

标题：

Large Scale GAN Training for High Fidelity Natural Image Synthesis

作者：

Andrew Brock* Heriot-Watt University

Jeff Donahue DeepMind

Karen Simonyan DeepMind

简介： 这篇文章研究了在超大规模的数据集(比如Imagenet)上面应用GAN，提出了“截断技巧”，并且用了几个新方法减小了训练不稳定的问题。极大提高了性能指标。

3 增大规模

增大batch size，发现可以用更少的循环达到更好的性能，但是模型变得不稳定，而且会完全崩溃。

增大每一层的宽度（channel个数）50%，加倍了参数。发现可以提升性能。但是增大深度会损害性能。

class embedding 在G中的 BatchNorm 层有大量的 weights，所以用共享的 embedding 代替每一个embedding有一层，这可以降低计算量和内存，提高训练速度。

将 latent space中的noise vector z 放入多个层而不只是初始层，这个背后的intuition是可以让G用latent space去影响多个层。这样做提高了内存消耗和计算量，但增加了训练速度。

3.1 “截断技巧”对于多样性和质量的权衡

对于先验分布 z ，运用“截断技巧”设定阈值，只要阈值之内的，然后把阈值之外的重新采样使其低于阈值。降低阈值会提高图像质量（IS），但是会降低多样性（FID先增后降）。较大的模型不适合截断，作者用正交正则化方法把G变平滑，则 z 就可以输出到好的空间中。作者还改进了正交正则化使其放松约束。

4.1 分析G对不稳定的影响

对于G，每个权重矩阵的前三个奇异值最重要。大多数层有良好的光谱，但是第一层表现不好，在模型崩溃时会发生爆炸。作者通过限制奇异值研究，发现不能确保稳定。

4.2 分析D对不稳定的影响

对于D，作者发现奇异值不会爆炸，但是会跳跃。作者通过梯度惩罚来平滑D，发现训练会变得平稳，但是图像质量(IS)会下降。最后作者发现D的LOSS会逐渐趋于0，但是在模型崩溃时会突然爆炸。作者又用testset验证了D发生了过拟合。

4.3 总结

模型崩溃不是D和G单独的原因，而是两者交互作用产生的原因。对D进行很强的约束可以确保训练稳定，但是会严重影响模型的性能。