

单据文字定位算法

问题定义

定位出票据中需要识别的关键信息，为后续的识别算法提供输入

样例：



机器编号: 661565721268

上海增值税电子普通发票

发票代码: 031001600211
发票号码: 52264999
开票日期: 2016 年 07 月 20 日
校验码: 81417 12727 25230 07744

购买方	名称: 麦当劳 (中国) 有限公司		密码区	5+06367>9239803/731723*<*/6*>03<027405401-2/>9+38-/7640-30>-/32-37*34803-031>234>/817<003<027405401-2/>>3>*				
	纳税人识别号: 地址、电话: 开户行及账号:							
货物或应税劳务、服务名称		规格型号	单位	数量	单价	金额	税率	税额
餐饮服务				1	408.02	408.02	6%	24.48
合计						¥408.02		¥24.48
价税合计(大写)		肆佰叁拾贰圆伍角零分					(小写) ¥432.50	
销售方	名称: 上海麦当劳食品有限公司		备注					
	纳税人识别号: 913100006072048513 地址、电话: 上海市黄浦区福州路8号埃力生国际大厦11层 021-33326100 开户行及账号: 招商银行上海分行新客站支行 211681476510001							

收款人: 复核: 开票人: 1450167 销售方: (章)

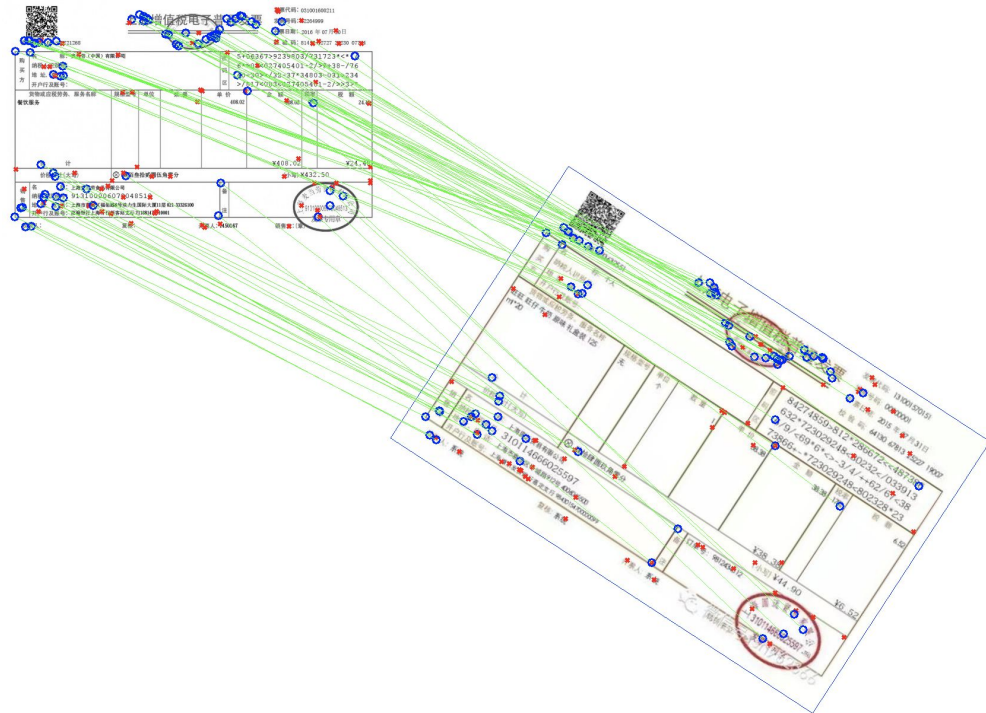


定位算法

- 模版匹配
 - a. 仅适用于版式固定的票据
 - b. 票据图像可以倾斜
 - c. 定位框可能不够紧致
 - d. 无需训练
- Faster R-CNN
 - a. 对票据版式无要求
 - b. 票据图像不可倾斜
 - c. 定位框比较紧致
 - d. 需要训练(每种版式单据或碎片的训练样本量最少需要500, 建议5000以上)

模版匹配

- 算法流程
 - a. 标记出模板上需要识别的关键信息区域
 - b. 提取模板特征点
 - c. 提取待匹配图片特征点
 - d. 进行特征点匹配, 计算出变换矩阵
 - e. 使用变换矩阵, 对待匹配图片进行变换
 - f. 截取待匹配图片上的关键信息区域
- 代码及文档
 - a. http://git.4paradigm.com/sunchang/cm_bc_ocr



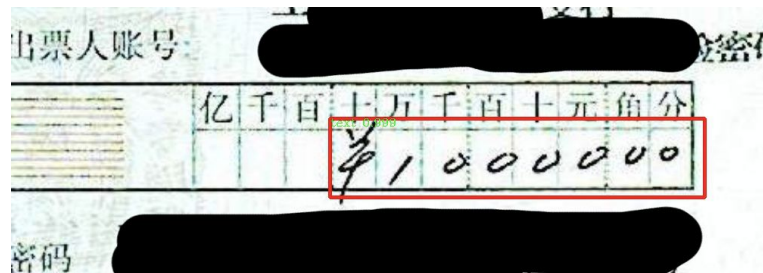
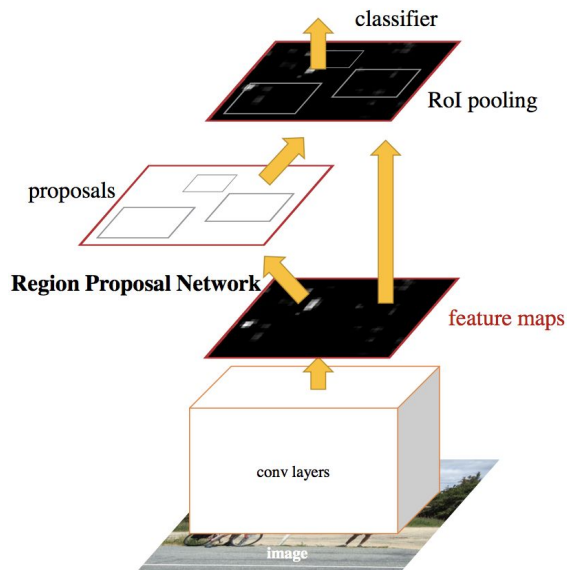
Faster R-CNN

- 算法流程

- 使用卷积网络提取特征
- 使用RPN网络提取候选框
 - 选定一些预先设定的Anchor(锚定框)
 - 分类:判断Anchor是前景还是背景
 - 回归:对Anchor的位置进行回归
 - 选择排名靠前的前景Anchor作为候选框
- 使用R-CNN网络进行目标分类
 - 分类:判断候选框是哪种类别
 - 回归:对候选框的位置进行回归

- 代码及文档

- <http://git.4paradigm.com/sunchang/deeptext>
- 江苏银行项目支持代码:目前仅支持单目标

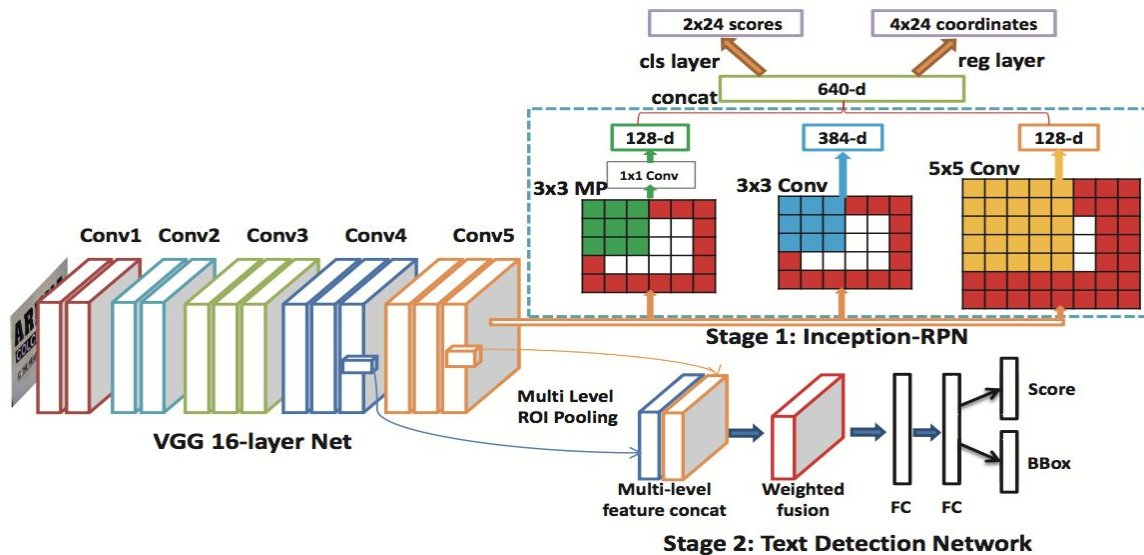


Deeptext论文对Faster RCNN的改进

1. 在卷积层之后增加了Inception结构

参加代码:http://git.4paradigm.com/sunchang/deeptext/blob/d4e1c90539bf4ff9772760d5a9f32d643f729fbc/rcnn/symbol/symbol_vgg.py#L370

2. Multi-Level ROI Pooling



对Deeptext的改进

- 不使用Multi-Level ROI Pooling
- 将pool4和relu5_3进行连接
- 然后进行1x1卷积
- 最后进行一次ROI Pooling

参见代码

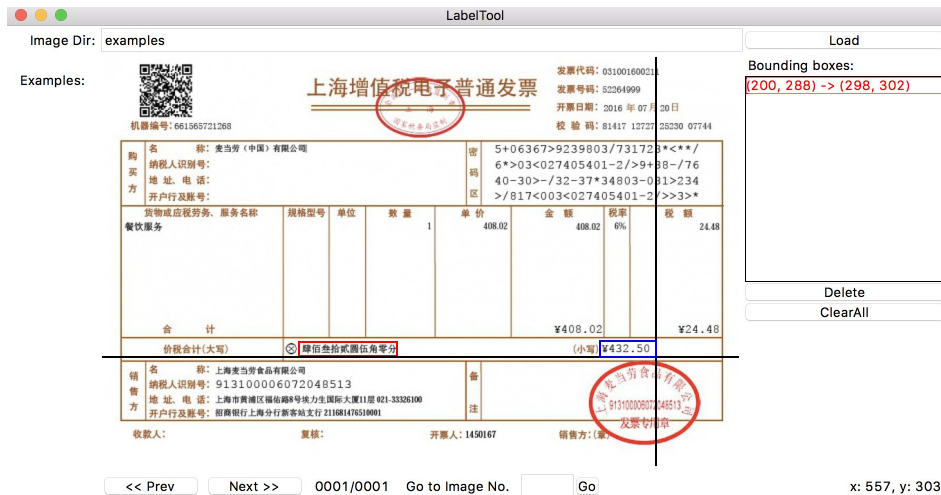
: http://git.4paradigm.com/sunchang/deeptext/blob/d4e1c90539bf4ff9772760d5a9f32d643f729fbc/rcnn/symbol_vgg.py#L419

使用流程

1. 判断单据是否为固定版式
2. 对固定版式单据，使用模版匹配法
 - a. 如果效果不好，针对割出来的区域进行Faster R-CNN精确定位
3. 对非固定版式单据或者碎片
 - a. 如果目标区域有较高的视觉区分度，则可以标记出部分区域
 - b. 如果目标区域无视觉区分度，则需要标记出所有文字区域
 - c. 使用Faster RCNN进行定位
 - d. 识别定位出的区域
 - e. 如果为多目标，需要根据位置关系或者文字内容 进行信息提取

工具使用方法

- 标注工具使用说明:
<http://git.4paradigm.com/sunchang/BBox-Label-Tool/blob/master/README.md>
- 模版匹配使用方法:
http://git.4paradigm.com/sunchang/cmbc_ocr/blob/master/proj_jsbank/README.md
- Faster RCNN使用方法:
http://git.4paradigm.com/sunchang/deeptext/blob/master/proj_jsbank/README.md



标注工具截屏

相关论文

1. R-CNN: Region-based Convolutional Neural Networks
2. Fast R-CNN
3. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks
4. DeepText: A Unified Framework for Text Proposal Generation and Text Detection in Natural Images

QA