Stack可译做堆叠，就是在GAN上面再放上一个GAN。

作者讲述自己的解决思路，原来难以生成高分辨率的图像，他们分解了这个问题，把生成高分辨率图片这个任务分解成两个更为简单的任务。就是文中说的一个GAN生成大致的形状和颜色，第二个GAN生成细节和修正错误，这比直接一个GAN生成更加容易。这似乎研究中解决一些任务的一种常用的方法，以前看的一篇论文，图像修复的，先修复缺失的轮廓边缘，再填充颜色。比直接一次完成好。就是把一个任务分成可以区分的两个任务，分别加以处理。

翻译：

从文字描述中生成具有照片般真实感的图片一直是计算机视觉中具有挑战性的一个课题，并且具有很多的应用。现有方法生成的图片可以粗略的反映出给定的描述，但是缺乏需要的细节和生动的对象局部。本文提出了stackGan，它从文字描述中生成具有照片般真实感的图像。1阶段的GAN根据文字描述草绘出一张简单形状和基本颜色符合的低分辨率图片，第二阶段的GAN把第一阶段的结果和文字条件作为输入，生成具有照片真实感的高分辨率的图片。二段GAN能够修正错误，并且通过增强过程给图片增加令人信服的细节。StackGAN生成的图片比其它方式生成的看上去更加可信。重要的一方面，这个模型可以生成256\*256的，其它模型最多知道了128.为了掩饰StackGAN的效果，我们基于CUB 和Oxford-102 数据集做了广泛的实验，这两个数据集包含了足够多的各种类别的对象，并且每类对象有很多的外观变体，被广泛用于图像生成分析。