



自然场景文本定位进展

阮少辉

卢俊国教授课题组

机器人与智能信息处理研究所
系统控制与信息处理教育部重点实验室
电子信息与电气工程学院自动化系
上海交通大学



自然场景文本检测

问题： 自然场景文本检测是要在复杂的自然场景图像中对其中的文本进行定位

应用： 智能交通系统，无人驾驶，助盲翻译系统，票据自动录入，基于内容的视频检索，网页信息过滤等等

难点： 背景复杂，网格，窗户，建筑等特征与文本很相似

文本布局多样，横向纵向倾斜等等

成像角度多样，文本会有不同程度的放射变换

光照、低分辨率等其他因素

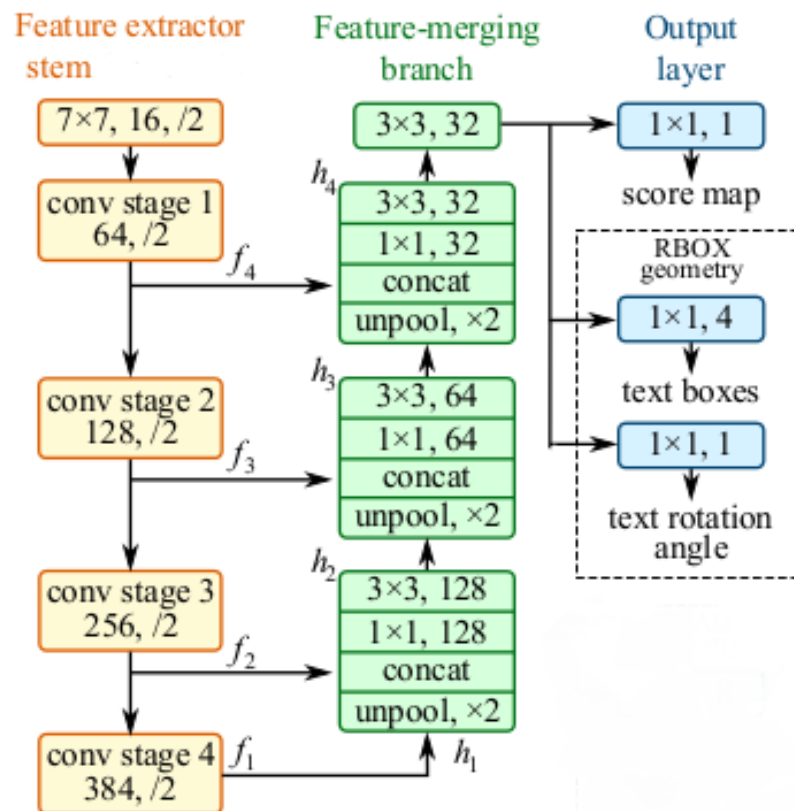




自然场景文本检测

方法：深度学习，提出一种基于FCN(全卷积)网络的倾斜文本检测模型

- 特征提取部分采用Resnet
- 抽取不同层次的特征采样之后进行特征融合
- 输出一个旋转角以及4个距离
- 端到端可训练
- 简洁的pipeline，只需要做一步nms的后处理





竞赛介绍

- ICDAR(International Conference on Document Analysis and Recognition) 是目前文档分析和识别领域的顶级会议。
- ICDAR分别在03, 11, 13, 15年举办Robust Reading Competition
- 几乎所有相关的论文都绕不开这个竞赛
- 竞赛一共有4个task: text localization、text segmentation、word recognition和end to end。
- 我们参加了task1, 也即text localization



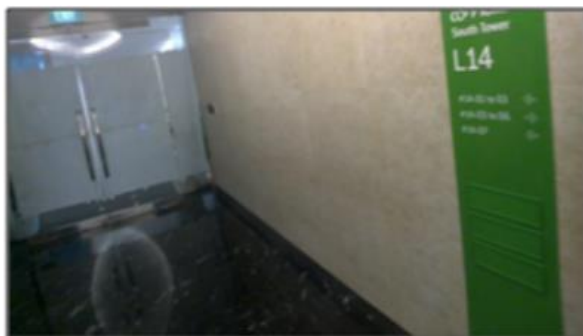
竞赛介绍

ICDAR2015 Incidental Scene Text:

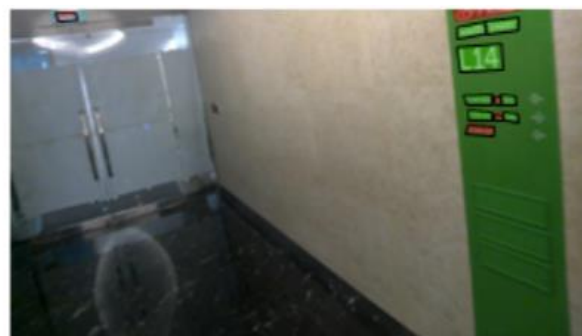
1. 该数据集包含 1000 张训练图片， 500 张测试图片
2. 通过Google Glass 进行采集，without people 's attention.
3. 数据集中图片运动模糊严重、文字为多角度、包含各种畸变。



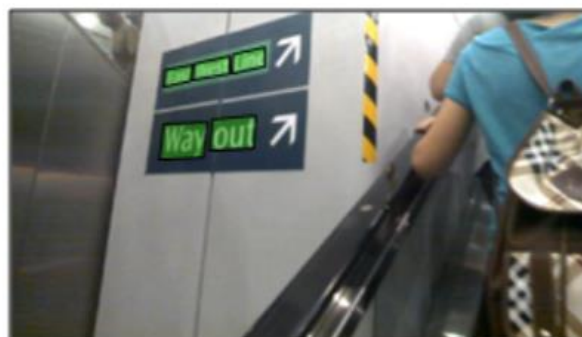
任务示例



img_1.jpg



img_2.jpg





自然场景文本检测

- 我们在目前最新的最难的ICDAR2015 incidental scene text dataset 上的task1 text localization取得了第 **10** 名的成绩。排名前列的都是知名的AI或者巨头公司。

Date	Method	Recall	Precision	Hmean
2017-07-12	Sensetime V2	84.98%	89.41%	87.14%
2017-07-12	Tencent-DPPR	81.80%	90.71%	86.03%
2017-07-04	Baidu IDL v3	81.99%	89.82%	85.73%
2017-05-13	SRC-B-Machine...	82.81%	88.66%	85.64%
2017-04-27	Sensetime_V1	83.92%	87.32%	85.59%
2017-07-26	TextBoxes++	80.79%	89.11%	84.75%
2017-08-26	CCFLAB_FTSN	80.02%	88.31%	83.96%
2017-02-17	NLPR-CASIA	82.76%	84.76%	83.75%
2017-06-11	SCUT_DMPNet...	83.73%	83.53%	83.63%
2017-08-18	Resnet_SJTU	78.29%	86.58%	82.23%
2017-08-18	Dahua-OCR v1	79.30%	85.16%	82.12%
2017-02-22	SRC-B-Machine...	79.78%	83.64%	81.67%





上海交通大学

SHANGHAI JIAO TONG UNIVERSITY



请领导和专家指正！
谢谢！

联系方式：卢俊国 教授 博导

Mobile: 13918206230

Email: jgLu@sjtu.edu.cn