

Universidade de Pernambuco

Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Dissertação de Mestrado

Área: Inteligência Computacional

Título: Avaliação e proposta de modelos fim a fim para transcrição de páginas de documentos

Orientador – Byron Leite Dantas Bezerra (byron.leite@upe.br)

Coorientador – Cleber Zanchettin (cz@cin.ufpe.br)

Descrição

Em sistemas OCR (*Optical Character Recognition*), a análise de layout e o reconhecimento de texto são de fundamental importância entre as etapas da extração de conteúdo automatizada de uma imagem de documento. Nesse contexto, o reconhecimento de documentos com layout simples pode ser considerado resolvido, enquanto os complexos, como os encontrados em revistas e alguns artigos técnicos, ainda representam um grande desafio [1].

Nos últimos anos, o desenvolvimento de algoritmos ponta a ponta para detecção e reconhecimento de texto a partir de imagens tem evoluído significativamente devido a utilização de modelos de aprendizado profundo (*Deep Learning*). Da mesma forma, a utilização de modelos de Inteligência Artificial em Processamento de Linguagem Natural (*Natural Language Processing*) possibilitou avanço na extração de informações a partir de texto estruturado.

Com a capacidade aprimorada desses modelos, foi possível substituir, por exemplo, os métodos clássicos de detecção, reconhecimento do texto e extração de informação de interesse por modelos fim a fim, que são treinados para ler a página inteira [2].

Nesse sentido, esse tipo de abordagem precisa ser ainda mais robusto para lidar com a complexidade da escrita cursiva e a grande diversidade dos documentos, principalmente com layout complexo, como documentos históricos e comerciais, como receitas médicas.

Dessa forma, o escopo deste projeto de mestrado tem como objetivo a investigação e desenvolvimento de modelos ponta a ponta cobrindo as etapas de detecção, reconhecimento e extração de textos em documentos de layout complexo.

A proposta envolve uma equipe multidisciplinar e faz parte do projeto de pesquisa e inovação “*Algoritmos e Modelos de Inteligência Artificial e Visão Computacional para Processamento Inteligente de Documentos*” fomentado pelo CNPQ, e em parceria com a empresa Di2Win (www.di2win.com).

Para conhecer mais sobre o orientador e seus temas de pesquisa, convido a assistir a entrevista [aqui](#).

Referências Bibliográficas

1. A. Antonacopoulos, D. Bridson, C. Papadopoulos, S. Pletschacher. A Realistic Dataset for Performance Evaluation of Document Layout Analysis. Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR2009), Barcelona, Spain, July 2009, pp. 296-300.
2. CASTRO, DAYVID; BEZERRA, BYRON LEITE DANTAS; ZANCHETTIN, CLEBER. An End-to-End Approach for Handwriting Recognition: From Handwritten Text Lines to Complete Pages. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2024, Seattle. IEEE, 2024, p.264, DOI: 10.1109/CVPRW63382.2024.00031.
3. Bezerra, B. L. D. ; Zanchettin, C., Toselli, A. H., and PIRLO, G. (Eds.). Handwritten: Recognition, Development and Analysis. New York: Nova Science Publishers, 2017.