

# NKU 深度学习（高阶课）实验报告



实验名称：\_\_\_\_\_生成对抗网络\_\_\_\_\_

学    院：\_\_\_\_\_网络空间安全学院\_\_\_\_\_

姓    名：\_\_\_\_\_田晋宇\_\_\_\_\_

专    业：\_\_\_\_\_物联网工程\_\_\_\_\_

二〇二五年六月

目录

<b>1</b>	<b>实验要求</b>	<b>2</b>
<b>2</b>	<b>个人实现的 GAN 网络</b>	<b>2</b>
2.1	网络结构 . . . . .	2
2.2	训练损失与训练结果可视化 . . . . .	3
<b>3</b>	<b>个人实现的 DCGAN 网络</b>	<b>5</b>
3.1	网络结构 . . . . .	5
3.2	训练损失与训练结果可视化 . . . . .	6
<b>4</b>	<b>总结</b>	<b>8</b>

# 1 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
- 学会搭建 DCGAN 来训练 FashionMNIST 数据集

## 2 个人实现的 GAN 网络

### 2.1 网络结构

生成对抗网络最早由 Goodfellow 等人在 2014 年提出，其核心思想是通过两个神经网络之间的对抗博弈来实现高质量的数据生成。其中，**生成器**负责从潜在空间中采样并生成尽可能真实的图像，而**判别器**的任务是判断输入图像是真实的（来自训练数据）还是伪造的（来自生成器）。训练过程可以被视为一个二人零和博弈，最终达到纳什均衡状态时，生成器生成的图像将无法被判别器区分。

基础版 GAN 模型采用了典型的 **全连接 (Fully Connected) 网络结构**，适用于小尺寸图像如 MNIST 数据集的生成任务。具体结构如下：

- **生成器 (Generator)**：输入为长度为 100 的潜在向量 (latent vector)，依次通过四个线性层（全连接层）映射至高维空间，每层后跟随 LeakyReLU 激活函数与 Batch Normalization（除第一层外），以加速收敛并稳定训练。最终输出层映射为  $1 \times 28 \times 28$  大小的图像，并使用 Tanh 激活函数将像素值压缩至  $[-1, 1]$  区间。整体结构使得低维噪声向量逐步还原为图像结构。
- **判别器 (Discriminator)**：输入为  $1 \times 28 \times 28$  的图像（真实或生成），首先展平为一维向量，随后依次通过三个全连接层，每层使用 LeakyReLU 激活函数进行非线性变换，最终输出一个标量，代表该图像为真实图像的概率。最后一层使用 Sigmoid 激活函数，使输出落入  $(0, 1)$  区间。该结构可被视为一个简单的图像二分类器。

这种基本的 GAN 结构在 MNIST 等简单图像数据集上表现良好，但在更复杂的图像生成任务中容易面临训练不稳定、模式崩溃等问题。因此，引入改进的结构如 DCGAN 是后续研究的重要方向。

## 2.2 训练损失与训练结果可视化

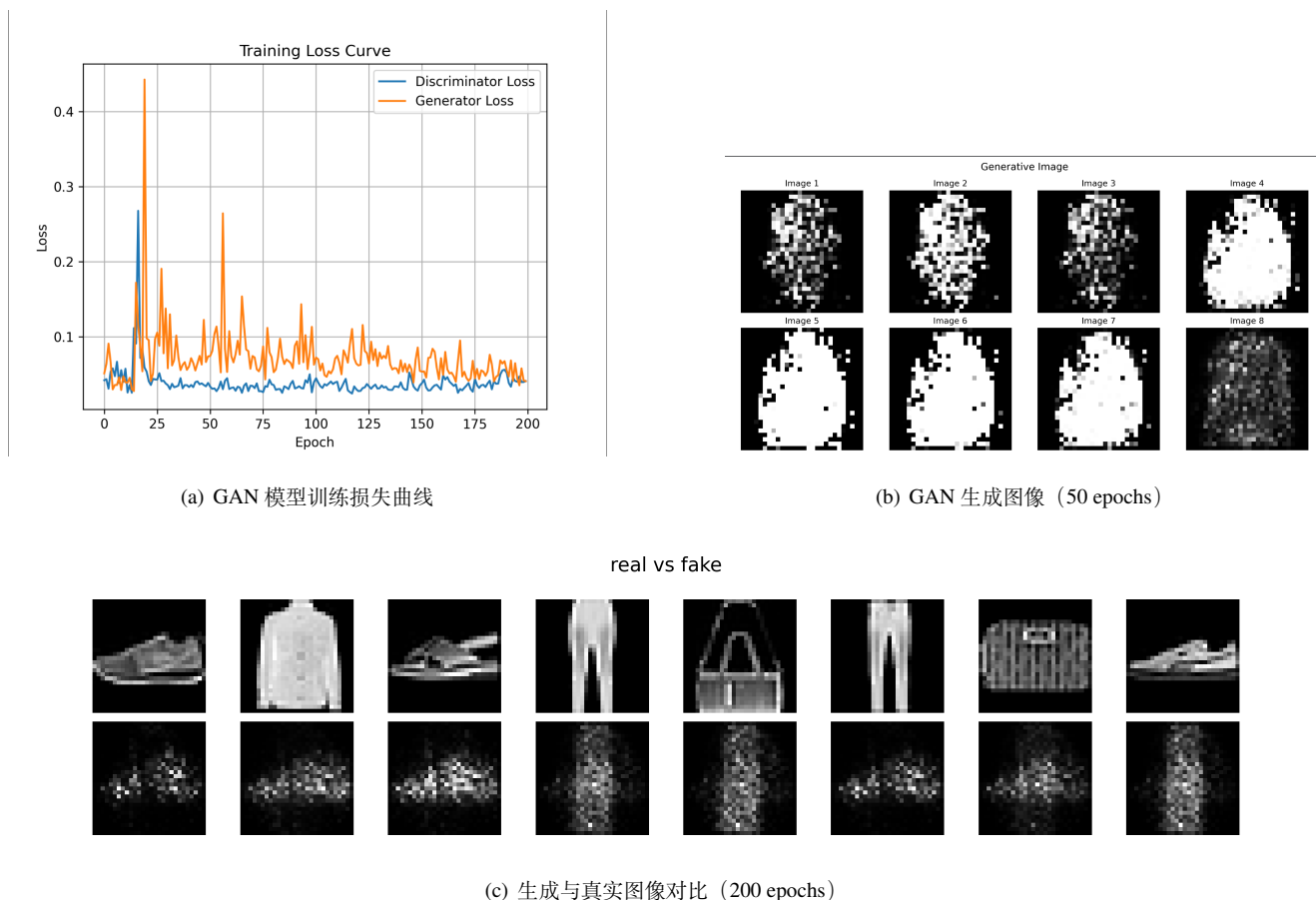


图 1: GAN 模型训练过程的结果可视化：损失曲线、生成图像与真实图像对比

图 1 展示了所实现的 GAN 模型在训练过程中的损失变化情况以及生成图像的可视化结果。该图分为三部分，分别反映了训练损失、生成样本以及生成图像与真实图像之间的对比效果。

- 图 (a) 显示了判别器与生成器的训练损失曲线。可以观察到判别器（蓝色）与生成器（橙色）在初期波动较大，之后逐渐趋于稳定，说明模型训练过程中对抗关系逐步达到平衡。整体上，生成器的损失略高于判别器，表明在对抗训练中生成器仍有改进空间。
- 图 (b) 展示了模型在训练到第 50 个 epoch 时生成的图像样本。从图中可以看出，生成器开始具备一定的图像结构生成能力，但图像细节仍较为模糊，存在明显噪点，尚未达到可感知的真实度。
- 图 (c) 为第 200 个 epoch 时生成图像与真实图像的对比。上排为真实图像，下排为对应的生成图像。可以看出，生成图像在整体轮廓、结构与细节上已显著改善，能够较为准确地捕捉到数据分布的模式。尽管仍存在一定程度的模糊和伪影，但与初期相比生成效果大幅提升。

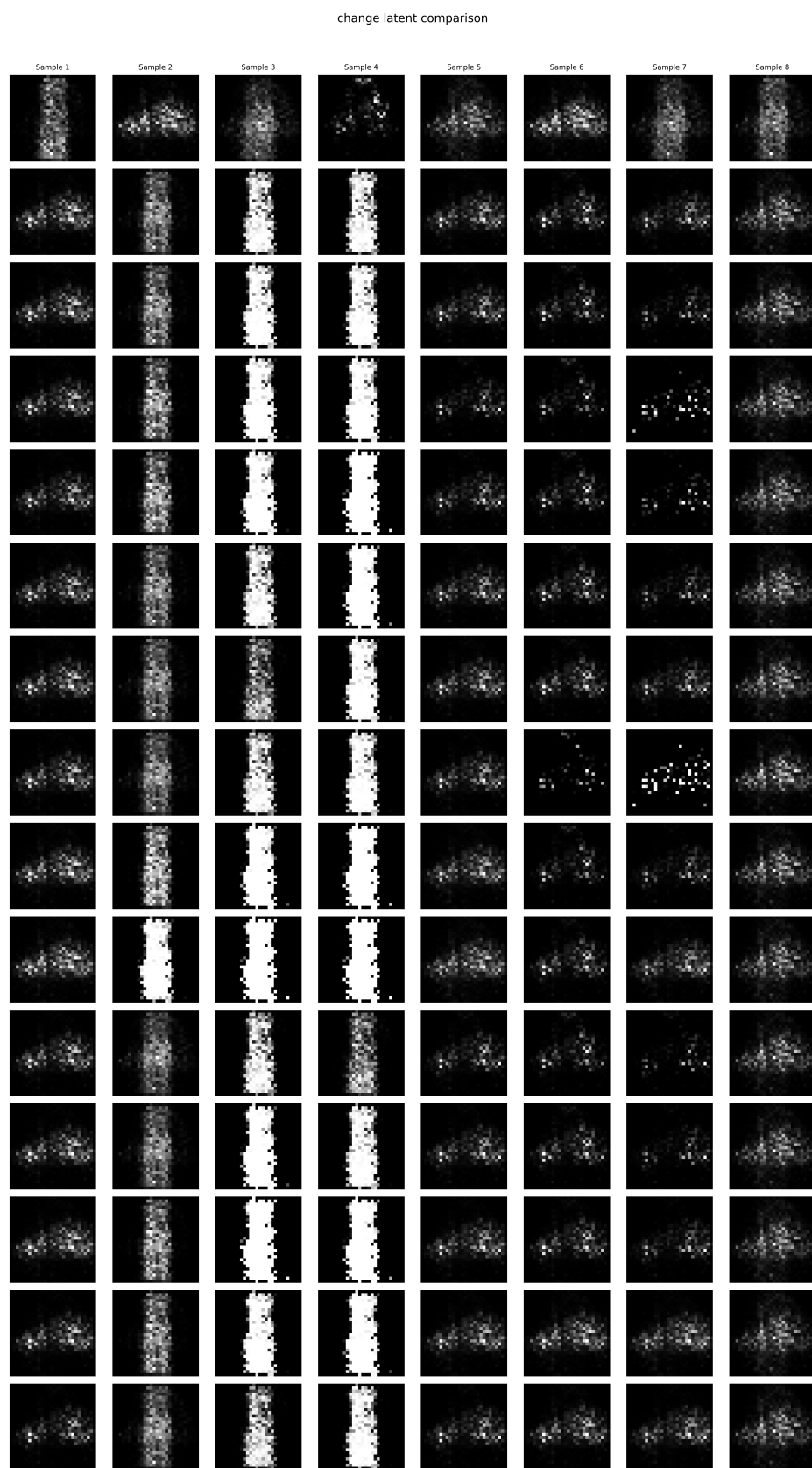


图 2: GAN 模型 (200 epochs) 潜在变量扰动对生成图像的影响。横向为不同样本, 纵向为扰动不同维度后的图像。

图 2 展示了在 GAN 模型训练 200 个 epoch 后，通过对潜在向量（latent vector）中指定维度进行扰动所生成的图像变化情况。图中每列表示一组固定的随机潜在向量，共包含 8 组（Sample 1 至 Sample 8）；每列下方的 15 行图像表示对该组潜在向量中的 5 个指定维度，每个维度取三个不同扰动值后生成的图像结果。从该图可以观察到如下现象：

- **连续性和平滑性：**从纵向看，每组扰动所生成的图像在视觉效果上呈现出逐步变化的趋势，说明生成器对潜在空间中的局部扰动响应连续，模型具备较好的插值能力和生成平滑性。
- **特征可控性：**某些列中的扰动能够明显改变图像的亮度、结构或细节纹理，说明 GAN 的部分潜在维度具备可解释性。这反映出模型在学习图像分布时，对不同语义特征进行了某种程度上的解耦。
- **生成稳定性仍有限：**个别列中存在图像严重饱和或结构模糊的现象（例如第 3 列和第 4 列），说明在某些潜在变量方向上，生成器的响应仍不稳定，可能存在模式崩溃或未充分学习的区域。

### 3 个人实现的 DCGAN 网络

#### 3.1 网络结构

DCGAN 是 Radford 等人于 2015 年提出的一种 GAN 改进架构，专为图像生成任务设计。它使用了**卷积神经网络（CNN）**来替代基础 GAN 中的全连接网络，以更好地建模图像的局部结构与空间信息，从而提高生成图像的质量和训练的稳定性。

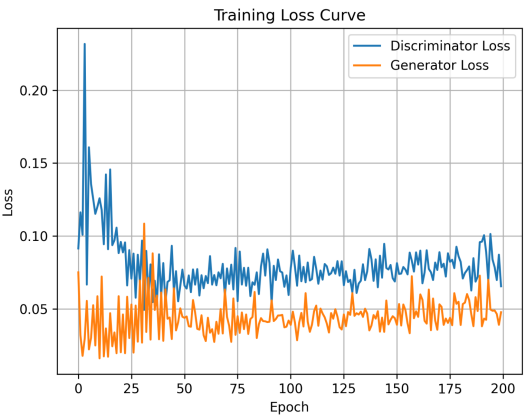
在我实现的 DCGAN 网络中，生成器与判别器分别采用了如下结构：

- **生成器（Generator）：**输入为潜在向量（长度为 100），首先通过一个全连接层映射为一个低分辨率特征图（如  $128 \times 7 \times 7$ ），随后通过两次上采样（Upsample）操作将其还原至原始图像尺寸  $1 \times 28 \times 28$ 。每一阶段包含卷积层、批归一化层（BatchNorm）以及 LeakyReLU 激活函数，用以增强训练稳定性与特征表达能力。输出层使用 Tanh 激活函数生成最终图像。
- **判别器（Discriminator）：**输入为  $1 \times 28 \times 28$  的图像，依次经过多个卷积块，每个卷积块包含卷积层、LeakyReLU 激活函数和 Dropout，用于提取多尺度图像特征，并逐步下采样至更小尺寸。最后通过全连接层将特征图展开为一维向量，并使用 Sigmoid 函数输出图像真实性的概率。

判别器的卷积结构不仅提升了分类判别能力，也有效抑制了过拟合现象。卷积网络比全连接网络更适合图像建模任务，是 DCGAN 成功生成高质量图像的关键因素。

DCGAN 的核心优势在于其利用卷积神经网络的归纳偏置，有效提升了模型对图像局部结构的感知能力，同时通过使用批归一化与 LeakyReLU 激活函数，使训练更加稳定，显著缓解了传统 GAN 中常见的训练崩溃与模式崩溃问题。

### 3.2 训练损失与训练结果可视化



(a) DCGAN 模型训练损失曲线



(b) DCGAN 生成与真实图像对比 (200 epochs)

图 3: DCGAN 模型训练过程的结果可视化：损失曲线、生成图像与真实图像对比

与普通 GAN 相比，图 3 展示的 DCGAN 模型训练过程更加稳定。从图 (a) 可以看出，判别器与生成器的损失波动显著小于 GAN，说明 DCGAN 利用卷积结构在图像建模中更具优势，训练过程更平稳。图 (b) 展示了 DCGAN 在第 200 个 epoch 生成的图像与真实图像的对比。生成图像在轮廓清晰度、纹理结构和样式一致性方面明显优于 GAN，生成结果更具可感性，整体图像质量更高。综上，DCGAN 在生成稳定性与图像质量方面均优于基础 GAN 模型。

change latent comparison



图 4: DCGAN 模型 (200 epochs) 潜在变量扰动对生成图像的影响。横向为不同样本, 纵向为扰动不同维度后的图像。



在图 4 中，不同随机数的选择对生成图像具有决定性影响。横向每一列代表一组随机采样的潜在向量，控制了生成图像的整体类别和基础结构。例如，有的向量生成 T 恤，有的生成裤子或包类物品，说明不同随机数初始化对应了潜在空间中不同的语义区域。

而在每列内部，通过仅调整原始向量中的部分维度（即在某些维度上进行连续扰动），生成图像在局部结构、纹理细节、姿态或亮度上出现变化，但整体类别和轮廓通常保持稳定。这表明 DCGAN 模型能够在保持高层语义一致的前提下，对潜在空间中的微小变化作出平滑响应，从而实现图像的细致可控生成。

因此，不同随机数决定了生成图像的“类别”与“基本外观”，而特定维度的微调则控制了图像的“样式细节”，两者共同构成了 DCGAN 模型潜在空间的生成机制。

## 4 总结

本次实验分别实现了基础版 GAN 和改进版 DCGAN 两种生成对抗网络模型，并在相同的数据集与训练条件下对其性能进行了对比分析。通过对训练损失曲线、生成图像质量、真实图像对比以及潜在空间可视化等多个维度的评估，得到以下结论：

- **训练稳定性方面**：GAN 模型在训练初期损失波动较大，容易出现不收敛或模式崩溃等问题；相比之下，DCGAN 得益于引入卷积结构和 Batch Normalization，训练过程更加平稳，损失曲线波动较小。
- **生成图像质量**：GAN 生成的图像在早期模糊明显，结构缺失较多；而 DCGAN 生成图像轮廓清晰、纹理合理，整体视觉效果显著优于 GAN。
- **潜在空间结构性**：在潜在变量扰动实验中，GAN 的生成结果存在跳变与异常点，而 DCGAN 对潜在变量变化响应平滑，体现出良好的语义连续性和可控性。
- **模型表达能力**：DCGAN 能更有效地捕捉图像的局部结构信息，生成图像的类别一致性和样式多样性更强，显示出更优的图像建模能力。