

ProGAN

notes written by h1astro

paper单位: NVIDIA

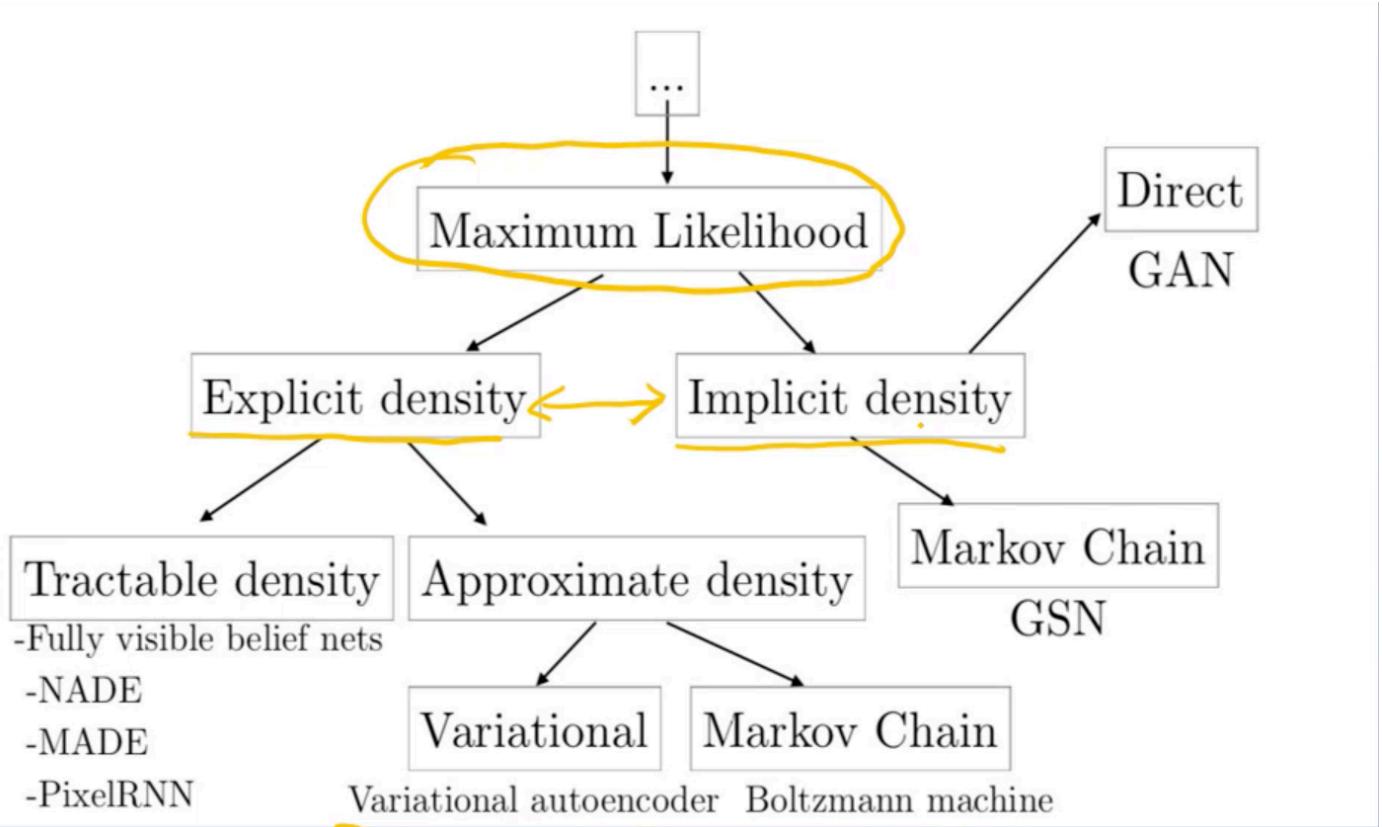
ICLR 2018

核心要点

1. 使用渐进的方式来训练生成器和判别器: 先从生成低分辨率图像开始, 然后不断增加模型层数来提升生成图像的细节
2. 方法能加速模型训练并大幅提升训练稳定性, 生成前所未有的高质量图像 (**1024*1024**)
3. 提出了一种简单方法来增加生成图像的多样性
4. 介绍了几种限制生成器和判别器之间不健康竞争的技巧
5. 提出了一种评价GAN生成效果的新方法, 包括对生成质量和多样性的衡量
6. 构建了一个CELEBA数据集的高清版本 HQ

研究背景

- 显示密度模型
 - 易解显性模型: 定义一个方便计算的密度分布, 主要的模型是Fully visible belief nets, 简称FVBN, 也被称作Auto-Regressive Network
 - 近似显性模型: 可以定义任意的密度分布, 使用近似方法来求解
- 隐性密度模型
 - GAN



- 神经自回归网络 (PixelRNN/CNN)

通过链式法则把联合概率分布分解为条件概率分布的乘积 使用神经网络来参数化每个P

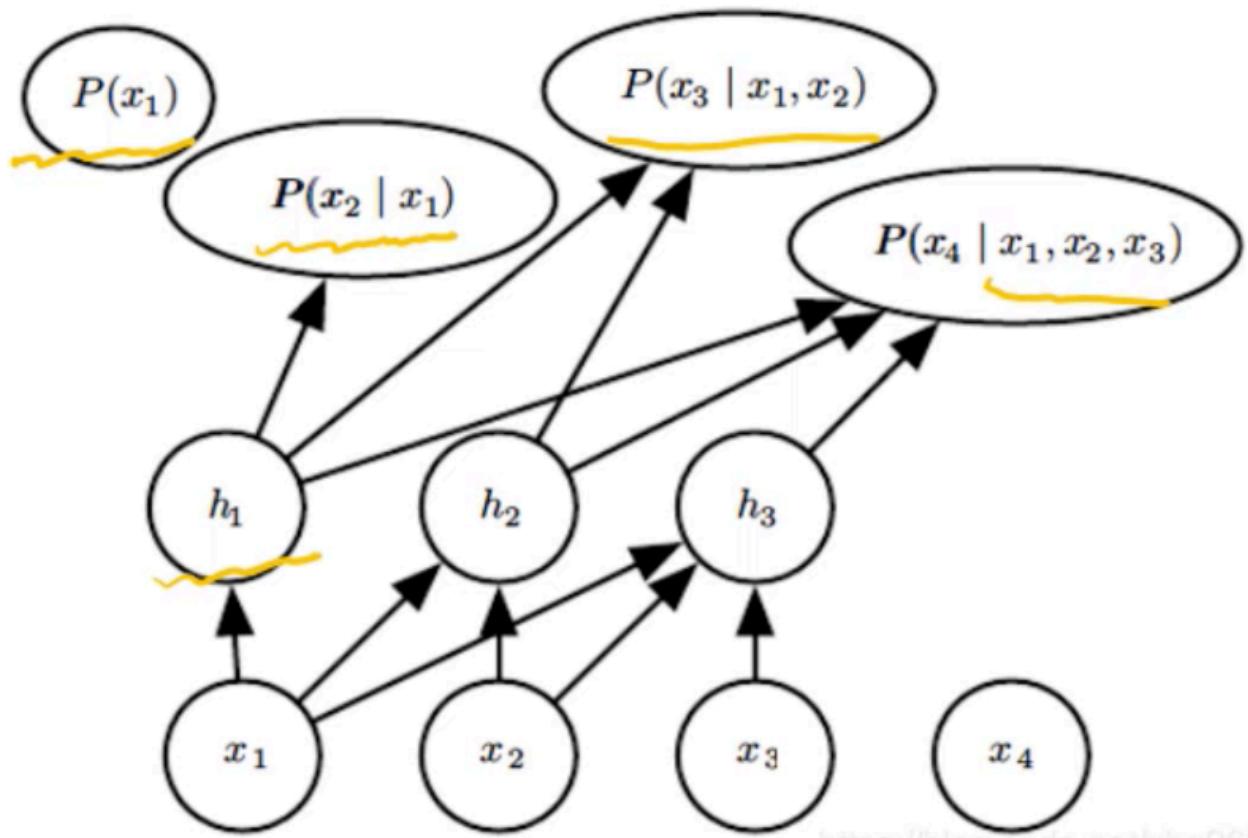
PixelRNN逐像素生成，效率很低，PixelCNN效果不如PixelRNN

- VAE-GAN

编码器：使 $P(z|x)$ 逼近分布 $P(z)$ ，比如标准正态分布，同时最小化生成器(解码器)和输入 x 的差距

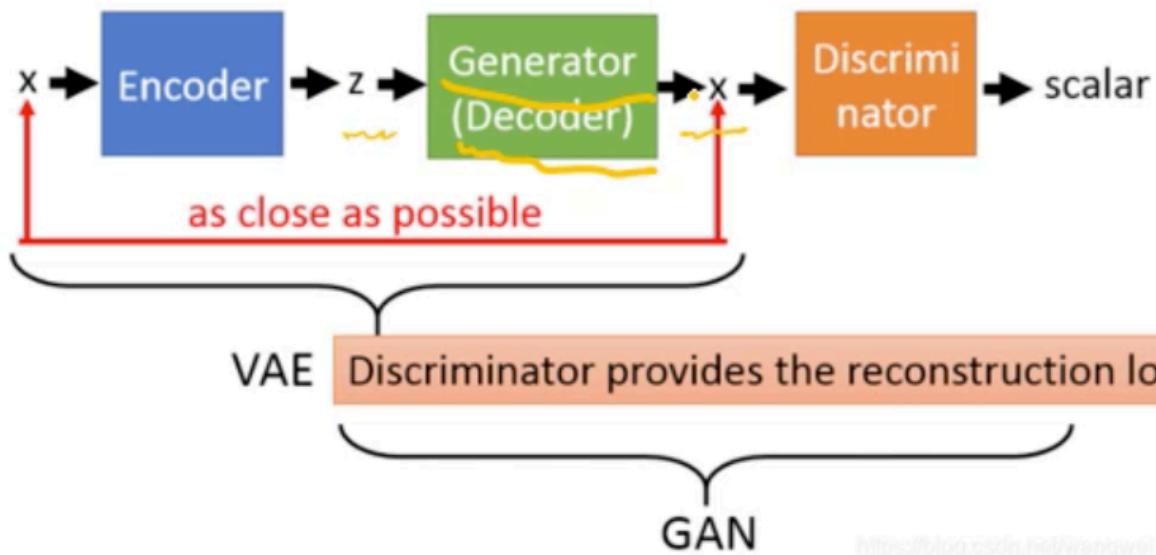
解码器：最小化输出和输入 x 的差距，同时要骗过判别器

判别器：给真实样本高分，给重建样本和生成样本低分



<https://blog.csdn.net/zhq9895>

- Minimize reconstruction error
- z close to normal
- Minimize reconstruction error
- Cheat discriminator
- Discriminate real, generated and reconstructed images



<http://blog.csdn.net/wangwei19871103>

GAN的损失函数

- f-divergence
 - JS散度(交叉熵)
 - LSGAN(MSE)
- Intergral probability metric。更好
 - Wasserstein距离

重点 重点来了!

$$\int_x p_g(x) f\left(\frac{p_{\text{data}}(x)}{p_g(x)}\right) dx$$

$$\mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)] + \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))]$$

$$\frac{1}{2} \mathbb{E}_{x \sim p_{\text{data}}(x)} [D(x) - a]^2 + \frac{1}{2} \mathbb{E}_{z \sim p_z(z)} [D(G(z)) - b]^2$$

$$\mathbb{E}_{x \sim p_{\text{data}}} [f(x)] - \mathbb{E}_{x \sim p_g} [f(x)]$$

$$\frac{1}{N} \sum_{i=1}^N D(x_i) - \frac{1}{N} \sum_{j=1}^N D(y_j)$$

图像生成的评价指标

- 可以评价生成样本的质量
- 可以评价生成样本的多样性，能发现过拟合、模式缺失、模式崩溃、直接记忆样本的问题
- 有界性，即输出的数值具有明确的上下界
- 给出的结果应当与人类感知一致
- 计算评价指标不应需要过多的样本
- 计算复杂度尽量低

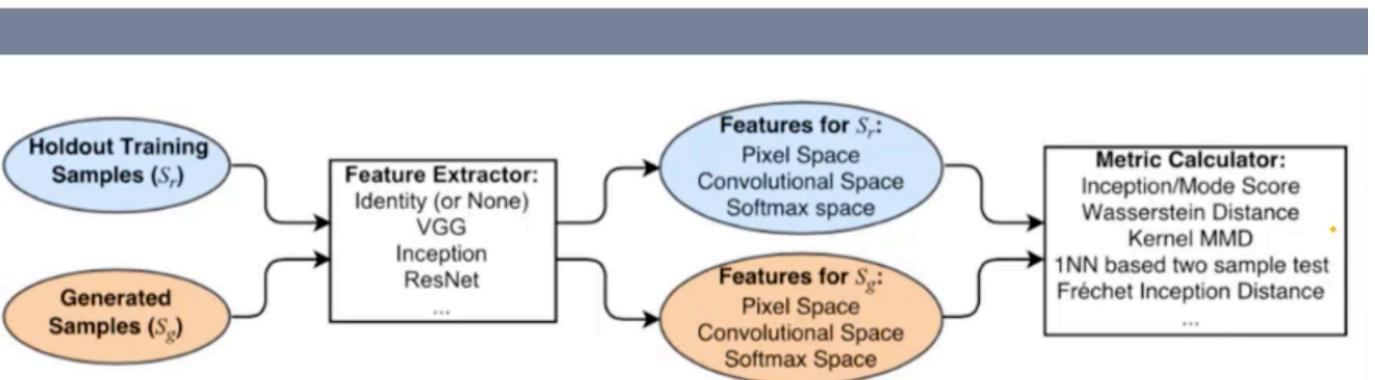
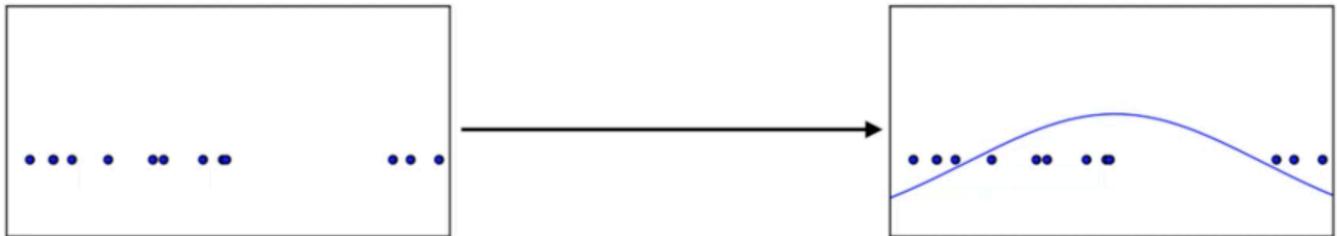


Figure 1: Typical sample based GAN evaluation methods.

[Https://zhuanlan.zhihu.com/p/109342043](https://zhuanlan.zhihu.com/p/109342043)

- Inception Score(IS)

$$\exp\left(\frac{1}{N} \sum_{i=1}^N D_{KL}(p(y|x^{(i)}) || \hat{p}(y))\right) \text{ KL散度越大越好}$$

- Frechet Inception Distance(FID)。 表征距离，描述线的差异

$$\|\mu_{data} - \mu_g\| + \text{tr}\left(\Sigma_{data} + \Sigma_g - 2(\Sigma_{data}\Sigma_g)^{\frac{1}{2}}\right)$$

- Maximum Mean Discrepancy (MMD) 迁移学习常用 最大均值差异

$$\mathbb{E}_{x,x' \sim p_{\text{data}}} [k(x, x')] - 2\mathbb{E}_{x \sim p_{\text{data}}, y \sim p_g} [k(x, y)] + \mathbb{E}_{y,y' \sim p_g} [k(y, y')]$$

- MS-SSIM

$$SSIM(X, Y) = [L_M(X, Y)]^{\alpha M} M \prod_{J=1}^M [C_J(X, Y)]^{\beta_j} [S_J(X, Y)]^{\gamma_j}$$



- 创建了首个大规模高清人脸数据库CelebA-HQ数据集，使得高清人脸生成的研究成为可能
- 首次生成了**1024*1024**分辨率的高清图像，确立了GAN在图像生成领域的绝对优势，大大加速了图像生成从实验室走向实际应用
- 从低分辨率逐次提升的策略缩短了训练所需的时间，训练速度提升2-6倍

