

## 读书笔记 2 用一个预训练模型提取特征

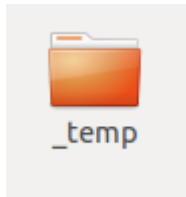
2014.7.21 薛开宇

本学习笔记的作用在于为后面打基础，没有什么实际的东西可以观测到，要可视化特征还要观看后面的教程。

### 2.1 制作一个数据库

先做一个临时文件夹存放东西。

```
mkdir examples/_temp
```



我们为两张在 images 文件夹的照片生成一个文件列表（默认为一张图片，cat1 是我随意加上去的）

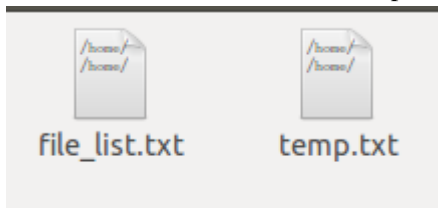


```
find `pwd`/examples/images -type f -exec echo { } \; > examples/_temp/temp.txt
```

我们将使用 imagedatalayer 预计标签之后的每一个文件名，所以让我们添加一个 0 到每一行的末尾

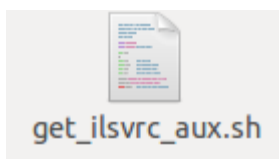
```
sed "s/$/ 0/" examples/_temp/temp.txt > examples/_temp/file_list.txt
```

这样，我们将得到两个文件 temp.txt 和 file\_list.txt。



### 2.2. 定义特征提取网络结构

在实践中，从一个数据集中减去均值图像对于提高分类准确性很重要，因此从 ILSVRC dataset 中下载均值图像数据库



```
data/ilsvrc12/get_ilsvrc_aux.sh
```

我们将使用其中的 data/ilsvrc212/imagenet\_mean.binaryproto 去定义网络结构。

将定义结构的文件 cope 到我们的临时文件夹。

```
cp examples/feature_extraction/imagenet_val.prototxt examples/_temp
```

然后，我们进入 `imagenet_val.prototxt` 更改路径。更改其中 `$CAFFE_DIR` 的地方  
下图是我做的改动。

```
image_data_param {
  source: "/home/xuekaiyu/caffe-master/examples/_temp/file_list.txt"
  mean_file: "/home/xuekaiyu/caffe-master/data/ilsrvrc12/imagenet_mean.binaryproto"
```

### 2.3.提取特征

执行指令

```
build/tools/extract_features.bin          examples/imagenet/caffe_reference_imagenet_model
examples/_temp/imagenet_val.prototxt fc7 examples/_temp/features 10
```

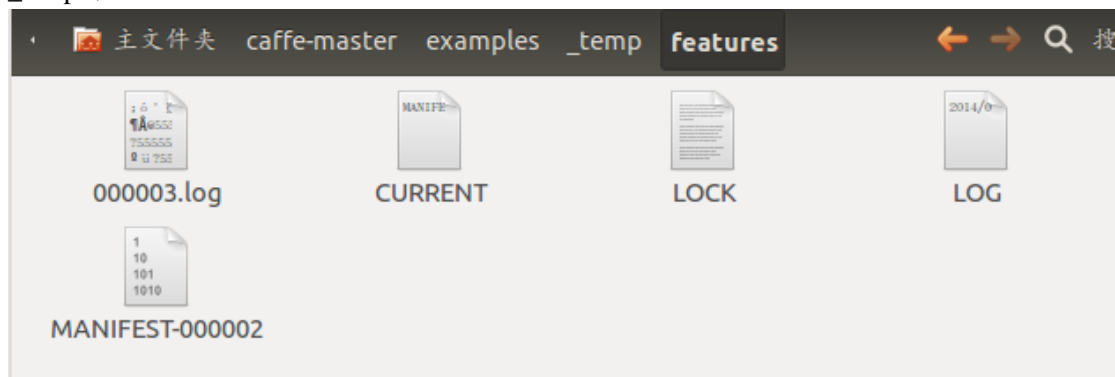
其中 `fc7` 是最高层的特征，我们也可以使用其他层提取，像 `conv5` 或 `pool3`

最后的参数是数据的批次

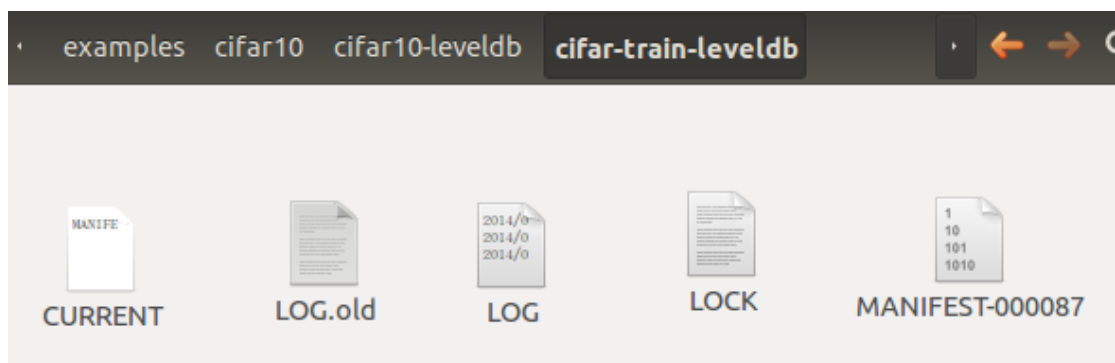
特征保存在 LevelDB `examples/_temp/features`，可以运用到其他代码了。

这里可以发现，特征文件中 5 个文件和 `cifar10` 中的 `leveldb` 是相似的。

`_temp` 下 `features`



`Cifar10` 下特征数据库



因此，这是我们训练时组建模型时必要的文件。

### 2.4.注意

当存在 `features` 文件夹时将出现错误，这时需要移除该文件夹。

资料来源: [caffe 官网教程](#)