字符分割

OCR技术：照片光学字符识别

步骤：

1、输入图片

2、Text detection扫描图片，找到有文字的图像区域

3、Character segmentation 字符分割

4、Character classification 字符分类

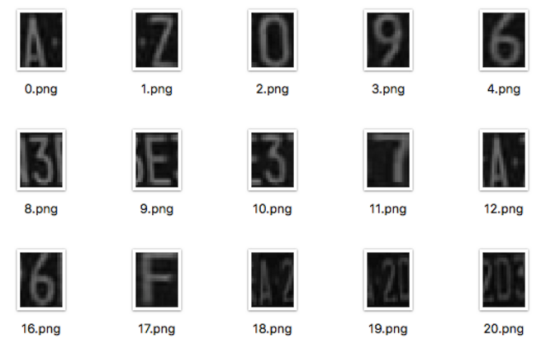
车牌识别(字符分割)：

我们之前采用的是垂直投影的方法对字符进行分割，其原理是将二值化图像上的白色像素点垂直累加起来，做出分布直方图，从而找到相邻字符的分界点进行分割。

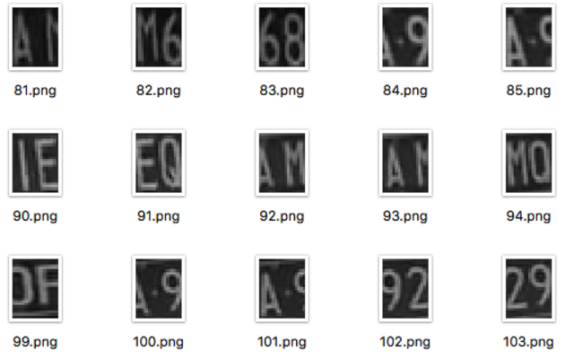


但是对于一些中文字符，分割的情况很不理想，如：川字等，所以我们对方法进行了改进。 使用基于滑动窗口的分割方法。利用一个训练好的正负样本分类器来进行在图像上滑动来获取最佳分割路径。将样本分为下面三类放到CNN模型中进行训练：

正样本：



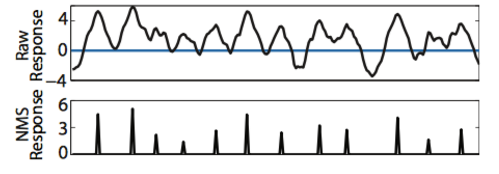
负样本：

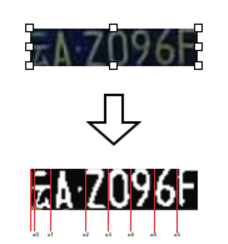


中文字符：



训练完得到一个三分类的分类器，我们把它运行在如下的车牌中，用滑动窗口进行扫描，使得第一个字符的中文概率+后6个字符的正样本概率最大，这个最大的位置即为最佳分割位置。





最后再将分割好的字符和数字放到字符识别的模型中进行识别。