

中图法分类号: TP309; TP399 文献标识码: A 文章编号: 1006-8961(2023)08-2221-02

论文引用格式: 刘成林, 金连文, 殷绪成, 白翔, 连宙辉, 彭良瑞, 刘禹良, 陈秀妍. 2023. 《中国图象图形学报》文档图像智能处理与识别专栏简介. 中国图象图形学报, 28(08):2221-2222[DOI:10.11834/jig.2300008]

《中国图象图形学报》 文档图像智能处理与识别专栏简介

刘成林¹, 金连文², 殷绪成³, 白翔⁴, 连宙辉⁵, 彭良瑞⁶, 刘禹良⁴, 陈秀妍⁷

1. 中国科学院自动化研究所多模态人工智能系统全国重点实验室, 北京 100190;
2. 华南理工大学电子与信息学院, 广州 510640;
3. 北京科技大学计算机与通信工程学院, 北京 100083;
4. 华中科技大学电子信息与通信学院, 武汉 430074;
5. 北京大学王选计算机研究所, 北京 100080;
6. 清华大学电子工程系, 北京 100084;
7. 《中国图象图形学报》编辑部, 北京 100190

文字和文档图像在工作及日常生活中无处不在, 文字承载的丰富而精确的语义信息, 在人工智能应用场景中非常重要, 如文档数字化、手写输入、图像检索、智能检测、工业自动化、机器人导航、信息安全、智能阅卷、教育信息化、电子政务、广告图片编辑、增强现实翻译等。因此, 文档图像分析与识别, 或称光学字符识别(OCR), 是极其重要的人工智能技术之一, 符合国家四个面向重大需求。

为了促进我国OCR领域的研究发展, 帮助相关研究人员了解此领域的重要学术前沿与技术进展, 《中国图象图形学报》邀请业内专家共同策划推出“文档图像智能处理与识别”专栏, 主要收录国内学者在文档图像智能处理与识别方面具有创新性、突破性的研究成果。以期为相关领域的研究人员提供参考。

经过严格评审, “文档图像智能处理与识别”专栏共收录10篇论文, 包括: 3篇综述, 1篇数据集论文, 6篇算法论文:

《文档智能分析与识别前沿: 回顾与展望》(作者: 刘成林*, 金连文, 白翔, 李晓辉, 殷飞)对文档智能分析与识别领域60多年的研究历史进行简要回顾, 重点对深度学习兴起以来的研究进展从几个主要技术环节(图像预处理、版面分析、场景文本检测、文本识别、图形符号识别和信息抽取)进行回顾和分析, 总结当前技术发展的主要特点和研究与应用中存在的不足, 最后从系统性能提升、应用扩展、学习能力增强3个角度提出未来研究方向。

《以文字为中心的图像理解技术综述》(作者: 张言, 李强, 申化文, 曾港艳, 周宇*, 马灿, 张远, 王伟平)对近年来以文字为中心的理解技术进行综述, 并按照理解认知程度, 将以文字为中心的理解任务划分成两类进行梳理, 分析了各任务、各类方法的优点和局限性。

《视觉信息抽取的深度学习方法综述》(作者: 林泽柠, 汪嘉鹏, 金连文*)对基于深度学习的视觉信息抽取算法进行归纳、流程简述和分析, 介绍了该领域的主流数据集和常用的评价指标, 并对这些算法在公开数据集上的性能进行对比。

《SCID: 用于富含视觉信息文档图像中信息提取任务的扫描中文票据数据集》(作者: 乔梁, 李再升, 程战战, 李玺*)收集、标注并公开发布了一个真实中文扫描票据数据集SCID, 展示了真实OCR技术应用场景的多项挑战, 可以为文档富视觉信息抽取相关技术领域研发和技术落地提供重要数据支持。

《针对文档图像的非对称式几何校正网络》(作者: 秦海, 李艺杰, 梁桥康*, 王耀南)提出一种集成文档区域定位与校正的轻量化几何校正网络AsymcNet, 端到端地实现文档图像的几何校正。通过实验验证了AsymcNet的有效性与先进性。

《多特征融合的敦煌古籍残片自动缀合》(作者: 郑玉彤, 李雪龙, 殷梓轩, 高歌, 翁彧*)提出了在碴口特征的基础上, 进行多层特征融合评估的古籍残片缀合方法。并基于构建的古籍残片数据集进行算法实现和模型评估, 有效提升了古籍残片缀合的效率。

《层级语义融合的场景文本检测》(作者:王紫霄,谢洪涛*,王裕鑫,张勇东)提出了层级语义融合的场景文本检测算法,旨在解决已有文本检测网络对于文字实例内的局部语义和文字实例间的全局语义提取能力不足的问题。

《用于在线手写公式合成的编解码网络》(作者:杨晨,杜俊*,薛莫白,张建树)提出基于编码—解码模型的在线数据生成方法,实现了从给定公式文本到在线轨迹点序列的生成,并且生成的数据可以更加灵活地对原有数据集进行增广。论文通过可视化结果展示了生成模型的性能,模型在不同难度的输入文本上都有良好表现,并且可以适应不同手写人风格。

《多损失融合网络的中国书法字体与风格分类》(作者:程文炎,周勇,陶承英,刘丽*,李志刚,邱桃荣)提出端到端融合多损失的孪生卷积神经网络,能同时解决中国书法字体和风格分类问题。网络由两个共享权重的分支构成,与传统孪生神经网络不同的是,本文将网络中每个分支扩展为一个分类网络。

《结合部首字形和层级结构的手写汉字纠错方法》(作者:李云青,杜俊*,胡鹏飞,张建树)从部首字形角度出发,引入了一个层级部首网络(HRN),通过注意力机制学习每个部首的表征与图像视觉特征相关区域,进而捕获包括部首字形信息的细粒度特征,这种方式能够对相似部首进行更好地区分,进而能够检测到错字中的细微错误。

我们期待广大读者和科技人员通过“文档图像智能处理与识别”专栏,能够更深入、更全面地了解该领域的最新方法和应用,吸引更多学者从事相关

研究并产生具有国际影响力的优秀成果,为本领域的发展做出新的贡献。

专栏编委会:

刘成林,中国科学院自动化研究所研究员,研究领域为模式识别、机器学习、文档分析与识别。

金连文,华南理工大学教授,研究领域为模式识别、文档分析与识别、信息抽取。

殷绪成,北京科技大学教授,研究领域为模式识别、文字识别、计算机视觉、人工智能芯片、工业智能与工业软件技术及应用。

白翔,华中科技大学教授,研究领域为模式识别、计算机视觉、文档分析与识别。

连宙辉,北京大学副教授,研究领域为计算机图形学、计算机视觉、人工智能。

彭良瑞,清华大学副研究员,研究领域为机器视觉与人机交互、智能图文信息处理等。

刘禹良,华中科技大学研究员,研究领域为文档图像智能。

专栏编辑:

陈秀妍,编辑,主要研究方向为学术出版和媒体传播等。E-mail:chenxy@aircas.ac.cn