

UNIVERSIDADE DE SÃO PAULO
Faculdade de Medicina de Ribeirão Preto
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto

HUGO CESAR PESSOTTI

**Desenvolvimento de um *Framework* para Classificação de
Doenças Pulmonares Difusas apoiado por Técnicas de
Recuperação de Imagens Baseada em Conteúdo e
Recuperação de Informação Textual**

RIBEIRÃO PRETO

2008

UNIVERSIDADE DE SÃO PAULO
Faculdade de Medicina de Ribeirão Preto
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto

HUGO CESAR PESSOTTI

**Desenvolvimento de um *Framework* para Classificação de
Doenças Pulmonares Difusas apoiado por Técnicas de
Recuperação de Imagens Baseada em Conteúdo e
Recuperação de Informação Textual**

Monografia apresentada a Faculdade de Filosofia
Ciências e Letras e a Faculdade de Medicina de
Ribeirão Preto como requisito parcial para obtenção
do grau de Bacharel em Informática Biomédica.

Orientadora: Profa. Dra. Alessandra Alaniz Macedo
Co-Orientador: Prof. Dr. Paulo Mazzoncini de Azevedo Marques

RIBEIRÃO PRETO
2008

Folha de Aprovação

HUGO CESAR PESSOTTI

**Desenvolvimento de um *Framework* para Classificação de Doenças
Pulmonares Difusas apoiado por Técnicas de Recuperação de Imagens
Baseada em Conteúdo e Recuperação de Informação Textual**

Monografia apresentada a Faculdade de Filosofia Ciências e Letras e a Faculdade de Medicina de Ribeirão Preto como requisito parcial para obtenção do grau de Bacharel em Informática Biomédica.

Aprovado em : _____

Conceito : _____

BANCA EXAMINADORA

Profa. Dra. Alessandra Alaniz Macedo
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/USP
Membro Titular

Prof. Dr. Jorge Elias Junior
Faculdade de Medicina de Ribeirão Preto/USP
Membro Titular

Prof. Dr. Luiz Otávio Murta Junior
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/USP
Membro Titular

"Não sabendo que era impossível, foi lá e fez."
-- Jean Cocteau

Aos meus pais, Sandra e José Roberto, pela excepcional educação que tive, pelo carinho, atenção e por sempre terem acreditado em minhas vitórias.

Agradecimentos

Agradeço aos meus pais e minhas irmãs por todos esses anos de boa convivência, essenciais para minha formação e educação, e por constituir um Lar acima de tudo.

A mim mesmo, por nunca duvidar de minha capacidade e por ter feito a escolha certa ao entrar no curso de Informática Biomédica.

A minha namorada e grande amiga, Lariza, pela confiança, carinho e dedicação.

Aos amigos do Laboratório de Informática em Saúde (LIS), por sempre criarem um ambiente ideal para trabalho e pesquisa, com seus momentos de descontração e estudo.

Aos amigos da Terceira Turma de Informática Biomédica, que seguiram a mesma caminhada durante esses quatro anos. Em particular, ao Fábio e Mário Sérgio, não esquecendo os “projetos secretos” que conduzimos durante as férias, por acreditarem firmemente em uma idéia que certamente virá a se concretizar no futuro.

A minha orientadora, Alessandra, pela paciência que teve comigo durante esses dois anos, pela seriedade que sempre levou nossos projetos e por sempre ter sido exigente com a qualidade de nossos trabalhos. Certamente aprendi mais coisas com ela do que poderia aprender durante meu curso de graduação.

Ao meu orientador Paulo Mazzoncini, pelas grandes idéias, discussões sempre produtivas e por sempre ter me recebido bem e doado um pouco de seu tempo nos momentos mais decisivos deste projeto.

Ao pessoal do Centro de Imagens do Hospital das Clínicas da FMRP, por terem fornecido as imagens utilizadas neste projeto. Em especial o mestrando Marcelo Ponciano, que fez o levantamento dos laudos e a separação das imagens de exames dos pacientes, e o Prof. Valdair Francisco Múglia, pelo apoio na classificação dos exemplos da base de imagens.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela bolsa de Iniciação Científica concedida, cujo projeto foi desenvolvido em paralelo a este trabalho de conclusão de curso.

Por fim, a todos que passaram e que irão passar pela minha vida, por deixarem em mim uma marca, cada um a sua maneira.

Resumo

Existem cerca de 180 patologias identificadas como Doenças Pulmonares Difusas (DPD). Grande parte dessas doenças apresenta características muito similares e geralmente são acompanhadas de sintomas não específicos, que dificultam seu diagnóstico. A modalidade de exame preferencial para o diagnóstico de DPD é a Tomografia Computadorizada de Alta Resolução, pois em muitos casos o uso de raios-X convencional não revela a presença da lesão. No entanto, devido ao grande número de imagens obtidas por esta modalidade, o diagnóstico torna-se um processo laborioso e sujeito a erros de interpretação. A Recuperação de Imagens Baseada em Conteúdo é uma técnica computacional que pode ser utilizada para ajudar os radiologistas na tarefa de analisar uma imagem médica, oferecendo-lhes uma segunda opinião e apontando estruturas de interesse que, em conjunto com técnicas de Recuperação de Informação, podem ser úteis para eliminar hipóteses e fornecer informações relevantes durante o diagnóstico, agregando conhecimento a sistemas médicos. Este trabalho busca explorar as áreas de recuperação de imagens baseada no conteúdo e recuperação de informação dentro do contexto de Doenças Pulmonares Difusas, para o desenvolvimento de um *Framework*¹. O trabalho foi desenvolvido em conjunto com médicos radiologistas do Serviço de Radiodiagnóstico do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (HCFMRP). Para validação do *framework* foi desenvolvida uma aplicação de prova de conceito a partir da integração dos componentes de recuperação de imagens e informação.

Palavras-chave: apoio ao diagnóstico, doenças pulmonares difusas, processamento de imagens médicas, recuperação de imagens baseada no conteúdo, recuperação de informação

¹ *Framework* é um termo utilizado para designar a tecnologia que oferece aos desenvolvedores de aplicações computacionais uma maneira poderosa de aumentar a reutilização de componentes de software [Johnson-1997].

Sumário

<i>CAPÍTULO 1</i> INTRODUÇÃO.....	1
<i>CAPÍTULO 2</i> FUNDAMENTOS TEÓRICOS.....	3
2.1. Doenças Pulmonares Difusas	3
2.2. Aquisição, Armazenamento e Exibição de Imagens de TCAR.....	5
2.3. Recuperação de Imagens Baseada em Conteúdo.....	6
2.4. Matrizes de Coocorrência e Descritores de Haralick.....	6
2.5. Transformada Discreta de Wavelets	8
2.6. K-Vizinhos Próximos	10
2.7. Redes Neurais Multicamada	11
2.8. Recuperação de Informação Textual	13
2.8.1. Pré-Processamento de Informação	14
2.8.2. Comparação de Documentos em Recuperação de Informação.....	15
<i>CAPÍTULO 3</i> ESPECIFICAÇÃO DOS REQUISITOS DO <i>FRAMEWORK</i>	16
3.1. Levantamento de Requisitos.....	16
3.1.1. Requisitos Funcionais	16
3.1.2. Requisitos Não-Funcionais	17
3.2. Modelos de Casos de Uso.....	18
3.2.1. Caso de Uso: Classificar Padrões em uma Imagem.....	18
3.2.2. Caso de Uso: Adicionar Exemplos à Base de Imagens.....	19
3.2.3. Caso de Uso: Busca de Imagens utilizando Palavras-Chave	20
3.2.4. Caso de Uso: Busca de Imagens similares utilizando um Laudo.....	20
3.2.5. Diagrama de Caso de Uso	21
<i>CAPÍTULO 4</i> PROJETO DOS COMPONENTES DO <i>FRAMEWORK</i>	22
4.1. Diagrama de Classes.....	22
4.2. Detalhamento das funcionalidades	23
<i>CAPÍTULO 5</i> EXPERIMENTAÇÃO E VALIDAÇÃO	25
5.1. Construção da Base de Imagens	25
5.2. Manipulação da Base de Imagens.....	26
5.2.1. Extração de Características	27
5.2.2. Classificação de Padrões	28
5.3. Experimentação da Base de Imagem considerando Recuperação de Imagens baseada em Conteúdo	28
5.4. Experimentação da Base de Imagens considerando Recuperação de Informação Textual	31
5.5. Experimentação da Base de Imagens considerando Recuperação de Imagens Baseada em Conteúdo apoiada por Recuperação de Informação Textual	34
<i>CAPÍTULO 6</i> CONCLUSÕES.....	37
<i>CAPÍTULO 7</i> REFERÊNCIAS	39

Lista de Figuras

FIGURA 1. COMPARATIVO DE PADRÕES RADIOLÓGICOS [UCHIYAMA-2003].	3
FIGURA 2. DIREÇÕES EM TORNO DE UM PIXEL COM DISTÂNCIA 1.	6
FIGURA 3. PSEUDO-ALGORITMO PARA CÁLCULO DA MATRIZ DE COCORÊNCIA EM UMA DADA DIREÇÃO E DISTÂNCIA.	7
FIGURA 4. REPRESENTAÇÃO BIDIMENSIONAL DOS FILTROS PASSA-BAIXA (L) E PASSA-ALTA (H).	8
FIGURA 5. PSEUDO-ALGORITMO PARA CÁLCULO DA DWT DE DAUBECHIES.	9
FIGURA 6. CLASSIFICAÇÃO VIA KNN COM DIFERENTES RAIOS DE ABRANGÊNCIA.	10
FIGURA 7. PSEUDO-ALGORITMO DO CLASSIFICADOR K-VIZINHOS PRÓXIMOS.	11
FIGURA 8. REDE NEURAL MULTICAMADAS.	11
FIGURA 9. PSEUDO-ALGORITMO UTILIZADO PARA TREINAR UMA REDE NEURAL MULTICAMADA.	13
FIGURA 10. DIGRAMA DE CASO DE USO REPRESENTANDO OS PRINCIPAIS EVENTOS.	21
FIGURA 11. DIAGRAMA DE CLASSES.	23
FIGURA 12. INTERFACE GRÁFICA PARA CRIAÇÃO DA BASE DE IMAGENS.	26
FIGURA 13. FLUXOGRAMA PARA A MANIPULAÇÃO DA BASE DE IMAGENS.	27
FIGURA 14. EXEMPLO DE GRADEADO SOBRE UMA IMAGEM.	27
FIGURA 15. CURVAS DE PRECISÃO E REVOCAÇÃO COMPARANDO EXTRATORES DE CARACTERÍSTICAS.	29
FIGURA 16. MATRIZ DE CONFUSÃO PARA O CLASSIFICADOR KNN E A COMBINAÇÃO DOS EXTRATORES.	30
FIGURA 17. MATRIZ DE CONFUSÃO PARA O CLASSIFICADOR MLP E A COMBINAÇÃO DOS EXTRATORES.	30
FIGURA 18. MATRIZ DE CONFUSÃO PARA O CLASSIFICADOR MLP E A COMBINAÇÃO DOS EXTRATORES, UTILIZANDO UM CONJUNTO MENOR DE PADRÕES.	31
FIGURA 19. CURVAS DE PRECISÃO-REVOCAÇÃO COMPARANDO O DESEMPENHO DA BUSCA TEXTUAL ENTRE OS DIFERENTES PADRÕES.	32
FIGURA 20. CURVAS DE <i>F-MEASURE</i> COMPARANDO O DESEMPENHO DA BUSCA TEXTUAL EM CONJUNTOS DE BUSCA DE TAMANHOS DISTINTOS.	33
FIGURA 21. MATRIZ DE CONFUSÃO PARA EXTRAÇÃO DE ATRIBUTOS VISUAIS E TEXTUAIS, UTILIZANDO A REPRESENTAÇÃO 0 E 1 PARA PALAVRAS-CHAVE.	35
FIGURA 22. MATRIZ DE CONFUSÃO PARA EXTRAÇÃO DE ATRIBUTOS VISUAIS E TEXTUAIS, UTILIZANDO A REPRESENTAÇÃO 0 E 0,25 PARA PALAVRAS-CHAVE.	36

Lista de Siglas

CBIR – *Content-Based Image Retrieval*

DICOM – *Digital Imaging and Communications in Medicine*

DPD – Doença Pulmonar Difusa

DWT – Transformada Discreta de Wavelets

HCFMRP – Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto

IR – *Information Retrieval*

KNN – *K-Nearest Neighbor*

MLP – *Multi-Layer Perceptron*

TCAR – Tomografia Computadorizada de Alta Resolução

UML – *Unified Modeling Language*

Capítulo 1

INTRODUÇÃO

As Doenças Pulmonares Difusas (DPD) formam um grupo de cerca de 180 patologias pulmonares que acometem o tecido intersticial do trato respiratório inferior e apresentam alta similaridade clínica, radiológica e fisiopatológica. Muitas dessas doenças são raras e geralmente acompanhadas de sintomas não específicos [Depeursinge-2007]. Seu diagnóstico é fortemente relacionado com as propriedades de textura do tecido afetado, devido ao caráter difuso destas patologias. [Azevedo-Marques-2006], [Martins-2005] e [Uchiyama-2003] sinalizam que o uso de atributos de textura, estatísticos e espectrais, têm se mostrado eficiente na detecção de doenças pulmonares difusas.

Embora o exame de raios-X tenha um custo considerável baixo e necessite de uma carga menor de radiação, esta não é a modalidade de exame indicada para o diagnóstico diferencial de DPD em função de sua baixa sensibilidade para variações sutis de padrões [Azevedo-Marques-2006; Azevedo-Marques-2007]. Em um estudo realizado por [Mathielson-1989], o diagnóstico de 118 pacientes com algum tipo de DPD foi avaliado corretamente em 93% dos casos através da tomografia computadorizada, enquanto a radiografia de tórax convencional permitiu uma taxa de acerto de 77%.

A Tomografia Computadorizada de Alta Resolução (TCAR) é um exame mais preciso que o exame de raios-X convencional para diagnósticos de DPD. A TCAR utiliza dados de cortes tridimensionais que evitam a superposição de órgãos internos durante a aquisição das imagens e permite uma análise detalhada do padrão de textura e distribuição do tecido pulmonar [Depeursinge-2007]. No entanto, o exame de TCAR gera uma quantidade significativa de planos e cortes de imagens a serem analisadas pelo radiologista, caracterizando-o como um processo trabalhoso e exaustivo. O diagnóstico de DPD é um grande desafio na área da radiologia [Doi-2005], mesmo para profissionais experientes, devido à variação de combinações de padrões radiológicos complexos e sua falta de correlação com os achados patológicos [Uchiyama-2003].

A idéia de utilizar o computador para auxiliar a análise de imagens radiográficas não é recente. Em 1964, [Meyers-1964] propôs um sistema para determinação automática da razão cárdio-torácica em radiografias de tórax. Em 1967, [Winsberg-1967] desenvolveu um sistema para análise automática de mamografias através de comparação bilateral. Porém, apesar desses esforços iniciais, esta área tornou-se uma linha de pesquisa na radiologia a partir dos

anos 80 [Doi-2005]. Técnicas de Inteligência Artificial têm sido aplicações com sucesso no auxílio ao diagnóstico de DPD, sendo os algoritmos mais comuns baseados em redes neurais artificiais [Ishida-1998; Uchiyama-2003; Azevedo-Marques-2007], redes bayesianas [Uchiyama-2003] e k-vizinhos próximos [Sluimer-2003].

Adicionalmente, os exames de TCAR são acompanhados por informações textuais inseridas pelo radiologista, além dos dados referentes ao formato DICOM, um padrão amplamente utilizado para armazenamento e transmissão de informações em imagens médicas. Estas informações são utilizadas pelo radiologista em conjunto com a análise das imagens, buscando eliminar hipóteses durante o diagnóstico diferencial. Nesse contexto, o uso de técnicas de Recuperação de Informação textual poderia ajudar a encontrar correlações entre as informações textuais e os padrões radiológicos, agregando conhecimento a sistemas médicos e complementando as hipóteses levantadas pelo radiologista.

Este trabalho tem como objetivo o desenvolvimento de um *framework* para análise quantitativa e classificação dos principais padrões radiológicos de Doenças Pulmonares Difusas em imagens de TCAR, apoiado por técnicas de Recuperação de Informação para correlacionar dados textuais presentes nos exames com os resultados a serem obtidos através de Recuperação de Imagens Baseadas em Conteúdo. Este *framework* deverá ser validado por meio da criação de uma aplicação de prova de conceito, utilizando matrizes de confusão e curvas de precisão-revocação² e *F-Measure*³ para avaliar o desempenho dos componentes desenvolvidos.

O restante desta monografia está organizado da seguinte maneira: o Capítulo 2 apresenta os fundamentos teóricos relacionados ao projeto, o Capítulo 3 trata da análise de requisitos do *framework*, o Capítulo 4 descreve os componentes desenvolvidos no projeto, o Capítulo 5 discute a metodologia e os resultados obtidos, o Capítulo 6 encerra o trabalho com as conclusões e vislumbra trabalhos futuros.

² Curvas de precisão-revocação analisam a precisão de algoritmos de classificação em função da porcentagem de documentos relevantes retornados, denominada revocação.

³ Curvas de F-Measure analisam a performance de algoritmos de classificação em função do tamanho do conjunto de busca, através da média harmônica entre as medidas de precisão e revocação.

Capítulo 2

FUNDAMENTOS TEÓRICOS

O estudo de doenças pulmonares difusas, imagens de TCAR, técnicas de recuperação de imagem baseada em conteúdo e recuperação de informação textual foram conceitos teóricos investigados para dar suporte ao trabalho apresentado nesta monografia. Esses conceitos e áreas são brevemente apresentados neste capítulo.

2.1. Doenças Pulmonares Difusas

Existem cerca de 180 patologias pulmonares classificadas no grupo de Doenças Pulmonares Difusas, que podem ser separadas em duas categorias: as doenças obstrutivas, caracterizadas pela limitação do fluxo de ar devido a uma obstrução parcial ou total das vias aéreas e as restritivas, caracterizadas pela redução do volume funcional dos pulmões, devido a distúrbios da parede torácica ou a ocorrência de doenças infiltrativas, agudas ou crônicas [Müller-2001]. Segundo [Martinez-1998], as DPDs classicamente englobavam os processos inflamatórios de caráter infiltrativo com evolução crônica ou subaguda, sendo mais tarde ampliado, incluindo-se as patologias de instalação aguda e também as patologias bronquiolares.

O diagnóstico de DPD é fortemente relacionado com as propriedades de textura do tecido afetado e os achados patológicos geralmente são descritos na forma de padrões radiológicos [Depeursinge-2007]. [Uchiyama-2003] destaca os padrões vidro fosco, opacidades reticulares e lineares, opacidades nodulares, favo de mel, áreas enfisematosas e consolidação pneumônica como sendo os principais, apresentados na Figura 1.

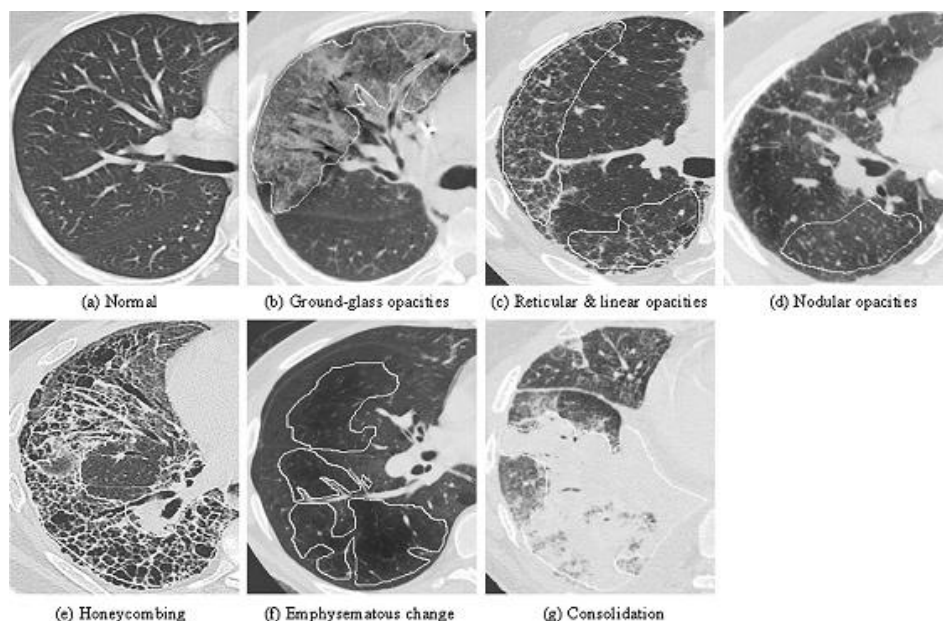


Figura 1. Comparativo de padrões radiológicos [Uchiyama-2003].

Os padrões radiológicos estudados neste projeto, conforme descritos no Consenso Brasileiro em TCAR do Tórax [Pereira-Silva-2005; Souza-2002], são:

- a) **Normal:** Considera-se como normal o padrão pulmonar homogêneo sem lesões aparentes, com todas suas estruturas bem definidas, desde os vasos sanguíneos até os brônquios.
- b) **Vidro Fosco:** Caracterizado pelo aumento da densidade do interstício pulmonar causado pela redução de ventilação dos espaços aéreos, seja pelo seu preenchimento parcial ou baixa inspiração, onde ainda é possível identificar os vasos e estruturas do tecido. A região torna-se esbranquiçada e menos definida, dando origem ao termo vidro fosco.
- c) **Opacidades Reticulares e Lineares:** É relacionado com o espessamento de estruturas do tecido pulmonar, resultando em uma imagem linear fina, alongada e com densidade mais elevada.
- d) **Opacidades Nodulares:** São estruturas aproximadamente arredondadas, totalmente ou parcialmente delimitadas, que apresentam densidade maior que o tecido intersticial em sua vizinhança. É chamado de micronódulo ou pequeno nódulo quando seu diâmetro é menor que 1 cm, sendo denominado nódulo caso esta medida não exceda 3 cm (acima disto, é denominado massa).
- e) **Favo de mel:** É caracterizado pela destruição de espaços aéreos periféricos, formando cistos pulmonares de paredes espessas e bem definidas com diâmetro variando de 0,3 a 1 cm, podendo atingir até 2,5 cm. Usualmente, são encontrados agrupados em camadas. Este padrão representa o estágio final da fibrose em parte do tecido intersticial.
- f) **Áreas Enfisematosas:** Caracterizadas pelo aumento permanente dos espaços aéreos, as áreas enfisematosas são resultantes da destruição das paredes alveolares, acarretando na diminuição da densidade do tecido intersticial, tornando a região mais escurecida na imagem tomográfica.
- g) **Consolidação Pneumônica:** É um padrão resultante do aumento acentuado da densidade do tecido pulmonar devido a diversos fatores, podendo incluir a substituição do ar alveolar por líquidos (resultante ou não de processos inflamatórios), sangue ou células (neoplásicas ou inflamatórias). A região torna-se esbranquiçada e não é mais possível identificar os vasos e estruturas do tecido, geralmente assumindo um caráter mais homogêneo.

2.2. Aquisição, Armazenamento e Exibição de Imagens de TCAR

A aquisição de imagens de TCAR é feita utilizando um equipamento que, através o mesmo princípio do raios-x convencional, é ajustado e construído para obter uma maior resolução durante a aquisição. Nesta modalidade, a espessura do corte é diminuída para aproximadamente 1-2mm, enquanto outros parâmetros internos são otimizados. Isto contribui para o aumento do número de cortes obtidos, ao custo de diminuir a velocidade do exame. O processamento é feito utilizando algoritmos específicos para reconstrução de imagens de alta resolução, usualmente envolvendo transformadas de Fourier, buscando diminuir a complexidade computacional devido ao número maior de cortes obtidos [Mayo-1991].

O armazenamento destas imagens segue o padrão DICOM, que define um conjunto de normas e regras para o tratamento, armazenamento e transmissão de informações médicas, mais especificamente em formato eletrônico. Estas regras formam um protocolo, permitindo a troca de informações entre equipamentos de imagens médicas e computadores, estabelecendo uma linguagem comum entre equipamentos de diferentes fabricantes que geralmente trabalham com formatos proprietários [Mildenberger-2002].

Toda imagem em formato DICOM contém um cabeçalho que, entre outras informações, armazena os dados da instituição que realizou o exame, a identificação e os dados antropométricos do paciente e especificações técnicas do equipamento médico utilizado. A imagem propriamente dita é armazenada com uma precisão de 16 bits, dos quais apenas 12 bits são efetivamente utilizados. Algumas implementações DICOM trabalham com o intervalo de 0 a 4096, enquanto outras utilizam o primeiro bit como sinal, com um intervalo de -2048 a 2048. Ao invés do pixel convencional, uma imagem DICOM armazena seus valores em uma escala própria. Em exames radiológicos, esta escala representa o coeficiente de atenuação de raios-X, que por sua vez representa a densidade do tecido.

A exibição de imagens de TCAR em formato DICOM é feita através da conversão da escala de atenuação em uma escala de tons de cinza, usualmente com 8 bits. Esta conversão é feita definindo-se um valor arbitrário na escala DICOM e uma janela ao seu redor. Em seguida, o limite inferior da janela é mapeado para o menor tom de cinza (preto) e o limite superior para o maior (branco). Desta forma, é possível exibir faixas específicas de atenuação, e, escolhendo corretamente os valores para o nível central e tamanho da janela, pode-se visualizar estruturas específicas, como por exemplo o pulmão. Esta operação de definição de valores é conhecida como janelamento e pode ser realizada sem a necessidade de aquisição de novas imagens.

2.3. Recuperação de Imagens Baseada em Conteúdo

Recuperação de Imagens baseada em Conteúdo (*Content-Based Image Retrieval - CBIR*) é uma área de pesquisa que busca criar sistemas com poder de busca e análise de imagens baseando-se em seu conteúdo: cor, textura e forma, que se manifestam através da distribuição dos pixels, unidades fundamentais formadoras das imagens digitais, que representam a quantização da luminosidade da imagem real ponto a ponto.

Em um sistema CBIR, cada imagem é representada por sua assinatura, ou vetor de características, a partir de um ou mais descritores. Descritores são funções que descrevem algum aspecto relacionado ao conteúdo de uma imagem. Esta assinatura é então enviada para uma base de dados, onde será indexada. As buscas são realizadas utilizando as assinaturas, e não as imagens propriamente ditas, sendo necessário um critério para definir a similaridade entre duas assinaturas e, por consequência, entre as imagens que representam. Essa similaridade é medida através de alguma função de distância, utilizada na tarefa de classificação de imagens.

2.4. Matrizes de Coocorrência e Descritores de Haralick

As matrizes de coocorrência fazem parte de um conjunto de técnicas de processamento de imagens voltadas para descrever a textura de uma imagem de uma maneira estatística. Esta abordagem leva em conta a distribuição e o relacionamento entre os níveis de cinza de uma imagem, armazenando a quantidade de pixels que apresentam um determinado nível de cinza e têm como vizinho outro pixel com outro nível de cinza [Haralick-1973].

Seja M a matriz de coocorrência de uma dada imagem. O relacionamento entre um par de níveis de cinza i e j pode ser calculado em quatro direções diferentes: 0, 45, 90 e 135 graus a partir do eixo x que passa através de um pixel (ver Figura 2). As outras direções (180, 225, 270 e 315 graus) podem ser obtidas a partir das quatro principais, pois seus vizinhos serão os mesmos, porém invertidos. Além da direção, é estabelecida uma distância entre o pixel central e sua vizinhança de interesse. Para pixels situados na borda, considera-se apenas os seus vizinhos que estão dentro da imagem.

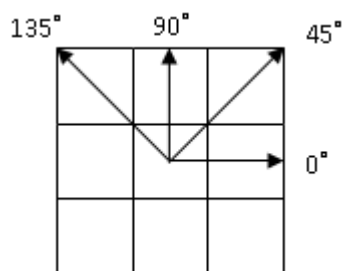


Figura 2. Direções em torno de um pixel com distância 1.

Dada uma distância δ e uma direção θ , a matriz de coocorrência M obtida a partir de uma imagem I de dimensões $n \times m$ e k níveis de cinza é calculada seguindo o pseudo-algoritmo apresentado na Figura 3:

```

função obterMatrizCooc(I,  $\delta$ ,  $\theta$ , k)

  para cada pixel (x,y) da imagem I
    i <- obterValorPixel(x, y)
    j <- obterValorPixelVizinho(x, y,  $\delta$ ,  $\theta$ )

    //construção da matriz de coocorrência
    M[i,j] <- M[i,j] + 1
    M[j,i] <- M[j,i] + 1
  fim para

fim

```

Figura 3. Pseudo-Algoritmo para cálculo da matriz de coocorrência em uma dada direção e distância.

A partir da matriz resultante das quatro direções é possível obter a matriz de coocorrência final através da soma de todas elas, seguida do cálculo de sua média. A seguir, a matriz é normalizada dividindo-a pela soma de todos seus elementos, de modo que a soma final de todos eles seja igual a 1.

Após obter a matriz de coocorrência, são calculadas métricas baseadas nos descritores de Haralick [Haralick-1973]. Em seu trabalho original, Haralick propôs 14 características (descritores), porém devido a correlação existente entre elas, apenas cinco são necessárias para obter bons resultados na classificação [Connors-1980]. Os cinco descritores mais utilizados são: energia (En), contraste (Con), correlação (Cor), entropia (Ent) e homogeneidade (Hom), definidos abaixo:

$$En = \sum_i \sum_j (M(i, j))^2 \quad (1)$$

$$Con = \sum_{k=0}^{m-1} k^2 \sum_{|i-j|=k} M(i, j) \quad (2)$$

$$Cor = \frac{1}{\sigma_i \sigma_j} \sum_i \sum_j i \cdot j \cdot M(i, j) - \mu_i \mu_j \quad (3)$$

$$Ent = \sum_i \sum_j M(i, j) \log(M(i, j)) \quad (4)$$

$$Hom = \sum_i \sum_j \frac{M(i, j)}{1 + (i - j)^2} \quad (5)$$

onde μ_i e σ_i são a média e variância horizontal e μ_j e σ_j a vertical.

2.5. Transformada Discreta de Wavelets

Outra abordagem comum no processamento de imagens é a espectral, representada pelo uso de transformadas discretas. Esta abordagem tem como objetivo transformar uma imagem (interpretada como um sinal) em uma função de frequência e tempo, denominada espectro.

Uma função bastante utilizada para o processamento de sinais é a Transformada de Fourier, que busca representar um sinal através da soma de funções senoidais, sendo apropriada para sinais periódicos. A Transformada de Fourier e suas transformadas derivadas (por exemplo, a Transformada Discreta do Cosseno) conseguem detectar a variação de tons da imagem através da análise dos picos de frequência, porém não armazenam informações sobre a localização. Além disso, para sinais não-periódicos (ex. imagens), a quantidade de elementos necessários para representá-los poderá ser maior que o sinal original.

A Transformada Discreta de Wavelets (DWT) busca contornar os problemas presentes na análise de sinais não-estacionários pela decomposição do sinal em diferentes escalas, utilizando bases de funções que são válidas apenas em determinados intervalos do sinal original. Esta abordagem não só permite que suas variações sejam detectadas, mas também localizadas [Daubechies-1992]. Seu funcionamento consiste na aplicação sucessiva de filtros de passa-alta e passa-baixa, dividindo o sinal original em duas componentes menores de alta e baixa frequência. Em uma imagem, este filtro é aplicado de maneira iterativa entre as linhas e colunas, dividindo-a em quatro imagens de frequências diferentes a cada iteração, representadas na Figura 4.

LL LL	LL HL	HL
LL LH	LL HH	
LH		HH

Figura 4. Representação bidimensional dos filtros passa-baixa (L) e passa-alta (H).

Existem várias transformadas de wavelets propostas, sendo a mais simples a DWT de Haar. Esta transformada divide o sinal de entrada em dois sinais de comprimento duas vezes menor que o original, baseando-se nas operações de soma e diferença. O primeiro é resultante da soma de pares sucessivos do sinal original e representa a componente de baixa frequência,

enquanto o segundo é obtido pela diferença entre os mesmos pares, representando a alta frequência [Mahmmoud-2007]. Repete-se esta seqüência de operações para todas as linhas de uma imagem e em seguida para todas as suas colunas, até obter uma imagem resultante da aplicação dos filtros conforme a figura anterior. As operações de soma e diferença podem ser ponderadas, dando origem a operações de média e diferença média.

No entanto, segundo [Wang-1997], a DWT de Haar não apresenta as propriedades de transição acentuada e rápida atenuação do sinal, impedindo que a imagem seja separada em componentes claras e distintas. Parte dessa limitação é decorrente da simplicidade do filtro, por usar apenas dois pontos sucessivos. A DWT de Daubechies [Daubechies-1992] é uma extensão da DWT de Haar, buscando preservar uma quantidade maior de informações do sinal original através do uso de janelas maiores que se sobrepõem.

Dada uma janela de tamanho k , a DWT de Daubechies obtida a partir de uma imagem I de dimensões $n \times m$ é calculada seguido o pseudo-algoritmo definido na Figura 5.

```

função obterDWTDaubechies(I, n, m, k)

//aplicação dos filtros nas linhas da imagem
para y de 1 a m
  para x de 1 a n passo 2
    s <- somaJanelaHorizontal(I, x, y, k)
    d <- diferençaJanelaHorizontal(I, x, y, k)
    temp[x/2,y] <- s
    temp[n/2+x/2,y] <- d
  fim para
fim para

//aplicação dos filtros nas colunas do passo anterior
para x de 1 a n
  para y de 1 a m passo 2
    s <- somaJanelaVertical(temp, x, y, k)
    d <- diferençaJanelaVertical(temp, x, y, k)

    saída[x,y/2] <- s
    saída[x, n/2+y/2] <- d
  fim para
fim para

fim

```

Figura 5. Pseudo-Algoritmo para cálculo da DWT de Daubechies.

Dada uma janela de tamanho 4, as equações de soma e diferença podem ser escritas como sendo:

$$s = h_0x_i + h_1x_{i+1} + h_2x_{i+2} + h_3x_{i+3} \quad (6)$$

$$d = h_3x_i - h_2x_{i+1} + h_1x_{i+2} - h_0x_{i+3} \quad (7)$$

onde i é o índice do elemento cuja janela será calculada e h são os coeficientes da DWT de Daubechies 4-tap, a seguir: $h_0 = 0.4830$; $h_1 = 0.8365$; $h_2 = 0.2241$; $h_3 = -0.1294$.

2.6. K-Vizinhos Próximos

Dentre todos os algoritmos utilizados na classificação e reconhecimento de padrões através de exemplos, o KNN (*K-Nearest Neighbor*) está entre os mais simples. O funcionamento deste algoritmo está ligado à idéia de que elementos de uma mesma classe geralmente estão mais próximos entre si, considerando uma função de distância e algum atributo para comparação.

Dado um elemento de classe desconhecida, o algoritmo tenta primeiro encontrar os seus k vizinhos mais próximos e utiliza suas classes para determinar as classes candidatas. A quantidade de vizinhos pode ser qualquer número inteiro maior que 1, porém é recomendável que seja um número ímpar para evitar situações de empate.

A medida de similaridade entre o elemento desconhecido e um de seus vizinhos é utilizada como peso para a classe correspondente, e se um ou mais vizinhos compartilham a mesma classe, esta classe recebe uma ponderação maior [Yang-1999]. Após obter todas as classes candidatas, estas são ordenadas de acordo com o peso recebido e uma é escolhida, usualmente a que atingiu a maior pontuação. A Figura 6 ilustra um possível espaço de classificação, partindo do círculo central e utilizando 3 e 5 vizinhos próximos.

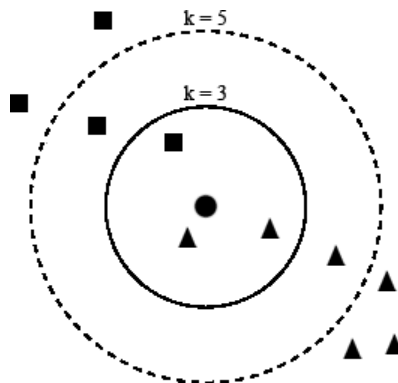


Figura 6. Classificação via kNN com diferentes raios de abrangência.

A similaridade entre duas imagens pode ser calculada através da função de distância Minkowski, descrita a seguir:

$$d(A, B) = \sqrt[p]{\sum_{i=1}^n |a_i - b_i|^p} \quad (8)$$

onde a e b são os vetores de características das imagens, n o número de atributos e p a norma da distância (para $p=2$, tem-se a distância Euclidiana). Quanto menor for a distância entre duas imagens, maior será a similaridade entre elas.

Considerando a distância Euclidiana e uma ponderação baseada somente no número de classes encontradas, a classificação de um elemento t utilizando um conjunto de exemplos x e k vizinhos pode ser feita utilizando pseudo-algoritmo apresentado na Figura 7.

```

função classificadorKNN(x, t, k)

  a <- vetorCaracteristica(t)

  para cada exemplo i
    b <- vetorCaracteristica(x)

    dist[i] <- 0
    para cada atributo j
      dist[i] <- dist[i] + potência(a[j] - b[j], 2)
    fim para

    dist[i] <- raizQuadrada(dist)
  fim para

  ordenar dist em ordem crescente

  //retorna a classe mais frequente dos vizinhos
  //considerando as k menores distâncias
  classe <- classeMajoritária(dist, k)

fim

```

Figura 7. Pseudo-Algoritmo do classificador K-Vizinhos Próximos.

2.7. Redes Neurais Multicamada

Também conhecida como MLP (*Multi-Layer Perceptron*), este tipo de Rede Neural é composta de elementos de processamento interconectados chamados neurônios [Haykin-2001] dispostos conforme a Figura 8.

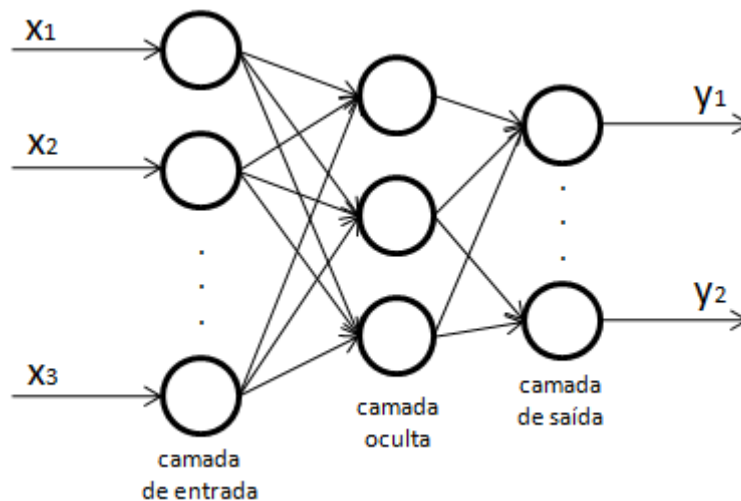


Figura 8. Rede Neural Multicamadas.

Classificadores baseados em MLP funcionam através do mapeamento de um conjunto de neurônios de entrada para um conjunto de neurônios de saída, utilizando uma camada oculta de neurônios que se conectam através de pesos. Cada neurônio i de uma camada é ligado a um neurônio j da camada anterior através de um peso w_{ij} , utilizado para representar a conexão entre eles.

De maneira geral, a saída y_j de um neurônio depende de suas conexões de entrada o_i , calculada através de uma soma ponderada de peso w_{ij} . Após o cálculo da contribuição de cada conexão de entrada, o resultado final é obtido através da aplicação de uma função de ativação, geralmente uma sigmoideal, conforme descrito nas equações abaixo:

$$y_j = \varphi(u_j(n)) \quad (9)$$

onde u_j representa a somatória ponderada:

$$u_j(n) = \sum_i w_{ji} o_i \quad (10)$$

utilizando a função de ativação logística:

$$\varphi(x) = \frac{1}{1 + e^{-ax}} \quad (11)$$

O treinamento de uma rede MLP é feito em duas etapas: *forward* e *backpropagation*. Na fase *forward*, os pesos w_{ij} são calculados partindo da camada de entrada, até chegar à camada de saída. Conhecida a classe d do exemplo fornecido na entrada, ao finalizar a fase *forward* é calculada a energia total do erro, segundo a equação a seguir.

$$E(n) = \frac{1}{2} \sum_j (e_j(n))^2 \quad (12)$$

onde e_j é o erro instantâneo, dado pela diferença entre o valor desejado d_j e a saída y_j :

$$e_j(n) = d_j(n) - y_j(n) \quad (13)$$

Caso a energia do erro seja maior que um valor pré-definido, todos pesos são corrigidos na fase *backpropagation* de acordo com a seguinte regra:

$$\Delta w_{ji}(n) = -\eta \delta_j(n) o_i(n) \quad (14)$$

onde η é a taxa de aprendizado da rede e δ_j representa o gradiente local do erro. Se o neurônio está na camada de saída, este gradiente é calculado por:

$$\delta_j(n) = \varphi'(u_j(n)) e_j(n) \quad (15)$$

e se estiver em uma camada oculta:

$$\delta_j(n) = \varphi'(u_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (16)$$

Uma vez que o erro seja minimizado segundo algum critério (calculando seu erro médio quadrático, por exemplo) e atinja um nível aceitável, a fase de treinamento é finalizada e os pesos finais podem ser usados para calcular a saída de um exemplo de classe desconhecida.

Seja x um conjunto de exemplos e η a taxa de aprendizado da rede. Simplificando as fases *forward* e *backpropagation*, o MLP pode ser descrito segundo o pseudo-código apresentado na Figura 9.

```

função MLP(x, η)

    //inicializa os pesos com valores aleatórios
    para cada peso w[i,j]
        w[i,j] <- rand()
    fim para

    para cada exemplo i
        d[i] <- classe(x[i])
    fim para

    enquanto (critério de parada não satisfeito) faça

        //fase forward
        para cada camada k partindo da camada de entrada
            para cada neurônio j da camada k
                calcular saída y[j]
            fim para
        fim para

        //fase backpropagation
        para cada camada k partindo da camada de saída
            para cada neurônio j da camada k
                para cada neurônio i da camada k-1
                    atualizar pesos w[j,i]
                fim para
            fim para
        fim para

    fim enquanto

fim

```

Figura 9. Pseudo-Algoritmo utilizado para treinar uma Rede Neural Multicamada.

2.8. Recuperação de Informação Textual

Pesquisadores da área de Recuperação de Informação (*Information Retrieval* - IR) investigam maneiras de manipular a representação, o armazenamento, a organização e o acesso à informação [Baeza-Yates-1999], sendo o foco a informação que usuários precisam. Para buscar a informação desejada torna-se necessária a modelagem da coleção de documentos, onde cada documento é representado por um conjunto de palavras-chave representativas (termos de índice).

Em geral, os pesquisadores tendem a focar esforços na identificação de estruturas organizacionais, conteúdo e significado. Assim, sistemas tradicionais de IR manipulam dados desestruturados e são incapazes de suportar a manipulação de outros tipos de mídias (por ex. imagens). Este projeto foge ao grupo de sistemas de IR tradicionais, uma vez que manipulará imagens médicas e laudos radiológicos. Nesse sentido, é necessária uma abordagem que seja capaz de manipular os laudos pelo uso de algoritmos de IR e imagens radiológicas através de técnicas de CBIR.

2.8.1. Pré-Processamento de Informação

As técnicas de IR baseiam-se amplamente no processamento textual de informações apresentadas em linguagem natural. O pré-processamento é uma tarefa importante de IR a qual inclui, por exemplo, remoção de *Stop-Words* e *Stemming*. Essas tarefas de pré-processamento são fundamentais para a adequação da entrada de dados para os algoritmos de IR.

Stop-Words são palavras que não apresentam um significado semântico dentro do contexto estudado, ou então simplesmente palavras que devem ser filtradas durante o processamento textual dos dados. Não existem algoritmos específicos para remoção de *stop-words*. Algoritmos devem ser criados a partir de listas de palavras específicas a serem removidas do texto que, geralmente, contém artigos, preposições e advérbios.

Esta técnica de pré-processamento é importante na diminuição do número de termos de um documento, reduzindo o custo computacional dos algoritmos de processamento sem alterar o significado semântico dos documentos. A lista de *Stop-Words* mais comum é mantida pelo *Snowball Project*, disponível em treze línguas, incluindo a portuguesa (<http://snowball.tartarus.org/>).

Stemming, ou radicalização, é uma técnica utilizada para extrair o radical de uma palavra através da remoção de sufixos. *Stemming* permite que palavras de mesmo significado sejam comparadas, independente de sua variação de gênero, número, grau e conjugação, como no caso de verbos. Esta técnica pode ser explorada de duas maneiras: (1) uso de um dicionário de radicais e (2) uso de algoritmos específicos para remover sufixos, baseado em regras. [Viera-2007] apresenta e compara os principais algoritmos de radicalização para a língua portuguesa, destacando a adaptação do algoritmo de Porter e de Orengo. Em especial, na língua portuguesa, deve ser observada a acentuação das palavras e a cedilha, pois em alguns casos o significado da palavra pode mudar significativamente (por ex. *contem* e *contém*).

O funcionamento geral do algoritmo de Porter é baseado na remoção de sufixos sob critérios que devem ser atendidos antes de sua aplicação, resumindo-se em cinco regras, sendo cada uma composta de uma lista própria de sufixos. Já o algoritmo de Orengo é baseado em regras de remoção de sufixos que podem eventualmente apresentar exceções, que devem ser tratadas antes de sua aplicação. No total, o algoritmo de Orengo apresenta 199 regras e 32 mil termos de exceções.

2.8.2. Comparação de Documentos em Recuperação de Informação

A similaridade entre documentos desestruturados geralmente é medida através do grau de similaridade de seus conteúdos, partindo do princípio que documentos que compartilham termos em comum têm uma maior probabilidade de serem similares. No entanto, nessa abordagem a semântica da informação é desprezada.

A importância de um termo para um documento pode ser determinada pela sua frequência. Esta importância aumenta proporcionalmente ao número de vezes que o termo aparece no documento e inversamente proporcional ao número de vezes que este mesmo termo aparece no restante dos documentos da coleção. Conseqüentemente, palavras comuns têm sua importância reduzida quanto mais aparecerem em outros documentos e palavras restritas ao contexto do documento têm seu peso aumentado. Esta medida de importância de termos é conhecida como TF-IDF (*term frequency-inverse document frequency*) e é amplamente utilizada em algoritmos de IR. Dado uma coleção de D documentos e um termo t_i , a medida TF-IDF é calculada segundo a equação a seguir:

$$TFIDF_i = \frac{n_i}{N} \log\left(\frac{D}{d_i}\right) \quad (17)$$

onde n_i é o número de vezes que o termo t_i aparece no documento, N o número total de termos do documento e d_i o número de documentos na coleção que apresentam o termo t_i . Sendo assim, o perfil de um documento pode ser definido como sendo um vetor de características onde cada posição representa um termo e sua importância no documento segundo a medida TF-IDF. Desta maneira, a similaridade entre dois documentos pode ser medida com base nos perfis de termos, normalmente pontuada por alguma função de *ranking* como, por exemplo, a medida de similaridade do cosseno, calculada pela equação:

$$\cos(d_i, d_j) = \frac{\sum_k d_{ki} d_{kj}}{\sqrt{\sum_k d_{ki}^2} \sqrt{\sum_k d_{kj}^2}} \quad (18)$$

onde k é o número de termos presentes na coleção de documentos e d_{ki} e d_{kj} representam a importância do termo t_k para os documentos d_i e d_j , respectivamente, utilizando a medida TF-IDF. O resultado é o cosseno do ângulo formado pelos dois vetores de pesos, onde valores negativos revelam que os documentos são opostos, um resultado próximo a zero indica que não há termos comuns entre eles. Os valores positivos indicam que os documentos são similares. Quanto mais próximo de 1, maior a similaridade entre os documentos, somente atingindo 1 quando os termos são equivalentes e aparecem com a mesma frequência.

ESPECIFICAÇÃO DOS REQUISITOS DO *FRAMEWORK*

O potencial de executar processos automatizados para apoiar tarefas diárias de radiologistas motiva o desenvolvimento deste trabalho. A formalização dessas tarefas em um modelo conceitual desvincula a necessidade de terminologia específica que cada software de apoio ao diagnóstico em radiologia utiliza. Este capítulo detalha o levantamento e a especificação de requisitos que os componentes de software do *framework* proposto deverão atender. O *framework* segue o formalismo de Orientação a Objetos modelados com a Linguagem de Modelagem Unificada (*Unified Modeling Language* - UML).

3.1. Levantamento de Requisitos

Em Engenharia de Software, levantamento de requisitos é uma técnica utilizada para esclarecer “o que” o sistema deve fazer e “como” as tarefas devem ser executadas, identificar os principais usuários e formalizar suas necessidades na forma de requisitos técnicos de software, descrever o ambiente técnico de suporte ao sistema e detectar precocemente problemas que poderão afetar a etapa de desenvolvimento [Pressman-2006].

Idealmente, a tarefa de levantamento de requisitos deve ser realizada antes que o desenvolvimento se inicie, e como produto tem-se a elaboração de um documento de requisitos, contendo descrições do sistema, casos de uso e diagramas que norteiam as etapas de desenvolvimento, validação, implantação e manutenção do sistema [Pressman-2006].

3.1.1. Requisitos Funcionais

Os requisitos funcionais representam as funcionalidades que um sistema deve atender, dividindo-se em requisitos evidentes e requisitos ocultos. Requisitos evidentes são aqueles que o usuário tem ciência de sua existência como, por exemplo, ampliar uma região de uma imagem. Já os ocultos são executados pelo sistema de maneira transparente ao usuário, geralmente servindo de suporte a requisitos evidentes. A extração de características de uma imagem pode ser considerada como sendo um requisito oculto, pois o usuário não sabe como e quando ele está sendo executada, porém ela é fundamental para que o sistema funcione corretamente.

Os sistemas suportados pelos componentes do *framework* proposto poderão se beneficiar de diversas funcionalidades, agrupadas segundo o tipo de tarefa que realizam. A seguir, são apresentados os requisitos funcionais mais importantes para o *framework*. Os requisitos evidentes foram marcados com um (E), enquanto os ocultos receberam a marcação (O).

Sistema

R1.1 Garantir controle de acesso mediante *login* e senha (E)

Extração e Quantificação de Características

R2.1 Obter a matriz de Co-Ocorrência de um conjunto de pixels (O)

R2.2 Calcular métricas a partir de uma matriz de Co-Ocorrência utilizando os cinco principais descritores de Haralick (O)

R2.3 Aplicar a transformada de Wavelets de Daubechies em um conjunto de pixels (O)

R2.4 Calcular métricas a partir de uma Wavelet de Daubechies (O)

R2.5 Gerar um vetor de características com base nas métricas calculadas (O)

Classificação

R3.1 Adicionar exemplos de padrões radiológicos à base de imagem (E)

R3.2 Gerar um classificador baseado em K-Vizinhos Próximos (O)

R3.2 Gerar um classificador baseado em Redes Neurais (O)

R3.3 Treinar um classificador utilizando um conjunto de exemplos (O)

R3.4 Validar um classificador utilizando um conjunto de exemplos (O)

R3.5 Obter a classificação de uma região de uma imagem utilizando um classificador (E)

Visualizador de Imagens

R4.1 Abrir uma seqüência de imagens dentro de um mesmo diretório (E)

R4.2 Abrir uma imagem no formato DICOM (E)

R4.3 Permitir a alteração do janelamento de uma imagem (E)

R4.4 Permitir a ampliação de uma região de uma imagem (E)

R4.5 Destacar uma região de uma imagem para exibir o resultado de sua classificação (E)

R4.5 Permitir a seleção de regiões de interesse (E)

R4.6 Obter os valores de pixel na escala DICOM de uma região de uma imagem (O)

R4.7 Subdividir uma imagem em regiões de mesmo tamanho (O)

Recuperação de Informações

R5.1 Obter os dados textuais do laudo de um paciente (O)

R5.2 Realizar uma busca por palavras-chave (E)

R5.3 Realizar uma busca por similaridade textual (E)

3.1.2. Requisitos Não-Funcionais

Requisitos não-funcionais definem os atributos de um sistema, geralmente especificando fatores de qualidade e restrições do sistema. Foram identificados três requisitos não-funcionais que o *framework* deverá seguir, sendo eles:

Integridade: Como a identidade dos pacientes envolvidos nos exames não pode ser retirada das imagens pelo fato de ser necessária para a conclusão do laudo, o acesso ao sistema deve ser restrito apenas aos radiologistas, não permitindo qualquer tipo de acesso externo.

Simplicidade: O sistema deverá funcionar da forma mais simples possível, para facilitar as tarefas que envolvam interação com o usuário. Qualquer dificuldade encontrada pode atrapalhar o radiologista e até mesmo o confundir, podendo levar a erros de diagnóstico.

Tempo de Resposta: O sistema deve ter um tempo de resposta psicologicamente aceitável. Se as tarefas de quantificação e classificação de padrões demorarem muito, deixarão de auxiliar o radiologista e irá aumentar ainda mais o tempo necessário para a análise de um conjunto de imagens.

Portabilidade: O sistema deverá ser compatível com diferentes arquiteturas de hardware e sistemas operacionais, sendo capaz de ser compilado e executado independente da plataforma utilizada, mantendo todas suas funcionalidades.

Modularidade: O projeto do sistema deve ser composto por diferentes módulos que, através de interfaces conhecidas e bem definidas, podem ser separados, trocados ou recombinados para dar origem a outros subsistemas, sendo cada módulo responsável por um conjunto de funcionalidades.

3.2. Modelos de Casos de Uso

Casos de Usos contam histórias de uso de um conjunto de funcionalidades em um determinado cenário, considerando as interações dos atores com essas funcionalidades. A seguir são apresentados os quatro principais eventos de um possível cenário de uso dos requisitos especificados para um sistema computacional de apoio ao diagnóstico em radiologia. Sequências alternativas serão consideradas, apresentando a resposta do sistema caso a ação do usuário seja diferente da esperada.

3.2.1. Caso de Uso: Classificar Padrões em uma Imagem

Resumo

Caso de Uso:	Classificar Padrões em uma Imagem
Atores:	Radiologista
Finalidade:	Classificar padrões em uma imagem DICOM
Visão geral:	O técnico obtém uma seqüência de imagens radiológicas em formato DICOM após a realização de um exame e as encaminha para o radiologista. Ao abrir a seqüência de imagens, o radiologista escolhe a mais representativa e seleciona a opção de Classificação de Padrões. O sistema localiza os padrões presentes na imagem e os destaca.

Seqüência típica de eventos

Ação do Ator	Resposta do Sistema
1. O técnico realiza um exame, obtém as imagens do paciente no formato DICOM e as encaminha para o radiologista.	
2. O radiologista abre a seqüência de imagens em formato DICOM.	3. O sistema exibe as imagens na tela.
4. O radiologista escolhe a imagem mais representativa da seqüência.	
5. O radiologista seleciona a opção de Classificação de Padrões.	6. O sistema processa a imagem em busca de padrões e destaca aqueles que encontrar.

Seqüências alternativas

Evento 2. O radiologista tenta abrir uma imagem que não está no formato DICOM. O Sistema deve emitir uma mensagem de aviso e rejeitar o arquivo.

Evento 5. O radiologista escolhe uma imagem que não contém nenhum padrão radiológico. O sistema deve emitir uma mensagem avisando que nada foi encontrado.

3.2.2. Caso de Uso: Adicionar Exemplos à Base de Imagens

Resumo

Caso de Uso:	Adicionar Exemplos à Base de Imagens
Atores:	Radiologista
Finalidade:	Adicionar uma imagem DICOM à base de imagens para ser utilizada por um classificador
Visão geral:	O radiologista solicita ao técnico uma seqüência de imagens do exame de um paciente já diagnosticado. O técnico obtém as imagens no formato DICOM e as encaminha para o radiologista. O radiologista abre a seqüência de imagens, seleciona a mais representativa e cria uma região ao redor do padrão radiológico utilizando a ferramenta de Região de Interesse. Criada a seleção, o radiologista seleciona a opção de Inserir Padrão, informa sua classificação e a adiciona à base de imagens.

Seqüência típica de eventos

Ação do Ator	Resposta do Sistema
1. O radiologista solicita ao técnico as imagens de um exame já realizado.	
2. O técnico obtém as imagens no Formato DICOM e as encaminha para o radiologista.	
3. O radiologista abre a seqüência de imagens em formato DICOM.	4. O sistema exibe as imagens na tela.
5. O radiologista escolhe a imagem mais representativa da seqüência.	
6. O radiologista seleciona a ferramenta de Região de Interesse e cria uma região ao redor do padrão radiológico.	7. O sistema destaca a área selecionada pelo radiologista.
8. O radiologista seleciona a opção de Inserir Padrão.	9. O sistema exibe a janela de Inserir Padrão.
10. O radiologista informa a classificação do padrão destacado.	11. O sistema insere o novo exemplo na base de imagens

Seqüências alternativas

Evento 3. O radiologista tenta abrir uma imagem que não está no formato DICOM. O sistema deve emitir uma mensagem de aviso e rejeitar o arquivo.

Evento 8. O radiologista seleciona a opção Inserir Padrão sem ter criado uma região. O sistema deve emitir uma mensagem de aviso e não abrir a janela de Inserir Padrão.

Evento 10. O radiologista não informa a classificação do padrão destacado. O sistema deve emitir uma mensagem de aviso e não adicionar o novo exemplo na base de imagens.

3.2.3. Caso de Uso: Busca de Imagens utilizando Palavras-Chave

Resumo

Caso de Uso:	Busca de Imagens utilizando Palavras-Chave
Atores:	Radiologista
Finalidade:	Retornar imagens a partir de seus laudos utilizando palavras-chave
Visão geral:	O radiologista determina as palavras-chave para realizar uma busca e as informa através da opção Busca por Palavra-Chave. O sistema retorna os resultados da busca sob a forma de imagens DICOM. O radiologista pode então selecionar os resultados mais relevantes para visualizar as imagens.

Seqüência típica de eventos

Ação do Ator	Resposta do Sistema
1. O radiologista determina as palavras-chave que deseja utilizar para realizar uma busca.	
2. O radiologista seleciona a opção Busca por Palavra-Chave.	3. O Sistema exibe a janela de Busca por Palavra-Chave.
3. O radiologista informa as palavras-chave para realizar a busca.	4. O Sistema retorna os resultados da busca sob a forma de imagens DICOM.
5. O radiologista seleciona os resultados mais relevantes e abre as imagens.	6. O Sistema exibe as imagens na tela.

Seqüências alternativas

Evento 3. O radiologista tenta iniciar uma busca sem ter informado ao menos uma palavra-chave. O sistema deve emitir uma mensagem de aviso e não realizar a busca.

Evento 5. O radiologista pode preferir salvar as imagens. O sistema deve perguntar o local desejado e salvar os resultados.

3.2.4. Caso de Uso: Busca de Imagens similares utilizando um Laudo

Resumo

Caso de Uso:	Busca de Imagens similares utilizando um Laudo
Atores:	Radiologista
Finalidade:	Retornar imagens similares a uma já existente a partir de seus laudos
Visão geral:	O radiologista solicita ao técnico o laudo de um paciente já diagnosticado. O técnico obtém o laudo e o encaminha para o radiologista. O radiologista informa o laudo ao sistema através da opção Busca de Imagens por Laudo. O sistema retorna os resultados da busca por similaridade sob a forma de imagens DICOM. O radiologista pode então selecionar os resultados mais relevantes para visualizar as imagens.

Seqüência típica de eventos

Ação do Ator	Resposta do Sistema
1. O radiologista solicita ao técnico o laudo de um exame já realizado.	
2. O radiologista seleciona a opção Busca de Imagens por Laudo.	3. O sistema exibe a janela de Busca de Imagens por Laudo.
4. O radiologista informa o laudo para realizar a busca por similaridade.	5. O sistema retorna os resultados da busca por similaridade sob a forma de imagens DICOM.
6. O radiologista seleciona os resultados mais relevantes e abre as imagens.	7. O sistema exibe as imagens na tela.

Seqüências alternativas

Evento 4. O radiologista tenta iniciar uma busca sem ter informado um laudo. O sistema deve emitir uma mensagem de aviso e não realizar a busca.

Evento 6. O radiologista pode preferir salvar as imagens. O sistema deve perguntar o local desejado e salvar os resultados.

3.2.5. Diagrama de Caso de Uso

Para ilustrar os Casos de Uso descritos anteriormente, foi criado um Diagrama de Caso de Uso, visualizado na Figura 10, representando o Ator principal (Radiologista) e os Casos de Uso com os quais esse ator está envolvido.

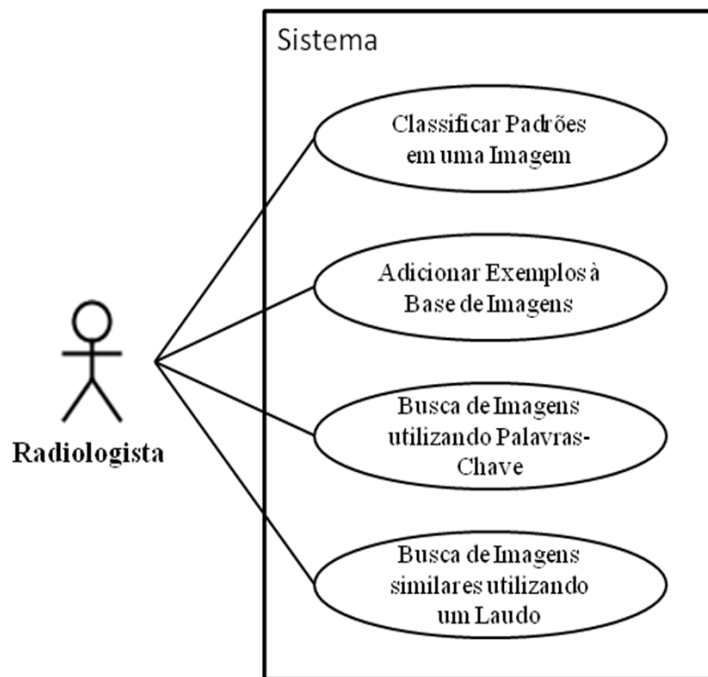


Figura 10. Diagrama de Caso de Uso representando os principais eventos.

Capítulo 4

PROJETO DOS COMPONENTES DO *FRAMEWORK*

O *framework* apresentado neste trabalho é uma tentativa de um projeto genérico para manipular laudos e imagens de DPD. O uso de técnicas de Orientação a Objetos torna o projeto modularizável, reutilizável e de fácil manutenção. Até o momento, o projeto apresenta um diagrama de classes, que é a base para a modelagem dos componentes do sistema, com suas características e funcionalidades. Esse diagrama de classes e as funcionalidades oferecidas pelo *framework* são apresentados neste capítulo.

4.1. Diagrama de Classes

Em Engenharia de Software, diagramas de classes apresentam e caracterizam objetos com seus inter-relacionamentos. Esse tipo de artefato de software de baixo nível de abstração guia as atividades de codificação posteriores. Para o diagrama de classes deste trabalho, foram definidas três interfaces para implementação dos componentes de Extração de Características, Classificação de Padrões e Recuperação de Informação. Interface é um conceito de Orientação a Objetos que define como um componente irá interagir com outros elementos. Qualquer componente que siga uma interface estará apto a interagir com os elementos do *framework*, sem que esse precise ser modificado.

Extrator: Interface que define componentes responsáveis pela extração de características (MatrizCoocorrência e Wavelets).

Classificador: Interface que define componentes responsáveis pela classificação de padrões (KVizinhosPróximos e RedeNeural).

RecuperadorInformação: Interface que define componentes responsáveis pela recuperação de informação.

Estas interfaces foram definidas com a intenção de garantir a modularidade do *framework*, de modo a permitir expansões futuras através do uso de outros extratores, classificadores ou recuperadores de informação que sigam as interfaces propostas. Além das interfaces, foi definida uma classe principal e uma classe de acesso a imagens DICOM, descritas a seguir.

Sistema: Classe central do *framework*, composta por interfaces de Extração de Características, Classificação de Padrões e Recuperação de Informação.

Visualizador: Classe relacionada a métodos de exibição e manipulação de imagens DICOM, suportada pela plataforma ImageJ.

Para ilustrar as classes e seus relacionamentos, foi criado o Diagrama de Classes apresentado na Figura 11.

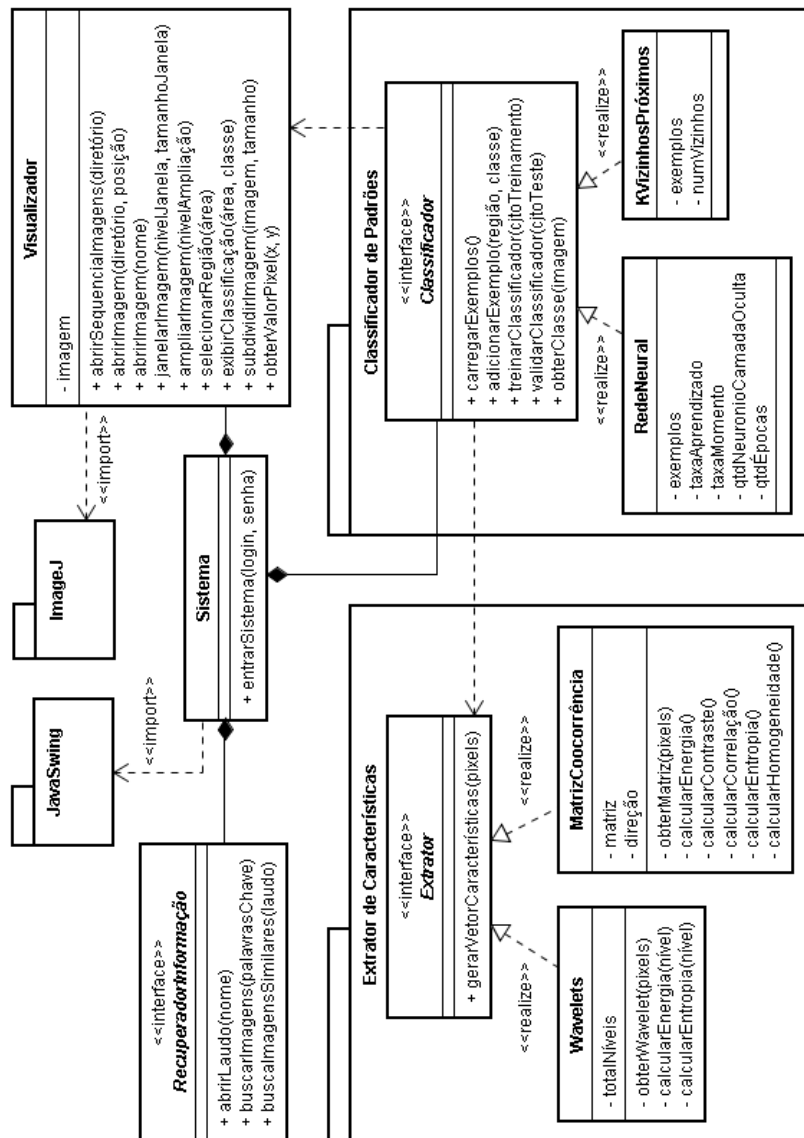


Figura 11. Diagrama de Classes.

4.2. Detalhamento das funcionalidades

Em termos de funcionalidades, as classes do *framework* foram organizadas seguindo o mesmo agrupamento proposto para o levantamento de requisitos no Capítulo 3. Os requisitos funcionais foram traduzidos em métodos, utilizados para realizar chamadas e executar as tarefas do usuário. A seguir, são apresentadas as classes e seus principais métodos.

Sistema

entrarSistema: garante acesso ao sistema mediante login e senha

Extrator

gerarVetorCaracterística: gera um vetor de características a partir de atributos extraídos da imagem

MatrizCoocorrência

obterMatriz: obtém a matriz de coocorrência de um dado conjunto de pixels
calcularEnergia: calcula a energia de uma matriz de coocorrência
calcularContraste: calcula o contraste de uma matriz de coocorrência
calcularCorrelação: calcula a correlação de uma matriz de coocorrência
calcularEntropia: calcula a entropia de uma matriz de coocorrência
calcularHomogeneidade: calcula a homogeneidade de uma matriz de coocorrência

Wavelets

obterWavelet: obtém a transformada de Wavelet de um conjunto de pixels
calcularEnergia: calcula a energia de um nível de uma Wavelet
calcularEntropia: calcula a entropia de um nível de uma Wavelet

Classificador

carregarExemplos: lê os vetores de características dos exemplos armazenados
adicionarExemplo: permite ao radiologista adicionar à base de imagem novos exemplos de padrões radiológicos
treinarClassificador: treina um classificador de padrões com base nos exemplos armazenados
validarClassificador: valida um classificador de padrões com base nos exemplos armazenados
obterClasse: permite ao radiologista obter a classificação uma região de interesse dentro dos padrões radiológicos treinados.

Visualizador

abrirSequenciaImagem: permite ao radiologista abrir uma seqüência de imagens dentro de um diretório específico
abrirImagem: permite ao radiologista abrir uma imagem em formato DICOM
janelarImagem: permite ao radiologista modificar o janelamento de uma imagem
ampliarImagem: permite ao radiologista ampliar parte de uma imagem
selecionarRegião: permite ao radiologista selecionar uma região de interesse da imagem
exibirClassificação: destaca uma região de uma imagem para exibir o resultado de sua classificação
subdividirImagem: subdivide uma imagem em regiões de mesmo tamanho, para serem utilizados nas tarefas de Extração de Características e Classificação de Padrões
obterValorPixel: obtém o valor de pixel na escala DICOM de um ponto da imagem

Recuperador de Informações

abrirLaudo: permite ao radiologista abrir o laudo do exame de um paciente
buscarPalavraChave: realiza uma busca por similaridade na base de imagens utilizando palavras-chave
buscaSimilaridade: realiza uma busca por similaridade na base de imagens utilizando seus laudos

Capítulo 5

EXPERIMENTAÇÃO E VALIDAÇÃO

Este capítulo descreve a construção e a manipulação da base de imagens preparada para experimentação dos componentes do *framework* desenvolvido. Os experimentos foram conduzidos considerando técnicas de recuperação de imagens baseada em conteúdo, recuperação de informação textual e sua união. A validação dos componentes desenvolvidos foi realizada a partir da análise de curvas de precisão-revocação e matrizes de confusão resultantes das tarefas de classificação de padrões.

5.1. Construção da Base de Imagens

Para armazenar os exemplos de padrões radiológicos foi utilizada uma base de imagens. Uma base de imagens funciona de maneira similar a uma base de dados, permitindo o armazenamento, remoção e localização de imagens. Neste projeto, uma imagem é identificada pelo seu cabeçalho DICOM. Além da imagem, armazenam-se também os limites da região que contém o padrão radiológico na forma de coordenadas cartesianas, considerando-se o canto superior esquerdo da imagem como sendo a origem. Desse modo, as coordenadas da região que contém um padrão podem ser corrigidas caso ele sido demarcado incorretamente.

A base de imagens utilizada neste projeto foi estruturada com o apoio do Prof. Dr. Valdair Francisco Múglia, com supervisão do co-orientador e o aluno de mestrado Marcelo Ponciano da Silva. A estruturação foi realizada de acordo com as seguintes etapas:

1. Levantamento de exames. Inicialmente, foram obtidos exames de tomografia de tórax do HCFMRP, dos quais se separaram o laudo e as séries de imagens em formato DICOM. Esta atividade foi realizada com o apoio do mestrando Marcelo Ponciano.

2. Escolha de imagens relevantes. Com a ajuda do Prof. Valdair foram escolhidas as imagens mais relevantes de cada série, considerando-se a presença de padrões radiológicos.

3. Identificação dos padrões. Após a escolha das imagens mais relevantes, foi solicitado ao radiologista que identificasse os padrões presentes em cada imagem utilizado uma ferramenta de seleção de regiões de interesse.

4. Criação da base de imagens. A partir das regiões demarcadas na etapa anterior, foi construída a base de imagens, armazenando o cabeçalho DICOM da imagem referente, os limites da região demarcada e os valores de pixel em seu interior. No total, foram obtidos 125 exemplos de padrões, divididos entre 15 de consolidação, opacidades nodulares e opacidades lineares e 20 de áreas enfisematosas, vidro fosco, favo de mel e o padrão normal.

Foi desenvolvido um sistema computacional que apresenta uma interface gráfica para esta etapa do projeto (ver Figura 12), contendo as funcionalidades necessárias para a criação da base de imagens. Além das funcionalidades previstas inicialmente, o radiologista que estava identificando os padrões solicitou a inclusão de um método para redefinir o janelamento da imagem para seu valor inicial.



Figura 12. Interface gráfica para criação da base de imagens.

5.2. Manipulação da Base de Imagens

A Figura 13 representa um fluxograma que apresenta as tarefas realizadas para a manipulação da base de imagens, descritas a seguir:

- (1) *Coleta da Imagem Médica*: obtém a imagem médica de referência em formato DICOM, preparada pelos procedimentos apresentados na seção 5.1;
- (2) *Extração de Atributos*: através de técnicas de processamento de imagens são extraídos os atributos de interesse da imagem;
- (3) *Busca de Imagens por Similaridade*: utilizando mecanismos de classificação, realiza uma busca por similaridade no banco de imagens;
- (4) *Coleta do Laudo Radiológico*: obtém os laudos radiológicos dos exames retornados pela busca por similaridade;
- (5) *Remoção de Stop-Words*: remoção de palavras sem significado semântico;
- (6) *Stemming*: redução de todas as palavras a seus radicais, facilitando a comparação entre termos;

- (7) *Extração de Palavras-Chave*: através do uso de uma lista de palavras-chave para cada padrão, identifica-se os termos mais relevantes do laudo;
- (8) *Associação Textual*: através das palavras-chave é feita a associação entre a imagem e seu laudo radiológico, adicionando esta informação ao vetor de características da imagem retornada na tarefa (3).

Os procedimentos de (4) a (8) são ser repetidos para cada imagem retornada na tarefa (3), formando a base para o uso de técnicas de IR no apoio a técnicas de CBIR. Os procedimentos (2) e (3) serão detalhados nas próximas subseções.

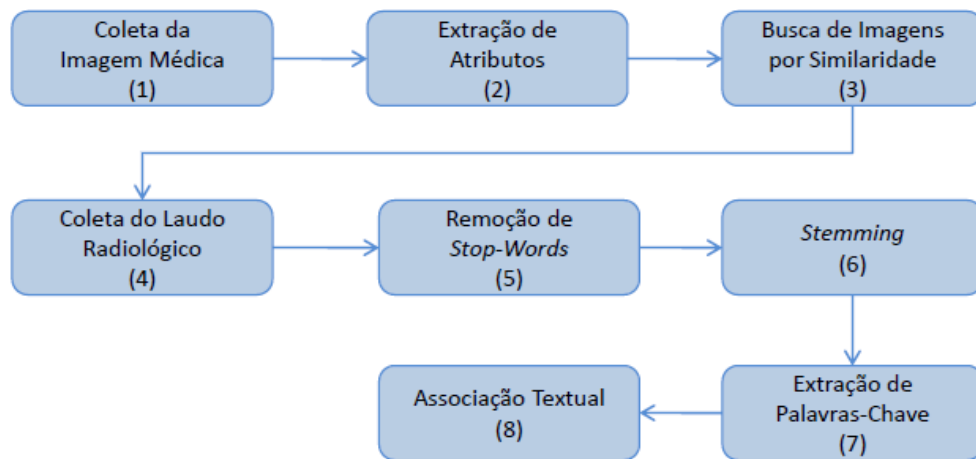


Figura 13. Fluxograma para a manipulação da base de imagens.

5.2.1. Extração de Características

Diferentemente dos sistemas tradicionais de quantificação e classificação de atributos em imagens, o foco da extração de características deste trabalho não é a imagem inteira, e sim sua divisão em pequenas regiões de interesse de mesmo tamanho, constituindo um gradeado sobre a imagem. Neste projeto, a altura e comprimento do elemento unitário do gradeado foram fixados em 20 pixels, como demonstrado na Figura 14.

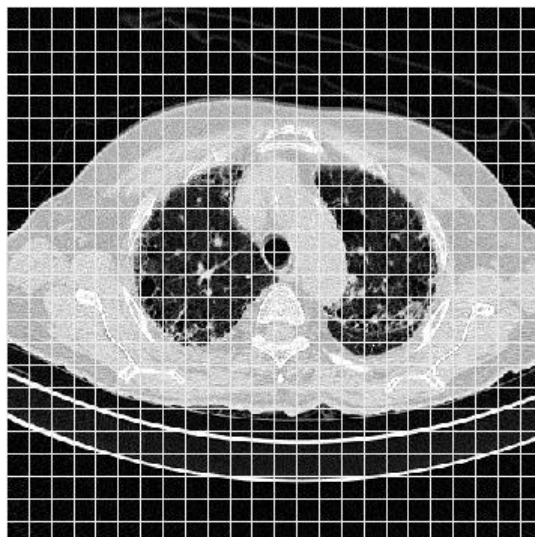


Figura 14. Exemplo de gradeado sobre uma imagem.

Para apoiar a extração de características, foram desenvolvidos dois extratores. O primeiro é apoiado por matrizes de coocorrência e o segundo por transformadas de Wavelets (DWT) de Daubechies 4-tap [Daubechies-1992]. Através da combinação por concatenação de vetores resultantes dos extratores é gerado um vetor de características contendo os cinco principais descritores de Haralick obtidos através da matriz de coocorrência e a energia e entropia de cada nível obtido pela DWT de Daubechies.

Devido ao pequeno tamanho do elemento constituinte do gradeado (20 pixels) dado a janela considerada (4-tap), foi utilizado apenas o primeiro nível da DWT de Daubechies para evitar a subdivisão excessiva da região a ser processada.

5.2.2. Classificação de Padrões

A Classificação de Padrões presentes em uma imagem deve ser feita sobre o vetor de características de cada elemento de seu gradeado, para tornar possível a quantificação e localização dos padrões a serem analisados. Foram considerados dois algoritmos de classificação: K-Vizinhos Próximos (KNN) e Rede Neural Multicamadas (MLP).

A estratégia utilizada para o treinamento e teste dos classificadores foi a Cross-Validação, onde o conjunto de exemplos é subdividido em n *folds* de mesmo tamanho, que constituem um subconjunto dos exemplos. Um dos *folds* é retirado do conjunto para ser posteriormente incorporado na etapa de teste, enquanto o algoritmo é treinado com os *folds* restantes. A escolha dos *folds* deve ser aleatória, para eliminar o viés da ordem que os exemplos são apresentados ao algoritmo. Repete-se o treinamento até que cada *fold* tenha sido utilizado na etapa de teste, armazenando-se os resultados de cada execução. Ao final do processamento de todos os *folds*, os resultados intermediários são combinados e uma matriz de confusão é criada.

5.3. Experimentação da Base de Imagem considerando Recuperação de Imagens baseada em Conteúdo

Para explorar a base de imagens, foi definido um componente de CBIR no *framework* proposto neste trabalho, sendo necessária a extração de características das imagens e a aplicação dos algoritmos de classificação. Conforme apresentado no fluxograma da Figura 13, a classificação de padrões utiliza a saída de dados da extração de atributos para realizar a busca por similaridade na base de imagens.

Embora tenha sido definido inicialmente que a fase de Extração de Características (Seção 5.2.1) seria composta pela combinação dos algoritmos de matriz de coocorrência e Wavelets de Daubechies (DWT), não havia sido realizada nenhuma comprovação prática de que esta era a melhor escolha para o projeto. Neste sentido, foi realizado um estudo da

precisão de classificação em três possíveis alternativas de extração de atributos: (1) matriz de coocorrência; (2) Wavelets de Daubechies; (3) Combinação de matriz de coocorrência e DWT através da concatenação dos vetores de características. Ao combinar mais de um algoritmo a quantidade de variáveis utilizadas para a tarefa de classificação aumenta, podendo ou não levar a um aumento da precisão.

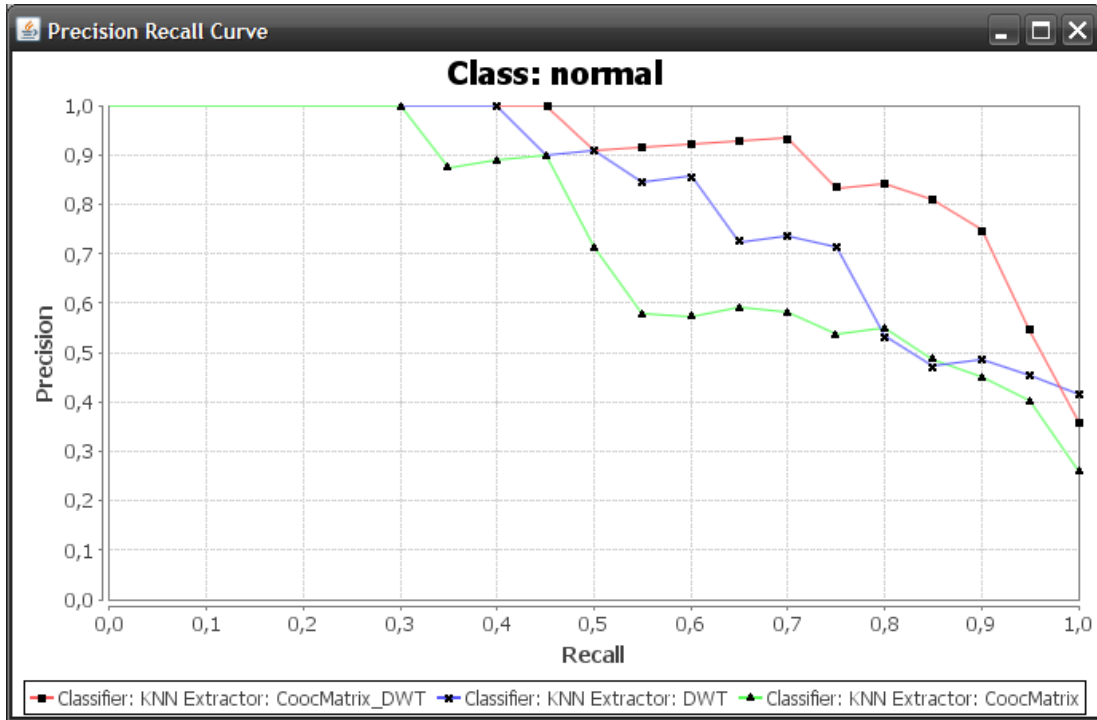


Figura 15. Curvas de Precisão e Revocação comparando Extratores de Características.

Conforme apresentado na Figura 15, a combinação dos dois algoritmos resultou em uma curva de precisão-revocação superior em relação aos cenários nos quais os algoritmos estavam separados, considerando-se o padrão Normal e o algoritmo de classificação KNN com vizinhança de tamanho 3. Embora tenha um custo computacional maior, o cálculo da matriz de Coocorrência e da DWT de uma imagem se mostrou mais apropriado para a classificação de padrões.

A avaliação dos algoritmos de classificação foi feita utilizando matrizes de confusão e medidas de precisão e revocação. Foram testados dois algoritmos, KNN e MLP (*Multi-Layer Perceptron*). Em ambos os casos foi utilizada a técnica de *cross-validation* como abordagem para a execução dos testes, utilizando 10 *folds*.

Classifier: KNN Extractor: CoocMatrix_DWT							
	consolidati...	emphysem...	ground-gla...	honeycom...	micronodul...	normal	reticular-lin...
consolidati...	5	0	7	0	0	0	3
emphysem...	0	15	0	0	1	3	1
ground-gla...	6	0	8	0	2	3	1
honeycom...	0	0	1	18	1	0	0
micronodul...	1	2	2	2	8	0	0
normal	0	4	0	1	0	15	0
reticular-lin...	1	0	3	2	0	1	8
	Precision	Recall					
consolidati...	0.38	0.33					
emphysem...	0.71	0.75					
ground-gla...	0.38	0.4					
honeycom...	0.78	0.9					
micronodul...	0.66	0.53					
normal	0.68	0.75					
reticular-lin...	0.61	0.53					
Classification Error: 38.4%							

Figura 16. Matriz de confusão para o Classificador KNN e a combinação dos Extratores.

A Figura 16 resume os resultados obtidos pelo algoritmo KNN. O erro de classificação foi de 38,4%, utilizando uma vizinhança de tamanho 3. A análise das medidas de precisão e revocação revelou que os erros estavam concentrados nos padrões de consolidação e vidro fosco, apresentando falsos negativos e falsos positivos mútuos, ou seja, houve uma dificuldade maior em classificar um exemplo entre estas duas classes.

O segundo algoritmo testado foi o MLP, com um taxa de aprendizado de 5% e 10 mil épocas de treinamento.

Classifier: MLP Extractor: CoocMatrix_DWT							
	consolidati...	emphysem...	ground-gla...	honeycom...	micronodul...	normal	reticular-lin...
consolidati...	6	0	6	0	0	0	3
emphysem...	0	17	1	0	0	1	1
ground-gla...	6	0	10	1	0	2	1
honeycom...	0	0	0	18	0	0	2
micronodul...	1	2	3	1	8	0	0
normal	0	2	1	0	0	16	1
reticular-lin...	2	0	1	0	1	1	10
	Precision	Recall					
consolidati...	0.4	0.4					
emphysem...	0.8	0.85					
ground-gla...	0.45	0.5					
honeycom...	0.9	0.9					
micronodul...	0.88	0.53					
normal	0.8	0.8					
reticular-lin...	0.55	0.66					
Classification Error: 32.0%							

Figura 17. Matriz de confusão para o Classificador MLP e a combinação dos Extratores

Conforme apresentado na Figura 17, o algoritmo MLP obteve um erro de classificação de 32%, novamente concentrando os erros nos padrões de consolidação e vidro fosco, de maneira semelhante ao algoritmo KNN.

Embora não seja possível determinar como os atributos são utilizados internamente pelo classificador, imagina-se que a densidade, identificada pelo brilho da imagem, tenha sido usada como ponto de corte na classificação entre consolidação, vidro fosco e o resto dos padrões, de tal maneira que os dois primeiros foram agrupados por apresentarem um aumento homogêneo de densidade ao longo do tecido. Com o objetivo de verificar a precisão do algoritmo sem a interferência destes padrões, foi realizado um outro teste com o algoritmo MLP sem estes padrões, mantendo-se os mesmos parâmetros.

	emphysem...	honeycom...	micronodul...	normal	reticular-lin...
emphysem...	16	0	2	2	0
honeycom...	0	18	2	0	0
micronodul...	2	2	9	1	1
normal	2	0	0	17	1
reticular-lin...	0	2	2	1	10
	Precision	Recall			
emphysem...	0.8	0.8			
honeycom...	0.81	0.9			
micronodul...	0.6	0.6			
normal	0.80	0.85			
reticular-lin...	0.83	0.66			
Classification Error: 22.22%					

Figura 18. Matriz de confusão para o Classificador MLP e a combinação dos Extratores, utilizando um conjunto menor de padrões

A Figura 18 apresenta os resultados da classificação considerando um conjunto menor de padrões. O erro de classificação diminuiu aproximadamente em 10%, comprovando que uma parte significativa do erro era devido à dificuldade de classificação entre os padrões de consolidação e vidro fosco.

Uma possível solução seria utilizar outros atributos durante a extração de características da imagem, de maneira a prover diferentes fontes de informação para classificar os padrões, podendo inclusive aumentar o vetor de características e, através da seleção de atributos, escolher os mais relevantes para um determinado conjunto de padrões.

5.4. Experimentação da Base de Imagens considerando Recuperação de Informação Textual

O desenvolvimento do componente de Recuperação de Informação do *framework* foi baseado em uma implementação *opensource* em Java do algoritmo de Porter para a língua portuguesa, denominada PTStemmer, disponível em <http://code.google.com/p/ptstemmer/>. Esta implementação tem a vantagem de apenas conter a funcionalidade de *stemming*, o que permitiu que o módulo fosse desenvolvido de maneira independente, uma vez que ela não exigia o uso de algoritmos específicos para processamento textual.

Inicialmente foi necessária a criação de uma base de documentos relevantes, contendo os laudos e sua classificação em termos representativos de padrões radiológicos. Foram selecionados 10 laudos de cada padrão estudado neste projeto, totalizando uma base de documentos com 60 instâncias.

Utilizando o método TF-IDF para ponderação dos termos e a medida do cosseno para estabelecer o *score* de distância entre os documentos de interesse e a base de laudos, foi realizada uma seqüência de buscas utilizando laudos representativos de cada padrão radiológico. Essa seqüência de buscas foi realizada com as seguintes expressões de consulta, onde o padrão em questão encontra-se sublinhado:

1. Consolidação pulmonar à esquerda compatível com processos inflamatório
2. Bolhas enfisematosas no terço superior esquerdo
3. Opacidade em vidro fosco compatível com fibrose cística
4. Fibrose pulmonar com áreas de favo de mel
5. Velamento micronodular sugestivo de processo inflamatório
6. Velamento reticular de etiologia a esclarecer

A avaliação do desempenho da busca foi feita com base em curvas de Precisão-Revocação e *F-Measure*. A medida *F-Measure*, também conhecida como *F1 score*, permite que a exatidão de um modelo seja determinada pela média harmônica das medidas de precisão e revocação de maneira pontual, a cada novo documento retornado pela busca.

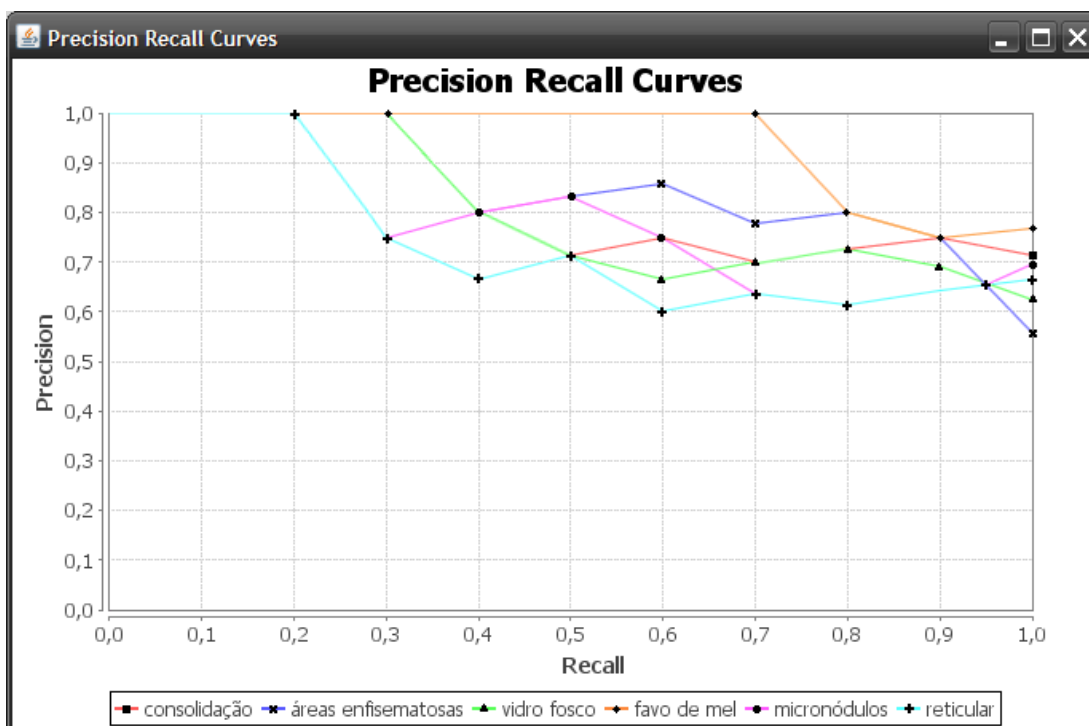


Figura 19. Curvas de Precisão-Revocação comparando o desempenho da busca textual entre os diferentes padrões

Em geral, todos os padrões apresentaram curvas de Precisão-Revocação semelhantes, conforme apresentado na Figura 19, com exceção do padrão favo de mel. Essa constatação sugere que os laudos referentes a este padrão sejam mais relacionados entre si, do ponto de vista textual. No entanto, isto não implica que os outros padrões apresentem descrições inconsistentes, pois o método empregado não analisa os laudos de forma semântica, ou seja, não considera o significado das palavras e seus sinônimos, o que pode produzir falsos negativos.

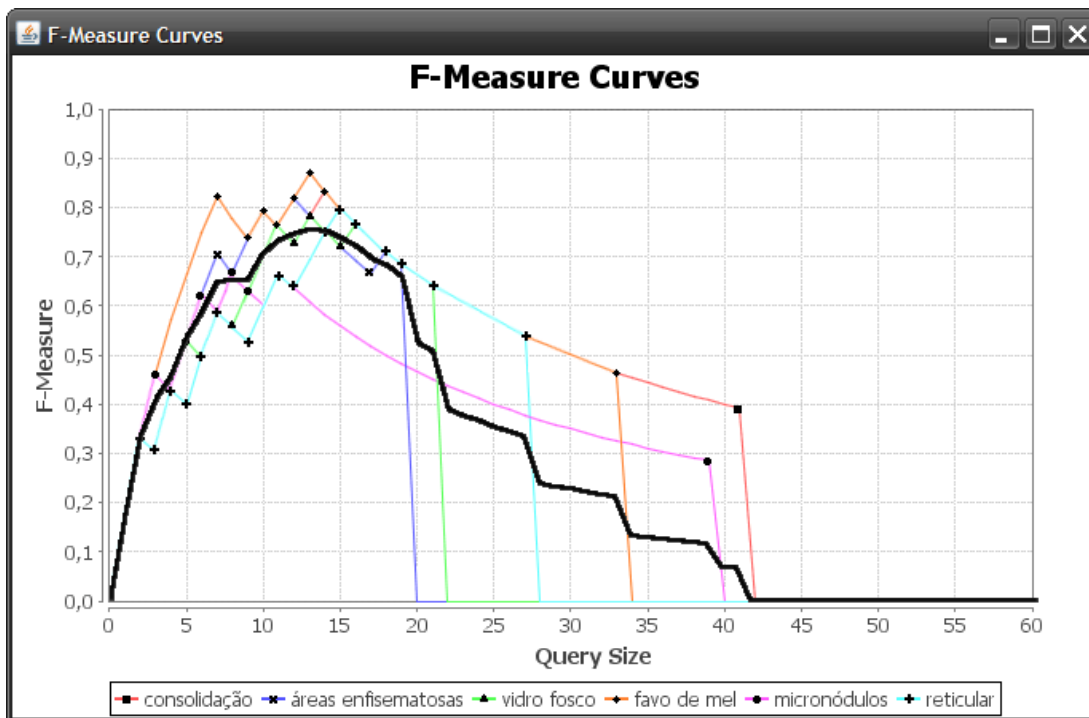


Figura 20. Curvas de *F-Measure* comparando o desempenho da busca textual em conjuntos de busca de tamanhos distintos

Considerando os conjuntos de busca contendo de 1 até o número total de laudos contidos na base de documentos relevantes, calculou-se a curva de *F-Measure* para cada padrão radiológico dentro dos vários conjuntos de busca de tamanhos distintos (ver Figura 20). A curva em destaque foi obtida pela média aritmética ponto a ponto das curvas dos padrões e representa a exatidão média do algoritmo de busca para cada conjunto de busca.

A maior exatidão obtida foi para o conjunto de busca de tamanho 13. Dado que o número de laudos relevantes por padrão radiológico é 10, é necessário um conjunto de busca maior que o número de documentos relevantes para obter resultados significativos, quando o ideal seria obter um conjunto menor. Isto se deve a baixa sensibilidade do método na presença de termos sinônimos em casos onde o mesmo diagnóstico pode ser escrito de maneiras diferentes, dificultando a sua comparação.

5.5. Experimentação da Base de Imagens considerando Recuperação de Imagens Baseada em Conteúdo apoiada por Recuperação de Informação Textual

Embora ainda não esteja claro qual mecanismo seria ideal para unir técnicas de recuperação de imagens e de informação, a partir do estudo realizado neste projeto levantaram-se dois pontos nos quais as duas técnicas se complementariam. O primeiro ponto refere-se a extração de características. A extração de características visuais baseia-se apenas em atributos quantitativos e, no caso de padrões visualmente semelhantes, só seria possível aumentar a taxa de acertos utilizando informações externas como, por exemplo, a opinião diagnóstica do médico. Imagina-se que o conhecimento do médico possa ser utilizado como entrada para o vetor de características, sob a forma de palavras-chave definidas pelo profissional. O segundo ponto trata da busca textual. A busca textual limita-se apenas ao laudo médico e, como ainda não há um padrão definido para a descrição de achados patológicos e terminologia radiológica [Pereira-Silva-2005], não é possível garantir que os laudos sempre seguirão o mesmo padrão necessário para a sua comparação. Imagina-se que o uso de características invariantes à interpretação humana, como a quantificação de atributos de uma imagem radiológica, ajudaria a melhorar a precisão dos algoritmos de categorização da informação.

A abordagem escolhida neste projeto foi o uso de técnicas de recuperação de informação para ampliar o vetor de características utilizado na busca de imagens por similaridade. Propõe-se que inicialmente seja feita uma busca por similaridade utilizando os atributos visuais da imagem e, a partir dos resultados, obter os laudos das imagens mais relevantes. Esses laudos podem ser considerados com uma pequena base de documentos, cujas palavras-chave são o nome dos padrões radiológicos de interesse. Para cada padrão radiológico foi definido um atributo binário no vetor de características da imagem, representado por 0 ou 1, onde 0 significa sua ausência e 1 indica que o padrão é citado no laudo, baseado na presença das palavras-chave. Para melhorar a precisão, assume-se que além dos nomes dos padrões serão também utilizados os sinônimos e termos associados mais frequentes.

As palavras-chave escolhidas para cada padrão radiológico foram:

1. Consolidação: Consolidação
2. Áreas Enfisematosas: Área Enfisematosa, Enfisema
3. Vidro Fosco: Vidro Fosco
4. Favo de mel: Favo de mel, Simplificação Pulmonar
5. Opacidades Nodulares: Nódulos, Micronódulos
6. Opacidades Reticulares e Lineares: Reticular, Espessamento, Linear, Trave

O método escolhido para a classificação foi o KNN com vizinhança de tamanho 3, utilizando vetores de classificação composto por 5 atributos de matriz de coocorrência (representados pelos Descritores de Haralick), 6 atributos de Wavelet de Daubechies (medidas de energia e entropia de 3 janelas) e 7 atributos textuais (representando os 6 padrões radiológicos, mais o normal). Dessa forma, pouco mais de um terço do vetor era composto de características textuais.

Foram realizados dois testes, o primeiro representando os atributos textuais com os valores 0 ou 1 e o segundo com um peso de 25%, ou seja, assumindo os valores 0 ou 0,25. Esta ponderação é importante, pois o uso de valores lógicos 0 e 1 torna a classificação enviesada ao passo que sempre atribuirá a distância máxima para atributos textuais que se diferem, enquanto em imagens isto apenas irá acontecer em situações extremas, dado que qualquer atributo visual é distribuído entre 0 e 1.

A matriz de confusão apresentada na Figura 21 foi obtida pelo método sem ponderação. Observa-se que a taxa de erros foi drasticamente reduzida para 8,8%, com poucos falsos negativos e falsos positivos. O problema de classificação entre os padrões de consolidação e vidro fosco foi eliminado após a adição da informação textual e de modo geral as taxas de precisão e revocação aumentaram consideravelmente.

	consolidati...	emphysem...	ground-gla...	honeycom...	micronodul...	normal	reticular-lin...
consolidati...	15	0	0	0	0	0	0
emphysem...	0	17	0	0	0	3	0
ground-gla...	0	0	19	0	0	1	0
honeycom...	0	0	0	19	0	1	0
micronodul...	0	2	0	0	13	0	0
normal	0	2	1	0	0	17	0
reticular-lin...	0	0	0	1	0	0	14
	Precision	Recall					
consolidati...	1.0	1.0					
emphysem...	0.80	0.85					
ground-gla...	0.95	0.95					
honeycom...	0.95	0.95					
micronodul...	1.0	0.86					
normal	0.77	0.85					
reticular-lin...	1.0	0.93					
Classification Error: 8.79%							

Figura 21. Matriz de Confusão para extração de atributos visuais e textuais, utilizando a representação 0 e 1 para palavras-chave

No entanto, não é desejável que a tarefa de classificação seja muito sensível aos laudos das imagens relevantes. Acredita-se que os principais aspectos negativos se resumem no baixo poder semântico do algoritmo de recuperação de informação utilizado, que considera, entre outras palavras, o “não” como sendo uma *stopword* (palavra a ser removida do documento

antes de seu processamento). Como um exemplo, a expressão “não se podendo descartar consolidação” seria reduzida a “podendo descartar consolidação”, alterando completamente o significado do laudo. Por se tratar de um documento em linguagem natural o laudo é rico em conteúdo semântico, exigindo o uso de algoritmos específicos para processamento de linguagem natural associado a ontologias e *thesaurus* para uniformizar a entrada de dados.

	consolidati...	emphysem...	ground-gla...	honeycom...	micronodul...	normal	reticular-lin...
consolidati...	15	0	0	0	0	0	0
emphysem...	0	16	0	0	1	3	0
ground-gla...	1	0	16	0	0	1	2
honeycom...	0	0	0	19	0	1	0
micronodul...	0	3	0	0	12	0	0
normal	0	2	1	0	0	17	0
reticular-lin...	0	0	0	1	0	0	14
	Precision	Recall					
consolidati...	0.93	1.0					
emphysem...	0.76	0.8					
ground-gla...	0.94	0.8					
honeycom...	0.95	0.95					
micronodul...	0.92	0.8					
normal	0.77	0.85					
reticular-lin...	0.87	0.93					
Classification Error: 12.8%							

Figura 22. Matriz de Confusão para extração de atributos visuais e textuais, utilizando a representação 0 e 0,25 para palavras-chave

Diminuindo o peso dos atributos textuais para 25%, obteve-se uma taxa de erro de 12,8%. As principais diferenças entre a matriz de confusão, apresentada na Figura 2, e as obtidas sem o uso de atributos textuais (ver Figuras 16 e 17) é a melhora na classificação entre consolidação e vidro fosco e um aumento significativo da precisão e revocação dos padrões de micronódulos e opacidades reticulares. Desta forma, o compromisso entre a classificação e o uso de laudo é reduzido, sem deixar de lado a contribuição do método de recuperação de informação em situações de alta similaridade visual, como ocorria entre a consolidação e vidro fosco.

Capítulo 6

CONCLUSÕES

A aplicação do *framework* em um sistema de classificação de padrões sobre uma base de imagens demonstrou que os componentes desenvolvidos podem ser integrados em uma única aplicação, sob a condição de seguirem as interfaces propostas. Devido à modularidade provida pelo *framework*, novos algoritmos de extração de atributos, classificação de padrões e recuperação de informação textual futuramente podem ser agregados ao sistema, sem a necessidade de sua remodelagem. A reorganização dos componentes pode dar origem a novas aplicações. O domínio da aplicação também pode ser alterado, bastando substituir a base de imagens. Para finalizar a definição do *framework*, os componentes do diagrama de classes deverão ser revistos considerando conceitos de Padrões de Projeto (*Design Patterns*).

O uso de técnicas de Recuperação de Informação mostrou-se bastante promissor, auxiliando a classificação de padrões radiológicos principalmente em situações onde os atributos visuais não são capazes de classificar com exatidão devido à alta similaridade visual. No entanto, ainda há diversos problemas a serem resolvidos. Em uma situação real, os laudos nem sempre seguem o mesmo padrão e, por serem documentos ricos em linguagem natural, exigem que o processamento textual seja feito em nível semântico. A extração de palavras-chave não é o método mais indicado para este caso, pois ainda que as palavras indiquem a presença de um padrão radiológico, a semântica do laudo pode dizer o contrário ou trazer um grau de incerteza. O próprio vocabulário radiológico pode variar entre os profissionais, sendo necessário o uso de *thesaurus* e ontologias para uniformizar a linguagem.

Outro ponto importante a ser considerado é a etapa de localização dos laudos. Neste projeto, os laudos de cada imagem tiveram que ser localizados manualmente, pois não há qualquer tipo de chave única que faz a ligação entre os laudos e os exames de imagem armazenados em mídia. Esta tarefa foi realizada usando o número do paciente e a data do exame, ambos presentes no cabeçalho DICOM. Esta não é a forma ideal de localização, pois não há integração efetiva entre os sistemas de informação da instituição com o seu centro de imagens. Recomenda-se o uso de um *worklist*, que não somente irá facilitar a localização do laudo de uma imagem, mas possibilitar a busca de exames anteriores e todas as informações do paciente, que também fazem parte do diagnóstico. Dessa forma, dados como o histórico do paciente poderiam ser incorporados ao sistema de auxílio ao diagnóstico, utilizando técnicas semelhantes de recuperação de informação

Além de resolver os problemas citados, pretende-se como trabalho futuro a aplicação do *framework* desenvolvido em sistemas de apoio ao diagnóstico. Nesse caso, o uso da extração de palavras-chave poderá auxiliar a tarefa de sugestão de diagnósticos, retornando para o radiologista sugestões baseadas em laudos de imagens similares já diagnosticadas por outros radiologistas, aumentando a robustez e a confiabilidade do sistema proporcionalmente à experiência do grupo de radiologistas. Para radiologistas com menos experiência ou residentes, um sistema resultante deste projeto poderá prover um ambiente de aprendizado eletrônico, fornecendo informações relevantes para o diagnóstico através, por exemplo, da extração de palavras-chave dos laudos de imagens similares a aquela sendo estudada.

Outro cenário considerado para a aplicação do *framework* desenvolvido seria em sistemas de auditoria de diagnóstico. Nesse caso, o sistema auxiliaria na observação do conteúdo escrito pelo radiologista para verificar se este segue o padrão semântico e sintático esperado para aquele tipo de imagem. Acredita-se poder realizar esse tipo de auditoria através da comparação com laudos já escritos para imagens similares. Uma variação poderia ser voltada para sistemas de vigilância em saúde, que alertaria radiologistas sobre possíveis erros de diagnóstico e problemas em massa apresentados pela população através do processamento das informações de laudos escritos por diferentes radiologistas de um mesmo grupo.

Capítulo 7

REFERÊNCIAS

[Azevedo-Marques-2006]

AZEVEDO-MARQUES, P. M.; TAHOSES, P. G.; IGLESIAS, I. et al. Automated Characterization of Interstitial Lung Disease on Postero-anterior Chest Radiographs. *International Journal of Computer Assisted Radiology and Surgery*, v. 1, p. 350-352, 2006.

[Azevedo-Marques-2007]

AZEVEDO-MARQUES, P. M.; AMBRÓSIO, P. E.; PEREIRA JR, R. R. et al. Characterization of interstitial lung disease in chest radiographs using SOM artificial neural network. *International Journal of Computer Assisted Radiology and Surgery*, v. 2, p. 368-370, 2007.

[Baeza-Yates-1999]

BAEZA-YATES, R.; RIBEIRO-NETO, B. Modern Information Retrieval. Addison Wesley, 1999, New York, NY.

[Connors-1980]

CONNERS, R. W.; HARLOW C. A. A theoretical comparison of texture algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, v. 2, n. 3, p. 204-222, 1980.

[Daubechies-1992]

DAUBECHIES I. Ten Lectures on Wavelets. *Society for Industrial and Applied Mathematics*, 1992.

[Depeursinge-2007]

DEPEURSINGE, A.; MÜLLER, H.; HIDKI, A. et al. Image-based diagnostic aid for interstitial lung disease with secondary data integration. *SPIE Medical Imaging*, v. 6512, 2007.

[Doi-2005]

DOI, K. Current status and future potential of computer-aided diagnosis in medical imaging. *British Journal of Radiology*, v. 78, p. S3-S19, 2005.

[Haralick-1973]

HARALICK, R. M.; SHANMUGAN, K.; DINSTEN, I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, v. 3, n. 6, p. 610-621, 1973.

[Haykin-2001]

HAYKIN, S. Redes Neurais: Princípios e Práticas, 2a. edição. Ed. Bookman, 2001.

[Ishida-1998]

ISHIDA, T.; KATSURAGAWA, S.; ASHIZAWA, K. et al. Application of artificial neural networks for quantitative analysis of image data in chest radiographs for detection of interstitial lung disease. *Journal of Digital Imaging*, v. 11, n. 4, p. 182-192, 1998.

[Johnson-1997]

JOHNSON, R. E. Frameworks = (Components + Patterns). *Communications of the ACM*, v. 40, n. 10, p. 39-42, 1997.

[Mahmmoud-2007]

MAHMMOUD M. I.; DESSOUKY, M. I. M.; DEYAB, S. et al. Comparison Between HAAR And Daubechies Wavelet Transformions on FPGA Technology, *Proceedings of World Academy of Science, Engineering and Technology*, v. 20, p. 23-29, 2007.

[Martinez-1998]

MARTINEZ, J. A. B. Doenças Intersticiais Pulmonares. *Medicina, Ribeirão Preto*, v.31, p. 247-256, 1998.

[Martins-2005]

MARTINS, E. R. S.; AZEVEDO-MARQUES, P. M.; OLIVEIRA, L. F. et al. Caracterização de lesões intersticiais de pulmão em radiograma de tórax utilizando análise local de textura. *Radiologia Brasileira*, v. 38, n. 6, p. 421-426, 2005.

[Mathielson-1989]

MATHIESON, J. R.; MAYO, J. R.; STAPLES, C. A. et al. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography. *Radiology*, v. 171, p. 111-116, 1989.

[Mayo-1991]

MAYO, J. R. High resolution computed tomography: Technical aspects. *Radiol Clin North Am.*, v. 29, n. 5, p. 1043-1049, 1991.

[Meyers-1964]

MEYERS, P. H.; NICE, C. M.; BECKER, H. C. et al. Automated computer analysis of radiographic images. *Radiology*, v. 83, p. 1029-1033, 1964.

[Mildenberger-2002]

MILDENBERGER, P.; EICHELBERG, M.; MARTIN, E. Introduction to the DICOM standard. *European Radiology*, v. 12, n. 4, p. 920-927, 2002.

[Müller-2001]

MÜLLER, N. L.; FRASER, R. S.; COLMAN, N. C. et al. Radiologic Diagnosis of Diseases of the Chest. W.B. Saunders Company, 2001.

[Pereira-Silva-2005]

PEREIRA-SILVA, J. L.; KAVAKAMA, J.; TERRA FILHO, M. et al. Consenso brasileiro sobre a terminologia dos descritores de tomografia computadorizada do tórax. *J Bras Pneumol*, v. 31, n. 2, p. 149-156, 2005.

[Pressman-2006]

PRESSMAN, R. S. Engenharia de Software, McGraw-Hill, 2006.

[Sluimer-2003]

SLUIMER, I. C.; van WAES, P. F.; VIERGEVER, M. A. et al. Computer-aided diagnosis in high resolution CT of the lungs. *Medical Physics*, v. 30, n. 12, p. 3081-3090, 2003.

[Souza-2002]

SOUZA JR, A. S.; ARAUJO NETO, C.; JASINOVODOLINSKY, D. et al. Terminologia para a descrição de tomografia computadorizada do tórax: Sugestões iniciais para um consenso brasileiro. *Radiologia Brasileira*, v. 35, n. 2, p. 125-128, 2002.

[Uchiyama-2003]

UCHIYAMA, Y.; KATSURAGAWA, S.; ABE, H. et al. Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. *Medical Physics*, v. 30, n. 9, p. 2440-2454, 2003.

[Viera-2007]

VIEIRA, A. F. G.; Virgil, J. Uma revisão dos algoritmos de radicalização em língua portuguesa. *Information Research*, v. 12, n. 3, 2007.

[Wang-1997]

WANG, J. Z.; WIEDERHOLD, G.; FIRSCHEIN, O. et al. Content-based image indexing and searching using Daubechies wavelets. *Int. J. Digit. Library*, v. 1, p. 311-328, 1997.

[Winsberg-1967]

WINSBERG, F.; ELKIN, M.; MACY, J. et al. Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. *Radiology*, v. 89, p. 211-215, 1967.

[Yang-1999]

YANG, Y.; LIU, X. A re-examination of text categorization methods. *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, p. 42-49, 1999.