

# Realimentação de Relevância em Memória para Consultas por Similaridade em Espaços Métricos

João Henrique V. do Carmo<sup>1</sup>, Carolina Yukari Veludo Watanabe<sup>1</sup>

<sup>1</sup>Departamento de Ciência da Computação – Universidade Federal de Rondônia (UNIR)  
Porto Velho – RO – Brasil

joaohvc1@gmail.com, carolina@unir.br

**Abstract.** *This article proposes a Relevance Feedback (RF) method for Content-Based Image Retrieval (CBIR) systems indexed by metric trees, such as Slim-Tree. Unlike the Turning Metrics Relevance Feedback (TMRF) method — which operates over secondary storage by requiring full metric-space reindexing or the use of weighted distances applied directly on the tree structure, thereby reducing pruning efficiency and increasing computational cost — the proposed approach performs RF entirely in primary memory by restricting similarity searches to a subset of  $P$  previously loaded elements retrieved via an initial unweighted query. Experimental results on the COVID-19 Radiography Database, using 5 RF iterations, demonstrate a reduction of up to 80% in query execution time compared to TMRF. Additionally, gains of up to 8 percentage points in precision were observed across the evaluated extractors, indicating a favorable trade-off between computational efficiency and retrieval quality.*

**Resumo.** *Este artigo propõe um método de Realimentação de Relevância (RF) para sistemas de Recuperação de Imagens Baseada em Conteúdo (CBIR) indexados por árvores métricas, como a Slim-Tree. Diferentemente do método Turning Metrics Relevance Feedback (TMRF) — que opera sobre armazenamento secundário ao demandar reindexação completa do espaço métrico ou aplicação de distâncias ponderadas diretamente na estrutura da árvore, comprometendo a eficiência da poda e elevando o custo computacional —, a abordagem proposta realiza a RF inteiramente em memória primária, restringindo a busca de similaridade a um subconjunto de  $P$  elementos previamente carregados por meio de uma consulta inicial sem ponderação. Os resultados experimentais sobre a base COVID-19 Radiography Database, com 5 iterações de RF, demonstram redução de até 80% no tempo de execução das consultas frente ao TMRF. Além disso, observou-se ganho de até 8 pontos percentuais na precisão nos extratores avaliados, evidenciando um trade-off favorável entre eficiência computacional e qualidade da recuperação.*

## 1. Introdução

A recuperação de imagens baseada em conteúdo (CBIR) utiliza características visuais como cor, forma e textura para indexar e buscar dados em grandes bases [Vishraj et al. 2022]. No entanto, esses sistemas enfrentam a lacuna semântica, onde as características de baixo nível extraídas computacionalmente não correspondem plenamente à percepção de alto nível do usuário [Yin et al. 2021]. Para mitigar esse problema,

técnicas de Realimentação de Relevância (RF) inserem o especialista no ciclo de busca, permitindo o ajuste iterativo dos resultados [Ahmed 2020].

Em bases de larga escala, o uso de Métodos de Acesso Métrico (MAMs), como a Slim-Tree, é essencial para garantir a eficiência das consultas por similaridade [Traina et al. 2002]. Contudo, a integração de RF em MAMs apresenta desafios significativos. A alteração dinâmica dos pesos das características durante a busca modifica a geometria do espaço métrico, o que frequentemente exige a reindexação completa da árvore [Marcacini et al. 2023].

O método Tuning Metrics Relevance Feedback (TMRF) propôs avanços na calibração dessas distâncias, mas o custo computacional de reestruturação do índice em armazenamento secundário permanece como um gargalo para aplicações em tempo real [Marcacini et al. 2023]. Diante desse cenário, este trabalho propõe uma abordagem de RF processada inteiramente em memória primária. O objetivo é realizar o refinamento da busca sobre um subconjunto de candidatos pré-recuperados, eliminando a necessidade de reindexação global ou de travessias custosas na árvore métrica com funções de distância modificadas.

## 2. Fundamentação Teórica

A arquitetura de um sistema CBIR convencional é dividida entre interface, processamento de consultas e base de imagens [Dhaygude and Kinariwala 2022]. O núcleo do processo reside na extração de características, que gera vetores numéricos (“assinaturas”) para representar o conteúdo visual [Marcacini et al. 2023].

Os MAMs organizam os dados baseando-se apenas na distância entre os elementos em um espaço métrico, que deve respeitar as propriedades de simetria, não negatividade e desigualdade triangular [Chen et al. 2022]. A Slim-Tree destaca-se entre os MAMs dinâmicos por sua capacidade de reduzir a sobreposição de nós através do algoritmo *Slim-Down*, otimizando o número de cálculos de distância e acessos ao disco durante consultas por abrangência ou de k-vizinhos mais próximos [Traina et al. 2002].

A RF permite que o sistema aprenda com as indicações de relevância fornecidas pelo usuário, refinando a medida de similaridade [Marcacini et al. 2023]. Uma abordagem comum é a ponderação de características, que atribui pesos distintos às dimensões do vetor de características para enfatizar atributos que discriminam melhor as classes de interesse [Tian 2018]. O aprendizado desses pesos pode ser realizado através do desvio padrão das amostras relevantes, conforme o algoritmo consolidado por [Rui et al. 1998].

Claude responded: O TMRF calibra a métrica de similaridade integrando RF à Slim-Tree, mas sem manter seus blocos em memória primária. O TMRF calibra a métrica de similaridade integrando RF à Slim-Tree, mas sem manter seus blocos em memória primária. Seu gargalo reside na aplicação de distâncias ponderadas à árvore já construída: sem reindexação, a eficiência da poda cai; com reindexação, o custo torna-se proibitivo. Métodos de RF em memória surgem como alternativa, restringindo o aprendizado de pesos aos candidatos mais promissores da consulta inicial.

## 3. Metodologia

Os experimentos foram conduzidos utilizando o repositório público *COVID-19 Radiography Database* [Rahman et al. 2022]. O conjunto de dados apresenta uma coleção

abrangente de radiografias torácicas, caracterizando um cenário real e desafiador para a recuperação de imagens baseada em conteúdo.

A representação visual foi realizada por cinco extratores de textura via *PyFeats* [Giakoumoglou 2021]: *Gray Histogram* (32 dimensões), *GLCM* [Haralick et al. 1973] (14 dimensões), *FOS* (16 dimensões), *LTE* (6 dimensões) e *SFM* (4 dimensões). Cada descritor gerou um índice independente com páginas de 1024 bytes, e todos os vetores foram normalizados por regularização L2 para garantir estabilidade das funções de distância.

O subconjunto de  $P$  elementos carregados em memória é determinado por uma consulta inicial à Slim-Tree via distância euclidiana sem ponderação, cujos retornos formam o *buffer* sobre o qual o Algoritmo 1 opera. A partir do *top-K* inicial — os  $K$  primeiros elementos já ordenados pela árvore —, o vetor de pesos  $W$  é calculado pelo desvio padrão inverso das dimensões dos elementos relevantes [Rui et al. 1998], e as iterações seguintes reordenam o *pool* em memória pela distância ponderada  $d_W$ , sem novos acessos à árvore.

---

**Algorithm 1** RF em Memória sobre a Slim-Tree

---

**Require:**  $Query, P, K, N_{iter}$ , índice Slim-Tree  $\mathcal{T}$

**Ensure:** *Top-K* refinado após  $N_{iter}$  iterações

```

1:  $Pool \leftarrow \mathcal{T}.knn(Query, P)$  ▷ busca inicial sem ponderação em disco
2:  $TopK \leftarrow Pool[1 .. K]$  ▷ primeiros  $K$  elementos já ordenados pela árvore
3:  $Relevant \leftarrow GetRelevant(TopK, Query)$  ▷ elementos retornados corretos
4:  $W \leftarrow ComputeWeights(Relevant)$  ▷ vetor de pesos da iteração 0
5: for  $i = 1$  to  $N_{iter} - 1$  do
6:    $TopK \leftarrow topKRetrieval(K, Pool, Query, W)$  ▷ busca ponderada em memória
7:    $Relevant \leftarrow GetRelevant(TopK, Query)$ 
8:   if  $Relevant \neq \emptyset$  then
9:      $W \leftarrow ComputeWeights(Relevant)$  ▷ atualiza pesos via desvio padrão
10:  end if
11: end for
12: return  $TopK$ 

```

---

A avaliação do método proposto foi segmentada em três baterias principais de experimentos:

1. **Avaliação de Desempenho Computacional e Precisão:** O primeiro conjunto de testes avaliou métricas críticas para Métodos de Acesso Métrico: tempo de execução, quantidade de acessos a disco e quantidade de funções de distância executadas. A eficácia foi validada rodando o algoritmo com  $P = 100$  elementos no *buffer* de memória e 5 iterações de RF para cada consulta. Cada uma das 4.000 imagens da base foi tratada como elemento de consulta. A qualidade da recuperação foi aferida por meio da métrica *Mean Average Precision* (MAP) e precisão, avaliadas nos cortes  $K = 10$  e  $K = 20$ , totalizando 4.000 execuções por extrator. Os valores da estratégia *Global* correspondem ao TMRF após reindexação com o vetor de pesos médio  $\bar{w}$  [Marcacini et al. 2023] para  $K = 10$ , e ao TMRF com distâncias ponderadas aplicadas diretamente na Slim-Tree sem reindexação

para  $K = 20$ . Os tempos para  $K = 20$  diferiram em menos de 1% em relação aos de  $K = 10$  — resultado esperado, pois o custo dominante da busca na Slim-Tree não depende de  $K$  — e são reportados apenas para  $K = 10$  por concisão.

2. **Análise de Trade-off Dimensional:** O segundo experimento consistiu na variação do hiperparâmetro  $P$  (quantidade de imagens submetidas à reavaliação de similaridade após a RF). O objetivo dessa variação foi construir uma curva de observação do *trade-off* existente entre o custo sistêmico (consumo de memória e tempo de processamento) e os ganhos marginais de precisão obtidos na recuperação.

Todos os algoritmos avaliados neste trabalho foram integralmente desenvolvidos na linguagem C++. Cumpre destacar que os resultados numéricos absolutos apresentados por [Marcacini et al. 2023] não puderam ser fielmente reproduzidos devido a dois fatores: o acesso restrito às sementes experimentais e à amostragem exata de imagens empregada no estudo base, e a ausência de tratamento por normalização L2 nos vetores de características de textura originais. Como a referida normalização foi introduzida nesta pesquisa para mitigar distorções de escala e estabilizar as funções de distância nos espaços métricos, tornou-se imperativo reexecutar tanto a abordagem base quanto as extensões propostas sob as mesmas condições de contorno, viabilizando uma análise comparativa rigorosa e justa.

## 4. Resultados

Os resultados são apresentados comparando a estratégia *Global* — equivalente ao TMRF [Marcacini et al. 2023] — e a estratégia *Em Memória* aqui proposta, sob as condições experimentais descritas na Seção 3.

### 4.1. Desempenho Computacional e Precisão

As Tabelas 1 e 2 apresentam os resultados de MAP e precisão para  $K = 10$  e  $K = 20$ , respectivamente. A Tabela 3 reporta os tempos de execução e a economia obtida para  $K = 10$ . Em todos os extratores e cortes avaliados, a abordagem em memória superou a estratégia global tanto em qualidade de recuperação quanto em eficiência computacional.

**Tabela 1. MAP e Precisão (Global vs. Em Memória) para  $K = 10$ . Médias sobre 4.000 consultas.**

| Extrator       | MAP Global | MAP Mem.      | Prec. Global | Prec. Mem.    |
|----------------|------------|---------------|--------------|---------------|
| Gray Histogram | 76,01%     | <b>80,31%</b> | 65,59%       | <b>70,09%</b> |
| FOS            | 64,75%     | <b>66,96%</b> | 51,58%       | <b>53,44%</b> |
| GLCM           | 54,15%     | <b>58,19%</b> | 41,42%       | <b>45,15%</b> |
| SFM            | 61,05%     | <b>62,12%</b> | 49,77%       | <b>50,60%</b> |
| LTE            | 58,42%     | <b>59,62%</b> | 45,91%       | <b>47,15%</b> |

Em termos de precisão, a abordagem em memória supera a estratégia global em todos os extratores e cortes avaliados. Para  $K = 10$  (Tabela 1), os ganhos de MAP frente ao TMRF pós-reindexação variam de +1,07 p.p. (SFM) a +4,30 p.p. (Gray Histogram). Para  $K = 20$  (Tabela 2), os ganhos são ainda mais expressivos: o extrator FOS atinge +7,53 p.p. de MAP e +8,22 p.p. de precisão, evidenciando que a RF em memória é especialmente eficaz quando o corte de recuperação é maior e a vizinhança inicial captura mais instâncias informativas.

**Tabela 2. MAP e Precisão (Global vs. Em Memória) para  $K = 20$ . Médias sobre 4.000 consultas.**

| Extrator       | MAP Global | MAP Mem.      | Prec. Global | Prec. Mem.    |
|----------------|------------|---------------|--------------|---------------|
| Gray Histogram | 72,53%     | <b>76,20%</b> | 62,98%       | <b>66,80%</b> |
| FOS            | 54,26%     | <b>61,79%</b> | 41,54%       | <b>49,76%</b> |
| GLCM           | 49,95%     | <b>54,14%</b> | 41,53%       | <b>43,40%</b> |
| SFM            | 55,20%     | <b>57,88%</b> | 46,68%       | <b>48,71%</b> |
| LTE            | 51,27%     | <b>55,23%</b> | 43,13%       | <b>45,20%</b> |

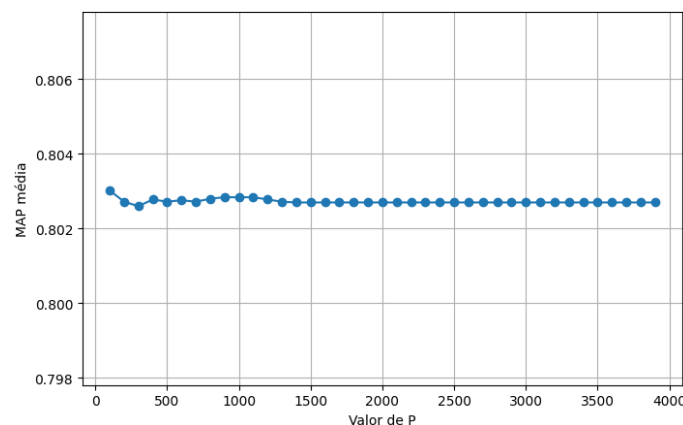
**Tabela 3. Tempos de execução total (s) e economia da abordagem em memória frente ao TMRF para  $K = 10$ . Médias sobre 4.000 consultas.**

| Extrator       | T. Global (s) | T. Mem. (s) | Economia      |
|----------------|---------------|-------------|---------------|
| Gray Histogram | 78,35         | 16,91       | <b>78,43%</b> |
| FOS            | 17,15         | 3,44        | <b>79,94%</b> |
| GLCM           | 12,45         | 2,63        | <b>78,88%</b> |
| SFM            | 1,28          | 0,49        | <b>61,72%</b> |
| LTE            | 4,96          | 1,03        | <b>79,23%</b> |

Em termos de eficiência (Tabela 3), a economia de tempo variou entre 61,72% (SFM) e 79,94% (FOS). O extrator SFM, de menor dimensionalidade (4 atributos), registrou a menor economia: seu custo basal de busca já é reduzido, diminuindo o ganho relativo da restrição ao buffer  $P$ . Os demais extratores convergem para economias próximas a 78–80%, indicando que o método escala bem com a dimensionalidade do espaço de características.

#### 4.2. Análise do *Trade-off* com o Parâmetro $P$

A variação do hiperparâmetro  $P$  (número de elementos submetidos à reavaliação em memória) permite controlar o equilíbrio entre o custo computacional e a qualidade da recuperação por similaridade.



**Figura 1. Comportamento do MAP em função do parâmetro  $P$ , demonstrando estabilidade e oscilações minúsculas de precisão entre  $P = 100$  e  $P = 4000$ .**

A análise experimental do impacto do parâmetro  $P$  revelou uma estabilização precoce na eficácia da recuperação. Ao variar o tamanho do subconjunto em memória a

partir de  $P = 100$ , com incrementos constantes até atingir o total da base de dados em  $P = 4000$ , constatou-se uma variação minúscula na precisão em todos os cenários avaliados, caracterizada por oscilações residuais insignificantes de aumento ou diminuição no MAP. Esse comportamento é evidenciado na Figura 1 para o extrator *gray histogram*, onde a qualidade da busca permanece praticamente constante ao longo de toda a escala.

Por outro lado, o custo computacional demonstra forte sensibilidade ao tamanho do subconjunto. Como a RF em memória consiste em reordenar o *pool* calculando  $P$  distâncias ponderadas por iteração, seu custo é estritamente proporcional a  $P$  — comportamento esperado de uma operação  $O(P)$  sem qualquer componente de indexação ou poda. Esses resultados validam a escolha de  $P = 100$  como configuração padrão ideal: ela preserva os ganhos de precisão da RF e minimiza o tempo de processamento em tempo real.

## 5. Considerações Finais

Este trabalho propôs um método de Realimentação de Relevância (RF) em memória para sistemas CBIR indexados por árvores métricas. A abordagem restringe o aprendizado de pesos ao subconjunto dos  $P$  candidatos mais próximos recuperados pela Slim-Tree via distância euclidiana sem ponderação, executando os ciclos de RF inteiramente em memória primária e eliminando a necessidade de reindexação ou de travessias com funções de distância modificadas em armazenamento secundário.

Os resultados experimentais sobre a base *COVID-19 Radiography Database* [Rahman et al. 2022] demonstraram que o método alcança economias de tempo entre 62% e 80% em relação ao TMRF com reindexação [Marcacini et al. 2023], superando-o em MAP em todos os extratores avaliados para  $K = 10$ . O extrator FOS destacou-se com 66,96% de MAP frente a 64,75% do TMRF pós-reindexação, enquanto o Gray Histogram exibiu a maior MAP absoluta (80,31% para  $K = 10$ , contra 76,01% do TMRF). A análise do hiperparâmetro  $P$  confirmou a existência de um ponto de estabilização, a partir do qual o aumento do buffer não produz melhoras expressivas na precisão, indicando que um valor moderado de  $P$  é suficiente para obter os benefícios semânticos da RF sem comprometer a eficiência.

Em termos de limitações, o método está sujeito ao chamado *viés de cobertura*: a qualidade da RF depende diretamente dos elementos presentes no buffer inicial. Se a consulta inicial recuperar poucos exemplares relevantes — situação mais provável em espaços de alta dimensionalidade ou com descritores de baixa discriminabilidade —, o vetor de pesos aprendido pode não capturar adequadamente a intenção do usuário.

Trabalhos futuros preveem: (i) a validação do método em bases de imagens de domínios genéricos — como conjuntos de fotografias naturais —, a fim de verificar se os ganhos obtidos em imagens radiográficas se generalizam para outros contextos visuais; (ii) a avaliação da escalabilidade do método em bases significativamente maiores, forçando a Slim-Tree a exceder a memória primária e permitindo uma análise da verdadeira fronteira computacional do algoritmo; (iii) a investigação de estratégias adaptativas de seleção do *buffer*, incluindo técnicas de diversificação dos candidatos carregados em memória; e (iv) a integração com estratégias de fusão de múltiplos descritores (*late fusion*), como direção promissora para ampliar a cobertura semântica sem aumentar o custo de indexação.

## Referências

- Ahmed, A. (2020). Implementing relevance feedback for content-based medical image retrieval. *IEEE Access*, 8:79969–79976.
- Chen, L., Gao, Y., Song, X., Li, Z., Zhu, Y., Miao, X., and Jensen, C. S. (2022). Indexing metric spaces for exact similarity search. *ACM Computing Surveys*, 55(6):1–39.
- Dhaygude, M. A. and Kinariwala, S. (2022). A literature survey on content-based information retrieval. *J. Comput. Technol. (JCT)*, 11:1–6.
- Giakoumoglou, N. (2021). PyFeats: Open source software for image feature extraction. <https://github.com/giakou4/pyfeats>.
- Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, (6):610–621.
- Marcacini, R. M. et al. (2023). Calibração de distância em métodos de acesso métrico por meio de realimentação de relevância. In *Proceedings of the 38th Brazilian Symposium on Databases (SBBD)*, pages 89–101.
- Rahman, T. et al. (2022). COVID-19 Radiography Database. Kaggle. Disponível em: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>.
- Rui, Y., Huang, T. S., Ortega, M., and Mehrotra, S. (1998). Relevance feedback: a power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(5):644–655.
- Tian, D. (2018). A review on relevance feedback for content-based image retrieval. *J. Inf. Hiding Multim. Signal Process.*, 9(1):108–119.
- Traina, C., Traina, A., Faloutsos, C., and Seeger, B. (2002). Fast indexing and visualization of metric data sets using slim-trees. *IEEE Transactions on Knowledge and Data Engineering*, 14(2):244–260.
- Vishraj, R., Gupta, S., and Singh, S. (2022). A comprehensive review of content-based image retrieval systems using deep learning and hand-crafted features in medical imaging. *Computers and Electrical Engineering*, 104:108450.
- Yin, C., Tang, J., Yuan, T., Xu, Z., and Wang, Y. (2021). Bridging the gap between semantic segmentation and instance segmentation. *IEEE Transactions on Multimedia*, 24:4183–4196.