

Universidade de Pernambuco

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

Proposta de Tese de Doutorado

Área: Inteligência Computacional

Título: Síntese de Escrita Manual para Treinamento de Modelos de Reconhecimento de Texto em Documentos Manuscritos

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

Coorientador – Arthur Flor de Sousa Neto (afsn@ecomppoli.br)

Descrição

A transcrição automática de textos manuscritos em imagens de documentos ainda apresenta desafios significativos, especialmente em cenários com escassez de dados e alta variabilidade de estilos de escrita. Embora sistemas HTR (*Handwritten Text Recognition*) tenham evoluído consideravelmente com o uso de modelos de aprendizado profundo, o desempenho de sistemas HTR em documentos complexos e escritos cursivos ainda é limitado. Isso se deve, em grande parte, à diversidade dos estilos de escrita e à falta de grandes volumes de dados anotados necessários para treinar modelos robustos [1][2].

Neste contexto, abordagens baseadas em síntese de imagens de texto manuscrito têm emergido como alternativas promissoras para contornar o problema da escassez de dados. Tais métodos têm como objetivo gerar imagens sintéticas realistas que preservem a variabilidade de conteúdo e estilo observada em dados reais. Métodos baseados em *Generative Adversarial Networks* (GANs) foram os mais explorados ao longo dos anos, seguidos por abordagens baseadas em *Diffusion Models* [3][4][5].

Apesar dos avanços recentes, a geração de textos manuscritos ainda enfrenta desafios significativos. A maioria dos modelos existentes se concentra em palavras isoladas, com poucas propostas abordando a coerência visual e linguística de sequências mais longas. Além disso, a geração de estruturas como linhas ou parágrafos sintéticos continua sendo rara e computacionalmente custosa [6].

Dessa forma, este projeto de doutorado propõe o desenvolvimento e integração de um modelo generativo baseado em *Diffusion Models* para a síntese de linhas manuscritas. O objetivo é produzir dados com variação de estilo, ruído realista e estrutura semântica coerente, visando apoiar tarefas de reconhecimento de escrita. Para conhecer mais sobre o orientador e seus temas de pesquisa, convido a assistir a entrevista [aqui](#).

Referências Bibliográficas

1. B. L. D. Bezerra, C. Zanchettin, A. H. Toselli, and G. Pirlo, "Handwriting: Recognition, Development and Analysis," Nova Science Publishers, 2017.
2. D. Coquenot, C. Chatelain, and T. Paquet, "End-to-End Handwritten Paragraph Text Recognition Using a Vertical Attention Network," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 508–524, 2023.
3. S. Fogel, H. Averbuch-Elor, S. Cohen, S. Mazor, and R. Litman, "ScrabbleGAN: Semi-Supervised Varying Length Handwritten Text Generation," in *Proc. CVPR*, 2020, pp. 4323–4332.
4. L. Kang, P. Riba, Y. Wang, M. Rusiñol, A. Fornés, and M. Villegas, "GANwriting: Content-Conditioned Generation of Styled Handwritten Word Images," in *Proc. ECCV*, 2020, pp. 273–289.
5. J. Gan, W. Wang, J. Leng, and X. Gao, "HiGAN+: Handwriting Imitation GAN with Disentangled Representations," *ACM Trans. Graph.*, vol. 42, no. 1, 2022.
6. A. F. S. Neto, B. L. D. Bezerra, G. C. D. Moura, and A. H. Toselli, "Data Augmentation for Offline Handwritten Text Recognition: A Systematic Literature Review," *SN Comput. Sci.*, 2024.