



**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA**

Autorizada pelo Decreto Federal nº 77.496 de 27/04/76

Recredenciamento pelo Decreto nº 17.228 de 25/11/2016



**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**

**COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA**

## **XXIII SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS SEMANA NACIONAL DE CIENTÍFICA E TECNOLÓGICA - 2019**

### **FUSÃO DE RANKINGS PARA RECUPERAÇÃO DE IMAGENS**

**Wanderson Bezerra da Silva<sup>1</sup>; Rodrigo Tripodi Calumby<sup>2</sup>**

1. Bolsista PIBIC/CNPq, Graduando em Engenharia da Computação, UEFS, e-mail: wbsilva@ecomp.uefs.br

2. Orientador, Departamento de Ciências Exatas, UEFS, e-mail: rtcalumby@uefs.br

**PALAVRAS-CHAVE:** fusão de rankings, recuperação de imagens

### **INTRODUÇÃO**

Dada a complexidade da consulta ou a pouca informação disponível para a configuração de um sistema de busca de imagens com realimentação de relevância, o conjunto inicial de resultados, anterior à primeira interação do usuário, pode não apresentar informações relevantes suficientes. Isso pode acarretar em um feedback com pouca informação e consequente dificultar a etapa de aprendizado. Por ser uma abordagem interativa, a influência do primeiro resultado propaga-se para as demais iterações, dado que o resultado inicial ruim limita a troca de informação entre o usuário e o sistema (Calumby *et al*, 2017) e consequentemente os modelos de aprendizado da intenção do usuário. Visando atenuar este problema, é imprescindível que o resultado da primeira interação seja o melhor possível.

Um abordagem proposta para este cenário, conhecida como agregação de rankings (Lin, 2010), é a combinação de resultados obtidos a partir de diferentes critérios de ranking, por exemplo, diferentes características visuais das imagens ou métricas de ranking baseadas no texto associado à elas. Esta fusão pode ser realizada de diferentes formas, incluindo algoritmos de agregação de rankings e algoritmos de reranqueamento (Mei *et al*, 2014). Utilizar métodos de fusão de rankings, de modo geral, permitem resultados superiores à utilização dos critérios de modo isolado ou com técnicas simples de combinação de escores de relevância. Considerando a importância de um bom conjunto inicial de resultados nos sistemas de recuperação interativa, este trabalho avalia experimentalmente a redução do *cold-start* por meio da exploração de métodos de agregação de rankings.

### **METODOLOGIA**

As técnicas de agregação de rankings podem ser divididas em duas categorias principais: baseadas em escores ou baseadas em ordem. No primeiro grupo, a função de agregação leva como entrada as informações de pontuação dos objetos de cada lista. Na segunda, apenas a ordem relativa entre os documentos é considerada (Muñoz *et al*, 2015). Dentre as técnicas baseadas em escores, pode-se destacar a família Comb\*

(Shaw & Fox, 1994) (e.g., CombMIN, CombMAX, CombSUM, CombMED, CombMNZ e CombANZ). Em relação aos métodos baseados em posição, pode-se citar o Median Rank Aggregation (MRA) (Fagin *et al*, 2003), Reciprocal Rank Fusion (RRF) (Cormack *et al*, 2009) e Borda Count (Young, 1974).

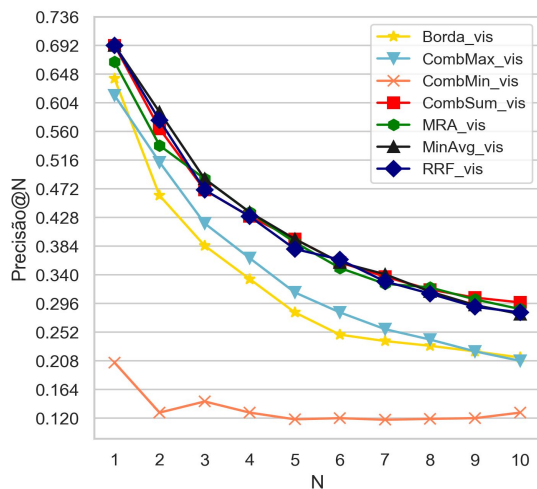
Para melhorar o ranking inicial de um sistema de recuperação interativa de imagens, (Calumby *et al*, 2014) apresenta um método de agregação baseado em escores. Neste método, a similaridade entre dois objetos é definida como o valor médio de todas as medidas de similaridade disponíveis considerando múltiplas características visuais e textuais. Para consultas com mais de um objeto, os itens de coleção são classificados com base no valor mínimo de distância para cada objeto da consulta.

Nesse trabalho foi utilizada a coleção ImageCLEF Photographic Retrieval Task que é formada por 20.000 imagens heterogêneas, em que cada imagem está associada a metadados textuais, como título, data e descrição dos conteúdos semânticos e visuais da imagem. Esta base de dados conta com 39 consultas compostas por um fragmento de texto e três imagens de exemplo para cada. Os experimentos basearam-se em sete descritores visuais globais (Penatti *et al*, 2012) baseados em cor (GCH, BIC, ACC e JAC) e textura (CCOM, LAS, e QCCH). Para a modalidade textual, utilizou-se seis medidas de similaridade entre o texto da consulta e a descrição das imagens, sendo elas: Cosseno, BM25 (Baeza-Yates & Ribeiro-Neto, 2008), Dice, Jaccard (Lewis *et al*, 2006), tf-idf-sum (dos Santos *et al*, 2009) e Bag-of-words (intersecção de termos normalizada). Foram aplicadas as técnicas de remoção de stop-words e stemização. Apenas os metadados em inglês foram considerados.

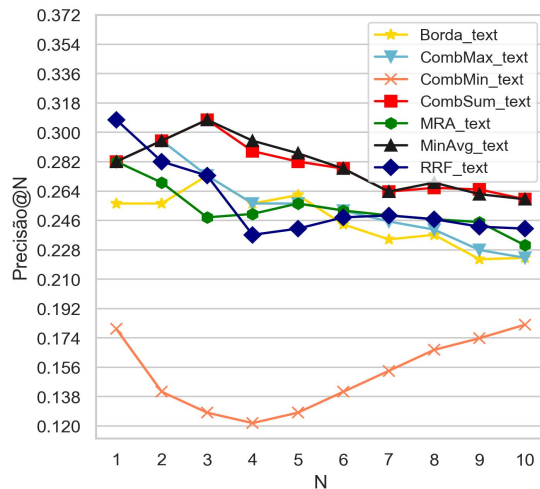
Nos experimentos foram utilizadas métodos de agregação baseados em escores e em ordem. Dentre as métricas baseadas em escores, aplicou-se os métodos da família Comb\*, sendo eles: CombMAX, CombMIN, CombSUM, CombANZ, CombMNZ, e CombMED. Dentre as medidas baseadas em ordem, considerou-se: MRA, RRF e Borda. Para cada um dos descritores, foi gerado um ranking das 20.000 imagens da base. Estes rankings foram usados como entrada para os métodos de agregação. Foram definidas 3 configurações de experimentos. A primeira considerou apenas os descritores visuais, enquanto na segunda considerou-se apenas os descritores textuais. O terceiro experimento utilizou uma configuração multimodal, utilizando visuais e textuais.

## **RESULTADOS E DISCUSSÃO**

Considerando que cada um dos rankings gerado pelos descritores contém todas as imagens de base de dados, o método CombSUM e suas derivações (CombMED, CombANZ e CombMNZ) apresentaram o mesmo resultado, visto que as derivações diferenciam-se especificamente pelo modo como levam em consideração a quantidade de rankings onde a imagem está presente. Assim, apresentamos aqui apenas os resultados do método CombSUM. Os rankings obtidos com os métodos de agregação foram comparados com aqueles obtidos com o método proposto por (Calumby *et al*, 2014) (aqui denominado MinAvg). Para avaliação, utilizou-se curvas de Precisão (P@N). A Figura 1 apresenta os resultados obtidos com os descritores visuais. A Figura 2 apresenta o comparativo considerando os descritores textuais.

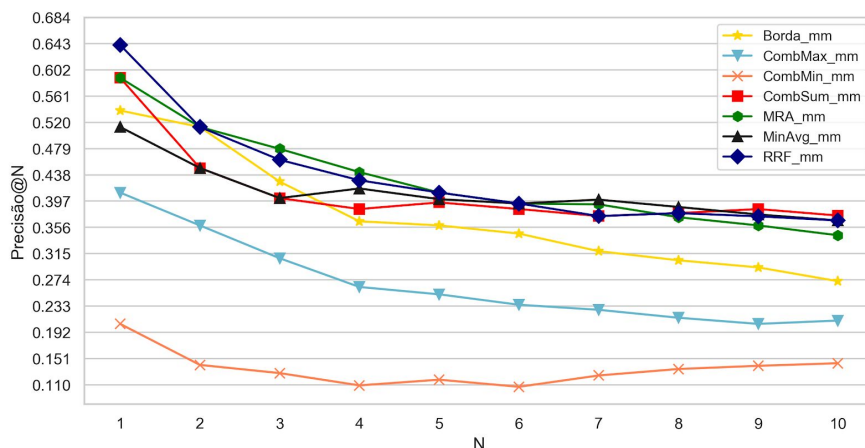


**Figura 1.** Comparativos dos resultados utilizando apenas descritores visuais.



**Figura 2.** Comparativos dos resultados utilizando apenas descritores textuais.

Os resultados demonstra que os métodos avaliados, quando utilizaram rankings baseados em apenas uma modalidade (visual ou textual), não conseguiram obter resultados superiores ao baseline. Entretanto, ao analisar os resultados na Figura 1, percebe-se há uma sobreposição em termos de  $P@N$  entre o MinAvg e os métodos CombSUM, MRA, RRF. Na Figura 2, pode-se perceber que também há uma equivalência entre o MinAvg e o CombSUM. A Figura 3 apresenta os resultados obtidos com cada um dos métodos de agregação com a configuração multimodal. Percebe-se que os métodos MRA e RRF apresentaram ganhos expressivos em termos de  $P@N$  em relação ao MinAvg nas primeiras posições do ranking. Além dos ganhos numéricos em termos de  $P@N$  nas primeiras posições, verificou-se com 95% de confiança por meio do teste de Wilcoxon que esses ganhos são estatisticamente significativos. Para as demais posições do rankings os métodos foram estatisticamente equivalentes.



**Figura 3.** Comparativos dos resultados para a configuração multimodal.

## CONSIDERAÇÕES FINAIS

Neste trabalho, apresentamos uma avaliação experimental de técnicas de agregação de rankings em cenários de buscas textuais, visuais e multimodais. Considerando a abordagem multimodal, resultados promissores foram alcançados em relação ao

baseline. Os ganhos no topo do ranking são benéficos ao usuário, pois uma maior presença de imagens relevantes nas primeiras posições permite ao usuário fornecer um bom feedback sem a necessidade de inspecionar para todos os itens. Vale destacar que nas configurações utilizadas nos experimentos, cada um dos descritores gerou um ranking utilizando todas as 20.000 imagens contidas na base de dados. Portanto, uma nova etapa de experimentação pode ser realizada para avaliar a qualidade dos métodos de agregação quando apenas os top-k itens são utilizados como entrada para os métodos de agregação.

## REFERÊNCIAS

- BAEZA-YATES, R. AND RIBEIRO-NETO, B. 2008. *Modern Information Retrieval: The Concepts and Technology Behind Search*. USA, 2nd edition.
- CALUMBY, R. T., DA S. TORRES, R., AND GONÇALVES, M. A. 2014. Multimodal retrieval with relevance feedback based on genetic programming. *MTAP*, (69):991–1019.
- CALUMBY, R. T., GONÇALVES, M. A., AND TORRES, R. da S. 2017. Diversity-based interactive learning meets multimodality. *Neurocomputing*, (259):159–175.
- CORMACK, G. V., CLARKE, C. L. A., AND BUETTCHER, S. 2009. Reciprocal rank fusion outperforms condorcet and individual rank learning methods. In *SIGIR*, pages 758–759. ACM.
- DOS SANTOS, K. C. L., DE ALMEIDA, H. M., GONÇALVES, M. A., AND DA SILVA TORRES, R. 2009. Recuperação de imagens da web utilizando múltiplas evidências textuais e programação genética. In *XXIV SBBD*, pages 91–105.
- FAGIN, R., KUMAR, R., AND SIVAKUMAR, D. 2003. Efficient similarity search and classification via rank aggregation. In *SIGMOD*, pages 301–312. ACM.
- LEWIS, J., OSSOWSKI, S., HICKS, J., ERRAMI, M., AND GARNER, H. R. 2006. Text similarity: an alternative way to search MEDLINE. *Bioinformatics*, 22(18):2298–2304.
- LIN, S. 2010. Rank aggregation methods. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(5):555–570.
- MEI, T., RUI, Y., LI, S., AND TIAN, Q. 2014. Multimedia search reranking: A literature survey. *ACM Comput. Surv.*
- PENATTI, O. A., VALLE, E., AND TORRES, R. da S. 2012. Comparative study of global color and texture descriptors for web image retrieval. *Journal of Visual Communication and Image Representation*, 23(2):359 – 380.
- SHAW, J. A. AND FOX, E. A. 1994. Combination of multiple searches. In *TREC-2*, pages 243–252.
- VARGAS MUÑOZ, J. A., DA SILVA TORRES, R., AND GONÇALVES, M. A. 2015. A soft computing approach for learning to aggregate rankings. In *CIKM*, pages 83–92.
- YOUNG, H. 1974. An axiomatization of borda's rule. *Journal of Economic Theory*, 9(1):43 – 52.