

# 阮晓虎

男/1986.7/湖北襄阳

Tel: 86-13124757621 E-mail: xiaohu\_ruan@163.com



## 教育背景

- 2011.09—2014.06    中科院半导体研究所    高速电路与神经网络实验室    硕士    电子信息工程  
• 研究方向：图像、视频处理，模式识别，机器学习相关算法
- 2007.09—2011.06    中央民族大学(211、985)    信息工程学院    本科    电子信息工程

## 工作经历

2017.06—now    同盾科技·反欺诈数据算法图像团队    资深算法工程师

[项目 1] 通用 OCR 算法，包括：广告图片、学信网信息、身份证、行驶证等场景文本检测和识别

- **工作 1：**基于 CTPN(ECCV2016 Weilin Huang .etc)算法上线通用文字检测服务，CTPN 算法主要考虑将图像中的文本行作为一个 sequence 的思路来做检测，根据这个 sequence 的行信息确定文本行的 y 方向位置，再经过文本行连接策略确定文本行位置，在对比 TextBoxes, Seglink(CVPR2017), SSD, 以及 segmentation based(FCN, EAST, mask-rcnn)等算法，从数据标注成本、效率、效果等综合考虑，最终选取 CTPN 进行训练，调试参数、上线部署，2k 规模测试集召回率 95.5%，准确率 93%，1080TI 11G 显存测试平均 QPS 26. 已达上线标准；
- **工作 2：**基于 CTPN 作者已开放测试源码基于 caffe 框架实现训练代码，实现 input\_data\_layer, bbox\_target\_layer, rpn\_cls\_loss\_layer, textline\_bbox\_loss\_layer 等相关 layer，并进行模型训练；
- **工作 3：**支持印尼身份证识别的 OCR 检测和识别。检测方案使用前文所述 CTPN，识别采用基于文本行的序列预测算法。调研 CRNN（使用 CNN 做特征提取，采用 bi-LSTM 学习序列的长短期依赖关系，使用 CTC Loss, Warp-CTC 作为损失函数学习模型参数）及 seq2seq+Attention model（CNN 处理文本行特征输入序列，使用带 Attention 的编码器解码器进行序列依赖关系的学习），最终根据分析及算法效果选用第二种算法进行模型训练、调优及部署，算法处于初步上线测试状态；
- **工作 4：**支持学信网（印刷体）OCR 识别中的检测：学信网的 OCR 检测要求 CPU 环境的快速算法，考虑学信网图像相对简洁（类似一般 PDF 文本文件），因此，采用 sobel 算子、二值化和横向结构的形态学算子（膨胀+腐蚀），以及二值化形状约束进行检测，最终算法在 CPU 机器部署，检测召回和精度均大于 99%，QPS: 52 左右，稳定运行时间大于 3 months；

[项目 2] 广告图片中二维码检测与识别

- **工作 1：**采用 SSD 算法检测+mobilenet\_v1 网络做分类的二维码检测任务，测试召回率 100%，审核量 30~50，现已稳定运行 6 months，并将该目标检测架构作为标准算法推广到其他业务场景（如：车辆检测，人体检测等）；

[项目 3] 用于深度学习算法的机器学习平台架构设计

- **工作：**作为图像组带头人，主导设计机器学习平台组件及架构，包括：数据存储，数据管理，标注平台，训练平台，模型管理，模型发布等功能组件，并推进相关项目小组实现和项目验收。目前，项目正在进展中；

2016.09—2017.05    阿里巴巴·iDST·拍立淘团队

资深算法工程师

[项目 1] 拍立淘类目及数据选品扩充

- **工作：**在寻求拍立淘图像搜索效果的提升方面：a). 对全淘商品搜索类目的划分做优化，对部分类目进行合并搜索，使得最终个别类目精度提升达 5~10 个百分点，搜索结果的同款率提升 1~2 个百分点，完成相关离线结果的验证；b). 通过选品数据的扩充，扩充叶子类目主图、sku 图以及对选品图像做 multicrop 等方式做 data expansion，完成离线实验分析，数据上线，daily 评测，预发效果验证等；
- **效果：**提高选品数据量及 query 类别的鲁棒性，各类目最终同款率提升 0.5~2 个百分点不等；（已上线）

[项目 2] 拍立淘个性化类目预测

- **工作：**利用实时搜索 pvlog 的分析及 GBDT 算法模型，对搜索类目进行辅助预测，实现离线实验效果分析和算法上线，包含：  
离线部分：用户短期+长期个性化特征构建、完成相应 ODPS MRJOB，并进行数据生产，从 pvlog 中提取训练及测试数据，离线模型训练及测试，数据及模型迭代，同款率分析等；

在线部分：特征构建及预测，实时 pvlog 到个性化特征计算（基于 Prosche 平台），短期和长期相关数据表 daily run 任务发布，分布式 GBDT 模型训练（基于 MPI 机器学习算法平台），预测模块 online 服务开发、部署、压测，特征服务 iGraph 表上线，与拍立淘 SP 端进行算法联调，最终各模块均达上线状态；

- **效果：**类目预测精度提升 7 个百分点，各类目在原来基础上同款率提升 1~5 个百分点不等；（已达成上线状态）

#### [项目 3] 万物识别相关

- **工作：**利用 Caffe 进行 tripletloss 训练，修改 LMDB 数据组织格式，在训练中（相应 data\_layer）从图像分类类目维度对样本进行过滤；

2014.02—2016.09 阿里巴巴·高德地图研发中心·图像算法组

图像算法工程师

#### [项目 1] 道路交通标志自动识别 Traffic Sign Recognition(TSR)项目

- **背景：**TSR 项目致力于道路交通标志信息的地图数据生产，由采集车得到的视频经过自动识别，直接输出红圆、蓝圆、黄三角、地面行车引导线、蓝绿方牌等交通标志的信息点，提高数据生产的效率，对算法的要求是：高召回、低误报，低时耗；

#### ➤ **算法：**

- “固定模式”类（红圆、蓝圆、黄三角）检测识别：采用 Adaboost 级联分类器提 Region Proposal，多种特征(Harr/HOG/LBP/PixelDiff)选择，建立自动样本迭代机制；采用 CNN 做候选框的分类，深度学习 CAFFE 框架应用，AlexNet、NIN 等网络结构训练；（已投产）
- 地面行车引导线检测识别：存在视角差异，采用图像消失点构建前方路面俯视图，在俯视图上采用上述固定内容标志检测方式；（已投产）
- 蓝绿路牌检测：问题的难度在于方牌尺寸和内容不固定，采用多特征融合+随机森林分类器；（已投产）
- 交通灯检测：限于高精图像横排、竖排圆形交通标志，颜色分割+SVM 分类器

#### ➤ **效果：**

- 测试集：众包、高精、智能硬件、DVR(不同地图数据采集方式得到的道路视频或图像)等数据来源，规模 25w 左右；
- “固定模式”类召回率达 98%，准确率 90~95%，百图错误率(FalseAlarm per hundred)约 0.25；
- 交通灯检测 DEMO 测试：召回 85%，准确 90%；（项目终止）

- **应用：**算法在高德数据生产线的表现，相比人工提效达 200%；

#### [项目 2] 蓝绿路牌内容识别

- **背景：**地图中道路路牌的矢量化数据生产是一个费时费力的环节，路牌内容的识别成为迫切需求；

- **算法：**传统机器学习 Cascade 与 Faster-rcnn 结合定位文字等信息，OCR 做文字识别；

#### [项目 3] 全景图像生成

- **背景：**为高德地图早期的街景图像生产提供服务，全景图像隐私保护处理；

#### ➤ **算法：**

- **街景图像处理**，包括图像匹配、拼接，球面映射，全景图生成，人脸、车牌检测等技术
- **鱼眼图像拼接**，相机参数、姿态角估计，warp，blending，重点攻克全景图拼接对齐问题[APAP, CVPR2015]
- **图像匹配**，SIFT 算法工程上的改进、图像匹配算法策略的实现，改进的 SIFT 算法在匹配效果不变的情况下，效率提升了一倍

## 科研经历

2012.09—2013.09 基于高维形象几何分析方法的仿生人脸图像处理与识别方法研究（国家自然科学基金重大研究计划项目） 主要负责人

- 人脸的自动定位技术，人脸图片的归一化
- 提出“一种基于特征匹配的人脸配准判断方法”，有效的保证了用于识别的人脸图片的有效性
- 人脸识别算法的探索，包含在 PCA、LDA、Local Gabor 等人脸识别算法上的探索

2012.06—2013.1 基于视觉仿生计算的主动安防监控技术 主要负责人

- 检测监控视频中存在的动态目标，识别不同的目标，完成不同目标的跟踪任务
- 算法均由 VC 编程实现，本工作将受“视频查询方法”专利保护

2011.10—2012.05 图片汉字识别·华侨大学(福建)合作项目 项目合作者

	<ul style="list-style-type: none"> <li>• 图片预处理、汉字切分、字块特征提取</li> <li>• 通用规范汉字表一级汉字和二级汉字识别，所有算法由 VC 编程实现</li> </ul>	
2009.09—2010.03	中央民族大学本科生科研项目（URTP）	项目负责人
	<ul style="list-style-type: none"> <li>• “小功率风力发电充电装置的设计与制作”，获得 2010 年 10 月期 URTP 优胜奖</li> </ul>	
2013.03—2013.05	基于色素标记的玉米单倍体高通量全自动分选系统	项目参与者
	<ul style="list-style-type: none"> <li>• 算法测试工作，项目申请算法演示</li> </ul>	

### 志愿活动

2007.10—2008.06	中央民族大学学生会	学生干部
	<ul style="list-style-type: none"> <li>• 调查管理勤工助学学生档案，搜集分配校内外勤工助学岗位，联系企业资助</li> <li>• 开办“明日之星—优秀校友座谈会”“青年创业发展交流会”等活动</li> </ul>	
2008.10—2011.06	中央民族大学数字网络教学终控中心	勤工助学&实习
	<ul style="list-style-type: none"> <li>• 负责与北京竞业达科技有限公司接洽，联合维护学校多媒体网络教学设备</li> </ul>	
2008.06—2008.09	2008 北京第 29 届奥林匹克运动会	志愿者
	<ul style="list-style-type: none"> <li>• “奥运服务站”语言翻译、外宾接待，“公益千万家”服务空巢老人等志愿活动</li> </ul>	
2009.07—2009.09	全国大学生电子设计大赛	参赛组负责人
	<ul style="list-style-type: none"> <li>• 参加校方“电子创新实验室”2 个月集中训练，模拟/数字电子线路设计，单片机</li> <li>• 参赛题目 F：宽带功率放大器设计，获得北京赛区三等奖</li> </ul>	

### 发表论文

• 专利	视频查询方法，申请号：201210567136.1，发文序号：2012122500611450
• 期刊论文	一种基于特征匹配的人脸配准判断方法，期刊：电子与信息学报(EI 检索)

### 技能特长

• 外语	大学英语四级 CET4（558/710），六级 CET6（502/710），雅思 IELTS（5.5/9）
• 计算机	计算机等级考试三级（网络技术）
• 编程语言	C/C++，python，shell，MATLAB
• 性格爱好	开朗随和，真诚率直，体育爱好：游泳，羽毛球，篮球