



## 游京翰

一个积极向上，兴趣广泛，热爱工作的年轻人

📅 24岁

☎ 136-1290-2871

✉ youjinghan2018@163.com

★ 2019年1月毕业

★ 中共党员

## 个人技能

- 深入研究OCR、图像合成技术
- 有开发深度学习、机器学习以及tensorflow、keras框架的经验
- 掌握python、C等编程语言
- 六级：441分(15年12月)

## 其他经历

- 大学担任卓越班班长、数学协会会长
- 从事全国农产品贸易洽谈会、全国武术运动会等志愿者服务

## 项目经验

### • 天池-ICPR MTWI 2018-网络图像的文本识别任务评测

**项目简介：**该竞赛对网络图片中的文本进行识别。**主要负责：**(1) 对本行区域进行ROI定位并对方向进行水平矫正。(2) 对训练集中的文本区域进行统计分析确定模型参数；(3) 在CNN+LSTM框架下，提出一个基于位置attention的文本图片序列识别模型，并用tensorpack神经网络包实现；(4) 提出了一个有效的合成图片方法，用于弥补训练样本的不足。**最终结果：**在评测测试集上编辑距离正确率达到80.4%，最终成绩为7/1370，在高校队伍中排名第二。

### • 多语言手写体识别引擎

**项目简介：**构建一个通用的多语言手写体识别引擎。**主要负责：**(1) 提出一个基于CNN+LSTM+CTC的文字图片序列识别模型，并在大数据集下用tensorflow实现多GPU数据并行训练框架；(2) 用n-gram语言模型进行后处理。**最终结果：**没有用到字典和语言模型修正的情况下，在ICDAR RIMES法语评测数据集上编辑距离正确率为96.5%，加泰诺尼亚语评测数据集编辑距离正确率为99%，另外中科院CASIA-HWDB中文单字评测数据集单字正确率为97.4%。

### • ICDAR2017-Information Extraction in Historical Handwritten Records评测

**项目简介：**这是OCR领域顶会ICDAR2017竞赛题目，任务是将加泰诺尼亚的历史婚姻手写记录识别转录出来，并对其中的人名地名以及所属关系进行命名实体标注。**主要负责：**(1) 用keras框架实现一个基于CNN和位置attention的文字图片序列识别模型；(2) 对人名地名及其所属关系的命名实体识别。**最终结果：**取得了95.46%的文字转录加命名实体识别正确率，并取得该竞赛词级别第一名。

### • 知识关联与推理类问题求解关键技术及系统

**项目简介：**国家863项目知识关联与推理类问题求解关键技术及系统的子模块。通过互联网获取的历史知识来完成高考历史简答题的作答。**主要负责：**(1) 对问题的材料中关键词做命名实体识别；(2) 利用Lucene搭建一个全文检索系统，用于检索答案。**最终结果：**实现了一个基于关键字抽取和全文检索的答题机器人。

2016.9-2019.1	哈尔滨工业大学(硕士)	计算机科学与技术 成绩排名前20%
2012.9-2016.7	天津工业大学(学士)	软件工程(卓越计划) 成绩排名前10%

## 奖项荣誉

---

- 竞赛获奖

1. 2018年天池-ICPR MTWI网络图像的文本识别评测7/1370，高校第二名
2. 2017年ICDAR-历史手写记录中的信息抽取评测（词级别）第一名
3. 2015年美国国际数学建模竞赛二等奖
4. 2014年全国大学生数学建模大赛（天津赛区）二等奖
5. 第五届蓝桥杯C程序设计（天津赛区）一等奖

- 学习获奖

1. 2017年哈尔滨工业大学一等奖学金
2. 2016年哈尔滨工业大学二等奖学金
3. 2015年计算机学院单项奖学金（社会活动类）
4. 2012-2014年校长二等奖学金（连续两年）

- 其他荣誉

1. 2012年驻马店市优秀志愿者
2. 2012-2013年度校优秀共青团员荣誉称号
3. 2013-2014、2014-2015年度校优秀大学生干部
4. 2014年校暑期社会实践“先进个人”荣誉称号
5. 2017年校优秀共产党员

## 自我评价

---

有较强的自控能力、做事情相对专注。乐于与他人合作，遇到问题积极与他人沟通。积极向上，喜欢篮球、跑步、游泳、读书、旅游等