

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA

ELIAS BORGES MACENA JÚNIOR

**Aplicação de Técnicas de
Content-Based Image Retrieval (CBIR)
em Imagens Radiográficas**

Goiânia
2016

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: ☒ **Dissertação** ☐ **Tese**

2. Identificação da Tese ou Dissertação

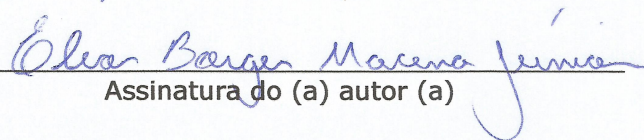
Nome completo do autor: Elias Borges Macena Júnior

Título do trabalho: Aplicação de Técnicas de Content-Based Image Retrieval (CBIR) em Imagens Radiográficas

3. Informações de acesso ao documento:

Concorda com a liberação total do documento ☒ SIM ☐ NÃO¹

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.


Assinatura do (a) autor (a)

Data: 29 / 08 / 16

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

ELIAS BORGES MACENA JÚNIOR

Aplicação de Técnicas de Content-Based Image Retrieval (CBIR) em Imagens Radiográficas

Dissertação apresentada ao Programa de Pós-Graduação do Instituto de Informática da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação.

Orientador: Prof. Dr. Ronaldo Martins da Costa

Goiânia
2016

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Macena Júnior, Elias

Aplicação de Técnicas de Content-Based Image Retrieval (CBIR)
em Imagens Radiográficas [manuscrito] / Elias Macena Júnior. -
2016.

LXXI, 71 f.: il.

Orientador: Prof. Dr. Ronaldo Martins da Costa.

Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto
de Informática (INF), Programa de Pós-Graduação em Ciência da
Computação, Goiânia, 2016.

Bibliografia. Apêndice.

Inclui fotografias, abreviaturas, gráfico, tabelas, lista de figuras,
lista de tabelas.

1. CBIR. 2. CAD. 3. Imagens Médicas. I. Martins da Costa,
Ronaldo, orient. II. Título.

CDU 004



ATA Nº 22/2016

**ATA DA SESSÃO DE JULGAMENTO DA DISSERTAÇÃO
DE MESTRADO DE ELIAS BORGES MACENA JUNIOR**

Aos trinta dias do mês de setembro de dois mil e dezesseis, às nove horas e trinta minutos, na sala 150 do Instituto de Informática da Universidade Federal de Goiás, Campus Samambaia, reuniu-se a banca examinadora designada na forma regimental pela Coordenação do Curso para julgar a dissertação de mestrado intitulada “**Aplicação de Técnicas de Content-Based Image Retrieval (CBIR) em Imagens Radiográficas**”, apresentada pelo aluno Elias Borges Macena Junior como parte dos requisitos necessários à obtenção do grau de Mestre em Ciência da Computação, área de concentração Ciência da Computação. A banca examinadora foi presidida pelo orientador do trabalho de dissertação, Professor Doutor Ronaldo Martins da Costa (INF/UFG), tendo como membros o Professor Doutor Leandro Luis Galdino de Oliveira (INF/UFG) e a Professora Doutora Fernanda Paula Yamamoto Silva (FO/UFG). Aberta a sessão, o candidato expôs seu trabalho. Em seguida, o aluno foi arguido pelos membros da banca e:

(☒) tendo demonstrado suficiência de conhecimento e capacidade de sistematização do tema de sua dissertação, a banca concluiu pela **aprovação** do candidato, sem restrições.

() tendo demonstrado suficiência de conhecimento e capacidade de sistematização do tema de sua dissertação, a banca concluiu pela **aprovação** do candidato, condicionado a satisfazer as exigências listadas na Folha de Modificação de Dissertação de Mestrado anexa à presente ata, no prazo máximo de 60 dias, a contar da presente data, ficando o professor orientador responsável por atestar o cumprimento dessas exigências.

() não tendo demonstrado suficiência de conhecimento e capacidade de sistematização do tema de sua dissertação, a banca concluiu pela **reprovação** do candidato.

Os trabalhos foram encerrados às 12 horas e trinta minutos. Nos termos do Regulamento Geral dos Cursos de Pós-Graduação desta Universidade, lavrou-se a presente ata que, lida e julgada conforme, segue assinada pelos membros da banca examinadora.

Prof. Dr. Ronaldo Martins da Costa

Prof. Dr. Leandro Luis Galdino de Oliveira

Profa. Dra. Fernanda Paula Yamamoto Silva

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador(a).

Elias Borges Macena Júnior

Graduou-se em Ciência da Computação na UFMT - Universidade Federal de Mato Grosso, campus de Barra do Garças. Durante sua graduação, foi monitor no curso de Ciência da Computação da UFMT. Durante o Mestrado, na UFG - Universidade Federal de Goiás, foi bolsista da CAPES e desenvolveu um trabalho de desenvolvimento de técnicas de recuperação de imagens baseado em conteúdo para auxílio ao diagnóstico sobre orientação de Ronaldo Martins da Costa

Dedico este trabalho a minha família, que me deu todo apoio durante este projeto

Agradecimentos

A todos aqueles que me ajudaram direta, ou indiretamente, na conclusão deste trabalho

Não! Tentar não. Faça ou não faça. Tentativa não há.

Mestre Yoda,
Star Wars: Episode V - The Empire Strikes Back.

Resumo

Macena Junior, E B. **Aplicação de Técnicas de Content-Based Image Retrieval (CBIR) em Imagens Radiográficas**. Goiânia, 2016. 71p. Dissertação de Mestrado. Instituto de Informática, Universidade Federal de Goiás.

Com o objetivo de melhorar o processo de diagnóstico vários centros de pesquisas têm focado no desenvolvimento de sistemas de informação aplicando poderosas técnicas de diagnóstico auxiliado por computador (Computer-Aided Diagnosis-CAD). Neste contexto, a criação de métodos de recuperação de imagens baseado em conteúdo (Content-based image retrieval - CBIR) é um passo importante para desenvolver um sistema CAD eficiente. Este trabalho propõe a validação de técnicas de recuperação com um método híbrido CBIR baseado em imagens médicas 2D. Os resultados das técnicas aplicadas, indicam uma taxa de acerto de 90,25% e ainda indicam um ganho de 35% no desempenho das técnicas, isto é, no tempo de busca e recuperação das imagens, abrindo caminho para o desenvolvimento de sistemas de informação mais eficientes para construção de sistemas de apoio ao diagnóstico genéricos.

Palavras-chave

CBIR, CAD, Imagens Médicas

Abstract

Macena Junior, E B. **Application of Techniques of Content- Based Image Retrieval (CBIR) in Radiographic Images**. Goiânia, 2016. 71p. MSc. Dissertation. Instituto de Informática, Universidade Federal de Goiás.

In order to improve the diagnostic process several research centers have focused on the development of information systems applying powerful techniques of computer-aided diagnosis (CAD). In this context, the creation of content-based image retrieval (CBIR) is an important step in developing an efficient CAD system. This work proposes the validation of recovery with a hybrid CBIR method based on 2D medical images. The results of the techniques applied, indicate a hit rate of 90.25% and indicate a gain of 35% in the performance of techniques, that is, the time search and retrieval of images, paving the way for the development of information systems more efficient to build support generic diagnostic systems.

Keywords

CBIR, CAD, Medical Images

Sumário

Lista de Figuras	13
Lista de Tabelas	14
1 Introdução	15
1.1 Recuperação de imagem baseada em conteúdo (<i>Content-Based Image Retrieval</i> - CBIR)	15
1.2 Justificativa	16
1.3 Trabalhos Relacionados	17
1.4 Objetivos	18
1.5 Estrutura da dissertação	18
2 Fundamentação Teórica	19
2.1 Processamento Digital de Imagens	19
2.1.1 Limiarização	21
2.2 Análise de Imagens	21
2.2.1 Representação e Descrição	22
2.2.2 Haralick	23
2.2.3 Equações Utilizadas	24
2.3 Reconhecimento de Padrões	27
2.3.1 K-vizinhos mais próximos (<i>K-Nearest Neighbors</i> - KNN)	28
2.3.2 Análise de Componentes Principais (<i>Principal Component Analysis</i> - PCA)	30
2.3.3 Metricas de Distância	31
2.4 Metodos de Validação	32
3 Metodologia	35
3.1 Armazenamento	35
3.2 Recuperação	36
3.2.1 Limiar Textual	37
3.2.2 Grupo de Imagens e Classificação	39
3.3 Detalhes do Sistema	40
4 Resultados	43
4.1 Radiografias de Pneumonia	43
4.2 Radiografias Periapicais	45
4.3 Resultados Finais	46
5 Conclusão	48

Referências Bibliográficas	49
A Apêndice 1	53

Lista de Figuras

1.1	Exemplo de CBIR.	16
2.1	(a) Imagem original (b) ROI pulmão direito (c) ROI pulmão esquerdo.	20
2.2	Funcionamento da Matriz de co-ocorrência.	24
2.3	Imagem com 4 níveis de cinza (0, 1, 2 e 3).	24
2.4	Matrizes de co-ocorrência geradas nos 4 ângulos.	25
2.5	Exemplos do KNN. (a) $K = 1$ a nova amostra pertence a classe azul. (b) $K = 2$ continua a classe azul devido a proximidade maior. (c) $K = 3$ Passa a pertencer a classe amarela, já que são duas amarelas mais próximas e apenas uma azul.	29
2.6	Exemplos de PCA. (a) Números de componentes igual ao das variáveis iniciais. (b) Projeção de 3 variáveis em duas componentes.	31
2.7	Diferença entre distância Euclidiana (linha verde) e distância Manhattan (linhas vermelha, azul e amarela).	32
2.8	Exemplo de 3 imagens e suas respectivas características.	33
2.9	Processo de validação do método leave-one-out.	34
3.1	Funcionamento do limiar textual.	38
3.2	Imagem Limiarizada.	40
3.3	Áreas de interesse extraídas da imagem original.	41
3.4	Exemplo de radiografia Periapical.	41
4.1	(a) Imagem de entrada (b) Primeiro resultado (c) Segundo Resultado.	45
4.2	Resultados dos testes em radiografias periapicais. (a) Imagem de entrada. (b) Recuperação da própria imagem dentro da base. (c) Primeira imagem mais semelhante presente na base.	46

Lista de Tabelas

4.1	Taxa de acerto na recuperação de imagens.	44
4.2	Tempos e taxa de acerto da metodologia.	44
4.3	Taxa de acerto na recuperação de imagens dentárias.	45
4.4	Taxas de Acerto com numero de componentes variando de 1 a 17.	47
4.5	Compilação dos resultados obtidos.	47

Introdução

Diagnóstico Auxiliado por Computador (*Computer-aided diagnosis* - CAD) é um sistema que visa aprimorar o diagnóstico servindo como uma segunda opinião ao especialista. Uma das vantagens de se utilizar este tipo de sistema é o fato dele não ser afetado por diversos fatores ao qual o especialista humano está sujeito, como distrações, cansaço, fadiga ocular, estresse.

O sistema CAD funciona imitando a percepção do especialista, no sistema construído, primeiro extraindo informações da imagem. O ato de extrair essas informações da imagem é chamado de extração de características. Na etapa seguinte o software deve tomar uma decisão com base nessas informações. Essa decisão é tomada a partir de classificação por aprendizagem supervisionada.

Esse processo de extração de características e comparação entre imagens é conhecido como recuperação de imagens baseado em conteúdo (*Content-Based Image Retrieval* - CBIR). Com uma técnica CBIR eficiente é possível desenvolver um sistema CAD eficiente. E é nesta parte que está o foco deste trabalho, desenvolver e validar uma técnica de recuperação eficiente genérica, isto é, que funcione em vários tipos de imagem, para apoio ao diagnóstico.

1.1 Recuperação de imagem baseada em conteúdo (*Content-Based Image Retrieval* - CBIR)

O termo recuperação de imagens baseada em conteúdo se originou em 1992, quando ele foi usado por T. Kato para descrever experimentos em recuperação automática de imagens de um banco de dados, com base nas cores e formas[11]. Desde então, o termo tem sido utilizado para descrever o processo de recuperação de imagens a partir de uma coleção com base em suas características. As técnicas, ferramentas e algoritmos utilizados são originários de várias áreas, como estatísticas, reconhecimento de padrões, processamento de sinais e visão computacional.

Devido ao avanço no armazenamento de dados e nas tecnologias de aquisição de imagem, os conjuntos de imagens se tornaram cada vez maiores. Para tornar possível lidar com tais conjuntos é necessário o desenvolvimento de ferramentas de CBIR que os administrem de forma eficiente[40].

Essas ferramentas consistem de duas principais funcionalidades: Extração de características da imagem e processamento das consultas.

O primeiro passo consiste em analisar uma imagem, e extrair características de interesse como como textura[36], forma, cor, dentre outros[17]. Existem várias características que podem ser utilizadas de acordo com o objetivo da técnica, no capítulo 2, será discutido detalhadamente como este processo é realizado.

O segundo passo é processar a consulta do usuário, através de comparações, isto é, comparar as características extraídas da imagem de consulta, com as imagens presentes na base, e através de alguma métrica, como a distância euclidiana, determinar as imagens mais próximas daquela usada como parâmetro, quanto menor a distância entre duas imagens, mais semelhantes elas serão. Na figura 1.1 pode ser visto um exemplo baseado na cores das imagens.



Figura 1.1: Exemplo de CBIR.

A utilização de técnicas de CBIR podem auxiliar na composição de sistemas CAD pois proveem aos casos similares a um determinado modelo fornecido como argumento de consulta [28] [37] [22]. Através dos casos similares encontrados o sistema é capaz de classificar a imagem usada como consulta, em outras palavras, é realizado o diagnóstico.

1.2 Justificativa

A principal motivação de estudos nessa área é otimizar o processo de diagnóstico, isto é, acelerando o processo, ao mesmo tempo em que se aumenta a qualidade do mesmo. Com isso, vários centros de pesquisa têm focado suas pesquisas no desenvolvimento de sistemas CAD mais poderosos [25] [22] [21]. Aplicações de radiologia, por exemplo, têm impulsionado o desenvolvimento de sistemas CAD e CBIR. A validação dos algoritmos de tais sistemas exige bases de imagens com casos variados e que conte-

nam estruturas de interesse procuradas pelos algoritmos. Porém em sua grande maioria, os bancos de imagens hoje disponíveis possuem várias limitações como ferramentas de busca ineficientes, ou ainda poucas variações de casos[25].

Uma vez definido um banco de imagens, são necessárias ferramentas que possam buscar e recuperar informações dentro do mesmo de maneira eficiente. Dada esta carência e demanda por sistemas CAD, justifica a condução de estudos que possam otimizar a busca e recuperação de imagens presentes nesses bancos de imagem, isto é, justifica a criação de técnicas de CBIR mais eficientes. Por consequência, a existência de técnicas de busca mais amplas e eficientes, pode também impulsionar o estudo e desenvolvimento de sistemas CAD mais eficientes.

1.3 Trabalhos Relacionados

Nesta seção serão apresentados alguns trabalhos com foco em CBIR e CAD. Esta seção serve para ressaltar a justificativa e o objetivo da realização deste trabalho.

Em seu estudo, Muller[28] identificou que a maioria das pesquisas sobre CBIR para aplicações médicas se concentram em algoritmos para extrair características de determinadas regiões do corpo humano ou de uma determinada patologia.

Yuan[42] construiu um banco de imagens bidimensionais (2D) de tomografia computadorizada do cérebro, todas as imagens foram fornecidas por hospitais e colaboradores.

Sanghavi[36] através das seguintes características de uma região de interesse (*Region of Interest* - ROI): Área do objeto, perímetro do objeto, descritor de Fourier e padrão binário local das imagens. O sistema era capaz de diagnosticar doenças relacionadas ao sangue como leucemia e malária.

O Spine Pathology and Image Retrieval System (SPIRS) [3] [24] [39] da U.S. National Library of Medicine fornece técnicas CBIR baseada em características da coluna vertebral mas trabalha apenas com imagens homogênea dentro da coleção.

Ramamurthy[32] usaram o detector de bordas de Canny para extrair características de imagens médicas e depois o Support Vector Machine (SVM) para classificá-las. Foram testados diferentes tipos de imagens, como imagens do cérebro e do peito. O principal problema deste método é que a medida que a base de imagens crescia a performance do algoritmo decrescia.

Além de Oliveira[30], Macedo[8] e Souza[38] que serviram como base para implementação de uma das técnicas presentes neste trabalho, para recuperação de radiografias com pneumonia. Os detalhes desta técnica serão melhor discutidas no capítulo 3.

Grande parte dos trabalhos da área são voltados para um único tipo de imagem, ou ainda, casos onde apesar da técnica trabalhar tipos diferentes de imagem, o grau de

eficiência da técnica diminui à medida que a base de imagens cresce.

1.4 Objetivos

Devido a limitação dos sistemas atuais é necessário definir uma abordagem onde a escalabilidade do sistema não atrapalhe a eficiência do mesmo.

Para isso além do CBIR sera realizado um pré-busca baseada em um limiar que podera definir o foco de busca das imagens de acordo com seu conteúdo, podendo assim, descartar o restante de imagens que não condizem com o tipo a ser buscado. Permitindo assim que a base armazene vários tipos de imagens.

Com base nessas premissas, o foco deste trabalho é desenvolver e validar técnicas de recuperação em vários tipos de imagem com foco ainda no tempo de recuperação das técnicas, utilizando uma base com imagens médicas 2D, criando assim uma técnica genérica que funcione em diferentes tipos de imagem. Este estudo pode contribuir para o desenvolvimento de sistemas CAD mais eficientes que auxiliem os profissionais de saúde.

1.5 Estrutura da dissertação

A fim de melhor apresentar os aspectos do trabalho e suas contribuições na criação de técnicas de armazenamento com foco no tempo de recuperação das mesmas, o texto foi organizado em 5 capítulos principais e suas respectivas seções e subseções.

O capítulo 1, esta introdução, tem como base os principios que levaram ao desenvolvimento do trabalho como o armazenamento e a recuperação das imagens, além de sua justificativa e seus objetivos, contando ainda com os trabalhos relacionados.

O capítulo 2 consiste da fundamentação teórica do trabalho, indo desde os descritores de textura e as características usadas para representar a imagem, até a métrica de distância usada para determinar a semelhança. Além disso serão apresentados alguns conceitos a respeito do armazenamento das imagens, conceitos esses que permitirão o uso do limiar que delimitará o escopo onde será realizada a busca.

O capítulo 3 é baseado na metodologia impregada neste trabalho para realização das buscas, mostrando passo a passo como ocorre o armazento, recuperação e exibição das imagens resultantes. Além disso sera mostrado alguns aspectos do software implementado que realiza tais ações.

No capítulo 4 serão discutidos os resultados obtidos, em todos os testes realizados com 3 tipos diferentes de imagem afim de demonstrar a eficiencia do método.

Por fim, o capítulo 5 mostra a conclusão e também sugestões para trabalhos futuros.

Fundamentação Teórica

Neste capítulo serão apresentados as fórmulas e conceitos aplicados a fim de resolver os problemas propostos neste trabalho.

2.1 Processamento Digital de Imagens

O interesse em métodos de processamento digital de imagens decorre de duas áreas principais de aplicação: melhoria da informação para a interpretação humana; e processamento de imagem para o armazenamento, transmissão e representação para a percepção autônoma[14].

Mas o que é uma imagem? Uma imagem pode ser definida como uma função bidimensional $f(x,y)$, onde a amplitude de f na coordenada (x,y) é chamada de intensidade ou nível de cinza da imagem naquele ponto. Se os valores de f são finitos e discretos então a imagem é chamada de imagem digital.

Uma imagem digital é composta de uma quantidade finita de elementos, cada um deles e um local único e valor atribuído. Estes elementos são conhecidos como Pixels.

O processamento de imagens é a etapa computacional que consiste na transformação de uma imagem de entrada em uma imagem de saída. Seu objetivo é melhorar certos aspectos da imagem de entrada a fim de facilitar etapas posteriores de um problema particular que a envolva.

Uma característica marcante dessa área é a multidisciplinaridade, pois além do conhecimento sobre técnicas de processamento, precisamos também adquirir conhecimento sobre a área onde as imagens serão utilizadas.

Não há consenso geral em quando o processamento digital de imagens termina e outras áreas começam como a análise de imagens e a visão computacional. Esta última tem como objetivo emular a visão humana, incluindo aprender e ser capaz de tomar decisões baseada em entradas visuais[14]. Este é o foco deste trabalho, e será melhor detalhado na seção 2.2.

O processamento é a primeira etapa a ser feita para qualquer imagem que venha a ser utilizada para classificação. Um exemplo disso é a conversão de uma imagem colorida

para uma imagem em níveis de cinza, pois todas as técnicas empregadas neste trabalho lidam com imagens deste tipo; logo, caso seja necessário realizar algum procedimento sobre uma imagem colorida é necessário antes convertê-la. Tal procedimento consiste em recuperar os valores RGB (canal vermelho, verde e azul) de todos os pixels da imagem e depois ponderá-los: $0.3R + 0.59G + 0.11B$. Esses valores estão relacionados à sensibilidade do olho humano.

Também é possível aplicar funções para determinar a região de interesse (Region of interest - ROI) ao qual será analisada, por exemplo, ao classificar uma imagem do pulmão, é definido o centro do pulmão esquerdo e direito como ROI. Na figura 2.1. pode ser visto como ocorre a divisão.

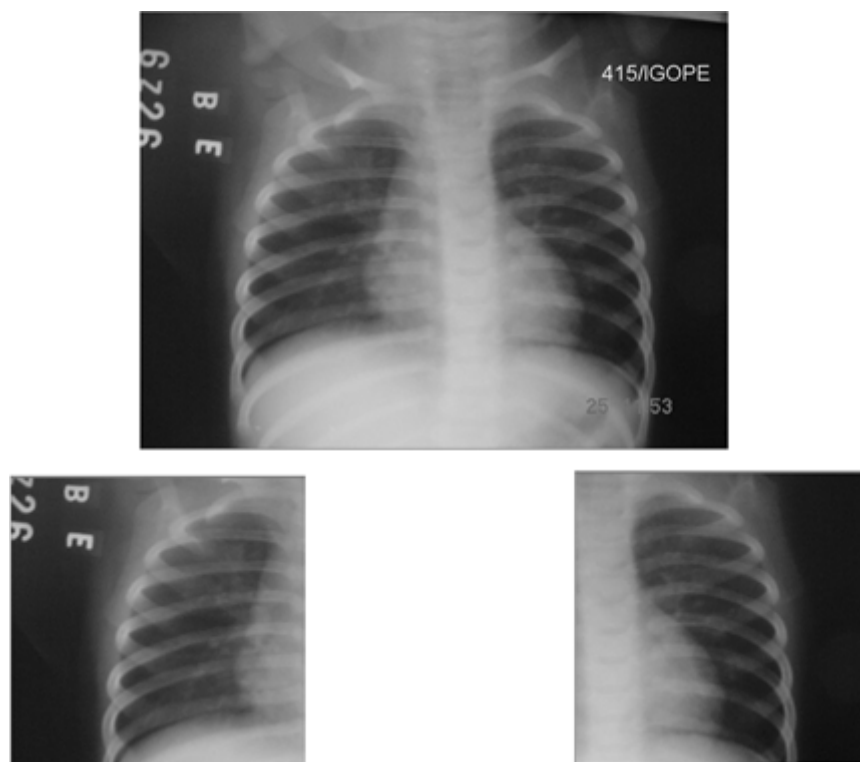


Figura 2.1: (a) Imagem original (b) ROI pulmão direito (c) ROI pulmão esquerdo.

É utilizado ainda um processo conhecido como limiarização para definir o ponto em que a imagem será cortada no eixo y (altura). Pode ainda ser adicionado novas ferramentas de processamento de imagem caso novas classes de imagens sejam adicionadas e requeiram um processamento inicial para definição da ROI. Podem também ser utilizadas imagens sem nenhum tipo de processamento inicial além da conversão para níveis de cinza, como ocorre nas imagens dentárias. No capítulo 3 será detalhado como ocorre este processo de corte enquanto na seção 2.1.1 será detalhado o processo de limiarização.

2.1.1 Limiarização

A limiarização tem um funcionamento simples: suponha uma imagem composta por objetos mais claros distribuídos em um fundo escuro de modo que os valores de intensidade da imagem estejam agrupados em dois grupos dominantes, como na figura 2.1. Uma maneira simples de extrair os objetos dessa imagem, ou ainda definir uma ROI, é encontrar um valor, ou limiar, T , que separa estes grupos de modo que a imagem segmentada $g(x,y)$ seja dada por:

$$g(x,y) = \begin{cases} 255, & \text{se } g(x,y) \geq T \\ 0, & \text{se } g(x,y) < T \end{cases} \quad (2-1)$$

Os valores 255 e 0 são referentes ao branco e preto respectivamente no sistema de cores RGB que normalmente varia de 0 a 255. Também pode ser usado 1 e 0 a fim de representar os objetos, onde 1 representa o branco (255) também chamado de objetos de frente da imagem enquanto o 0 representa o fundo da imagem.

O modelo usado neste trabalho é conhecido como limiarização global, pois o valor de T é constante e aplicado a toda a imagem. Existem várias maneiras de estabelecer esse valor T . Neste trabalho este valor é escolhido por meio da função:

$$T = \frac{1}{2}(m_1 + m_2) \quad (2-2)$$

Onde m_1 e m_2 são os valores de intensidade médios dos grupos separados por um T_i , neste caso calculado através da média entre o maior e o menor nível de intensidade presente na imagem. Em outras palavras, usamos a equação 2-1 para separarmos a imagem em dois grupos através de um limiar inicial escolhido da média entre o maior e o menor valor de intensidade. Uma vez separados esses grupos, calculamos os valores de intensidade médios de cada grupo, estes serão os valores de m_1 e m_2 . Após calcularmos m_1 e m_2 calculamos T a partir dos mesmos e comparamos esse novo T com o antigo T , também chamado aqui de T_i . Este processo iterativo é repetido até que $T_i = T$.

2.2 Análise de Imagens

Como dito na seção 2.1 o foco deste trabalho é emular a visão humana, fazendo inferências a partir de entradas visuais, isto é, a partir de uma imagem digital de entrada tomar uma decisão. As principais funções necessárias para realização desta tarefa são:

- **Representação e Descrição:** O processo de representação utiliza estruturas para armazenar e manipular os objetos extraídos de uma imagem. Enquanto o processo

de descrição extrai características que possam ser utilizadas na discriminação entre classes de objetos.

- Reconhecimento e interpretação: O processo de reconhecimento identifica os objetos da imagem, baseado nas características extraídas. O processo de interpretação consiste em atribuir um significado ao conjunto de objetos reconhecidos, isto é, classifica-los.

Na seção 2.2.1 será detalhado o processo de extração de características enquanto a seção 2.3 focará no processo de reconhecimento.

2.2.1 Representação e Descrição

Uma vez que uma imagem tenha sido processada, o próximo passo é encontrar uma maneira de representar suas características, externas, como suas fronteiras e bordas, ou internas, através dos pixels que compõem a imagem. Escolher a maneira de representar a imagem é apenas uma parte do processo. Uma vez determinada a maneira é preciso encontrar uma forma de descrever a imagem baseada na representação escolhida.

Uma representação externa deve ser utilizada quando o foco é encontrar características relacionadas a formas presentes na imagem, como por exemplo encontrar círculos presentes na imagem. Enquanto uma representação interna tem como foco encontrar propriedades específicas, como cor ou textura. Em alguns casos é necessário usar os dois modelos de representação[14].

A representação externa é baseada em algoritmos de detecção de bordas, como a limiarização descrita na seção 2.1.1, que é capaz de criar bordas separando objetos através da divisão binária onde certos pixels recebem o valor 255 e outros o valor 0 (branco e preto respectivamente). O foco deste trabalho reside na representação interna, onde a imagem é descrita através de suas cores ou sua textura.

A representação de imagens através da cor é uma ferramenta poderosa. Rodrigues[34], por exemplo, propôs um método de classificação de cachaça que identifica os tipos de madeira e o tempo de envelhecimento utilizando algoritmos de visão computacional e reconhecimento de padrões. Os atributos entrada do algoritmo são informações do modelo de cores RGB. O problema deste tipo de abordagem, é que imagens médicas, como tomografias e radiografias são em tons de cinza, logo este tipo de abordagem enfrentaria dificuldades. Com base nisso o foco foi na textura das imagens.

Os descritores de textura são uma importante abordagem para representar uma imagem. Textura pode ser definida como um conjunto de determinados padrões que podem apresentar um formato repetitivo ou não nas imagens. Encontrar padrões ajuda na interpretação de uma imagem. A análise de textura trabalha através de propriedades da imagem, como rugosidade, suavidade e regularidade.

A noção de textura parece depender de três elementos:

- Alguma ordem local é repetida sobre uma região que é maior que o tamanho dessa ordem.
- A ordem consiste do arranjo não-aleatório de partes elementares.
- As partes são entidades uniformes com as mesmas dimensões dentro da região da textura. [16]

A análise de desenvolvida neste trabalho deriva do artigo de Haralick[15] e será abordada na seção 2.2.2.

2.2.2 Haralick

O Trabalho de Haralick[15] foi desenvolver um conjunto de características que tornaria possível classificar uma imagem com base nelas. Com base nisso, foi natural pensar em um modelo que usasse 3 elementos padrões que a visão humana utiliza, que são as características espectrais, contextuais e texturais.

As características espectrais estão relacionadas as bandas visíveis e invisíveis do espectro eletromagnético. Enquanto as texturais tem informações sobre a distribuição das características espectrais. Já as contextuais são a respeito da área a qual a imagem a ser analisada pertence. Segundo Haralick[15] quando são imagens em tons de cinza que serão analisadas independentemente por uma máquina, então a textura e os tons de cinza são as áreas mais importantes.

O descritor de textura sugerido por Haralick[15] tem como base assumir que a informação a respeito da textura de uma imagem I qualquer está contida na relação espacial global que os tons de cinza têm uns com os outros. Isto significa que as informações são adequadamente especificadas pelo conjunto de matrizes de co-ocorrência dos tons de cinza.

Esta matriz tem dimensões 256×256 para imagens com 256 tons de cinza e cada célula funciona como um contador, isto é, a célula (i, j) armazena quantas vezes a cor i e a cor j estão presentes na imagem em pixels separados por uma distância d . Neste trabalho será definida $d = 1$. Com base na figura 2.2 podem ser feitas algumas observações:

- Os pixels 1 e 5 são os vizinhos mais próximos do pixel central (*) que tem orientação horizontal 0° .
- Os pixels 3 e 7 são os vizinhos mais próximos do pixel central que tem orientação 90° .
- Os pixels 4 e 8 são os vizinhos mais próximos do pixel central com orientação de 45° .
- Os pixels 2 e 6 são os vizinhos mais próximos do pixel central com orientação 135° .

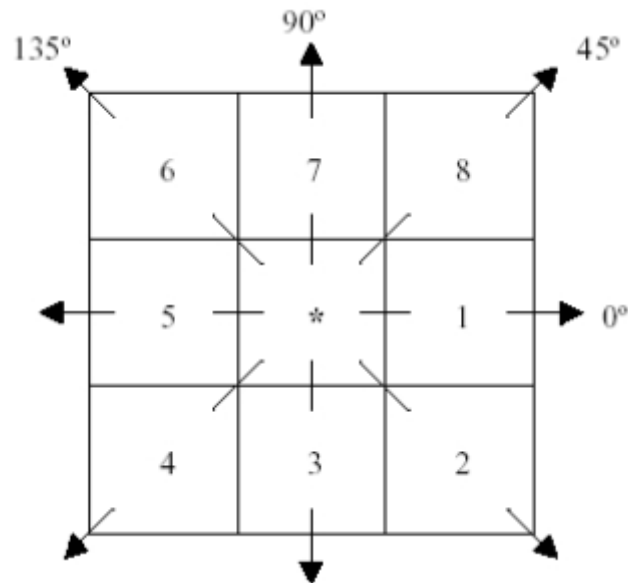


Figura 2.2: Funcionamento da Matriz de co-ocorrência.

Uma vez definidas as vizinhanças presentes na figura 2.2, A matriz de co-ocorrência será formada contando os pares em função do ângulo analisado (0°, 45°, 90° ou 135°) e distância. Considere a imagem presente na figura 2.3.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figura 2.3: Imagem com 4 níveis de cinza (0, 1, 2 e 3).

Com base no fato da figura 2.3 ter apenas 4 níveis de cinza, serão necessárias 4 matrizes (0°, 45°, 90° e 135°) com dimensões 4x4 a fim de computar as vizinhanças. Na figura 2.4 pode ser vista as 4 matrizes geradas utilizando como base a figura 2.3.

Uma vez calculadas as matrizes de co-ocorrência, é necessário calcular as características que representarão a imagem, essas equações serão detalhadas na seção 2.2.3. Tais equações são sugeridas por Haralick[15] e extraídas de cada uma das matrizes de co-ocorrência referentes a cada imagem que venha a ser analisada.

2.2.3 Equações Utilizadas

A seguir serão mostradas as equações que representam as imagens e que serão utilizadas para determinar a semelhança entre cada uma.

$$\begin{aligned}
 m_{0^\circ} &= \begin{bmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} & m_{45^\circ} &= \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\
 m_{90^\circ} &= \begin{bmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} & m_{135^\circ} &= \begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}
 \end{aligned}$$

Figura 2.4: Matrizes de co-ocorrência geradas nos 4 ângulos.

Coefficiente de variação: Tem como foco analisar o peso do desvio-padrão em relação à média dos valores dos pixels das imagens. Dado pela equação:

$$CV = \frac{\sigma}{\mu} \quad (2-3)$$

Onde σ é o desvio padrão e μ é a média dos coeficientes.

Contraste: O contraste caracteriza-se pela diferença entre os tons de cinza. Baixo contraste ocorre quando há pequena diferença entre os níveis de cinza dos pixels localizados em uma região contígua da imagem. É dada pela equação:

$$Contraste = \sum_i \sum_j (i - j)^2 * p(i, j) \quad (2-4)$$

Onde $p(i, j)$ é (i, j) -ésima probabilidade de ocorrência de um nível de cinza na imagem.

Correlação: A correlação mede a dependência linear entre os tons de cinza presentes em uma imagem. Valores altos de correlação dados pela equação, próximos de um, indicam a existência de relação linear entre os pares de níveis de cinza.

$$Correlacao = \frac{\sum_i \sum_j (ij) * p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (2-5)$$

Onde $p(i, j)$ é (i, j) -ésima probabilidade de ocorrência de um nível de cinza na imagem e $\sigma_x, \sigma_y, \mu_x, \mu_y$ são os desvios padrão e médias em x e y respectivamente.

Desvio médio: O Desvio Médio é uma medida da dispersão dos dados em relação à média de uma sequência, o "afastamento" em relação a essa média. Esta medida representa a média das distâncias entre cada elemento da amostra e seu valor médio.

Descrito pela equação:

$$DesvMedio = \frac{\sum_i \sum_j |p(i, j) - \bar{x}|}{N} \quad (2-6)$$

Onde N é o número total de pixels na imagem e \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Variância: Descreve quanto os valores de tons de cinza estão dispersos em torno da média. É dada pela equação:

$$V = \frac{\sum_i \sum_j |(i, j) - \bar{x}|}{N} \quad (2-7)$$

Onde V é a variância de $P(i, j)$, isto é, a variância no pixel (i, j) .

Diferença da variância:

$$difVar = V(p(x - y)) \quad (2-8)$$

$p(x - y)(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = |i - j|$.

Diferença da entropia:

$$difEntropia = -\sum_i p(x - y)(i) \log(p(x - y)(i)) \quad (2-9)$$

Energia: Também conhecida como segundo momento angular, expressa a uniformidade de uma textura, assumindo o valor máximo 1 e é expressa pela equação:

$$Energia = \sum_i \sum_j p(i, j)^2 \quad (2-10)$$

Energia média: É a energia da imagem dividida pelo número total de pixels. Dado pela equação:

$$EnergiaMedia = Energia/N \quad (2-11)$$

Entropia: A medida de entropia expressa a desordem contida na textura. Quando uma imagem não é uniforme, as entradas apresentam valores próximos de zero o que resulta em valores altos de entropia. É dada pela equação:

$$Entropia = \sum_i \sum_j p(i, j) \log(p(i, j)) \quad (2-12)$$

Momento de diferenças inversas: Quando a concentração dos valores na diagonal da matriz de co-ocorrência for máxima, o Momento da Diferença Inverso (MDI) atinge o valor máximo. É dado pela equação:

$$MDI = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (2-13)$$

Resíduo médio: É a diferença média entre os dados observados e os valores ajustados pelo modelo (calculados). É dado pela equação:

$$ResMedio = \sum_i \sum_j ((i - j) - \bar{x})^2 \quad (2-14)$$

Onde \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Soma média:

$$SumMedia = \sum_i i p_{(x+y)}(i) \quad (2-15)$$

$p_{(x+y)}(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = i + j$.

Somatório da entropia:

$$SumEntropia = \sum_i p_{(x+y)}(i) \log(p_{(x+y)}(i)) \quad (2-16)$$

Somatório da variância:

$$SumVar = \sum_i (i - SumEntropia)^2 p_{(x+y)}(i) \quad (2-17)$$

Suavidade:

$$SV = 1 - \frac{1}{1 + \mu^2} \quad (2-18)$$

Desvio padrão: É a raiz quadrada da variância, dado pela equação:

$$\sigma = \sqrt{V} \quad (2-19)$$

Uma vez que todas as equações que representam as características tenham sido computadas, é necessário passar à segunda fase, que é reconhecer padrões e classificar as imagens com base nelas. Na seção 2.3 será detalhado este processo.

2.3 Reconhecimento de Padrões

O reconhecimento de padrões é o ato de tomar uma decisão baseada em dados de entrada. Dentro da proposta deste trabalho seria de acordo com as características extraídas de uma imagem, classifica-la de acordo com as classes existentes.

Isto ocorre, pois, o conjunto de dados é particionado em grupos, separados por características específicas, de maneira que objetos dentro de um grupo são semelhantes entre si ao passo que diferem de objetos de outro grupo. O que deve ser feito então é construir um modelo que saiba dizer a qual dos grupos existentes um objeto pertence, isto é, classifica-lo.

O processo é dividido em duas etapas principais: a primeira é a extração de características que representam a imagem, discutida na seção 2.2; a segunda que será detalhada nesta seção, é a construção de um modelo que seja capaz de classificar as amostras.

Segundo Duda[10], para os tipos de objetos a classificar (ou descrever) podemos utilizar algumas abordagens como:

- Abordagem estatística - A abordagem clássica, também chamada de "Teoria da Decisão". Assume que as características das classes são conduzidas por determinados modelos de probabilidade
- Abordagem sintática - Procura descrever a estrutura dos padrões usando inter-relações de características básicas também chamadas de primitivas.
- Abordagem neuronal - Abordagem tipo "caixa preta" que procura determinar um mapeamento ótimo entre entradas e saídas baseada em modelos de neurônios.
- Abordagem difusa - Usa a teoria dos conjuntos difusos para modelizar grau de incerteza inerente a características e a classificações.

O modelo a ser adotado para reconhecimento de padrões neste trabalho é conhecido como K-vizinhos mais próximos (*K-Nearest Neighbors* - KNN) e será melhor detalhado na seção 2.3.1

2.3.1 K-vizinhos mais próximos (*K-Nearest Neighbors* - KNN)

O KNN foi introduzido pela primeira vez por Fix[13] em seu estudo de análise discriminante. É um algoritmo de aprendizagem lazy não-paramétrico, isto é, não é computada uma função de classificação, usando a chamada "aprendizagem baseada em instância".

O KNN classifica novas amostras de acordo com o relacionamento da nova amostra com o conjunto de dados [27]. Esse relacionamento é representado como pontos em um espaço euclidiano, ou seja, a distância entre eles. A formulação do método KNN exige a definição de uma métrica de distância e a quantidade de vizinhos k que será usado na classificação. A distância euclidiana é comumente utilizada, enquanto que k é definido empiricamente[2][19][12]. Outras métricas podem ser usadas além da distância Euclidiana, na seção 2.3.3 será discutido a respeito das métricas de distância.

Uma vez calculada a distância entre a nova amostra e as amostras já conhecidas presentes na base, a nova amostra será classificadas como pertencente a classe à qual ela tenha k vizinhos mais próximos. Na figura 2.5 pode ser visto alguns exemplos de KNN.

Nesse caso, a amostra X será classificada como pertencente a classe A, pois a maioria dos k vizinhos mais próximos são pertencentes a classe A.

De acordo com Bhatia[5] e Miloud-Aquidate[26] as vantagens do KNN são:

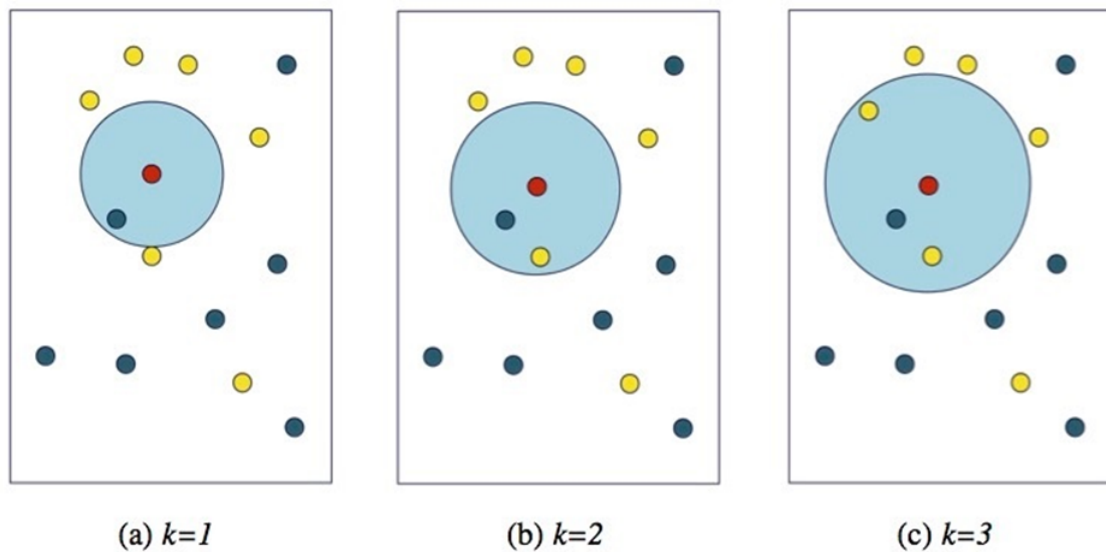


Figura 2.5: Exemplos do KNN. (a) $K = 1$ a nova amostra pertence a classe azul. (b) $K = 2$ continua a classe azul devido a proximidade maior. (c) $K = 3$ Passa a pertencer a classe amarela, já que são duas amarelas mais próximas e apenas uma azul.

- Implementação muito simples;
- Ótimo em conjuntos de treino com dimensão elevada;
- Usa informação local, o que pode facilmente levar à implementação de comportamentos adaptativos;
- Implementação facilmente passível de paralelização.

Já suas desvantagens consistem em:

- Exige muito espaço de armazenamento para o conjunto de treino;
- Computacionalmente muito exigente,
- Altamente susceptível à curva da dimensionalidade, onde as dimensões maiores exigem um número de exemplos exponencialmente maior no conjunto de treino

Além desses fatores, a escolha do k também influencia no resultado geral. Neste trabalho foi definido o $k = 1$, ou seja, apenas o vizinho mais próximo é levado em consideração. A escolha deste valor se deve ao fato de valores altos de k terem um custo computacional alto. Outro fator levado em consideração é o fato do KNN exigir muito espaço de armazenamento do conjunto de treino, por isso é usado a técnica de redução de dimensionalidade de análise de componentes principais, com isso ao invés de 17 características a serem calculadas, apenas o número de componentes resultantes será levado em consideração no cálculo de distância. A análise de componentes principais (*Principal Component Analysis* - PCA) será discutida na seção 2.3.2.

2.3.2 Análise de Componentes Principais (*Principal Component Analysis - PCA*)

Análise de Componentes Principais (PCA) é uma técnica que analisa uma tabela de dados em que as observações são descritas por variáveis dependentes intercorrelacionadas quantitativas. Seu objetivo é extrair as informações importantes da tabela a fim de representá-las como um conjunto de novas variáveis ortogonais chamadas de componentes principais[1].

Sua origem pode ser associada a Pearson[31], como também Cauchy[7], ou ainda Jordan[20] mas sua instanciação moderna foi formalizada por Hotelling[18], que criou o termo componente principal (*principal component*).

A partir de p variáveis $x_1, x_2, x_3, \dots, x_p$, obtêm-se as variáveis não correlacionadas $y_1, y_2, y_3, \dots, y_p$ através de uma transformação linear que maximiza a variância dos dados, onde y_i é a i -ésima componente principal. As variáveis y , também chamadas de componentes principais, são ordenadas de forma que $v(y_1) > v(y_2) > v(y_3) > \dots > v(y_p)$ e $v(y)$ é a variância.

Os passos para calcular as componentes principais são:

- Obter os dados ou as M amostras de vetores de dimensão n ;
- Calcular a média ou o valor médio destes dados;
- Subtrair a média de todos os itens de dados;
- Calcular a matriz de covariância S usando todas as subtrações. Ela é o resultado da média do produto de cada subtração por ela mesma e terá dimensão $n \times n$;
- Calcular os auto valores λ e auto vetores w da matriz de covariância.
- Arranjar a matriz da Transformada de Hotteling (cujas linhas são formadas a partir dos auto vetores da matriz de covariância arranjados de modo que a primeira linha, o elemento $(0,0)$, seja o auto vetor correspondente ao maior autovalor, e assim sucessivamente até que a última linha corresponda ao menor autovalor.

Seja S a matriz de covariância das amostras M , w os auto vetores e λ os auto valores, tem-se que:

$$Sw = \lambda w \quad (2-20)$$

O auto vetor w com o maior auto valor λ associado, corresponde à componente principal do conjunto de dados usado. Mantendo a ordem $\lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_p$. Na 3.4 pode ser visto este processo.

A qualidade do modelo de PCA pode ser avaliada usando técnicas de validação cruzada (cross-validation) tais como o bootstrap e leave-one-out. Na seção 2.4 serão detalhadas as técnicas de validação.

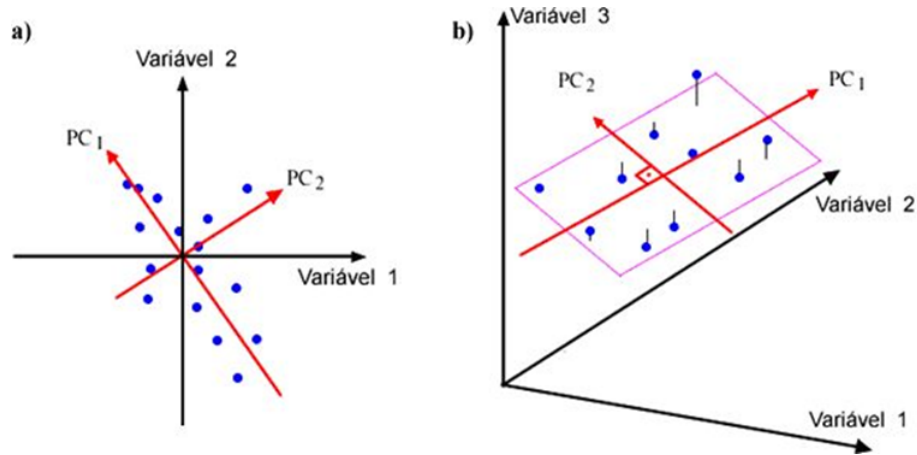


Figura 2.6: Exemplos de PCA. (a) Números de componentes igual ao das variáveis iniciais. (b) Projeção de 3 variáveis em duas componentes.

2.3.3 Métricas de Distância

A fim de determinar a semelhança entre imagens, é necessário que se use algum tipo de função que seja capaz de calcular a distância entre as duas, isto é, o qual próximas cada uma das características que representam a imagem estão uma das outras.

Para isso é preciso utilizar algum tipo de função que calcule a distância d entre as imagens x e y . Neste caso a distância entre cada uma das 17 características que compõem a imagem, ou a distância entre as componentes principais calculadas.

Existem várias maneiras de se realizar tal cálculo, como a distância Euclidiana d_e que dá o comprimento de um segmento de reta entre dois pontos [9] dada pela equação:

$$d_e(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_n - y_n)^2} \quad (2-21)$$

O problema em usar esse tipo de cálculo é a complexidade e o custo computacional que é exigido, pois há várias potências e uma raiz quadrada. Considerando um cenário com uma quantidade alta de imagens, gerariam um tempo de execução extenso que diminuiria o desempenho do classificador. Com isso em mente é necessário procurar uma alternativa.

Uma maneira de resolver esse problema é usando a geometria do táxi, também chamada de distância Manhattan, que consiste apenas de somas e subtrações, na equação 2-22 pode ser visto o mesmo cálculo de distância empregado na equação 2-21 dessa vez usando a métrica de Manhattan[9].

$$d_m(x, y) = |x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + \dots + |x_n - y_n| \quad (2-22)$$

Esta métrica, diferente da distância Euclidiana, dá uma aproximação do valor real de distância, na figura 2.7 pode ser visto a diferença entre as duas.

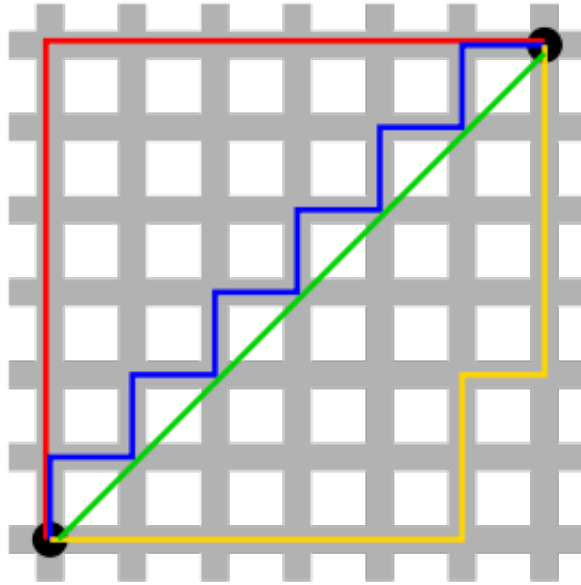


Figura 2.7: Diferença entre distância Euclidiana (linha verde) e distância Manhattan (linhas vermelha, azul e amarela).

Uma vez calculada a distância entre a imagem de partida e todas as outras é possível definir a semelhança, isto é, quais imagens estão mais próximas da imagem parâmetro. Na figura 2.8 pode ser visto 3 exemplos de imagens, X, Y e Z com suas respectivas características. Se os valores de suas características fossem aplicados na equação 2-22 $d_m(x,y)$ e $d_m(x,z)$ os valores seriam 32 e 120 respectivamente, o que significa que a imagem mais semelhante a imagem X é a imagem Y.

2.4 Metodos de Validação

Uma vez que todos os métodos tenham sido implementados e o classificador esteja funcional, é necessário validar seus resultados, isto é, mostrar que os resultados estão corretos.

Existem vários problemas que podem ocorrer, por exemplo: um algoritmo que tem 90% de acerto. À primeira vista uma taxa de acerto de 90% parece algo bom, mas pode não ser o caso. Imagine que existam 100 amostras, 90 da classe A e 10 da classe B, se o algoritmo classificar todas as amostras como sendo da classe A, automaticamente ele terá 90% de acerto, mas ao mesmo tempo terá 100% de erro nas amostras da classe B.

O cenário ideal é que não haja *overfitting*, isto é, o classificador se adapte demais a um determinado conjunto de dados avaliando ele com alta eficiência, mas com desempenho comprometido com qualquer outro conjunto que seja diferente do primeiro. Principalmente quando o objetivo é criar um classificador genérico que deve se adaptar a vários conjuntos de imagem, como é o caso deste trabalho.

	Imagem X	Imagem Y	Imagem Z
Característica 1	5	5	2
Característica 2	6	8	0
Característica 3	8	7	15
Característica 4	5	4	13
Característica 5	10	11	5
Característica 6	9	11	0
Característica 7	8	6	19
Característica 8	7	5	16
Característica 9	2	0	14
Característica 10	1	3	0
Característica 11	10	12	2
Característica 12	15	19	3
Característica 13	18	16	8
Característica 14	2	0	7
Característica 15	3	4	9
Característica 16	4	3	6
Característica 17	6	1	0

Figura 2.8: Exemplo de 3 imagens e suas respectivas características.

Uma maneira de se fazer isso é utilizando validação cruzada (*cross-validation*). Validação cruzada é uma técnica que visa avaliar o quanto um classificador consegue generalizar a partir de um conjunto de dados. A ideia principal da técnica consiste em dividir um conjunto M de amostras em N novos conjuntos, mutuamente exclusivos, aleatoriamente distribuídos. Em um determinado momento um desses conjuntos é utilizado para validação do classificador enquanto os outros são utilizados para realizar o aprendizado do mesmo. Este processo é repetido várias vezes, em cada uma delas um novo conjunto é utilizado para validação dos dados, ao final o resultado é a taxa de acerto médio dos K testes realizados, por exemplo: São realizados 5 testes de validação com as respectivas taxas de acerto, 75%, 80%, 65%, 60%, 90%, o resultado final do algoritmo será de 74% de acerto.

Um exemplo de validação cruzada é o *k-fold* onde o conjunto M é dividido em k novos conjuntos seguindo o mesmo princípio básico de um para testes e os outros para aprendizado. Neste trabalho foi empregado uma variação do *k-fold* chamado de *leave-one-out*. No *leave-one-out* $k = M$, isto é, cada conjunto N criado tem exatamente uma amostra. Na figura 2.9 pode ser visto como ocorre este processo.

O método *leave-one-out* tem como vantagem o fato de investigar todos os dados individualmente garantindo assim a confiabilidade do método proposto. Em contrapartida, tem alto custo computacional já que cada amostra será testada contra todas as outras.

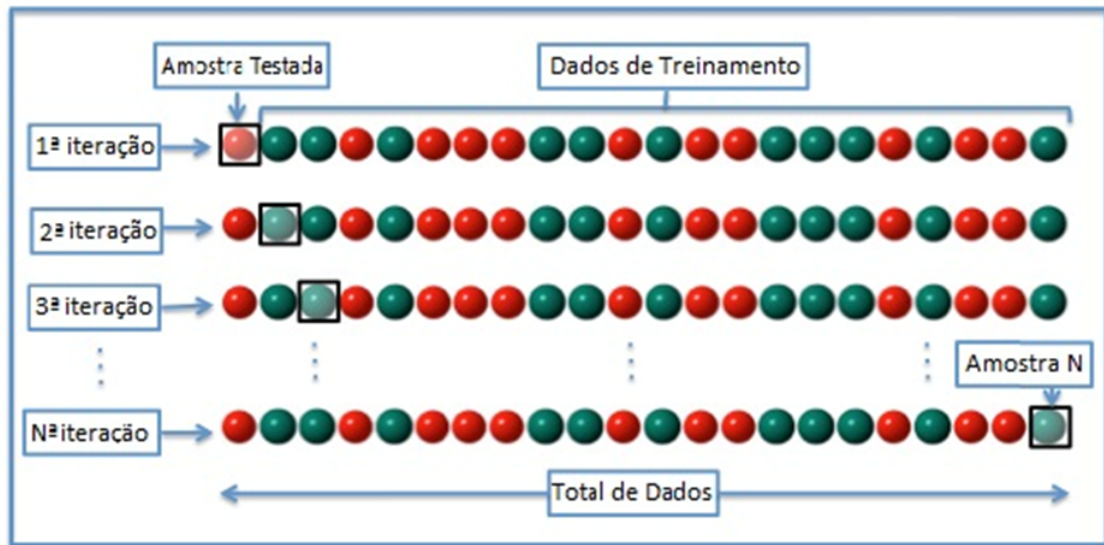


Figura 2.9: Processo de validação do método leave-one-out.

Devido ao uso do limiar textual que será detalhado no capítulo 3 esse problema é amenizado.

Metodologia

Este capítulo é dedicado a mostrar como as técnicas discutidas no capítulo 2 interagem de maneira a produzir os resultados propostos, além de discorrer a respeito de problemas enfrentados neste processo.

A primeira parte deste capítulo é dedicada a mostrar como ocorre o armazenamento das imagens, além de ressaltar a importância de uma boa armazenagem no desenvolvimento de um sistema CAD.

A segunda parte foca na parte de recuperação e classificação das imagens, mostrando como é feito o processo de definição do limiar, a redução do escopo de buscas e por último a classificação baseada na semelhança, a fim de determinar o sucesso ou falha de cada busca.

Por fim, são detalhados algumas particularidades e dificuldades enfrentadas pelo modelo proposto, deixando sugestões de melhoras que podem ser implementadas.

3.1 Armazenamento

Yuan[42] afirma ser fundamental para sistemas CAD e de CBIR que a base de imagens seja definida de forma eficiente capaz de guardar informações relevantes das imagens para posterior recuperação. Logo o armazenamento é parte fundamental deste trabalho.

Em se tratando das imagens propriamente ditas existem 3 tipos de informações associadas a elas que são importantes e devem ser armazenadas:

- Informações independentes de conteúdo: informações que não estão presentes na imagem, mas são ligadas a ela, como nome, data, etc.
- Informações descritivas de conteúdo: informações descritivas da imagem do ponto de vista semântico, por exemplo, se é um exame do coração, o tipo de exame realizado, etc.
- Informações dependentes de conteúdo: informações do conteúdo físico da imagem, como cor, textura, forma, etc.

Cada uma dessas informações tem importância quando se fala de sistemas CAD, por exemplo: Ao saber a data que uma imagem foi adicionada a base, tem se ideia do tempo que um paciente apresenta o problema descrito na imagem. Além disso, as informações descritivas de conteúdo têm papel crucial no uso do limiar textual que será apresentado na seção 3.2.1.

Outro fator importante ao armazenar várias informações relacionadas a imagem, é para o auxílio de pacientes que tenham problemas parecidos com os presentes na base, isto é, ao recuperar a imagem mais semelhante a imagem parâmetro, também será recuperado todas as informações anexadas a ela, que podem incluir diagnóstico e tratamento utilizado.

Os bancos de dados contêm campos para tratar os arquivos multimídia, como as imagens, tornando o armazenamento uma etapa trivial[4][6][29][35]. Uma vez que o banco tenha sido definido corretamente de maneira que as informações importantes sejam armazenadas junto com a imagem, a grande dificuldade e foco dos principais trabalhos publicados atualmente está em como recuperar essas informações de maneira eficiente.

Para solucionar esta dificuldade, desenvolvemos uma metodologia que a cada vez que uma nova imagem é adicionada ao banco, todas as características, discutidas no capítulo 2, que serão utilizadas para recuperação serão extraídas, calculadas e armazenadas no banco, isto é feito para evitar que as mesmas operações sejam realizadas repetidas vezes, a cada nova busca, impactando diretamente na velocidade de recuperação das imagens. Além disso, é definido um limiar textual que permite delimitar o escopo de imagens que serão submetidas ao CBIR. Reduzindo assim a quantidade de comparações e o tempo de busca de cada execução, no capítulo 4 será mostrado os ganhos gerados pelo uso deste recurso.

3.2 Recuperação

Uma vez que as imagens estejam armazenadas na base, é necessária uma maneira de busca-las eficientemente. O método como as imagens são recuperadas é conhecido como busca por similaridade na qual se inclui técnicas de CBIR. Funciona usando uma imagem como parâmetro de busca para recuperar imagens que se assemelham a ela. A busca funciona através de 3 passos:

- Extração de Características: Primeiramente a imagem é analisada e as características de interesse são extraídas através de algum descritor.
- Indexação de Características (Vetor de Características): Uma vez extraídas, é montado um vetor de características com todas as informações extraídas (tais como cor, dimensão das estruturas encontradas, formas identificadas, etc.) da imagem.

- Comparação: Ao final é realizada a comparação entre os diversos vetores de características que foram extraídos das imagens, é nessa etapa que são definidas as imagens semelhantes que devem ser recuperadas.

Como mostrado no capítulo 2, as características de interesse são baseadas na textura da imagem, e o vetor de características é montado com base nas equações mostradas na seção 2.2.3 enquanto a comparação ocorre usando a métrica de distância Manhattan.

Pode ainda, de acordo com a imagem, ocorrer uma pequena variação na extração de características, por exemplo: As imagens de pulmão carregam dois vetores, um referente ao pulmão esquerdo e outra ao pulmão direito, ambos com as mesmas características; caso seja necessário compará-las com imagens com apenas um vetor, as dentárias, apenas o primeiro é comparado. Isto não ocorre devido ao limiar textual.

Mas apenas este processo não é suficiente para garantir a eficiência da técnica, e principalmente o desempenho. Uma vez que a comparação entre diversas imagens é algo com alto custo computacional. Uma simples inspeção sequencial pode ser aceitável em bases pequenas, mas em bases com milhões, ou talvez bilhões de imagens diferentes, um bom algoritmo de indexação é imperativo[23].

Uma maneira de tornar o método mais eficiente é mesclando buscas textuais com as técnicas de CBIR. Segundo Arampatzis[4] buscas textuais produzem resultados melhores do que buscas por similaridade. Arampatzis[4] sugere que ao combinar as duas produziriam melhores resultados e diminuiriam os custos computacionais. Usando esse conceito como base é criado uma etapa que precede o CBIR com um limiar textual, que reduz o número de imagens nas quais serão aplicados os cálculos para extração de características e cálculo de distância.

3.2.1 Limiar Textual

Enquanto imagens contém informações que não podem ser simplesmente descritas por palavras, descritores textuais são importantes para recuperação de resultados relevantes, mas em compensação contém pouca informação visual[41].

Uma maneira de contornar este problema, seria combinar as duas técnicas, de maneira ao resultado final ser uma fusão dos resultados da busca textual com o CBIR. Apesar de ser um método comprovadamente robusto, existem algumas questões importantes a respeito:

- O peso apropriado atribuído a cada modalidade e a definição da normalização/combinção dos dados não são tarefas triviais, podem exigir dados de treinamento.

- Se os resultados forem avaliados apenas por inspeção visual, o método não é formalmente bom; isto ocorre porque as vezes os resultados textuais podem piorar os resultados visuais que serão exibidos ao final.

Ao invés disso, o ideal é fazer uma abordagem que aplique esse processo em duas fases, ao invés de fundir as duas. Em outras palavras, a fim de reduzir o tamanho do conjunto de dados, um ranking pode ser definido através de uma característica específica ou um rank. Este processo também é conhecido como limiarização. Isso melhora a eficiência do CBIR ao reduzir operações computacionalmente caras. Três fatores determinam qual é o limiar correto:

- O número de itens relevantes no conjunto de dados
- A qualidade da classificação
- A medida que o limiar deseja otimizar [33]

No escopo deste trabalho, uma maneira de determinar este limiar seria definindo informações importantes no contexto do paciente, como por exemplo, o tipo de exame a ser usado na busca, como uma tomografia ou radiografia, ou ainda caso existisse, um grupo de risco que apresentaria problemas parecidos entre todos os pacientes, o que delimitaria ainda mais o escopo e aumentaria a qualidade e número de itens relevantes no conjunto de dados. Isso garantiria que não ocorreria comparações do tipo: Uma tomografia sendo comparada a uma radiografia. Na figura 3.1 tem-se um exemplo do funcionamento do limiar textual.

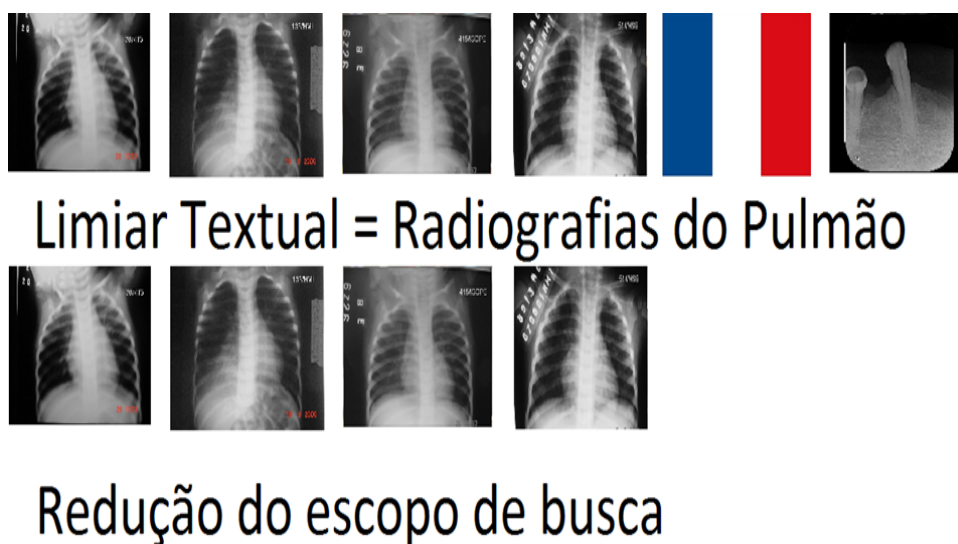


Figura 3.1: Funcionamento do limiar textual.

Uma vez que o limiar textual é aplicado, apenas as imagens que pertencem ao conjunto de interesse são submetidas as técnicas de CBIR descritas no capítulo 2. Apenas esse processo já consegue produzir uma melhora na eficiência do método proposto, mas

existe ainda uma fase intermediária entre esta etapa e a aplicação do CBIR que visa melhorar ainda mais a eficiência do método proposto. Este processo será mostrado na seção 3.2.2.

A maneira como o limiar textual é aplicado varia de acordo com a forma de armazenamento das imagens, por exemplo: Se for uma estrutura de diretórios, o limiar deve definir a pasta em qual procurar as imagens; caso seja em uma tabela de um banco de dados, basta criar um campo de tipo que difere entre as imagens e fazer a busca baseado nesse campo.

3.2.2 Grupo de Imagens e Classificação

A partir do momento que o limiar textual é capaz de definir o grupo de imagens, deve-se agora aplicar o CBIR para determinar dentro daquele grupo quais as imagens mais semelhantes entre si. Neste momento entra um passo intermediário que tem como objetivo aprimorar os resultados e melhorar a eficiência do método proposto.

Ao trabalhar com a ideia de um classificador genérico deve-se levar em conta que ele trabalhe com qualquer tipo de imagem, e isto normalmente diminui sua eficácia como mostrado na seção 1.3. O cenário ideal seria então, que de alguma maneira ele pudesse se adaptar ao grupo de imagens que ele deve classificar de maneira dinâmica.

Com base nesta premissa, é usada a combinação limiar textual + PCA, isto é, uma vez que o limiar textual reduz o conjunto de imagens, criando um subconjunto com apenas imagens do mesmo tipo em relação a qual será usada como parâmetro de busca, o PCA é então calculado com base nesse subconjunto. Uma vez que o PCA é calculado usando um subconjunto com apenas imagens do mesmo tipo da imagem de interesse, é garantido que as componentes geradas terão o máximo de representatividade, e irão gerar bons resultados. A quantidade de componentes utilizada é discutida no capítulo 4.

Uma vez calculado o PCA e definido a quantidade de componentes a ser utilizada, apenas essas componentes serão usadas no cálculo de distância, e a partir daí a imagem mais semelhante ($k = 1$) será definida como resultado e servirá como base para a classificação da imagem parâmetro.

À primeira vista, pode parecer um processo que não produziria melhoras significantes, mas existem particularidades em cada imagem, por isso apenas compará-las entre si não seria eficiente. As imagens do pulmão por exemplo, tem sua ROI definida dinamicamente através do uso de limiarização binária.

Na seção 3.3 serão discutidos alguns detalhes específicos do sistema, como essas diferenças entre as imagens, e o impacto que elas geram a eficiência e desempenho do mesmo.

3.3 Detalhes do Sistema

Cada classe de imagem tem suas particularidades, e num sistema que visa classificar imagens, essas características devem ser levadas em conta. Como no caso de radiografias de pulmão, que podem ter infiltrações em cada pulmão separadamente, e por isso, é necessário olhar cada um deles individualmente.

Este processo é feito para que o método seja capaz de analisar pulmão esquerdo e direito separadamente. Primeiramente, a imagem é dividida ao meio, gerando lado esquerdo e lado direito, após isso, é necessário selecionar apenas o interior do pulmão, que é onde estão as informações que definirão se a imagem tem ou não pneumonia.

O segundo passo é então, separar o interior do pulmão, visivelmente mais escuro dos ossos, parte mais clara. A limiarização cumpre esta etapa. Na figura 3.2 pode ser visto o resultado deste processo.



Figura 3.2: *Imagem Limiarizada.*

Uma vez limiarizada a imagem é possível determinar a altura onde se encontra o pulmão a cada nova imagem, descartando zonas sem informação. Na figura 3.3, pode ser visto o resultado deste processo.

Este processo é feito cada vez que uma nova imagem deve ser armazenada na base, e difere a cada nova imagem. Por exemplo imagens dentárias, tem como ROI a própria imagem, uma vez que os dentes ocupam toda a imagem. Na figura 3.4 pode ser vista esta diferença.

Apesar desse pré-processamento ser feito no armazenamento, estas informações impactam diretamente na recuperação. Com base nisso, é reforçado a motivação do uso do limiar textual juntamente com o PCA. Se o PCA for usado num subgrupo com imagens completamente diferentes entre si, sua eficiência será prejudicada.

No momento o sistema ainda funciona usando apenas uma estrutura básica de diretórios com poucas informações associadas a cada imagem, mas se mostra capaz de

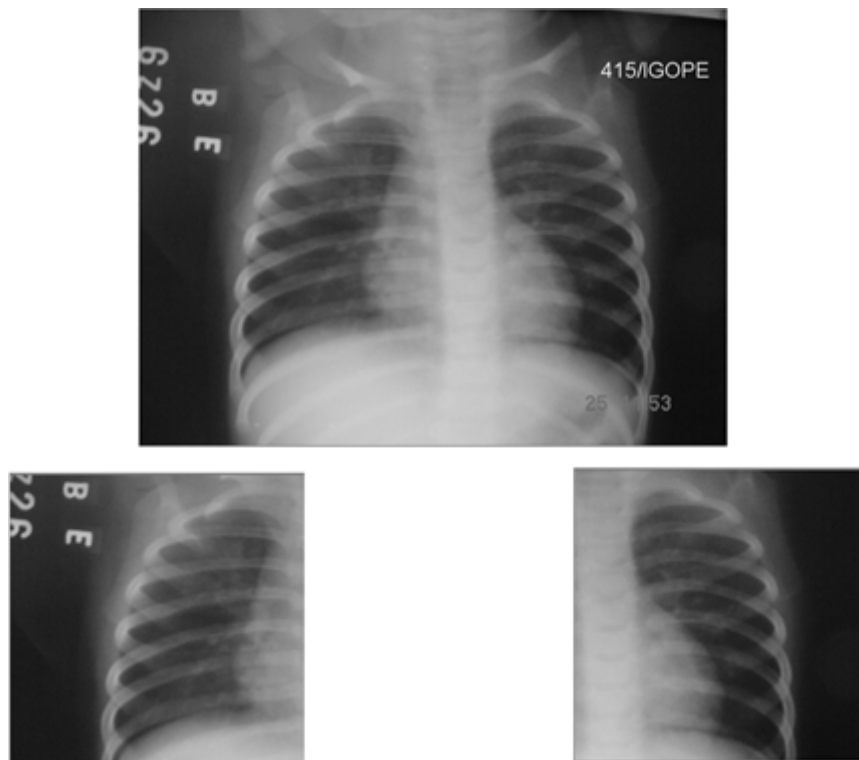


Figura 3.3: *Áreas de interesse extraídas da imagem original.*

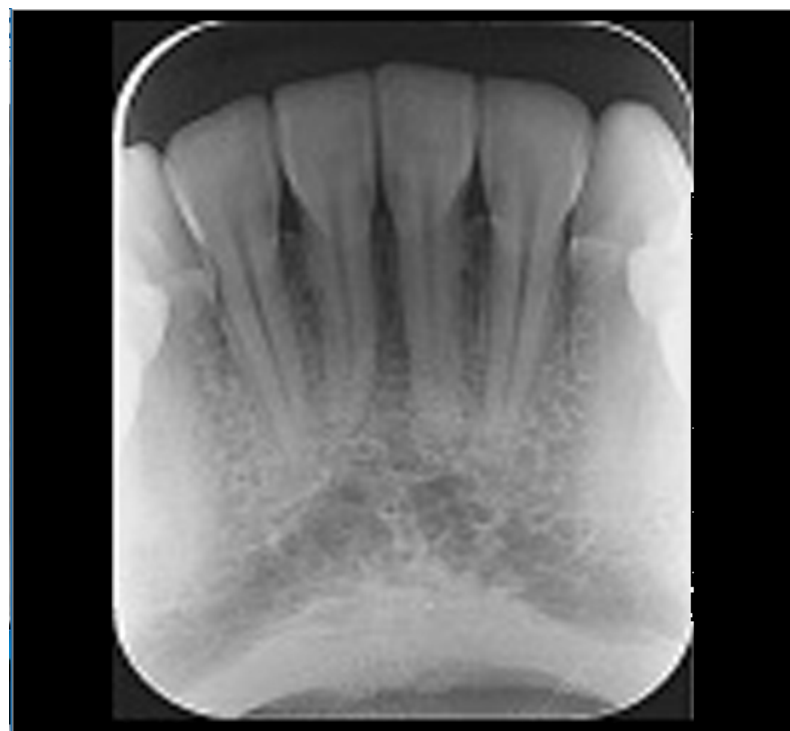


Figura 3.4: *Exemplo de radiografia Periapical.*

classifica-las e recuperá-las de maneira eficiente, além de toda e qualquer informação associada a imagem recuperada durante a busca.

O método proposto permite dois tipos de busca, uma pelo próprio usuário, caso já existam exames dele na base, ele é capaz de recuperá-los. Esse método é chamado de busca pela própria imagem. Além disso, ele também é capaz de procurar pelo exame mais semelhante presente na base ao usado como parâmetro, descartando cópias caso existam. Este segundo método serve como base para classificação.

Resultados

Neste capítulo serão mostrados os resultados obtidos através dos testes realizados com dois tipos diferentes de imagem: Radiografias de pneumonia e Radiografias periapicais.

Em todos os casos foram utilizados testes com cross-validation usando o método leave-one-out devido aos conjuntos de imagens serem limitados. Além disso foi convenção usar 2 componentes nos testes, pois foi o valor que teve a menor variância entre as classes (Pulmão e dentes) ao mesmo tempo que teve o melhor resultado médio com uma taxa de acerto geral 90,25%

Na seção 4.1 serão discutidos os resultados referentes as radiografias de pneumonia, enquanto a seção 4.2 será destinado as radiografias periapicais. Por fim a seção 4.3 será feita uma consideração final a respeito dos resultados obtidos nos dois tipos de imagem.

4.1 Radiografias de Pneumonia

A primeira etapa, foi testar o algoritmo implementado para recuperação das radiografias de pneumonia. Nesta etapa foram usadas 110 imagens radiográficas além das 192 imagens dentro da base com conteúdo completamente diferente, neste caso as radiografias periapicais.

Uma vez que em 100% dos casos foi possível identificar a imagem de busca dentro das bases, a etapa seguinte foi testar usando cross-validation as outras imagens contidas dentro da base e verificar se o diagnóstico presente na imagem de busca era o mesmo na imagem resultante. Foi utilizado o método leave-one-out para realizar os testes.

A fim de validar os resultados, todas as imagens radiográficas que continham algum grau de pneumonia foram avaliados pelo médico Altino V. De Rezende Filho, cirurgião geral, CRM-GO 8156. As imagens foram classificadas de 2 formas, primeiro com o pulmão com o maior nível de infiltração: esquerdo, direito, ou ambos.

Depois foi determinado o grau de infiltração do pulmão, avaliado no item anterior, classificada como: leve, média ou grave Não confundir nível de infiltração

com nível de pneumonia, uma infiltração grave pode ter uma pneumonia fácil de tratar, enquanto uma infiltração leve, pode ser um caso muito mais complexo. Os resultados obtidos foram comparados a sua avaliação.

A tabela 4.1 mostra a taxa de acerto do algoritmo em relação a classificação realizada pelo médico:

Tabela 4.1: *Taxa de acerto na recuperação de imagens.*

Tipo de Busca	Taxa de Acerto
Taxa de Acerto na busca da própria imagem	100%
Taxa de Acerto do método leave-one-out (PCA = 2)	90,91%

Como mostrado na tabela 4.1 em 90,91% dos casos, a imagem recuperada, tinha o mesmo diagnóstico, ou uma pequena variação do mesmo, por exemplo, mesmo lado de pulmão infiltrado, e no máximo um nível de diferença no grau de infiltração.

Na tabela 4.2 é mostrado a comparação dos resultados utilizando o algoritmo diretamente na base de imagens ou utilizando um limiar textual. Tanto na etapa de recuperação de informações da base, como na etapa de comparação e cálculo de distância entre as imagens recuperadas foi obtida uma melhora de desempenho.

Tabela 4.2: *Tempos e taxa de acerto da metodologia.*

Função	Tempos sem Limiar	Tempo com Limiar	Melhora no desempenho
Recuperação	0.05 Segundos	0.015 Segundos	70%
Comparação	0.02 Segundos	0.015 Segundos	25%
Total	0.07 Segundos	0.03 Segundos	58%

Arampatzis[4], usando um limiar dinâmico, conseguiu aumentos de eficiência de no máximo 22.9% em relação ao melhor tipo de busca, em seu estudo a busca textual simples comparada a busca em dois estágios (limiar + CBIR), usando a base de imagens da Wikipédia. Em nosso trabalho, os tempos de recuperação da metodologia empregada atingiram 58% de melhora no tempo na recuperação. Note que essa melhora aumenta conforme o número de tipos diferentes de imagens adicionadas na base também aumenta.

A figura 4.1 mostra um exemplo dos resultados visuais obtidos com a busca. Usando a figura 4.1 (a) como parâmetro de busca, os resultados obtidos foram: Sua cópia presente na base representada pela figura 4.1 (b), e logo após, a figura 4.1 (c), que representa a imagem mais semelhante a imagem parâmetro. A figura 4.1 (c) e a figura 4.1 (a) são ambas avaliadas como sem pneumonia.

Após os resultados obtidos com os testes nas radiografias do pulmão, foram realizados testes em radiografias periapicais, usando os mesmos métodos de validação.

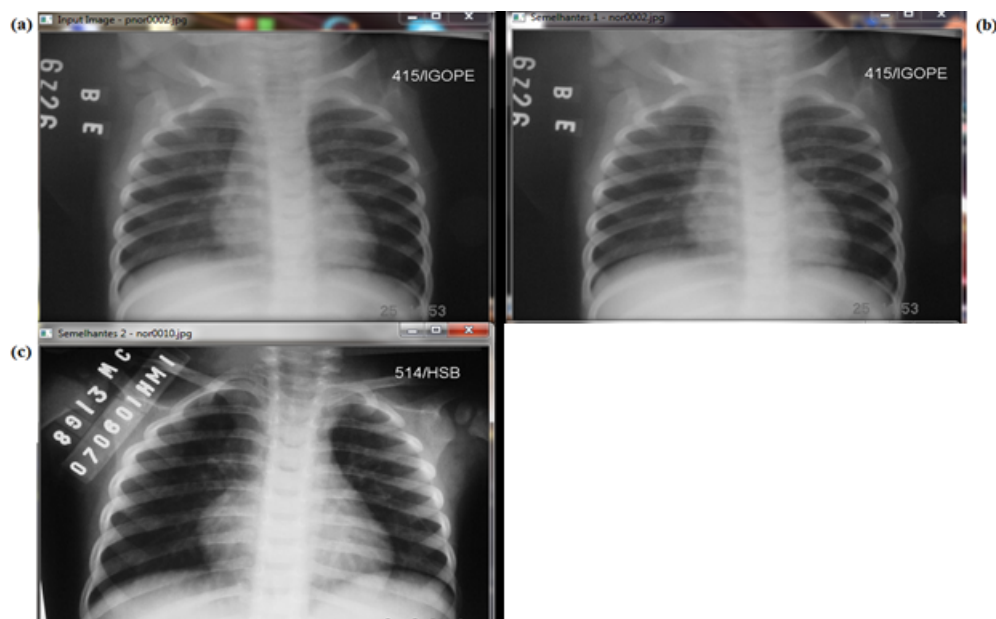


Figura 4.1: (a) Imagem de entrada (b) Primeiro resultado (c) Segundo Resultado.

4.2 Radiografias Periapicais

As imagens dentárias foram classificadas em duas classes principais, dentição normal, ou anormal. Dentro da segunda classe, elas se dividiam em 3 subclasses: aparelho, implante e edentulo. Em um primeiro momento, será avaliado apenas em duas hipóteses, dentição normal ou não. Este trabalho de classificação foi realizado em parceria com a professora Dra. Fernanda Paula Yamamoto Silva da Faculdade de Odontologia - UFG.

Assim como nas radiografias do pulmão, em 100% dos casos foi possível identificar a imagem parâmetro dentro da base de dados. Posteriormente foi realizada a classificação em normal ou anormal. Na tabela 4.3 é mostrado as taxas de acerto da técnica.

Tabela 4.3: Taxa de acerto na recuperação de imagens dentárias.

Tipo de Busca	Taxa de Acerto
Taxa de Acerto na busca da própria imagem	100%
Taxa de Acerto do método leave-one-out (PCA = 2)	89,58%

A figura 4.2 mostra os resultados obtidos usando os mesmos testes de validação usados nas radiografias de pulmão. Como pode ser visto na figura 4.2 foi possível não apenas recuperar uma cópia de si presente na base, como a imagem mais semelhante também tem a mesma classificação, neste caso dentição normal.

O ganho de desempenho obtido ao usar o limiar também se fez presente. Foi obtido um ganho total de 30% de ganho de desempenho. O principal fator do ganho menor, é relativo a quantidade maior de imagens, mas também ao fato da resolução

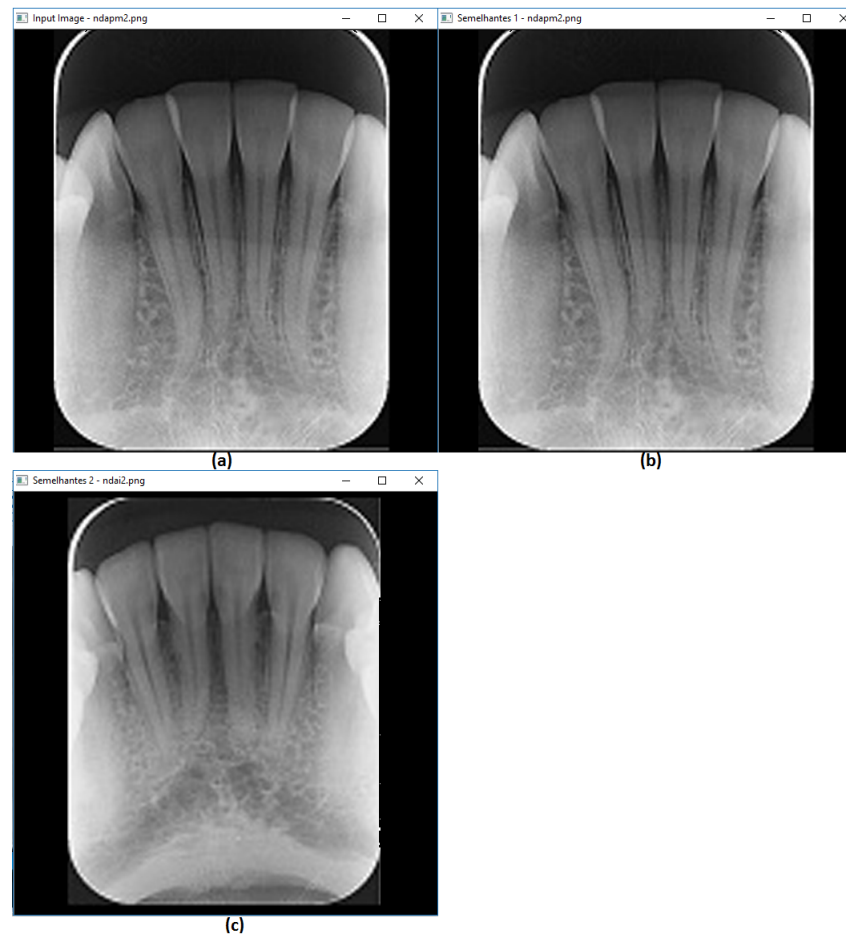


Figura 4.2: Resultados dos testes em radiografias periapicais. (a) Imagem de entrada. (b) Recuperação da própria imagem dentro da base. (c) Primeira imagem mais semelhante presente na base.

das radiografias periapicais também ser maior, logo o tempo médio de execução das comparações é maior.

4.3 Resultados Finais

Em ambos os casos, foram testados todos os valores possíveis de componentes, indo de 1 até 17 (número total de características extraídas). A ideia era achar o número de componentes com menor variância entre os resultados, já que o objetivo do trabalho é desenvolver um classificador geral, por isso mesmo, que o número de componentes pudesse mudar a cada classe, o componente escolhido foi aquele com a melhor média e menor variância. Na tabela 4.4 é mostrado as taxas de acerto em cada classe de imagens indo de 1 a 17 componentes.

Apesar de individualmente 3 e 1 componentes terem melhores resultados nas imagens de pulmão e dentárias, respectivamente, a média de ambas as classes não foi a

Tabela 4.4: *Taxas de Acerto com numero de componentes variando de 1 a 17.*

Componentes	Acerto - Pulmão	Acerto - Dentes	Acerto - Média
1	85,45%	93,23%	89,34%
2	90,91%	89,58%	90,25%
3	91,82%	86,98%	89,40%
4	87,27%	85,42%	86,34%
5	83,64%	82,81%	83,22%
6	81,82%	83,85%	82,84%
7	81,82%	84,38%	83,10%
8	80,00%	84,38%	82,19%
9	79,09%	83,85%	81,47%
10	80,00%	83,85%	81,93%
11	79,09%	83,33%	81,21%
12	80,00%	83,85%	81,93%
13	80,00%	83,85%	81,93%
14	80,00%	83,85%	81,93%
15	80,00%	83,85%	81,93%
16	80,00%	83,85%	81,93%
17	83,64%	80,21%	81,92%

melhor possível. Pensando em um classificador que deve ter a menor interferência possível no tipo de imagem a ser analisado, já que é um classificador genérico, foi escolhido 2 componentes para testar ambas as classes. A tabela 4.5 mostra uma compilação dos resultados obtidos nos dois métodos, além de compilação obtida através da média dos dois primeiros.

Tabela 4.5: *Compilação dos resultados obtidos.*

Imagens	Acerto 1ª Imagem	Acerto 2ª Imagem	Desempenho
Pulmão	100%	90.91%	58%
Dentes	100%	89.58%	30%
Total	100%	90.25%	44%

A ideia é que a medida que novas classes de imagens sejam adicionadas, a taxa de acerto se mantenha estável, sempre próxima aos 90%. O limiar se torna ainda mais importante, pois o objetivo do PCA é otimizar os resultados em função das imagens selecionadas no escopo de busca, e uma vez que o limiar garante que apenas imagens de interesse serão selecionadas, o PCA maximizará os resultados.

O motivo de terem sido selecionadas imagens tão diferentes entre si, como as radiografias periapicais e de pulmão, foi para testar se o sistema pode se adaptar a tais diferenças, o que representaria que ele é pouco influenciável pelo conteúdo da imagem que será classificada. Como os resultados mostraram o objetivo foi alcançado.

Conclusão

A técnica de CBIR implementada tem como maior contribuição auxiliar no diagnóstico e tratamento de pessoas que apresentem problemas semelhantes aos contidos na base, uma vez que a comparação visual dos exames poderá ser útil ao avaliar as condições de cada paciente e o melhor tipo de tratamento a ser utilizado.

Este trabalho abre a possibilidade de implementação de um sistema de informação CAD que pode ser utilizado na busca por casos semelhantes em exames baseados em imagens. Tal sistema de informação tornaria possível identificar rapidamente casos semelhantes com base nas imagens, mesmo em bases extensas com variados tipos de imagem servindo assim de suporte aos especialistas servindo como segunda opinião. Uma vez que o profissional de saúde defina o diagnóstico mais adequado, pois através dos resultados poderá ter acesso a informações como qual tratamento foi mais efetivo nos casos semelhantes que foram recuperados, poderá maximizar os resultados de recuperação dos pacientes.

O sistema também pode auxiliar na busca por casos raros, doenças ou situações incomuns, através da similaridade de outras imagens presentes na base, trazendo assim várias informações associadas a ela como condições do paciente, e tipo de tratamento utilizado anteriormente, mesmo que os casos ocorram em locais diferentes.

O sistema ainda permite que outros tipos de imagem sejam adicionados, onde devido ao limiar textual combinado ao PCA o sistema poderá classificar vários tipos de imagem sem perder a eficiência da técnica. Uma vez que seja validado o uso em imagens 2D pretende-se implementar também técnicas com foco em imagens 3D no sistema.

Como sugestão a trabalhos futuros pode ainda ser implementado uma interface gráfica para facilitar o uso além de definir uma base de dados completa com várias informações que possam ser anexadas a imagem.

Referências Bibliográficas

- [1] ABDI, H.; WILLIAMS, L. J. **Principal component analysis**. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2:433–459, 2010.
- [2] ANKERST, M.; KASTENMÜLLER, G.; KRIEGER, H.-P.; SEIDL, T. **Nearest neighbor classification in 3d protein databases**. *ISMB*, 99:34–43, 2000.
- [3] ANTANI, S.; LONG, L. R.; THOMA, G. R. **Content-based image retrieval for large biomedical image archives**. *Proc 11th World Cong Medical Informatics*, p. 829–833, 2004.
- [4] ARAMPATZIS, A.; ZAGORIS, K.; CHATZICHRISTOFIS, S. A. **Dynamic two-stage image retrieval from large multimedia databases**. *Information Processing & Management*, 49:274–285, 2012.
- [5] BHATIA, N.; VANDANA. **Survey of nearest neighbor techniques**. *(IJCSIS) International Journal of Computer Science and Information Security*, 8:302–305, 2010.
- [6] BORGES, E. N.; CONY, C. A.; DE FREITAS, A. L. C. **Utilização de algoritmos de busca por similaridade em sgbd's comerciais**. *Anais do IV Simpósio de Informática da Região Centro do RS - SIRC/RS*, p. 1–7, 2005.
- [7] CAUCHY, A.-L. **Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes**. *Oeuvres Complètes (Ileme Serie)*, 9:174–195, 1829.
- [8] DE MACEDO, S. O. **Desenvolvimento de um sistema de auxílio ao diagnóstico de pneumonia na infância utilizando visão computacional**. Master's thesis, Instituto de Informática, Universidade Federal de Goiás, 2012.
- [9] DEZA, M. M.; DEZA, E. **Encyclopedia of Distances**. Springer, New York, 2009.
- [10] DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. Wiley, 2001.
- [11] EAKINS, J.; GRAHAM, M.; FRANKLIN, T. **Content-based image retrieval**. *Library and Information Briefings*, 85:1–15, 1999.

- [12] ERTÖZ, L.; STEINBACH, M.; KUMAR, V. **A new shared nearest neighbor clustering algorithm and its applications.** *Workshop on Clustering High Dimensional Data and its Applications at 2nd SIAM International Conference on Data Mining*, p. 105–115, 2002.
- [13] FIX, E.; HODGES, J. L. **Discriminatory analysis, nonparametric discrimination: Consistency properties.** *US Air Force School of Aviation Medicine, Technical Report 4*, 3:477+, 1951.
- [14] GONZALEZ, R. C.; WOODS, R. E. **Processamento Digital De Imagens - 3ª Ed.** Pearson Education - Br, 2011.
- [15] HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. **Textural features for image classification.** *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3, NO. 6:610–621, 1973.
- [16] HAWKINS, J. K. **Textural properties for pattern recognition.** *Picture Processing and Psychopictorics*, ed. by B.S. Lipkin, A. Rosenfield, Academic Press, p. 347–370, 1970.
- [17] HEBBAR, H.; U.C, N.; MUSHIGERI, S. **Content based image retrieval based on cumulative distribution function - a performance evaluation.** *International Journal of Computer Applications*, 81:16–22, 2013.
- [18] HOTELLING, H. **Analysis of a complex of statistical variables into principal components.** *Journal of Educational Psychology*, 25:417–441, 1933.
- [19] JAIN, A. K.; DUIN, R. P.; MAO, J. **Statistical pattern recognition: A review. pattern analysis and machine intelligence.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:4–37, 2000.
- [20] JORDAN, C. **Memoire sur les formes bilineaires.** *Journal de Mathematiques Pures et Appliquees*, 19:35–54, 1974.
- [21] KACHOURI, R.; DJEMAL, K.; MAAREF, H.; MASMOUDI, D. S.; DERBEL, N. **Feature extraction and relevance evaluation for heterogeneous image database recognition.** *First Workshops on Image Processing Theory, Tools and Applications*, p. 1–6, 2008.
- [22] KIM, J.; CAI, W.; FENG, D.; WU, H. **A new way for multidimensional medical data management: volume of interest (voi)-based retrieval of medical images with visual and functional features.** *IEEE Transactions on Information Technology in Biomedicine*, 10:598–607, 2006.

- [23] LI, X.; CHEN, L.; ZHANG, L.; LIN, F.; MA, W.-Y. **Image annotation by large-scale content-based image retrieval**. *14th ACM international conference on Multimedia*, p. 607–610, 2006.
- [24] LONG, L. R.; ANTANI, S. K.; THOMA, G. R. **Image informatics at a national research center**. *Computerized Medical Imaging and Graphics*, 29:171–193, 2005.
- [25] MATHEUS, B. R. N.; SCHIABEL, H. **Online mammographic images database for development and comparison of cad schemes**. *Journal of Digital Imaging*, 24:500–506, 2011.
- [26] MILOUD-AOUIDATE, A.; BABA-ALI, A. R. **Survey of nearest neighbor condensing techniques**. *(IJACSA) International Journal of Advanced Computer Science and Applications*, 2:59–64, 2011.
- [27] MITCHELL, T. **Machine Learning**. McGraw Hill, 1997.
- [28] MÜLLER, H.; MICHOUX, N.; BANDON, D.; GEISSBUHLER, A. **A review of content-based image retrieval systems in medical applications-clinical benefits and future directions**. *International Journal of Medical Informatics*, 73:1–23, 2004.
- [29] MUNESAWANG, P.; GEISSBUHLER, A. **Next-generation multimedia database retrieval**. *IEEE Multimedia*, 14:106–107, 2006.
- [30] OLIVEIRA, L. L. G.; E SILVA, S. A.; RIBEIRO, L. H. V.; DE OLIVEIRA, R. M.; COELHO, C. J.; ANDRADE, A. L. S. S. **Computer-aided diagnosis in chest radiography for detection of childhood pneumonia**. *In International Journal of Medical Informatics*, 77:555–564, 2008.
- [31] PEARSON, K. **On lines and planes of closest fit to systems of points in space**. *Philosophical Magazine*, 2:559–572, 1901.
- [32] RAMAMURTHY, B.; CHANDRAN, K. R. **Content based image retrieval for medical images using canny edge detection algorithm**. *International Journal of Computer Applications*, 17:32–37, 2011.
- [33] ROBERTSON, S.; HULL, D. A. **The trec-9 filtering track final report**. *Proceedings of The Ninth Text REtrieval Conference, TREC*, p. 25–40, 2000.
- [34] RODRIGUES, B. U. **Reconhecimento do tipo de cachaça utilizando visão computacional e reconhecimento de padrões**. Master's thesis, Instituto de Informática, Universidade Federal de Goiás, 2015.

- [35] SAMET, H. **Techniques for similarity searching in multimedia databases.** *VLDB Endowment*, 3:1649–1650, 2010.
- [36] SANGHAVI, J.; KAYANDE, D. **Content based image retrieval (cbir) system for diagnosis of blood related diseases.** *IJCA Proceedings on National Conference on Innovative Paradigms in Engineering & Technology 2013*, 1:11–15, 2013.
- [37] SIADAT, M.-R.; SOLTANIAN-ZADEH, H.; FOTOUHI, F.; ELISEVICH, K. **Content-based image database system for epilepsy.** *Computer Methods and Programs in Biomedicine*, 79:209–226, 2005.
- [38] SOUSA, R. T.; MARQUES, O.; SOARES, F. A. A. M. N.; JR, I. I. G.; DE OLIVEIRA, L. L. G.; SPOTO, E. S. **Comparative performance analysis of machine learning classifiers in detection of childhood pneumonia using chest radiographs.** *Procedia Computer Science*, 18:2579–2582, 2013.
- [39] THOMA, G. R.; LONG, L. R.; ANTANI, S. **Biomedical imaging research and development: Knowledge from images in the medical enterprise.** Technical report, LHNBCB Board of Scientific Counselors, 2006.
- [40] TORRES, R. S.; FALCÃO, A. X. **Content-based image retrieval: Theory and applications.** *Revista de Informática Teórica e Aplicada - RITA*, 13:161–185, 2006.
- [41] VAN LEUKEN, R. H.; GARCIA, L.; OLIVARES, X.; VAN ZWOL, R. **Visual diversification of image search results.** *18th international conference on World wide web*, p. 341–350, 2009.
- [42] YUAN, K.; TIAN, Z.; ZOU, J.; BAI, Y.; YOU, Q. **Brain ct image database building for computer-aided diagnosis using content-based image retrieval.** *Information Processing and Management: an International Journal*, 47:176–185, 2011.

Apêndice 1

Artigos produzidos durante a pesquisa:

- Information System using Computer-Aided Diagnosis and Content-Based Image Retrieval. In: XI Workshop de Visão Computacional, 2015, São Carlos - SP. v. 1. p. 50-55.
- Analysis of Techniques of the Content-Based Image Retrieval to construct an Information System of the Computer-Aided Diagnosis. In: Americas Conference on Information Systems (AMCIS), 2015, Porto Rico. v. 1. p. 1-12.

Information System using Computer-Aided Diagnosis and Content-Based Image Retrieval

Elias Borges Macena Junior, Ronaldo Martins da Costa,
Leandro Luis Galdino de Oliveira, Rogerio Lopes Salvini
Instituto de Informática
Universidade Federal de Goiás – UFG
{ jrlinked, ronaldomc12, leandroluis.inf,
rogerio.salvini}@gmail.com

Fátima de Lourdes dos Santos Nunes Marques
Escola de Artes, Ciências e Humanidades
Universidade de São Paulo – USP
fatima.nunes@usp.br

Abstract—To improve the process of diagnosis with improved quality several studies center have focused their research on the development of information systems applying powerful computer-aided diagnosis techniques (CAD). In this context, the creation of databases to images with appropriate structures, capable of store images from different imaging systems with different characteristics and specifications is the first step to develop an efficient CAD system. This work proposes the validation of storage techniques and recovery with a hybrid CBIR method based on 2D medical images. The results of the techniques applied indicate a gain of 41% in image retrieval speed, opening way for the development of more efficient information systems to support diagnosis.

Keywords—Computer-Aided Diagnosis, Content-Based Retrieve

I. INTRODUÇÃO

Bancos de imagens médicas constituem uma necessidade fundamental para a avaliação e desenvolvimento de sistemas de Diagnóstico Auxiliado por Computador (CAD). A utilização de técnicas de CBIR podem auxiliar na composição deste sistemas pois proveem aos profissionais da saúde casos similares a um determinado modelo fornecido como argumento de consulta[1][2][3].

Na tentativa de aprimorar o processo do diagnóstico, vários centros de pesquisa têm direcionado seus trabalhos para o desenvolvimento de sistemas CAD [4]. Segundo [3] e [5], aplicações de radiologia têm impulsionado o desenvolvimento de sistemas CAD e *Content Based Image Retrieval* (CBIR). A validação dos algoritmos de tais sistemas exige bases de imagens com casos variados e que contenham estruturas de interesse procuradas pelos algoritmos. Porém em sua grande maioria, os bancos de imagens possuem limitações como ferramentas de busca, imagens de baixa qualidade, pouca quantidade de imagens ou poucas variações de casos [4].

Neste contexto, a criação de um banco de imagens com estruturas capazes de armazenar adequadamente imagens providas de diferentes sistemas de captura, com diferentes características e especificações, é o primeiro passo para o desenvolvimento de um eficiente sistema CAD. [6] afirma ser fundamental para sistemas CAD e de CBIR que a base de imagens seja definida de forma eficiente capaz de armazenar informações relevantes das imagens para posterior recuperação.

Existem diversas técnicas para armazenar informações de uma imagem médica. [6] aplicou com sucesso Non-Negative Tensor Factorization (NTF), Filtros Gabor com Non-Negative Tensor Factorization (NMF) e Principal Components Analysis

(PCA) para extrair características de imagens e armazená-las em um banco de imagens bidimensionais (2D).

Também é possível observar cada vez mais presente nos centros médicos, centros de pesquisa e hospitais a utilização de equipamentos capazes de gerar imagens 3D que podem aumentar a precisão do diagnóstico. Paralelamente é possível observar, porém em menor escala, trabalhos que avaliam características 3D em imagens médicas [7][8][9]. No entanto, a grande maioria dos trabalhos que realizam armazenamento ou busca em imagens 3D passam pelo processo de reconstrução das mesmas o que não é uma operação simples[10][11][12][13].

Dada esta carência e demanda por bancos de imagens médicas, sejam eles 2D ou 3D, justifica a condução de estudos que possam sugerir especificações para a construção de bancos de imagens médicas 2D ou 3D. Uma correta definição pode impulsionar a implementação de bancos de imagens para aplicações médicas e também o desenvolvimento de sistemas CAD mais eficientes.

Com base nessas premissas, o foco deste trabalho é desenvolver e validar técnicas de armazenamento e principalmente tempo de recuperação utilizando CBIR em uma base com imagens médicas 2D. Este estudo pode contribuir para o desenvolvimento de sistemas CAD mais eficientes que auxiliem os profissionais de saúde.

II. METODOLOGIA

A. Armazenamento

Bancos de dados atualmente trabalham com arquivos multimídia, usando um campo do tipo BLOB (Binary Large Object - Grande Objeto binário). Trata-se um campo criado para o armazenamento de qualquer tipo de informação em formato binário.

Os bancos de dados contém campos para tratar os arquivos multimídia, como as imagens, tornando o armazenamento uma etapa trivial [14][15][16][17]. A grande dificuldade e foco dos principais trabalhos publicados atualmente está em como recuperar essas informações de maneira eficiente.

Para solucionar esta dificuldade, desenvolvemos uma metodologia que a cada vez que uma nova imagem é adicionada ao banco, todas as características que serão utilizadas para recuperação serão extraídas, calculadas e armazenadas no banco, isto é feito para evitar que as mesmas operações sejam realizadas repetidas vezes, a cada nova busca, impactando diretamente na velocidade de recuperação das imagens. Além delas, outras informações poderão ser armazenadas junto da imagem, como nome do paciente, sexo,

idade, e qualquer informação que venha a ser útil para a recuperação.

B. Recuperação

A maneira usada para recuperar imagens em bancos de dados é conhecida como busca por similaridade na qual se inclui técnicas de CBIR. Funciona usando uma imagem como parâmetro de busca para recuperar as que se assemelham a ela. O CBIR funciona através de 3 passos:

- Extração de Características (a imagem é analisada e as características de interesse são extraídas através de algum descritor).
- Indexação de Características (Vetor de Características) Uma vez extraídas, é construído um vetor de características com as informações extraídas (ex: cor, dimensão, formas, etc.) da imagem.
- Comparação (É realizada a comparação entre os diversos vetores de características extraídos, nessa etapa que são definidas as imagens semelhantes que devem ser recuperadas).

Existem várias técnicas de CBIR, cada uma baseada em algum tipo de descritor, ou característica da imagem foco. Para implementar uma técnica de CBIR eficiente é preciso conhecer o tipo de imagem foco, nesse caso, serão imagens radiográficas do pulmão, com ou sem pneumonia. Na Fig. 1 pode ser visto um exemplo de imagem que será usada no desenvolvimento da técnica de CBIR a ser usada no banco de imagens. Uma vez estabelecida a base pretende-se expandi-la para outros tipos de imagens.

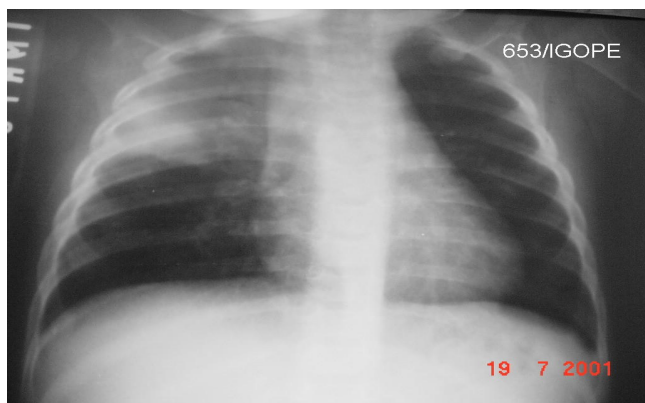


Fig. 1. Exemplo de imagem usada nos testes

Segundo [18] e [19] a melhor forma de trabalhar essas imagens é separando o pulmão esquerdo e direito e analisando-os separadamente. Isto é feito não apenas para descartar áreas sem informações importantes, mas também para gerar subimagens quadradas e tornar possível a aplicação da Transformada Wavelet de Haar. A separação é feita de forma automática pelo algoritmo.

A Transformada de Haar permite extrair informações da imagem tanto no domínio de tempo como no domínio de frequência, permitindo que sejam analisadas informações globais como também detalhes da imagem. A transformação de uma imagem consiste nas sucessivas passagens de filtros

passas-altas e passas-baixas, decompondo os sinais em detalhes e aproximações. A Fig.2 apresenta como uma imagem é decomposta pela transformada.

Cada imagem é decomposta em 3 direções, horizontal, vertical e diagonal, uma nova imagem chamada de coeficiente de aproximação (CA_1) é gerada. Essas 4 imagens formam o nível de D1 de decomposição da imagem original. Para mais detalhes, basta repetir o processo, usando a imagem CA_1 criando assim o nível de decomposição D2, com mais 4 imagens, D2 Horizontal, D2 Vertical, D2 Diagonal e CA_2 . Cada imagem é decomposta nos níveis D1, D2 e D3 após serem divididas em duas subimagens. A Fig. 3 mostra a imagem original, e as duas subimagens resultantes após a divisão enquanto a Fig. 4 mostra a Fig. 3(b) e a Fig. 3(c) no nível D1 de decomposição.

Nenhum outro tipo de pré-processamento é necessário além de dividir as imagens em duas partes e da decomposição pela Transformada de Haar.

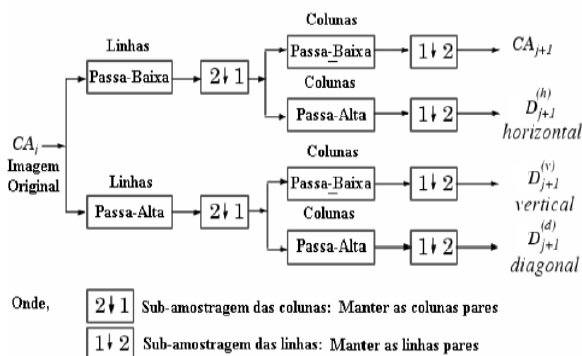


Fig. 2. Funcionamento da Transformada

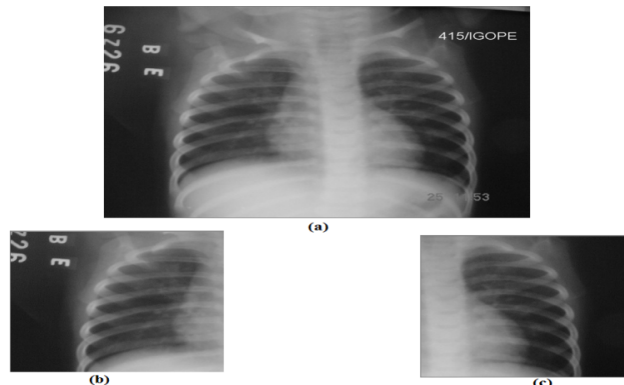


Fig. 3. (a) Imagem original (b) Subimagem do pulmão esquerdo (c) Subimagem do pulmão direito

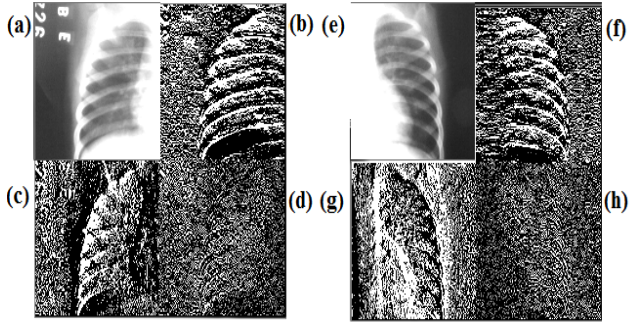


Fig. 4. (a) CA₁ (b) D1 Horizontal (c) D1 Vertical (d) D1 Diagonal referentes ao pulmão esquerdo (e) CA₁ (f) D1 Horizontal (g) D1 Vertical (h) D1 Diagonal referentes ao pulmão direito

Uma vez aplicada a transformada, executam-se as equações matemáticas que produzem informações estatísticas dos coeficientes em cada sub-imagem. As equações matemáticas apresentadas na Tab. I estão baseadas no trabalho de [20].

TABLE I. EQUAÇÕES UTILIZADAS PARA CONSTRUÇÃO DO VETOR DE CARACTERÍSTICAS

Descrição	Equação
Coefficiente de variação	$CV = \frac{\sigma}{\mu}$
Contraste	$Contraste = \sum_i \sum_j (i - j)^2 * p(i, j)$
Correlação	$Correlação = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
Desvio Médio	$DesvMedio = \frac{\sum_i \sum_j p(i, j) - \bar{x} }{N}$
Variância	$V = \frac{\sum_i \sum_j (i, j) - \bar{x} }{N}$
Somatória da variância	$SumVariância = \sum_i (i - SumEntropia)^2 p_{x+y}(i)$
Diferença da variância	$diffVariância = V(p_{x-y})$
Entropia	$Entropia = \sum_i \sum_j p(i, j) \log(p(i, j))$
Somatória da entropia	$SumEntropia = - \sum_i p_{x+y}(i) \log(p_{x+y}(i))$
Diferença da entropia	$diffEntropia = - \sum_i p_{x-y}(i) \log(p_{x-y}(i))$
Energia	$Energia = \sum_i \sum_j p(i, j)^2$
Energia Média	$EnergiaMedia = \frac{Energia}{N}$
Momento de diferenças inversas	$MDI = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$
Resíduo médio	$ResMedio = \sum_i \sum_j ((i - j) - \bar{x})^2$
Soma média	$SumMedia = \sum_i ip_{x+y}(i)$
Suavidade	$SV = 1 - \frac{1}{1 + \mu^2}$
Desvio padrão	$\sigma = \sqrt{V}$

Essas informações são indexadas no vetor de características de cada imagem. Cada uma delas terá 18 valores (9 de cada pulmão, direito e esquerdo).

Uma vez que as características sejam armazenadas, é realizada o cálculo da distância de Manhattan, dado pela equação (1) entre cada imagem, isto é, a distância entre cada valor de cada característica. A imagem mais próxima da imagem de consulta é retornada. A distância de Manhattan foi escolhida por ser mais barata computacionalmente, pois usa apenas somas e subtrações para realizar o cálculo.

$$DistManhattan = |x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n| \quad (1)$$

Para cada nova imagem que fosse usada, esse cálculo seria repetido. Ao final a imagem com a menor distância calculada seria a imagem mais semelhante a original.

Para cada imagem na base, serão geradas 18 sub-imagens com 17 características cada uma, gerando um total de 306 valores a serem usados, o que tem um alto custo computacional. Numa base com 1000 imagens, seriam necessários 306 mil comparações. Apenas implementar o CBIR não é eficiente, é preciso aumentar a eficiência da técnica e diminuir os custos computacionais.

Uma maneira de tornar o método mais eficiente é mesclando buscas textuais com as técnicas de CBIR. Segundo [14] buscas textuais produzem resultados melhores do que buscas por similaridade. [14] sugere que combinar as duas produziria melhores resultados e diminuiriam os custos computacionais. Usando esse conceito como base é criado uma etapa que precede o CBIR com um limiar textual, que reduz o número de imagens nas quais serão aplicados os cálculos para extração de características e cálculo de distância.

Dentro do banco de imagens existem radiografias do pulmão e radiografias de outras partes do corpo, inseridas propositalmente. Comparar essas radiografias seria um desperdício. Um limiar textual que defina o foco apenas para exames do pulmão reduz o escopo de comparações tornando o algoritmo mais eficiente.

Um limiar textual também permite a mudança do algoritmo a ser utilizado, maximizando a taxa de acerto. Desta maneira é possível definir um algoritmo para cada tipo de exame armazenado na base.

III. RESULTADOS

Foram usadas 80 imagens radiográficas e outras 80 imagens aleatórias com o mesmo tamanho, mas conteúdo completamente diferente. Na Fig.5 Temos um exemplo de imagem aleatória. Isto serve para mostrar que o método é capaz de ignorar imagens que não tem relação com a imagem foco do teste, permitindo que outros tipos de imagens sejam adicionados sem comprometer a eficiência, pois não entram na média de tempo. Para testar a eficiência da técnica de CBIR, todas as imagens em algum momento foram usadas como parâmetro de busca. O algoritmo apresentou 100% de acerto na busca da imagem. Em 100% dos casos, a segunda imagem mais semelhante era uma variação da imagem original, porém do mesmo caso/paciente/diagnóstico. Isto é verificado devido as informações associadas a imagem que também são

recuperadas.



Fig. 5. Exemplo de imagem aleatória retirada da internet.

Comparando a utilização do algoritmo aplicado diretamente na base de imagens ou utilizando um limiar textual, obteve-se melhora de desempenho. A etapa de comparação e cálculo de distância entre as imagens recuperadas também teve uma melhora de desempenho. Num total de 400 testes, cada imagem foi verificada cinco vezes, os resultados podem ser vistos na Tab. II.

TABLE II. TEMPOS E TAXA DE ACERTO DA METODOLOGIA

Técnica	Tempo sem Limiar	Tempo com Limiar	Melhora
Recuperação das informações da base	0.035 Segundos	0.015 Segundos	58%
Comparação entre vetores de característica e Exibição das Imagens	0.016 Segundos	0.015 Segundos	7%
Total	0.051 Segundos	0.03 Segundos	41%
Taxa de Acerto	100%	100%	

Os resultados exibidos na Tab. II foram obtidos usando um notebook com processador i7-3630QM de 2.4GHz com 8GB de memória RAM.

A Referência [14], usando um limiar dinâmico, conseguiu aumentos de eficiência de no máximo 22.9% em relação ao melhor tipo de busca, em seu estudo a busca textual simples comparada a busca em dois estágios (limiar + CBIR), usando a base de imagens da Wikipédia. Em nosso trabalho, os tempos de recuperação da metodologia empregada atingiram 41% de melhora no tempo na recuperação.

Dessa maneira, como mencionado, é possível que outros tipos de imagens sejam adicionados na base, sem que as técnicas já implementadas sejam prejudicadas, abrindo um escopo cada vez maior no tipo de imagens suportadas.

IV. CONCLUSÃO

A base de imagens médicas, e a técnica de CBIR implementada para recuperação das mesmas tem como maior contribuição o fato de auxiliar no diagnóstico e tratamento de pessoas que apresentem exames semelhantes aos armazenados na base, uma vez que a comparação visual dos exames feito

pelo CBIR poderá ser útil ao avaliar as condições de cada paciente e o tipo de tratamento a ser utilizado.

Este trabalho abre possibilidades de construção de um sistema de informação CAD que pode ser utilizado por clínicas ou hospitais na busca por casos semelhantes em exames baseados em imagens. Tal sistema de informação oferecia suporte aos profissionais tornando os diagnósticos mais precisos, pois seria possível identificar rapidamente casos semelhantes com base nas imagens.

Pretende-se ainda implementar outros algoritmos para outros tipos de imagem, onde a técnica usada para recuperação seria escolhidos através de um limiar textual, cobrindo uma quantidade cada vez maior de tipos de imagem sem perder a eficiência da técnica.

REFERENCES

- [1] Müller, H., Michoux, N., Bandon, D., and Geissbuhler, A. (2004). A review of content-based image retrieval systems in medical applications—Clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1), p. 1–23.
- [2] Siadat, M., Soltanian-Zadeh, H., Fotouhi, F., and Elisevich, K. (2005). Content-based image database system for epilepsy. *Computer Methods and Programs in Biomedicine*, 79(3), p. 209–226.
- [3] Kim, J., Cai, W., Feng, D., and Wu, H. (2006). A new way for multidimensional medical data management: Volume of interest (VOI)-based retrieval of medical images with visual and functional features. *IEEE Transactions on Information Technology in Biomedicine*, 10(3), p. 598–606.
- [4] Matheus, B. R. N. and Schiabel, H. (2010). Online Mammographic Images Database for Development and Comparison of CAD Schemes. *Journal of Digital Imaging*, p. 1618-727X.
- [5] Kachouri, R., Djemal, K., Maaref, H., Masmoudi, D.S. and Derbel, N. (2008). Feature extraction and relevance evaluation for heterogeneous image database recognition. *Image Processing Theory, Tools and Applications. IPTA 2008. First Workshops on*. p. 1-6.
- [6] Yuan, K., Tian, Z., Zou, J., Bai, Y. and You, Q. (2011). Brain CT image database building for computer-aided diagnosis using content-based image retrieval. *Information Processing and Management. Vol.47(2)*, p. 176-185.
- [7] Frank H., Olaf K., Horst K. H. and Heinz-Otto P. (2011). Interactive 3D medical image segmentation with energy-minimizing implicit functions. *Visual Computing in Biology and Medicine - Computers & Graphics. Volume 35, Issue 2*, p. 275–287.
- [8] Danzhou L., Hua, K.A. and Sugaya, K. (2006). A Framework for Web-Based Interactive Applications of High-Resolution 3D Medical Image Data. *Computer-Based Medical Systems. CBMS 2006. 19th IEEE International Symposium on. Digital Object Identifier: 10.1109/CBMS.2006.12*. p. 119-124.
- [9] Chen T., Wang Y., Zhang H. and Xiao C. (2010). An embedded 3D medical image processing and visualization platform based on dual-core processor. *Intelligent Control and Automation (WCICA), 8th World Congress on. Digital Object Identifier: 10.1109/WCICA.2010.5554388*, p. 2936-2941.
- [10] Bergamasco, L.C.C., Nunes, F.L.S. (2013). Applying Distance Histogram to retrieve 3D cardiac medical models, In: *The American Medical Informatics Association Annual Symposium*.
- [11] Chen, Z. Q., Zou, K. S., Ip, W. H. and Chan, C. Y. (2011). 3D model retrieval based on fuzzy weighted shape distributions. In: *Advanced Materials Research. [S.l.]: Trans Tech Publications. v. 201-203*, p. 1678–1681.
- [12] Vranic, D. V. (2001). 3D Model Retrieval. Tese (PhD Thesis) — University of Leipzig, Germany.

- [13] Yang, F. and Leng, B. (2007). OFS: A feature selection method for shape-based 3D model retrieval. In: Proceeding of 10th IEEE International Conference on Computer-Aided Design and Computer Graphics. Beijing, China: IEEE Computer Society. p. 114–119.
- [14] Arampatzis, A., Zagoris, K., and Chatzichristofis, S. A. (2011). Dynamic two-stage image retrieval from large multimedia databases. In Information Processing and Management.
- [15] Borges, E. N., Cony, C. A., and de Freitas, A. L. C. (2005). Utilização de algoritmos de busca por similaridade em sgbd's comerciais.
- [16] Muneesawang, P. and L. Guan, M. (2006). Database retrieval: A human-centered approach. In IEEE Computer Society.
- [17] Samet, H. (2010). Techniques for similarity searching in multimedia databases. In The 36th International Conference on Very Large Data Bases.
- [18] Oliveira, L. L. G., e Silva, S. A., Ribeiro, L. H. V., de Oliveira, R. M., Coelho, C. J., and Andrade, A. L. S. S. (2008). Computer-aided diagnosis in chest radiography for detection of childhood pneumonia. In International Journal of Medical Informatics.
- [19] de Macedo, S. O. (2012). Desenvolvimento de um sistema de auxílio ao diagnóstico de pneumonia na infância utilizando visão computacional. Mestrado, Instituto de Informática, Universidade Federal de Goiás.
- [20] Haralick, R., Shanmugam, K and Dinstein, I. (1973). Textural Features for Image Classification. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, vol. SMC-3, NO. 6.

APÊNDICE I

Coefficiente de variação: Tem como foco analisar o peso do desvio-padrão em relação a média dos valores dos pixels das imagens. Dado pela equação:

$$CV = \frac{\sigma}{\mu}$$

Onde σ é o desvio padrão e μ é a média dos coeficientes.

Contraste: O contraste caracteriza-se pela diferença entre os tons de cinza. Baixo contraste ocorre quando há pequena diferença entre os níveis de cinza dos pixels localizados em uma região contígua da imagem. É dada pela equação:

$$Contraste = \sum_i \sum_j (i - j)^2 * p(i, j)$$

Onde $p(i, j)$ é (i, j) -ésima probabilidade de ocorrência de um nível de cinza na imagem.

Correlação: A correlação mede a dependência linear entre os tons de cinza presentes em uma imagem. Valores altos de correlação dados pela equação, próximos de um, indicam a existência de relação linear entre os pares de níveis de cinza.

$$Correlação = \frac{\sum_i \sum_j (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Onde $p(i, j)$ é (i, j) -ésima probabilidade de ocorrência de um nível de cinza na imagem e $\sigma_x, \sigma_y, \mu_x$ e μ_y são os desvios padrão e médias em x e y respectivamente.

Desvio médio: O Desvio Médio é uma medida da dispersão dos dados em relação à média de uma sequência, o “afastamento” em relação a essa média. Esta medida representa a média das distâncias entre cada elemento da amostra e seu valor médio. Descrito pela equação:

$$DesvMedio = \frac{\sum_i \sum_j |p(i, j) - \bar{x}|}{N}$$

Onde N é o número total de pixels na imagem e \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Variância: Descreve quanto os valores de tons de cinza estão dispersos em torno da média. É dada pela equação:

$$V = \frac{\sum_i \sum_j |(i, j) - \bar{x}|}{N}$$

Onde V é a variância de $P(i, j)$

Diferença da variância:

$$difVariância = V(p_{x-y})$$

$p_{x-y}(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = |i - j|$

Diferença da entropia:

$$difEntropia = - \sum_i p_{x-y}(i) \log\{p_{x-y}(i)\}$$

Energia: Também conhecida como segundo momento angular, expressa a uniformidade de uma textura, assumindo o valor máximo 1 e é expressa pela equação:

$$Energia = \sum_i \sum_j p(i, j)^2$$

Energia média: É a energia da imagem dividida pelo número total de pixels. Dado pela equação:

$$EnergiaMedia = Energia/N$$

Entropia: A medida de entropia expressa a desordem contida na textura. Quando uma imagem não é uniforme, as entradas apresentam valores próximos de zero o que resulta em valores altos de entropia. É dada pela equação:

$$Entropia = \sum_i \sum_j p(i, j) \log(p(i, j))$$

Momento de diferenças inversas: Quando a concentração dos valores na diagonal da matriz de co-ocorrência for máxima, o Momento da Diferença Inverso (MDI) atinge o valor máximo. É dado pela equação:

$$MDI = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$$

Resíduo médio: É a diferença média entre os dados observados e os valores ajustados pelo modelo (calculados). É dado pela equação:

$$ResMedio = \sum_i \sum_j ((i - j) - \bar{x})^2$$

Onde \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Soma média:

$$SumMedia = \sum_i ip_{x+y}(i)$$

$p_{x+y}(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = i + j$

Somatório da entropia:

$$SumEntropia = - \sum_i p_{x+y}(i) \log(p_{x+y}(i))$$

Somatório da variância:

$$SumVariancia = \sum_i (i - SumEntropia)^2 p_{x+y}(i)$$

Suavidade:

$$SV = 1 - \frac{1}{1 + \mu^2}$$

Desvio padrão: É a raiz quadrada da variância, dado pela equação:

$$\sigma = \sqrt{V}$$

Analysis of Techniques of the Content-Based Image Retrieval to construct an Information System of the Computer-Aided Diagnosis

Emergent Research Forum papers - Research-In-Progress

Elias Borges Macena Júnior
Rogério Lopes Salvini
Leandro Luís Galdino de Oliveira
Ronaldo Martins da Costa
Instituto de Informática
Universidade Federal de Goiás
{[rlinked](#), [rogerio.salvini](#), [leandroluis.inf](#),
[ronaldomc12](#)}@gmail.com

Fátima de Lourdes dos Santos
Nunes Marque
Escola de Artes, Ciências e Humanidades
Universidade de São Paulo - USP
fatima.nunes@usp.br

Abstract

To improve the process of diagnosis with improved quality several researches center have focused their research on the development of information systems applying powerful computer-aided diagnosis techniques (CAD). In this context, the creation of databases to images with appropriate structures, capable of store images from different imaging systems with different characteristics and specifications is the first step to develop an efficient CAD system. This work proposes the validation of storage techniques and recovery with a hybrid CBIR method based on 2D medical images. The results of the techniques applied indicate a gain of 41% in image retrieval speed, opening way for the development of more efficient information systems to support diagnosis.

Keywords

Informática Médica, Recuperação de Imagens, Sistema de Informação Hospitalar, CBIR, CAD

Introdução

Bancos de imagens médicas constituem uma necessidade fundamental para a avaliação e desenvolvimento de sistemas de Diagnóstico Auxiliado por Computador (CAD). A utilização de técnicas de CBIR podem auxiliar na composição deste sistemas pois proveem aos usuários (profissionais da saúde) casos similares a um determinado modelo fornecido como argumento de consulta[Müller et al. 2004][Siadat et al. 2005][Kim et al. 2006].

Na tentativa de acelerar o processo do diagnóstico com ganho de qualidade, vários centros de pesquisa têm focado suas pesquisas no desenvolvimento de sistemas CAD mais poderosos [Matheus and Schiabel 2010]. Segundo [Kim et al. 2006] e [Kachouri et al 2008], além de várias aplicações que fazem uso de bancos de imagens, aplicações de radiologia têm impulsionado o desenvolvimento de sistemas CAD e *Content Based Image Retrieval* (CBIR). A validação dos algoritmos de tais sistemas exige bases de imagens com casos variados e que contenham estruturas de interesse procuradas pelos algoritmos. Porém em sua grande maioria, os bancos de imagens hoje disponíveis possuem várias limitações como ferramentas de busca, imagens de baixa qualidade, pouca quantidade de imagens ou ainda poucas variações de casos [Matheus and Schiabel 2010].

Neste contexto, a criação de um banco de imagens com estruturas capazes de armazenar adequadamente imagens providas de diferentes sistemas de captura, com diferentes características e especificações, é o primeiro passo para o desenvolvimento de um eficiente sistema CAD. [Yuan et al. 2011] afirma ser

fundamental para sistemas CAD e de CBIR que a base de imagens seja definida de forma eficiente capaz de guardar informações relevantes das imagens para posterior recuperação.

Em seu estudo, [Müller et al. 2004] identificou que a maioria das pesquisas sobre CBIR para aplicações médicas se concentram em algoritmos para extrair características de determinadas regiões do corpo humano ou de uma determinada patologia, ignorando a construção do banco de imagens que deveria ser primeiro passo a ser executado.

Existem diversas técnicas para armazenar informações de uma imagem médica. [Yuan et al. 2011] aplicou com sucesso Non-Negative Tensor Factorization (NTF), Filtros Gabor com Non-Negative Tensor Factorization (NMF) e Principal Components Analysis (PCA) para extrair características de imagens e armazená-las em um banco de imagens bidimensionais (2D).

Existem trabalhos que desenvolveram bases de imagens específicas para determinadas áreas: [Matheus and Schiabel 2010] implementou um banco de imagens 2D de mamografias, as imagens são todas inseridas por especialistas. Usuários diversos podem solicitar acesso ao banco pela internet. Todas as operações no banco são supervisionadas por um administrador que garante a qualidade do banco. [Yuan et al. 2011] construiu um banco de imagens bidimensionais (2D) de tomografia computadorizada do cérebro, todas as imagens foram fornecidas por hospitais e colaboradores.

Também é possível observar cada vez mais presente nos centros médicos, centros de pesquisa e hospitais a utilização de equipamentos capazes de gerar imagens 3D que podem aumentar a precisão do diagnóstico. Paralelamente é possível observar, porém em menor escala, trabalhos que avaliam características 3D em imagens médicas [Frank et al. 2011][Danzhou et al. 2006][Chen et al. 2010]. No entanto, a grande maioria dos trabalhos que realizam armazenamento ou busca em imagens 3D passam pelo processo de reconstrução das mesmas o que não é uma operação simples[Bergamasco and Nunes 2013][Chen et al. 2011][Vranic 2001][Yang and Leng 2007].

Dada esta carência e demanda por bancos de imagens médicas, sejam eles 2D ou 3D, justifica a condução de estudos que possam sugerir especificações para a construção de bancos de imagens médicas 2D ou 3D. Uma correta definição de parâmetros para a construção de bancos de imagens 2D ou 3D pode impulsionar a implementação de bancos de imagens, especialmente para aplicações médicas. Por consequência, a existência de mais bancos de imagens médicas, pode também impulsionar o estudo e desenvolvimento de sistemas CAD mais eficientes.

Com base nessas premissas, o foco deste trabalho é desenvolver e validar técnicas de armazenamento e principalmente tempo de recuperação utilizando CBIR em uma base com imagens médicas 2D. Este estudo pode contribuir para o desenvolvimento de sistemas CAD mais eficientes que auxiliem os profissionais de saúde.

Metodologia

Armazenamento

Bancos de dados atualmente trabalham com arquivos multimídia, usando um campo do tipo BLOB (Binary Large Object - Grande Objeto binário). Trata-se um campo criado para o armazenamento de qualquer tipo de informação em formato binário dentro de uma tabela em um Banco de dados.

Os bancos de dados contêm campos para tratar os arquivos multimídia, como as imagens, tornando o armazenamento uma etapa trivial. [Arampatzis et al. 2011], [Borges et al. 2005], [Muneesawang and L. Guan 2006] e [Samet 2010]. A grande dificuldade e foco dos principais trabalhos publicados atualmente está em como recuperar essas informações de maneira eficiente.

Para solucionar esta dificuldade, desenvolvemos uma metodologia que a cada vez que uma nova imagem é adicionada ao banco, todas as características que serão utilizadas para recuperação serão extraídas, calculadas e armazenadas no banco, isto é feito para evitar que as mesmas operações sejam realizadas repetidas vezes, a cada nova busca, impactando diretamente na velocidade de recuperação das imagens. Na seção de Recuperação serão discutidas quais são e como calcular cada uma destas características que representam a imagem. Além delas, outras informações poderão ser armazenadas junto da imagem, como nome do paciente, sexo, idade, e qualquer outra informação que possa vir a ser útil para sua recuperação.

Recuperação

A maneira usada para recuperar imagens em bancos de dados é conhecida como busca por similaridade na qual se inclui técnicas de CBIR. Funciona usando uma imagem como parâmetro de busca para recuperar imagens que se assemelham a ela. O CBIR funciona através de 3 passos:

- **Extração de Características**
Primeiramente a imagem é analisada e as características de interesse são extraídas através de algum descritor.
- **Indexação de Características (Vetor de Características)**
Uma vez extraídas, é montado um vetor de características com todas as informações extraídas (tais como cor, dimensão das estruturas encontradas, formas identificadas, etc.) da imagem.
- **Comparação**
Ao final é realizada a comparação entre os diversos vetores de características que foram extraídos das imagens, é nessa etapa que são definidas as imagens semelhantes que devem ser recuperadas.

Existem várias técnicas de CBIR, baseadas em descritores, ou característica da imagem. Para implementar uma técnica de CBIR eficiente é preciso conhecer o tipo de imagem, nesse caso, serão imagens radiográficas do pulmão, com ou sem pneumonia. A Figura 1. Exemplo de imagem usada nos testes. Uma vez estabelecida a base pretende-se expandi-la para outros tipos de imagens.

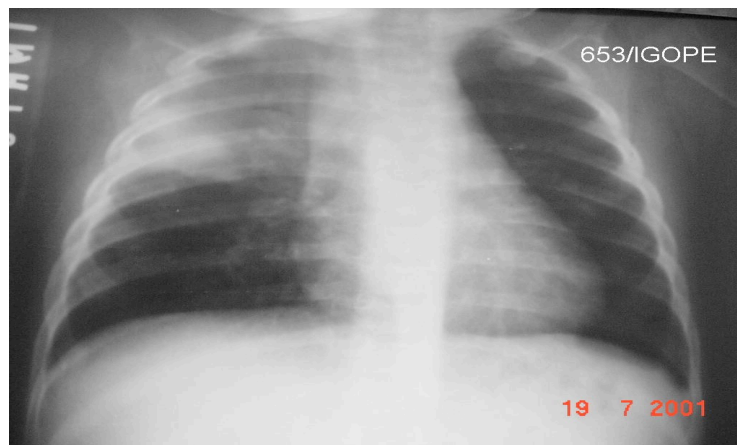


Figura 1. Exemplo de imagem usada nos testes

Segundo [Oliveira et al. 2008] e [de Macedo 2012] a melhor forma de trabalhar essas imagens é separando o pulmão esquerdo e direito e analisando-os separadamente. Isto é feito para descartar áreas sem informações importantes e também para gerar subimagens quadradas tornando possível a aplicação da Transformada Wavelet de Haar.

A Transformada de Haar permite extrair informações da imagem no domínio de tempo e no domínio de frequência, permitindo que sejam analisadas informações globais e detalhes da imagem. A transformação de uma imagem consiste em sucessivas passagens de filtros passas-alta e passas-baixa, decompondo os sinais em detalhes e aproximações. A Figura 2 apresenta como uma imagem é decomposta pela transformada.

Cada imagem é decomposta em 3 direções, horizontal, vertical e diagonal, e uma nova imagem chamada de coeficiente de aproximação (CA_1) é gerada. Essas 4 imagens formam o nível de D1 de decomposição da imagem original. Para um maior detalhamento, basta repetir o processo, usando a imagem CA_1 criando assim o nível de decomposição D2, com mais 4 imagens, D2 Horizontal, D2 Vertical, D2 Diagonal e CA_2 . Cada imagem é decomposta nos níveis D1, D2 e D3 após serem divididas em duas subimagens. A Figura 3

mostra a imagem original, e as duas subimagens resultantes após a divisão enquanto a Figura 4 mostra a Figura 3(b) e a Figura 3(c) no nível D1 de decomposição.

Nenhum outro tipo de pré-processamento é necessário além de dividir as imagens em duas partes e da decomposição pela Transformada de Haar.

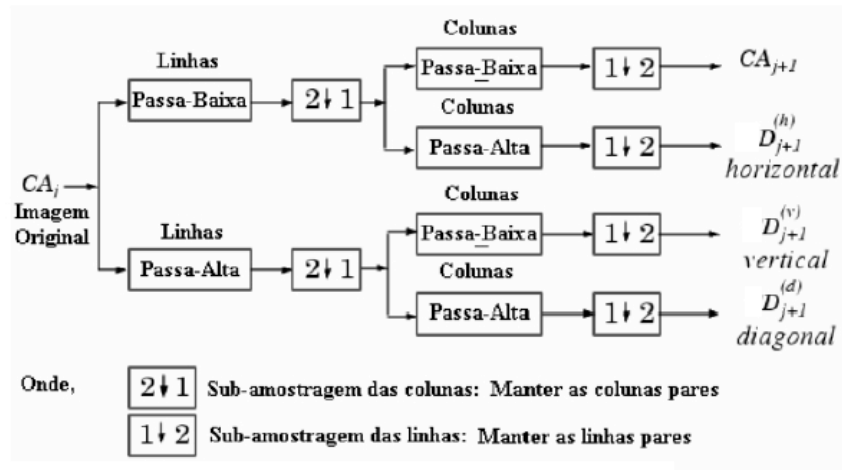


Figura 2. Funcionamento da Transformada

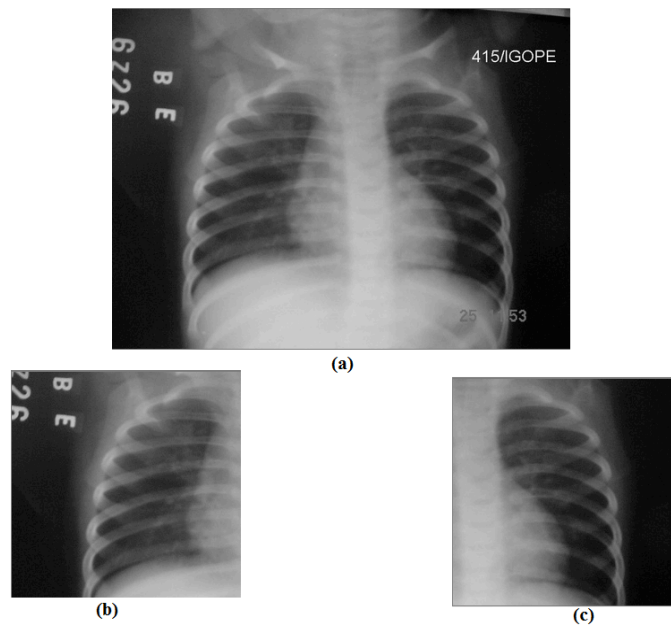


Figura 3. (a) Imagem original (b) Subimagem do pulmão esquerdo (c) Subimagem do pulmão direito

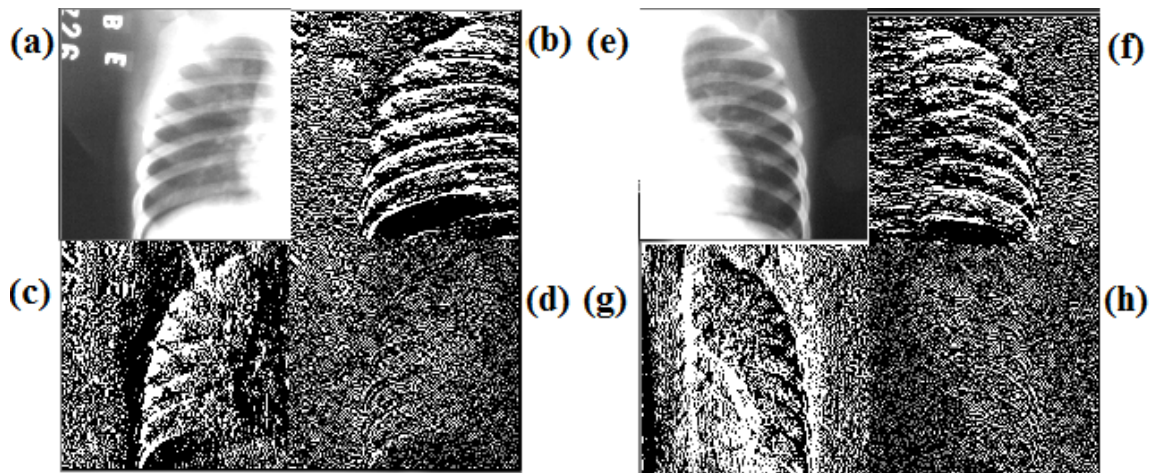


Figura 4. (a) CA, (b) D1 Horizontal (c) D1 Vertical (d) D1 Diagonal referentes ao pulmão esquerdo (e) CA, (f) D1 Horizontal (g) D1 Vertical (h) D1 Diagonal referentes ao pulmão direito

Uma vez aplicada a transformada, executam-se as equações matemáticas que produzem informações estatísticas dos coeficientes em cada sub-imagem. As equações matemáticas apresentadas na Tabela 1 estão baseadas no trabalho de [Haralick et al. 1973] e são descritas no Apêndice I deste documento.

Descrição	Equação
Coeficiente de variação	$CV = \frac{\sigma}{\mu}$
Contraste	$Contraste = \sum_i \sum_j (i - j)^2 * p(i, j)$
Correlação	$Correlação = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
Desvio Médio	$DesvMedio = \frac{\sum_i \sum_j p(i, j) - \bar{x} }{N}$
Variância	$V = \frac{\sum_i \sum_j (i, j) - \bar{x} }{N}$
Somatória da variância	$SumVariância = \sum_i (i - SumEntropia)^2 p_{x+y}(i)$
Diferença da variância	$difVariância = V(p_{x-y})$
Entropia	$Entropia = \sum_i \sum_j p(i, j) \log(p(i, j))$
Somatória da entropia	$SumEntropia = - \sum_i p_{x+y}(i) \log(p_{x+y}(i))$
Diferença da entropia	$difEntropia = - \sum_i p_{x-y}(i) \log\{p_{x-y}(i)\}$
Energia	$Energia = \sum_i \sum_j p(i, j)^2$

Energia Média	$EnergiaMedia = \frac{Energia}{N}$
Momento de diferenças inversas	$MDI = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$
Resíduo médio	$ResMedio = \sum_i \sum_j ((i - j) - \bar{x})^2$
Soma média	$SumMedia = \sum_i ip_{x+y}(i)$
Suavidade	$SV = 1 - \frac{1}{1 + \mu^2}$
Desvio padrão	$\sigma = \sqrt{V}$

Tabela 1. Equações utilizadas para construção do vetor de características

Essas informações são indexadas no vetor de características de cada imagem. Cada uma delas terá 18 valores (9 de cada pulmão, direito e esquerdo).

Uma vez que as características sejam armazenadas, é realizada o calculo da distância de Manhattan, dado pela equação (18) entre cada imagem, isto é, a distância entre cada valor de cada característica. A imagem mais próxima da imagem de consulta é retornada .

$$DistManhattan = |x1 - y1| + |x2 - y2| + \dots + |xn - yn| \quad (18)$$

A Figura 5 apresenta valores hipotéticos extraídos de uma imagem. Substituindo estes valores na equação 18 obtemos a distância Manhattan entre as imagens A e B. Os diversos valores de X e Y neste trabalho são substituídos pelas características apresentadas nas equações 1 a 17 presentes Tabela 1.

Atributos Imagem A	Atributos Imagem B
x1 = 10	y1 = 2
x2 = 5	y2 = 9
x3 = 5	y3 = 8
x4 = 6	y4 = 3
x5 = 3	y5 = 7

Figura 5. Atributos a serem usados no cálculo de distância

$$DistManhattan = |10 - 2| + |5 - 9| + |5 - 8| + |6 - 3| + |3 - 7|$$

A distância Manhattan entre as Imagens A e B presentes na Figura 5 será de 22. Para cada nova imagem C, D e E que fossem usadas, esse calculo seria refeito. Ao final a imagem com a menor distância calculada em relação a imagem A, seria a imagem mais semelhante a ela.

Para cada imagem na base, serão geradas 18 sub-imagens com 17 características cada uma, gerando um total de 306 valores a serem usados na equação 18, o que tem um alto custo computacional. Numa base com 1000 imagens, seriam necessários 306 mil comparações. Apenas implementar o CBIR não é suficientemente eficiente, é preciso aumentar a eficiência da técnica e diminuir os custos computacionais das comparações.

Uma maneira de tornar o método mais eficiente é mesclando buscas textuais com as técnicas de CBIR. Segundo [Arampatzis et al. 2011] buscas textuais produzem resultados melhores do que buscas por similaridade. [Arampatzis et al. 2011] sugere que ao combinar as duas produziriam melhores resultados e diminuiriam os custos computacionais. Usando esse conceito como base é criado uma etapa que precede o CBIR com um limiar textual, que reduz o número de imagens nas quais serão aplicados os cálculos para extração de características e calculo de distância.

Dentro do banco de imagens existem além de radiografias do pulmão, radiografias de outras partes do corpo, inseridas propositalmente, comparar essas radiografias com uma radiografia do pulmão seria um desperdício. Um limiar textual que definisse como foco apenas exames do pulmão reduziria o escopo de comparações tornando o algoritmo mais eficiente.

Um limiar textual também permite a mudança do algoritmo a ser utilizado, maximizando a taxa de acerto. Desta maneira é possível definir um algoritmo para cada tipo de exame armazenado na base. A Figura 6 apresenta um exemplo do Limiar. Além de radiografias de pulmão, existem radiografias diferentes na base. Após aplicação do limiar textual, apenas as radiografias do pulmão entram na comparação.

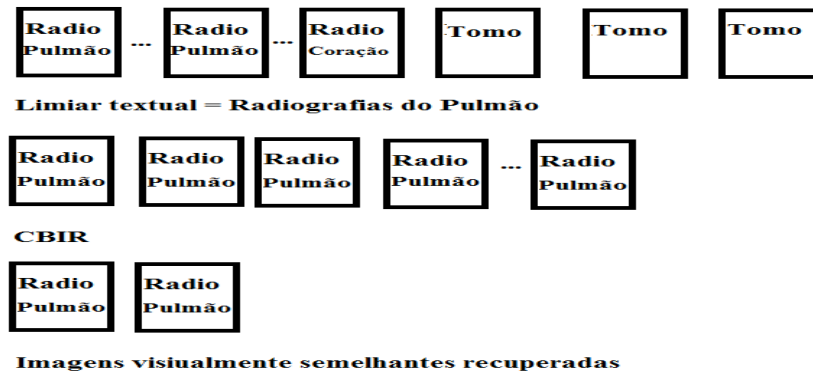


Figura 6. Exemplo do Limiar em funcionamento

Resultados

Foram usadas 80 imagens radiográficas e outras 80 imagens aleatórias com o mesmo tamanho, mas conteúdo completamente diferente. Para testar a eficiência da técnica de CBIR, todas as imagens em algum momento foram usadas como parâmetro de busca. O algoritmo apresentou 100% de acerto na busca da imagem. Em 100% dos casos, a segunda imagem mais semelhante era uma variação da imagem original, porém do mesmo caso/paciente/diagnóstico.

A Figura 7 mostra os resultados. Usando a Figura 7(a) como parâmetro de busca, as três imagens mais próximas a ela dentro da base de imagens foram: Sua cópia presente na base representada pela Figura 7(b), e logo após, a Figura 7(c) seguida da Figura 7(d).

Comparando a utilização do algoritmo aplicado diretamente na base de imagens ou utilizando um limiar textual, obteve-se melhora de desempenho. A etapa de comparação e cálculo de distância entre as imagens recuperadas também teve uma melhora de desempenho. Num total de 400 testes, cada imagem foi verificada cinco vezes, os resultados podem ser vistos na Tabela 1:

Técnica	Tempos sem Limiar	Tempo com Limiar	Melhora no desempenho
Recuperação das informações da base	0.035 Segundos	0.015 Segundos	58%
Comparação entre vetores de característica e Exibição das Imagens	0.016 Segundos	0.015 Segundos	7%
Total	0.051 Segundos	0.03 Segundos	41%
Taxa de Acerto	100%	100%	

Tabela 1. Tempos e taxa de acerto da metodologia

[Arampatzis et al. 2011], usando um limiar dinâmico, conseguiu aumentos de eficiência de no máximo 22.9% em relação ao melhor tipo de busca, em seu estudo a busca textual simples comparada a busca em dois estágios (limiar + CBIR), usando a base de imagens da Wikipédia. Em nosso trabalho, os tempos de recuperação da metodologia empregada atingiram 41% de melhora no tempo na recuperação.

Dessa maneira, como mencionado na seção de recuperação, é possível que outros tipos de imagens sejam adicionadas na base, sem que as técnicas já implementadas sejam prejudicadas, abrindo um escopo cada vez maior no tipo de imagens suportadas.

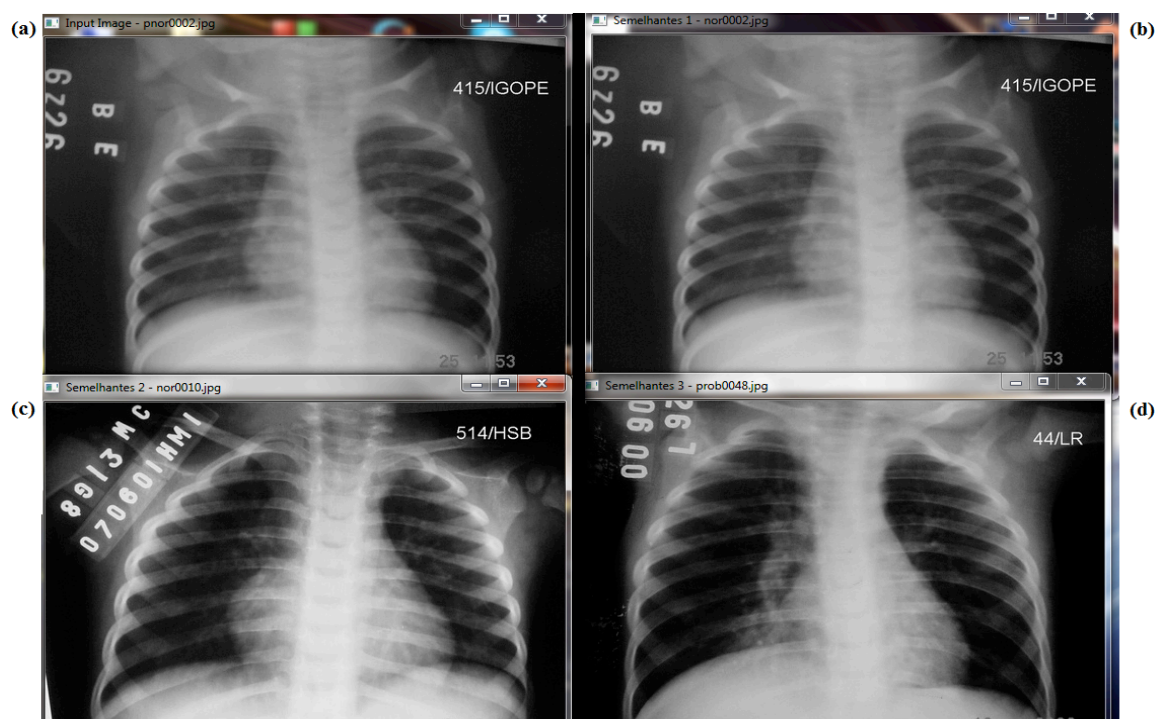


Figura 7 (a) Imagem de entrada (b) Primeiro resultado (c) Segundo Resultado (d) Terceiro resultado definidos pelo nível de semelhança

Conclusão

A base de imagens médicas, e a técnica de CBIR implementada para recuperação das mesmas tem como maior contribuição o fato de auxiliar no diagnóstico e tratamento de pessoas que apresentem exames semelhantes aos armazenados na base, uma vez que a comparação visual dos exames feito pelo CBIR poderá ser útil ao avaliar as condições de cada paciente e o tipo de tratamento a ser utilizado.

Este trabalho abre possibilidades de construção de um sistema de informação CAD que pode ser utilizado por clínicas ou hospitais na busca por casos semelhantes em exames baseados em imagens. Tal sistema de informação ofereceria suporte aos profissionais tornando os diagnósticos mais precisos, pois seria possível identificar rapidamente casos semelhantes com base nas imagens, mesmo em bases extensas com variados tipos de imagem.

Uma vez que o profissional de saúde recupere um, ou vários casos semelhantes, poderá definir o diagnóstico mais adequado, além de poder verificar qual tratamento teve melhor resultado nos casos que foram recuperados, para determinar o melhor ao paciente em questão e maximizar os resultados.

Pretende-se ainda implementar outros algoritmos para outros tipos de imagem, onde a técnica usada para recuperação seria escolhidos através de um limiar textual, cobrindo uma quantidade cada vez maior de tipos de imagem sem perder a eficiência da técnica. Uma vez que seja validado o uso em imagens 2D pretende-se implementar também técnicas com foco em imagens 3D no sistema.

REFERÊNCIAS

- Arampatzis, A., Zagoris, K., and Chatzichristofis, S. A. (2011). Dynamic two-stage image retrieval from large multimedia databases. In *Information Processing and Management*.
- Bergamasco, L.C.C., Nunes, F.L.S. (2013). Applying Distance Histogram to retrieve 3D cardiac medical models, In: *The American Medical Informatics Association Annual Symposium*.
- Borges, E. N., Cony, C. A., and de Freitas, A. L. C. (2005). Utilização de algoritmos de busca por similaridade em sgbd's comerciais.
- da Silva, C. Y. V. W. and Traina, A. J. M. Recuperação de imagens médicas por conteúdo utilizando wavelets e pca. In *Sociedade Brasileira de Informática em Saúde*.
- Danzhou L., Hua, K.A. and Sugaya, K. (2006). A Framework for Web-Based Interactive Applications of High-Resolution 3D Medical Image Data. *Computer-Based Medical Systems. CBMS 2006. 19th IEEE International Symposium on*. Digital Object Identifier: 10.1109/CBMS.2006.12. p. 119-124.
- de Macedo, S. O. (2012). Desenvolvimento de um sistema de auxílio ao diagnóstico de pneumonia na infância utilizando visão computacional. Mestrado, Instituto de Informática, Universidade Federal de Goiás.
- Chen T., Wang Y., Zhang H. and Xiao C. (2010). An embedded 3D medical image processing and visualization platform based on dual-core processor. *Intelligent Control and Automation (WCICA), 8th World Congress on*. Digital Object Identifier: 10.1109/WCICA.2010.5554388, p. 2936-2941.
- Chen, Z. Q., Zou, K. S., Ip, W. H. and Chan, C. Y. (2011). 3D model retrieval based on fuzzy weighted shape distributions. In: *Advanced Materials Research*. [S.l.]: Trans Tech Publications. v. 201-203, p. 1678-1681.
- Frank H., Olaf K., Horst K. H. and Heinz-Otto P. (2011). Interactive 3D medical image segmentation with energy-minimizing implicit functions. *Visual Computing in Biology and Medicine - Computers & Graphics*. Volume 35, Issue 2, p. 275-287.
- Haralick, R., Shanmugam, K. and Dinstein, I. (1973). Textural Features for Image Classification. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS*, vol. SMC-3, NO. 6.
- Hebbbar, H., C. N. U., and Mushigeri, S. (2013). Content based image retrieval based on cumulative distribution function a performance evaluation. In *International Journal of Computer Applications*.
- Kachouri, R., Djemal, K., Maaref, H., Masmoudi, D.S. and Derbel, N. (2008). Feature extraction and relevance evaluation for heterogeneous image database recognition. *Image Processing Theory, Tools and Applications. IPTA 2008. First Workshops on*. p. 1-6.
- Kim, J., Cai, W., Feng, D., and Wu, H. (2006). A new way for multidimensional medical data management: Volume of interest (VOI)-based retrieval of medical images with visual and functional features. *IEEE Transactions on Information Technology in Biomedicine*, 10(3), p. 598-606.
- Matheus, B. R. N. and Schiabel, H. (2010). Online Mammographic Images Database for Development and Comparison of CAD Schemes. *Journal of Digital Imaging*, p. 1618-727X.
- Müller, H., Michoux, N., Bandon, D., and Geissbühler, A. (2004). A review of content-based image retrieval systems in medical applications—Clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1), p. 1-23.
- Muneesawang, P. and L. Guan, M. (2006). Database retrieval: A human-centered approach. In *IEEE Computer Society*.
- Oliveira, L. L. G., e Silva, S. A., Ribeiro, L. H. V., de Oliveira, R. M., Coelho, C. J., and Andrade, A. L. S. S. (2008). Computer-aided diagnosis in chest radiography for detection of childhood pneumonia. In *International Journal of Medical Informatics*.
- Samet, H. (2010). Techniques for similarity searching in multimedia databases. In *The 36th International Conference on Very Large Data Bases*.
- Sanghavi, J. and Kayande, D. (2013). Content based image retrieval (cbir) system for diagnosis of blood related diseases. In *National Conference on Innovative Paradigms in Engineering e Technology*.
- Siadat, M., Soltanian-Zadeh, H., Fotouhi, F., and Elisevich, K. (2005). Content-based image database system for epilepsy. *Computer Methods and Programs in Biomedicine*, 79(3), p. 209-226.
- Vranic, D. V. (2001). 3D Model Retrieval. Tese (PhD Thesis) — University of Leipezig, Germany.
- Yang, F. and Leng, B. (2007). OFS: A feature selection method for shape-based 3D model retrieval. In: *Proceeding of 10th IEEE International Conference on Computer-Aided Design and Computer Graphics*. Beijing, China: IEEE Computer Society. p. 114-119.

Yuan, K., Tian, Z., Zou, J., Bai, Y. and You, Q. (2011). Brain CT image database building for computer-aided diagnosis using content-based image retrieval. Information Processing and Management. Vol.47(2), p. 176-185.

Apêndice I

Coeficiente de variação: Tem como foco analisar o peso do desvio-padrão em relação a média dos valores dos pixels das imagens. Dado pela equação:

$$CV = \frac{\sigma}{\mu}$$

Onde σ é o desvio padrão e μ é a média dos coeficientes.

Contraste: O contraste caracteriza-se pela diferença entre os tons de cinza. Baixo contraste ocorre quando há pequena diferença entre os níveis de cinza dos pixels localizados em uma região contígua da imagem. É dada pela equação:

$$Contraste = \sum_i \sum_j (i - j)^2 * p(i, j)$$

Onde $p(i, j)$ é (i,j)-ésima probabilidade de ocorrência de um nível de cinza na imagem.

Correlação: A correlação mede a dependência linear entre os tons de cinza presentes em uma imagem. Valores altos de correlação dados pela equação, próximos de um, indicam a existência de relação linear entre os pares de níveis de cinza.

$$Correlação = \frac{\sum_i \sum_j (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Onde $p(i, j)$ é (i,j)-ésima probabilidade de ocorrência de um nível de cinza na imagem e $\sigma_x, \sigma_y, \mu_x$ e μ_y são os desvios padrão e médias em x e y respectivamente.

Desvio médio: O Desvio Médio é uma medida da dispersão dos dados em relação à média de uma sequência, o “afastamento” em relação a essa média. Esta medida representa a média das distâncias entre cada elemento da amostra e seu valor médio. Descrito pela equação:

$$DesvMedio = \frac{\sum_i \sum_j |p(i, j) - \bar{x}|}{N}$$

Onde N é o número total de pixels na imagem e \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Variância: Descreve quanto os valores de tons de cinza estão dispersos em torno da média. É dada pela equação:

$$V = \frac{\sum_i \sum_j |(i, j) - \bar{x}|}{N}$$

Onde V é a variância de $P(i, j)$

Diferença da variância:

$$difVariância = V(p_{x-y})$$

$p_{x-y}(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = |i - j|$

Diferença da entropia:

$$difEntropia = - \sum_i p_{x-y}(i) \log\{p_{x-y}(i)\}$$

Energia: Também conhecida como segundo momento angular, expressa a uniformidade de uma textura, assumindo o valor máximo 1 e é expressa pela equação:

$$Energia = \sum_i \sum_j p(i, j)^2$$

Energia média: É a energia da imagem dividida pelo número total de pixels. Dado pela equação:

$$EnergiaMedia = Energia/N$$

Entropia: A medida de entropia expressa a desordem contida na textura. Quando uma imagem não é uniforme, as entradas apresentam valores próximos de zero o que resulta em valores altos de entropia. É dada pela equação:

$$Entropia = \sum_i \sum_j p(i, j) \log(p(i, j))$$

Momento de diferenças inversas: Quando a concentração dos valores na diagonal da matriz de co-ocorrência for máxima, o Momento da Diferença Inverso (MDI) atinge o valor máximo. É dado pela equação:

$$MDI = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$$

Resíduo médio: É a diferença média entre os dados observados e os valores ajustados pelo modelo (calculados). É dado pela equação:

$$ResMedio = \sum_i \sum_j ((i - j) - \bar{x})^2$$

Onde \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Soma média:

$$SumMedia = \sum_i i p_{x+y}(i)$$

$p_{x+y}(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = i + j$

Somatório da entropia:

$$SumEntropia = - \sum_i p_{x+y}(i) \log(p_{x+y}(i))$$

Somatório da variância:

$$SumVariancia = \sum_i (i - SumEntropia)^2 p_{x+y}(i)$$

Suavidade:

Analysis of Techniques of the Content-Based Image Retrieval to construct an Information System of the Computer-Aided Diagnosis

$$SV = 1 - \frac{1}{1 + \mu^2}$$

Desvio padrão: É a raiz quadrada da variância, dado pela equação:

$$\sigma = \sqrt{V}$$