



Universidade de Pernambuco (UPE)
Escola Politécnica de Pernambuco (POLI)
Instituto de Ciências Biológicas (ICB)

Coordenação de Pós-Graduação em Engenharia de Sistemas

Proposta de Dissertação de Mestrado

Área: Cibernética

Linha de Pesquisa: **Modelagem e simulação de sistemas inteligentes e embarcados**

Título Provisório: *Uso de LLMs multimodais na curadoria de imagens para construção de datasets de plantas da Caatinga e da Mata Atlântica*

Orientador: Rodrigo de Paula Monteiro.

Co-orientador: Carmelo José Albanez Bastos Filho

Descrição:

A preservação de ecossistemas como a Caatinga e a Mata Atlântica é fundamental para a manutenção da biodiversidade, do equilíbrio climático e da qualidade de vida das populações humanas que dependem diretamente de seus recursos [1]. Esses biomas, vêm sofrendo com intensos processos de degradação, resultando na perda de espécies nativas, no empobrecimento do solo e na redução da capacidade de regeneração natural. A restauração dessas áreas, por meio de práticas como o plantio de mudas nativas, constitui uma estratégia indispensável para a recuperação dos serviços ecossistêmicos e para a mitigação das mudanças climáticas [2].

Nesse cenário, a correta identificação das espécies vegetais desempenha um papel central, tanto para orientar ações de restauração quanto para subsidiar políticas de conservação. Contudo, a identificação manual de espécies, sobretudo em estágios iniciais de crescimento, apresenta limitações devido à semelhança morfológica entre indivíduos e à necessidade de especialistas treinados [3]. O emprego de ferramentas de inteligência artificial tem se consolidado como uma abordagem promissora para o reconhecimento automático de espécies a partir de registros de imagens. Entretanto, a eficácia desses métodos está intrinsecamente condicionada à disponibilidade de bases de dados robustas e devidamente estruturadas, uma vez que a acurácia e a capacidade de generalização dos modelos dependem diretamente da diversidade, representatividade e consistência dos exemplos utilizados em seu treinamento[4].

Apesar de seu potencial, a construção de bases de dados nesse contexto enfrenta desafios significativos. Entre eles estão a dificuldade de coleta de imagens representativas em ambientes naturais, a presença de ruído decorrente de variações de luminosidade e fundo, além da necessidade de anotações precisas, que demandam tempo e conhecimento especializado [5].



Universidade de Pernambuco (UPE)
Escola Politécnica de Pernambuco (POLI)
Instituto de Ciências Biológicas (ICB)

Diante desse cenário, este projeto propõe o uso de descritores gerados por modelos de linguagem multimodais (LLMs multimodais) para a triagem de imagens que contenham plantas com características pré-definidas. A ideia central é empregar tais modelos como ferramenta auxiliar na organização e estruturação de bases de dados, reduzindo o esforço manual na seleção e categorização de imagens. Essa abordagem busca acelerar o processo de construção de datasets robustos e representativos, indispensáveis ao treinamento de modelos de identificação de espécies, contribuindo assim para o avanço de pesquisas em monitoramento da biodiversidade e para a efetividade de ações de preservação e restauração ambiental.

Referências:

- [1] MILLER, G. T.; SPOOLMAN, S. E. Environmental science. 16. ed. Boston: Cengage Learning, 2019.
- [2] ANTONGIOVANNI, M. et al. Prioritizing restoration in the Caatinga: A landscape resilience and biodiversity approach. Journal of Applied Ecology, v. 59, n. 6, p. 1424-1437, 2022. DOI: <https://doi.org/10.1111/1365-2664.1413>.
- [3] LORENZI, H. Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. 6. ed. Nova Odessa: Instituto Plantarum, 2016. v. 1.
- [4] WÄLDCHEN, J.; MÄDER, P. Plant species identification using computer vision techniques: A systematic literature review. Ecological Informatics, v. 40, p. 50-59, 2017. DOI: <https://doi.org/10.1016/j.ecoinf.2017.05.005>
- [5] XU, J.; CHEN, Y.; LIU, X. Plant image recognition with deep learning: A review. Computers and Electronics in Agriculture, v. 212, p. 108072, 2023. DOI: <https://doi.org/10.1016/j.compag.2023.108072>