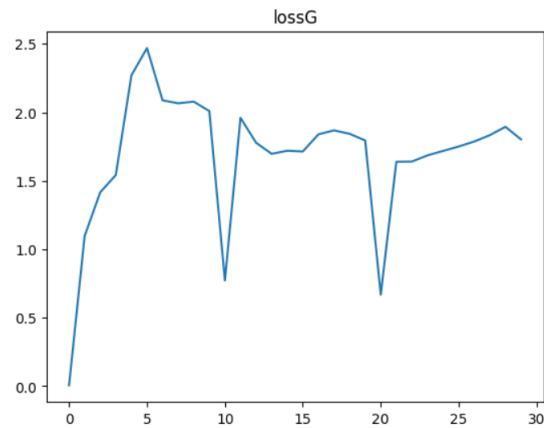
```
Discriminator(  
    (fc1): Linear(in_features=784, out_features=128, bias=True)  
    (nonlin1): LeakyReLU(negative_slope=0.2)  
    (fc2): Linear(in_features=128, out_features=1, bias=True)  
)  
Generator(  
    (fc1): Linear(in_features=100, out_features=128, bias=True)  
    (nonlin1): LeakyReLU(negative_slope=0.2)  
    (fc2): Linear(in_features=128, out_features=784, bias=True)  
)
```

图 1: Caption

三、 loss 曲线

如图所示，下图分别为分类器和生成器的 loss 曲线：





四、 随机数生成图

```

1 f = torch.randn(8,100,device=device)
2 x = G(f)
3 show_imgs(x)

```

生成结果如图6所示

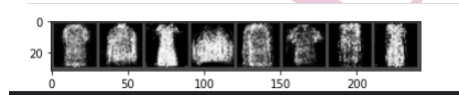


图 2: Caption

针对自定义的 100 个随机数，自由挑选 5 个随机数，查看调整每个随机数时，生成图像的变化（每个随机数调整 3 次，共生成 15x8 张图），总结调整每个随机数时，生成图像发生的变化。这里只需要针对我们生成的每一个 tensor 张量更改其中的随机数即可。

（一） 修改一

修改第二个随机数为 0, 10, 100，如下图（从上到下依次为未修改原图，0, 10, 100），生成结果如图6所示

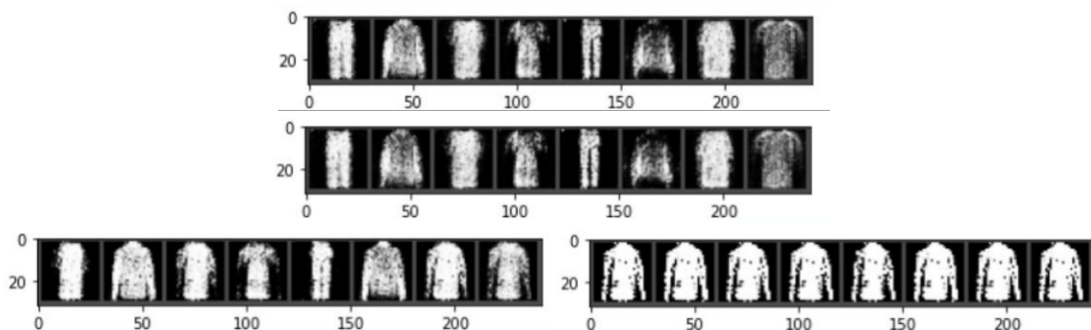


图 3: Caption

发现随着第二个随机数的增大，长袖上衣的图像越来越明显，说明第二个随机数与此类分布关系紧密。

(二) 修改二

修改第七十一个随机数为-10, 10, 100, 如下图 (从上到下依次为 -10, 原图, 10, 100), 生成结果如图6所示

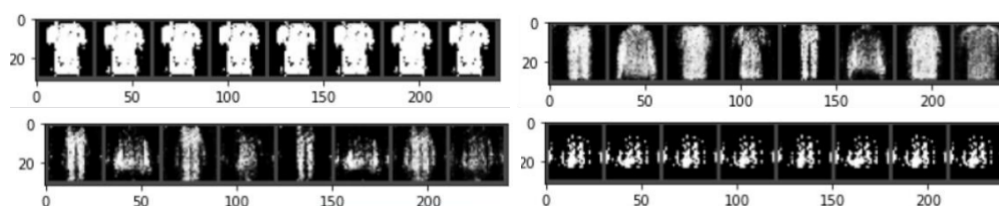


图 4: Caption

可以看到, 随着第 71 个随机数的减小, 白色短袖轮廓特征的图像越来越明显, 随着该随机数的增大, 轮廓越来越不明显。

(三) 修改三

修改第一百个随机数为-50, -10, 10, 如下图 (从上到下依次为 10, 原图, -10, -50), 生成结果如图6所示

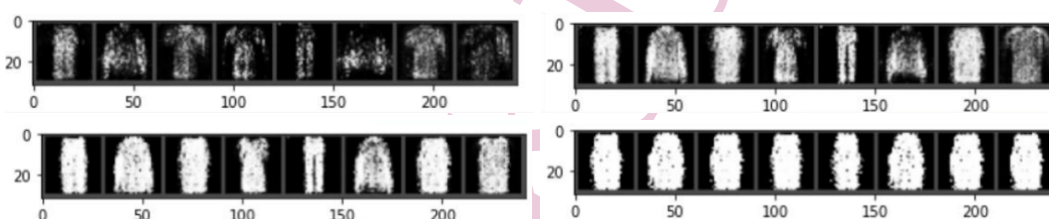


图 5: Caption

可以看到, 随着第 100 个随机数的减小, 白色长方体越来越明显, 随着该随机数的减小, 该特征越来越不明显, 导致几乎所有图片越来越黑。

总结

那么为什么改变噪声的情况会对结果产生不同的影响呢? 其实答案就在 GAN 网络中的生成器中。

我们将输入的随机噪声称为潜在因子。机器学习中的潜在因子通常彼此独立, 以简化模型训练过程。在 GAN 中, 潜在因子 z 的分布应与真实图像的潜在因子分布类似。如果从正态或均匀分布中采样 z , 则优化后的模型可能需要 z 来嵌入类型和风格以外的信息。这里就潜在地为我们指明了 GAN 中的潜在因子实际上代表的就是一个一个不同的特征, 在原始 GAN 的定义中, 是把图片看作高维空间上的点, 而你的训练集也是一系列点, 那么他们在高维空间上就服从一定的分布。那么我们在高维空间上就可以将随机噪声的值看作为高维度上的点, 这些点在各自的隐空间中具有各自的语义和信息, 因此擅自改变噪声的值, 带来的就是网络能读取的信息的改变, 如下例子很好证明这一点:

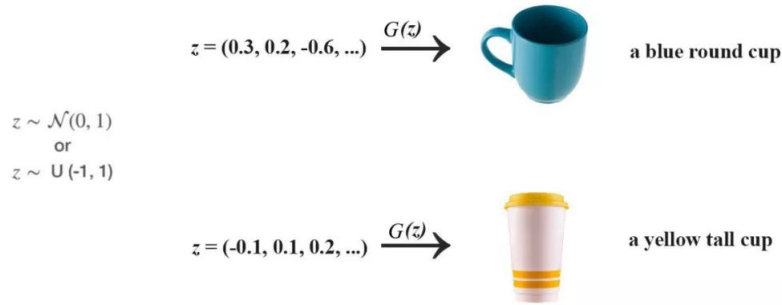


图 6: Caption

而且根据上述实验结果，我们可以大胆猜测，第七十一个随机数隐含的信息是 T 恤，它越小，T 恤越明显；第 100 个随机数则隐含裤子的语义，越小裤子的轮廓越明显；第二个随机数则隐含长袖的语义，越大长袖的轮廓越明显。

那么如何去深究这些随机噪声在隐空间上的含义呢？styleGAN 告诉我们答案：它把高斯分布的噪声（下图 b）使用 MLP 扭曲成隐含的特征空间（下图 a），以便更好地生成图像。

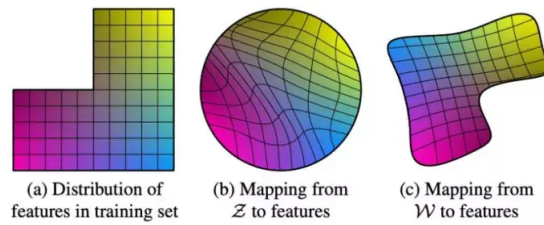


图 7: Caption

而基于风格的生成器为了保证潜在因子 Z 的作用，使用单独学得的仿射运算 A 在每一层中转换 w。转换后的 w 将作为风格信息作用于空间数据。

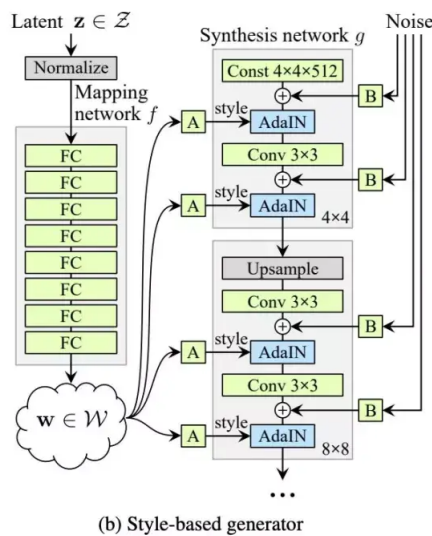


图 8: Caption