

郝杰东

✉ jdhao@hotmail.com · ☎ (+86) 132-6152-3068 · in [jdhao](#) · 🏠 <https://jdhao.github.io> · 📍 广东深圳

🎓 教育背景

工学硕士，中国科学院自动化研究所，模式识别与智能系统	2014.09 – 2018.06
工学学士，中南大学，自动化	2010.09 – 2014.06

⚙️ 技能

- 编程相关
 - 熟悉 Python 使用 (代码风格良好)，熟悉使用 Flask, uwsgi 以及 docker 搭建一整套在线服务流程；了解 C++, lua 的使用；熟悉常用的算法以及数据结构
 - 熟悉 Linux 命令行下开发环境，熟练使用 Zsh, bash 下常用命令行工具，熟悉简单的 shell 脚本编程；熟练使用 Git (branch, merge, rebase, stash) 进行代码管理，熟悉 make 编译工具；熟悉 docker 镜像构建、打包、使用
 - 精通 Vim/Neovim 使用 (日常主力开发工具)，熟悉 PyCharm, Sublime Text 使用
- 模式识别与机器学习
 - 熟悉 OCR 文字识别技术 (模型的训练，语料库收集与数据合成) 和文字识别评测方法；熟悉图片文档重建整套流程技术 (目标检测，复杂排版后处理，PPT 生成)；熟悉基于内容的图片检索相关技术及流程；熟悉视频指纹/去重算法以及相关流程
 - 熟练使用 PyTorch 深度学习框架
- 语言水平
 - 熟练阅读英文文献及资料，能够迅速学习并掌握最新技术
 - 大学英语等级考试：四级 650，六级 613

👥 项目经历

vivo 内容理解平台作者去重和内容去重项目	2021 年 1 月 – 至今
-------------------------------	-----------------

算法开发与工程部署 项目负责人

项目背景：合并内容提供方 (CP) 提供的重复作者，建立统一作者体系；发现 CP 提供内容中的重复视频，减少视频信息流推荐视频重复率，提升用户体验；满足业务运营补品与内容监控需求。

- 负责视频去重算法的研发，算法评测指标设计与实现，离线测试框架的搭建，badcase 的分析优化；参与服务架构设计，负责视频去重中控，视频去重服务，视频帧特征提取服务的开发、部署。
- 主要涉及技术：视频截帧与音频提取，视频帧特征提取，视频特征检索，视频去重后处理，视频序列匹配，音频去重，视频黑边/蒙版检测，视频 gid 生成等技术。
- 服务月调用量超三千万，具体指标，召回率：98%，误杀率：百万分 15 (前值，召回 70%，误杀百万分之 40)

vivo 智慧视觉 PPT 重建项目	2020 年 1 月 – 2020 年 12 月
---------------------------	--------------------------

算法开发与工程部署 项目负责人

用户拍一张 PPT 文档图片 (拍摄投影或屏幕等场景)，经过算法服务处理，将原始图片转为可编辑的文档，图片，表格、文字可编辑；文档背景，文字大小、颜色、段落，排版近似还原为原始图片效果；用户可在手机上先预览生成的 PPT 效果，再生成 PPT。

- 全面负责项目整体进展管理 (数据收集，质量目标确定，排期，开发，测试对接与解 bug)，独立负责所有算法相关开发工作，以及部分 PPT 重建服务的部署工程工作。
- 主要涉及技术：基于 maskRCNN 的图片元素检测，经过复杂后处理 (maskRCNN 检测结果后处理，OCR 结果拼接以及组成段落，文字行字号、行间距的精细化确定，文字与图片，表格等其他元素

的精细化排版处理, 表格识别解析), 最后使用 `python-pptx` 将解析结果生成为可编辑的 PowerPoint 文件。

- 图片重建 PPT 通过率 (测试人工查验): 87%。

vivo 智慧视觉扫 WiFi 项目

2019 年 6 月 – 2020 年 2 月

算法开发与工程部署 项目负责人

项目目的: 用户拍摄含有 WiFi 账号密码的图片, 智能识别其中的账号密码信息, 一键连 WiFi。

- 负责项目整体进展管理 (数据收集, 质量目标确定, 排期, 开发, 测试对接与解 bug), 并负责其中的 WiFi 信息解析模块, 负责 WiFi 识别服务的在线部署。
- 可以准确识别并解析出实际生活场景中复杂图片, 包括一行, 多行等复杂排布, 打印、手写等不同字体风格, 以及带有多角度旋转的 WiFi 牌。
- 账号密码识别准确率, 账号: 87%, 密码: 85%

vivo 名片 OCR 以及通用 OCR 项目

2018 年 10 月 – 2019 年 6 月

算法开发

- 负责竖排文字识别 (模型 CRNN), 由于竖排文字图像稀少, 同时开发了竖排文字图像合成工具 (使用维基百科和 THUNews 等多种语料库, 支持繁简字体, 可合成多种背景和颜色, 带有多重噪声, 模糊, 形变, 裁剪等的文字图片)。名片场景, CER: 0.06, 整行准确率 0.85; 通用 OCR 场景竖排文字识别, CER: 0.14 左右 (识别难度较大, 文字排布不规则, 艺术字, 模糊, 尺寸小)
- 设计并训练文档/名片方向检测模型, 带权重分类模型, 准确率 0.99 以上
- 为了识别文本行的方向, 提升文本识别准确率, 设计并训练文本行方向分类模型, 分类准确率 0.97 左右

i 其他

- GitHub: <https://github.com/jdhao>
- 技术博客: <https://jdhao.github.io/> (历史访问量超 260 万, 月 PV 3 万)
- Stack Overflow: 积分值 16000+, 排名前 0.09% (总人数 1600 万+), 常回答与 PyTorch, Vim, Python 等相关问题

论文

- Jiedong Hao, Yafei Wen, Jie Deng, Jun Gan, Shuai Ren, Hui Tan and Xiaoxin Chen, EEM: An End-to-end Evaluation Metric for Scene Text Detection and Recognition. ICDAR 2021
- Jiedong Hao, Jing Dong, Wei Wang and Tieniu Tan, DeepFirearm: Learning Discriminative Feature Representation for Fine-Grained Firearm Retrieval. ICPR 2018 (**Best Scientific Paper Award**)