



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: Telemática  
Linha de Pesquisa: **Processamento e transmissão digital da informação**  
Título Provisório: Combinações de versões modificadas do Algoritmo K-means: Aplicações em Processamento de Sinais e Imagens  
Orientador: Verusca Severo de Lima  
Co-orientador: Francisco Madeiro Bernardino Junior

### **Descrição:**

O algoritmo *K-means* (KM), também conhecido como Algoritmo *Lloyd* Generalizado (GLA), é considerado uma das técnicas mais amplamente utilizadas em problemas de agrupamento de dados (*clustering*) [1, 2]. O KM agrupa pontos de dados em  $K$  clusters, usando a pesquisa do vizinho mais próximo, em que cada ponto é alocado no *cluster* do centróide mais próximo a partir de uma medida de distância [3, 4]. O centróide é calculado como a média de todos os pontos de dados atribuídos ao *cluster*. O agrupamento é feito de forma a minimizar a soma das distâncias quadradas entre cada ponto e o centróide mais próximo [5].

O KM tem sido usado em diversos cenários de aplicação, tais como: compressão de imagens [6, 7], previsão da vida útil de motores aeronáuticos [8], previsão de doenças cardíacas [9], localização de veículos [10], segmentação de imagens médicas [11, 12], entre outros.

O projeto de mestrado ora proposto tem como objetivo principal a avaliação de combinações de versões modificadas do algoritmo *K-means*, em aplicações da área de processamento de sinais e imagens, as quais podem incluir: compressão de imagens, transmissão digital de sinais e imagens e agrupamento de dados.

### **Referências Bibliográficas:**

[1] A. M. Ikotun, *et al.* “*K-means* clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data”, *Information Sciences*, vol. 622, 2023.



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

- [2] Y. Linde, A. Buzo, e R. Gray, “An algorithm for vector quantizer design”, IEEE Transactions on Communications, vol. 28, n°. 1, pp. 84–95, 1980.
- [3] P. Fränti, e S. Sieranoja, “How much can  $K$ -means be improved by using better initialization and repeats?”, Pattern Recognition, vol. 93, pp. 95-112, 2019.
- [4] C. Tirnauca, *et al.* “Global optimality in  $K$ -means clustering”, Information Sciences, vols. 439–440, pp. 79-94, 2018.
- [5] Q. An, *et al.* “A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges”, Sensors, vol. 23, n°. 9, 2023.
- [6] Severo, V.; *et al.* “On the initialization of swarm intelligence algorithms for vector quantization codebook design”, Sensors, vol. 24, n°. 8, 2024.
- [7] C. Fonseca, F.A. Ferreira, e F. Madeiro, “Vector quantization codebook design based on fish school search algorithm”. Applied Soft Computing, vol. 73, pp. 958–968, 2018.
- [8] J. Liu, *et al.* “Prediction of remaining useful life of multi-stage aero-engine based on clustering and LSTM fusion”, Reliability Engineering & System Safety, vol. 214, 2021.
- [9] R.C., Ripan, *et al.* “A data-driven heart disease prediction model through  $K$ -means clustering-based anomaly detection”, SN Computer Science, vol. 2, n°. 112, pp. 1-12, 2021.
- [10] L., Xiangze, *et al.* “Location information collection and optimization for agricultural vehicle based on UWB”, Nongye Jixie Xuebao/Transactions of the Chinese Society of Agricultural Machinery, vol. 49, n°. 10, 2018.
- [11] R. Khilkhal and M. Ismael, "Brain tumor segmentation utilizing thresholding and  $K$ -means", 2022 Muthanna International Conference on Engineering Science and Technology (MICEST), Samawah, Iraq, pp. 43-48, 2022.
- [e] J. Qiao, *et al.* “Data on MRI brain lesion segmentation using  $K$ -means and Gaussian Mixture Model-Expectation Maximization”, Data in Brief, vol. 27, 104628, 2019.