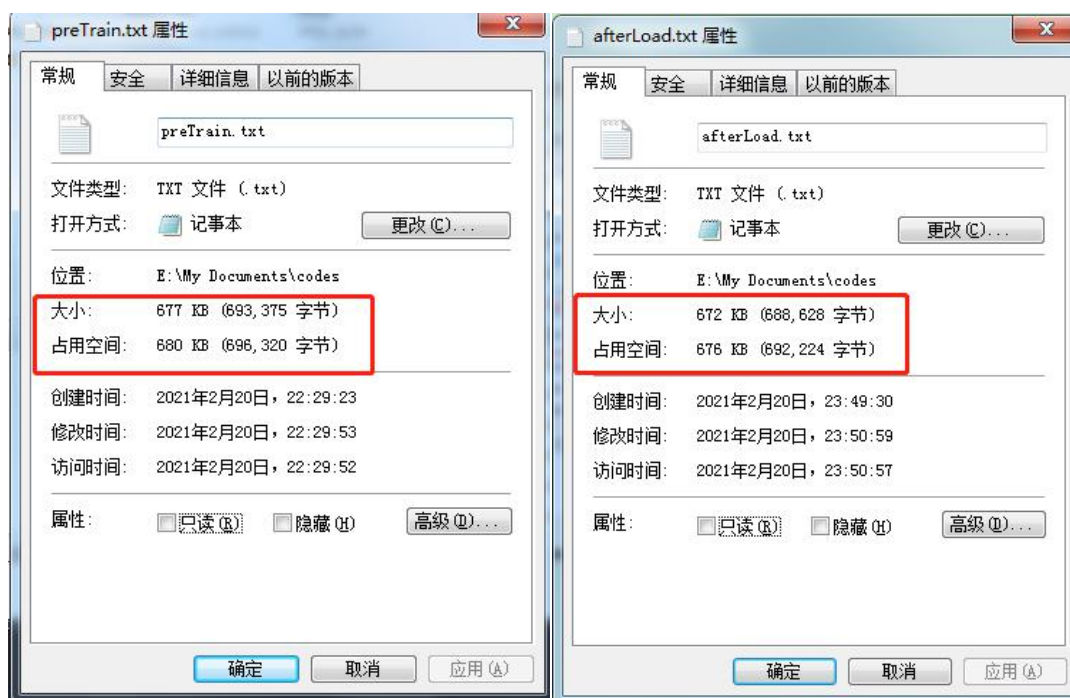


证据 1

pretrain 的参数和载入之后输出的参数两者文件的大小几乎一致,相差 4kb。



证据 2

通过打印每一层的名字,然后根据每一层的名字去查找两个打印文本中的值进行一一对比,发现共同有的层的参数都是相同的,唯一的差别就是 afterLoad 的模型参数中每一个卷积层, bn 层之后都会多一个 num_batches_tracked 层,然后通过查找资料,发现以下结论:

pytorch0.4.1 之后在 BN 层加入了 track_running_stats 这个参数,这个参数的作用如下:

训练时用来统计训练时的 forward 过的 min-batch 数目,每经过一个 min-batch, track_running_stats+=1

如果没有指定 momentum, 则使用 1/num_batches_tracked 作为因数来计算均值和方差(running mean and variance)。

链接: https://blog.csdn.net/qq_34132310/article/details/107384294

之所以会出现 pretrain 没有这个参数是因为预训练模型是在 pytorch0.3 情况下训练出来的,且这个参数并不会对训练产生什么影响。

证据 3

Lr=0.0001, 且他超参数都相同的情况下, 导入 (上图) 和没导入 (下图) 训练的结果曲线。

