



马洁锋

✉ jfma@mail.ustc.edu.cn

☎ (+86) 173-0569-1836

in JeffreyMa

📄 github.com/jfma-USTC

Python

Pytorch

Linux

Flask

Javascript

🎓 教育背景

中国科学技术大学 2020.09 - 2025.06 博士学位
语音及语言信息处理国家工程研究中心, GPA:3.83/4.3, 系排名:4/65, 导师: [杜俊](#) 副教授

中国科学技术大学 2016.09 - 2020.06 学士学位
信息科学技术学院人工智能英才班, GPA:3.87/4.3, 系排名:4/91

🔬 科研/项目经历

多模态内容理解 科大讯飞 AI 研究院 2023.02 - 至今

- **任务介绍:** 利用多模态生成式模型统一处理若干项多模态内容理解任务。
- **研究内容:** 研究基于文本、图像、视频、音频等模态输入的生成式内容理解框架, 采用视觉/音频/文本 Transformer Encoder 模型对图像/视频/音频/文本进行表征学习, 构成用户输入的查询对象, 通过用户提供的自然语言作为查询内容, 并采用 Transformer Decoder 模型联合前述查询对象和查询内容进行跨模态特征交互, 使得模型能够给出文本形式的任务目标输出。相关工作进行中。

智能文档结构化 科大讯飞 AI 研究院 2022.03 - 2022.12

- **任务介绍:** 从文档图像中恢复出文本内容及其层级信息。
- **研究内容:** 参与采用基于局部注意的图网络进行句级文档预训练相关工作, 验证多模态特征表示方式在下游任务上的提升; 提出单页可跨页索引的建模方式, 使文档结构化能力扩展到超长文档场景; 基于多模态表征提出针对数据集自适应的解码策略, 相较基线系统取得 56.7% 的性能提升。
- **成果总结:** 提出论文结构化数据集及解决方案 HRDoc¹, 相关论文已被 *AAAI 2023 (Oral)*¹ 接收, [开源代码](#); 新的预训练方式 GraphDoc² 在多项信息抽取下游任务上达到 SOTA, 相关论文已被 *TMM*² 接收, [论文](#), [开源代码](#)。

零样本文档信息抽取 腾讯优图实验室 2020.10 - 2021.10

- **任务介绍:** 实际场景下文档信息抽取存在未见关键词、关键词缺失等情况, 希望设计一种具备处理零样本能力的抽取方案。
- **研究内容:** 构建基于 Prompt 的 V-K 序列生成模型, 大大降低对关键词的依赖。
- **成果总结:** 在 SROIE 等五项数据集上取得 SOTA 结果, 并在内部多版式零样本数据集 LastDoc4000 上显著超越 LayoutXLM 等模型, 在未见版式 / 未见关键词场景下取得 30.32% 的 F1 提升。相关论文 QGN³ 已被会议 *ACM-MM 2022 (Oral)*³ 接收, [论文](#)。

生成式文档信息抽取 腾讯优图实验室 2020.10 - 2021.10

- **任务介绍:** 票据场景下的信息抽取面临 OCR 错误、检测框粘连、复合字段等难题, 希望设计一种鲁棒的抽取方案, 可以做到对上述问题的较好处理。
- **研究内容:** 构建文档场景下的生成式预训练语言模型, 将下游任务重新定义为 K-V 序列生成任务。
- **成果总结:** 在 SROIE/CORD 榜单上取得第一名的成绩, 在 OCR 错误 / 阅读语序错误场景下分别取得 31.1% / 25.3% 的 F1 提升。相关论文 GMN⁴ 已被会议 *NAACL 2022 (Oral)*⁴ 接收, [论文](#)。

基于 Kaldi 工具库的 2D-HMM 实现 2019.09 - 2020.11

- **任务介绍:** 将 1D-HMM OCR 扩展至 2D 场景, 用二维空间隐状态来建模汉字的结构。
- **研究内容:** 实现二维隐马尔可夫模型的状态建模、前后向算法、EM 算法等训练推理过程。

- **成果总结:** 实现基于 Kaldi 的 2D-HMM 模型, 在 HWDB 数据集上相较于 1D-HMM 模型汉字错误率下降 50%。相关工作已被会议 *ICIG 2021*⁵ 接收, [论文](#), [开源代码](#)。

♡ 获奖情况

第一名, 第五届中国创新挑战赛智慧教育专题赛“教育手写公式识别”比赛 (1/250)	2020.10
奖学金, 2018 年中国科学技术大学校优秀学生奖学金——金奖 (Top 1%)	2018
奖学金, 2019/2017 年中国科学技术大学校优秀学生奖学金——银奖 (Top 10%)	2017 / 2019

i 已发表论文

- 1: **Jiefeng Ma**, et al. HRDoc: Dataset and Baseline Method Toward Hierarchical Reconstruction of Document Structures, AAAI 2023 (CCF-A, **Oral**)
- 2: Zhenrong Zhang, **Jiefeng Ma**, et al. Multimodal Pre-training Based on Graph Attention Network for Document Understanding, TMM (中科院 **JCR-1** 区)
- 3: Haoyu Cao, Xin Li, **Jiefeng Ma**, et al. QGN: Query-driven Generative Network for Document Information Extraction, ACM-MM 2022 (CCF-A, **Oral**)
- 4: Haoyu Cao*, **Jiefeng Ma***, et al. GMN: Generative Multi-modal Network for Practical Document Information Extraction, NAACL 2022 (CCF-C, **Oral**)
- 5: **Jiefeng Ma**, et al. An Open-source Library of 2D-GMM-HMM Based on Kaldi Toolkit and Its Application to Handwritten Chinese Character Recognition, ICIG 2021

i 在投论文

- 1: Pengfei Hu, **Jiefeng Ma**, et al. Count, Decode and Fetch: A New Approach to Handwritten Chinese Character Error Correction (**submitted to ICCV 2023**)
- 2: Zhenrong Zhang, **Jiefeng Ma**, et al. SEMv2: Table Line Detection Based on Conditional Convolution (**submitted to ICCV 2023**)