**训练**

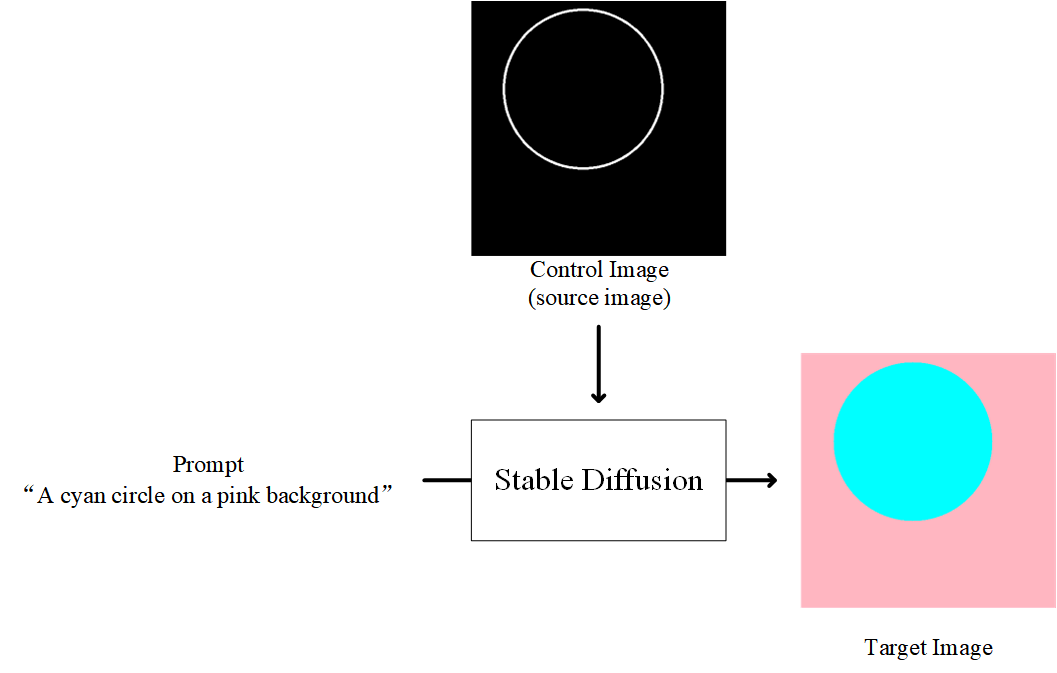
接下来是训练部分。训练需要的数据集结构如下：

Source/：源图片，相当于控制条件。对于线稿上色，就需要提供线稿图

Target/：目标图片，也就是ground truth 图片，原图

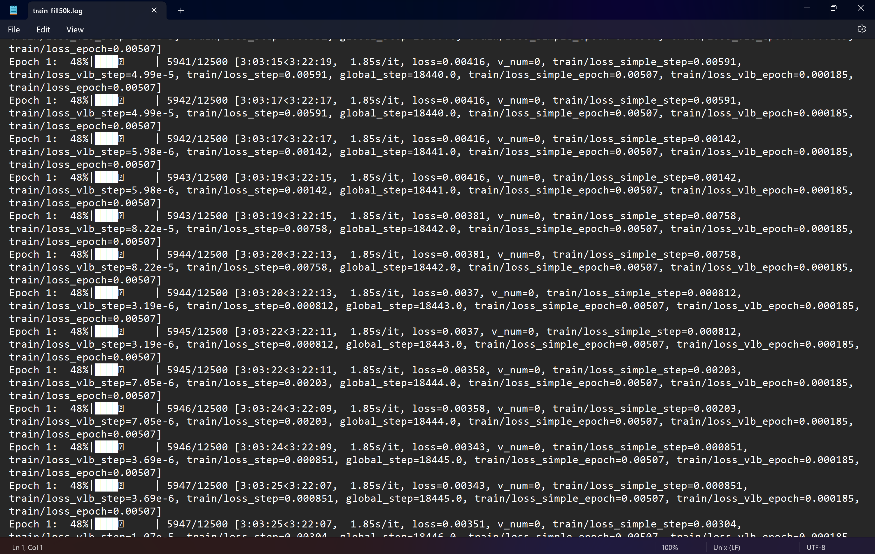
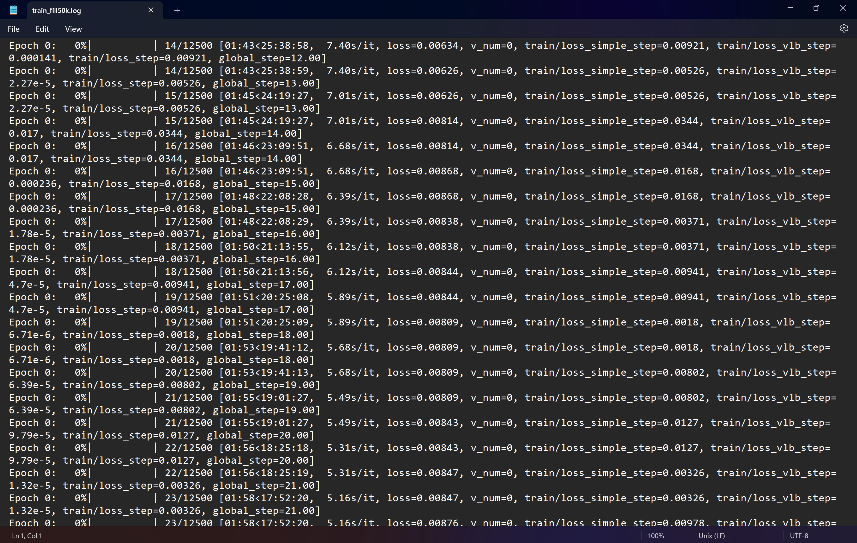
prompt.json：提示说明文字

我们想控制 SD 模型用颜色填充circle，Prompt包含我们target的一些文字描述。SD模型是已经训练好的，它已经知道提示文字的含义，但不知道“控制图像（源图像）”的含义。我们的目标就是让他知道控制条件的含义，然后生成期望的图片。



我们目前缺少合适的数据集，所以我们只能使用作者提供的测试数据集 fill50k，共5万个图片对。这个数据集不是专门针对Canny模型或者Lineart\_Anime模型的，并且同时我们训练的时长不够，所以最终训练的模型的效果非常不理想。

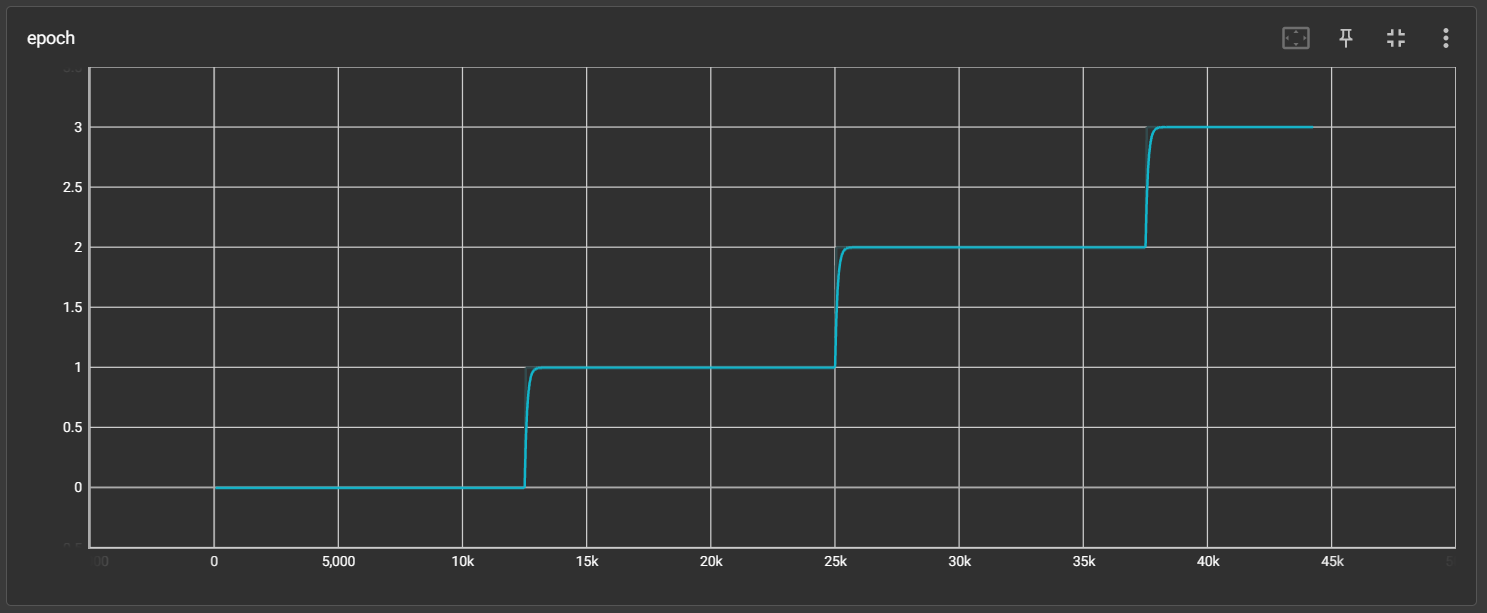
我们是在batch\_size=4，learning\_rate=10^5条件下进行训练的。训练前会将原来的SD模型（7个g）复制，并输出为一个包含了SD和ControlNet的模型文件control\_sd15\_ini.ckpt（5个g），训练时锁定SD模型，仅训练ControlNet部分。以下是我们训练中的部分过程输出，以及利用tensorboard得到的训练数据图像。

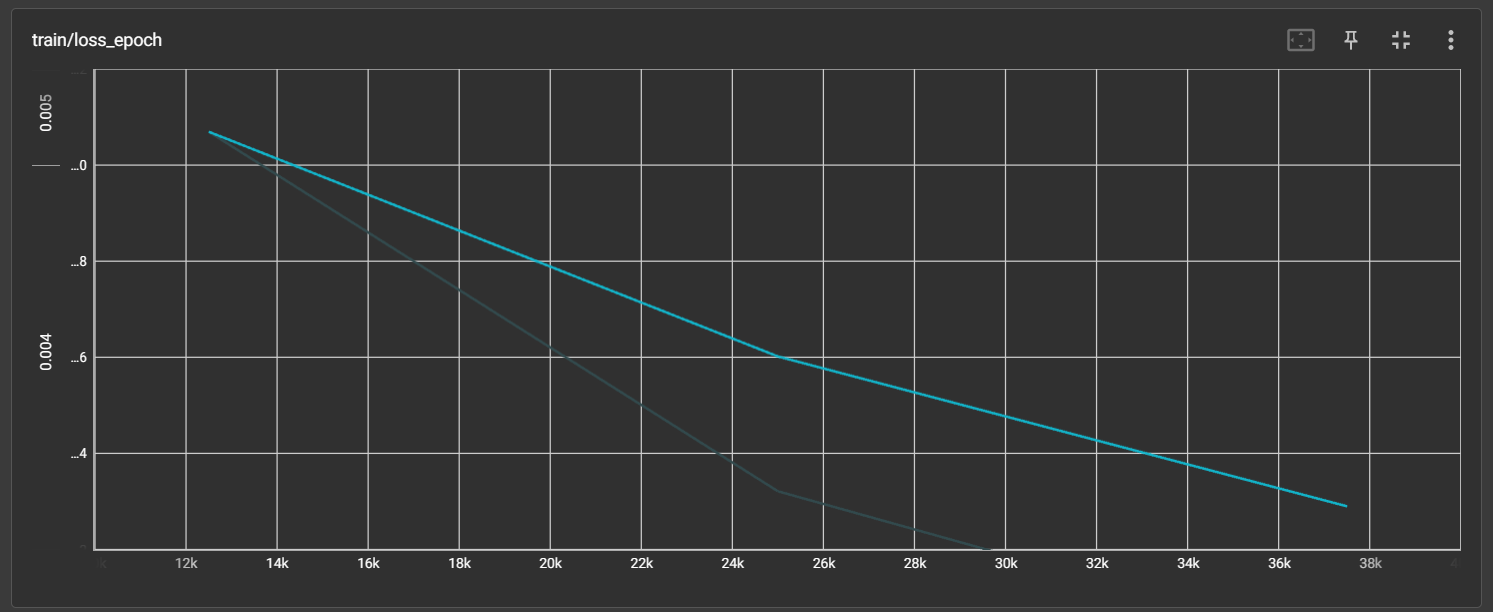


由于时间关系，我们只执行了3个半的epoch。（注意图中的数据的含义，epoch就是所有图片都执行了一次流程。Batch\_size就是一次处理4张图片。因为有50k的图，batch\_size=4，所以一次epoch会执行12.5k个steps）

分析图像：

（3个epoch）





（应该是上一步epoch的平均loss？可以发现…）

图表

描述已自动生成

（每一个step的loss变化。可以发现…）

最后是训练结果。可以看到整体效果很不理想。只能勉强看出大致的形状有一定的相似关系。找到合适的线稿数据集，并训练充分的话，应该是能够取得较好的效果的。