**实验目的：**

实现基于AlexNet的RCNN进行图片中的目标识别。

**实验原理：**

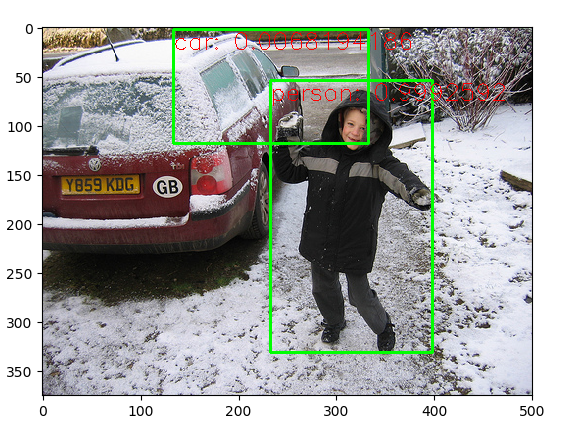
RCNN的过程要点可以简单的表示为：

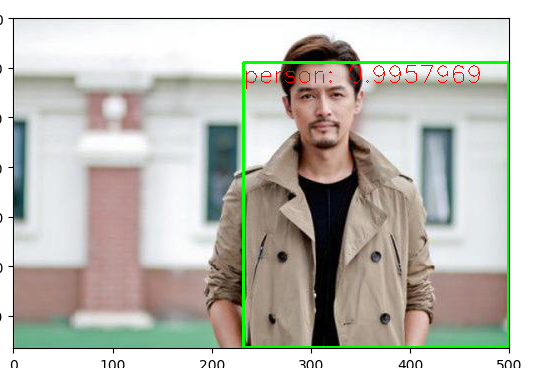
* 利用selective search方法对输入图像进行region of interest ROI提取
* 对每个候选区域，使用深度网络提取特征
* 特征送入每一类的SVM 分类器，判别是否属于该类
* 使用回归器精细修正候选框位置

具体可参考提供的参考资料《RCNN目标识别》或者伯克利大学的Ross Girshick等人于2014年CVPR上的经典论文《Rich feature hierarchies for Accurate Object Detection and Segmentation》。

**输入：**

本次练习的输入是一张图片，输出是进行目标识别之后，对各类目标加上边框并且进行分类识别。





**任务：**

在这个练习中，你将实现基于AlexNet的神经网络，以对图片进行目标识别。

AlexNet是在ImageNet 2012比赛中的最优网络，R-CNN灵感就来源于它。

* alexnet.py: AlexNet网络结构的定义。
* finetune.py: AlexNet微调主程序。
* datagenerator.py:加载用于微调的训练集和验证集。

(以上三个文件是改进AlexNet时需要用到的代码文件，下载地址：https://github.com/kratzert/finetune\_alexnet\_with\_tensorflow)

微调使用的数据集是pascalvoc\_2012，下载地址：http://host.robots.ox.ac.uk/pascal/VOC/voc2012/VOCtrainval\_11-May-2012.tar

**需要实现以下代码：**

* tool.py:非极大值抑制(NMS)、交并比(IOU)算法实现、selective\_search接口调用。
* ckpt2npy.py:将finetune.py生成的校验点(CKPT)转化为可供numpy模块加载的权重矩阵。
* generate\_finetune\_data.py:解析VOC PASCALL 2012数据集，生成微调alexnet的训练集和验证集。
* train\_svm.py:训练用于分类的SVM分类器。
* demo.py:模型测试

**提示：**

RCNN的主要缺点包括训练速度慢和检测速度慢，而且很耗内存。可以通过仅训练VOC数据中的少数类别（如人、车等）来减少计算量。在训练时需要使用GPU，请训练好Model保存以便于检查。

因为selective search在框的生成和选择上有一定随机性，所以多次运行时结果会不同。

**代码检查标准：**

能够以较大正确概率进行目标识别。