



南開大學
Nankai University

网络空间安全学院
深度学习（高阶课）实验报告

生成对抗网络实验报告

姓名：王舒瑀
学号：2212777
专业：物联网工程

2025 年 6 月 3 日

目录

| | |
|--|----------|
| 1 实验要求 | 2 |
| 2 实验内容 | 2 |
| 2.1 基于 GAN 网络训练 FashionMNIST 数据集 | 2 |
| 2.1.1 网络结构 | 2 |
| 2.1.2 训练结果 | 2 |
| 2.1.3 潜向量维度可视化分析 | 3 |
| 2.2 基于 DCGAN 网络训练 FashionMNIST 数据集 | 5 |
| 2.2.1 网络结构 | 5 |
| 2.2.2 训练结果 | 6 |
| 2.2.3 结果对比 | 7 |

1 实验要求

1. 掌握 GAN 原理
2. 学会使用 PyTorch 搭建 GAN 网络来训练 FashionMNIST 数据集
3. 学会搭建 DCGAN 来训练 FashionMNIST 数据集

2 实验内容

2.1 基于 GAN 网络训练 FashionMNIST 数据集

本次实验使用 GAN（生成对抗网络）对 FashionMNIST 数据集进行训练。FashionMNIST 数据集包含 60,000 张训练图像和 10,000 张测试图像，每张图像为 28×28 像素的灰度图像，代表 10 类不同的服饰类别。

2.1.1 网络结构

GAN 网络由两个主要部分组成：生成器（Generator）和判别器（Discriminator）。生成器的任务是从随机噪声生成逼真的图像，而判别器的任务则是区分真实图像与生成图像。

```
Discriminator(  
    (fc1): Linear(in_features=784, out_features=256, bias=True)  
    (lr1): LeakyReLU(negative_slope=0.2)  
    (fc2): Linear(in_features=256, out_features=1, bias=True)  
)  
Generator(  
    (fc1): Linear(in_features=100, out_features=256, bias=True)  
    (lr2): LeakyReLU(negative_slope=0.2)  
    (fc2): Linear(in_features=256, out_features=784, bias=True)  
)
```

图 2.1: GAN 网络结构

如图 2.1 所示，本部分实验所用的 GAN 网络结构中，生成器采用全连接和反卷积层逐步提升图像尺寸至 28×28 ，判别器则通过卷积层和全连接层判别图像真伪。

以下是自定义随机数生成的一组八个图像：

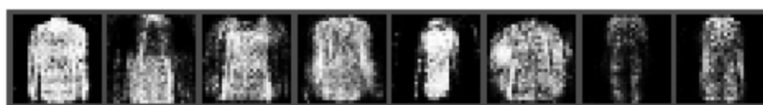


图 2.2: GAN 生成图

通过上述措施，可以有效缓解 DCGAN 训练中判别器过强的问题，促进生成器与判别器之间的良性竞争，从而改善生成效果。

2.1.2 训练结果

GAN 网络经过 30 个 epochs 的训练后，生成的服饰图像逐渐逼近真实图像，如图 2.3 所示。初始阶段生成图像模糊且特征不明显，随着训练迭代的增加，生成图像的质量显著提升，各类服饰的结构和细节逐渐显现。

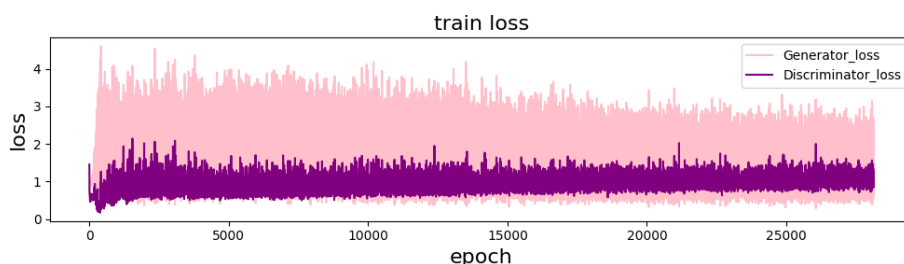


图 2.3: GAN 网络训练图像

2.1.3 潜向量维度可视化分析

为了探究潜空间中各维度所对应的“语义”或“属性”，我们对训练好的生成器进行以下实验流程：

1. 先固定一组 100 条随机噪声向量，每次取其中 8 条生成基准图 2.10。2. 随机挑取 5 条噪声向量，对每条向量在某一维度（如维度 11、23）分别做三次微调 ($\Delta z = -4, 0, +4$)，每次复制该向量生成 8 张图，共 $5 \times 3 = 15$ 组；3. 统计并记录每组图像的平均亮度 (mean_lum) 与对比度 (std_contrast)，并结合可视化结果进行语义归纳。

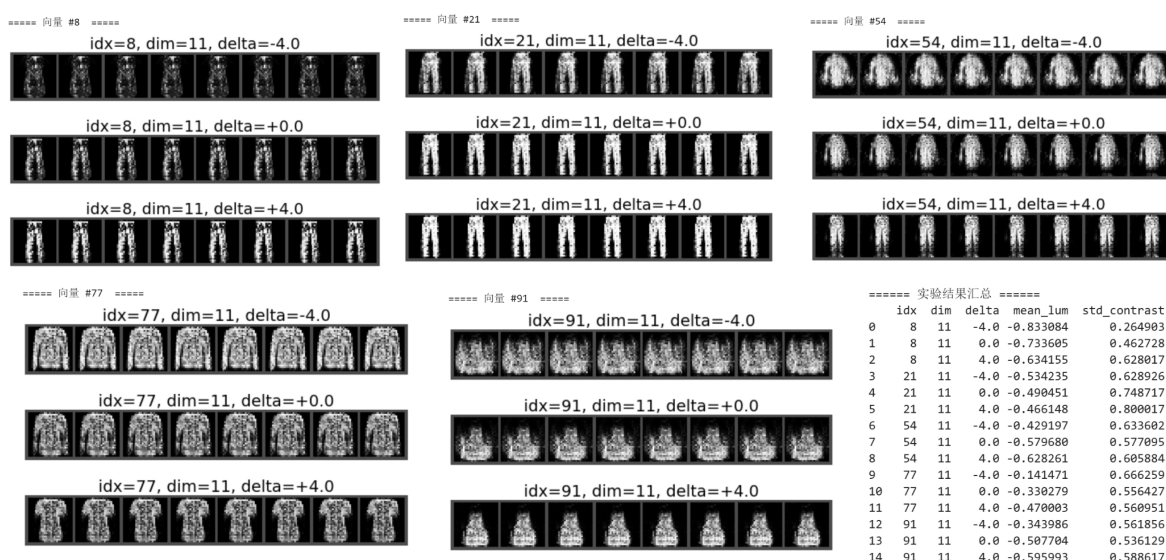


图 2.4: GAN 网络维度 11 微调效果

维度 11 调整效果 如图 2.11 所示，对潜向量第 11 维做 $\Delta z = \{-4, 0, +4\}$ 调节时，生成图像在“亮度”和“纹理清晰度”上产生了显著变化：

- $\Delta z = -4$ ：平均亮度最低（较暗），对比度也最低，图像纹理平滑，边缘模糊；
- $\Delta z = 0$ ：基准状态，各向量的亮度和对比度位于中间值；
- $\Delta z = +4$ ：平均亮度最高（较亮），对比度最高，纹理细节丰富，轮廓锐利。

从定量结果来看，所有 5 条向量在第 11 维上都呈现“亮度与对比度随 Δz 线性上升”的一致性趋势，表明第 11 维对应了一个“整体提亮+纹理锐化”的语义方向。

维度 11 扩展扫描

为了进一步验证第 11 维在潜空间中对“亮度+纹理锐化”语义的控制能力，在原先的 $\Delta z = \{-4, 0, +4\}$ 之外，又增加了更大幅度的 $\Delta z = \pm 20$ 扫描。

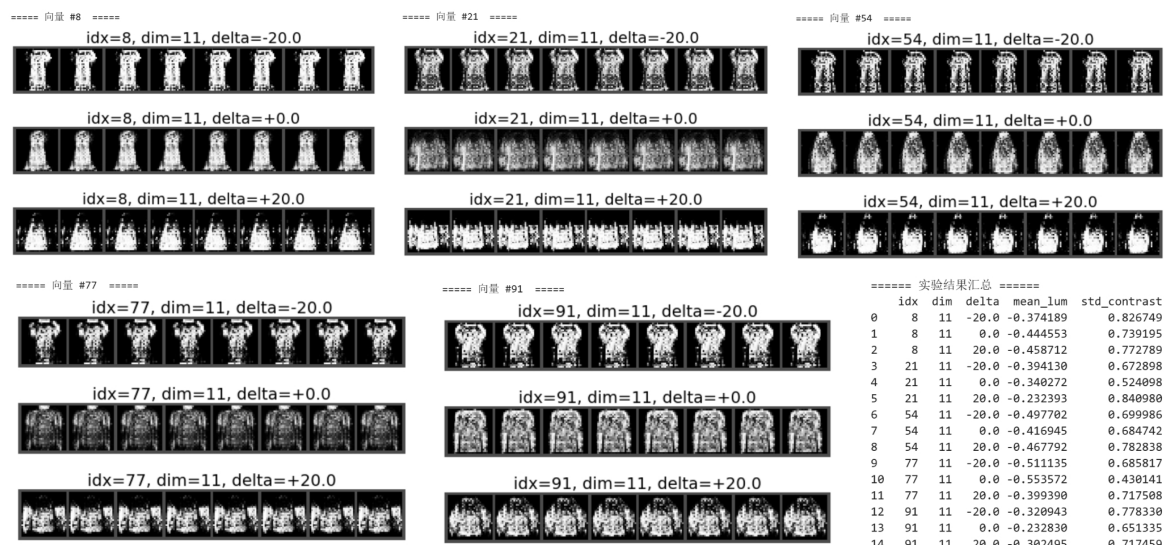


图 2.5: GAN 网络维度 11 大幅度微调效果

随着 $|\Delta z|$ 从 4 上升到 20，**平均亮度** (mean_lum) 和**对比度** (std_contrast) 都呈现出更强的线性趋势，同时进一步放大这些特点，过大幅度会导致图像局部出现轻微失真或噪点“爆炸”。由此可证，第 11 维在潜空间中可能对应一个“光照+纹理强度”方向，控制亮度和清晰度。亮度与对比度随 Δz 近似线性变化

维度 23 调整效果 如图 2.6 所示，对潜向量第 23 维做相同幅度微调时，生成图像发生了“类别/形状”层面的转换，而不仅仅是亮度或纹理变化。以 idx=54 为例： $\Delta z = -4$ ：倾向生成上衣 (T-shirt、毛衣等)； $\Delta z = 0$ ：生成鞋履 (短靴、运动鞋等)； $\Delta z = +4$ ：生成下装 (裤子、连衣裙等) 或配件 (手提包、双肩包等)。此时亮度和对比度指标波动微小 (Δ 小于 0.1)，但类别发生了跳跃，说明第 23 维对于类别这样的特征更为敏感

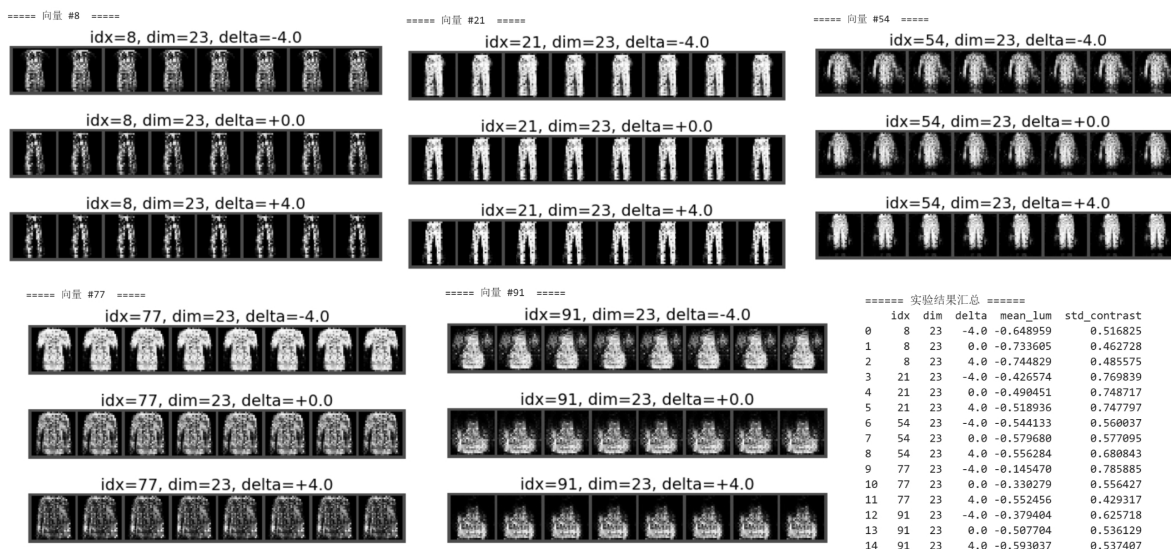


图 2.6: GAN 网络维度 23 微调效果

形状变化的成因分析 在对第 11 维施加大幅度 Δz （如 ± 20 ）时，除了亮度和纹理的明显变化，我们还观测到生成图的整体轮廓与局部形状都会发生可见的变化。这背后的主要原因有两点：

1. **语义维度的混合性**：GAN 潜空间中的每个维度并非严格“纯亮度”或“纯纹理”，而是一个混合了多种高阶特征的方向。生成器内部各层的特征变换会将同一维度的信号映射到不同的语义子空间——包括光照、纹理、乃至形状。因此，当 $|\Delta z|$ 足够大时，这个维度携带的“几何”成分也被激活，导致物体轮廓拉伸、收缩或出现扭曲。
2. **分布之外的失真效应**：极端的 Δz 会使噪声向量离开训练时的分布支持区域，生成器在该区域往往工作不稳定，容易出现形状畸变或局部“爆裂”噪点。这类失真看似“形状变化”，实则是模型对超出其学习经验范围输入响应过冲。

2.2 基于 DCGAN 网络训练 FashionMNIST 数据集

为进一步提高生成图像的质量，本实验使用了 DCGAN（深度卷积生成对抗网络）结构进行训练。与传统 GAN 不同，DCGAN 使用深度卷积网络作为生成器和判别器，大大提高了网络的稳定性和生成图像的质量。

2.2.1 网络结构

图 2.7 展示了 DCGAN 的具体网络结构，生成器通过反卷积层有效地学习图像特征，而判别器则利用卷积层高效地辨别图像真伪。

```
AttnDecoderRNN(
  (embedding): Embedding(2991, 512)
  (attention): BahdanauAttention(
    (W_a): Linear(in_features=512, out_features=512, bias=True)
    (U_a): Linear(in_features=512, out_features=512, bias=True)
    (V_a): Linear(in_features=512, out_features=1, bias=True)
  )
  (gru): GRU(1024, 512, batch_first=True)
  (out): Linear(in_features=512, out_features=2991, bias=True)
  (dropout): Dropout(p=0.1, inplace=False)
)
```

图 2.7: DCGAN 网络结构图

DCGAN 网络在构建和训练中遇到了一些训练问题：

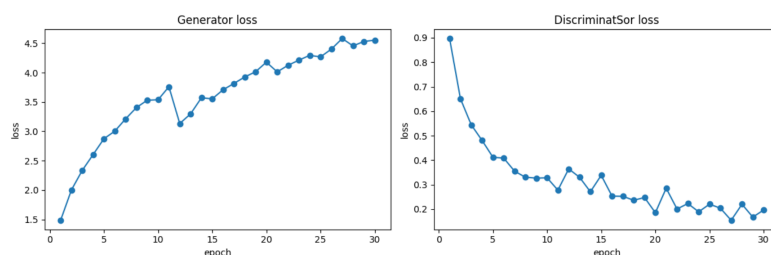


图 2.8: 判别器过强问题图像

由图 2.8 可以明显看出，DCGAN 网络在训练过程中出现了判别器过强的问题。具体表现为：

- **生成器损失逐步增大**：从图中左侧 Generator loss 曲线可以看出，随训练 epoch 增加，生成器的损失逐渐增加，尤其是在第 10 个 epoch 之后明显持续上升。这表明生成器在生成欺骗判别器的假图像时困难越来越大，生成图像质量无法有效提升甚至可能退化。
- **判别器损失迅速下降后趋于稳定**：右侧 Discriminator loss 图像显示，在训练的前 10 个 epoch，判别器损失迅速降低，说明判别器迅速适应并轻易地识别生成器生成的假图像，从而导致生成器无法有效欺骗判别器。

原因分析及改进

判别器过强主要可能源于以下几种原因：首先可能是判别器网络容量过大，能够轻易识别生成器所产生的图像。且生成器网络容量相对不足，生成的图像质量难以有效提高。其次也可能是判别器的训练步骤或学习率设定不合理，可能学习速度过快。故调整了网络结构与容量，减小判别器网络的复杂性和容量，学习策略调整方面适当降低判别器的学习率，防止其过快适应生成器的输出。

2.2.2 训练结果

通过上述措施，可以有效缓解 DCGAN 训练中判别器过强的问题，促进生成器与判别器之间的良性竞争，从而改善生成效果。

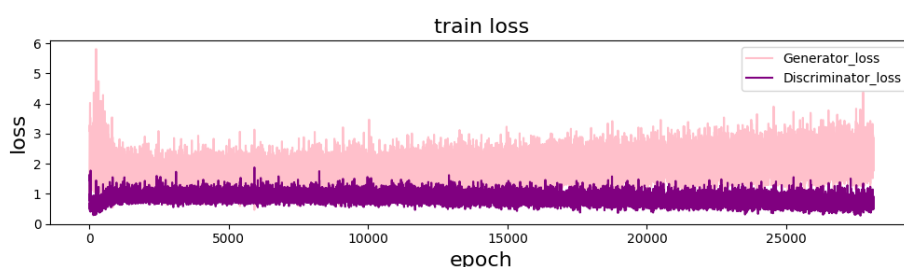


图 2.9: DCGAN 训练曲线

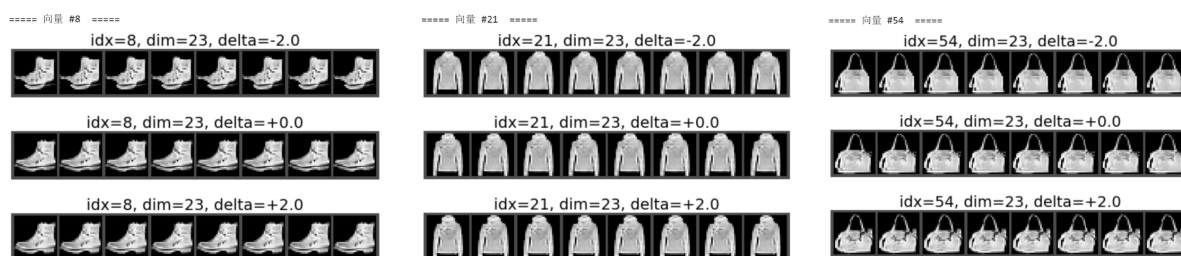
以下是训练结束后自定义随机数生成的图像实例：



图 2.10: DCGAN 生成图

经过上述改进后，生成器和判别器损失曲线趋于平稳，对抗过程更加均衡。第 50 个 epoch 后固定噪声生成的样本如图 2.10 所示，图像结构更加清晰、细节更丰富。

以下是维度 23 上进行调整后的图像：



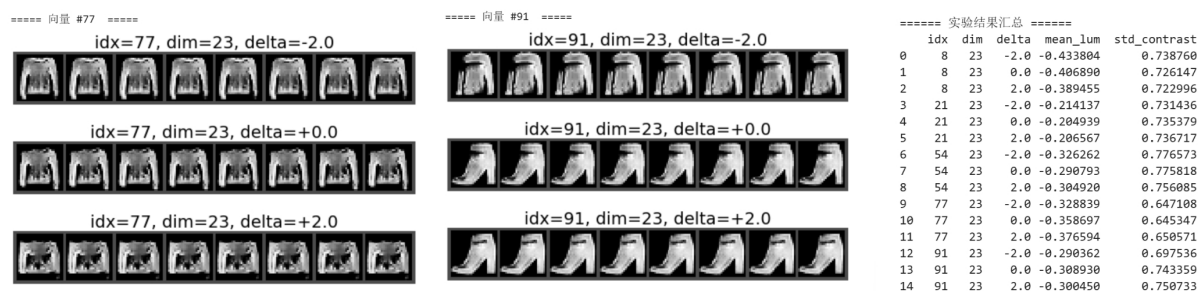


图 2.11: DCGAN 网络维度 23 调整

可以看到,在 DCGAN 网络上调整时,量化指标几乎没有发生明显的“亮度 / 对比度”波动(mean_lum 都在 $-0.26 \sim -0.46$ 之间, std_contrast 也只在 $0.69 \sim 0.75$ 之间小幅上下),但从可视化来看,模型的变化在生成的类别形状上有体现。特别的,DCGAN 出现类别较大变化的值比 GAN 小,可以看到 Δz 为 2 时已经有很明显的形状变化,而 GAN 在 Δz 为 4 时形状只有某些图片由一定明显变化。这与两个网络的不同结构有关。

2.2.3 结果对比

通过对比 GAN 与 DCGAN 在 FashionMNIST 数据集上的表现,并对潜在空间中的若干维度进行了可视化分析,可以发现:

- **DCGAN 优于普通 GAN:** 深度卷积结构提升了图像质量和训练稳定性,使得生成样本在细节和纹理上更胜一筹。
- **潜空间具备可控语义方向:** 无监督训练即可在线性维度上找到控制亮度(第 11 维)、类别切换(第 23 维)等属性的方向。
- **幅度与簇间距离:** Δz 的大小决定了属性变化的幅度,从同簇内细节到跨簇类别均可通过线性扫描实现。