

Resultados Complementares

Nesta seção é abordado resultados complementares da Seção 5 do artigo *Calibração de Distância em Métodos de Acesso Métrico por meio de Realimentação de Relevância*. Em nosso segundo experimento abordamos sobre o efeito da inclusão da função de distância ponderada no método de acesso métrico, como complemento, realizamos também esta análise sobre um *Toy Dataset* com 50.000 elementos gerada por distribuição gaussiana, adicionando também a análise comparando o número de cálculos de distâncias e acessos ao disco entre ponderar constantemente a realização de cada consulta e ponderar em períodos intervalados a cada cinquenta consultas. Na Figura 1 é possível observar que a ponderação não constante da função de distância possui uma pequena variação menor de realizações de cálculos. A razão de não ter grande diferença acontece pois em ambos os casos como não ocorre uma reindexação da árvore métrica, a árvore métrica é recalculada para o ultimo vetor pesos informado como parâmetro para a consulta.

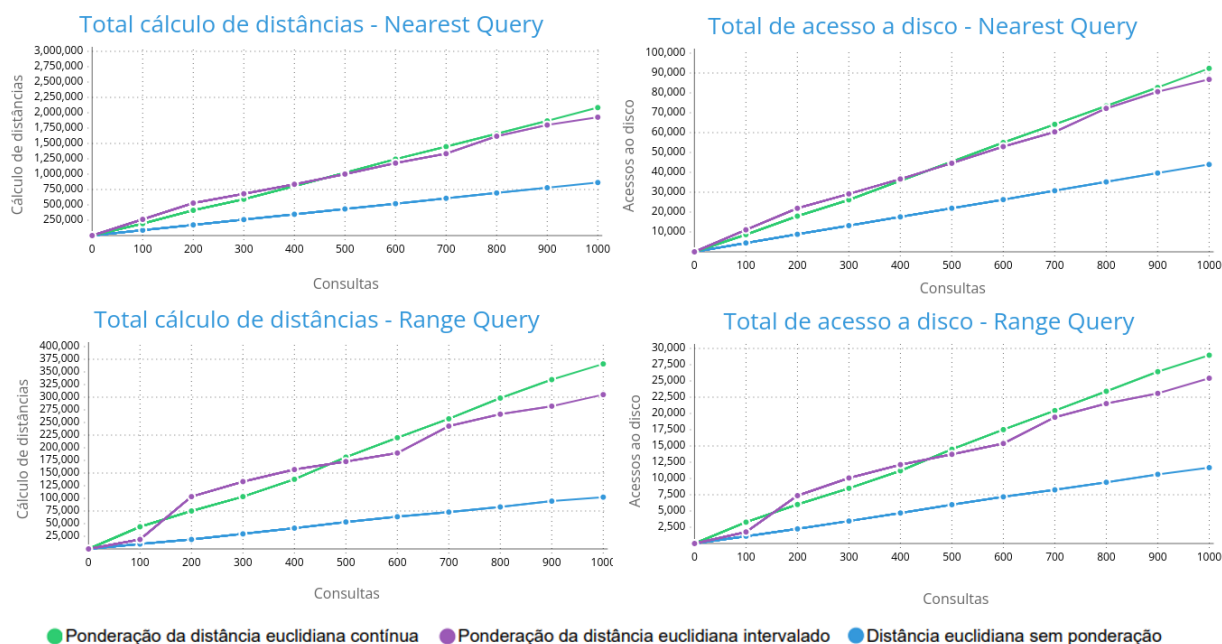


Figura 1 – Comparação dos cálculos de distâncias e acessos a disco entre a função de distância padrão e ponderação da função de distância.

Na Tabela 4 comparamos a **acurácia** (ACC) e **F1-score** (F1) obtidas entre a base de dados original e a base dados após a realização de um ciclo de RR, demonstrando como a realimentação de relevância que aprende um vetor de pesos consegue aprimorar um espaço métrico, aplicando uma melhor semântica na organização dos elementos. Para isto, é realizado a validação cruzada de até 10 pastas utilizando o classificador KNN. É possível observar que a RR consegue um ganho considerável na maioria das combinações, obtendo 83,75% de acurácia no extrator GLCM com um ganho de 47% em comparação ao não uso de Realimentação de Relevância na base de COVID-19.

Tabela 4. Acurácia e F1-Score por descritor de textura, sem RR e com RR Desvio Padrão

| Dataset | Extrator | Sem RR | | RR Desvio Padrão | | Ganho (%) | |
|-------------|-----------|---------|--------|------------------|--------|-----------|--------|
| | | ACC (%) | F1 (%) | ACC (%) | F1 (%) | ACC (%) | F1 (%) |
| COREL-1000 | Gray Hist | 54,55 | 53,61 | 57,11 | 56,18 | 5 | 5 |
| | GLCM | 80 | 80,62 | 80 | 80,33 | 0 | 0 |
| COVID-19 | GLCM | 57,12 | 57,04 | 83,75 | 83,78 | 47 | 47 |
| | FOS | 66,29 | 66,12 | 73,75 | 72,89 | 11 | 10 |
| | SFM | 68,04 | 67,65 | 80,58 | 80,12 | 18 | 18 |
| | LTE | 58,83 | 58,22 | 70,58 | 68,9 | 20 | 18 |
| BUSI-BREAST | Gray Hist | 63,5 | 66,07 | 69 | 67,36 | 9 | 2 |
| | GLCM | 60 | 61,57 | 71,25 | 65,09 | 19 | 6 |
| | FOS | 58,75 | 61 | 61,87 | 62,36 | 5 | 1 |
| | SFM | 57,49 | 61 | 67,5 | 68,52 | 17 | 12 |
| | LTE | 65,62 | 66,19 | 70,71 | 69,88 | 8 | 6 |

Em nosso terceiro experimento realizamos a análise da técnica RR Desvio Padrão através de ciclos de realimentação calculando o *Mean Average Precision* nas *top 10* e *20* imagens considerando todos elementos das bases de imagens analisadas (COREL-1000, COVID-19 e BUSI-BREAST). Como complemento a esta análise, realizamos o *Mean Average Precision* nas *top 10* e *20* imagens em todas as bases de imagens analisando o aprimoramento da precisão separados por cada classe de cada base de imagens. Na Figura 2 é possível observar que a técnica consegue aprimorar ou empatar a precisão em todas as classes analisadas, nunca piorando a precisão.

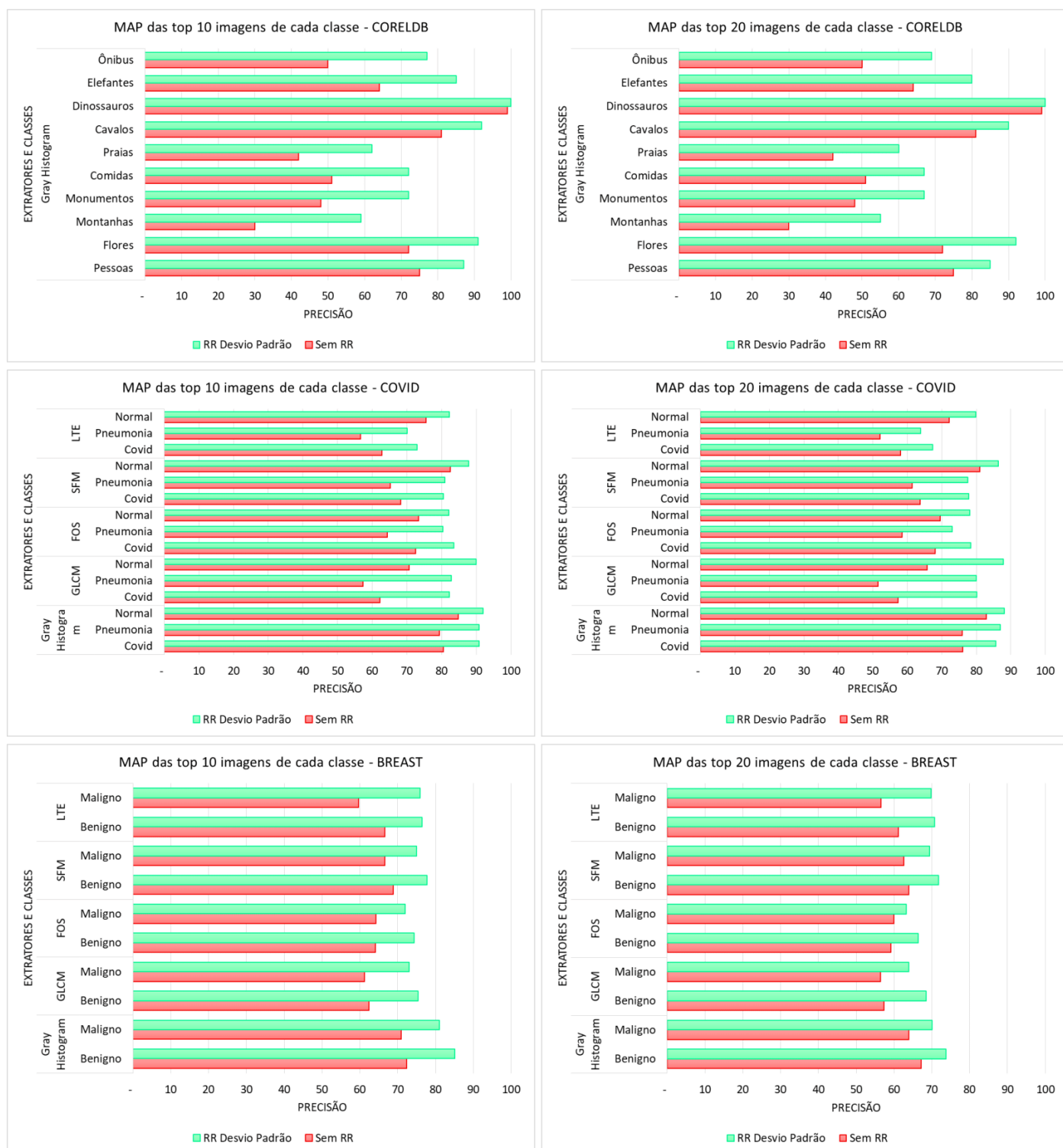


Figura 2 - Mean Average Precision das top 10 e 20 imagens por classe nas bases de imagens em análise.

Em nosso último experimento realizamos a visualização das bases de imagens utilizando T-SNE comparando o espaço métrico original com o espaço métrico transformado pela RR Desvio Padrão sobre a base de imagens de COVID. Em complemento a esta análise, realizamos a visualização comparando os espaços métricos transformados em todas as bases de imagens analisadas (COREL-1000, COVID-19, BUSI-BREAST), sendo possível observar o melhoramento semântico obtido pela técnica como apresentado na Figura 3.

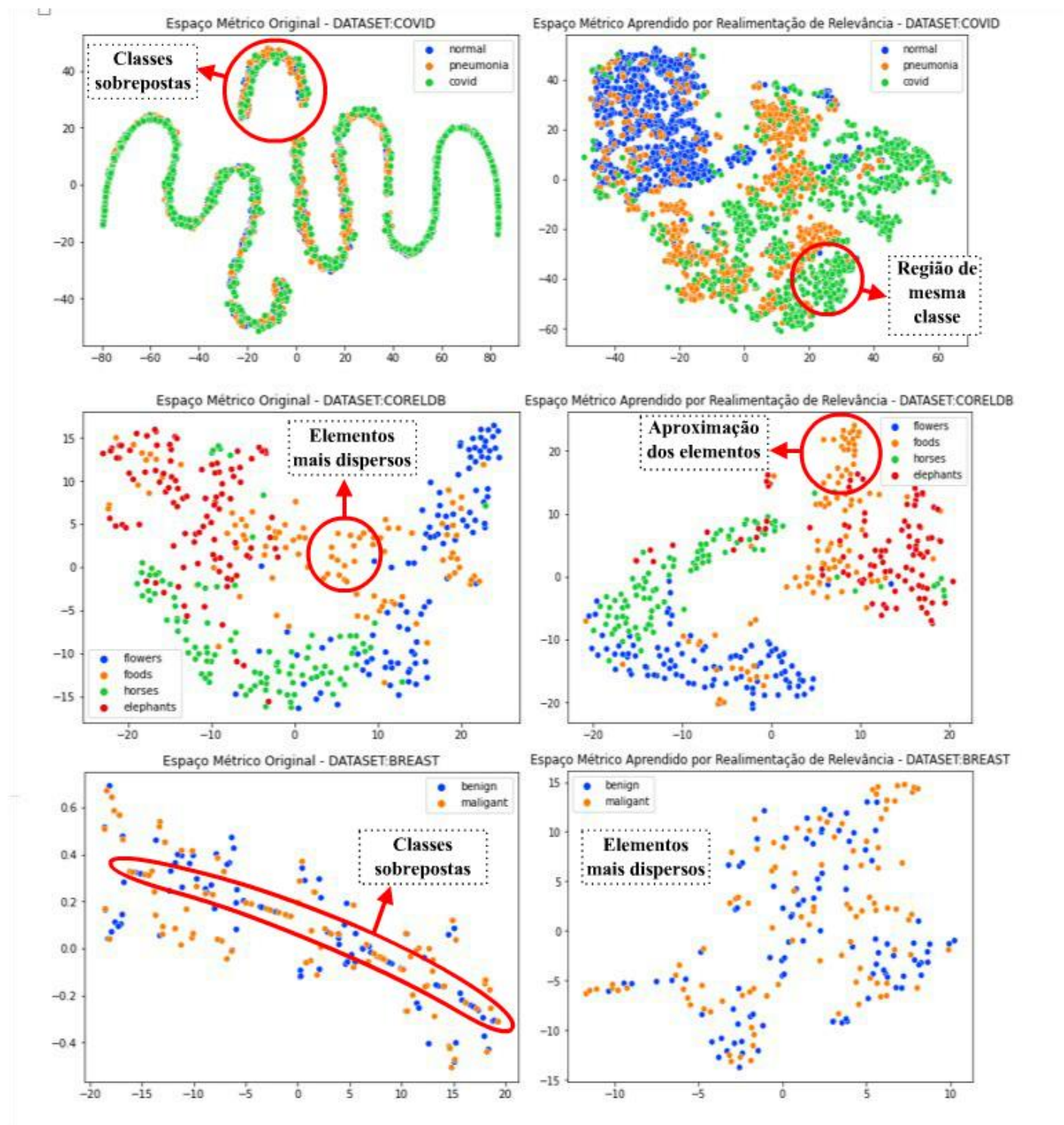


Figura 3 - Visualização da modificação do espaço métrico obtido pela técnica RR Desvio Padrão após 5 ciclos de realimentação, em cada base de imagens em análise.