

$$\text{BigGAN} = \left(\begin{array}{c} \text{大规模数据集和大规模网络结构} \\ \text{TT, SL, skip 2, Ortho} \end{array} \right) + \text{GAN}$$

G, D, 实际上是算力的胜利.

1. 在集群中用不同数量机器分别训练 G, D
2. 机器数量如何合理化分?
3. G, D 间有数据交互, 如何实现高效的数据交换!
4. 对 G, D 本身的分布式训练该如何考量?