

Classificador de imagens de pulmão utilizando wavelets de Haar e distância de Mahalanobis

Rafaelo P. da Rosa ¹

Marcelo A. Porto ¹

Alice Kozakevicius ¹

Data de submissão: 21.03.2015

Data de aceitação: 01.11.2015

Resumo: Um método de extração de características e reconhecimento de padrões de imagens de tomografia de pulmão é proposto a partir da aplicação da transformada wavelet de Haar 2D. Como medida de similaridade, a distância de Mahalanobis é utilizada por considerar a correlação entre os dados, o que se torna relevante, uma vez que a distinção visual entre as imagens é pequena. O algoritmo foi testado utilizando-se diferentes vetores de características, bem como diferentes agrupamentos de imagens. Foi obtida uma taxa de acerto superior a 85% ao se considerar duas classes.

Abstract: A method to extract features and to recognize patterns from lung images obtained by computerized tomography is proposed by applying the 2D Haar wavelet transform. To measure the similarity between images, the Mahalanobis distance was used because it considers the data correlation, which exists since the visual distinction among the image groups is generally small. The algorithm was tested with several feature vectors and with different classes of images. The proposed system was able to classify the images with an accuracy higher than 85% considering only two classes.

1 Introdução

O crescente número de imagens digitais geradas diariamente em clínicas e hospitais possibilita que profissionais da saúde utilizem esses recursos em sistemas para gerar diagnósticos mais precisos e rápidos, processo conhecido como *computer-aided diagnosis (CAD)*. Entretanto, o volume de dados disponíveis requer que estes sejam disponibilizados de modo a serem úteis para o usuário, dependendo das suas necessidades de aplicação. Os sistemas de *content-based image retrieval (CBIR)*, ou busca de imagens por conteúdo, são uma alternativa

¹Programa de Pós-graduação em Informática, UFSM, Av. Roraima, 1000, CEP: 97105-900, CT, Santa Maria - RS
{rafaelo.pinheiro1988@gmail.com}
{porto@inf.ufsm.br}
{alicek@ufsm.br}

que possibilita o usuário obter imagens de um determinado conjunto que sejam semelhantes a uma imagem fornecida, sem a necessidade de se utilizar informações textuais [18]. O CBIR baseia-se no processo de extração de características relevantes que podem ser utilizadas para determinar semelhança entre imagens e classificá-las. Considerando que a relevância de uma característica é sensível ao contexto a ser analisado, existe a dificuldade em se desenvolver sistemas genéricos de CBIR, isto é, que possam ser aplicáveis a qualquer conjunto de imagens. Entre os contextos mais específicos abordados em problemas de CBIR há o das imagens diagnósticas [1], que justifica sua importância pela necessidade de se recuperar imagens em bases de dados extensas a partir de uma determinada imagem de busca (*query image*), seja para comparar peculiaridades em comum entre elas, como por exemplo um determinado tipo de nódulo, ou exames de um mesmo paciente que se referem a uma mesma estrutura corporal para se gerar um histórico. A grande dificuldade em se construir um sistema eficiente de CBIR reside no conceito de semelhança estar associado à percepção humana, um critério subjetivo que muitas vezes pode não corresponder ao que foi implementado num algoritmo para automatizar o processo através do uso de uma máquina [8].

Desta forma, sistemas de CBIR são geralmente orientados a problemas específicos, dada esta dificuldade de se construir um sistema genérico capaz de gerenciar todo e qualquer tipo de imagem. Isso porque, o mesmo critério de semelhança pode não ser aplicável ou relevante para todos os tipos de imagem. No caso de imagens diagnósticas, por exemplo, não há sentido em considerar um critério como cor, pois as imagens em geral são monocromáticas. Uma alternativa para se determinar critérios de similaridade em imagens são as funções wavelets e suas transformadas, abordadas em trabalhos da área como em [12], no qual foi utilizada uma versão modificada da wavelet de Haar na extração de características. Após esta etapa os dados foram agrupados em clusters através do algoritmo K-Means no contexto de imagens de textura. Em [14], o método de CBIR foi construído a partir do conceito de energia, utilizando-se três wavelets diferentes para provar sua eficiência.

O uso de wavelets para extração de características é uma metodologia frequentemente mencionada na literatura recente, tanto no contexto das imagens biomédicas [3], quanto nos casos de classificação de texturas [16]. Determinadas famílias de wavelets apresentam vantagens sobre outras dependendo do problema abordado. As wavelets de Gabor, por exemplo, possuem propriedades que facilitam a identificação e extração de texturas [13], enquanto que wavelets de Haar possuem a vantagem de ter um menor custo computacional [20] e dispensarem qualquer tipo de extrapolação para o tratamento dos valores de fronteira [9].

O objetivo principal deste trabalho é dar continuidade ao estudo do potencial das transformadas wavelets no processo de extração de características de imagens biomédicas. Através dos dados inicialmente transformados pela transformada de Haar bidimensional, pode-se estabelecer uma medida de similaridade entre imagens de tomografia de pulmão. Aqui a distância de Mahalanobis [4] é utilizada para se obter esta medida e desempenha um papel

importante, pois permite que a correlação entre os dados seja levada em consideração. A partir do momento em que uma imagem de busca (query image) é comparada com imagens de um banco de dados, é possível gerar um resultado contendo as imagens do banco mais semelhantes em relação àquela que foi fornecida, de acordo com o critério de semelhança estipulado. Para dar continuidade à discussão, este trabalho está organizado da seguinte maneira: na Seção 2 são apresentados alguns trabalhos relacionados; na Seção 3, conceitos básicos sobre wavelets de Haar são abordados e na Seção 4 são apresentadas noções e conceitos sobre a distância de Mahalanobis, ingrediente fundamental na determinação do critério de semelhança a ser explorado na metodologia aqui apresentada. Finalmente nas Seções 5 e 6 são descritos o algoritmo proposto, como ele foi utilizado e os resultados obtidos.

2 Trabalhos relacionados

Recentemente, diversos autores propuseram soluções para o problema de extração de características utilizando wavelets em diferentes contextos. Dua et. al [5] apresentaram um método baseado em diferentes filtros de transformada wavelet para classificar imagens de glaucoma através da energia de sub-bandas da transformada. Shurong et. al [10] utilizaram uma estrutura piramidal construída a partir de três diferentes tipos de transformada wavelet para extração de textura em imagens diagnósticas. Sarker e Iqbal [15] propuseram uma abordagem híbrida para sistemas de CBIR que consiste na utilização de wavelets de Haar para extração de textura e dos chamados *color moments* para extração de cores.

É possível observar que a maioria dos trabalhos citados, cuja extração de características faz uso de alguma transformada wavelet, utiliza a distância euclidiana como métrica de comparação entre os vetores de características. Contudo, dependendo da situação essa distância pode não apresentar o melhor resultado, principalmente se houver forte correlação entre os valores obtidos. Nesse caso, alternativas como a distância de Mahalanobis [4], que será investigada no contexto de aplicações de sistemas CBIR no presente trabalho, podem ser mais viáveis para o problema.

3 Wavelets de Haar

As transformadas wavelets possuem uma variedade de aplicações no campo das imagens, mais especificamente na edição, compressão e busca por conteúdo de acordo com Stollnitz et. al [20]. Neste contexto, as transformadas discretas produzem representações de imagens em diferentes escalas, sendo que em cada nível da transformada, um bloco de coeficientes de escala associado a uma imagem grosseira é gerado, além de três outros blocos de coeficientes wavelets (ou de detalhe), referentes à variação de informação nas direções horizontal, vertical e diagonal. Quanto maior for o número de níveis de decomposição, maior

será o número de blocos de detalhes complementares para recuperar a imagem original a partir do bloco associado à imagem mais grosseira, definida no último nível de decomposição. As wavelets ortonormais possuem a vantagem de possibilitar uma representação mais compacta, dado o fato de que a redundância estatística entre pixels é eliminada [21]. Entre as diversas famílias de wavelets ortonormais existentes, a mais simples em termos de algoritmo e de representação de dados é a de Haar. Isso porque as funções-base dessas wavelets são representadas por filtros com o menor comprimento possível, $h_k = [1/\sqrt{2}, 1/\sqrt{2}]$ e $g_k = [1/\sqrt{2}, -1/\sqrt{2}]$. As wavelets de Haar são descritas a partir de dilatações e translações da função escala $\phi(x)$ e wavelet $\psi(x)$, sendo a transformada unidimensional definida através da convolução da função analisada pelos filtros h_k e g_k . Na formulação decimada da transformada wavelet discreta (TWD), os coeficientes de escala são reduzidos pela metade a cada nível de decomposição.

A simplicidade das wavelets de Haar traz algumas vantagens como o fato de possuírem um baixo custo computacional, não precisarem de valores extrapolados para o cálculo da transformada nas vizinhanças das bordas da imagem [9] e serem eficientes na detecção de variações abruptas no sinal ou imagem analisados. O cálculo da TWD decimada de Haar para um vetor de tamanho n é descrito a partir dos seguintes passos [12]: (1) Calcula-se a média dos valores dois a dois disjuntos, ou seja, haverá um total de $n/2$ médias; (2) Calcula-se a diferença entre a média e cada um dos valores utilizados para formá-la. Haverá então um total de $n/2$ diferenças; (3) A primeira metade do vetor é preenchida com as médias; (4) A segunda metade do vetor é preenchida com as diferenças; (5) O processo é repetido de maneira recursiva sempre na primeira metade do vetor até que a mesma seja unitária.

Uma limitação das TWDs é que esse procedimento parte do princípio de que o tamanho n da amostra inicial é uma potência de 2. Em algumas situações, como no caso deste trabalho, essa não é uma limitação real, pois é fácil de se obter imagens com dimensão sendo potência de 2. No entanto, no caso da wavelet de Haar, esta limitação pode ainda ser evitada, uma vez que com apenas poucos valores extrapolados, a transformada pode ser facilmente calculada para um quantidade par de pontos qualquer. Outro aspecto a ser observado é a questão de se utilizar os filtros h_k e g_k na forma normalizada, o que faz com que as bases ortogonais para a representação dos dados sejam, na verdade, ortonormais, o que traz facilidades em processos como o da compressão [20].

Existem duas abordagens principais para o processo de decomposição para a TWD-2D: standard e non-standard [20]. Neste trabalho, o algoritmo na forma non-standard é considerado, que consiste em aplicar um nível da transformada em todas as linhas e depois em todas as colunas da imagem. Para os demais níveis, segue-se o mesmo procedimento, até o nível mais grosseiro. O algoritmo non-standard [20] é descrito pelo painel (a) da Figura 1.

3.1 Extração de características utilizando Haar

Em aplicações envolvendo imagens, as aplicações do filtro passa-baixa h_k (cálculo das médias) resultam em um "borramento" (*blurring*), i.e., perda de resolução da imagem a cada nível. Em contrapartida, os coeficientes wavelets (diferenças) captam as variações contidas nos contornos. Estes são extraídos a partir da aplicação do filtro passa-alta g_k em cada par de valores consecutivos, nas três direções: horizontal (H), vertical (V) e diagonal (D). Ao se aplicar a transformada 2D em três níveis, haverá um bloco de coeficientes de escala no nível mais grosseiro, denotado por C3, e três subconjuntos formados pelos blocos de detalhes nas três direções H_j, V_j, D_j , sendo $j = 1, 2, 3$ indicando o nível da transformação, Figura 1.

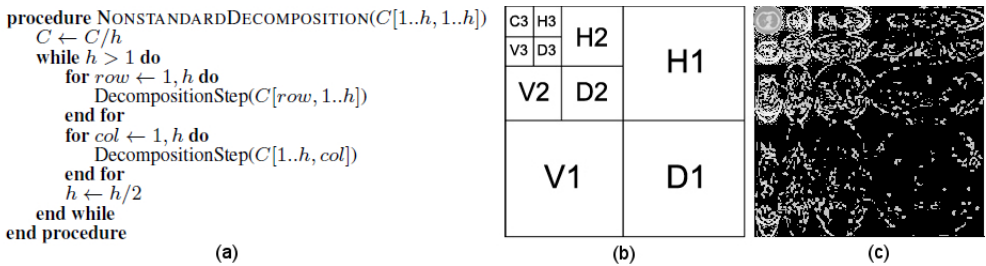


Figura 1: TWD-2D 3 níveis: (a) algoritmo non-standard; (b) estrutura de blocos; (c) TWD-2D de imagem biomédica com coeficientes wavelets com contraste realçado.

É possível extrair medidas estatísticas entre os dados de cada nível e utilizá-las para compor um vetor de características que irá representar a imagem analisada. Entre essas medidas, consideram-se o valor médio, a energia, o desvio padrão, o valor mínimo e o valor máximo, sendo cada uma delas aplicáveis tanto aos blocos de coeficientes de escala quanto aos de coeficientes wavelet.

4 Distância de Mahalanobis

A distância euclidiana é uma medida muito usual, no entanto, pode não ser a mais adequada para se calcular distâncias entre grupos de objetos, especialmente quando estes possuem uma forte correlação entre si. É nesse contexto que surge como alternativa a distância de Mahalanobis, que considera a correlação existente entre os elementos em questão. A distância de Mahalanobis pode ser aplicada a elementos de um espaço n dimensional qualquer, sendo que, com relação à aplicação em foco neste trabalho, cada dimensão corresponde a uma característica diferente, que será então utilizada como um atributo para classificação de padrões e, conseqüentemente, tomada de decisão. Para que o cálculo seja realizado, é neces-

sário que cada grupo possua o mesmo número de características. Porém, não é preciso que todos os grupos tenham o mesmo número de objetos. Esses detalhes serão exemplificados na seção 6 com as variáveis do problema abordado.

A distância de Mahalanobis possui algumas vantagens que se destacam sobre a distância euclidiana. Ela considera que a variância em cada direção é diferente, além de considerar a covariância entre as variáveis [22]. A correlação entre os dados pode ser calculada a partir da matriz de variância-covariância do vetor original, no entanto, este cálculo pode apresentar certos problemas como o fato de muitos dados serem redundantes ou correlacionados, caso haja um conjunto muito grande de variáveis. Esse problema é chamado de multicolinearidade e resulta numa matriz singular de variância-covariância e que não pode ser invertida para que o cálculo da distância seja efetuado. Outra limitação consiste no fato de que o número de objetos no conjunto inicial deve ser maior do que ou igual ao número de variáveis, sendo em alguns casos necessário reduzir o número de características a serem consideradas para satisfazer essa propriedade [11].

Uma das aplicações para a distância de Mahalanobis é o reconhecimento de padrões. A partir de um conjunto de treinamento de indivíduos organizados em classes, é possível construir um modelo matemático baseado nessa distância a fim de determinar a que classe um novo indivíduo pertence, partindo do princípio de que ele pertence a uma e somente uma das classes já existentes. Esse modelo, conhecido como Análise Discriminante Linear (LDA), tem como objetivo principal separar amostras de grupos distintos a partir da maximização da separabilidade entre classes enquanto minimiza a variabilidade dentro de uma classe [7] e é representado na equação 1,

$$cf_K(x_i) = (x_i - \bar{x}_K)C_{pooled}^{-1}(x_i - \bar{x}_K)^T - 2 \ln \pi_K. \quad (1)$$

K é uma determinada classe, C_{pooled}^{-1} é a matriz de variância-covariância de diferentes classes e x_i é um novo indivíduo. $cf_K(x_i)$ é uma pontuação de classificação, de modo que a classe K que apresentar a maior pontuação é aquela à qual o indivíduo x_i tem mais chance de pertencer. O segundo termo da equação, $2 \ln \pi_K$ é a probabilidade que precisa ser considerada no caso de classes terem um número diferente de indivíduos [11]. Essa medida é utilizada como base para classificar as imagens do conjunto analisado.

Para que seja possível a aplicação desta distância no reconhecimento de padrões, é necessário que a matriz combinada (*pooled*) inversa seja criada primeiramente. Para a geração desta matriz, leva-se em consideração todas as classes do conjunto de treinamento, e todos os elementos para cada grupo ao qual pertencem. Todos os elementos para cada classe devem ter o mesmo número de variáveis (mesmo número de colunas), enquanto que o número de elementos (número de linhas) pode variar entre os grupos. É calculada a média para cada uma das variáveis de cada classe e este valor é então subtraído do valor correspondente à variável

para cada um dos elementos, gerando uma matriz \hat{X}_g para cada grupo g .

A seguir, na equação 2, é dada a matriz de covariância C_g para cada um dos grupos indexados por g , sendo n_g o número de imagens correspondente. A operação $(.)$ é o produto entre duas matrizes e \hat{X}_g é a matriz criada para o grupo g . Todas as matrizes de covariância geradas terão o mesmo tamanho, tanto em linhas como em colunas, o qual corresponde ao número de variáveis utilizadas.

$$C_g = \frac{1}{n_g}(\hat{X}_g^T \cdot \hat{X}_g). \quad (2)$$

Assim, através da equação 3, a matriz *pooled* obtida é entendida como uma combinação de todas as matrizes de covariância geradas no passo anterior. Esta matriz é obtida por combinação linear das matrizes de covariância c_i para cada grupo i , sendo as constantes da combinação dadas por pesos médios n_i/n associados ao número de representantes de cada classe (n_i), com (r, c) correspondente ao índice da matriz de covariância *pooled* s . n é o número total de elementos do conjunto. Na última etapa, é feito o cálculo da matriz *pooled* inversa, utilizando-se a biblioteca desenvolvida por [2].

$$s_i(r, c) = \frac{1}{n} \sum_{i=1}^g n_i c_i(r, s). \quad (3)$$

A partir desse resultado, é possível aplicar a distância de Mahalanobis para classificar um elemento como pertencente a uma determinada classe, utilizando a matriz inversa calculada anteriormente. Esta distância é calculada K vezes, sendo K o número de grupos que constituem o conjunto de treinamento. A cada iteração deste cálculo, é gerado o vetor que corresponde à diferença entre os valores das variáveis do elemento analisado com os valores médios de cada uma das variáveis do grupo atual na iteração. Este vetor é, então, aplicado na equação 1. O valor de π_K corresponde ao peso aplicado para o grupo da iteração K .

O resultado gerado para cada uma das iterações (para cada grupo) pode ser entendido como uma pontuação que determinará a qual grupo (classe) o elemento pertence. Assim, o maior valor produzido entre as iterações corresponderá ao grupo no qual o elemento provavelmente pertença. Essa medida foi utilizada como base para classificar as imagens analisadas.

5 Algoritmo

O algoritmo desenvolvido neste trabalho pode ser dividido em duas etapas: uma destinada a realizar o pré-processamento e outra para a classificação. Todo o desenvolvimento

do sistema foi implementado utilizando-se a linguagem de programação Java. Utilizou-se o framework da JAI (Java Advanced Imaging) para o tratamento das imagens. A transformada wavelet de Haar 2D não normalizada, com estrutura piramidal, foi aplicada para a extração de características a partir dos coeficientes obtidos pela aplicação da transformada.

5.1 Pré-processamento

Antes da devida classificação, o sistema necessita passar por um pré-processamento. Nesta fase, cada um dos grupos ou classes são analisados, permitindo que seja aplicada a transformada wavelet de Haar 2D em cada uma das imagens de cada classe, gerando um vetor de características por imagem. As características e o tamanho escolhido para este vetor serão abordados detalhadamente na seção 6, sendo ele composto de valores de energia dos coeficientes de detalhe e da média dos coeficientes de escala do nível mais grosseiro.

A etapa de pré-processamento é destinada para a formação da matriz *pooled* inversa, usando, para isso, todas as imagens pertencentes a cada uma das classes do conjunto de treinamento. Para a geração desta matriz inversa, é necessário, em um primeiro momento, especificar como as matrizes para cada um dos grupos devem ser formadas. Ao invés de utilizar-se variáveis para cada uma das colunas de uma matriz, como relatado na seção 4, utilizam-se características extraídas com a aplicação da transformada wavelet sobre cada uma das imagens, como as energias nas direções vertical, horizontal e diagonal, além da média dos coeficientes de escala. Assim sendo, assumindo que o vetor de características seria formado por quatro valores, cada grupo teria matrizes com quatro colunas. Desta forma, com a especificação das matrizes definida, gera-se a matriz *pooled* inversa a partir dos cálculos apresentados na seção 4. Para cada grupo, é criado um vetor separado, com as médias de cada uma das características escolhidas. Esses vetores criados, um para cada classe, serão utilizados tanto na geração da matriz *pooled* inversa quanto na etapa de classificação.

5.2 Classificação

Nesta etapa, é realizada a classificação do elemento de entrada como pertencente a um determinado grupo. O dado de entrada, neste caso, é uma imagem de tomografia de pulmão, a qual deseja-se saber a qual grupo ela pertence. Esta classificação é feita através da distância de Mahalanobis, baseando-se no valor de pontuação resultante, o que leva à determinação da classe a qual o dado de entrada pertence.

Inicialmente, aplica-se a transformada wavelet de Haar sobre a imagem de entrada, gerando um vetor de características com a mesma dimensão dos vetores gerados na etapa de pré-processamento, ou seja, com as mesmas características. Este vetor é então passado para o sistema que determina a classificação, conforme explicado no final da seção 4.

6 Metodologia

A questão da escolha de quais e quantos atributos serão considerados é extremamente relevante, pois está associada à performance do classificador escolhido. A seguir, serão abordados os experimentos realizados neste trabalho, como a escolha das características que irão compor os vetores de atributos, bem como os resultados encontrados. Inicialmente, será feita uma breve apresentação do conjunto de imagens utilizado. O motivo de utilizar-se imagens de tomografia vai ao encontro com o que foi abordado na seção 3.1, pois os pixels de borda nessas imagens são todos redundantes, permitindo com que a aplicação da transformada de Haar seja a mais adequada para o caso.

6.1 Conjunto de imagens

O banco de imagens utilizado neste trabalho foi o Computed Tomography Emphysema Database [19], adquirido a partir de um estudo feito no Department of Respiratory Medicine, no hospital universitário de Gentofte [17] consistindo de imagens de tomografia computadorizada (CT) de tórax. Foi utilizado o equipamento GE (General Electric) (LightSpeed QX/i; GE Medical Systems, Milwaukee, WI, USA), formado por um detector de 4 linhas. Cada imagem (*slice*), do conjunto de imagens produzidas foi gerada com uma resolução de 0.78 x 0.78 mm, com espessura de 1.25mm, voltagem do tubo de 140kV e corrente de 200mAs. Os slices foram reconstruídos usando um algoritmo de resolução espacial alta [19].

De acordo com [19], para a produção deste banco, 39 sujeitos foram submetidos ao estudo, gerando cerca de 115 imagens de pulmão, divididas nas alturas inferior, média e superior. Destes 39 sujeitos, 9 eram não fumantes, 10 eram fumantes sem COPD (doença pulmonar obstrutiva crônica) e 20 eram fumantes com COPD categorizados conforme a severidade do problema em moderado, severo e muito severo. O diagnóstico para estes pacientes foi realizado por dois profissionais da saúde: um radiologista e um pneumologista com experiência em imagens de CT. Como forma de simplificar os testes realizados para a ideia proposta, utilizaram-se apenas as imagens da altura inferior do pulmão, as quais são classificadas em seis classes: "No Emphysema"(19 imagens), "Minimal"(7 imagens), "Mild"(4 imagens), "Moderate"(3 imagens), "Severe"(2 imagens) e "VerySevere"(2 imagens).

6.2 Experimentos e resultados

A seguir, são apresentados os experimentos realizados. Inicialmente, são fixados quatro atributos e efetuados testes com diferentes vetores de características, extraídas a partir dos blocos da TWD-2D de Haar. O primeiro teste, denotado por "Level 1", consiste em formar um vetor com as características referentes às energias nas direções horizontal, vertical e diagonal ($EH1$, $EV1$, $ED1$), além da média dos coeficientes de escala para o primeiro nível

da TWD-2D ($MC1$). Assim, $x_i = (EH1, EV1, ED1, MC1)$.

No segundo e terceiro testes, denotados por "Level 2" e "Level 3", o vetor de atributos é formado com estas mesmas características, mas adquiridas no segundo e terceiro níveis da TWD-2D, respectivamente. Com isso, tenta-se verificar em qual nível da transformada estão contidas informações mais significativas para a classificação. No quarto teste, denominado "Direção H", o vetor é construído apenas com os valores das energias na direção horizontal, extraídas do primeiro, segundo e terceiro níveis, além da média dos coeficientes de escala do terceiro nível. No quinto e sexto testes, o vetor é construído com a mesma estratégia do teste anterior, porém com as características capturadas nas direções vertical e diagonal, denotados respectivamente de "Direção V" e "Direção D".

É importante observar a variação das energias EV , EH e ED , calculadas através dos coeficientes H , V e D de cada nível, em função das diferentes classes analisadas. A Figura 2 apresenta representantes das classes "Minimal", "Moderate" e "VerySevere", além das energias médias nos 3 níveis da TWD-2D para cada uma dessas classes. Nota-se uma variação significativa das energias para cada nível para cada classe. Por exemplo, a energia EH para os 3 níveis decresce entre as classes, no entanto a energia correspondente aos coeficientes V é crescente em cada nível. Já a energia ED em cada nível tem variação menos acentuada. Estas variações são atributos que contribuem no desempenho do classificador.

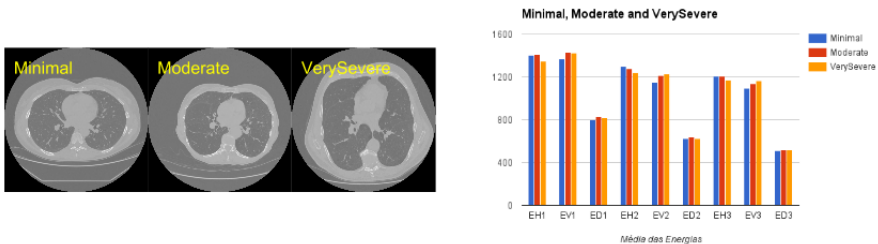


Figura 2: À esquerda, imagens de 3 classes. À direita, a energia média de cada bloco V , H e D da TWD-2D nos 3 níveis analisados.

Todos os testes desenhados nesta seção são baseados na técnica estatística *Leave-one-out*, na qual retira-se uma amostra do conjunto de dados que pertence a alguma das classes já conhecidas. Esta amostra passa, então, a ser a imagem de teste, comparada em relação aos elementos restantes em cada uma das classes. Com isso pode-se medir a precisão das classes e o quão distintas elas realmente são. Uma vez que o número de imagens disponíveis no banco de dados é pequeno e não homogêneo para cada classe, foi necessário um reagrupamento das classes originais apresentadas na seção 6.1 de tal forma que se pudesse realizar os experimentos. O novo reagrupamento é mostrado na Tabela 1. Inicialmente, realizaram-se testes utilizando a distância euclidiana ao invés de Mahalanobis com o objetivo de analisar os resultados obtidos por esta medida de similaridade. Os testes feitos foram os mesmos descri-

Classe	Número de Imagens	Agrupamento
NoEmphysema	19	NoEmphysema
Minimal	11	Minimal + Mild
Severe	7	Moderate + Severe + VerySevere

Tabela 1: Agrupamento de classes de imagens similares.

tos acima, utilizando-se um vetor de características de quatro elementos, mais precisamente a energia nas direções vertical, horizontal e diagonal e a média dos coeficientes de escala do nível mais grosseiro. Os resultados são apresentados na Tabela 2. A baixa performance de classificação apresentada nesta tabela corrobora para a não utilização da distância euclidiana.

Classe	Level 1	Level 2	Level 3	Direção H	Direção V	Direção D
No Emphysema	47.36%	42.10%	36.84%	52.63%	42.10%	52.63%
Minimal	18.18%	9.09%	27.27%	18.18%	9.09%	27.27%
Severe	14.28%	14.28%	0	28.57%	0	42.85%

Tabela 2: Precisão do algoritmo via distância euclidiana aplicado aos grupos.

A Tabela 3 apresenta os resultados quanto à precisão de acertos, gerados pelos seis testes descritos acima. Verifica-se que os maiores acertos para a classe "No Emphysema" foram gerados por vetores de características formados pelo primeiro e terceiro níveis da transformada, além das características coletadas na direção diagonal. Já a classe "Minimal" obteve mais acertos no terceiro nível e na direção diagonal. Percebeu-se também que a classe "Severe" teve uma maior precisão de acertos no vetor criado na direção vertical. Verifica-se ainda que os resultados obtidos com esta medida de similaridade são superiores quando comparados aos resultados gerados pela distância euclidiana da Tabela 2. Após a análise

Classe	Level 1	Level 2	Level 3	Direção H	Direção V	Direção D
No Emphysema	89.47%	78.94%	89.47%	73.68%	84.21%	89.47%
Minimal	45.45%	45.45%	54.54%	36.36%	9.09%	54.54%
Severe	28.57%	42.85%	28.57%	0	71.42%	0

Tabela 3: Precisão do algoritmo aplicado aos grupos.

da Tabela 3, criou-se um vetor de características com a expectativa de se aumentar o número de acertos para as três classes em questão. Desta forma, construiu-se um vetor híbrido com seis elementos, no qual foram mescladas as características geradas pelo nível 1 e pela direção vertical. Os seis elementos deste vetor são: energias nas direções horizontais, verticais e diagonais do primeiro nível da transformada, as energias na direção

vertical no segundo e terceiro nível e a média dos coeficientes de escala do terceiro nível, $x_i = (EH1, EV1, ED1, EV2, EV3, MC3)$. Os resultados para o teste com este vetor são apresentados na Figura 3. Nesta figura, é possível observar que houve um aumento na taxa de acerto da classe "Severe" para 85.71%. A classe "No Emphysema" apresentou a mesma porcentagem de acerto do teste para o primeiro nível, ou seja 89.47%, enquanto que a classe "Minimal" teve uma precisão de 54.54%.

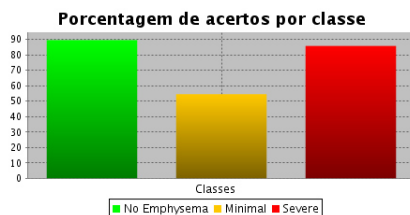


Figura 3: Taxa de acerto para a aplicação do vetor de características híbrido de seis elementos sobre as classes "No Emphysema", "Minimal" e "Severe".

Devido à restrição imposta pela falta de imagens distintas em cada uma das classes, realizou-se um novo experimento, no qual foram reagrupadas as 6 classes mencionadas anteriormente em apenas duas classes distintas, "No Emphysema" e "Severe", apresentado na Tabela 4. A Figura 4 mostra os resultados com a aplicação do vetor de características híbrido, formado anteriormente, sobre o conjunto de imagens. Houve uma melhoria quanto à taxa de acerto, pois o sistema acertou a classificação de imagens da classe "No Emphysema" em 93.33% dos casos, enquanto que para a classe "Severe", o sistema acertou 85.71%. Como os

Classes	Número de Imagens	Agrupamento
NoEmphysema	30	NoEmphysema + Minimal + Mild
Severe	7	Moderate + Severe + VerySevere

Tabela 4: Agrupamento com 2 classes de imagens.

resultados apresentados na Figura 4a apresentaram resultados competitivos quanto à taxa de acerto para as duas classes em questão, um último experimento foi desenhado para analisar a taxa de acerto entre as classes "No Emphysema" e "Minimal", agrupadas de acordo com a Tabela 5. Com a utilização do mesmo vetor híbrido com os seis elementos formados anteriormente, foram obtidos os resultados da Figura 4b. Houve um aumento na taxa de acerto da classe "Minimal", passando a ser de 63.63%, enquanto que a classe "No Emphysema" teve uma taxa de 89.47%.

A Tabela 6 apresenta os resultados da Figura 4a, sob o ponto de vista de verdadeiros-positivos (TP) e falsos-negativos (FN). Os cruzamentos TP x TP e FN x FN significam que

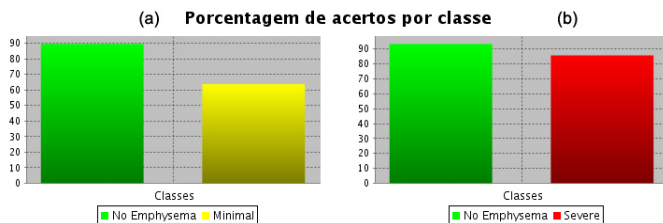


Figura 4: Taxa de acerto com o vetor de características híbrido de 6 elementos sobre as classes "No Emphysema" e "Minimal"(a), e "No Emphysema" e "Severe"(b).

Classes	Número de Imagens	Agrupamento
NoEmphysema	19	NoEmphysema
Minimal	11	Minimal + Mild

Tabela 5: Agrupamento das classes.

	TP	FN
TP	6	1
FN	2	28

Tabela 6: Tabela com TP e FN.

o sistema desenvolvido acertou a classificação das imagens, de acordo com o diagnóstico realizado sobre elas. O cruzamento FN x TP significa que a imagem foi diagnosticada como saudável e o sistema a classificou como pertencente a grupo de risco. Por fim, a relação TP x FN corresponde a imagens que foram diagnosticadas com presença de enfisema enquanto que o sistema as classificou como saudáveis, sendo este considerado o pior resultado para o algoritmo proposto, uma vez que um paciente poderia estar com um elevado grau de enfisema e o sistema o classificaria como saudável. Pela Tabela 6, o sistema acertou 34 casos, errando apenas três vezes. Duas imagens saudáveis foram classificadas como de risco, enquanto que apenas uma das imagens referentes a um grupo de risco foi classificada como saudável.

Comparativamente, o algoritmo apresentado possui maior eficiência que vetores de características cuja similaridade é medida através da distância euclidiana, a exemplo do vetor construído a partir de um GFB (Gaussian Filter Bank) apresentado em [6], cuja eficiência em termos de acertos na classificação do mesmo banco de imagens foi de 61.3%.

7 Conclusão

Neste trabalho, desenvolveu-se um algoritmo de classificação de imagens médicas de pulmão, retiradas de um banco de imagens de enfisema pulmonar. Esta classificação é baseada na distância de Mahalanobis aplicada no reconhecimento de padrões. Utilizou-se a transformada wavelet de Haar 2D para a extração de características, energia dos coeficientes de detalhe e média dos coeficientes de escala, como forma de auxiliar no processo de classifi-

cação. Os resultados mostraram um considerável nível de acerto por parte do sistema. Como trabalhos futuros, outros bancos de imagens poderão ser incorporados ao sistema a fim de validá-lo com ainda maior confiabilidade. A implementação também pode considerar outros métodos de classificação como o K-Means.

8 Agradecimentos

Os autores agradecem ao professor Marcos d'Ornellas (UFSM) pela indicação do banco de imagens utilizado, à FAPERGS PG 1873-25.51/13-0 e FAPERGS Bolsa Mestrado nº 0382-2551/14-9, e ao Programa de Pós-Graduação em Informática (PPGI) da UFSM.

Referências

- [1] Ceyhun Burak Akgul, Daniel L. Rubin, Sandy Napel, Christopher F. Beaulieu, Hayit Greenspan, and Burak Acar. Content-based image retrieval in radiology: Current status and future directions. *J. Digital Imaging*, pages 208–222, 2011.
- [2] A. Amini. Matrix operations in java. <http://www.codeproject.com/Articles/405128/Matrix-operations-in-Java>, 2013. (2015-01-08).
- [3] Mohammad Asadzadeh, Eliza Hashemi, and Alice Kozakevicius. On efficiency of combined daubechies wavelets and statistical parameters applied in mammography. *Applied and Computational Mathematics: an international journal*, 12(3):289–306, 2013.
- [4] Xavier De Coster, Charles De Groote, Arnaud Destiné, Pierre Deville, Laurent Lamouline, Thibault Leruitte, and Vincent Nuttin. Mahalanobis distance, jaro-winkler distance and ndollar in usigesture. 1.
- [5] S. Dua, U.R. Acharya, P. Chowriappa, and S.V. Sree. Wavelet-based energy features for glaucomatous image classification. *Information Technology in Biomedicine, IEEE Transactions on*, 16(1):80–87, Jan 2012.
- [6] Mehrdad J Gangeh, Lauge Sørensen, Saher B Shaker, Mohamed S Kamel, and Marleen De Bruijne. Multiple classifier systems in texton-based approach for the classification of ct images of lung. In *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*, pages 153–163. Springer, 2011.
- [7] Gilson A Giraldi, Paulo S Rodrigues, Edson C Kitani, and Carlos E Thomaz. Dimensionality reduction, classification and reconstruction problems in statistical learning approaches. *Revista de Informática Teórica e Aplicada*, 15(1):137–169, 2008.
- [8] P.W. Huang and S.K. Dai. Image retrieval by texture similarity. *Pattern Recognition*, 36(3):665 – 679, 2003.

- [9] Alice Kozakevicius and Alex A Schmidt. Wavelet transform with special boundary treatment for 1d data. *Computational and Applied Mathematics*, 32(3):447–457, 2013.
- [10] Shurong Liu, Kun Han, Zhibin Song, and Misheng Li. Texture characteristic extraction of medical images based on pyramid structure wavelet transform. In *Computer Design and Applications (ICCD), 2010 International Conference on*, volume 1, pages V1–342–V1–345, June 2010.
- [11] R. De Maesschalck, D. Jouan-Rimbaud, and D.L. Massart. The mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, 50(1):1 – 18, 2000.
- [12] M.A. Pasnur and Malge P.S. Image retrieval using modified haar wavelet transform and k means clustering. *International Journal of Emerging Technology and Advanced Engineering*, 3, 2013.
- [13] Fernanda Rispoli Quartieri, Jacob Scharcanski, Letícia Vieira Guimarães, and Adalberto Schuck Junior. Representação e classificação de texturas da íris baseada na transformada ótima de gabor. *Revista de informática teórica e aplicada. Porto Alegre. Vol. 15, n. 2 (2008), p. 105-120*, 2008.
- [14] K. Rajakumar and S. Muttan. Medical image retrieval using energy efficient wavelet transform. In *Computing Communication and Networking Technologies (ICCCNT), 2010 International Conference on*, pages 1–5, July 2010.
- [15] Md. Iqbal Hasan Sarker and Md. Shahed Iqbal. Content-based image retrieval using haar wavelet transform and color moment. *Smart CR*, pages 155–165, 2013.
- [16] Nicu Sebe and Michael S Lew. Wavelet based texture classification. In *Pattern Recognition. Proceedings. 15th International Conference on*, volume 3. IEEE, 2000.
- [17] Saher B Shaker, Karin A von Wachenfeldt, Susanne Larsson, Iréne Mile, Sofia Persdotter, Magnus Dahlbäck, Per Broberg, Berend Stoel, Karen S Bach, Marianne Hestad, et al. Identification of patients with chronic obstructive pulmonary disease (copd) by measurement of plasma biomarkers. *The clinical respiratory journal*, 2(1):17–25, 2008.
- [18] Linda Shapiro and George C Stockman. Computer vision. ed: *Prentice Hall*, 2001.
- [19] Lauge Sørensen, Saher B. Shaker, and Marleen de Bruijne. Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Transactions on Medical Imaging*, 29(2):559–569, Feb 2010.
- [20] Eric J. Stollnitz, Tony D. Derose, and David H. Salesin. Wavelets for computer graphics: A primer - part 1. *IEEE Computer Graphics and Applications*, 15:76–84, 1995.

- [21] A.J.M. Traina, C.A.B. Castanon, and Jr. Traina, C. Multiwavemed: a system for medical image retrieval through wavelets transformations. In *Computer-Based Medical Systems, 2003. Proceedings. 16th IEEE Symposium*, pages 150–155, June 2003.
- [22] R. Wicklin. What is mahalanobis distance? <http://blogs.sas.com/content/iml/2012/02/15/what-is-mahalanobis-distance>, 2012. (2014-12-12).