



[首页](#)[资讯](#)[深度资源](#)[产业视频](#)[GMIS峰会](#)[AI 商用搜索](#)

[登录/注册](#)

使用Inception V3 实现指定类型图像识别

By [机器之心](#) 2017年1月24日 16:12

作者介绍：老刘(弦)，互联网码农，从业7年有余 曾就职于人人网，网信金融，京东金融。主攻互联网安全体系，大规模分布式系统，数据挖掘等方向，现在蒙特利尔进行学术深造。

最近一直在人肉审图片，现在用的模型也一直是人肉调节的，根据经验设置的weights，每天的误报率大概10% - 20%，好在总量并不大，肉眼五分钟内搞定。不过作为码农，有个多年养成的习惯：非常讨厌周而复始的干同一件事情。于是就把Tensorflow的Inception V3拿出来作为图像审核的第二道关卡。此文将介绍如何构建一个自己的图像识别程序，识别的准确度取决于样本类型和训练参数，鉴于我也没深挖，且已经够用，这篇不讲优化之类的，纯工程，照着操作就能出结果。

说实话，我并没看inception的结构图：



只是看了下它在ImageNet上的表现结果，觉得还不错，完全够用了，而且这个图像审核没有人脸或者物体识别那么高的要求，而且还需要一定的过拟合。

代码我已经从Tensorflow的model里拆出来，提交到了我自己的github: https://github.com/Cyber-Neuron/inception_v3

先说说里面的几个坑，完全是按照官方教程运行完后发现的，然后Google之，发现无数人进坑，然而却没什么答案，TF更新太快了。

1. 官方只给出了Training 和 Validation 的代码，并没有给出分类代码
2. 训练完成后的模型是checkpoint 文件，可以很方便的进行增量训练，但是用checkpoint文件去做预测的样例很少，而且没有固化.pb文件成功的案例可以参考。
3. 在上述问题搞定后，发现作者为了提升Training和Validation的速度，引入了多线程机制，但是，但是，但是，没有考虑到123456经过多线程输出会变成562314这种序列，而且完全随机，这就意味着预测结果和文件名不能一一对上，Validation用的是统计结果，并不关心预测结果的顺序T_T，但是如果做预测，顺序乱了还有个毛线用啊！
4. 需要修改官方源码让它支持同时输出文件名和预测结果

5. 集群上没有装Cuda 8.0，我无法用GPU+分布式训练模型，一颗CPU跑20个小时才完成2000次迭代，幸运的是，已经过拟和了，达到了使用要求

上述几个问题，完全是把原理弄明白了以后发现的，此时再想换别的方案已然不值得，因为时间已经消耗在上面了，迁移成本略高（这就是为啥各家都在开源自己的深度学习框架，要不然大家都用Google，都去diy它的模块，其它家就死了）

目前俺这个版本算是网上第一个完整填满坑且能实用的版本(use inception v3 to retrain your own model的例子，教程里从未给出预测的代码)。

```
2017-01-23 20:47:05.912472: starting evaluation on (train).  
2017-01-23 20:47:29.402914: precision @ 1 = 1.0000 recall @ 5 = 1.0000 [1000 samples]
```

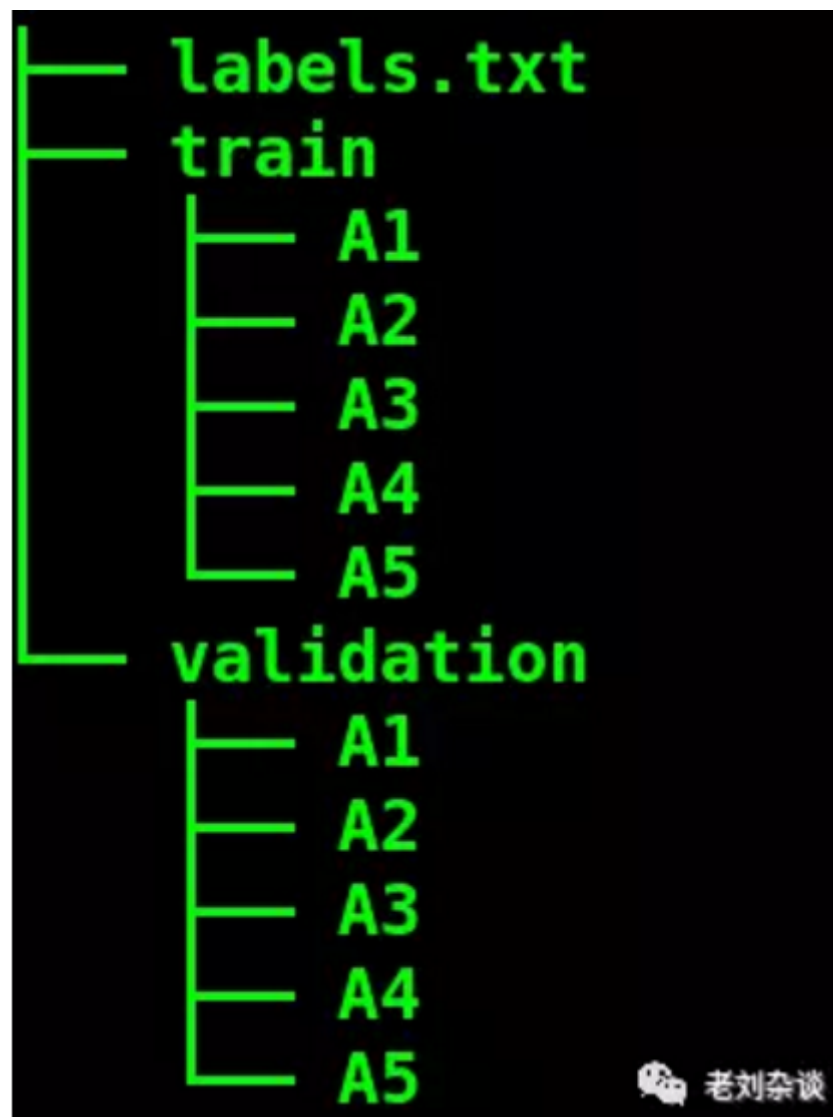
可以看到evaluation的结果已经100%的准确度了，这也就意味着过拟合，过拟合并非坏事，在特殊情况下可以很好的完成找相同物体的任务。

首先安装Bazel和Tensorflow,这些在官方的教程里面都有，不写了。

要想完成增量训练，首先准备好训练数据，目录格式如下：

```
$ less sample.tar.gz  
rw-r--r-- worker/worker 823 2017-01-21 fish2.0/A1/953275.jpg  
rw-r--r-- worker/worker 823 2017-01-21 fish2.0/A2/1067563.jpg  
rw-r--r-- worker/worker 823 2017-01-21 fish2.0/A3/2353330.jpg  
rw-r--r-- worker/worker 823 2017-01-21 fish2.0/A4/1664412.jpg  
rw-r--r-- worker/worker 823 2017-01-21 fish2.0/A5/1254412.jpg
```

目录A1 - 5是类别，里面的jpg是训练数据。准备好数据后执行 sh preprocess.sh和 sh build_data.sh，执行完毕后会生成一个raw-data的目录，里面有自动生成的training 和 validation的文件夹，其中labels.txt对应A1-A5五个分类。



还会生成便于TF读取的文件

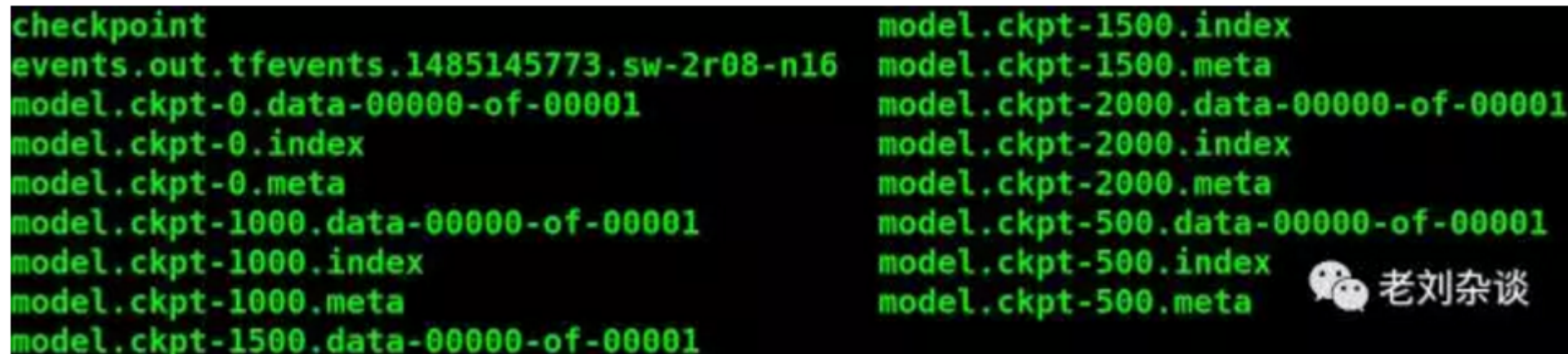
```
train-00000-of-00002  validation-00000-of-00002  
train-00001-of-00002  validation-00001-of-00002
```

接下来下载已经训练好的模型（可以参考相关教程<https://github.com/tensorflow/models/tree/master/inception>），由于我们是增量训练，所以节省了很多时间。

执行sh train.sh 经过若干个小时的等待（不知为啥GPU集群的文件系统出了问题，ls下都是龟速，所以我只能等待了漫长的20多个小时才得到了满意的模型）

完成后得到了如下图的文件：

```
checkpoint
events.out.tfevents.1485145773.sw-2r08-n16
model.ckpt-0.data-00000-of-00001
model.ckpt-0.index
model.ckpt-0.meta
model.ckpt-1000.data-00000-of-00001
model.ckpt-1000.index
model.ckpt-1000.meta
model.ckpt-1500.data-00000-of-00001
model.ckpt-1500.index
model.ckpt-1500.meta
model.ckpt-2000.data-00000-of-00001
model.ckpt-2000.index
model.ckpt-2000.meta
model.ckpt-500.data-00000-of-00001
model.ckpt-500.index
model.ckpt-500.meta
```



这些文件是训练好的模型，可以用来做评估了。

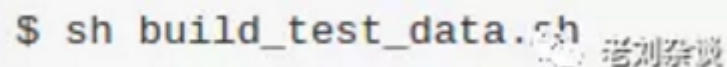
执行sh eval.sh即可得到评估结果。

以下便是俺diy的部分了：

1. 生成用于预测的机读数据，
2. 完成预测

准备好测试数据，放入上述结构的目录

```
$ sh build_test_data.sh
```



然后就可以得到机读文件

```
$ ls  
test-00000-of-00001
```

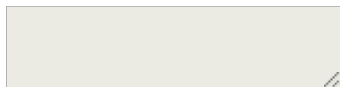
然后执行预测指令，可以看到分类的结果

```
$ sh predict.sh 10  
2017-01-23 18:16:54.551870: starting evaluation on ().  
1167803.jpg 2  
1343160.jpg 1  
2369878.jpg 5  
1343563.jpg 1  
1210248.jpg 5  
1220936.jpg 5  
2241516.jpg 5  
1440936.jpg 1  
1265147.jpg 1  
2305191.jpg 1
```

此刻心情格外爽朗 哈哈,匆匆记下来，防止以后忘了。

声明：本文由机器之心经授权转载自微信公号《老刘杂谈》，作者：老刘(弦)，禁止二次转载。

[图像识别工程GitHub计算机视觉](#)



[提交评论](#)

登录后参与评论[去登录](#)



[关于我们寻求报道商务合作](#)

©2017版权所有 机器之心（北京）科技有限公司

京 ICP 备 12027496

全球人工智能信息服务

友情链接

[Synced Global](#)[机器之心](#) [Medium](#) [博客PaperWeekly](#)[网易智能动脉网](#)



联系电话：+86 010-57150141

联系邮箱：contact@jiqizhixin.com