

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**Desenvolvimento de Método de Acesso Métrico com
Calibração de Distância Perceptual para Consultas por
Similaridade de Imagens Médicas**

Renato Gomes Marcacini

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências
de Computação e Matemática Computacional (PPG-C²MC)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Renato Gomes Marcacini

Desenvolvimento de Método de Acesso Métrico com Calibração de Distância Perceptual para Consultas por Similaridade de Imagens Médicas

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientadora: Profa. Dra. Agma Juci Machado Traina

USP – São Carlos
Janeiro de 2024

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

G313d Gomes Marcacini, Renato
 Desenvolvimento de Método de Acesso Métrico com
 Calibração de Distância Perceptual para Consultas por
 Similaridade de Imagens Médicas / Renato Gomes
 Marcacini; orientadora Agma Juci Machado Traina. --
 São Carlos, 2024.
 100 p.

 Dissertação (Mestrado - Programa de Pós-Graduação
 em Ciências de Computação e Matemática
 Computacional) -- Instituto de Ciências Matemáticas
 e de Computação, Universidade de São Paulo, 2024.

 1. Método de Acesso Métrico. 2. Realimentação de
 Relevância. 3. Funções de Distância Ponderada. 4.
 Consultas por Similaridade. 5. CBIR. I. Juci
 Machado Traina, Agma, orient. II. Título.

Renato Gomes Marcacini

**Development of a Metric Access Method with Perceptual
Distance Calibration for Medical Image Similarity Queries**

Master dissertation submitted to the Instituto de
Ciências Matemáticas e de Computação – ICMC-
USP, in partial fulfillment of the requirements for the
degree of the Master Program in Computer Science
and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and
Computational Mathematics

Advisor: Profa. Dra. Agma Juci Machado Traina

USP – São Carlos
January 2024

Dedico este trabalho aos meus pais, cujo apoio e encorajamento inabaláveis tornaram possível esta jornada acadêmica.

AGRADECIMENTOS

Gostária de agradecer aos meus pais e familiares pelo apoio incondicional em cada etapa da minha jornada acadêmica. Seu amor, incentivo e compreensão foram a base fundamental para que eu pudesse enfrentar os desafios e alcançar meus objetivos.

Agradeço a minha orientadora Agma Juci Machado Traina, sou imensamente grato por sua orientação, paciência e acolhimento durante esta jornada acadêmica. Minha gratidão por todos ensinamentos e conselhos.

Também desejo agradecer ao meu amigo Willian Dener de Oliveira, cuja colaboração foi vital para o êxito desta dissertação. Suas discussões e ideias compartilhadas durante os momentos de estudo e pesquisa foram essenciais para realização desta dissertação.

Agradeço Grupo de Bases de Dados e de Imagens (GBDi), cujo ambiente de pesquisa e troca de conhecimentos proporcionou uma experiência gratificante e inspiradora. O apoio e o compartilhamento de ideias com os membros do grupo foram cruciais para o desenvolvimento deste estudo.

Por fim, agradeço também às agências de fomento Coordenação de Aperfeiçoamento de Pessoal de Níveis Superior (CAPES) e Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) sob número do processo 2021/00366-1 e 2016/17078-0, por concederem a bolsa de estudos que viabilizou financeiramente a realização deste mestrado. Agradeço por acreditarem em meu potencial e por investirem na minha formação acadêmica.

*“O trabalho duro supera o talento quando o talento não trabalha duro.”
(Tim Notke)*

RESUMO

MARCACINI, R. G. **Desenvolvimento de Método de Acesso Métrico com Calibração de Distância Perceptual para Consultas por Similaridade de Imagens Médicas**. 2024. 100 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2024.

Indexar e recuperar informações organizadas em bases de dados, de modo eficiente e preciso para responder às consultas solicitadas pelos usuários, são pilares do processo de engenharia e ciência de dados. No contexto médico, o tratamento e extração de semântica do dado complexo do tipo imagem, são fundamentais para tomadas de decisões. Esse tipo de tratamento demanda extrair a essência (características) desses dados e usualmente efetuar consultas por similaridade sobre tais características, em vez de usar o dado complexo em si. Os Métodos de Acesso Métrico (MAMs) foram desenvolvidos como ferramentas fundamentais para processar consultas por similaridade de modo eficiente apoiando os sistemas de gerenciamento de bases de dados. Os MAMs utilizam funções de distância fixas para realizar a construção da árvore métrica, o que por sua vez impede que um MAM consiga indexar os elementos utilizando duas ou mais funções de distância na mesma indexação. Em muitos casos, o uso de funções de distância tradicionais falha em oferecer respostas compatíveis com a percepção de distância dos usuários. Esta pesquisa em nível de Mestrado desenvolveu mecanismos para tratar esse problema, desenvolvendo uma abordagem que permita incluir funções de distâncias ponderadas para o processamento de consultas por similaridade em MAM. Um vetor de pesos corretamente aprendido, permite ponderar funções de distâncias e aprimorar a semântica dos dados, trazendo maior precisão ao processamento de consultas. Deve-se, para isso, utilizar recursos de Realimentação de Relevância (RR), visando capturar a percepção do usuário com relação à semelhança entre imagens. Este trabalho propõe dois métodos que tratam o aprimoramento do espaço de características e inclusão de funções de distância ponderadas no MAM *Slim-Tree*. O método *Fusion Relevance Feedback* (FRF) aplica um pré-processamento combinando extratores de características tradicionais e aprimorando o espaço de características utilizando RR, alcançando acurácias equivalentes e superiores em relação as técnicas de aprendizado profundo. O método *Tuning Metrics Relevance Feedback* (TMRF) infere vetores de pesos no MAM *Slim-Tree* e apresenta uma metodologia de reindexação que mantém a estrutura otimizada com o espaço métrico aprimorando. As análises realizadas demonstram que o método TMRF aprimora o espaço métrico do conjunto de dados e mantém o MAM eficiente, sendo 70% mais rápida em relação a estratégias sequenciais, com ganhos expressivos em termos de acurácia de até 42% através de aprendizado por RR.

Palavras-chave: Método de Acesso Métrico, MAM, Realimentação de Relevância, CBIR, Funções de Distância Ponderada, Consultas por Similaridade.

ABSTRACT

MARCACINI, R. G. **Development of a Metric Access Method with Perceptual Distance Calibration for Medical Image Similarity Queries**. 2024. 100 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2024.

Indexing and retrieving information organized in databases efficiently and accurately to respond to queries requested by users and specialists are pillars of the engineering and data science process. In the medical context, complex image-type data's management and semantic extraction are fundamental for decision-making. This type of management requires extracting the essence (features) of these data and usually carrying out similar queries on such features, instead of using the complex data itself. The Metric Access Methods (MAMs) were developed as fundamental tools to process similarity queries supporting database management systems efficiently. MAMs use fixed distance functions to build the metric tree, which in turn prevents a MAM from being able to index elements using two or more distance functions in the same index. In many cases, the use of traditional distance functions fails to provide answers compatible with users' perception of distance. This Masters level research developed novel mechanisms to deal with this issue, developing an approach that allows including weighted distance functions for processing similarity queries in MAM. A properly learned vector of weights allows weighting distance functions and improving data semantics, enhancing the query processing accuracy. For this intent, resources of Relevance Feedback (RF) were used, to capture the user's perception regarding the similarity between images. This work proposes two methods to enhance the feature space and include weighted distance functions in the Slim-Tree MAM. The method Fusion Relevance Feedback (FRF) applies a pre-processing step combining traditional feature extractors and improving the feature space using RF, achieving equivalent and superior accuracies compared to deep learning techniques. The Tuning Metrics Relevance Feedback (TMRF) method infers weight vectors in the Slim-Tree and presents a reindexing methodology that keeps the structure optimized with the metric space improving. The analyses demonstrated that the TMRF method improves the metric space of the data set and keeps the MAM efficient, being 70% faster in relation to sequential strategies, with expressive gains in accuracy of up to 42% through learning by RF.

Keywords: Metric Access Method, MAM, Relevance Feedback, CBIR, Weighted Distance Functions, Semantic similarity queries.

LISTA DE ILUSTRAÇÕES

Figura 1 – Arquitetura de um sistema CBIR convencional composta por três módulos: Interface, Processamento de Consultas e Base de Imagens, ilustrando o fluxo de recuperação de dados complexos	31
Figura 2 – Processo de extração de características e geração de um conjunto de vetores de características a partir de um conjunto de imagens	32
Figura 3 – Exemplo de histograma de cor RGB mostrando a distribuição das intensidades de vermelho, verde e azul na imagem	33
Figura 4 – Exemplo de histograma em níveis de cinza representando a distribuição de intensidades de tons de cinza na imagem	33
Figura 5 – Exemplo de segmentação de imagem baseada na frequência de cor utilizando o algoritmo <i>K-means</i> com 2 clusters, destacando a separação de regiões distintas com cores semelhantes	34
Figura 6 – Exemplo de imagens com características de texturas diferentes	35
Figura 7 – Exemplo de consulta por abrangência. O elemento central em preto representa o centro de consulta, enquanto o círculo ao seu redor define o raio especificado na consulta. Todos os elementos dentro deste círculo são recuperados como resultados da consulta	40
Figura 8 – Exemplo de consulta aos k-Vizinhos mais próximos. O elemento central em preto representa o centro de consulta e os quatro elementos mais próximos ao centro de consulta são destacados como resultados da consulta, onde $k = 4$	40
Figura 9 – Ilustração dos elementos e abrangência da região circunscrita considerando o Centro da Consulta C e raio R , para 3 funções de distância: Euclidiana, Manhattan e Chebychev	42
Figura 10 – Transformação do espaço de busca da função de distância euclidiana ponderada, aprimorando o resultado de uma consulta por similaridade	43
Figura 11 – Exemplo de uma <i>Slim-tree</i> organizando 17 elementos, utilizando espaço bidimensional e distância Euclidiana. Os nós são organizados em três níveis, e representados como círculos devido ao uso da distância Euclidiana	46
Figura 12 – Representação hierárquica (árvore) de uma <i>Slim-tree</i> organizando 17 elementos, para os dados da Figura 11. Observe que a <i>Slim-tree</i> não possui elementos no nó raiz	46

Figura 13 – Exemplo do funcionamento do <i>Fat Factor</i> no método de acesso métrico <i>Slim-Tree</i> , demonstrando a redução de sobreposição em uma região para otimizar a recuperação de dados	48
Figura 14 – Arquitetura de um sistema CBIR com integração de um módulo de realimentação de relevância no Processamento de Consultas, visando a capacidade de aprimorar a eficácia dos resultados	52
Figura 15 – Ciclo de um processo de Realimentação de Relevância, destacando a interação iterativa do usuário que é incorporado a um processo de aprendizado de pesos, convertendo em melhores resultados	53
Figura 16 – Reformulação de consultas em realimentação de relevância, no qual a subtração do conjunto de resultados relevantes do conjunto de resultados não-relevantes leva à criação de uma nova consulta aprimorada	54
Figura 17 – Exemplo da técnica de movimentação do centro de consulta, na qual (a) representa a consulta inicial e (b) representa o novo centro de consultas com o objetivo de se aproximar de elementos relevantes	55
Figura 18 – Processo de técnicas de realimentação de relevância que refinam a medida de similaridade, onde o refinamento são alcançados por meio do aprendizado de pesos, transformando um conjunto de dados em um espaço métrico aprimorado	57
Figura 19 – Exemplo de uma matriz de confusão que destaca as métricas para avaliar o desempenho de um modelo de classificação	61
Figura 20 – Exemplo do procedimento de captura de uma imagem de Raio-X, com a mão do paciente posicionada para a passagem dos feixes de raios-X, permitindo a medição da absorção pelos tecidos e a obtenção de uma imagem para diagnóstico	65
Figura 21 – Exemplo de procedimento para a captura de uma imagem de ultrassom, obtida por meio do uso de ondas sonoras de alta frequência que são emitidas e refletidas pelos tecidos do corpo para visualização médica	66
Figura 22 – Exemplos de imagens pertencentes às três categorias da base de imagens de COVID-19	67
Figura 23 – Exemplos de imagens pertencentes as duas categorias da base de imagens de mamografia	68
Figura 24 – Exemplos de imagens da base de imagens do Corel DB	69
Figura 25 – Diagrama da abordagem <i>Fusion Relevance Feedback</i> , na qual é utilizado combinações de extratores de características em conjunto de realimentação de relevância para aprimorar o espaço de características do conjunto no pré-processamento	72
Figura 26 – Seleção de amostras no pré-processamento do conjunto de dados, onde a posição e a correspondência de classe dos resultados são avaliadas em relação a imagem de consulta	74

Figura 27 – Gráfico de barras que compara as acurácias por descritor de textura, Sem RR e com RR Automática.	76
Figura 28 – Diagrama da abordagem <i>Tuning Metrics Relevance Feedback</i> , na qual destaca a eficácia da abordagem proposta por utilizar métodos de acesso métrico e realimentação de relevância em relação as abordagens convencionais que utilizam matrizes de distâncias	78
Figura 29 – Processo de geração do vetor de pesos sobre a matriz de características de imagens relevantes	79
Figura 30 – Exemplo do cálculo do vetor de pesos médio que reflete a relevância de uma janela de vetores de pesos do histórico de consultas que armazenaram vetores de pesos	80
Figura 31 – Exemplo do cálculo do vetor de pesos médio que reflete os pesos médios das dimensões de um conjunto de vetores de pesos	81
Figura 32 – Processo para realização da reindexação do MAM, onde o conjunto de dados ponderados pelo vetor de pesos médio é passado por transformações do espaço de características e clusterização, até ser avaliado pela métrica de <i>Sillhoutte</i> para validar a reindexação	83
Figura 33 – Tempo de consulta em milissegundos recuperando em incrementos de 5% da base de dados, até atingir 100% da base de imagens de COVID-19, comparando o desempenho das consultas normais, consultas ponderadas e de recuperação sequencial ao longo do processo de recuperação	85
Figura 34 – Quantidade de cálculos de distâncias e acesso ao disco entre função de distância com e sem ponderação.	86
Figura 35 – Exemplo de análise para validação da reindexação do MAM, onde a linha pontilhada em (a) é o valor médio das <i>Silhouettes</i> calculadas para cada <i>cluster</i> (b), indicando a medida geral da qualidade da clusterização.	87
Figura 36 – Visualização do espaço métrico original (a) e do espaço métrico modificado (b) pela técnica RR Desvio Padrão após 5 ciclos de realimentação, indicando a melhoria semântica	87

LISTA DE TABELAS

Tabela 1 – Resumo das principais características dos MAMs Dinâmicos.	49
Tabela 2 – <i>Revisão da Literatura de abordagens de realimentação de relevância que aprendem métricas de similaridade em matriz de distâncias</i>	60
Tabela 3 – Parâmetros e dimensões dos extratores de características.	67
Tabela 4 – <i>Extratores de texturas para análise do método FRF</i>	74
Tabela 5 – <i>Comparação da ACC e F1 por descritor de textura, Sem RR e com RR Automática.</i>	75
Tabela 6 – <i>Mean Average Precision</i> sobre os top 10 e 20 imagens, de cada extrator sem RR e com RR Desvio Padrão em diferentes bases de imagens.	85

LISTA DE ABREVIATURAS E SIGLAS

CBIR	<i>Sistemas de Recuperação de Imagens por Conteúdo</i>
EB	<i>Exabytes</i>
FN	False Negative
FOS	<i>First Order Statistical</i>
FP	False Positive
FRF	<i>Fusion Relevance Feedback</i>
GLCM	<i>Gray Level Co-occurrence Matrix</i>
HSV	<i>Hue, Saturation, Value</i>
LMNN	<i>Large Margin Nearest Neighbor</i>
LTE	<i>Texture Energy Measure</i>
MAMs	Métodos de Acesso Métrico
MAP	<i>Mean Average Precision</i>
QPM	Query Point Movement
RGB	<i>Red, Green, Blue</i>
RR	Realimentação de Relevância
SFM	<i>Statistical Feature Matrix</i>
SGBDs	Sistemas de Gerenciamento de Banco de Dados
SVM-AL	<i>Support Vector Machine Active</i>
t-SNE	<i>t-Distributed Stochastic Neighbor Embedding</i>
TMRF	<i>Tuning Metrics Relevance Feedback</i>
TN	True Negative
TP	True Positive

SUMÁRIO

1	INTRODUÇÃO	25
1.1	Motivação	25
1.2	Objetivos	27
1.3	Organização da Dissertação	28
2	RECUPERAÇÃO DE IMAGEM POR CONTEÚDO	29
2.1	Visão Geral	29
2.2	Arquitetura de um Sistema CBIR	30
2.3	Extração de Características	31
2.4	Considerações Finais	37
3	CONSULTAS POR SIMILARIDADE	39
3.1	Visão Geral	39
3.2	Funções de Distância e Espaço Métrico	41
3.3	Funções de Distância Ponderada	42
3.4	Métodos de Acesso Métrico	43
3.4.1	<i>Slim-Tree</i>	45
3.5	Percepção do Usuário em Consultas por Similaridade	49
3.6	Considerações Finais	50
4	REALIMENTAÇÃO DE RELEVÂNCIA E TRABALHOS RELACIONADOS	51
4.1	Visão Geral	51
4.2	Reformulação de Consultas	54
4.3	Refinamento da Medida de Similaridade	56
4.4	Trabalhos Relacionados e Descrição do Problema	57
4.5	Métricas de Avaliação	60
4.6	Considerações Finais	64
5	AQUISIÇÃO E DESCRIÇÃO DAS BASES DE IMAGENS	65
5.1	Visão Geral	65
5.2	Análise de Imagens de COVID-19 e Mamografias e Gerais	67
5.3	Considerações Finais	68

6	MÉTODO FRF (FUSION RELEVANCE FEEDBACK)	71
6.1	Visão Geral	71
6.2	Combinação de Extratores de Características <i>Early Fusion</i>	72
6.3	Aprendizado de Pesos Automático	73
6.4	Experimento	74
6.5	Resultados e Discussões	75
6.6	Considerações Finais	76
7	MÉTODO TMRF (TUNING METRICS RELEVANCE FEEDBACK)	77
7.1	Visão Geral	77
7.2	Realimentação de Relevância por Desvio Padrão	78
7.3	Histórico de Consultas e Seleção do Vetor de Pesos	80
7.4	Inclusão de Funções de Distância Ponderadas na <i>Slim-Tree</i>	81
7.5	Reindexação da <i>Slim-Tree</i>	82
7.6	Experimentos	84
7.7	Resultados e Discussões	84
7.8	Considerações Finais	88
8	CONCLUSÃO	89
8.1	Considerações Finais	89
8.2	Contribuições	89
8.3	Trabalhos Futuros	90
	REFERÊNCIAS	93

INTRODUÇÃO

1.1 Motivação

O processo de analisar e tomar decisões baseadas em exames de pacientes é primordial para o diagnóstico e tratamento corretos de pacientes. O avanço tecnológico de dispositivos que realizam exames de imagens, tais como tomografia computadorizada, ressonância magnética e ultrassom, aliado à diminuição dos custos desses exames, tem também contribuído para o crescimento rápido da requisição de tais exames. Anualmente, a produção de novos dados de saúde tem sido da ordem de 150 *Exabytes* (EB (*exabyte*)= 10^{18} bytes), com taxa de crescimento de 48% anualmente ¹, a qual tinha previsão de chegar a 175.000 EB em 2025 antes de ocorrer a pandemia de COVID-19 (YUNITA; SANTOSO; HASIBUAN, 2022; DASH *et al.*, 2019). Logo, percebe-se que é necessário contar com sistemas que permitam armazenar e recuperar as informações dos pacientes e os exames de imagens de modo eficiente e eficaz.

O processo de analisar uma imagem, que é realizado pelo especialista médico, conta com sua experiência e treinamento desenvolvidos ao longo dos anos. Uma parte importante desse treinamento baseia-se na análise e comparação de imagens semelhantes, ressaltando pontos de interesse e possíveis alterações e anomalias. Um aspecto desse processo é que a memória visual do especialista é treinada para detectar padrões que o levem a indicar possíveis achados na imagem. Porém, os seres humanos possuem capacidade de memória limitada, e muitas vezes ao analisar uma imagem recordam-se de casos semelhantes que já viram no passado, entretanto não se recordam de quais análises foram feitas, ou de possíveis resultados. Segundo (VISHRAJ; GUPTA; SINGH, 2022; YIN *et al.*, 2021), é nesse ponto que os sistemas computacionais de recuperação de imagens baseada em seu conteúdo (do inglês *Content-based Image Retrieval* - CBIR) atuam, ao buscarem na base de imagens aquelas que são similares à uma imagem que está sendo analisada e demais dados associados a elas e ao paciente, levando em consideração o

¹ Nature Medicine | Vol 25 | Jan 2019 | pp. 24–29 | www.nature.com/naturemedicine)

tempo de busca e a precisão das consultas por similaridade.

Muitas das pesquisas da literatura ressaltam e demonstram a eficácia das técnicas de Realimentação de Relevância (RR) para aumentar o desempenho geral das consultas aumentando a semântica dos resultados (MOSBAH, 2023; AHMED, 2020; TIAN, 2018; MOHANAN; RAJU, 2017). A RR, nesse contexto, implica em um processo interativo e iterativo, onde o sistema interage com o usuário ou especialista para refinar e ajustar os resultados das consultas através de conjuntos de amostras selecionadas pelos usuários que servem como aprendizado. A ênfase nesse processo é garantir que as imagens recuperadas sejam semanticamente relevantes para a intenção do usuário. Isso resulta em uma melhor correspondência entre os resultados retornados e as expectativas do usuário, criando uma experiência mais satisfatória e produtiva. Entretanto, apesar dos benefícios da RR, a abordagem predominante na literatura para a recuperação de dados ainda envolve cálculos entre pares de elementos, frequentemente expressos por meio de matrizes de distância. Essas matrizes armazenam as medidas de distância ou similaridade entre cada par de elementos, e os algoritmos de recuperação usam essas matrizes para classificar e ordenar os resultados. Essa abordagem, embora seja útil para pequenos conjuntos de dados, enfrenta limitações quando se trata da recuperação eficiente de grandes volumes de dados complexos. Isso envolve a exploração de métodos que possam lidar eficientemente com a busca em grandes volumes de dados, levando em consideração não apenas a precisão do sistema, mas também o custo computacional.

Os Métodos de Acesso Métrico (MAM) foram desenvolvidos para otimizar o processamento de consultas por similaridade a dados complexos, principalmente para grandes conjuntos de dados. Por exemplo, em uma aplicação de imagens médicas, a agilidade no processamento é essencial, já que é necessário analisar uma grande quantidade de imagens em tempo hábil. Com a ajuda de um MAM, é possível comparar rapidamente as características de uma base de dados de imagens médicas, permitindo que o sistema ou especialista identifique rapidamente possíveis diagnósticos e tratamentos previamente armazenados, proporcionando uma referência para diagnósticos atuais. Ou seja, um MAM acelera o processo de recuperação de dados que são utilizados em sistemas de análise, melhorando o suporte à decisão clínica para o tratamento do paciente. Os MAMs envolvem a definição de uma métrica de similaridade para construção da indexação, a qual é utilizada para determinar quais elementos estão mais próximos de um determinado elemento de consulta, o que por sua vez impede que, ordinariamente, um MAM consiga indexar os elementos utilizando duas ou mais funções de distância em uma mesma indexação (RAZENTE; BARIONI; SILVA, 2022; LI *et al.*, 2018; SILVA, 2009).

O uso de um MAM que permita a recuperação otimizada de dados complexos que contemple a semântica de similaridade real dos dados em um espaço métrico e que possa adequar a função de distância, são fundamentais para a aceitação de sistemas CBIR pelos especialistas. Portanto, soluções baseadas na inclusão de funções de distância ponderadas no MAM juntamente com técnicas de RR, que aprendem pesos para ajustar medidas de similaridade, são essenciais

para obter consultas de similaridade e tempos de processamentos mais eficazes pelos MAMs em relação às estratégias de recuperação por matriz de distância. Esse projeto de mestrado visa atuar nessa direção, com a análise e desenvolvimento de novos algoritmos e técnicas de RR que inserem o especialista humano no refinamento da consulta, e que deverão ser integrados a métodos de acesso métricos, que são a base para o processamento de consultas por similaridade, de modo a apoiar a recuperação de imagens por conteúdo.

Partimos do pressuposto que esse aprendizado, por meio de alteração das funções de distância ajustadas pela indicação dos usuários aos métodos de acesso, permitirão que as consultas sejam respondidas de modo mais próximo aos anseios dos especialistas, além de manter a recuperação de consultas em um tempo hábil. Ou seja, com tal procedimento será possível diminuir a lacuna (*gap*) semântica (AHMED, 2020; WEI; AMIT, 2019; DESERNO; ANTANI; LONG, 2009) inerente a processos de recuperação de dados complexos e imagens por conteúdo, que é um dos maiores problemas desse tipo de abordagem computacional.

1.2 Objetivos

O objetivo desta pesquisa em nível de Mestrado é o desenvolvimento de tecnologia para a construção de métodos de acesso métrico que incluam a percepção do usuário no processamento de consultas por similaridade. Ou seja, tal tecnologia deverá ser materializada em um novo método de acesso métrico perceptual que utilize um passo de pré-processamento de calibragem de funções de distância ponderadas e características de imagens. Para isto, este trabalho propôs uma nova abordagem denominada *Tuning Metrics Relevance Feedback* (TMRF), que introduz a melhoria semântica do espaço de características e reindexa o espaço métrico mantendo a otimização da estrutura de indexação. Ainda neste trabalho, foi desenvolvida outra abordagem denominada *Fusion Relevance Feedback* (FRF) que combina extratores de características de textura em conjunto com RR para aprimorar o espaço de características. Mais especificamente, os objetivos nesta pesquisa são:

- Definir um procedimento de otimização de espaços de características por meio de realimentação de relevância que utilizem matrizes de características do *feedback* do usuário.
- Elaborar uma estratégia de inserção de vetores de pesos para o ajuste de funções de distância ponderadas no método de acesso métrico.
- Analises do impacto de custo computacional da calibração de distância dinâmica no método de acesso métrico.
- Elaboração de uma metodologia para definir quando realizar a reindexação do método de acesso métrico que mantenha a estrutura otimizada e com o espaço de características aprimorado.

A validação dos resultados foi feita utilizando a acurácia, o *Mean Average Precision* e o coeficiente de *Silhouette* sobre três conjuntos de base de imagens. Sendo duas dessas bases relacionado a imagens médicas e uma base de imagens gerais para validar a generalização do método proposto.

1.3 Organização da Dissertação

A dissertação é organizado na seguinte forma: no [Capítulo 1](#), são expostos o tema e a motivação, juntamente dos objetivos principais. Os três capítulos seguintes apresentam os conceitos específicos abordados no trabalho e as principais referências relacionadas, sendo os tópicos de bases de dados, recuperação de imagens por conteúdo e descritores de texturas utilizados nos resultados preliminares abordado no [Capítulo 2](#). O conteúdo relacionado a consultas por similaridade, os principais métodos de acesso métricos e suas vantagens e desvantagens são abordados no [Capítulo 3](#). As técnicas de captura de percepção utilizando realimentação de relevância, incluindo as abordagens que otimizam o espaço métrico, são apresentadas no [Capítulo 4](#). As descrições das bases de imagens utilizadas para validar os métodos proposto nessa pesquisa são apresentadas no [Capítulo 5](#). O primeiro método proposto *Fusion Relevance Feedback* (FRF) junto com experimentos para avaliar a abordagem é apresentado no [Capítulo 6](#). O segundo método proposto *Tuning Metrics Relevance Feedback* (TMRF), discussão do método e experimentos para avaliar a abordagem é apresentado no [Capítulo 7](#). Por fim, a conclusão, principais contribuições e discussão de trabalhos futuros desta pesquisa é abordado no [Capítulo 8](#).

RECUPERAÇÃO DE IMAGEM POR CONTEÚDO

2.1 Visão Geral

No contexto de recuperação de informação, a literatura oferece diversos tópicos para a construção de sistemas que aprimoram as tomadas de decisões para atender às expectativas do usuário. Uma das etapas para uma boa metodologia é a seção de pré-processamento com a definição dos métodos apropriados (WILLEMINK *et al.*, 2020; REZENDE, 2003). Existe uma grande dificuldade para que os resultados recuperados correspondam à expectativa real do usuário, levando geralmente ao problema de lacuna semântica (*semantic gap*) (YIN *et al.*, 2021; LI *et al.*, 2016). Uma forma eficiente para atender às expectativas do usuário é por meio de realimentação de relevância (AHMED, 2020; ZHANG *et al.*, 2007). Técnicas de realimentação de relevância necessitam que o usuário forneça informações essenciais, como rotular imagens como relevantes ou irrelevantes para ajustar a consulta a sua expectativa.

As aplicações com imagens médicas são vastas e possuem um ciclo de uso longo. Esse ciclo inicia na coleta dos dados e imagens até o diagnóstico e tratamento/acompanhamento do paciente, devendo ser guardadas por 20 anos segundo legislação brasileira indicada em Resolução do Conselho Federal de Medicina (CFM, 2007).

Um dos principais problemas para o desenvolvimento e implementação de sistemas computacionais na prática clínica é a indisponibilidade de grandes quantidades de dados de treinamento para treinar algoritmos de inteligência artificial. Além disso, há a necessidade que os dados sejam devidamente apurados, analisados e rotulados para que esses algoritmos tenham impacto positivo. Para a utilização das imagens médicas, alguns procedimentos são definidos buscando a estruturação de forma homogênea desses dados. Os dados devem ser acessados, consultados e anonimizados adequadamente e armazenados com segurança (WILLEMINK *et al.*,

2020). A qualidade e a quantidade das imagens variam conforme o domínio do projeto. Desta forma, a estruturação dos dados ocorre de forma a padronizar o formato e a dimensão dessas imagens e rotulá-los com seu real significado, antecipando melhoria para os algoritmos de busca.

2.2 Arquitetura de um Sistema CBIR

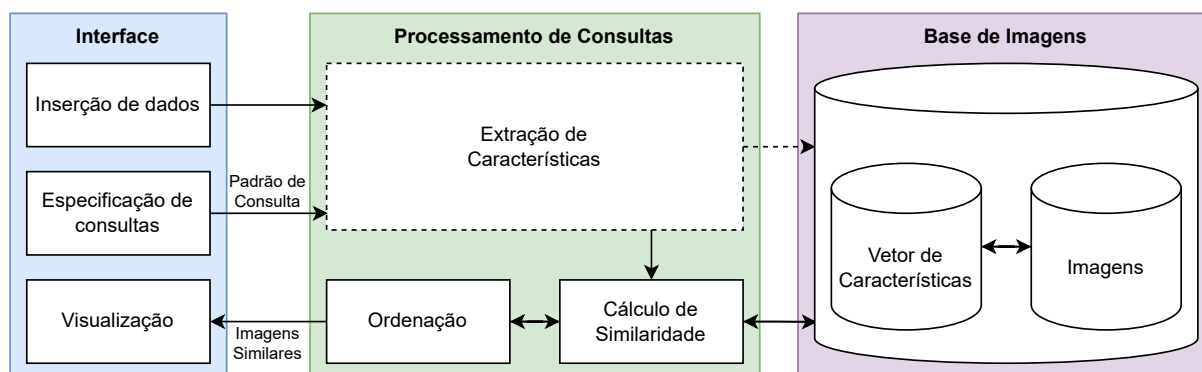
Sistemas de Recuperação de Imagens por Conteúdo (CBIR) são construídos com arquiteturas utilizadas para recuperação de dados complexos, como por exemplo imagens médicas. Em outras palavras, esses sistemas são projetados para localizar e recuperar imagens com base nas características visuais das próprias imagens, em vez de depender de metadados ou descrições textuais. O objetivo é encontrar imagens que possuam características visuais semelhantes à imagem de consulta, classificando-as por ordem de relevância. A arquitetura de um sistema CBIR é organizada de modo a responder de forma precisa e otimizada consultas que são realizadas pelos os usuários (DHAYGUDE; KINARIWALA, 2022). Para isso a arquitetura é dividida em 3 camadas: Interface, Processamento de Consultas e Base de Imagens.

A camada de **Interface** é o módulo de interação do usuário com o sistema. Ela permite ao usuário fazer especificações de consultas, como determinar qual tipo de descritor será usado como base de recuperação, incluir novos dados para armazenamento no banco de dados e visualizar os resultados da busca, que são ranqueados com base na similaridade com a consulta realizada.

A camada de **Processamento de Consultas** é o módulo em que é realizada a extração de características e o cálculo de similaridade da consulta realizada pelo usuário com os resultados mais similares. Essa é uma camada de extrema importância, pois requer um estudo de caso para otimização e funcionamento em diferentes domínios. Nesse estudo de caso, são encontradas informações para definir o tipo de extrator de características e as funções de distância a fim de obter um sistema com alta precisão.

A camada da **Base de Imagens** é o módulo de armazenamento de dados complexos. Geralmente, este módulo é passado por um pré-processamento local, extraíndo as características de um conjunto de dados utilizado de estudo de caso. Deste processamento, são armazenados os vetores de características que representam o dado complexo. Essas características são recuperadas através de um SGBD que contenha um método de acesso métrico para realizar operações de busca por similaridade. A [Figura 1](#) apresenta a arquitetura de um sistema CBIR convencional, demonstrando como é a organização e a comunicação de cada camada, ilustrando o fluxo de inserção e recuperação de dados.

Figura 1 – Arquitetura de um sistema CBIR convencional composta por três módulos: Interface, Processamento de Consultas e Base de Imagens, ilustrando o fluxo de recuperação de dados complexos



Fonte: Elaborada pelo autor.

2.3 Extração de Características

Uma imagem é representada digitalmente por uma matriz ou mapa de pixels, cada pixel possui um conjunto de valores numéricos que representam a parte visual da imagem em baixo nível. A extração de característica é um método de captura do conteúdo visual de uma imagem para indexação e recuperação em bases de dados. Este processo gera um vetor de características que funciona como uma assinatura da imagem, esta informação é relevante para resolver a tarefa computacional relacionada a uma determinada aplicação (DEVAREDDI; SRIKRISHNA, 2022; HARALICK; SHANMUGAM; DINSTEIN, 1973). As técnicas de extração de características são divididas em globais e locais.

As técnicas de extração de características globais extraem informação sobre a matriz inteira de uma imagem não considerando o conteúdo interno ou partes da imagem, os histogramas de cor ou níveis de cinzas são exemplos muito utilizados para este tipo de extração. Os histogramas possuem a vantagem de demandarem baixo custo de computacional, porém dependendo do domínio da aplicação, estes histogramas podem não representar semanticamente as imagens acarretando em muitas imagens irrelevantes durante as consultas, pois traduzem representações muitas vezes ambíguas. Ou seja, duas imagens diferentes podem ter o mesmo histograma (LATIF *et al.*, 2019).

As técnicas de extração de características locais usualmente extraem informação sobre um processo de segmentação das imagens. O resultado desta segmentação são regiões de interesse de uma imagem que são utilizadas para aumentar o nível semântico durante as consultas (LATIF *et al.*, 2019).

Para aplicações considerando o domínio de imagens, no geral os extratores de características podem ser englobados em três tipos de grandes grupos de classificação: características de

cor, características de forma e características de textura. Os algoritmos de extração de características analisam exclusivamente um desses grupos, para se ter uma análise simultânea dessas características geralmente é aplicada a combinação de diferentes extratores gerando um novo vetor de características, em vários casos esta combinação aumenta a qualidade de precisão nas consultas, porém pode aumentar significativamente o número de dimensões do vetor combinado.

A [Figura 2](#) ilustra o processo de extração de características de várias imagens, onde ao final do processo é gerado um vetor de característica único para cada imagem extraída. Esses vetores de características possuem a mesma dimensão e funcionam como uma assinatura exclusiva para cada imagem, a qual pode ser usada para armazenamento e cálculo de similaridade entre as imagens.

Figura 2 – Processo de extração de características e geração de um conjunto de vetores de características a partir de um conjunto de imagens



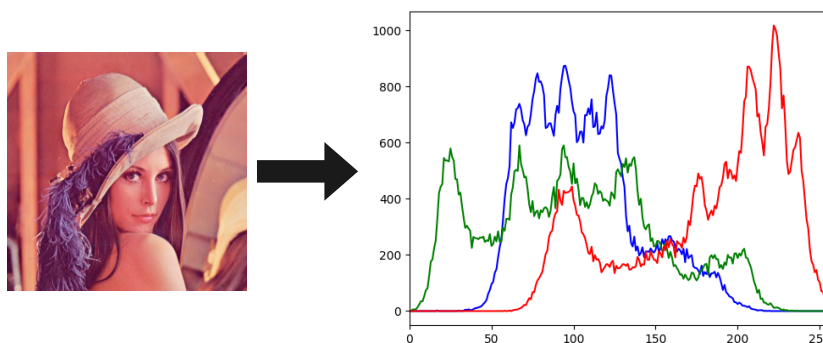
Fonte: Elaborada pelo autor.

Cor

O histograma de cores representa a distribuição das cores ou a frequência de cada cor existente em cada pixel de uma imagem. É amplamente utilizado por ser computacionalmente barato e possui a vantagem de ser invariante a rotação e escala. Porém o histograma de cores pode apresentar alguns problemas por não considerar informação espacial dos pixels em uma imagem. Imagens coloridas são armazenadas em 3 atributos primários formando um espaço de cor. Existem vários espaços de cores como o *Red, Green, Blue* (RGB) e o *Hue, Saturation, Value* (HSV). O padrão de representação mais utilizado para dispositivos de exibição é o RGB que inclui a composição das cores vermelha, verde e azul, já o HSV representa a matiz, a saturação e o brilho (LI *et al.*, 2022).

A extração de características de histograma de cor é um processo que envolve a conversão da imagem para o espaço de cores apropriado, a criação de histogramas para cada canal de cor, conta a ocorrência de cada valor de cor na imagem e, em seguida, organizar essas contagens em um vetor de características. O vetor de características resultante representa de forma compacta a distribuição de cores na imagem e pode ser utilizado em várias aplicações de processamento de imagens. A [Figura 3](#) representa dois histogramas de intensidade de cor RGB das imagens, para a camada RGB cada pixel possui um valor entre 0 e 255.

Figura 3 – Exemplo de histograma de cor RGB mostrando a distribuição das intensidades de vermelho, verde e azul na imagem



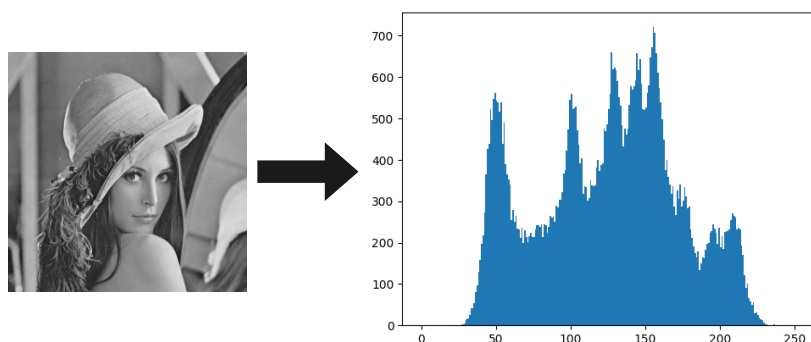
Fonte: Elaborada pelo autor.

Histograma de Cinza

O histograma de cinza é uma representação visual da distribuição das intensidades de cinza em uma imagem. Ao contrário do histograma de cores, que considera os componentes de cor (como vermelho, verde e azul), o histograma de cinza se concentra exclusivamente nas intensidades de cinza, que variam de preto (valor de intensidade mínimo) a branco (valor de intensidade máximo) (LI *et al.*, 2022).

O processo de criação de um histograma de cinza envolve a conversão da imagem colorida para escala de cinza, onde cada pixel é representado apenas por um valor de intensidade. Esses valores de intensidade são então contabilizados e organizados em um histograma, que mostra a frequência de ocorrência de cada valor de intensidade na imagem. A Figura 4 representa um histograma de cinza da imagem, onde cada pixel possui um valor entre 0 e 255 em apenas um canal de cor.

Figura 4 – Exemplo de histograma em níveis de cinza representando a distribuição de intensidades de tons de cinza na imagem



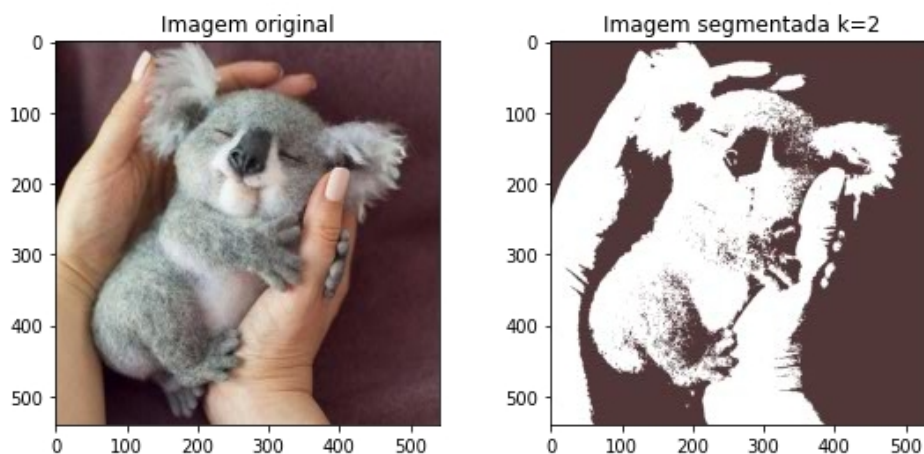
Fonte: Elaborada pelo autor.

Forma

As formas de objetos presentes nas imagens são usualmente obtidas após um processo de segmentação da imagem. Porém não há um processo de segmentação automatizado que atenda de modo otimizado a qualquer domínio de imagens (TAJBAKHSI *et al.*, 2020; ADORAM; LEW, 1999). Uma segmentação deve favorecer a execução correta de algoritmos de extração de características a fim de aumentar a precisão de uma consulta. Para isto, pode ser preciso realizar uma análise mais aprofundada junto ao especialista para determinar quais regiões são mais importantes e o tipo de pré-processamento que melhor caracterizam essas regiões (GONZALEZ; WOODS, 2018). Existem vários algoritmos e métodos disponíveis para a extração de características de forma e dependem do tipo de informação que deseja obter.

Um algoritmo que trabalha com agrupamento de elementos e pode ser utilizado como base de segmentação é o *K-Means*. A extração de características de forma utilizando o processo por *K-means* é um método para identificar e quantificar as propriedades geométricas e estruturais dos contornos ou silhuetas de objetos em uma imagem, enquanto também levando em consideração a informação de cor. O resultado deste método de extração é o agrupamento de pixels semelhantes com base na cor e, em seguida, extrair características de forma desses grupos. A Figura 5 apresenta uma segmentação de uma imagem utilizando a técnica *K-Means* separando as cores da imagem em k clusters, cada cluster encontra na imagem a cor mais frequente filtrando o espaço de cor para realizar uma segmentação. A Seção 4.5 detalhará a técnica *K-Means* e sua aplicação em análise de dados.

Figura 5 – Exemplo de segmentação de imagem baseada na frequência de cor utilizando o algoritmo *K-means* com 2 clusters, destacando a separação de regiões distintas com cores semelhantes



Fonte: Elaborada pelo autor.

Textura

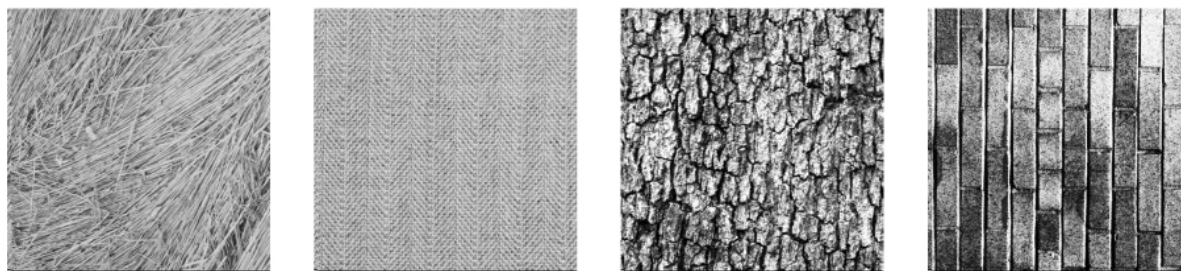
A extração de características de textura refere-se ao processo de identificar e quantificar padrões ou propriedades texturais presentes em uma imagem. A textura em uma imagem pode

ser definida como a repetição de padrões visuais em diferentes escalas ou direções. Apesar de não existir uma definição clara de “textura”, na literatura, a textura é abordada e diferenciada pelo olhar humano através de características como fina ou grosseira, lisa ou áspera, homogênea ou não homogênea, entre outros.

Embora a extração de textura em uma imagem possa ser extraída utilizando o histograma de cores das imagens, as técnicas de extração de textura de uma região da imagem é mais comumente determinada de acordo com a distribuição dos níveis de cinzas pelos pixels na região de uma imagem. O histograma de cinza é menos sensível a variações de cor do que o histograma de cor. Ao trabalhar com o histograma de cinza, as características de textura extraídas são baseadas nas variações de intensidade de cinza, que são mais robustas e consistentes em diferentes condições de iluminação e variações de cor.

As três principais abordagens para extração de características de textura são: estatísticas, estruturais e espectrais (ARMI; FEKRI-ERSHAD, 2019; CEZAR, 2005). A análise estatística é caracterizada pela variação dos atributos de intensidade, contraste, granularidade e direção, onde é considerada a distribuição dos tons de cinzas e o inter-relacionamento entre eles. A análise estrutural descreve texturas através de primitivas de texturas das relações espaciais entre várias primitivas no decorrer de regras de localização. A Figura 6 apresenta exemplos de tipos variados de texturas.

Figura 6 – Exemplo de imagens com características de texturas diferentes



Fonte: Brodatz (2020).

First Order Statistical

O *First Order Statistical* (FOS) ou Estatística de Primeira Ordem é um conjunto de medidas estatísticas básicas aplicadas a uma imagem para descrever suas propriedades de intensidade dos *pixels*. Essas medidas incluem extração de estatísticas simples, como média, desvio padrão, mediana, valor mínimo e máximo, entre outras. Nesta técnica, não é considerada a informação de textura relacionada à distribuição de intensidade da vizinhança de cada pixel da imagem, e sim uma taxa de ocorrência de cada intensidade de pixel em uma imagem (ANTONELLI; SIMONE; VIOLA, 2023; HARALICK; SHANMUGAM; DINSTEIN, 1973).

Gray Level Co-occurrence Matrix

Proposto por (HARALICK; SHANMUGAM; DINSTEIN, 1973), o *Gray Level Co-occurrence Matrix* (GLCM) ou Matriz de Coocorrência de Níveis de Cinza é um método estatístico para descrição de texturas. Este método gera uma matriz de ordem ($N_g \times N_g$) que armazena a probabilidade de que dois valores de intensidade de cinza estejam envolvidos por uma determinada relação espacial. Essa matriz é usada para capturar informações sobre a distribuição de padrões de intensidade dos pixels em diferentes direções e distâncias.

Diferentes medidas estatísticas são extraídas para caracterizar a textura de uma imagem, em um total de quatorze medidas apresentadas por Haralick, sendo elas: contraste, homogeneidade, energia, correlação, soma das médias, soma da variação, entropia, soma da entropia, diferença da média, informação mútua, medidas de correlação angular, medidas de inércia e medidas de correlação máxima.

As medidas fornecem informações sobre a distribuição espacial de padrões de intensidade dos *pixels* na imagem. Este método de análise de textura se preocupa com a detecção de aspectos de uma imagem que são invariantes a rotação. Para isso, as GLCM são calculadas extraindo a intensidade de *pixel* nos ângulos 0° , 45° , 90° , 135° graus com valor de distância usualmente de 1, 2 ou 3 do pixel central.

Statistical Feature Matrix

Proposto por (WU; CHEN, 1992), a *Statistical Feature Matrix* (SFM) é uma matriz numérica que captura informações estatísticas relevantes do conjunto de dados em questão. Ela é construída através do cálculo de várias medidas estatísticas de segunda ordem de pares de pixels com várias distâncias.

Após calcular as medidas estatísticas, elas são organizadas na matriz, onde cada linha representa uma observação ou objeto do conjunto de dados, e cada coluna corresponde a uma medida estatística específica. O tamanho de matriz depende da distância máxima usada em vez do número da intensidade de cinza e pode ser expandida com facilidade, os atributos extraídos desta técnica são quatro, sendo elas a aspereza, rugosidade, periodicidade e contraste.

Texture Energy Measure

Proposto por (LAWS, 1980), a *Texture Energy Measure* (LTE) é uma abordagem que mede a variação dentro de uma janela de tamanho fixo, é uma análise muito rápida que requer algumas convoluções com máscaras de coeficiente inteiro pequeno. Dada uma imagem f ocorre um pré-processamento modificando os pixels da mesma para média zero, este procedimento subtrai a imagem original por uma nova imagem média. Em seguida, a imagem média é calculada deslizando uma máscara de janela menor (por exemplo 5×5) sobre a imagem original f , em cada pixel é calculado o valor médio dentro da vizinhança da máscara. Os atributos extraídos deste métodos são seis, sendo a energia da textura do kernel LL, EE, SS e a média da energia da textura LE, ES e LS.

2.4 Considerações Finais

Este capítulo apresentou os conceitos principais relacionados à recuperação de imagens por conteúdo, bem como técnicas de descrever as imagens para que possam ser indexadas e posteriormente recuperadas por conteúdo. O objetivo foi trazer uma visão geral ao leitor e não de ser exaustivo, cobrindo todo o panorama da área.

CONSULTAS POR SIMILARIDADE

3.1 Visão Geral

Os Sistemas de Gerenciamento de Banco de Dados (SGBDs) tradicionais oferecem recursos para a realização de consultas sobre dados armazenados, utilizando relação de igualdade e de ordem total. Porém, geralmente este tipo de consulta é realizado em dados convencionais como textos curtos e números. Entretanto, no contexto de dados complexos (por exemplo: imagens, sons e áudios), a comparação entre dois ou mais elementos utilizando técnicas de consultas como relação de igualdade e ordem total, não se aplica, pois dificilmente duas imagens ou áudios (ou mesmo outro dado complexo) serão exatamente iguais quando obtidos em instantes diferentes. A consulta por similaridade consiste em procurar elementos em um conjunto de objetos utilizando um critério de similaridade. Nesse contexto, as consultas por similaridade ocorrem através de cálculos em um espaço métrico, que se refere a uma representação matemática que descreve como a distância entre dois pontos é medida (DHAYGUDE; KINARIWALA, 2022).

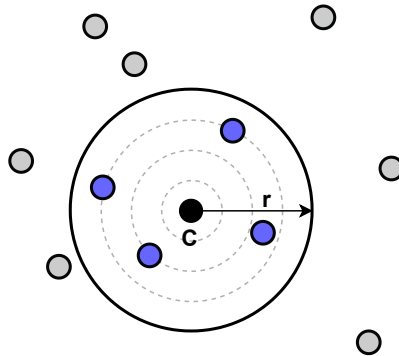
Na literatura, as consultas por similaridade possuem pelo menos dois tipos fundamentais: as consultas por abrangência e as consultas aos k-vizinhos mais próximos (FIGUEROA; CAMARENA-IBARROLA; VALERO, 2022; BÖHM; BERCHTOLD; KEIM, 2001; CHÁVEZ *et al.*, 2001). A definição formal dessas consultas é apresentada a seguir:

Nas consultas por abrangência, o usuário especifica um raio (ou limite de distância) e um ponto de referência no espaço métrico. O objetivo é encontrar todos os elementos dentro deste raio de busca em relação ao ponto de referência.

Consulta por Abrangência: As consultas por abrangência são consultas por similaridade nas quais um usuário especifica um ponto de referência conhecido como centro de consulta e um valor do raio de recuperação. O objetivo é recuperar todos os elementos no espaço métrico que estejam dentro desse raio em relação ao centro de consulta. Na [Figura 7](#), temos um exemplo de consulta por abrangência em um espaço métrico bidimensional onde cada ponto representa

um elemento no espaço métrico. O elemento central, destacado em preto, representa nosso centro de consulta. Os círculo em torno desse elemento central representa o raio r especificado na consulta. Os elementos dentro deste círculo são os elementos que satisfazem a consulta e são recuperados como resultados.

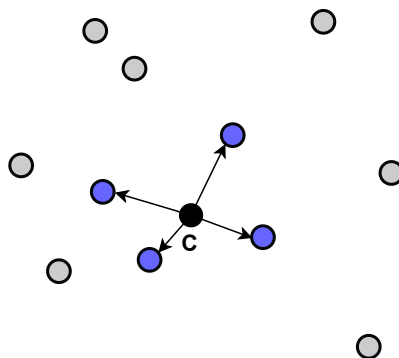
Figura 7 – Exemplo de consulta por abrangência. O elemento central em preto representa o centro de consulta, enquanto o círculo ao seu redor define o raio especificado na consulta. Todos os elementos dentro deste círculo são recuperados como resultados da consulta



Fonte: Elaborada pelo autor.

Consulta por k-Vizinhos mais próximos: A consulta por k-Vizinhos mais próximos é usada para encontrar os k elementos mais próximos a um elemento de consulta específico no espaço métrico. Na Figura 8 temos um espaço métrico bidimensional com pontos representando elementos. O elemento central, destacado em preto, é nosso centro de consulta. Neste caso, especificamos que desejamos encontrar os quatro vizinhos mais próximos ($k = 4$). Os quatro pontos mais próximos ao elementos de consulta são destacados, e esses são os resultados de uma consulta por k-Vizinhos mais próximos.

Figura 8 – Exemplo de consulta aos k-Vizinhos mais próximos. O elemento central em preto representa o centro de consulta e os quatro elementos mais próximos ao centro de consulta são destacados como resultados da consulta, onde $k = 4$



$k = 4$

Fonte: Elaborada pelo autor.

A utilização desses operadores têm como base responder às consultas mais usuais sobre dados complexos, como por exemplo: “Considerando a imagem de Raio-X de pulmão de João da Silva, retorne as 5 imagens mais semelhantes a essa imagem de entrada”, ou “Retorne todas as imagens que difiram em até 10% da imagem de Raio-X de tórax de João da Silva, indicada como entrada para a consulta”. Além disso, como tais consultas são realizadas com frequência, é importante que o gerenciador de bases de dados possa realizá-las com agilidade. Nesse sentido, entram os métodos de acesso que permitem indexar e recuperar dados complexos e apoiam o processamento de consultas por similaridade de modo eficiente.

3.2 Funções de Distância e Espaço Métrico

Os sistemas de recuperação por conteúdo demandam mecanismos para responder a consultas por similaridade e retornar elementos que mais se assemelham a um elemento de consulta. As funções de distância medem a dissimilaridade (o oposto da similaridade) entre dois elementos de um conjunto. Ou seja, ao retornar distância zero, indica que os elementos são iguais, e quanto maior for o valor da distância, mais dissimilares os dois elementos são. Há uma grande variedade de funções de distância fornecidas na literatura (CHEN *et al.*, 2022). As mais utilizadas são as funções da família Minkowski (L_p), entre elas a distância Manhattan (L_1), Euclidiana (L_2) e Chebychev (L_∞), essas funções são representadas por:

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}$$

Um **Espaço Métrico** é definido como um par $M = S, d()$, onde S é o domínio de dados (chaves de indexação) e $d()$ é uma função de distância que satisfaz as seguintes propriedades:

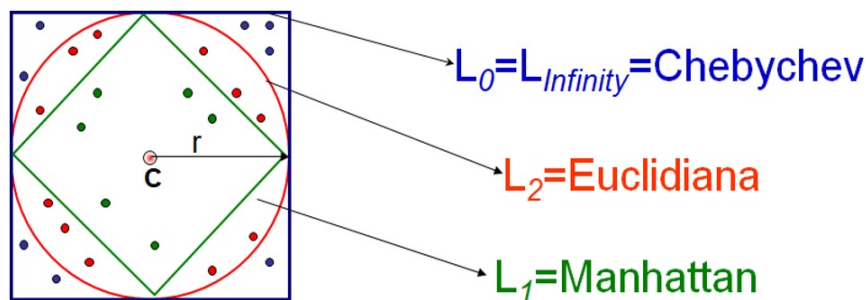
- Simetria: $d(s_1, s_2) = d(s_2, s_1)$;
- Não negatividade: $0 < d(s_1, s_2) < \infty$, se $(s_1 \neq s_2)$;
- Desigualdade triangular: $d(s_1, s_2) \leq d(s_1, s_3) + d(s_3, s_2)$;

onde: s_1, s_2, s_3 são elementos pertencentes à S . As propriedades de um espaço métrico são importantes para a montagem da estrutura de dados e do processamento de operações de busca.

A função de distância tem influência direta na precisão de uma consulta. Desta forma, ela deve ser escolhida e avaliada conforme os descritores do domínio a ser trabalhado. Observe na [Figura 9](#), que ilustra o mecanismo de três funções de distância das mais usuais, em um espaço bi-dimensional. Nota-se que o elemento central C e o raio r indicam 3 regiões com elementos variados, dependendo da função de distância empregada à consulta, que visa retornar todos os elementos que estão a no máximo distância r do elemento C. Ao considerar a função L_1 (Distância

Manhattan) somente os elementos verdes fazem parte da resposta. Ao considerar a função L_2 (Distância Euclidiana) os elementos verdes e vermelhos integram a resposta. Finalmente, ao considerar a função L_∞ (Distância Chebychev) os elementos verdes, vermelhos e azuis, em sua totalidade, integram a resposta. Desse modo, percebe-se que a escolha de uma função de distância implica fortemente no conjunto resposta a esse tipo de consulta (por similaridade), que é processado por meio de cálculos de distância.

Figura 9 – Ilustração dos elementos e abrangência da região circunscrita considerando o Centro da Consulta C e raio R, para 3 funções de distância: Euclidiana, Manhattan e Chebychev



Fonte: Elaborada pelo autor.

3.3 Funções de Distância Ponderada

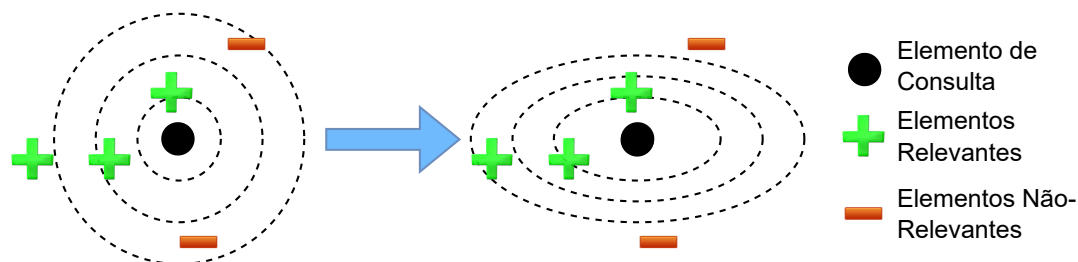
As funções de distância ponderada são uma ferramenta poderosa no aprendizado de métricas. Este tipo de métrica de similaridade dinâmica permite calcular a distância entre dois elementos, levando em consideração a importância de cada atributo, ou seja, são atribuídos diferentes pesos às diferentes dimensões nos dados. A ponderação permite ao usuário adaptar a função de distância às necessidades e preferências específicas de uma determinada tarefa, atribuindo pesos maiores às dimensões mais importantes, fazendo com que melhore os resultados de uma consulta (GHODRATNAMA; MOGHADDAM, 2021; TIAN, 2018). As funções de distância ponderadas de Minkowski são representadas pela equação (3.1), w_i é o peso aplicado sobre a dimensão, na qual a soma de todos os pesos deve sempre manter o valor igual a um, ou seja, $\sum w_i = 1$, esta função é representada por:

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt[p]{\sum_{i=1}^n w_i |x_i - y_i|^p} \quad (3.1)$$

As funções de distância ponderada também são usadas em técnicas de recuperação de dados complexos como k-vizinhos mais próximos (KNN). No KNN, o algoritmo procura os k pontos de dados mais próximos de uma determinada entrada e, em seguida, prevê a saída com base na classe majoritária desses k pontos. Ao usar uma função de distância ponderada, podemos dar mais peso às dimensões mais importantes, alterando o formato de abrangência da região

de busca, desta forma melhorando a precisão de uma consulta por similaridade. A [Figura 10](#) apresenta o efeito de ponderação de uma função de distância euclidiana por pesos, “esticando” o formato de busca melhorando o resultado de uma consulta por similaridade.

Figura 10 – Transformação do espaço de busca da função de distância euclidiana ponderada, aprimorando o resultado de uma consulta por similaridade



Fonte: Elaborada pelo autor.

3.4 Métodos de Acesso Métrico

Os Métodos de Acesso Métrico (MAMs) utilizam funções de distância para organizar os elementos na base de dados. Para isso o MAM assume que os elementos estão em um espaço métrico e os organiza utilizando suas relações de similaridade. Com foco na rapidez do processamento das consultas, os MAMs buscam reduzir o número de cálculos de distância e de acessos a disco para a realização de consultas. A métrica (distância) é usada para quantificar quão similar são dois elementos, e assim permite que se expressem consultas baseadas em similaridade. Um objeto complexo (por exemplo, uma imagem) passa a ser manipulada (indexada/recuperada) com base em seus descritores. Um descritor é formado por um par <vetor de características e função de distância> que deve representar o conteúdo do dado complexo permitindo sua comparação mais apropriada para um determinado contexto ([MUFARROHA; ANAMISA; HAPSANI, 2020](#); [TORRES; FALCAO, 2006](#)). Por exemplo, em imagens pode-se obter descritores baseados em características de cor, forma e textura e uma função de distância para compará-los. Desse modo, cada descritor forma um espaço métrico ao se definir uma função de distância adequada para o vetor de características extraído da imagem ou do dado complexo. Os espaços métricos fazem parte do apoio teórico para a modelagem das consultas por similaridade.

Vários trabalhos apresentam propostas de MAMs na literatura. O primeiro trabalho na literatura foi proposto por Burkhard e Keller ([BURKHARD; KELLER, 1973](#)) denominado BK-Tree, que foi a primeira estrutura a apresentar formas de particionamento recursivas do espaço métrico, que se tornou a base para construção de MAMs. A ideia geral dessas estruturas consiste na escolha de um elemento representante e a utilização de uma função de distância para particionar os demais elementos em vários subconjuntos.

As *BK-Tree* fazem parte dos MAMs estáticos, isto é, estruturas que realizam em apenas

uma passagem inicial sobre o conjunto de dados, a indexação do conjunto total de elementos e não permitem inserções e remoções de elementos após a sua construção. Ou seja, o custo de inserção/atualização posterior é tão alto ou equivalente a construir a estrutura toda novamente. A partir dessa estrutura, outros MAMs conhecidos na literatura foram sendo desenvolvidos, alguns exemplos são: *VP-Tree* (YANILOS, 1993), *MVP-Tree* (BOZKAYA; OZSOYOGLU, 1997), *M-Tree* (CIACCIA; PATELLA; ZEZULA, 1997), *Slim-Tree* (TRAINA *et al.*, 2000), *DBM-Tree* (VIEIRA *et al.*, 2006), *Omni-technique* (TRAINA *et al.*, 2007), *MM-Tree* (POLA; TRAINA; TRAINA, 2007), *Onion-Tree* (CARÉLO *et al.*, 2011), *NOBH-Tree* (POLA; JR; TRAINA, 2014), *UPB-Tree/UPB-Forest* (CHEN *et al.*, 2017) e *xBR⁺-trees* (ROUMELIS *et al.*, 2018).

Os MAMs foram evoluindo ao longo dos anos para lidar com alta dimensionalidade e se tornando estruturas dinâmicas, isto é, permitindo inserções e remoções de elementos após a estrutura já ter sido construída, uma característica fundamental para ser utilizada em um SGBDR (Sistema de Gerenciamento de Banco de Dados Relacional). Vale também ressaltar que os MAMs podem ser classificados como armazenamento em memória primária ou secundária.

Em (CIACCIA; PATELLA; ZEZULA, 1997) foi apresentado o primeiro MAM dinâmico na literatura, denominado *M-Tree*. Foi proposto para organizar e buscar grandes conjuntos em memória secundária. Como a maior parte dos MAMs, a *M-tree* utiliza a propriedade da desigualdade triangular de espaços métricos, usado para podar o espaço de busca aumentando o desempenho das consultas.

Em (POLA; TRAINA; TRAINA, 2007) foi proposto um MAM dinâmico denominado *MM-Tree*, com o objetivo de responder às consultas por similaridade de forma rápida e que possam ser mantidos em memória primária, para aplicações que não demandem persistência. Todos os nós possuem os mesmos componentes, não havendo diferença entre nós folhas e nós índices, cada nó armazena no máximo dois elementos que são representantes de sua região. A *Onion-tree* (CARÉLO *et al.*, 2011) é uma extensão da *MM-tree*, onde múltiplas regiões no formato de “bolas” são usadas para aumentar o leque da árvore, e sua estrutura tem como base algoritmos para calcular o melhor número de partições de dados em cada nó.

Em (TRAINA *et al.*, 2001; TRAINA *et al.*, 2007), foi proposto uma técnica denominada *Omni-Technique* que amplifica o poder de poda da estrutura em árvore de um MAM, e gera um novo método de acesso métrico, da chamada *Omni-Family*. Logo no início da construção da estrutura são selecionados vários elementos do conjunto para tornarem-se representantes globais, denominados de *Omni-Foci*. Os elementos são escolhidos na indexação, a fim de melhorar a otimização do acesso ao espaço métrico diminuindo a quantidade de cálculo de distância necessário para responder às consultas por similaridade.

Em (VIEIRA *et al.*, 2006) foi proposto um MAM dinâmico, balanceado por densidade de dados denominado *DBM-Tree*. A *DBM-Tree* trata-se de uma estrutura hierárquica que cresce de cima para baixo e foi o primeiro MAM a controlar a sobreposição propondo um balanceamento flexível que organiza a estrutura na inserção de novos elementos. A *NOBH-Tree* (POLA; JR;

(TRAINA, 2014) é na realidade um conjunto de árvores métricas baseadas em memória principal visando reduzir o número de cálculos de distância necessários para responder às consultas por similaridade. O espaço métrico é dividido em várias regiões usando elementos como pivôs e evitando a intersecção de nós. Porém, é uma estrutura para memória principal somente.

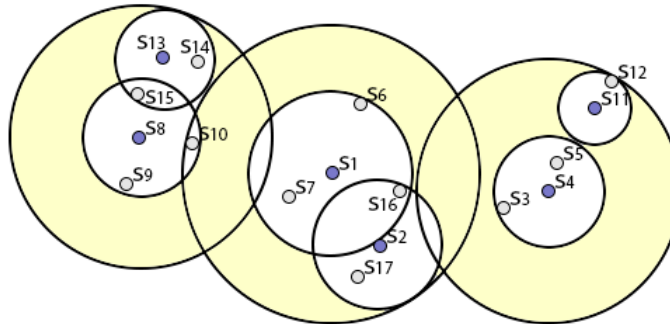
Em (ROUMELIS *et al.*, 2018) foi proposta a xBR^+ -tree. Trata-se de um MAM com objetivo de realizar carregamento em massa (*bulk-loading*) de conjuntos de dados de forma rápida, ou seja, é uma abordagem que é otimizada para se trabalhar com grandes conjuntos de dados. Nesta estrutura, o *bulk-loading* é útil quando um índice precisa ser construído pela primeira vez, pois através desta metodologia sabe-se os dados com antecedência e se são conjuntos de dados grandes. As xBR^+ -trees são balanceadas e baseadas em disco, que subdivide o espaço métrico em quadrantes (*Quadtree Family*), também possui um algoritmo que faz uso da memória principal disponível visando melhorar o desempenho no processamento de consultas.

3.4.1 Slim-Tree

Em (TRAINA *et al.*, 2002), foi apresentado um MAM chamado *Slim-Tree*, que é uma extensão aprimorada do *M-Tree*. O objetivo da estrutura *Slim-Tree* é indexar os objetos que compõem o espaço métrico de forma a reduzir a sobreposição de nós que podem causar perda de desempenho. Para isso, possui o algoritmo *Slim-Down*, que permite a reorganização da estrutura a qualquer momento durante sua construção. A *Slim-Tree* compartilha características similares com uma *M-Tree*, como a propriedade da desigualdade triangular, que também é utilizada para otimizar as consultas. A diferença está no método de subdivisão da árvore e na escolha das subárvores. Tanto a *Slim-Tree* quanto a *M-Tree* não possuem algoritmos de remoção após a construção da estrutura, mas sinalizam que os elementos estão logicamente removidos (BUENO, 2009).

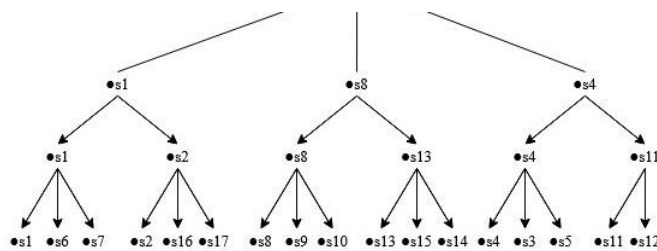
A alocação de um novo elemento em uma *Slim-Tree* verifica se o nó possui espaço para indexação, no caso de não houver espaço, é aplicado um algoritmo de alocação de um novo nó no mesmo nível da árvore, redistribuindo os elementos entre estes nós. A Figura 11 e apresenta a estrutura espacial de uma *Slim-Tree* com 3 níveis. Observa-se que os círculos na cor branca são nós folhas e os círculos na cor cinza são os nós índices. Os elementos representantes são associados aos pontos centrais na cor preta e os elementos restantes na cor cinza. Na Figura 12, a estrutura lógica representa a árvore associada à estrutura espacial.

Figura 11 – Exemplo de uma *Slim-tree* organizando 17 elementos, utilizando espaço bidimensional e distância Euclidiana. Os nós são organizados em três níveis, e representados como círculos devido ao uso da distância Euclidiana



Fonte: Elaborada pelo autor.

Figura 12 – Representação hierárquica (árvore) de uma *Slim-tree* organizando 17 elementos, para os dados da Figura 11. Observe que a *Slim-tree* não possui elementos no nó raiz



Fonte: Elaborada pelo autor.

A *Slim-Tree* é um MAM que mantém a organização dos elementos em uma estrutura hierárquica, isto é, a construção de sua estrutura ocorre a partir da raiz. O algoritmo de construção da árvore possui o critério de localizar um nó representante que possua a área de cobertura que inclui o novo elemento. Caso nenhum nó satisfaça este critério é selecionado o nó com o centro mais próximo do elemento. Em caso de haver empates na seleção de um nó representante, é executado o algoritmo *ChooseSubTree* para realizar o desempate. Para isto, a *Slim-Tree* possui três implementações para o algoritmo:

- **random**: Seleciona aleatoriamente entre o conjunto de nós representantes.
- **mindist**: É calculada a distância entre os nós representantes, o nó com a menor distância é selecionado.
- **minocup**: Entre os nós representantes, é escolhido aquele com menor ocupação.

Quando um novo nó tenta ser alocado em um nó no qual seu espaço de armazenamento atingiu sua capacidade máxima, o nó representante é particionado em dois e seus elementos são

redistribuídos para um novo nó representante que se torna um nó índice. A divisão de um nó representante em uma *Slim-Tree* possui três implementações possíveis:

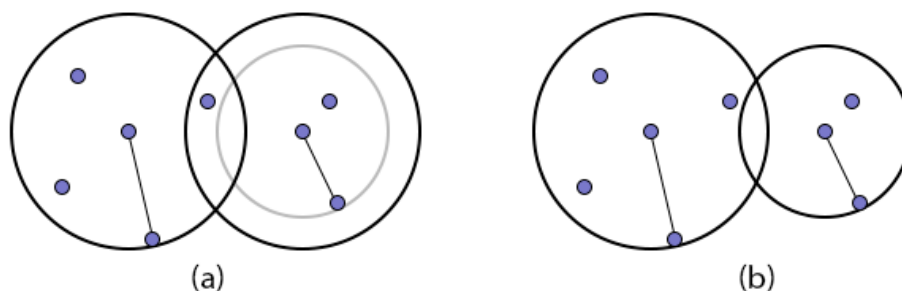
- **random**: Dois elementos são aleatoriamente selecionados e os elementos existentes são distribuídos entre eles. A cada par de nós selecionados, é calculada a distância entre eles e o novo centro, em seguida é escolhido o nó com a menor distância.
- **minMax**: Todos os pares de elementos são considerados possíveis novos representantes. Para cada par, um algoritmo linear distribui os elementos entre os representantes. O elemento representante que minimiza a área de cobertura é escolhido.
- **MST**: Uma árvore geradora mínima (*Minimal Spanning Tree*) dos elementos é criada e o maior arco é retirado. Os agrupamentos formados definem os novos nós representantes com seus conjuntos de elementos.

Em geral, esta estrutura mantém a árvore balanceada para minimizar a altura, a fim de otimizar e diminuir o número de acessos ao disco. Porém, na maioria dos MAMs dinâmicos podem ocorrer a sobreposição de regiões estabelecidas pelos nós que pertencem ao mesmo nível da árvore, sendo necessária a análise desses nós em uma busca por profundidade e em largura. A *Slim-Tree* foi proposta de forma a conseguir medir a sobreposição dos nós e a reduzir a sobreposição de regiões nos espaços métricos. Para isto, esta estrutura fornece o fator de sobreposição *Fat-Factor* e o algoritmo de reorganização denominado *Slim-Down*.

O fator de sobreposição *Fat-Factor* é a medida de avaliação do grau de sobreposição que os nós da árvore possuem. Este fator foi utilizado para o desenvolvimento do algoritmo de reorganização *Slim-Down*. Este algoritmo tem como objetivos a redução da sobreposição entre os nós que ocorre a partir da troca dos elementos entre os nós folhas. A seguir são apresentados os passos do ciclo de funcionamento do algoritmo, também exemplificado na [Figura 13](#).

1. Para cada nó i no nível h da árvore, encontre o elemento c mais distante do elemento representante.
2. Selecione o nó j vizinho do i , na qual a área de cobertura inclua o elemento c . Se o nó j selecionado existir e não estiver cheio, remova o elemento c do nó i e insira no nó j .
3. Se o nó i não estiver vazio, corrija o raio do nó i , caso contrário exclua o nó.
4. Os passos de 1 a 3 são aplicados sequencialmente sobre todos os nós de um determinado nível da árvore. Após este ciclo, se um elemento mudou para outro nó, é repetido novamente os passos de 1 a 3.

Figura 13 – Exemplo do funcionamento do *Fat Factor* no método de acesso métrico *Slim-Tree*, demonstrando a redução de sobreposição em uma região para otimizar a recuperação de dados



Fonte: Elaborada pelo autor.

A escolha da *Slim-Tree* como estrutura para incluir funções de distâncias ponderada reside no seu recurso de balanceamento da árvore métrica. Pois a distribuição dos dados pode ser afetada pelas transformações da função de distância. A *Slim-Tree* é especialmente projetada para lidar com essas mudanças na geometria do espaço métrico. O balanceamento da árvore é um fator crucial na eficiência da estrutura.

A *Slim-Tree*, por meio de técnicas avançadas de particionamento de dados e reorganização, mantém a árvore balanceada. Isso significa que a profundidade da árvore é controlada, evitando ramos extremamente longos e assegurando que o tempo de busca seja mais previsível e reduzido.

A Tabela 1 apresenta um resumo das principais características dos MAMs (índices de acesso multidimensionais) dinâmicos. A tabela destaca nove tipos de MAMs dinâmicos: *M-Tree*, *Slim-Tree*, *MM-Tree*, *Onion-Tree*, *Omni-Family*, *DBM-Tree*, *NOBH-Tree* e *xBR+Trees*. O campo **Dinamicidade** indica que todos os MAMs são capazes de lidar com atualizações dinâmicas nos dados. Quanto ao **Armazenamento**, alguns MAMs utilizam memória secundária, enquanto outros utilizam memória primária ou uma combinação de ambas. A **Divisão do Espaço** refere-se à forma como o espaço é dividido pelos MAMs para organizar os dados, como bolas, anéis ou quadrantes. A **Função de Distância Flexível** indica se os MAMs suportam funções de distância flexíveis para medir a similaridade entre os objetos. O campo **Aprende com usuário** indica se os MAMs possuem métodos em sua construção que permitem aprender com as preferências do usuário, ou seja, transformação do espaço métrico de forma dinâmica. Os diferentes algoritmos suportados pelos MAMs são listados, incluindo inserção, otimização, remoção e busca. O **Tratamento de Sobreposição** indica se os MAMs permitem sobreposição parcial ou se evitam completamente a sobreposição. Por fim, a **Construção** indica se a árvore métrica dos MAMs são construídos de baixo para cima ou de cima para baixo. Essas informações são essenciais para entender e escolher o MAM mais adequado para cada aplicação, na qual nesta pesquisa, a *Slim-Tree* foi escolhida por possuir algoritmos de tratamento de sobreposição devido a calibração de distância do espaço métrico.

Tabela 1 – Resumo das principais características dos MAMs Dinâmicos.

	M-Tree	Slim-Tree	MM-Tree	Onion-Tree	Omni-Family	DBM-Tree	NOBH-Tree	xBR ⁺ -trees
Dinamicidade	Dinâmico	Dinâmico	Dinâmico	Dinâmico	Dinâmico	Dinâmico	Dinâmico	Dinâmico
Armazenamento	Memória Secundária	Memória Secundária	Memória Primária	Memória Primária	Memória Primária	Memória Secundária	Memória Primária	Memória Principal e Secundária
Divisão do Espaço	Bola	Bola	Bola	Bola	Anel	Bola	Bola	Quadrantes
Função de Distância Flexível	Não	Não	Não	Não	Não	Não	Não	Não
Aprende com usuário	Não	Não	Não	Não	Não	Não	Não	Não
Algoritmos	Inserção e Busca	Inserção, Otimização e Busca	Inserção, Semi-otimização e Busca	Inserção, Substituição de pivôs e Busca	Inserção, Otimização e Busca	Inserção, Otimização, Remoção e Busca	Inserção e Busca	Inserção, Otimização e Busca
Tratamento de Sobreposição	Não	Sim	Sem sobreposição	Sem sobreposição	Sim, parcial	Sim, parcial	Sem sobreposição	Sim, parcial
Construção	Top-down	Top-down	Top-down	Bottom-up		Top-down	Top-down	Bottom-up

3.5 Percepção do Usuário em Consultas por Similaridade

Os perfis dos usuários podem ser utilizados para a análise semântica dos dados de um domínio específico. A captura desses perfis pode ser realizada de modo estático ou dinâmico. Essa captura estática é realizada pela entrada manual ou pela importação feita pelo especialista desses dados, bem como das informações de análises complementares que podem ser adicionadas posteriormente ao perfil do especialista registrado. Já a captura dinâmica dos perfis é realizada durante o processo de uma busca por similaridade definindo os parâmetros que o usuário deseja de uma consulta. Na área da medicina, diferentes usuários com níveis diferentes de especialização podem aplicar informações e análises em imagens e alimentar sistemas CBIR (*Content-based Image Retrieval* ou Recuperação de Imagens Baseada em Conteúdo) (VISHRAJ; GUPTA; SINGH, 2022; NESSO *et al.*, 2018; AZEVEDO-MARQUES *et al.*, 2017). As informações são utilizadas para destacar a busca em um subconjunto específico de imagens armazenadas no banco de dados.

Em sistemas CBIR são estudados diversos descritores que representam o conteúdo intrínseco das imagens. A base dos sistemas CBIR é definir quais características devem ser empregadas para identificar adequadamente uma imagem. Tradicionalmente as principais são as propriedades de cor, textura e forma (GONZALEZ; WOODS, 2018). Um dos maiores desafios é encontrar uma forma de reduzir a descontinuidade semântica, ou seja, a divergência entre a compreensão humana sobre imagens e a compreensão que pode ser obtida pelo sistema computacional (YIN *et al.*, 2021; SHRIVASTAVA; BHOYAR; ZADGAONKAR, 2017; DESERNO; ANTANI; LONG, 2009). As imagens podem ter significados diferentes para cada pessoa e depende do conhecimento e experiência que cada um tem ao analisar uma imagem. Desta forma, a precisão em uma consulta por similaridade está associada à percepção do usuário.

A descontinuidade semântica pode ser reduzida pelo uso de diferentes descritores nas imagens, permitindo ao usuário diversas opções, de forma que as consultas satisfaçam melhor às suas expectativas (AHMED, 2020; PONCIANO-SILVA *et al.*, 2009). A utilização de um ou mais descritores depende da aplicação e preferência do usuário que realiza a análise das imagens. Várias técnicas têm sido estudadas com o objetivo de tornar a interação do usuário com o sistema o mais acessível possível, ou seja, ao especificar uma consulta, retirar do usuário a tarefa de ajustar manualmente os valores de parâmetros de baixo nível e possibilitar o refinamento de forma transparente (BEDO *et al.*, 2016).

Em (LI; CHANG; WU, 2003) foi proposta uma métrica denominada função de distância parcial, que visa mensurar a similaridade perceptual em imagens através de uma função de distância, e que auxilia a redução da descontinuidade semântica, em domínios de imagens do cotidiano, tais como paisagens, flores, carros, etc. Conseguir quantificar e caracterizar a similaridade perceptual, pode aumentar a precisão na modificação dos pesos da função de distância e criar uma métrica que se aproxime do desejo do usuário. Desta forma, um sistema ideal deve conseguir retornar resultados que satisfaça a expectativa do usuário, bem como mensurar a similaridade perceptual que dê suporte e tenha consistência com a percepção humana.

3.6 Considerações Finais

Neste capítulo foi realizada uma revisão das principais fundamentações da literatura em consulta por similaridade e abordagens de métodos de acesso métrico, sendo observadas as técnicas de otimização e avaliação de qualidade para cada uma delas. Foi também abordada a percepção do usuário em sistemas médicos.

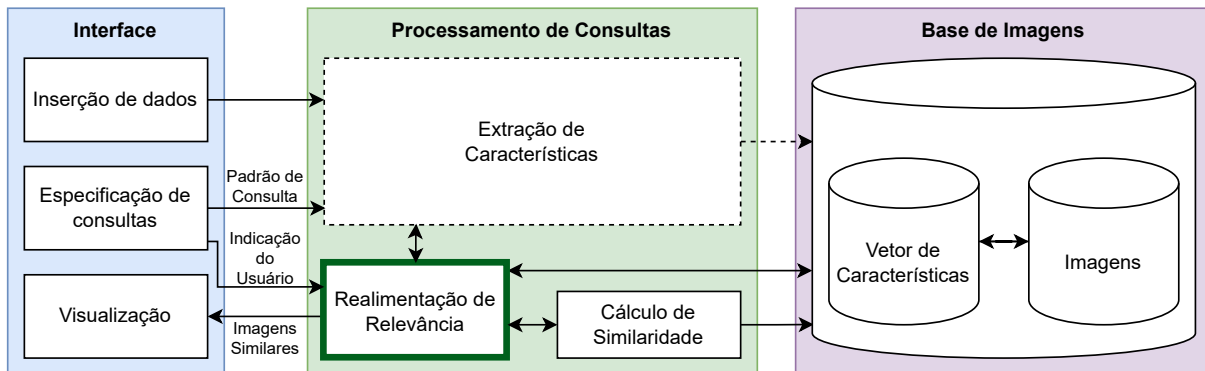
REALIMENTAÇÃO DE RELEVÂNCIA E TRABALHOS RELACIONADOS

4.1 Visão Geral

A área de recuperação de informação apresenta diversos problemas em aberto. O principal problema está na dificuldade de representar todo conteúdo presente em uma imagem, ou seja, ocorre a descontinuidade semântica entre as características de baixo nível extraídas automaticamente e a resposta esperada pelo usuário. Segundo (AHMED, 2020; DEPEURSINGE *et al.*, 2011), o problema está na falta de ferramentas interativas que permitam ao especialista avaliar e testar o que é realmente similar em uma imagem considerando sua percepção. Assim, em uma consulta, utilizando a relação entre a quantidade de resultados significativos retornados e o total de resultados que deveriam ser retornados, pode-se construir curvas de precisão e revocação, que auxiliam na análise de eficácia da recuperação de dados utilizando os descritores dos dados complexos (imagens nesse projeto).

Nos sistemas CBIR, os tratamentos de consultas por similaridade devem ter um nível de precisão o mais alto e próximo possível do esperado pelo usuário, o que pode ser alcançado envolvendo os critérios da percepção do especialista do domínio. Para isso, é adicionado um novo submódulo referente à **Realimentação de Relevância** na camada de **Processamento de Consultas**, a qual fica responsável pela comunicação direta com a **Interface** onde o usuário consegue informar sua percepção do que é relevante ou não em uma consulta, assim permitindo a integração das técnicas de realimentação de relevância para diminuição do *gap* semântico. A **Figura 14** apresenta a arquitetura de um sistema CBIR com adição do submódulo de **Realimentação de Relevância**.

Figura 14 – Arquitetura de um sistema CBIR com integração de um módulo de realimentação de relevância no Processamento de Consultas, visando a capacidade de aprimorar a eficácia dos resultados



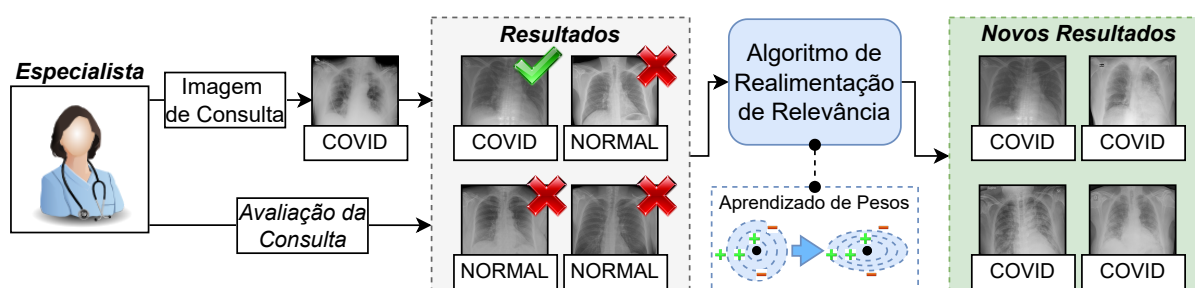
Fonte: Elaborada pelo autor.

Um sistema de recuperação de imagens por conteúdo é baseado na comparação de seus elementos por meio de seus descritores, sendo que as imagens são recuperadas por meio de consultas por similaridade. Ou seja, dada como entrada uma base de imagens, o usuário deseja recuperar as imagens mais similares (segundo o seu entendimento pessoal) a um elemento de consulta. Essa percepção de similaridade pode variar de usuário para usuário, considerando diversos fatores, inclusive seu grau de expertise, interesse e foco no momento, entre outros.

Nesse projeto, deve-se incluir o *feedback* do usuário, de modo apropriado e eficiente, para que a percepção do usuário seja realmente aproveitada para realizar reindexações do método de acesso métrico. Ou seja, Ao mesmo tempo que a consulta será reprocessada com o ajuste fino sobre as amostras fornecidas, o algoritmo deverá trabalhar sobre o histórico de ciclos de realimentação para escolher a melhor métrica para o conjunto de dados.

Em geral, as técnicas de RR atribui diferentes níveis de relevância aos atributos presentes no vetor de características do conjunto de dados. Esses atributos representam a semântica dos dados complexos, em que cada atributo é um valor numérico que representa uma característica específica. Quando uma abordagem demonstra alta eficácia semântica, significa que ela é capaz de fornecer respostas precisas e relevantes, levando em consideração o contexto e a intenção do usuário. No processo de uma RR, o usuário que realiza a consulta, julga as imagens retornadas como relevantes ou irrelevantes. Em uma próxima iteração, pode-se gerar novas consultas retornando um conjunto mais próximo da expectativa do usuário. Esse ciclo de refinamento de consultas com a indicação do usuário de quais respostas são relevantes e quais não são relevantes é o núcleo do processo de realimentação de relevância, conforme é exemplificado na [Figura 15](#).

Figura 15 – Ciclo de um processo de Realimentação de Relevância, destacando a interação iterativa do usuário que é incorporado a um processo de aprendizado de pesos, convertendo em melhores resultados



Fonte: Elaborada pelo autor.

A recuperação de informação ocorre através do descritor do dado complexo, que é a combinação de um vetor de características e uma função de distância que calcula a similaridade entre os elementos. Após a realização da primeira consulta feita pelo usuário, a função de distância ponderada pode receber modificações em seus pesos modificando o vetor de características para diminuir a descontinuidade semântica. A ponderação atribui um peso para cada dimensão do vetor de características, a fim de destacar ou reduzir a influência de cada dimensão. A literatura possui várias abordagens como as apresentadas nos trabalhos de (RUI *et al.*, 1998; LIU *et al.*, 2007; WAN; LIU, 2006; TIAN, 2018; AHMED, 2020), que utilizam as informações explícitas ou implícitas da consulta original, onde o usuário realimenta o sistema julgando os resultados relevantes e irrelevantes, permitindo o ajuste iterativo da busca. Desta forma, é possível fazer com que a consulta por similaridade utilize a composição desses vetores de características, aumentando a precisão a cada nova iteração.

A literatura apresenta diversas abordagens para técnicas de RR. As principais metodologias de forma resumida são:

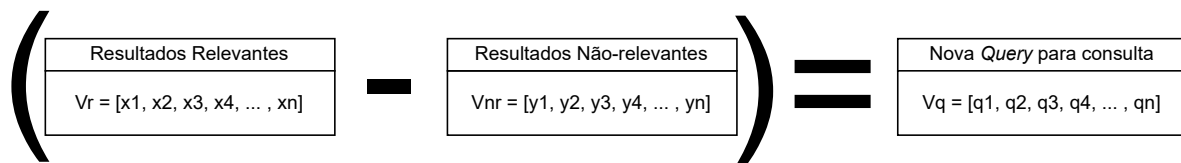
- **Reformulação de Consultas:** Têm o objetivo de determinar um ponto ideal de referência no espaço métrico, aproximando-se de elementos positivos e afastando-se de elementos negativos.
- **Refinamento da Medida de Similaridade:** Trabalha com vetores de ponderação das características que auxiliam no cálculo de similaridade de imagens.
- **Aprendizado de Máquina:** Trabalha com amostras rotuladas para o modelo melhorar sua capacidade de fazer previsões e realizar classificações de dados.

4.2 Reformulação de Consultas

A reformulação de consultas é uma abordagem de RR que a partir do *feedback* de elementos fornecido pelo usuário, é gerado um novo elemento de consulta baseado na percepção do que é relevante para o usuário. Em dados complexos, este elemento é um vetor de características que é utilizado como uma nova consulta movimentando o centro de uma consulta por similaridade, a fim de que o novo centro se aproxime de elementos mais relevantes.

Existem diversas formas de gerar uma consulta com base no *feedback* do usuário. Porém todas partem do princípio do julgamento de elementos relevantes e não relevantes que iniciam a partir de uma primeira consulta. Este julgamento pode ser capturado automaticamente pelo sistema ou pela interação com o usuário. Cada elemento julgado é um vetor de característica que representa este dado complexo, imagem neste caso, com o nível semântico baixo (valores reais). De acordo com (AHMED, 2020; GULDOGAN; GABBOUJ, 2009), as estratégias de recuperação de informação que utilizam técnicas de realimentação de relevância são classificadas pelo refinamento da consulta e são baseadas no modelo de representação vetorial (SALTON; MCGILL, 1983). As estratégias base que aplicam a técnica de RR criam uma “pseudo-imagem” a partir das imagens julgadas pelo usuário, e executam uma nova consulta ou múltiplas consultas com esta imagem a fim de que o resultado esteja mais próximo das expectativas do usuário. Comumente a nova consulta é obtida pela diferença entre o conjunto do vetor de características de elementos positivo e negativos como mostrado na Figura 16. Esta estratégia foi utilizada em (KARAMTI *et al.*, 2018; MARTINET; CHIARAMELLA; MULHEM, 2011).

Figura 16 – Reformulação de consultas em realimentação de relevância, no qual a subtração do conjunto de resultados relevantes do conjunto de resultados não-relevantes leva à criação de uma nova consulta aprimorada

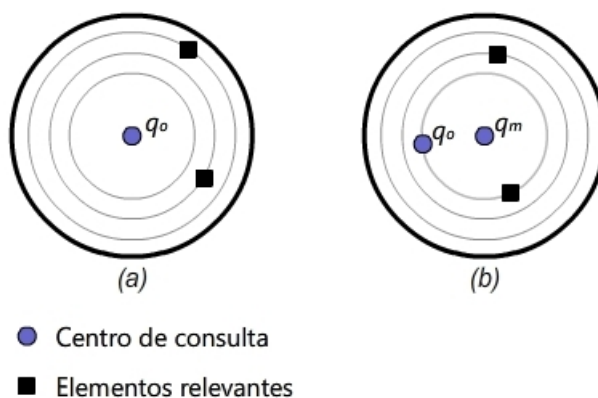


Fonte: Elaborada pelo autor.

A Movimentação do Centro de Consulta (do inglês *Query Point Movement* - QPM) é uma das técnicas mais utilizadas para realimentação de relevância que otimizam uma consulta, pois trata-se de uma estratégia que tenta adequar a consulta o mais perto das expectativas do usuário (SU *et al.*, 2022; NOVAES *et al.*, 2021; ROCCHIO, 1971). Seu funcionamento acontece na movimentação do centro de consulta original, isto é, a partir da realimentação fornecida pelo usuário, constrói-se um novo centro que se aproxima de elementos relevantes e se afasta de elementos irrelevantes. A Figura 17 (a) representa a consulta inicial e a Figura 17 (b) ilustra o novo centro de consulta após a aplicação da técnica, aprimorando o resultado da consulta.

Em geral, o resultado da movimentação do centro de consulta é realizado pela maximização da diferença entre o vetor de características da consulta original e a média dos vetores de características das imagens rotuladas como relevantes.

Figura 17 – Exemplo da técnica de movimentação do centro de consulta, na qual (a) representa a consulta inicial e (b) representa o novo centro de consultas com o objetivo de se aproximar de elementos relevantes



Fonte: Elaborada pelo autor.

A cada nova iteração é aplicada a realimentação de relevância, e a percepção do usuário pode determinar o grau de relevância de cada uma. Sendo a realimentação composta por exemplos positivos e negativos, a fórmula de Rocchio é empregada para a movimentação do centro de consulta. Esta fórmula é apresentada a seguir, onde Q_0 é o vetor de características inicial da consulta, Q_1 é o novo vetor de características gerado, r é o número de elementos relevantes, s é o número de elementos irrelevantes, R_i é o i -ésimo vetor de características relevante e S_i é o i -ésimo vetor de características de elementos irrelevantes.

$$Q_1 = Q_0 + \frac{1}{r} \sum_{i=1}^r R_i - \frac{1}{s} \sum_{i=1}^s S_i$$

Na literatura, há estudos sobre qual abordagem entre realimentação positiva e negativa possui maior efeito para melhorar uma consulta. Em (RUI; HUANG; MEHROTRA, 1997) é abordado somente o uso de realimentação positiva para gerar uma nova consulta. Rui, Huang e Mehrotra (1997) afirma que a melhoria da precisão após utilizar somente elementos positivos é limitado, desaparecendo após poucas iterações, esta abordagem é utilizada em (BANERJEE *et al.*, 2018; MARKONIS; SCHAER; MÜLLER, 2016). Em (MULLER *et al.*, 2000) é feito um estudo sobre o uso de realimentação negativa, Muller *et al.* (2000) afirma que algumas técnicas mostram uma grande melhora na consulta utilizando apenas feedback de elementos não relevantes desde que sejam elementos julgados corretamente, caso contrário pode até degradar uma consulta. Esta abordagem é utilizada em (WANG; FANG; ZHAI, 2008; WANG; FANG;

ZHAI, 2007). Em (YAN; HAUPTMANN; JIN, 2003; KHERFI; ZIOU; BERNARDI, 2002) é abordado o impacto da contribuição do uso da realimentação negativa e como utilizá-la em conjunto com a realimentação positiva. Verifica-se que o uso combinado dessas abordagens apresentam melhores resultados (XU; WANG; MAO, 2017; SU *et al.*, 2003).

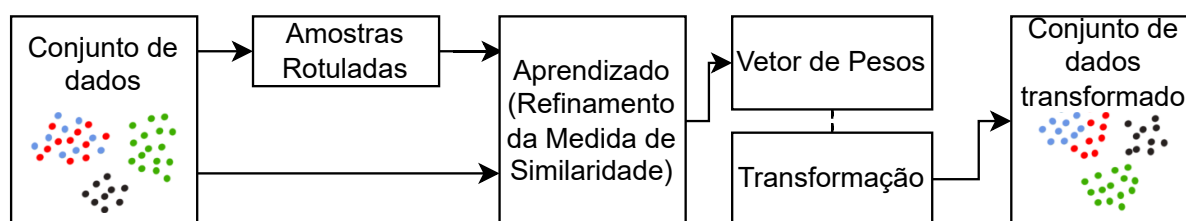
Vale ressaltar que as técnicas de otimização de consulta alteram somente o centro de consulta em uma busca, ou seja, não refinam a métrica de similaridade ou realizam modificações na árvore métrica construída pelos métodos de acesso métrico. Desta forma, as formas de armazenar a percepção do usuário em uma consulta para serem reutilizadas em novas consultas são minimizadas, fazendo com que o usuário passe sempre pelo processo de realizar um ciclo de realimentação para cada nova consulta realizada. Portanto, este tipo de abordagem não é a melhor opção para organizar e armazenar novas árvores métricas no MAM.

4.3 Refinamento da Medida de Similaridade

Otimização de função de distância ou refinamento de similaridade é uma abordagem combinada com as técnicas de realimentação de relevância a partir do feedback realizado pelo usuário. São realizadas transformações no formato de busca da função de distância, de forma que os resultados de uma consulta evite resultados não relevantes. Esta abordagem permite ao usuário ajustar (ponderar os pesos) de funções de distância modificando todo formato de busca do espaço métrico. Adaptando esta abordagem com os MAMs, é possível o armazenamento das especificações do usuário e otimizar novas consultas. No contexto médico, é possível criar mecanismos para que especialistas experientes indiquem esses pesos das funções de distância que serão utilizadas nos métodos de acesso métrico para responder **consultas por similaridade** e apoiar a tomada de decisões dos profissionais.

As técnicas que possuem o objetivo de refinar a medida de similaridade seguem uma mesma metodologia de aprender um vetor de pesos que consiga atribuir mais relevância às características do dado complexo, de forma a atribuir maior similaridade entre dados de mesma classe. Para isto, são utilizadas amostras rotuladas, geralmente fornecidas diretamente pelo usuário. Em seguida, as amostras são passadas como entrada para o algoritmo de aprendizado de pesos realizar o processo de encontrar uma métrica que realiza uma boa separação dos elementos baseado no *feedback* fornecido. No final do processo, o vetor de pesos aprendido é utilizado para ponderar toda a matriz de características dos dados, gerando um novo espaço métrico considerado mais eficaz para responder consultas por similaridade. A Figura 18 apresenta o processo de técnicas de realimentação de relevância que refinam a medida de similaridade.

Figura 18 – Processo de técnicas de realimentação de relevância que refinam a medida de similaridade, onde o refinamento são alcançados por meio do aprendizado de pesos, transformando um conjunto de dados em um espaço métrico aprimorado



Fonte: Elaborada pelo autor.

Os primeiros trabalhos a utilizar esta abordagem de refinamento da medida de similaridade foram propostos nos sistemas MARS (PORKAEW; CHAKRABARTI, 1999) e *MindReader* (ISHIKAWA; SUBRAMANYA; FALOUTSOS, 1998), por utilizarem métodos estatísticos como o desvio padrão para obtenção dos pesos, que é uma abordagem consolidada e bastante utilizada na literatura. Outros trabalhos que utilizam esta abordagem são (BHOSLE; KOKARE, 2020; CHÁVEZ *et al.*, 2013; SILVA *et al.*, 2012).

O desvio padrão é um conceito em estatística que fornece informação sobre a dispersão dos dados e mede a distância da média dos elementos em relação à média simples. Desta forma, é possível medir se os elementos estão bem concentrados ou se estão dispersos. O sistema MARS (PORKAEW; CHAKRABARTI, 1999) baseia os pesos sobre o inverso do desvio padrão de elementos relevantes, enquanto o sistema *MindReader* (ISHIKAWA; SUBRAMANYA; FALOUTSOS, 1998) considera pesos proporcionais ao inverso da variância entre os atributos relevantes do vetor de característica.

Vale ressaltar que a otimização de uma consulta que modifica a função de distância cria um novo espaço métrico, ou seja, no contexto de métodos de acessos métricos, gera uma nova árvore métrica onde os nós são os elementos e as arestas são as distâncias por similaridade entre esses elementos. Assim, modificar a função de distância a cada interação ou consulta realizada no sistema, e em seguida armazenar esse novo espaço métrico gerado, resultará em um aumento do custo computacional em relação ao espaço de armazenamento de um método de acesso métrico, bem como ampliará o tempo de resposta à uma consulta, devido a construções de novos métodos de acesso métrico em tempo real. Para isso é necessária uma metodologia que reindexe os métodos de acesso métrico que mantenham a otimização do espaço métrico com distâncias ponderadas de forma que não degrade o tempo de resposta para o usuário.

4.4 Trabalhos Relacionados e Descrição do Problema

A precisão de consultas por similaridade é importante para a tomada de decisão por especialistas de um domínio, assim também como o tempo de processamento destas consultas é

um fator relevante. Os trabalhos da literatura detalham algumas técnicas para ampliar a precisão de consultas por similaridade.

No trabalho de (RUI *et al.*, 1998), os autores definiram uma heurística com atualização de parâmetros baseada em otimização da matriz de distâncias, e descobriram que o método baseado em otimização com elementos relevantes alcança maior precisão. O objetivo principal é aprender com o *feedback* dos usuários sobre a qualidade dos resultados retornados e, em seguida, ajustar o algoritmo de busca para fornecer resultados mais relevantes. O método utiliza o desvio padrão das pontuações de similaridade em conjunto de três pesos pré definidos para realizar transformações na matriz de características de todo conjunto de dados. O desvio padrão é uma medida que indica quão uniforme são os dados do conjunto ressaltando os atributos dos vetores de características. Apesar de o desvio padrão ser considerado uma medida de baixa complexidade, na técnica proposta de (RUI *et al.*, 1998), a utilização de múltiplos pesos em conjunto ao cálculo do desvio padrão para cada elemento relevante exige recursos computacionais consideráveis para lidar com grandes volumes de dados. Isso pode tornar a técnica menos prática em cenários de busca em tempo real ou com muitos usuários simultâneos, sendo comumente utilizado como técnica para otimizar os dados em pré-processamento utilizando amostras rotuladas.

No trabalho de (CHANG; KAMATAKI; CHEN, 2009), a distorção do espaço de características utilizando vetores de deslocamento médio foi proposta como uma maneira de preencher a lacuna entre a semântica de alto nível e os recursos de baixo nível. O vetor de deslocamento médio representa a direção e a magnitude do deslocamento necessário para mover cada elemento da base de dados em direção aos elementos relevantes. Para cada elemento do conjunto de dados, o algoritmo calcula o centro de massa ou média ponderada das amostras que estão dentro de uma determinada janela de busca ao redor desse elemento, encontrando de forma eficaz concentrações de elementos de mesma classe. Entretanto, é um algoritmo comumente utilizado no pré-processamento de dados. Pois o algoritmo é computacionalmente intensivo, especialmente em grandes conjuntos de dados, uma vez que é necessário calcular o centro de massa para cada ponto em cada iteração.

Em (WEINBERGER; SAUL, 2009), foi proposto o algoritmo de aprendizado de métrica *Large Margin Nearest Neighbor* (LMNN). O objetivo principal do LMNN é aprender uma métrica que aumente a separação entre exemplos de diferentes classes e diminua a distância entre exemplos da mesma classe. Essa abordagem permite que algoritmos de vizinhos mais próximos aproveitem melhor a matriz de características, aumentando o desempenho geral em problemas de classificação. A medida de distância utilizada é a distância de Mahalanobis, que leva em consideração a correlação entre as variáveis e é uma métrica amplamente empregada para problemas estatísticos e de classificação. O LMNN pode ser computacionalmente intensivo, especialmente em grandes conjuntos de dados. O processo de otimização requer múltiplas iterações e a manipulação de matrizes de alta dimensão em todo conjunto de dados.

Em (BRESSAN; BUGATTI; SAITO, 2019), a utilização de *Support Vector Machine*

Active (SVM-AL) com RR define um hiperplano separador entre os elementos rotulados como relevantes em imagens de mamografias, procurando uma maximização das distâncias entre os vetores de suporte de cada classe. O aprendizado ativo, é uma abordagem onde o algoritmo seleciona quais amostras que o modelo acredita serem mais informativas de dados não rotulados para adicionar ao conjunto rotulado, com o objetivo de melhorar o desempenho do modelo. A técnica SVM-AL envolve um processo iterativo de seleção de amostras, rotulagem e re-treinamento do modelo a cada iteração na base de dados.

No trabalho de (KUMARAN *et al.*, 2021), é utilizada a abordagem de aprendizado de máquina que visa melhorar o desempenho da classificação supervisionada por meio da criação de métricas de distância otimizadas e por meio do uso de aprendizado ativo para aquisição de dados relevantes. A técnica proposta utiliza abordagens de otimização para encontrar métricas de distância ótimas que generalizam a métrica de Mahalanobis. Após calcular uma métrica ótima, o método é capaz de recomendar regiões precisas para amostragem de dados, visando melhorar o desempenho da classificação. Da mesma forma de técnicas aprendizado de métrica, o processo de otimização utiliza múltiplas iterações sobre a matriz de características de todo conjunto de dados para ajustar a métrica.

Segundo a literatura anteriormente listada, as estratégias apoiam-se no desenvolvimento de técnicas eficientes para aumentar a precisão de consultas a fim de evitar falsos positivos. Entretanto, a maioria dessas estratégias demandam alto custo computacional por não fazerem uso de estruturas de indexação que otimizam a recuperação de dados complexos. Com foco neste cenário, esta pesquisa apresenta uma metodologia que inclui a técnica de RR para ponderar a função de distância no MAM Slim-Tree, analisando a eficiência do RR em melhorar a precisão da consulta e também o preço pago relativo ao tempo de processamento de MAMs ponderados em comparação à recuperação sequencial. A Tabela 2 apresenta as principais abordagens de RR que modificam a métrica de similaridade aprimorando o espaço métrico e as características.

Tabela 2 – Revisão da Literatura de abordagens de realimentação de relevância que aprendem métricas de similaridade em matriz de distâncias

Literatura	Refinamento da Medida de Similaridade	Melhora a Semântica	Utiliza poucas amostras	Aprende somente pela matriz de feedback	Utiliza MAM
(RUI <i>et al.</i> , 1998)	✓	✓	✓	✗	✗
(CHANG; KAMATAKI; CHEN, 2009)	✓	✓	✓	✗	✗
(WEINBERGER; SAUL, 2009)	✓	✓	✗	✓	✗
(BRESSAN; BUGATTI; SAITO, 2019)	✓	✓	✗	✓	✗
(KUMARAN <i>et al.</i> , 2021)	✓	✓	✗	✓	✗
<i>Tuning Metrics Relevance Feedback</i> (Nossa Pesquisa)	✓	✓	✓	✓	✓

4.5 Métricas de Avaliação

Análises Quantitativas

Matriz de Confusão

Uma matriz de confusão é uma tabela que é usada para descrever o desempenho de um modelo de classificação em um problema de aprendizado supervisionado. Ela mostra a relação entre as classes reais dos dados e as classes previstas pelo modelo. Para análises preditivas, a matriz de confusão é utilizada para mostrar em um algoritmo de classificação a frequência de elementos falsos positivos, falsos negativos, verdadeiros positivos e verdadeiros negativos.

- **Verdadeiro positivo** (*true positive* — TP): ocorre quando, no conjunto de elementos, o elemento foi classificado corretamente de acordo com a classe desejada.
- **Falso positivo** (*false positive* — FP): ocorre quando, no conjunto de elementos, o elemento foi classificado incorretamente de acordo com a classe desejada.
- **Verdadeiro Negativo** (*true negative* — TN): ocorre quando, no conjunto de elementos, o elemento foi classificado corretamente de acordo com a classe não desejada.

- **Falso negativo (*false negative* — FN):** ocorre quando, no conjunto de elementos, o elemento foi classificado incorretamente de acordo com a classe não desejada.

A matriz de confusão é organizada em linhas e colunas, onde cada linha representa a classe verdadeira de um conjunto de dados e cada coluna representa a classe prevista pelo modelo. A [Figura 19](#) apresenta um exemplo de uma matriz de confusão para um problema de classificação binária.

Figura 19 – Exemplo de uma matriz de confusão que destaca as métricas para avaliar o desempenho de um modelo de classificação

		Elemento Verdadeiro	
		Classe Positiva	Classe Negativa
Elemento Previsto	Classe Positiva	Verdadeiro Positivo (TP)	Falso Positivo (FP)
	Classe Negativa	Falso Negativo (FN)	Verdadeiro Negativo (TN)

Fonte: Elaborada pelo autor.

Precisão e Revocação

Em sistemas para recuperação por similaridade é preciso um método para mensurar a precisão das consultas. As duas medidas mais utilizadas neste contexto são a **precisão** (*precision*) e **revocação** (*recall*) que utilizam a matriz de confusão para a análise estatística. A precisão é a razão entre o número de elementos relevantes e o total de elementos recuperados pela consulta, enquanto a revocação é a razão entre o número de elementos relevantes retornados pela consulta e o total de elementos relevantes da base.

Dado um conjunto de elementos relevantes TP e elementos não relevantes FP , a fórmula da precisão é representada por:

$$P = \left(\frac{TP}{TP + FP} \right)$$

Dado um conjunto de elementos relevantes TP e elementos não relevantes que não foram retornados como resultado FN , a fórmula da revocação é representada por:

$$R = \left(\frac{TP}{TP + FN} \right)$$

Mean Average Precision

O **Mean Average Precision** (MAP) é uma métrica amplamente utilizada na área de recuperação de informação e aprendizado de máquina para avaliar a eficácia de algoritmos de

classificação e sistemas de busca. O MAP é particularmente útil quando se lida com problemas de classificação em que várias classes estão presentes e a ordem de relevância dessas classes é importante. Para isto, primeiro deve-se primeiro obter o *Average Precision*.

O *Average Precision* é calculado para cada consulta individualmente, considerando os resultados retornados em uma consulta. Ele mede a capacidade do sistema em retornar imagens relevantes em uma ordem que corresponda à sua relevância esperada. O *Average Precision* avalia a qualidade do *ranking* produzido pelo sistema.

O cálculo do *Average Precision* envolve a obtenção da **precisão** em cada ponto em que uma imagem relevante é recuperada na lista de resultados. Essa medida é então multiplicada pela relevância da imagem em questão. A soma dessas multiplicações é realizada em todos os pontos onde imagens relevantes são encontrados, e o resultado é dividido pelo número total de imagens relevantes presentes na coleção de dados. O valor resultante é a *Average Precision* para aquela consulta específica.

Para obter o MAP, calcula-se a *Average Precision* para cada consulta e, em seguida, calcula-se a média desses valores para todas as consultas presentes no conjunto de teste. Essa média reflete a eficácia geral do sistema de busca ou classificação em todo o conjunto de dados.

Acurácia e F1-Score

A **acurácia** e a **F1-Score** são duas métricas comumente usadas para avaliar o desempenho de modelos de classificação. A acurácia é uma medida simples e intuitiva que calcula a proporção de amostras corretamente classificadas em relação ao total de amostras. Ela é calculada dividindo o número de classificações corretas (verdadeiros positivos e verdadeiros negativos) pelo número total de amostras. A fórmula para a acurácia é representada por:

$$Accuracy = \frac{TN + TP}{TP + FN + TN + FP}$$

A **F1-Score** é uma métrica que combina a precisão e a revocação em uma única medida resumida. A **F1-Score** é a média harmônica da precisão e da revocação e fornece uma medida balanceada entre as duas. É especialmente útil quando há um desequilíbrio significativo entre as classes ou quando os falsos positivos e falsos negativos têm importâncias diferentes. A fórmula para a **F1-Score** é representada por:

$$F1 - Score = \frac{2 * precision * recall}{precision + recall}$$

Análises Qualitativas

Visualização de Dados por *t-Distributed Stochastic Neighbor Embedding*

O *t-Distributed Stochastic Neighbor Embedding* (t-SNE) é um algoritmo de aprendizado de máquina não supervisionado usado principalmente para visualização de dados de alta dimen-

são. Esta técnica foi proposta por (MAATEN; HINTON, 2008) e é amplamente utilizada em áreas como análise exploratória de dados, visualização de dados e redução de dimensionalidade.

O objetivo do t-SNE é mapear os dados de alta dimensão para um espaço de menor dimensão (geralmente 2D ou 3D) de modo que as relações estruturais entre os dados sejam preservadas tanto quanto possível. A t-SNE é particularmente útil quando se deseja visualizar dados complexos e entender o espaço métrico formado do conjunto de dados e avaliar a semântica da estrutura subjacente.

K-Means e Coeficiente de Silhouette

O *K-means* é um algoritmo de *clustering* (agrupamento) que busca agrupar um conjunto de objetos em K grupos (*clusters*) distintos, de forma que os objetos dentro de um mesmo grupo sejam semelhantes entre si e diferentes dos objetos de outros grupos (AHMED; SERAJ; ISLAM, 2020). O número de *clusters*, K , é definido previamente pelo usuário ou quando se sabe a quantidade de classes do conjunto de dados. O algoritmo funciona iterativamente, atribuindo os objetos aos *clusters* de acordo com a proximidade de suas características em relação aos centróides (pontos representativos) de cada *cluster*. O objetivo é minimizar a variância dentro de cada *cluster*.

O coeficiente de Silhouette é uma métrica de avaliação de clusterização que mede a qualidade e a coesão dos *clusters* obtidos em um conjunto de dados. Ele fornece uma medida de quão bem cada objeto se encaixa no seu próprio *cluster* em relação aos outros *clusters*. O coeficiente é calculado para cada objeto individualmente, levando em consideração duas medidas: a distância média entre o objeto e todos os outros objetos do mesmo *cluster* (a) e a distância média entre o objeto e todos os objetos de um *cluster* vizinho mais próximo (b). Essas distâncias podem ser calculadas usando qualquer métrica de distância, como a distância euclidiana. O coeficiente de *Silhouette* (s) para um objeto é calculado pela equação abaixo.

$$s = (b - a) / \max(a, b)$$

O coeficiente de *Silhouette* não é usado para avaliar a qualidade do espaço métrico em si, mas sim para avaliar a qualidade da clusterização em um espaço métrico, o que indica a qualidade semântica dos elementos dos *clusters*, a qual também podemos relacionar como avaliação da qualidade do espaço métrico (SHAHAPURE; NICHOLAS, 2020; ROUSSEEUW, 1987). O coeficiente de *Silhouette* mede a coesão dos *clusters* em relação à separação entre os *clusters* em um conjunto de dados específico. O valor do coeficiente de Silhouette varia de -1 a 1. Um valor próximo de 1 indica que o objeto está bem ajustado ao seu *cluster* e mal ajustado aos *clusters* vizinhos, o que é desejável. Um valor próximo de -1 indica o oposto, ou seja, o objeto está mal ajustado ao seu *cluster* e bem ajustado aos *clusters* vizinhos. Um valor próximo de 0 indica que o objeto está em ou próximo da fronteira entre dois *clusters*.

4.6 Considerações Finais

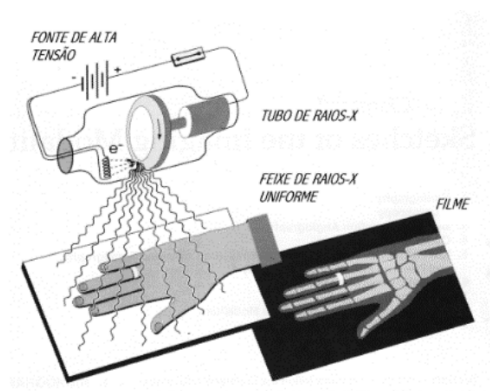
Neste capítulo foi realizada uma revisão das principais fundamentações da literatura em técnicas de realimentação de relevância. Dentro deste tema foram abordadas as métricas de avaliação, as desvantagens trazidas pelo *gap* semântico e a otimização trazidas por técnicas de realimentação de relevância que alteram o espaço métrico para serem otimizadas por métodos de acesso métrico. Este capítulo não teve por objetivo esgotar o assunto, pois existem outras vertentes que poderiam ser consideradas, mas ampliariam significativamente o escopo do trabalho. Esse não é o objetivo dos autores, mas é interessante ressaltar que técnicas de diversidade complementares às de similaridade e mesmo de aprendizado profundo poderiam ser também agregadas à proposta, em outro nível de interesse do usuário e especialista.

AQUISIÇÃO E DESCRIÇÃO DAS BASES DE IMAGENS

5.1 Visão Geral

As imagens médicas, como as de Raio-X, desempenham um papel fundamental no diagnóstico e tratamento de uma ampla variedade de doenças e condições médicas. Essas imagens fornecem informações valiosas sobre a estrutura e composição dos tecidos do corpo humano. Com o avanço da tecnologia, os sistemas de aquisição de imagens médicas têm se tornado cada vez mais sofisticados, permitindo uma visualização mais precisa e detalhada do corpo humano. As imagens de Raio-X são um dos tipos mais comuns de imagens médicas utilizadas na prática clínica. Elas são obtidas ao passar feixes de raios-X através do corpo do paciente e capturar a absorção desses raios pelos tecidos. A [Figura 20](#) apresenta o exemplo de captura de uma imagem de Raio-X pulmonar.

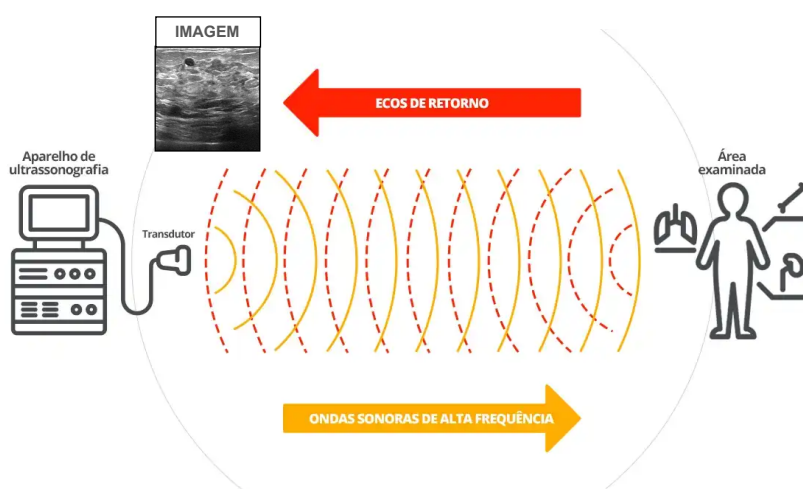
Figura 20 – Exemplo do procedimento de captura de uma imagem de Raio-X, com a mão do paciente posicionada para a passagem dos feixes de raios-X, permitindo a medição da absorção pelos tecidos e a obtenção de uma imagem para diagnóstico



Fonte: [Dimenstein et al. \(2022\)](#).

As imagens de ultrassom são obtidas por meio do uso de ondas sonoras de alta frequência que são emitidas e refletidas pelos tecidos do corpo. Essas ondas sonoras são convertidas em imagens em tempo real, permitindo uma visualização dinâmica dos órgãos, tecidos e fluxo sanguíneo. O ultrassom pode fornecer informações adicionais sobre a textura, a forma e a vascularização dos tecidos mamários, auxiliando na identificação de lesões suspeitas, como nódulos ou massas. A [Figura 21](#) apresenta o exemplo de captura de uma imagem de ultrassom do tipo mamografia.

Figura 21 – Exemplo de procedimento para a captura de uma imagem de ultrassom, obtida por meio do uso de ondas sonoras de alta frequência que são emitidas e refletidas pelos tecidos do corpo para visualização médica



Fonte: Elaborada pelo autor.

No entanto, a mera aquisição das imagens de Raio-X e imagens de ultrassom não é suficiente para uma análise completa e precisa. A extração de características desempenha um papel fundamental na interpretação e compreensão dessas imagens e também necessitam do aumento semântico de certas características para contribuir a tomadas de decisões.

Nesta pesquisa foram realizadas análises em duas bases de imagens médicas, sendo a primeira de imagens de Raio-X Pulmonares de COVID-19 e segunda uma base de imagens de ultrassom de mamas. A extração de características por meio de extratores de textura possibilita uma análise mais precisa e automatizada das imagens médicas.

No contexto de imagens pulmonares relacionadas ao COVID-19 e de imagens de ultrassom para identificação de tumores em mamas, as características de textura têm sido amplamente utilizadas ([BENAOUALI et al., 2022](#); [REZAZADEH; JAFARIAN; KORD, 2022](#); [BARDHAN; ROGA, 2021](#); [PEREIRA et al., 2020](#)). Desta forma, em ambas as bases de imagens adquiridas, foram extraídas características de nove extratores de textura e um extrator de histograma de cinzas que foram abordados na [Seção 2.3](#) e utilizados para experimentos do método desenvolvido nesta pesquisa. Para extração de características, os extratores e seus respectivos parâmetros são apresentados na [Tabela 3](#).

Tabela 3 – Parâmetros e dimensões dos extratores de características.

Extratores	Parâmetros	Dimensões
FOS	default parameters	16
GLCM	default parameters	28
SFM	Lr=Lc=4	4
LTE	3	6
Gray Histogram	default parameters	32

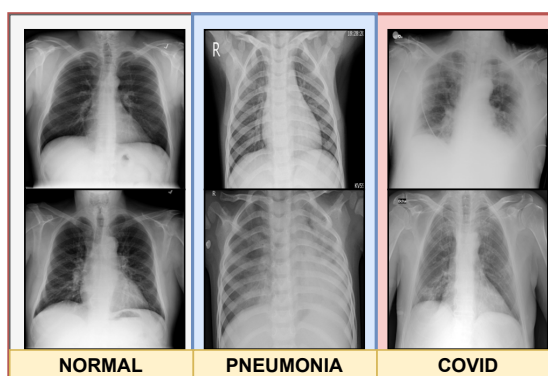
5.2 Análise de Imagens de COVID-19 e Mamografias e Gerais

COVID - 19

As imagens de Raio-X pulmonares são um recurso valioso no contexto da COVID-19. Essas imagens podem revelar opacidades, infiltrados e padrões característicos associados à infecção pelo vírus. As radiografias torácicas podem ser usadas para detectar e monitorar a progressão da doença, bem como para avaliar a resposta ao tratamento. Através da visualização das imagens de Raio-X pulmonares, os médicos podem identificar áreas afetadas, avaliar a extensão da infecção e planejar o tratamento adequado.

A imagens de COVID-19 em análise são provenientes de uma base pública vencedora do *COVID-19 Dataset Award da Kaggle Community*. A base é constituída de imagens de Raio-X pulmonar separadas em três categorias, imagens de **COVID-19**, imagens de **pneumonia** e **normais**, disponível em (RAHMAN, 2022). Foram retiradas um total de 3000 imagens separando 1000 imagens para cada categoria, todas as imagens estão no formato JPG, com resolução de 299x299 pixels. A Figura 22 apresenta imagens de exemplos para cada categorias da base de imagens de COVID-19.

Figura 22 – Exemplos de imagens pertencentes às três categorias da base de imagens de COVID-19



Fonte: Adaptada de Rahman (2022).

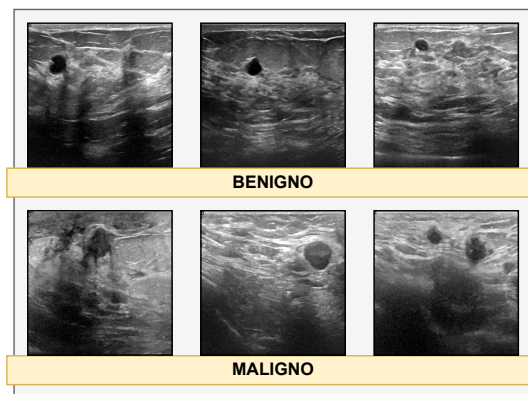
Mamografias

As imagens de ultrassom específicas da mama, são utilizadas para a detecção precoce e

diagnóstico de câncer de mama. Essas imagens podem revelar microcalcificações, massas ou distorções estruturais na mama, que podem indicar a presença de tumores malignos. A realização regular de mamografias é fundamental para a detecção precoce do câncer de mama, pois permite o diagnóstico em estágios iniciais, aumentando significativamente as chances de sucesso do tratamento.

A base de imagens de ultrassom de mama denominada BUSI-BREAST é uma base pública separada em duas classes, imagens de ultrassom de mamas **benignas** e **malignas**, de um total de 200 imagens separando 100 imagens para cada categoria, disponível em (AL-DHABYANI *et al.*, 2020). A Figura 24 apresenta imagens de exemplos para cada categorias da base de imagens de mamografia.

Figura 23 – Exemplos de imagens pertencentes as duas categorias da base de imagens de mamografia



Fonte: Adaptada de Al-Dhabyani *et al.* (2020).

Corel DB

A base Corel DB é uma base pública composto por 1000 imagens coloridas, que são divididas em 10 categorias diferentes, disponível em (WANG; LI; WIEDERHOLD, 2001). Cada categoria representa um tema ou conceito específico, como animais, praias, montanhas, monumentos, entre outros. As imagens desta base foi montada para serem representativas e diversas em termos de conteúdo visual. Neste projeto, esta base também foi incluída nas análises utilizando o extrator de histograma de cinzas, a fim de validar a generalização dos algoritmos presentes nesta pesquisa. A Figura 24 apresenta imagens de exemplos da base de imagens Corel DB.

5.3 Considerações Finais

Neste capítulo foram apresentados os conceitos básicos de imagens médicas muito utilizados como imagens de Raio-X e imagens de ultrassom. Dentro deste tema, foram abordadas as duas bases de imagens médicas e uma base generalizada utilizados para estudo de caso nesta

Figura 24 – Exemplos de imagens da base de imagens do Corel DB



Fonte: [Weihs et al. \(2014\)](#).

pesquisa de mestrado e os tipos de extratores de características mais comumente utilizados para análise.

MÉTODO FRF (FUSION RELEVANCE FEEDBACK)

6.1 Visão Geral

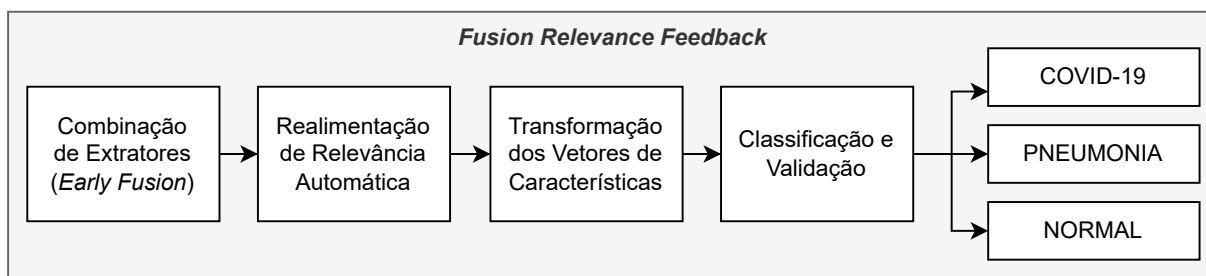
Nesta pesquisa é proposta a *Fusion Relevance Feedback* (FRF), que visa identificar qual descritor tem o melhor desempenho no domínio das imagens de Raio-X por meio da combinação de extratores de textura com o refinamento desses extratores através da ponderação de pesos utilizando RR automática. Isto é, o sistema faz a RR sem exigir esforço por parte do usuário, aprimorando o processo de modo transparente para o usuário.

O objetivo foi desenvolver uma abordagem que aumenta a acurácia de extratores para classificar imagens de Raio-X pulmonar visando o cenário da COVID-19, sem demandar esforço adicional do usuário. Por meio de testes exaustivos de combinação de extratores e aprendizado de pesos com RR, pode-se obter novos extratores que apresentam resultados ótimos para classificação de dados.

O principal objetivo desta abordagem é de realizar RR já no pré-processamento do conjunto de dados, eliminando a necessidade de interação constante com os usuários. Isso permite que o sistema ajuste seus parâmetros e estratégias internas para maximizar a qualidade dos resultados sem requerer um esforço significativo por parte do usuário final. Isso também significa que a abordagem FRF tem a vantagem de adaptar-se a diferentes conjuntos de dados e cenários de aplicação.

A abordagem FRF permite a geração de novos conjuntos de descritores otimizados, resultando em extratores que demonstraram um desempenho significativamente aprimorado na classificação de dados de Raio-X pulmonar relacionados à COVID-19. O diagrama da metodologia deste método é apresentado na [Figura 25](#).

Figura 25 – Diagrama da abordagem *Fusion Relevance Feedback*, na qual é utilizado combinações de extratores de características em conjunto de realimentação de relevância para aprimorar o espaço de características do conjunto no pré-processamento



Fonte: Elaborada pelo autor.

6.2 Combinação de Extratores de Características *Early Fusion*

A combinação de extratores de características é uma abordagem amplamente adotada em aplicações de recuperação de imagens, frequentemente reconhecida como uma técnica de *Early Fusion*. Essa estratégia é implementada antes do cálculo das distâncias entre os elementos em um conjunto de dados e envolve a fusão de diferentes vetores de características para criar uma nova representação do dado complexo. O propósito dessa abordagem é potencializar a compreensão das características presentes nas imagens e, conseqüentemente, melhorar a capacidade do sistema de recuperação em reconhecer padrões relevantes.

A realização da combinação de extratores pode ser efetuada por meio de diferentes operações entre os vetores, como adição ou multiplicação dos elementos. Uma das formas mais sólidas e reconhecidas de *Early Fusion* é a concatenação dos vetores de características em um único vetor. No contexto deste estudo, optou-se pela concatenação dos vetores como o método de combinação. Esse processo envolve a formação de grupos compostos por dois e três extratores de textura, combinados sem repetição. Essa abordagem resultou em um total de 120 novas combinações únicas que foram submetidas a análise.

A escolha pela concatenação como técnica de combinação é baseada por sua capacidade de manter a informação original dos extratores de textura. Ao unir os vetores de características, a concatenação resulta em um vetor mais longo, incorporando as contribuições individuais de cada extrator. Esse vetor ampliado reflete uma representação mais abrangente das características presentes nas imagens, permitindo que o sistema de recuperação explore uma gama mais ampla de informações para serem refinados e selecionados durante a análise.

A formação de grupos de dois e três extratores, sem repetição, resulta em uma variedade significativa de combinações. Isso é importante, pois permite examinar várias configurações possíveis para determinar quais combinações de extratores de textura podem oferecer um

desempenho otimizado na classificação de imagens de Raio-X pulmonar relacionadas à COVID-19. Essa abordagem sistemática possibilita uma análise abrangente das diferentes maneiras pelas quais as características podem ser combinadas, identificando aquelas que são mais eficazes na discriminação entre diferentes classes de imagens.

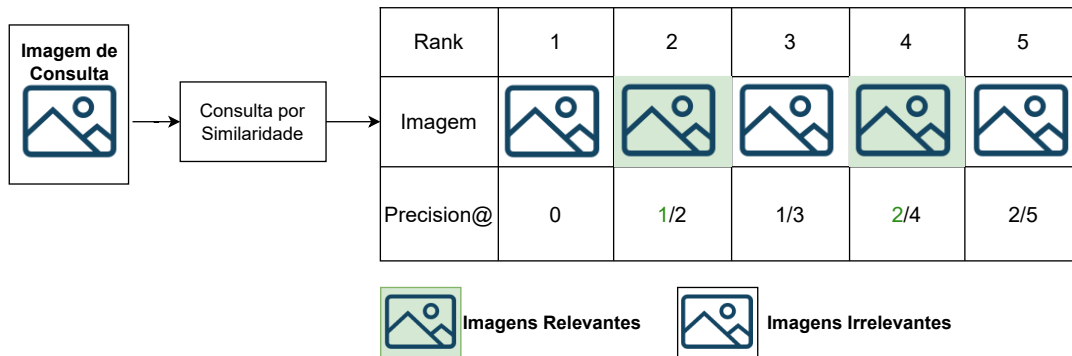
6.3 Aprendizado de Pesos Automático

A base das técnicas de RR reside na ideia de que, a partir de um conjunto de amostras coletadas a partir de uma consulta de similaridade, o sistema é capaz de identificar padrões de características relevantes nas imagens. Consequentemente, à medida que o sistema continua a receber novas consultas, ele pode refinar seus resultados gradativamente, gerando melhorias sucessivas. Uma das abordagens que incorpora as técnicas de RR é a ponderação de pesos. Nesse método, as dimensões do vetor de características são ponderadas com base no inverso do desvio padrão de cada dimensão das imagens rotuladas como relevantes. A utilização de medidas estatísticas como o desvio padrão e a variância inversa permite avaliar a dispersão dos elementos e aprender pesos com base nas amostras de imagens disponíveis para o treinamento.

O trabalho de (RUI *et al.*, 1998) demonstrou que o uso de amostras relevantes para a RR resulta em um aumento significativo na precisão na primeira iteração em comparação com outras abordagens. No entanto, a melhoria na precisão tende a estabilizar após poucas iterações. Nesse contexto, a primeira iteração é crucial para aprimorar os extratores de características de forma automática, com o objetivo de evitar a propagação de erros ao longo das iterações subsequentes.

O aprendizado automático dos pesos é realizado por meio da aplicação do método de avaliação *Mean Average Precision*. Esse método de avaliação simula consultas de similaridade, permitindo a análise e captura da quantidade de imagens relevantes que são recuperadas em consultas de similaridade durante a etapa de pré-processamento. Nesta etapa, as imagens que são consideradas relevantes são utilizadas como *feedback* para a técnica de RR gerar os vetores de pesos. O método FRF atribui um valor a cada imagem para rotular suas classes como COVID, pneumonia ou normal. Posteriormente, é realizada uma seleção exaustiva aleatória de uma das classes, que é utilizada como uma amostra para a RR. Esse processo é repetido várias vezes, selecionando diferentes conjuntos de imagens da mesma classe, a fim de encontrar a amostra que proporciona resultados significativos na RR. A [Figura 26](#) apresenta a seleção de imagens relevantes utilizando a avaliação MAP. As imagens relevantes representam o conjunto de *feedback* para técnica de RR gerar o vetor de pesos para transformar o espaço de características referente a consulta por similaridade sendo realizada.

Figura 26 – Seleção de amostras no pré-processamento do conjunto de dados, onde a posição e a correspondência de classe dos resultados são avaliadas em relação a imagem de consulta



Fonte: Elaborada pelo autor.

Para cada imagem selecionada, é calculado o inverso do desvio padrão ao longo do eixo do vetor de características, gerando assim um vetor de pesos normalizado. A abordagem FRF aprende um peso para cada dimensão no espaço da função L_p , refinando assim a medida de similaridade. O vetor de pesos resultante é então utilizado para transformar os vetores de características, realizando a multiplicação da matriz de características. Isso tem o efeito de enfatizar os atributos que são considerados relevantes de acordo com os pesos calculados. Em essência, essa etapa visa realçar as características que têm uma maior contribuição na distinção entre as diferentes classes de imagens, aumentando assim a sensibilidade do sistema na detecção de padrões específicos.

6.4 Experimento

Além dos quatro extratores de texturas abordados na [Seção 2.3](#), foram incluídos mais cinco extratores para gerar mais combinações e análise, totalizando um total de nove extratores de texturas para analisar a base de dados de Raio-X COVID-19. Para o experimento, foram selecionadas 750 amostras de forma aleatória da base de COVID-19, sendo 250 imagens da classe COVID, 250 imagens da classe de pneumonia e 250 imagens da classe normal. De forma resumida a [Tabela 4](#) apresenta os nove extratores utilizados no experimento.

Tabela 4 – Extratores de texturas para análise do método FRF

Extrator	Dimensão
<i>First Order Statistics</i> (FOS)	16
<i>Gray Level Co-occurrence Matrix</i> (GLCM)	28
<i>Gray Level Difference Statistics</i> (GLDS)	5
<i>Neighborhood Gray Tone Difference Matrix</i> (NGTDM)	5
<i>Statistical Feature Matrix</i> (SFM)	4
<i>Law's Texture Energy Measures</i> (LTE)	6
<i>Fourier Power Spectrum</i> (FPS)	2
<i>Local Binary Pattern</i> (LBP)	59
<i>Local Phase Quantization</i> (LPQ)	255

A performance da FRF foi avaliada a partir da realização das seguintes etapas:

- Seleção dos melhores descritores utilizando a classificação por *K Nearest Neighbor* (KNN) utilizando a função de distância ponderada euclidiana com validação cruzada de até 10 vezes.
- Análise do ganho semântico nos extratores de texturas abordados, utilizando a técnica **RR Automática** em relação a precisão do conjunto original.

6.5 Resultados e Discussões

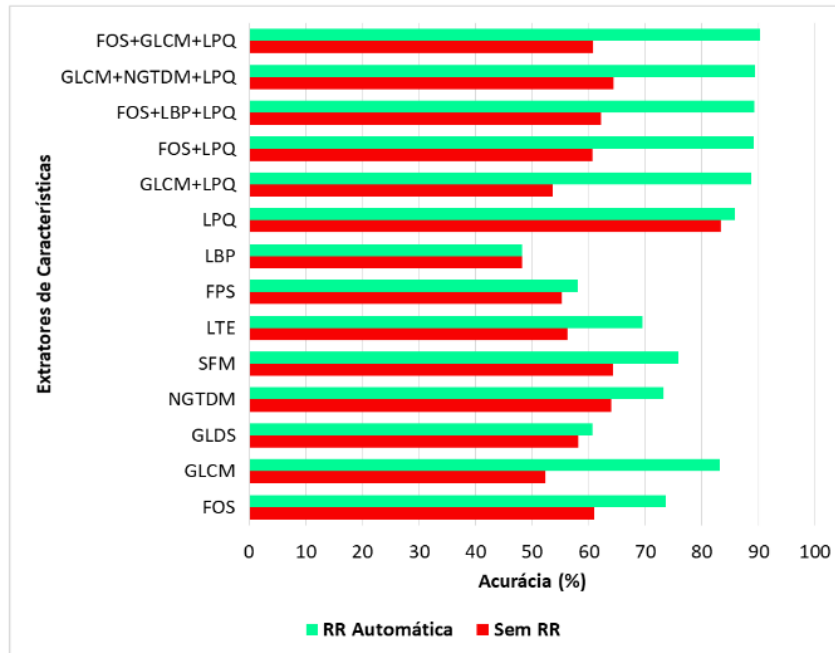
Após a exploração de todas as possíveis combinações entre grupos de dois e três extratores, um total de 120 combinações únicas é obtido. O sistema, então, realiza uma seleção automática de um subconjunto a partir da base de imagens para treinamento dos pesos, utilizando a técnica de RR que se baseia no inverso do desvio padrão.

A [Tabela 5](#) apresenta os resultados de acurácia (ACC) e F1-Score (F1) alcançados pela classificação utilizando cada extrator original individualmente, assim como as melhores combinações obtidas após a aplicação da técnica de RR Automática. A tabela faz uma comparação entre as abordagens com e sem o uso de RR na base de imagens. A análise dos resultados revela que a aplicação da técnica de RR automática resulta em ganhos consideráveis em termos de precisão na maioria das combinações. Este mesmo resultado pode ser observado na visualização gráfica na [Figura 27](#). Um exemplo destacado é a combinação **FOS+GLCM+LPQ**, que alcança uma acurácia de 90,27%. Isso evidencia que a utilização da técnica de RR automática é eficaz em melhorar significativamente o desempenho da classificação em comparação com as abordagens que não incorporam a RR.

Tabela 5 – Comparação da ACC e F1 por descritor de textura, Sem RR e com RR Automática.

Extrator	Sem RR		RR Automática		
	ACC (%)	F1 (%)	ACC (%)	F1 (%)	Ganho (%)
FOS	60,93	60,59	73,60	73,58	20,79
GLCM	52,40	51,68	83,20	83,13	58,77
GLDS	58,13	57,68	60,66	60,43	4,35
NGTDM	64,00	63,82	73,20	73,00	14,37
SFM	61,19	61,06	75,86	75,76	23,97
LTE	56,26	55,77	69,46	69,22	23,46
FPS	55,33	55,19	58,00	57,46	4,82
LBP	48,26	47,63	48,26	47,70	0
LPQ	83,33	83,34	85,86	85,95	3,03
GLCM+LPQ	53,60	52,92	88,67	88,66	65,42
FOS+LPQ	60,67	60,33	89,07	89,12	46,81
FOS+LBP+LPQ	62,13	61,85	89,20	89,27	43,56
GLCM+NGTDM+LPQ	64,40	64,24	89,33	89,35	38,74
FOS+GLCM+LPQ	60,80	60,23	90,27	90,32	48,47

Figura 27 – Gráfico de barras que compara as acurácias por descritor de textura, Sem RR e com RR Automática.



Fonte: Elaborada pelo autor.

6.6 Considerações Finais

Nesta capítulo, foi apresentado o método *Fusion Relevance Feedback* (FRF) que aprimora o espaço de características utilizando realimentação de relevância automática sobre combinações de extratores de características. Os resultados destacam a capacidade da técnica de RR em identificar padrões relevantes e ponderar os pesos das características dos extratores de forma a maximizar a acurácia da classificação. Essa abordagem automatizada de ajuste de pesos contribui para uma melhoria consistente dos resultados de classificação, demonstrando sua eficácia em ambientes complexos de processamento de imagens.

MÉTODO TMRF (TUNING METRICS RELEVANCE FEEDBACK)

7.1 Visão Geral

Os MAMs envolvem a definição de uma métrica de similaridade para a construção da indexação. Essa métrica é utilizada para determinar quais elementos estão mais próximos de um determinado elemento de consulta. No entanto, um MAM não consegue indexar os elementos utilizando duas ou mais funções de distância em uma mesma indexação.

A utilização desses métodos de indexação e recuperação otimizada de dados complexos é essencial para sistemas de Recuperação Baseada em Conteúdo (CBIR). Para isso, é necessário ser capaz de modificar e aprimorar a semântica do espaço métrico indexado. Portanto, soluções que combinam o uso de funções de distâncias ponderadas em MAMs com técnicas de Realimentação de Relevância (RR) que aprendem pesos são fundamentais para obter consultas por similaridade mais eficazes, ao mesmo tempo em que otimizam o tempo de consulta dos MAMs em comparação com modelos de recuperação por matriz de distâncias.

Nesta pesquisa, foi desenvolvido o método *Tuning Metrics Relevance Feedback* (TMRF), que tem como objetivo incluir o processo de RR denominado RR Desvio Padrão. Esse processo de RR aprende vetores de pesos que são capazes de ponderar métricas de similaridade diretamente no MAM *Slim-Tree*. Nas abordagens convencionais, o processo de RR acaba ponderando toda a matriz de características, tanto do *feedback* informado pelo usuário quanto da matriz de características da base de dados, o que gera um gargalo no tempo de processamento das operações de consulta em dados complexos.

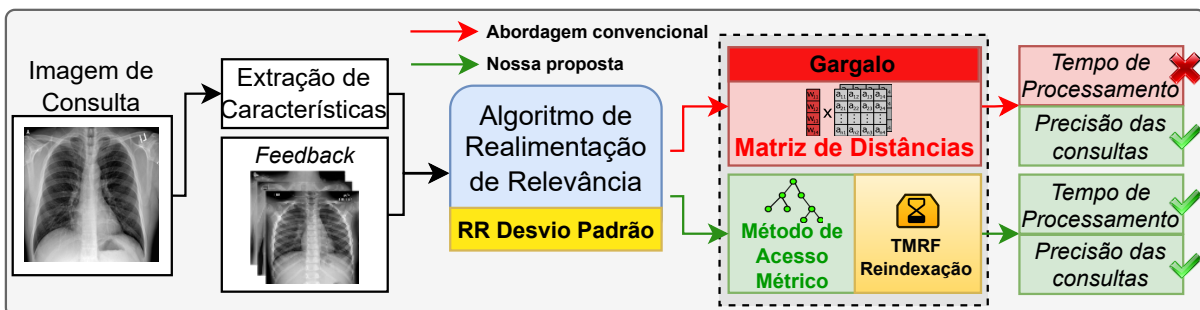
Nosso objetivo foi desenvolver uma metodologia que inclua funções de distâncias ponderadas na *Slim-Tree*, permitindo que a métrica de similaridade seja modificada dinamicamente durante os ciclos de RR realizados pelos usuários. A ideia é aproveitar a eficiência de tempo

de processamento de recuperação dessa estrutura e o aprimoramento do espaço métrico através do RR Desvio Padrão. Isso é importante porque o MAM acaba usando funções de distâncias fixas para indexar os elementos, o que pode limitar a precisão dos resultados. A metodologia foi modularizada a fim de realizar os seguintes procedimentos.

- Otimização do algoritmo de realimentação de relevância para atuar sob a matriz de *feedback*.
- Seleção do vetor de pesos mais relevante para reindexação.
- Inclusão de funções de distâncias ponderadas na *Slim-Tree*.
- Especificar a necessidade de reindexar a *Slim-Tree*.

A Figura 28 demonstra um diagrama comparando o processo de aprendizado do RR para retornar uma consulta por similaridade, mostrando o gargalo que ocorre nas operações de matrizes de distância nas abordagens convencionais, e o método TMRF que integra o MAM para lidar com o processamento de forma mais eficaz.

Figura 28 – Diagrama da abordagem *Tuning Metrics Relevance Feedback*, na qual destaca a eficácia da abordagem proposta por utilizar métodos de acesso métrico e realimentação de relevância em relação as abordagens convencionais que utilizam matrizes de distâncias



Fonte: Elaborada pelo autor.

7.2 Realimentação de Relevância por Desvio Padrão

Para realizar as análises neste artigo, adotamos o algoritmo proposto por (RUI *et al.*, 1998), conforme detalhado na Capítulo 4. No entanto, efetuamos uma modificação na forma de cálculo original, visando melhorar a eficiência do processo.

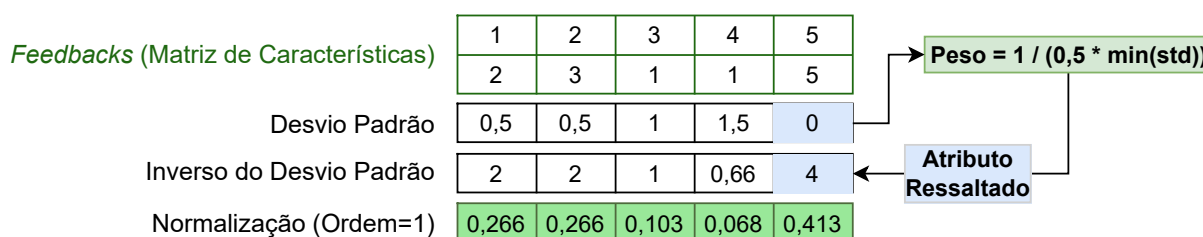
No algoritmo original, eram utilizados três vetores de pesos pré-definidos para distribuir os pesos de cada camada do algoritmo. Cada um desses vetores era responsável por destacar diferentes atributos na matriz de características. No entanto, essa abordagem resultava em múltiplos cálculos na matriz, o que tornava o tempo de processamento ineficiente, especialmente em conjuntos de dados extensos.

No nosso algoritmo revisado, denominado **RR Desvio Padrão**, propomos uma abordagem mais eficiente. Em vez de utilizar três vetores de pesos, calculamos uma única matriz de pesos com base na variância da matriz de características das imagens relevantes. Esse novo método reduz consideravelmente a quantidade de cálculos necessários. O procedimento do aprendizado de pesos da RR Desvio Padrão é descrito pelo passos a seguir:

- A matriz de características é gerada através do *feedback* de entrada que é considerado como elementos relevantes.
- Para cada coluna da matriz de características é calculado o seu desvio padrão. O resultado será um vetor de desvio padrão de mesma dimensionalidade do vetor de características do conjunto de dados.
- Para cada dimensão do vetor do desvio padrão é calculado o inverso do desvio padrão. Se durante o cálculo, o valor do desvio padrão for igual 0, é selecionado o atributo com o menor valor no vetor de desvio padrão e é aplicado um peso pré-definido (*weight*) para ressaltar essa dimensão específica, onde $weight = 1 / (0.5 * \min(std))$. Isso nos permite avaliar a importância de cada atributo em relação aos demais.
- Por fim, é realizada a etapa de normalização sobre o vetor de pesos obtido. Isso é feito para garantir que a soma de todos os pesos seja igual a 1 ($\sum wi = 1$). Essa normalização é importante para que cada peso contribua proporcionalmente para a similaridade final entre as imagens e as consultas.

Ao adotar o algoritmo RR Desvio Padrão, buscamos obter uma técnica mais eficaz e eficiente para calcular os pesos dos atributos. Essa abordagem permite uma melhoria significativa no tempo de processamento em comparação com o algoritmo original, sem comprometer a qualidade dos resultados. A [Figura 29](#) ilustra um exemplo do processo para obter um vetor de pesos a partir da matriz de características com dois *feedbacks* que representam imagens relevantes. Ao final é obtido o vetor de pesos deste ciclo de realimentação.

Figura 29 – Processo de geração do vetor de pesos sobre a matriz de características de imagens relevantes



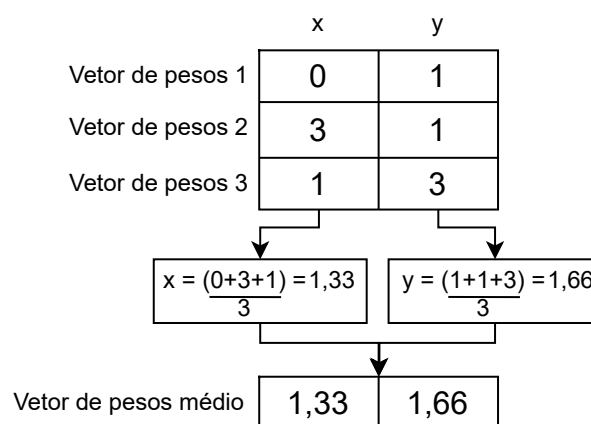
Fonte: Elaborada pelo autor.

7.3 Histórico de Consultas e Seleção do Vetor de Pesos

A escolha do vetor de pesos para refinar a medida de similaridade é essencial para aprimorar o espaço métrico de forma que dê as devidas relevâncias para cada categoria ou classe do conjunto de dados a ser trabalhado. Vale ressaltar que a realização do ciclo de realimentação têm foco em aprimorar a semântica baseada na classe da imagem de consulta em que está sendo processada. Isto significa que em um conjunto de vetores de pesos passados por ciclos de RR, cada vetor de pesos teve um aprendizado ressaltando atributos diferentes referente a suas consultas realizadas. Desta forma, definir qual vetor de pesos deve ser utilizado a médio e longo prazo para reestruturar um método de acesso métrico é um processo determinístico.

A metodologia desenvolvida para selecionar o vetor de pesos utiliza o histórico de consultas por similaridade que foram ativadas por RR, na qual foram armazenados os vetores de pesos obtidos. Utiliza-se uma janela deslizante sobre os vetores de pesos para realizar cálculos de vetores de pesos médios, que representam uma média geral de ponderação das características. Esta média geral têm como objetivo dar a mesma prioridade de relevância para todas as classes de um conjunto de dados. O exemplo ilustrado na [Figura 30](#) consiste no cálculo do vetor de pesos médio a partir de três vetores de pesos de duas dimensões. Para obter o vetor de pesos médio, somamos as componentes correspondentes de cada vetor de pesos e dividimos pelo número total de vetores. No exemplo apresentado, o vetor de pesos médio representa um ponto médio que equilibra as direções e magnitudes dos três vetores de pesos originais.

Figura 30 – Exemplo do cálculo do vetor de pesos médio que reflete a relevância de uma janela de vetores de pesos do histórico de consultas que armazenaram vetores de pesos



Fonte: Elaborada pelo autor.

A [Figura 31](#) demonstra o exemplo plotado do ponto médio que representa o vetor de pesos médio. É possível observar que o vetor de pesos médio fica no centro entre os vetores de pesos, representando a média geral.

Figura 31 – Exemplo do cálculo do vetor de pesos médio que reflete os pesos médios das dimensões de um conjunto de vetores de pesos



Fonte: Elaborada pelo autor.

7.4 Inclusão de Funções de Distância Ponderadas na *Slim-Tree*

Como apresentado na [Subseção 3.4.1](#) sobre a *Slim-Tree*, essa estrutura de MAM é amplamente reconhecida por sua eficiência na realização de consultas por similaridade em conjuntos de dados complexos. No entanto, a abordagem padrão de indexação utilizando funções de distâncias fixas pode limitar a capacidade da estrutura de fornecer respostas semânticas precisas para as consultas. O método TMRF demonstra que é possível aprimorar a eficácia semântica da estrutura ao incluir funções de distâncias ponderadas.

No contexto da indexação de dados complexos, a ponderação de funções de distância desempenha um papel crucial na atribuição de pesos diferentes aos atributos, de acordo com sua relevância para a tarefa de consulta. Essa ponderação permite que certos atributos sejam enfatizados ou atenuados durante o cálculo da similaridade entre os elementos, resultando em respostas mais coerentes e semanticamente ricas. No entanto, é importante reconhecer que a inclusão da ponderação de funções de distância adiciona uma carga computacional adicional devido à necessidade de recalculá-las com base nos pesos atribuídos. Embora esse recálculo possa introduzir um pequeno aumento no tempo de processamento, o ganho semântico geralmente compensa esse custo adicional.

Para incluir funções de distância ponderadas na *Slim-Tree*, é necessário modificar o cálculo da distância entre dois elementos utilizando as funções ponderadas L_p conforme descrito na [Seção 3.3](#). Além disso, a *Slim-Tree* precisa ser adaptada para permitir que os usuários forneçam um vetor de pesos para cada atributo de forma dinâmica. Para isso, durante a construção dos índices da *Slim-Tree*, é definida uma métrica de similaridade ponderada, onde todos os valores do vetor de pesos são inicialmente iguais a um. A inclusão do vetor de pesos é feita adicionando

um novo parâmetro ao método de consulta da *Slim-Tree*, permitindo que o usuário especifique os pesos para cada atributo, os quais são aprendidos por meio da abordagem **RR Desvio Padrão**, conforme detalhado na [Seção 7.2](#). Durante o cálculo de distâncias na *Slim-Tree*, o vetor de pesos deve multiplicar cada dimensão do espaço pela sua respectiva ponderação antes de calcular a distância. Esse procedimento é aplicado nos ciclos de RR até que ocorra uma reindexação da estrutura.

O [Algoritmo 1](#) apresenta o pseudocódigo principal para calcular a distância entre dois elementos levando em consideração a ponderação de pesos utilizando a distância Euclidiana como exemplo. Esse algoritmo mostra como os pesos são aplicados a cada atributo e como a distância é recalculada com base na ponderação definida. Dessa forma, após o aprendizado dos pesos por meio da abordagem RR Desvio Padrão, o vetor de ponderação resultante pode ser fornecido como parâmetro para realizar os cálculos de distâncias ponderadas.

No caso em que nenhum vetor de pesos é fornecido como parâmetro, todos os pesos são inicializados com o valor padrão igual a um. Isso permite que a *Slim-Tree* atue como uma função de distância padrão, sem a inclusão de ponderação, caso seja desejado.

Algoritmo 1 – Calibração da distância Euclidiana entre dois objetos na *Slim-Tree*

Requer: $Obj1, Obj2$

Assegure: \sqrt{D}

$W \leftarrow GetWeights()$

▷ Captura o último vetor de pesos aprendido

função GETDISTANCE($Obj1, Obj2$)

se W is empty **então**

 ▷ Se vetor de pesos for vazio, é uma distância padrão

para each integer i in $Obj1.size()$ **faça**

$W_i \leftarrow 1$

fim para

fim se

para each integer i in $Obj1.size()$ **faça**

$Tmp \leftarrow (Obj1.GetFeatures()[i] - Obj2.GetFeatures()[i])$

$D \leftarrow D + Tmp^2 * W_i$

 ▷ Calibração das distâncias

fim para

fim função

7.5 Reindexação da *Slim-Tree*

A metodologia proposta para determinar o momento ideal de reindexar um MAM após o aprimoramento da medida de similaridade ao longo dos ciclos de realimentação de relevância é baseada em análises do histórico de consultas por similaridade ativadas pela RR. Essa abordagem permite identificar o momento em que uma reindexação se faz necessária para obter um espaço métrico transformado mais adequado.

O MAM *Slim-Tree* é implementado em arquivos de disco. Dessa forma, mais de uma árvore pode ser criada e armazenada em disco e, com essa flexibilidade, a árvore armazenada no arquivo pode ser constantemente atualizada. Se uma árvore para um determinado extrator

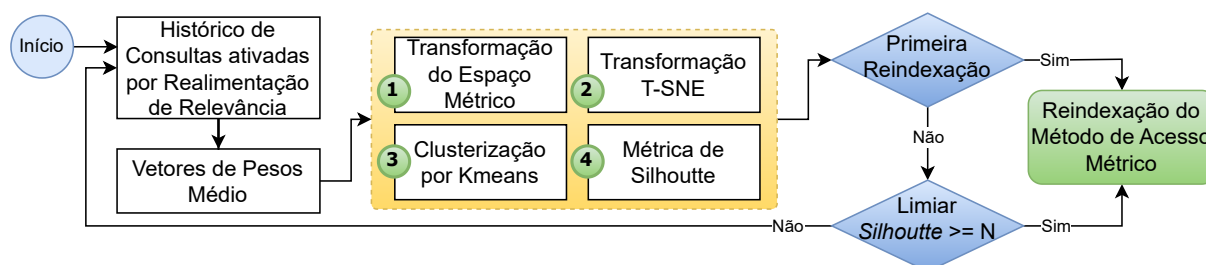
contém características extraídas de toda a base de imagem, esta árvore pode ser consultada por qualquer usuário sem construir uma nova árvore. Portanto, este módulo de geração de árvore de índice de características permite atualizar a árvore ou criar uma nova árvore contendo todas as características da base de imagens. Este tipo de implementação nos leva a um problema a ser respondido: *Quando e de que forma é o melhor momento para reindexar um MAM depois do usuário aprimorar a medida de similaridade ao longo de ciclos de realimentação de relevância?*

O primeiro passo consiste em utilizar o histórico e a seleção do vetor de pesos médio, pois representa o vetor de pesos mais relevante para reindexar o MAM [Seção 7.3](#). Com o objetivo de obter um espaço métrico transformado e reduzir a dimensionalidade, é aplicado o algoritmo t-SNE (*t-Distributed Stochastic Neighbor Embedding*) sobre o espaço métrico transformado pelo vetor de pesos médio. A transformação t-SNE permite realizar análise interpretativa do espaço métrico gerado.

Após a redução de dimensionalidade com t-SNE, é realizado o agrupamento dos dados utilizando o algoritmo *k-means*. O agrupamento resultante do *k-means* é útil para avaliar a qualidade da representação do espaço métrico transformado. Com base no coeficiente de *Silhouette*, a primeira reindexação no MAM ocorre utilizando o espaço métrico ponderado que apresenta o maior coeficiente de *Silhouette*. À medida que o histórico de consultas no sistema aumenta, novas reindexações são realizadas somente quando o coeficiente de *Silhouette* calculado para o espaço métrico mais recente supera o valor da última reindexação.

Essa metodologia permite determinar o momento adequado para reindexar o MAM, garantindo que a estrutura esteja atualizada com o espaço métrico transformado mais eficiente e com melhor qualidade de agrupamento dos elementos. A [Figura 32](#) ilustra o diagrama dessa metodologia, apresentando as etapas desde a coleta dos vetores de pesos até a realização das reindexações com base no coeficiente de *Silhouette*.

Figura 32 – Processo para realização da reindexação do MAM, onde o conjunto de dados ponderados pelo vetor de pesos médio é passado por transformações do espaço de características e clusterização, até ser avaliado pela métrica de *Silhouette* para validar a reindexação



Fonte: Elaborada pelo autor.

Essa abordagem proporciona uma forma inteligente e automatizada de reindexação do MAM, garantindo que a longo prazo, a estrutura mantenha sua eficiência de recuperação, ao mesmo tempo em que se beneficia do aprimoramento da precisão das consultas por similaridade

aprendidas pela RR Desvio Padrão.

7.6 Experimentos

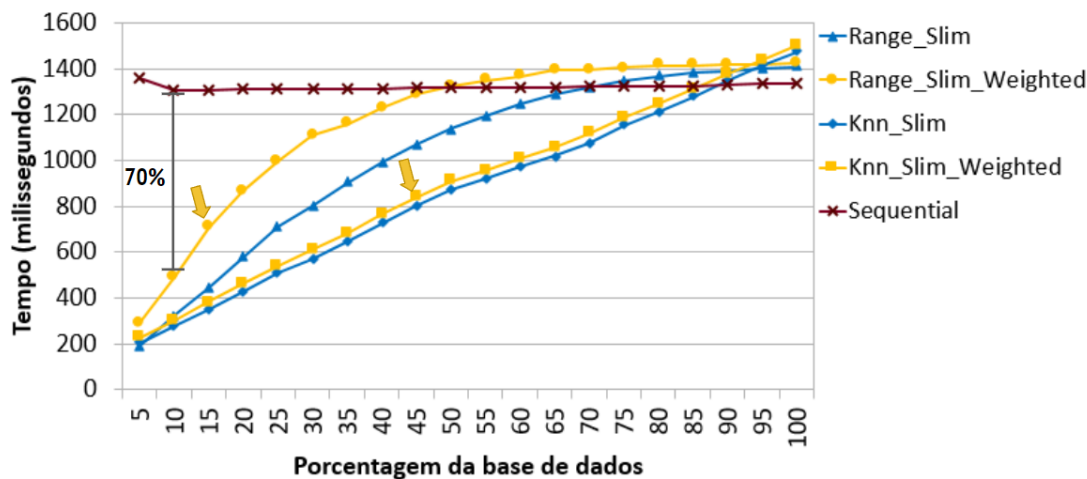
A performance da TMRF foi avaliada a partir da realização das seguintes etapas:

1. Análise do ganho semântico nos extratores de texturas abordados, utilizando a técnica **RR Desvio Padrão** em relação a precisão do conjunto original.
2. Análise qualitativa do ganho semântico da técnica **RR Desvio Padrão** através de visualização de dados.
3. Análise do aumento do custo computacional a curto prazo devido a inclusão de funções de distância ponderadas na *Slim-Tree*.
4. Comparação do custo computacional de tempo de recuperação da base de imagem entre o MAM com função de distância ponderada e abordagem sequencial.
5. Exemplo de análise do determinismo da reindexação do MAM através do coeficiente e gráfico de *Silhouette*.

7.7 Resultados e Discussões

A implementação de uma função de distância ponderada em um MAM visa ampliar a capacidade semântica da estrutura na recuperação de dados. Verificamos o tempo gasto pelas consultas por similaridade até retornar 100% da base de dados COVID-19 a qual possui a maior quantidade de elementos, comparando o tempo de consulta do MAM com o de busca sequencial, que é a abordagem passível de comparação com outros métodos. É possível observar que mesmo com o aumento de cálculos de distâncias e acessos a disco pelos métodos de buscas utilizando funções de distâncias ponderadas como a *Range_Slim_Weighted* e *kNN_Slim_Weighted*, o tempo gasto com o uso do MAM ainda é muito menor em relação ao sequencial, sendo 70% mais rápido recuperando 10% de toda a base, como apresentado na [Figura 33](#). Vale ressaltar que quando o MAM é reindexado, o tempo gasto nas consultas e o número de cálculos de distância e acessos a disco se equiparam aos da *Range_Slim* e *kNN_Slim* por não realizarem a reconstrução da árvore métrica.

Figura 33 – Tempo de consulta em milissegundos recuperando em incrementos de 5% da base de dados, até atingir 100% da base de imagens de COVID-19, comparando o desempenho das consultas normais, consultas ponderadas e de recuperação sequencial ao longo do processo de recuperação



Tempo de consulta em milissegundos sobre a base de COVID-19.

Fonte: Elaborada pelo autor.

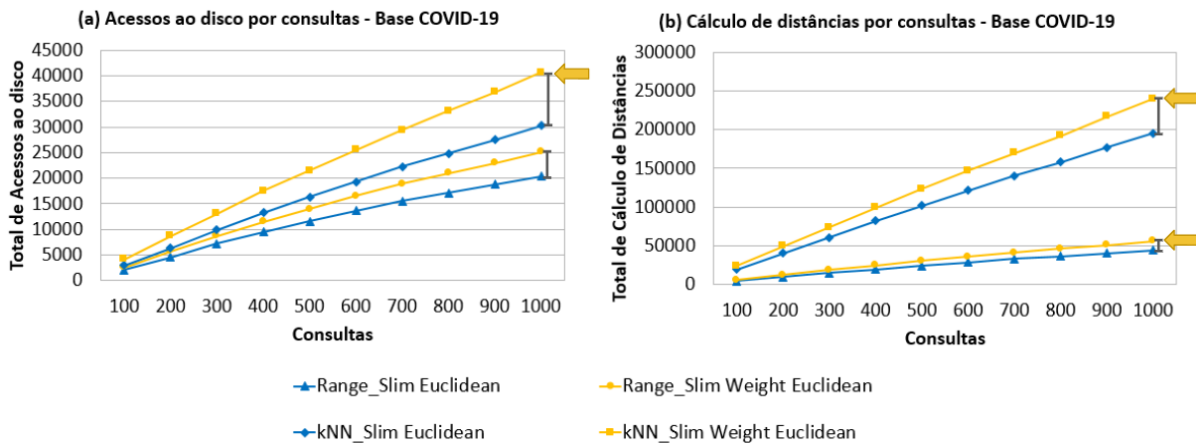
A RR visa aprimorar o potencial semântico das consultas por similaridade. Na [Tabela 6](#) comparamos o efeito de ciclos de RR de até 5 iterações sobre consultas nas bases de dados, cujo objetivo é demonstrar como a RR que aprende um vetor de pesos consegue reorganizar os elementos melhorando significativamente o *rank* de uma consulta. Para isto são calculados o *Average Precision* do ciclo de realimentação e *Mean Average Precision* sobre as *top 10* e *20* imagens nas bases de dados. É possível observar que a ponderação da função de distância melhora significativamente a precisão das consultas por similaridade nos diferentes descritores usados nos experimentos, obtendo um ganho de até 42% nas *top 20* imagens do extrator GLCM.

Tabela 6 – *Mean Average Precision* sobre os *top 10* e *20* imagens, de cada extrator sem RR e com RR Desvio Padrão em diferentes bases de imagens.

Dataset	Extrator	Sem RR		RR Desvio Padrão		Ganho (%)	
		P@10	P@20	P@10	P@20	P@10	P@20
COREL-1000	Gray Hist	61,10	56,13	79,65	71,80	30	28
	Gray Hist	81,57	78,26	91,21	86,88	12	11
COVID-19	GLCM	63,36	58,18	84,96	82,70	34	42
	FOS	70,10	65,37	81,95	76,50	17	17
	SFM	71,96	68,72	83,10	80,62	15	17
	LTE	64,96	60,78	75,07	70,39	16	16
	Gray Hist	71,68	65,57	83,07	72,00	16	10
BUSI-BREAST	GLCM	61,82	56,88	74,28	66,20	20	16
	FOS	64,20	59,55	73,12	64,8	14	9
	SFM	67,74	63,23	76,41	70,58	13	12
	LTE	63	58,86	76,17	70,23	21	19
	Gray Hist	63	58,86	76,17	70,23	21	19

A Figura 34 apresenta a comparação da quantidade de acessos a disco (Figura 34 (a)) e cálculos de distâncias (Figura 34 (b)) da função de distância ponderada em relação à função de distância padrão no MAM, na base COVID-19. Apesar do aumento do custo computacional das consultas **Range_Slim Weight Euclidean** e **kNN_Slim Weight Euclidean**, a acurácia e precisão das consultas por similaridade é significativamente aumentada pela calibração das distâncias utilizando a técnica RR, na qual ao longo de comparação da métrica *Silhouette* é realizada a reindexação do MAM, atingido a mesma quantidade de números de cálculos e acesso a disco do **Range_Slim Euclidean** e **kNN_Slim Euclidean**. Para a análise de acurácia na recuperação e tempo computacional foi utilizado um computador com o sistema operacional Ubuntu 24.04, memória de 8G RAM e processador (8 CPUs) 2.0GHz, a as análises foram executadas em memória secundária SDD.

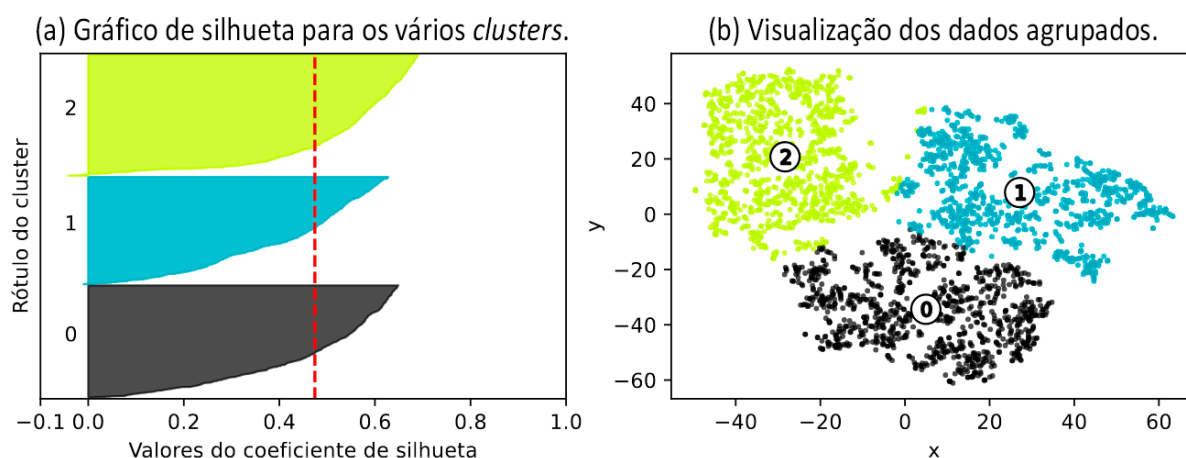
Figura 34 – Quantidade de cálculos de distâncias e acesso ao disco entre função de distância com e sem ponderação.



Fonte: Elaborada pelo autor.

A Figura 35 apresenta a análise de Silhueta (*Silhouette*) sobre um vetor de pesos aprendido pelo nosso método *RR Desvio Padrão* para determinar a reindexação do MAM. A Silhueta fornece uma maneira de avaliar parâmetros como o número de *clusters* encontrados. O gráfico de Silhueta apresentado na (Figura 35 (a)) mostra que o espaço métrico transformado é uma escolha boa para os dados fornecidos, devido à presença de *clusters* com pontuações de Silhueta acima da média e também devido a pequenas flutuações no tamanho do gráfico de Silhueta com uma média do coeficiente de Silhueta igual a 0.46. Na (Figura 35 (b)) é observado como a clusterização conseguiu realizar a separabilidade das classes baseado na coesão do elementos no *cluster*, que é utilizado para avaliar o coeficiente de Silhueta.

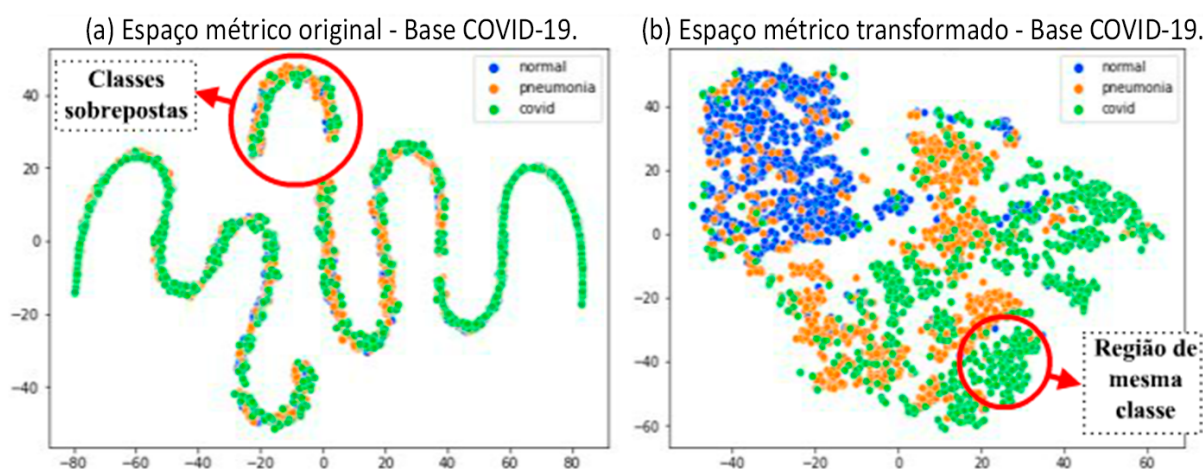
Figura 35 – Exemplo de análise para validação da reindexação do MAM, onde a linha pontilhada em (a) é o valor médio das *Silhouettes* calculadas para cada *cluster* (b), indicando a medida geral da qualidade da clusterização.



Fonte: Elaborada pelo autor.

A Figura 36 demonstra visualmente, por t-SNE, a melhoria semântica e de acurácia do espaço métrico original (Figura 36 (a)), utilizando a técnica de *RR com desvio padrão* do extrator de textura GLCM sobre a base de imagens de COVID-19. A abordagem desenvolvida juntamente com os algoritmos desta pesquisa estão disponíveis no Github¹. Na (Figura 36 (b)), é possível observar a melhoria visual em regiões com elementos de mesma categoria, resultado do ajuste do espaço métrico com a técnica proposta.

Figura 36 – Visualização do espaço métrico original (a) e do espaço métrico modificado (b) pela técnica RR Desvio Padrão após 5 ciclos de realimentação, indicando a melhoria semântica



Fonte: Elaborada pelo autor.

¹ Resultados complementares, disponíveis em: <https://github.com/renatomarcacini/Tuning-Metrics-Relevance-Feedback>

7.8 Considerações Finais

Neste capítulo, foi apresentado o método principal da pesquisa de mestrado denominada *Tuning Metrics Relevance Feedback* (TMRF). Foi demonstrada a metodologia empregada para se obter o aprimoramento do espaço métrico e formas de manter a estabilidade da estrutura definindo quando reindexar o método de acesso métrico. Os resultados destacam a capacidade da técnica de RR Desvio Padrão em identificar padrões relevantes e selecionar vetores de pesos estáveis para transformar o espaço de características de forma a maximizar a precisão das consultas por similaridade. Além disso, essa abordagem contribui para manter a eficiência dos métodos de acesso métrico superiores as abordagens convencionais de recuperação sequencial.

CONCLUSÃO

8.1 Considerações Finais

Essa dissertação de mestrado visou apresentar e desenvolver uma nova abordagem para o aprimoramento de sistemas CBIR. Essa abordagem se baseia na calibração de distância em métodos de acesso métrico por meio de refinamento da medida de similaridade por realimentação de relevância. As novas técnicas apresentadas avançaram no estado da arte relativo à inclusão de funções de distância ponderadas em métodos de acesso métrico que permite o aprimoramento semântico de conjunto de dados aplicado ao contexto médio.

A aplicação do método *Fusion Relevance Feedback* (FRF) proposta neste trabalho, atua no aprimoramento de combinações de extratores de características utilizando técnica de realimentação de relevância automática e fornece o melhor descritor para uma base de dados complexo.

A aplicação do método *Tuning Metrics Relevance Feedback* (TMRF) proposta neste trabalho, atua tanto no aprimoramento de técnica de realimentação de relevância e na escolha do vetor de pesos, quanto define o momento da reindexação do método de acesso métrico mantendo a otimização da estrutura de recuperação de dados. A estrutura modular do método também permite a alteração dos parâmetros desejados de forma direta e simples.

8.2 Contribuições

Este trabalho apresentou duas contribuições principais. A primeira dela foi o método *Fusion Relevance Feedback* (FRF) que atua no pré-processamento do conjunto de dados aprimorando o espaço de características de forma automática utilizando Realimentação de Relevância (RR).

A segunda contribuição deste trabalho foi o método *Tuning Metrics Relevance Feedback*

(TMRF), que é um método que infere vetores de pesos no método de acesso métrico *Slim-Tree* e define uma metodologia de reindexação para manter o método de acesso métrico otimizado em conjunto do espaço métrico aprimorado pela RR Desvio Padrão.

No período do desenvolvimento da pesquisa foram produzidos os seguintes artigos:

MARCACINI, Renato Gomes e TRAINA, Agma Juci Machado. *Realimentação de relevância para aprimorar o processamento de consultas por similaridade sobre bases de imagens de COVID-19*. 2022, Anais [...] São Paulo: SBIS, 2022. p. 190-195. Disponível em: <<http://sbis.org.br/anais-estendidos-cbis22>>.

MARCACINI, Renato Gomes; DE OLIVEIRA, Willian Dener; TRAINA, Agma Juci Machado. *Calibração de Distância em Métodos de Acesso Métrico por meio de Realimentação de Relevância*. In: SIMPÓSIO BRASILEIRO DE BANCO DE DADOS (SBBB), 38. , 2023, Belo Horizonte/MG. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023. p. 89-101. ISSN 2763-8979. DOI: <<https://doi.org/10.5753/sbbd.2023.231731>>.

8.3 Trabalhos Futuros

Há diversas formas de aprimorar as propostas aqui apresentadas. Entre elas pode-se citar:

- Ao invés de recalcular diretamente a função de distância ponderada na *Slim-Tree*, utilizar o conjunto de dados retornados de uma consulta e aplicar o ciclo de realimentação de relevância em memória na seção de usuário. Este estudo é interessante para diminuir o custo computacional aplicado ao método de acesso métrico e verificar se um aprendizado em uma amostra de dados pode representar a semântica da base de dados inteira e gerar métodos de acesso métricos específicos para cada perfil de usuário.
- Comparar e analisar se o aprendizado no conjunto de dados em memória é suficiente para aprimorar o espaço métrico. Esta análise é interessante, pois considerando a base de imagens inteira, durante o ciclo de realimentação, novas imagens seriam retornadas ao *top K*. Entretanto, com o conjunto de dados em memória, o aprendizado só é realizado no conjunto da consulta realizada. Aprimorar e selecionar o vetor de pesos durante as *top K* imagens retornadas durante a consulta em memória.
- Incluir a função de distância Mahalanobis na *Slim-Tree*. As funções de distância Mahalanobis são poderosas para aprimorar um espaço métrico, pois consideram as correlações dos atributos do vetor de características. Entretanto, por utilizarem matrizes, são de alto custo computacional para cada transformação em técnica de realimentação de relevância. Estudos para encontrar formas de utilizar somente a diagonal principal da matriz como vetor de pesos e calibrar a distância no método de acesso métrico podem ser essenciais.

- Utilizar a visualização de dados para auxiliar o aprendizado da técnica de realimentação de relevância. Na qual o usuário possa fazer uso da visualização de dados através de um grafo e selecionar o *feedback* de forma mais semântica.

REFERÊNCIAS

- ADORAM, M.; LEW, M. S. Irus: image retrieval using shape. In: IEEE. **Proceedings IEEE International Conference on Multimedia Computing and Systems**. [S.l.], 1999. v. 2, p. 597–602. Citado na página 34.
- AHMED, A. Implementing relevance feedback for content-based medical image retrieval. **IEEE Access**, IEEE, v. 8, p. 79969–79976, 2020. Citado nas páginas 26, 27, 29, 50, 51, 53 e 54.
- AHMED, M.; SERAJ, R.; ISLAM, S. M. S. The k-means algorithm: A comprehensive survey and performance evaluation. **Electronics**, MDPI, v. 9, n. 8, p. 1295, 2020. Citado na página 63.
- AL-DHABYANI, W.; GOMAA, M.; KHALED, H.; FAHMY, A. Dataset of breast ultrasound images. **Data in brief**, Elsevier, v. 28, p. 104863, 2020. Citado na página 68.
- ANTONELLI, L.; SIMONE, V. D.; VIOLA, M. Segmenting mr images through texture extraction and multiplicative components optimization. In: SPRINGER. **International Conference on Scale Space and Variational Methods in Computer Vision**. [S.l.], 2023. p. 511–521. Citado na página 35.
- ARMI, L.; FEKRI-ERSHAD, S. Texture image analysis and texture classification methods-a review. **arXiv preprint arXiv:1904.06554**, 2019. Citado na página 35.
- AZEVEDO-MARQUES, P. M. D.; MENCATTINI, A.; SALMERI, M.; RANGAYYAN, R. M. **Medical image analysis and informatics: computer-aided diagnosis and therapy**. [S.l.]: CRC Press, 2017. Citado na página 49.
- BANERJEE, I.; KURTZ, C.; DEVORAH, A. E.; DO, B.; RUBIN, D. L.; BEAULIEU, C. F. Relevance feedback for enhancing content based image retrieval and automatic prediction of semantic image features: Application to bone tumor radiographs. **Journal of biomedical informatics**, Elsevier, v. 84, p. 123–135, 2018. Citado na página 55.
- BARDHAN, S.; ROGA, S. Feature based automated detection of covid-19 from chest x-ray images. **Emerging Technologies During the Era of COVID-19 Pandemic**, Nature Publishing Group, v. 348, p. 115, 2021. Citado na página 66.
- BEDO, M. V. N.; SANTOS, D. P. D.; PONCIANO-SILVA, M.; AZEVEDO-MARQUES, P. M. de; TRAINA, C. *et al.* Endowing a content-based medical image retrieval system with perceptual similarity using ensemble strategy. **Journal of digital imaging**, Springer, v. 29, n. 1, p. 22–37, 2016. Citado na página 50.
- BENAOUALI, M.; BENTOUMI, M.; TOUATI, M.; AHMED, A. T.; MIMI, M. Segmentation and classification of benign and malignant breast tumors via texture characterization from ultrasound images. In: IEEE. **2022 7th International Conference on Image and Signal Processing and their Applications (ISPA)**. [S.l.], 2022. p. 1–4. Citado na página 66.
- BHOSLE, N.; KOKARE, M. Random forest-based active learning for content-based image retrieval. **International Journal of Intelligent Information and Database Systems**, Inderscience Publishers (IEL), v. 13, n. 1, p. 72–88, 2020. Citado na página 57.

- BÖHM, C.; BERCHTOLD, S.; KEIM, D. A. Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, v. 33, n. 3, p. 322–373, 2001. Citado na página 39.
- BOZKAYA, T.; OZSOYOGLU, M. Distance-based indexing for high-dimensional metric spaces. In: **Proceedings of the 1997 ACM SIGMOD international conference on Management of data**. [S.l.: s.n.], 1997. p. 357–368. Citado na página 44.
- BRESSAN, R. S.; BUGATTI, P. H.; SAITO, P. T. Breast cancer diagnosis through active learning in content-based image retrieval. **Neurocomputing**, Elsevier, v. 357, p. 1–10, 2019. Citado nas páginas 58 e 60.
- BRODATZ. **Brodatz Textures**. [S.l.], 2020. Disponível em: <<https://www.ux.uis.no/~tranden/brodatz.html>>. Citado na página 35.
- BUENO, R. **Tratamento de tempo e dinamicidade em dados representados em espaços métricos**. Tese (Doutorado) — Universidade de São Paulo, 2009. Citado na página 45.
- BURKHARD, W. A.; KELLER, R. M. Some approaches to best-match file searching. **Communications of the ACM**, ACM New York, NY, USA, v. 16, n. 4, p. 230–236, 1973. Citado na página 43.
- CARÉLO, C. C. M.; POLA, I. R. V.; CIFERRI, R. R.; TRAINA, A. J. M.; JR, C. T.; CIFERRI, C. D. de A. Slicing the metric space to provide quick indexing of complex data in the main memory. **Information Systems**, Elsevier, v. 36, n. 1, p. 79–98, 2011. Citado na página 44.
- CEZAR, J. Desenvolvimento de métodos para extração, comparação e análise de características intrínsecas de imagens médicas, visando à recuperação perceptual por conteúdo. Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2005. Citado na página 35.
- CFM. Resolução do conselho federal de medicina. **Resolução n.º 1.821/07**, p. 1–6, 2007. Available online at: <<https://sistemas.cfm.org.br/normas/visualizar/resolucoes/BR/2007/1821>>. Disponível em: <<https://sistemas.cfm.org.br/normas/visualizar/resolucoes/BR/2007/1821>>. Citado na página 29.
- CHANG, Y.-J.; KAMATAKI, K.; CHEN, T. Mean shift feature space warping for relevance feedback. In: IEEE. **2009 16th IEEE International Conference on Image Processing (ICIP)**. [S.l.], 2009. p. 1849–1852. Citado nas páginas 58 e 60.
- CHÁVEZ, E.; NAVARRO, G.; BAEZA-YATES, R.; MARROQUÍN, J. L. Searching in metric spaces. **ACM computing surveys (CSUR)**, ACM New York, NY, USA, v. 33, n. 3, p. 273–321, 2001. Citado na página 39.
- CHÁVEZ, R. O.; ESCALANTE, H. J.; GÓMEZ, M. Montes-y; SUCAR, L. E. Multimodal markov random field for image reranking based on relevance feedback. **International Scholarly Research Notices**, Hindawi, v. 2013, 2013. Citado na página 57.
- CHEN, L.; GAO, Y.; SONG, X.; LI, Z.; ZHU, Y.; MIAO, X.; JENSEN, C. S. Indexing metric spaces for exact similarity search. **ACM Computing Surveys**, ACM New York, NY, v. 55, n. 6, p. 1–39, 2022. Citado na página 41.
- CHEN, L.; GAO, Y.; ZHONG, A.; JENSEN, C. S.; CHEN, G.; ZHENG, B. Indexing metric uncertain data for range queries and range joins. **The VLDB Journal**, Springer, v. 26, n. 4, p. 585–610, 2017. Citado na página 44.

- CIACCIA, P.; PATELLA, M.; ZEZULA, P. M-tree: An efficient access method for similarity search in metric spaces. In: CITESEER. **Vldb**. [S.l.], 1997. v. 97, p. 426–435. Citado na página [44](#).
- DASH, S. S.; SHAKYAWAR, S.; SHARMA, M.; KAUSHIK, S. **Big data in healthcare: management, analysis and future prospects**. [S.l.], 2019. v. 6. Citado na página [25](#).
- DEPEURSINGE, A.; FISCHER, B.; MÜLLER, H.; DESERNO, T. M. Suppl 1: prototypes for content-based image retrieval in clinical practice. **The open medical informatics journal**, Bentham Science Publishers, v. 5, p. 58, 2011. Citado na página [51](#).
- DESERNO, T. M.; ANTANI, S.; LONG, R. Ontology of gaps in content-based image retrieval. **Journal of digital imaging**, Springer, v. 22, n. 2, p. 202–215, 2009. Citado nas páginas [27](#) e [49](#).
- DEVAREDDI, R. B.; SRIKRISHNA, A. Review on content-based image retrieval models for efficient feature extraction for data analysis. In: IEEE. **2022 International Conference on Electronics and Renewable Systems (ICEARS)**. [S.l.], 2022. p. 969–980. Citado na página [31](#).
- DHAYGUDE, M. A.; KINARIWALA, S. A literature survey on content-based information retrieval. **J. Comput. Technol.(JCT)**, v. 11, p. 1–6, 2022. Citado nas páginas [30](#) e [39](#).
- DIMENSTEIN, R.; BOTELHO, M. Z.; LYKAWKA, R.; MEDEIROS, R. F. de; PELEGRINI, F. R.; BORTOLOTO, M. V.; MAZZOLA, A. A.; GRACHI, B. E.; HIRATA, T.; BACELAR, A. Bases físicas e tecnológicas em diagnóstico por imagem. **ABFM**, Colégio Brasileiro de Radiologia e Diagnóstico por Imagem, 2022. Citado na página [65](#).
- FIGUEROA, K.; CAMARENA-IBARROLA, A.; VALERO, L. Permutation based algorithm improved by classes for similarity searching. **Computación y Sistemas**, Instituto Politécnico Nacional, Centro de Investigación en Computación, v. 26, n. 1, p. 71–79, 2022. Citado na página [39](#).
- GHODRATNAMA, S.; MOGHADDAM, H. A. Content-based image retrieval using feature weighting and c-means clustering in a multi-label classification framework. **Pattern Analysis and Applications**, Springer, v. 24, p. 1–10, 2021. Citado na página [42](#).
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. [S.l.]: Pearson, 2018. v. 4. Citado nas páginas [34](#) e [49](#).
- GULDOGAN, E.; GABBOUJ, M. Dynamic feature weights with relevance feedback in content-based image retrieval. In: IEEE. **2009 24th International Symposium on Computer and Information Sciences**. [S.l.], 2009. p. 56–59. Citado na página [54](#).
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. **IEEE Transactions on systems, man, and cybernetics**, Ieee, n. 6, p. 610–621, 1973. Citado nas páginas [31](#), [35](#) e [36](#).
- ISHIKAWA, Y.; SUBRAMANYA, R.; FALOUTSOS, C. Mindreader: Querying databases through multiple examples. Carnegie Mellon University, 1998. Citado na página [57](#).
- KARAMTI, H.; TMAR, M.; VISANI, M.; URRUTY, T.; GARGOURI, F. Vector space model adaptation and pseudo relevance feedback for content-based image retrieval. **Multimedia Tools and Applications**, Springer, v. 77, n. 5, p. 5475–5501, 2018. Citado na página [54](#).

- KHERFI, M. L.; ZIOU, D.; BERNARDI, A. Learning from negative example in relevance feedback for content-based image retrieval. In: IEEE. **Object recognition supported by user interaction for service robots**. [S.l.], 2002. v. 2, p. 933–936. Citado na página 56.
- KUMARAN, K.; PAPAGEORGIOU, D. J.; TAKAC, M.; LUEG, L.; SAHINIDIS, N. V. Active metric learning for supervised classification. **Computers & Chemical Engineering**, Elsevier, v. 144, p. 107132, 2021. Citado nas páginas 59 e 60.
- LATIF, A.; RASHEED, A.; SAJID, U.; AHMED, J.; ALI, N.; RATYAL, N. I.; ZAFAR, B.; DAR, S. H.; SAJID, M.; KHALIL, T. *et al.* Content-based image retrieval and feature extraction: a comprehensive review. **Mathematical problems in engineering**, Hindawi, v. 2019, 2019. Citado na página 31.
- LAWS, K. I. Rapid texture identification. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Image processing for missile guidance**. [S.l.], 1980. v. 238, p. 376–381. Citado na página 36.
- LI, B.; CHANG, E.; WU, Y. Discovery of a perceptual distance function for measuring image similarity. **Multimedia systems**, Springer, v. 8, n. 6, p. 512–522, 2003. Citado na página 50.
- LI, X.; LI, C.; RAHAMAN, M. M.; SUN, H.; LI, X.; WU, J.; YAO, Y.; GRZEGORZEK, M. A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches. **Artificial Intelligence Review**, Springer, v. 55, n. 6, p. 4809–4878, 2022. Citado nas páginas 32 e 33.
- LI, X.; URICCHIO, T.; BALLAN, L.; BERTINI, M.; SNOEK, C. G.; BIMBO, A. D. Socializing the semantic gap: A comparative survey on image tag assignment, refinement, and retrieval. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, v. 49, n. 1, p. 1–39, 2016. Citado na página 29.
- LI, Z.; ZHANG, X.; MÜLLER, H.; ZHANG, S. Large-scale retrieval for medical image analytics: A comprehensive review. **Medical image analysis**, Elsevier, v. 43, p. 66–84, 2018. Citado na página 26.
- LIU, Y.; ZHANG, D.; LU, G.; MA, W.-Y. A survey of content-based image retrieval with high-level semantics. **Pattern recognition**, Elsevier, v. 40, n. 1, p. 262–282, 2007. Citado na página 53.
- MAATEN, L. Van der; HINTON, G. Visualizing data using t-sne. **Journal of machine learning research**, v. 9, n. 11, 2008. Citado na página 63.
- MARKONIS, D.; SCHAER, R.; MÜLLER, H. Evaluating multimodal relevance feedback techniques for medical image retrieval. **Information Retrieval Journal**, Springer, v. 19, n. 1-2, p. 100–112, 2016. Citado na página 55.
- MARTINET, J.; CHIARAMELLA, Y.; MULHEM, P. A relational vector space model using an advanced weighting scheme for image retrieval. **Information processing & management**, Elsevier, v. 47, n. 3, p. 391–414, 2011. Citado na página 54.
- MOHANAN, A.; RAJU, S. A survey on different relevance feedback techniques in content based image retrieval. **Int. Res. J. Eng. Technol**, v. 4, n. 2, p. 582–585, 2017. Citado na página 26.

MOSBAH, M. Query refinement into information retrieval systems: An overview. **Journal of Information and Organizational Sciences**, Fakultet organizacije i informatike Sveučilišta u Zagrebu, v. 47, n. 1, p. 133–151, 2023. Citado na página 26.

MUFARROHA, F.; ANAMISA, D.; HAPSANI, A. Content based image retrieval using two color feature extraction. In: IOP PUBLISHING. **Journal of physics: conference series**. [S.l.], 2020. v. 1569, n. 3, p. 032072. Citado na página 43.

MULLER, H.; MULLER, W.; MARCHAND-MAILLET, S.; PUN, T.; SQUIRE, D. M. Strategies for positive and negative relevance feedback in image retrieval. In: IEEE. **Proceedings 15th International Conference on Pattern Recognition. ICPR-2000**. [S.l.], 2000. v. 1, p. 1043–1046. Citado na página 55.

NESSO, M. R.; CAZZOLATO, M. T.; SCABORA, L. C.; OLIVEIRA, P. H.; SPADON, G.; SOUZA, J. A. de; OLIVEIRA, W. D.; CHINO, D. Y.; RODRIGUES, J. F.; TRAINA, A. J. *et al.* Rafiki: Retrieval-based application for imaging and knowledge investigation. In: IEEE. **2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.], 2018. p. 71–76. Citado na página 49.

NOVAES, J. V. d. O.; SANTOS, L. F. D.; CARVALHO, L. O.; OLIVEIRA, D. d.; BÊDO, M. V. N.; TRAINA, A. J. M.; JUNIOR, C. T. J-eda: a workbench for tuning similarity and diversity search parameters in content-based image retrieval. **Journal of Information and Data Management-JIDM**, v. 12, n. 2, p. 126–137, 2021. Citado na página 54.

PEREIRA, R. M.; BERTOLINI, D.; TEIXEIRA, L. O.; JR, C. N. S.; COSTA, Y. M. Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 194, p. 105532, 2020. Citado na página 66.

POLA, I. R. V.; JR, C. T.; TRAINA, A. J. M. The nobh-tree: Improving in-memory metric access methods by using metric hyperplanes with non-overlapping nodes. **Data & Knowledge Engineering**, Elsevier, v. 94, p. 65–88, 2014. Citado nas páginas 44 e 45.

POLA, I. R. V.; TRAINA, C.; TRAINA, A. J. M. The mm-tree: A memory-based metric tree without overlap between nodes. In: SPRINGER. **East European Conference on Advances in Databases and Information Systems**. [S.l.], 2007. p. 157–171. Citado na página 44.

PONCIANO-SILVA, M.; TRAINA, A. J.; AZEVEDO-MARQUES, P. M.; FELIPE, J. C.; TRAINA, C. Including the perceptual parameter to tune the retrieval ability of pulmonary cbir systems. In: IEEE. **2009 22nd IEEE International Symposium on Computer-Based Medical Systems**. [S.l.], 2009. p. 1–8. Citado na página 50.

PORKAEW, K.; CHAKRABARTI, K. Query refinement for multimedia similarity retrieval in mars. In: **Proceedings of the seventh ACM international conference on Multimedia (Part 1)**. [S.l.: s.n.], 1999. p. 235–238. Citado na página 57.

RAHMAN, T. **Covid-19 radiography database**. 2022. Available online at: <<https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>>. Disponível em: <<https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>>. Citado na página 67.

RAZENDE, H.; BARIONI, M. C. N.; SILVA, Y. N. Storing data once in m-trees and pm-trees: Revisiting the building principles of metric access methods. **Information Systems**, Elsevier, v. 104, p. 101896, 2022. Citado na página 26.

REZAZADEH, A.; JAFARIAN, Y.; KORD, A. Explainable ensemble machine learning for breast cancer diagnosis based on ultrasound image texture features. **Forecasting**, MDPI, v. 4, n. 1, p. 262–274, 2022. Citado na página 66.

REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. [S.l.]: Editora Manole Ltda, 2003. Citado na página 29.

ROCCHIO, J. Relevance feedback in information retrieval. **The Smart retrieval system-experiments in automatic document processing**, Prentice Hall, p. 313–323, 1971. Citado na página 54.

ROUMELIS, G.; VASSILAKOPOULOS, M.; CORRAL, A.; MANOLOPOULOS, Y. An efficient algorithm for bulk-loading xbr+-trees. **Computer Standards & Interfaces**, Elsevier, v. 57, p. 83–100, 2018. Citado nas páginas 44 e 45.

ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of computational and applied mathematics**, Elsevier, v. 20, p. 53–65, 1987. Citado na página 63.

RUI, Y.; HUANG, T. S.; MEHROTRA, S. Relevance feedback techniques in interactive content-based image retrieval. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Storage and Retrieval for Image and Video Databases VI**. [S.l.], 1997. v. 3312, p. 25–36. Citado na página 55.

RUI, Y.; HUANG, T. S.; ORTEGA, M.; MEHROTRA, S. Relevance feedback: a power tool for interactive content-based image retrieval. **IEEE Transactions on circuits and systems for video technology**, IEEE, v. 8, n. 5, p. 644–655, 1998. Citado nas páginas 53, 58, 60, 73 e 78.

SALTON, G.; MCGILL, M. J. **Introduction to modern information retrieval**. [S.l.]: mcgraw-hill, 1983. Citado na página 54.

SHAHAPURE, K. R.; NICHOLAS, C. Cluster quality analysis using silhouette score. In: IEEE. **2020 IEEE 7th international conference on data science and advanced analytics (DSAA)**. [S.l.], 2020. p. 747–748. Citado na página 63.

SHRIVASTAVA, P.; BHOYAR, K.; ZADGAONKAR, A. Bridging the semantic gap with human perception based features for scene categorization. **International Journal of Intelligent Computing and Cybernetics**, Emerald Publishing Limited, 2017. Citado na página 49.

SILVA, A. T. da; SANTOS, J. A. dos; FALCÃO, A. X.; TORRES, R. d. S.; MAGALHÃES, L. P. Incorporating multiple distance spaces in optimum-path forest classification to improve feedback-based learning. **Computer Vision and Image Understanding**, Elsevier, v. 116, n. 4, p. 510–523, 2012. Citado na página 57.

SILVA, M. P. d. **Processamento de consultas por similaridade em imagens médicas visando à recuperação perceptual guiada pelo usuário**. Tese (Doutorado) — Universidade de São Paulo, 2009. Citado na página 26.

- SU, J.-H.; HONG, T.-P.; CHEN, Y.-T.; CHIN, C.-Y. High-performance content-based music retrieval via automated navigation and semantic features. **Engineering Applications of Artificial Intelligence**, Elsevier, v. 115, p. 105267, 2022. Citado na página 54.
- SU, Z.; ZHANG, H.; LI, S.; MA, S. Relevance feedback in content-based image retrieval: Bayesian framework, feature subspaces, and progressive learning. **IEEE transactions on image processing**, IEEE, v. 12, n. 8, p. 924–937, 2003. Citado na página 56.
- TAJBAKHS, N.; JEYASEELAN, L.; LI, Q.; CHIANG, J. N.; WU, Z.; DING, X. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. **Medical Image Analysis**, Elsevier, v. 63, p. 101693, 2020. Citado na página 34.
- TIAN, D. A review on relevance feedback for content-based image retrieval. **J. Inf. Hiding Multim. Signal Process.**, v. 9, n. 1, p. 108–119, 2018. Citado nas páginas 26, 42 e 53.
- TORRES, R. da S.; FALCAO, A. X. Content-based image retrieval: theory and applications. **RITA**, v. 13, n. 2, p. 161–185, 2006. Citado na página 43.
- TRAINA, A.; TRAINA, C.; FALOUTSOS, C. *et al.* Similarity search without tears: the omni-family of all-purpose access methods. In: IEEE. **Proceedings 17th International Conference on Data Engineering**. [S.l.], 2001. p. 623–630. Citado na página 44.
- TRAINA, C.; FILHO, R. F. S.; TRAINA, A. J.; VIEIRA, M. R.; FALOUTSOS, C. The omni-family of all-purpose access methods: a simple and effective way to make similarity search more efficient. **The VLDB Journal**, Springer, v. 16, n. 4, p. 483–505, 2007. Citado na página 44.
- TRAINA, C.; TRAINA, A.; FALOUTSOS, C.; SEEGER, B. Fast indexing and visualization of metric data sets using slim-trees. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 14, n. 2, p. 244–260, 2002. Citado na página 45.
- TRAINA, C.; TRAINA, A.; SEEGER, B.; FALOUTSOS, C. Slim-trees: High performance metric trees minimizing overlap between nodes. In: SPRINGER. **International Conference on Extending Database Technology**. [S.l.], 2000. p. 51–65. Citado na página 44.
- VIEIRA, M. R.; JR, C. T.; CHINO, F. J.; TRAINA, A. J. Dbm-tree: Trading height-balancing for performance in metric access methods. **Journal of the Brazilian Computer Society**, SciELO Brasil, v. 11, p. 37–51, 2006. Citado na página 44.
- VISHRAJ, R.; GUPTA, S.; SINGH, S. A comprehensive review of content-based image retrieval systems using deep learning and hand-crafted features in medical imaging: Research challenges and future directions. **Computers and Electrical Engineering**, Elsevier, v. 104, p. 108450, 2022. Citado nas páginas 25 e 49.
- WAN, C.; LIU, M. Content-based audio retrieval with relevance feedback. **Pattern recognition letters**, Elsevier, v. 27, n. 2, p. 85–92, 2006. Citado na página 53.
- WANG, J. Z.; LI, J.; WIEDERHOLD, G. Simplicity: Semantics-sensitive integrated matching for picture libraries. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 23, n. 9, p. 947–963, 2001. Citado na página 68.
- WANG, X.; FANG, H.; ZHAI, C. A study of methods for negative relevance feedback. In: **Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval**. [S.l.: s.n.], 2008. p. 219–226. Citado nas páginas 55 e 56.

- WANG, X. X.; FANG, H.; ZHAI, C. Improve retrieval accuracy for difficult queries using negative feedback. In: **Proceedings of the sixteenth ACM conference on Conference on information and knowledge management**. [S.l.: s.n.], 2007. p. 991–994. Citado nas páginas 55 e 56.
- WEI, M.; AMIT, N. Leveraging hyperupcalls to bridge the semantic gap: An application perspective. **IEEE Data Eng. Bull.**, v. 42, n. 1, p. 22–35, 2019. Citado na página 27.
- WEIHS, P.; CHIHEB, B.; NCIR; GHOURABI, M. E.; LIMAM, M.; KARAA, A.; ZOU, C.; FRANCO, C.; WEIHS, C.; PORZIO, G.; PUCETTI, G.; TALEB, H.; AROUI, M.-A.; AYADI, M.; AZAIZ, N.; NADIA, E.; GHAZZELI, N.; GIUDICI, P.; GÖB, R. **Proceedings of MSDM 2014**. [S.l.: s.n.], 2014. Citado na página 69.
- WEINBERGER, K. Q.; SAUL, L. K. Distance metric learning for large margin nearest neighbor classification. **Journal of machine learning research**, v. 10, n. 2, 2009. Citado nas páginas 58 e 60.
- WILLEMINK, M. J.; KOSZEK, W. A.; HARDELL, C.; WU, J.; FLEISCHMANN, D.; HARVEY, H.; FOLIO, L. R.; SUMMERS, R. M.; RUBIN, D. L.; LUNGREN, M. P. Preparing medical imaging data for machine learning. **Radiology**, Radiological Society of North America, v. 295, n. 1, p. 4–15, 2020. Citado nas páginas 29 e 30.
- WU, C.-M.; CHEN, Y.-C. Statistical feature matrix for texture analysis. **CVGIP: Graphical Models and Image Processing**, Elsevier, v. 54, n. 5, p. 407–419, 1992. Citado na página 36.
- XU, H.; WANG, J.-y.; MAO, L. Relevance feedback for content-based image retrieval using deep learning. In: IEEE. **2017 2nd International Conference on Image, Vision and Computing (ICIVC)**. [S.l.], 2017. p. 629–633. Citado na página 56.
- YAN, R.; HAUPTMANN, A. G.; JIN, R. Negative pseudo-relevance feedback in content-based video retrieval. In: **Proceedings of the eleventh ACM international conference on Multimedia**. [S.l.: s.n.], 2003. p. 343–346. Citado na página 56.
- YIANILOS, P. N. Data structures and algorithms for nearest neighbor search in general metric spaces. In: **Soda**. [S.l.: s.n.], 1993. v. 93, n. 194, p. 311–21. Citado na página 44.
- YIN, C.; TANG, J.; YUAN, T.; XU, Z.; WANG, Y. Bridging the gap between semantic segmentation and instance segmentation. **IEEE Transactions on Multimedia**, IEEE, v. 24, p. 4183–4196, 2021. Citado nas páginas 25, 29 e 49.
- YUNITA, A.; SANTOSO, H. B.; HASIBUAN, Z. A. ‘everything is data’: towards one big data ecosystem using multiple sources of data on higher education in indonesia. **Journal of Big Data**, SpringerOpen, v. 9, n. 1, p. 1–22, 2022. Citado na página 25.
- ZHANG, Z.; JI, R.; YAO, H.; XU, P.; WANG, J. Random sampling svm based soft query expansion for image retrieval. In: IEEE. **Fourth International Conference on Image and Graphics (ICIG 2007)**. [S.l.], 2007. p. 805–809. Citado na página 29.

