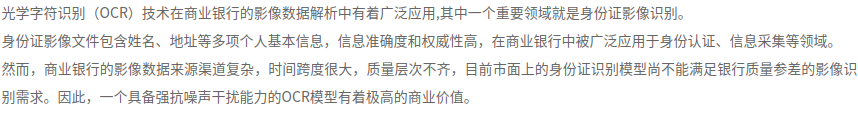
**毕业设计论文进度汇报（2.10-2.16）**

姓名：高尚 学号：20161104492

1. **目前论文的进度**

论文题目为：《基于OCR的身份证要素提取设计与实现》

[题目背景](https://www.datafountain.cn/competitions/346)：



论文正文部分包括：

1. 绪论：光学字符识别技术目前的发展状况，论文采用的技术方案。 完成
2. 图片预处理：原始身份证扫描件中提取有效区域及去除水印。 完成区域提取
3. 文本处理：定位图片中的文本和识别文本。
4. 结果纠正：爬取全国行政区和身份证号码的对应关系进行结果纠正
5. 方案展望：总结方案中的不足和后续可提升的地方。
6. **剩余工作量**

上述的（二）（三）（四）（五）待完成，剩余约70%

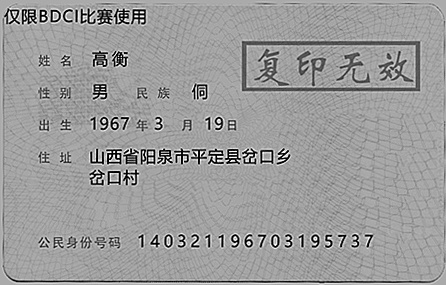
1. **遇到的问题**

原始数据提取的有效区域不都是正面角度，有些图片会出现颠倒造成后续文本识别的困难。

例如：

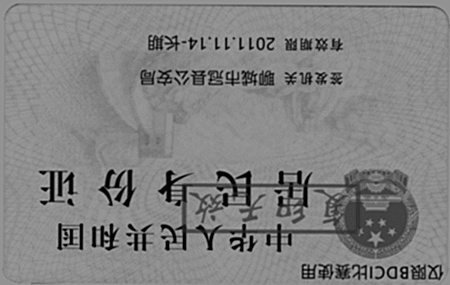
正常情况：

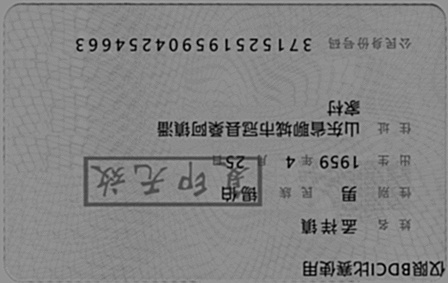






失败情况：







1. **预计解决方案**

根据左上角的“仅限BDCI比赛使用”写一个模块去判断图像正反，若反则翻转图像。