# 使用faster rcnn训练自己的数据

## 数据准备

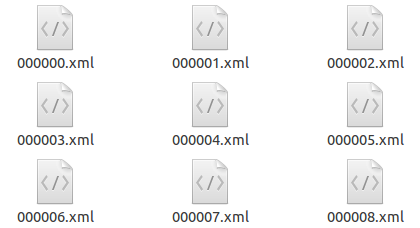
原版的faster rcnn是使用VOC2007数据集进行训练的，为方便起见，将自己的数据集组织成VOC2007的数据格式。

基本的形式如下，在VOC2007文件夹下包括3个文件夹，Annotations、ImageSets、JPEGImages：



### 1.1 Annotations目录

Annotations目录下存放标注文件，以xml文件格式存储：

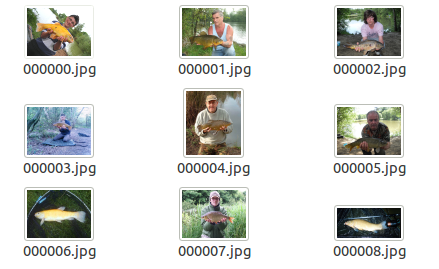


以某一个xml文件为例，包括以下信息：



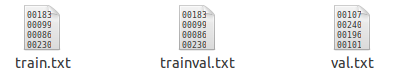
### 1.2 JPEGImages目录

保存图像，文件名与标注文件对应，格式为jpg，不是jpeg：

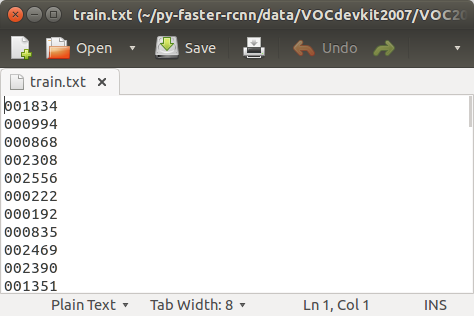


### 1.3 ImageSets目录

该文件夹下保存训练、验证和测试用的图像列表文件，对于faster-rcnn训练，用到的有train.txt、trainval.txt和val.txt文件：



内容形式如下：



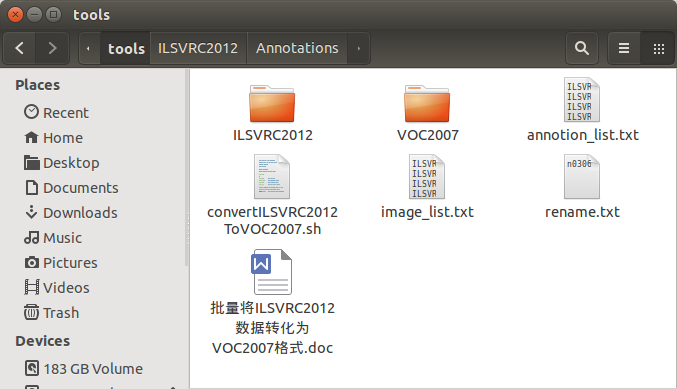
### 1.4 ILSVRC2012ToVOC2007工具

ILSVRC2012数据集包括了1000个类别的图像数据，可以使用该数据集进行训练数据，但是该数据集的组织形式和标注格式与VOC2007数据集有所差别，因而不能直接用来训练。本人编写了一个ILSVRC2012ToVOC2007工具，可以方便的从ILSVRC2012数据集转化到VOC2007数据格式，以下介绍该工具的使用。

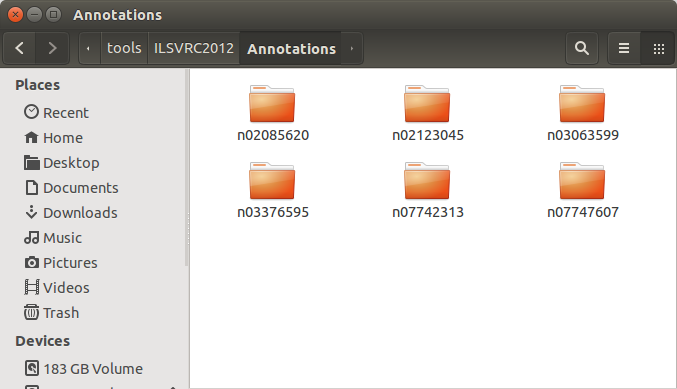
ILSVRC2012ToVOC2007文件夹下包括两个文件，分别保存工具的源码和编译后生成的可执行文件：

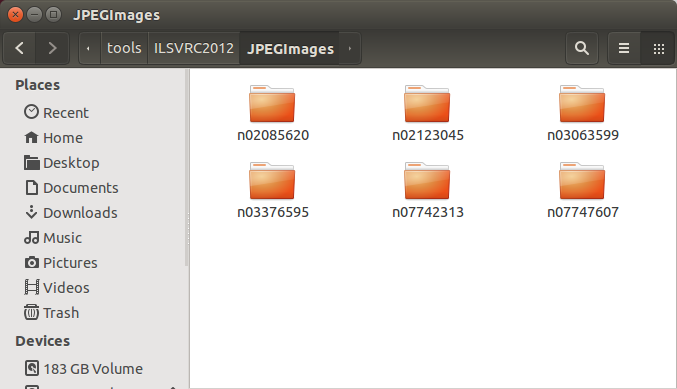


tools文件夹下包括以下文件及文件夹：



将ILSVRC2012中的图像文件和标签文件拷贝到ILSVRC2012/Annotations目录和ILSVRC2012/JPEGImages目录下，图像和标签一一对应。



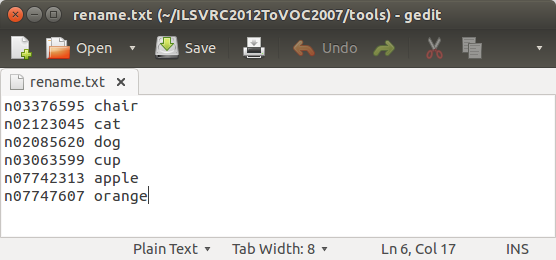


制作标签列表文件和图像列表文件，分别列出每个图像和对应标签的路径：

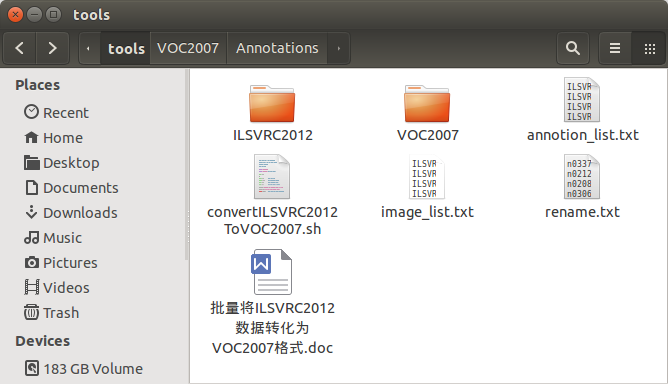




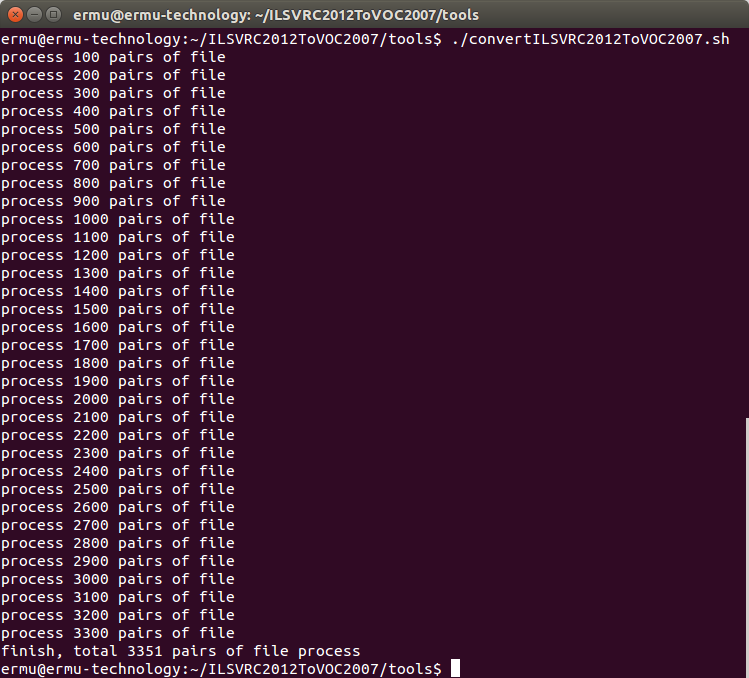
制作重命名文件，因为ILSVRC2012中，并没有明显的标记出物体的名称，而是以编码代替，为了更加直观，我们制作一个重命名文件，将其转换为物体的名称：



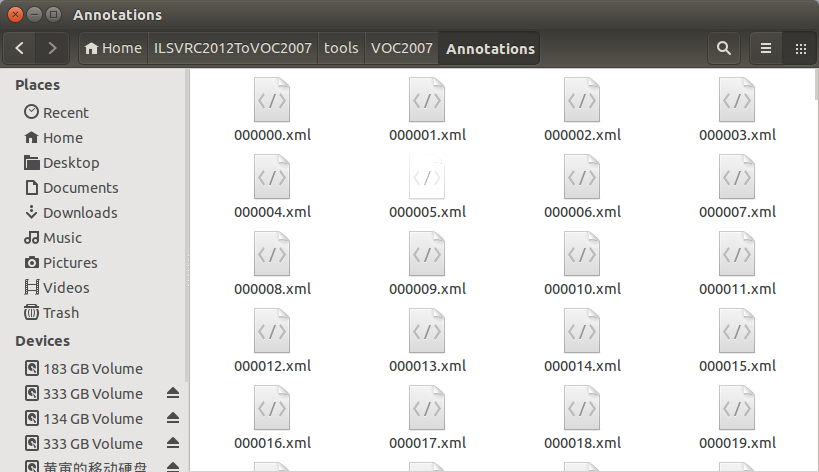
最后tools文件夹下的文件如下：



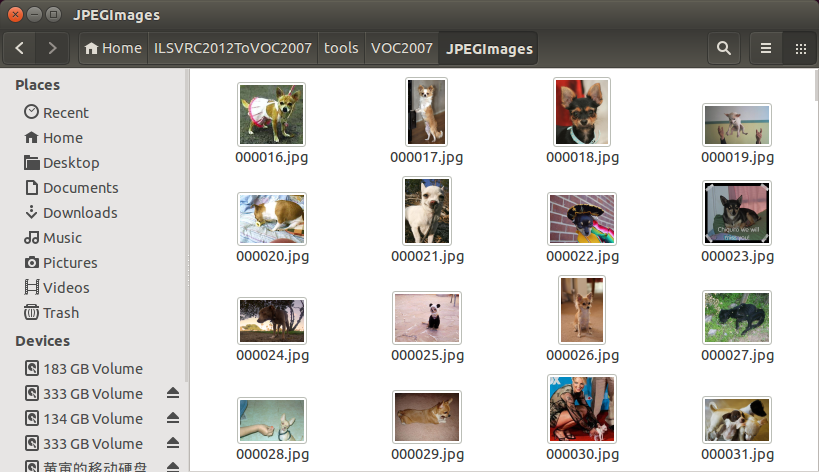
执行./convertILSVRC2012ToVOC2007.sh命令：



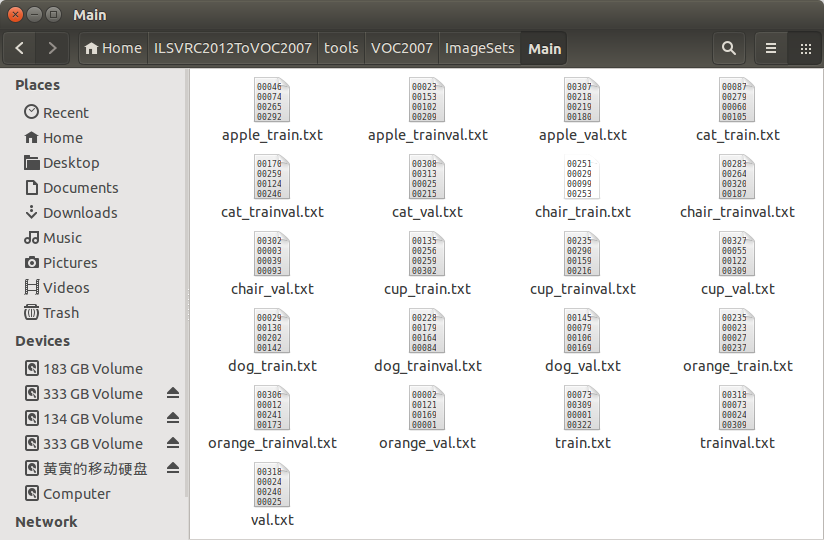
VOC2007/Annotations目录下：



VOC2007/JPEGImages目录下：



VOC2007/ImageSets/Main目录下：



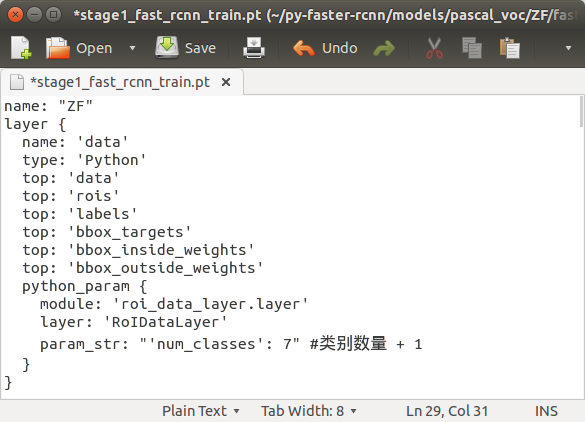
到这里，VOC2007格式的训练数据就生成了，将VOC2007拷贝到py-faster-rcnn/data/VOCdevkit2007目录下。

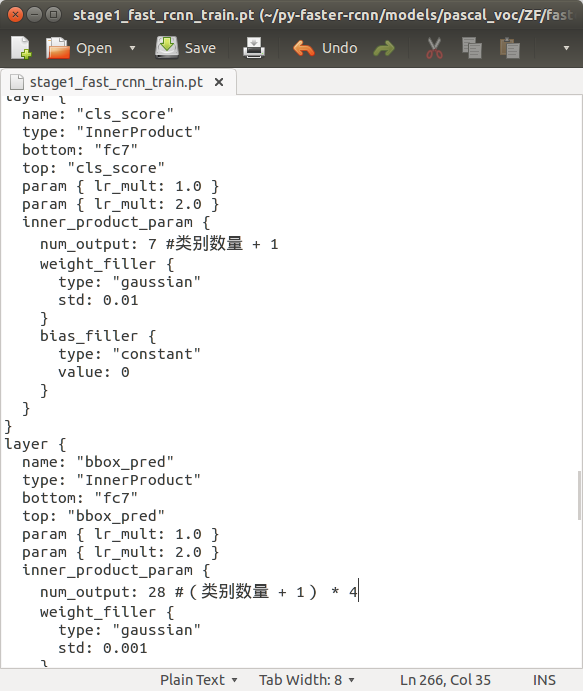
## 修改必要文件

根据我们需要检测的类别数量，对官方版本的部分文件进行必要修改，例如，这里我们需要检测6个类别，那么，加上背景，共有7类。另外，对于每个类要预测4个坐标点，则预测的坐标点数量为28个。

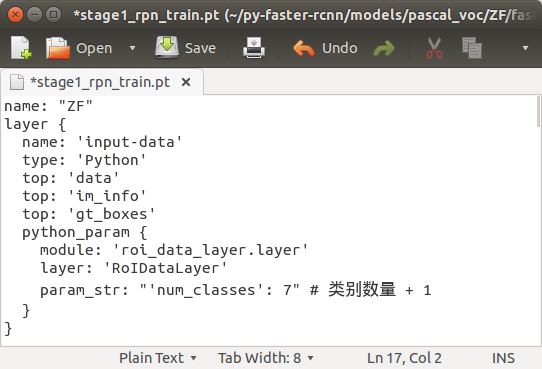
### 2.1 models/pascal\_voc/ZF/faster\_rcnn\_alt\_opt目录下：

1. stage1\_fast\_rcnn\_train.pt文件：

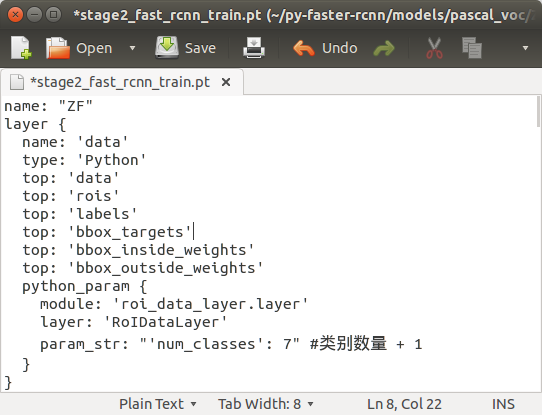


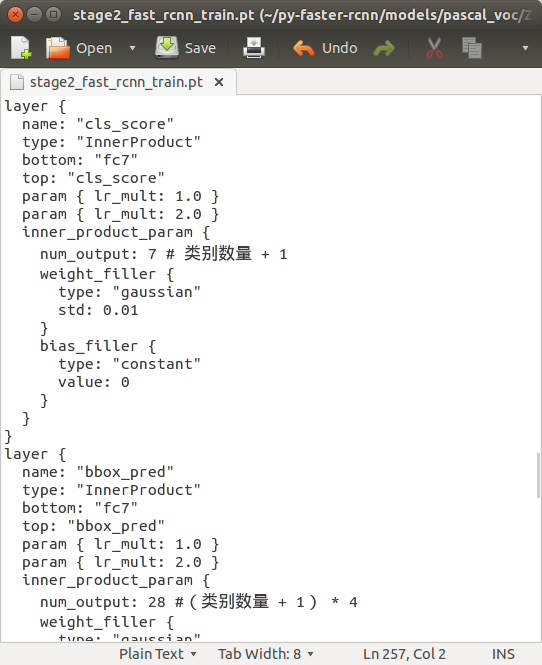


1. stage1\_rpn\_train.pt文件：

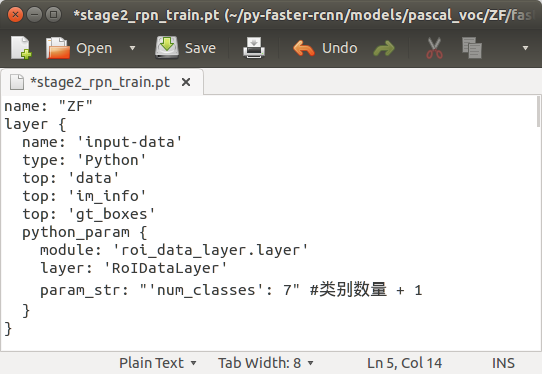


1. stage2\_fast\_rcnn\_train.pt文件

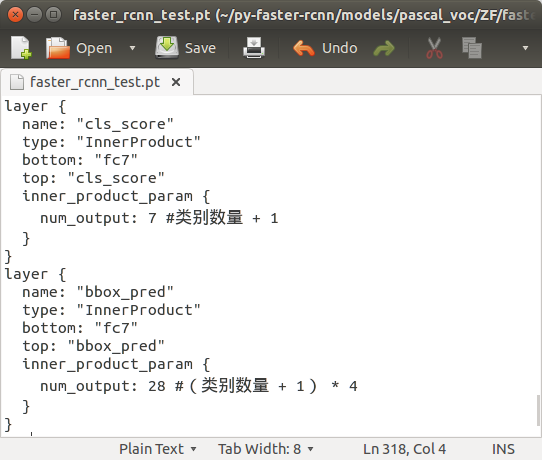




1. stage2\_rpn\_train.pt文件



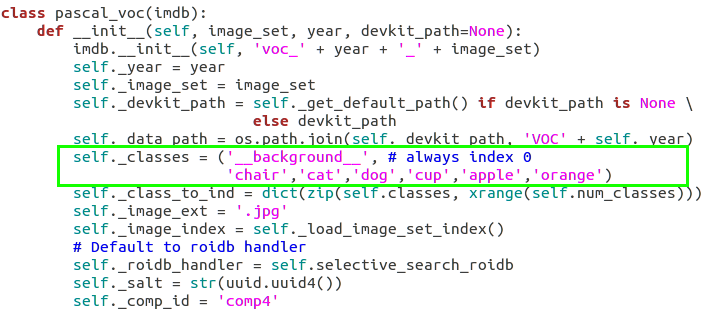
1. faster\_rcnn\_test.pt文件



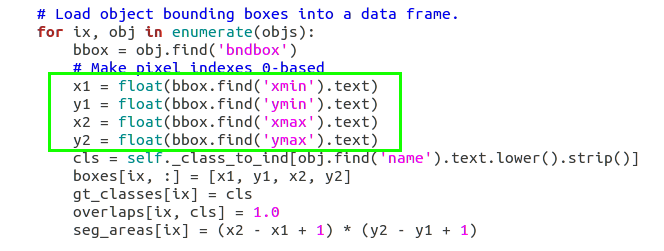
### 2.2 py-faster-rcnn/lib/datasets目录下：

1. pascal\_voc.py

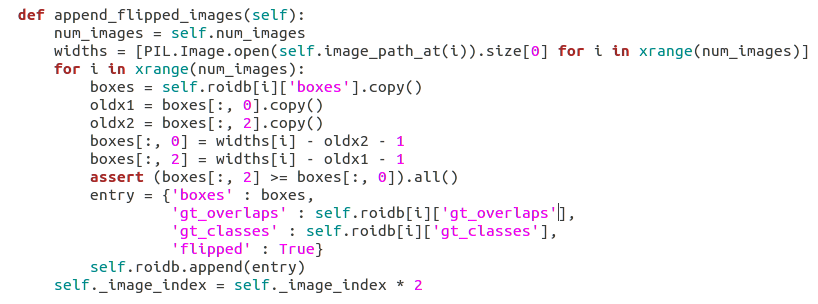
把类别名改成自己的类别名称，这里不需要注意各个类别的顺序，随便排列即可。



另外，由于在原始的VOC2007数据中，候选框的坐标是从1开始的，而不是0开始的，所以在程序中，有一个将候选框各坐标减1的操作，而这里，我们使用的是ILSVRC2012中的数据，它对候选框的标注坐标是从0开始的，因此，我们要把减1操作去掉：



1. imdb.py



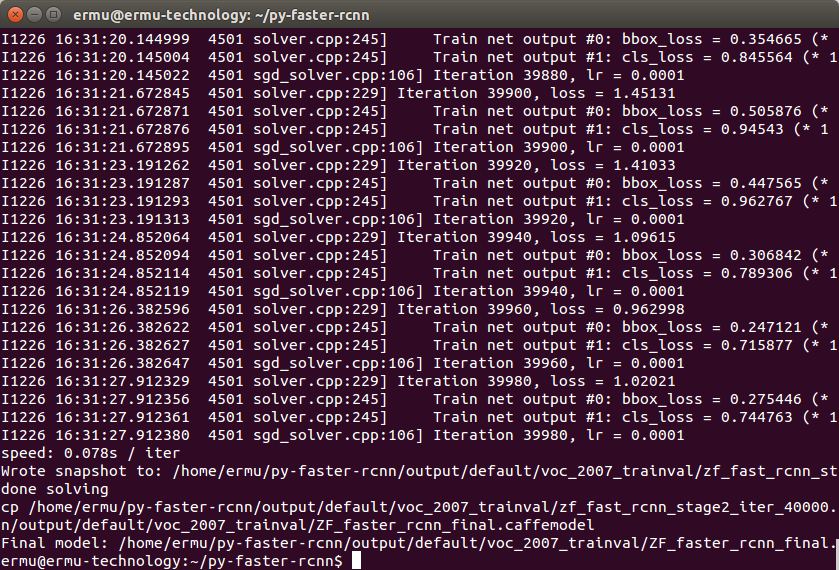
## 训练

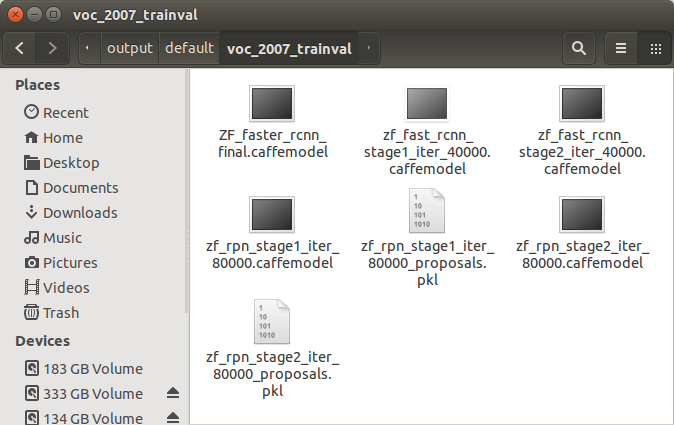
确保py-faster-rcnn目录下没有output文件夹，若有，将其改名或删除。

确保py-faster-rcnn/data目录下没有cache目录，若有，将其改名或删除。

在py-faster-rcnn目录下执行：python ./tools/train\_faster\_rcnn\_alt\_opt.py --gpu 0 --net\_name ZF --weights data/imagenet\_models/ZF.v2.caffemodel --imdb voc\_2007\_trainval --cfg experiments/cfgs/faster\_rcnn\_alt\_opt.yml

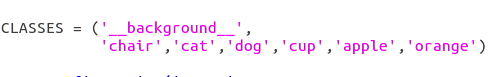
训练大概花了6个小时，训练结果保存在output目录下。





## 测试

将训练结果文件ZF\_faster\_rcnn\_final.caffemodel拷贝到py-faster-rcnn/data/faster\_rcnn\_models文件夹下，并修改py-faster-rcnn/tools文件夹下的demo.py文件，将类别名称改成自己的类别：



并根据测试图像，修改相应的图像名称，在tools目录下，执行 python demo.py



