

BigGAN

notes written by h1astro

核心要点

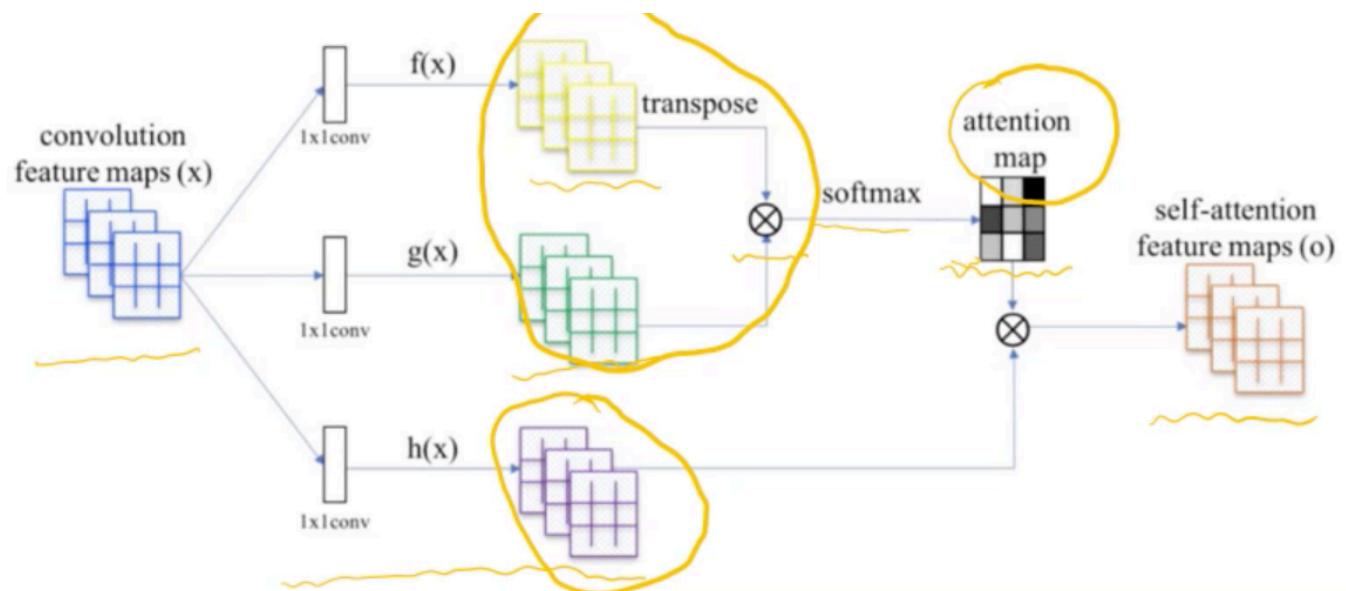
1. 基于复杂数据集(如ImageNet)生成高分辨率的多类别图像仍旧是一个非常困难的目标
2. 为此, 训练了现有最大规模的GAN, 并研究了这种规模下GAN训练的不稳定性
3. 在生成器上应用正交正则化使得它能够进行隐空间的截断, 从而可以调节生成器输入的方差, 实现了对生成图像保真度和多样性之间平衡的良好控制
4. BigGAN成为了目前类别条件图像生成领域的新SOA模型
5. 使用ImageNet进行128x128分辨率的训练时, BigGAN的IS得分为**166.5**, FID得分**7.4**

谱归一化

- 卷积操作可以看成是矩阵乘法, 只需要约束, 每一层卷积核的参数W, 使它是Lipschitz连续的, 就可以满足整个神经网络的Lipschitz连续性
- 矩阵L2范数, 又被称为矩阵的谱范数 $\|W\|_2 = \max_{x \neq 0} \frac{\|Wx\|}{\|x\|} = \sqrt{\lambda_{\max}(\text{eig}(A^T \times A))}$
- 谱归一化, 是另一种让函数满足Lipschitz连续性的方式, 因为经过谱归一化之后, 神经网络的每一层权重, 都满足 $\|A_x\| \leq K\|x\|, \forall x \in \mathbb{R}^n$
- 在每一次训练迭代中对神经网络的每一层作奇异值(SVD)分解求解谱范数是不现实的, 因此采用幂迭代的方法去迭代得到奇异值的近似解

SAGAN (self-attention GAN)

- self-attention机制是Non-local Neural Networks提出的, 能更好地学习到全局特征间的依赖关系
- 基于SNGAN, 使用了谱归一化
- 原始的谱归一化只在判别器中使用, 而SAGAN中谱归一化同时用在了判别器和生成器上
- 除了生成器和判别器的最后一层外, 每个卷积/反卷积单元都会使用谱归一化, 生成器同时还保留了BN层
- SAGAN还用到了《cGANs With Projection Discriminator》提出的conditional normalization和projection in the discriminator



条件判别

- concat自由度非常大，增加了函数的假设空间，导致性能变差

$$y = \frac{x - \mathbb{E}[x]}{\sqrt{\text{Var}[x] + \epsilon}} \cdot \gamma + \beta$$

- Condition Batch Normalization是另一种融合条件信息的方式，作者将它应用在生成器的每一层
- 其中 γ 和 β 是把图片的类别信息输入一个浅层网络求的，从而不同类别的图片，将对于不同的BN层参数
- 判别器中，首先提取输入图像的特征，然后再分成两路：一路与编码后的类别标签 y 做点乘，另一路映射成一维向量，最后两路结果相加，作为神经网络最终的输出，越大代表输入越真实

参数初始化

- 常数初始化
- 均匀分布初始化
- 正态分布初始化
- Xavier均匀/正态分布初始化
- Kaiming均匀/正态分布初始化
- 单位矩阵初始化
- 正交矩阵初始化 $AA^T=E$

研究成果

- 将多类别自然图像生成的效果大幅提升到了令人震惊的水准
- 深入研究了GAN训练稳定性与性能之间的平衡



研究意义

- 证明了大batch和大网络对于GAN的重要性
- 积累了大规模GAN网络训练的工程经验
- 大幅推进了多类别图像生成的SOA得分