**一、准备工作**

* **/home/zr/data/VOCdevkit/下新建文件夹 MyWork**

1. **MyWork**文件夹下新建 Annotations 、ImageSets、 JPEGImages 、 Labels 四个文件夹；其中 JPEGImages 和 Labels 两个文件夹内各自有子文件夹 1 ，用于LabelTool标记时候，输入到Image Dir：栏目下，即可导入 JPEGImages/1 文件夹中仍未进行标记的图片，标记后的标签文件同步保存在Labels/1 文件夹下，txt文件命名与图片名称相同。
2. ImageSets 文件夹下包含 Main 子文件夹。

* **/home/caffe/data/下新建文件夹 MyWork**

1. 该文件夹下包含create\_data\_list.sh 、create\_data\_crew.sh 和 labelmap\_crew.prototxt三个文件，根据需求进行适应性修改。

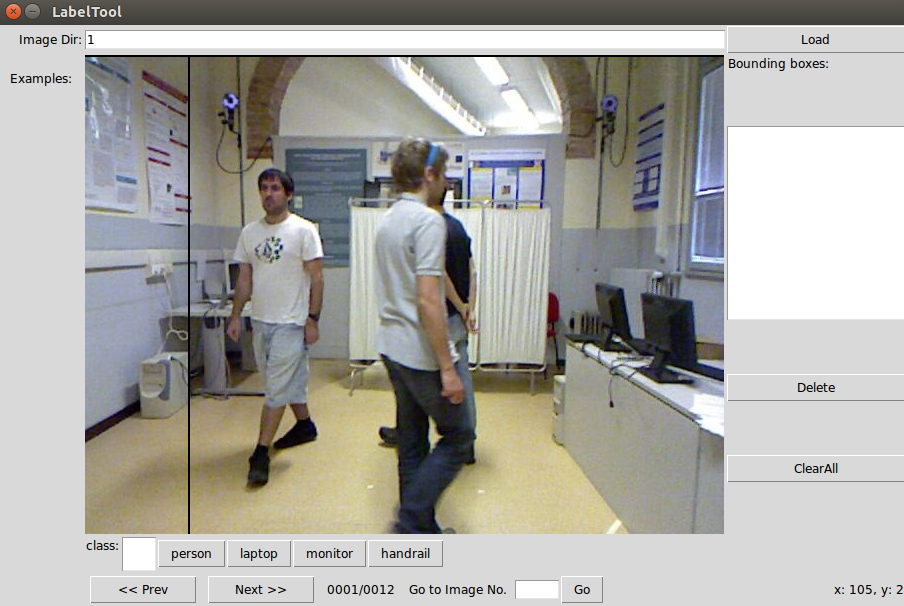
**二、数据集的标注**

1、使用路径/home/zr/data/VOCdevkit/MyWork目录下的数据集标注工具SSD\_CrewDetect.py来实现；

2、切到所在路径/home/zr/data/VOCdevkit/MyWork目录下，执行指令

python SSD\_CrewDetect.py

可打开标注软件如下图：



在Image dir下输入1，点击 load，完成未标注的图片的加载。

然后选择class为person，再进行画框，画完后，点击Next>>则完成保存，继续后续未标注图片的标注。

注意：

（1）先选择类别，再画框；

（2） 一张图片标注完后, 点击Next>>按钮, 标注下一张图片,  图片label成功后,会在Labels对应目录下生成与图片文件名对应的label文件，如果点击Prev>>则会将之前标注好的图片清空，则需要重新标注；

（3）如果目录中的图片已经标注过,点击load时不会被重新加载进来；方便中断，下次可继续标注未标注图片。

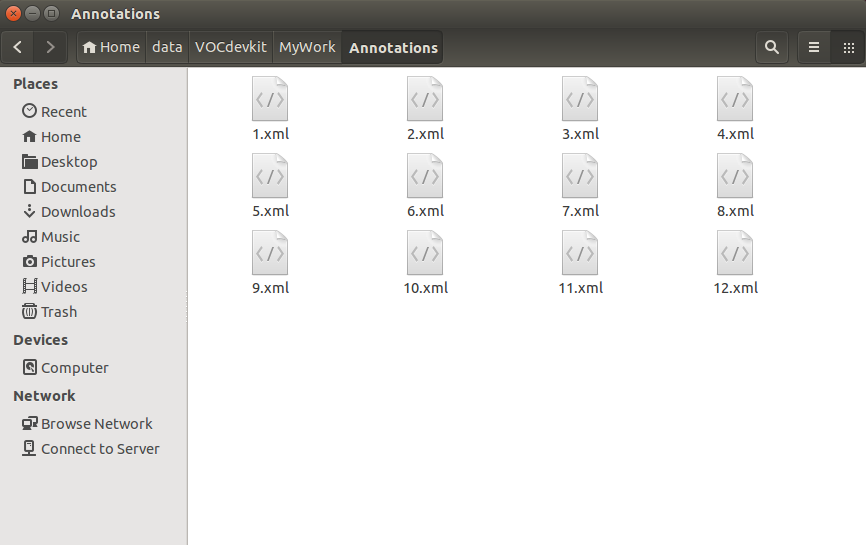
生成的txt文档存储的标签内容格式为：

object\_number  
className x1min y1min x1max y1max  
classname x2min y2min x2max y2max  
...

**三、数据集转换为xml格式**

1、使用路径/home/zr/data/VOCdevkit/MyWork目录下的createXml.py程序来实现；

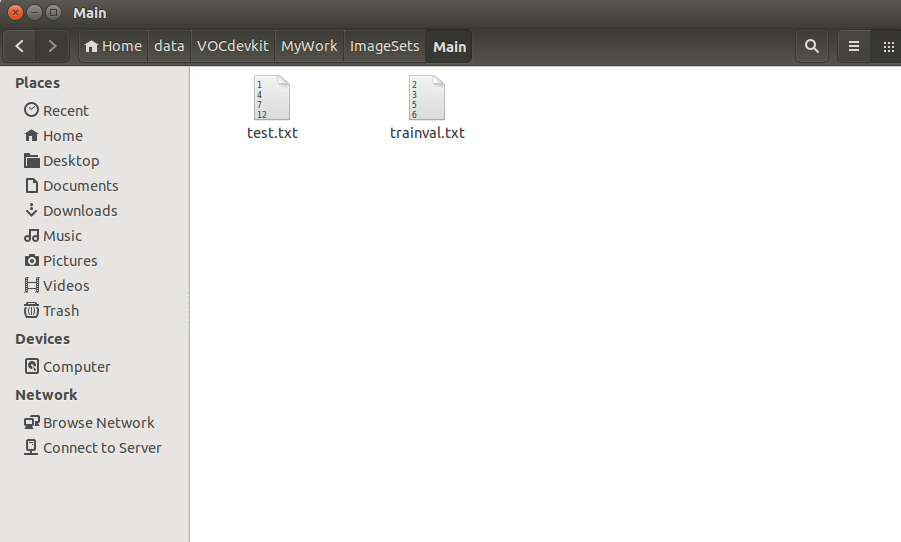
2、切换到/home/zr/data/VOCdevkit/MyWork目录下，执行指令./createXml.py实现数据转化为xml格式，可见在/home/zr/data/VOCdevkit/MyWork/Annotations目录下生成了对应的.xml格式文件。



**四、生成训练集和测试集标识文件**

1、使用路径/home/zr/data/VOCdevkit/MyWork目录下createTest.py程序来实现；

2、切换到/home/zr/data/VOCdevkit/MyWork目录下，执行指令./createTest.py 1 3250 800实现训练、测试数据集标识文件生成，可见在/home/zr/data/VOCdevkit/MyWork/ImageSets/Main目录下生成了标识文件。命令参数中，1表示起始图片序号，3250表示结束图片序号，800为测试集的图像数目。

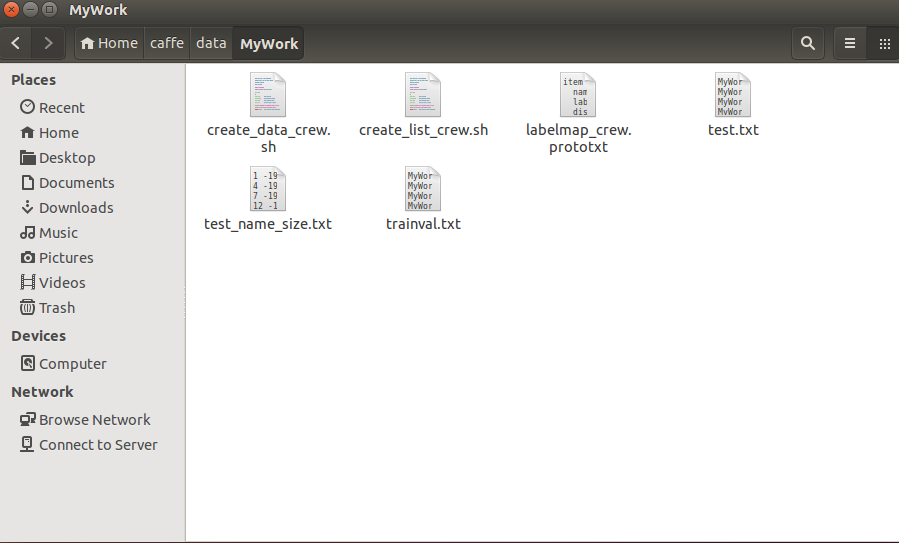


**五、生成lmdb数据集**

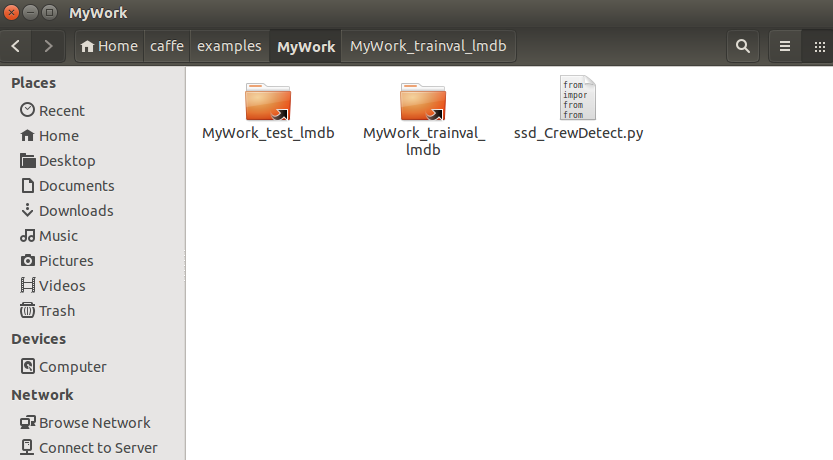
1、使用/home/zr/caffe/data /MyWork下的create\_list\_crew.sh 和 create\_data\_crew.sh 程序实现；

2、使用路径/home/zr/data/VOCdevkit/MyWork路径：

（1）先运行./create\_list\_crew.sh指令，在/home/zr/caffe/data/MyWork目录下生成了trainval.txt 、 test.txt 、test\_name\_size.txt 三个文件。



（2）再运行./create\_data\_crew.sh指令，在/home/zr/data/VOCdevkit/MyWork/lmdb目录下生成了MyWork\_trainval\_lmdb 文件夹，里面包含了data.mdb 、 lock.mdb两个文件；生成了MyWork\_test\_lmdb文件夹，里面同样包含了data.mdb 、 lock.mdb两个文件；同时在/home/zr/caffe/examples/MyWork下生成了lmdb数据集的链接（不一定需要，因为训练网络中数据集可直接地址定位到/home/zr/data/VOCdevkit/MyWork/lmdb地址下）。



**六、修改labelmap\_crew.prototxt文件**

修改后为：

item {

name: "none\_of\_the\_above"

label: 0

display\_name: "background"

}

item {

name: "person"

label: 1

display\_name: "person"

}

注意：labelmap\_crew.prototxt文件中使用的名称，应该和数据集标注工具SSD\_CrewDetect.py使用的标签名称完全一致，包括大小写！

**至此，完成自己数据集制作，可用其进行网络的训练应用！**