

一、GAN 算法

1.什么是 GAN(隐式密度模型)

GAN 是一种深度神经网络架构，由一个生成网络和一个判别网络组成。生成网络产生“假”数据，并试图欺骗判别网络；判别网络对生成数据进行真伪鉴别，试图正确识别所有“假”数据。在训练迭代的过程中，两个网络持续的进化和对抗，直到达到平衡状态，判别网络无法识别“假”数据，训练结束。

2.GAN 原理

生成器主要用来学习真实图像分布从而让自身生成的图像更加真实，以骗过判别器。

判别器则需要对接受图片进行真假判别。

在训练过程中，生成器努力地让生成的图像更加真实，而判别器则努力地去识别出图像的真假，这个过程相当于一个二人博弈，随着时间的推移，生成器和判别器在不断地进行对抗。最终两个网络达到了一个动态均衡：生成器生成的图像接近于真实图像分布，而判别器识别不出真假图像，对于给定图像的预测为真的概率基本接近 0.5（相当于随机猜测类别）。

二、算法的流程

1.GAN 算法的流程

假设我们有两个网络 G (Generator) 和 D (Discriminator)。

- G 是一个生成图片的网络，它接受一个随机的噪声 z ，提高这个噪声生成图片，记作 $G(z)$ 。

- D 是一个判别网络，判别一张图片是不是“真实的”，它的输入参数为 x ， x 代表一张图片，输出 $D(x)$ 代表 x 为真实图片的概率，如果为 1，就代表是真实的，若为 0，代表不是真实的。

在训练过程中，将随机噪声输入生成网络 G，得到生成的图片；判别器接收生成的图片和真实的图片，并尽量将两者区分开来。在这个计算过程中，能否正确区分生成的图片和真实的图片将作为判别器的损失，而能否生成近似真实的图片并使得判别器将生成的图片判定为真将作为生成器的损失。这样的话，生成器的损失是通过判别器的输出来计算的，而判别器的输出是一个概率值，我们可以通过交叉熵计算。

2.损失函数

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p(z)} [\log (1 - D(G(z)))]$$

公式中的 x 代表真实图片。

z 代表输入 G 网络的噪声。

$G(z)$ 表示由 z 在 G 网络生成的图片。

$D(x)$ 表示 D 网络判断图片是否真实的概率。

对公式的注释：

对于判别器 D ，我们希望它可以正确识别真实数据（即更 $D(x)$ 输出更接近于 1），这便是 GAN 公式前半部分：

$$E_{x \sim p_{data}(x)} [\log D(x)]$$

故该公式的含义是：判别器判断出真实数据的概率。我们的优化目标是希望这个概率越大越好。也就是说，对于服从 P_{data} 分布的图片 x ，判别器应该给出的预测结果时 $D(x) = 1$

再来看后半部分：

$$E_{z \sim p(z)} [\log (1 - D(G(z)))]$$

其中 $p(z)$ 表示噪声的分布

对于判别器 D 来说，如果其输入的是生成的数据，即 $D(G(z))$ 判别器的目标是最小化该值，希望它被判断为 0，即希望 $\log (1 - D(G(z)))$ 越大越好。

对于生成器 G 来说，它希望生成的数据被判别器识别为真，即 $D(G(z)) = 1$ ，也即 $\log (1 - D(G(z)))$ 越小越好。

可以看到判别器 D 和生成器 G 对 $\log (1 - D(G(z)))$ 优化目标是相反的。

总结来说，对于判别器 D ，我们希望最大化 $\log D(x)$ 和 $\log (1 - D(G(z)))$ ，从而最大化 $V(D, G)$ 。

对于生成器而言，其目标是最小化 $\log (1 - D(G(z)))$ ，从而最大化 $V(D, G)$ 目标。

3.GAN 应用领域

图像生成、图像增强、风格化和艺术的图像创造、声音的转换