



Universidade Federal de Uberlândia
Faculdade de Engenharia Elétrica
Graduação em Engenharia Biomédica

ANA CLARA CASTRO PIMENTEL SILVA ARAÚJO

**Comparação de famílias *wavelets* para *denoising* de
imagens mamográficas digitais**

Uberlândia
2018

ANA CLARA CASTRO PIMENTEL SILVA ARAÚJO

**Comparação de famílias *wavelets* para *denoising* de
imagens mamográficas digitais**

Trabalho apresentado como requisito parcial de
avaliação na disciplina Trabalho de Conclusão
de Curso do Curso de Engenharia Biomédica da
Universidade Federal de Uberlândia.

Orientador: Prof.a Dra. Ana Cláudia Patrocínio

Assinatura do orientador

Uberlândia

2018

Dedico este trabalho a todas as mulheres que passaram por essa dura experiência e a seus familiares, que juntos pela fé superaram esse momento de dor.

AGRADECIMENTOS

Sou grata, primeiramente, à Deus pelo presente que foi cursar engenharia biomédica na UFU. Aos meus professores da UFU, em especial Ana Cláudia Patrocínio por me orientar com tanto carinho e paciência e, também por ter proposto um tema tão relevante para este trabalho. Sou grata aos meus pais e à minha avó Nenê pelo amor, apoio e paciência e por terem me ensinado a ser uma pessoa com valores e ver o próximo com amor. Ao Victor e sua família pelo carinho e apoio durante minha jornada. Sou grata aos meus amigos, colegas de turma e ao pessoal do BioLab pelo carinho, ajuda e incentivo. E à todas as pessoas que de alguma forma colaboraram com este trabalho.

RESUMO

Há um grande interesse da comunidade científica, juntamente com as organizações de saúde, em melhorar o diagnóstico do câncer de mama, visto que é uma das causas principais da morte de mulheres adultas. A técnica mais utilizada para isso, principalmente para estágios iniciais, é a análise da imagem mamográfica da paciente, portanto melhorar a qualidade dessas imagens é uma tarefa fundamental e diretamente ligada ao diagnóstico correto. Para isso, são utilizadas ferramentas matemáticas de processamento para diminuir a quantidade de ruídos presentes na imagem final. Neste trabalho discutimos o uso de diversas famílias de *wavelets* como técnica de *denoising* e comparamos os melhores resultados para cada tipo de padrão de mama, separadas segundo a classificação proposta pelo BIRADS™. Um banco de dados contendo quarenta imagens, dez de cada padrão, foi processado no MatLab a fim de obter a quantificação da relação sinal-ruído (SNR) assim como relação sinal-ruído de pico (PSNR). Esses valores de SNR e PSNR foram usados como ferramenta para avaliar os melhores desempenhos. Na literatura quanto maiores os valores de SNR e PSNR, melhor a relação entre sinal-ruído e, portanto, menos ruidosa é a imagem final. Inicialmente, realizou-se 26 testes em quatro imagens (uma de cada padrão de mama) com diferentes famílias *wavelets* e níveis. Com o resultado desses testes notou-se que, independentemente do padrão de mama, os piores e melhores resultados eram comuns às quatro imagens. Dessa forma, oito *wavelets* foram selecionadas para o processamento das demais imagens. Os resultados das quarenta imagens foram analisados e o melhor desempenho foi obtido pela *wavelet* coiflet 3 no nível 1: maior SNR em 82,5% das imagens e maior PSNR em 92,5%. Não mostrou relação com algum padrão de mama específico.

Palavras-chave: mamografia, *wavelets*, *denoising*.

ABSTRACT

There is a big interest of the scientific community with health organizations, to improve the diagnosis of breast cancer, since it is one of the main causes of death in adult women. The most used technique for this, especially for initial stages, is the mammographyc image analysis of the patient, so improving the quality of these images is a fundamental task and directly linked to the correct diagnosis. For this, mathematical processing tools are used to reduce the amount of noise present in the final image. In this paper we discuss the use of several wavelet families as a denoising technique and compare the best results for each type of breast pattern, separated according to the classification proposed by BIRADS TM. A database containing forty images, ten of each pattern, was processed in MatLab in order to obtain signal-to-noise ratio (SNR) quantification as well as peak signal-to-noise ratio (PSNR). These values of SNR and PSNR were used as a tool to evaluate the best performances. In the literature, the higher the SNR and PSNR values, the better the signal-to-noise ratio and, therefore, the less noisy is the final image. Initially, 26 tests were performed on four images (one of each breast pattern) with different wavelet families and levels. With the result of these tests it was noticed that, regardless of breast pattern, the worst and best results were common to all four images. Thus, eight wavelets were selected for the processing of the other images. The results of the forty images were analyzed and the best performance was obtained by the coiflet 3 wavelet at level 1: higher SNR in 82.5% of the images and higher PSNR in 92.5%. It was not related to any specific breast pattern.

Keywords: *mammographyc, wavelets, denoising.*

LISTA DE ILUSTRAÇÕES

Figura 1: Exemplos de cada padrão de mama.....	17
Figura 2: Janela inicial do Matlab, comando wavemenu na Command Window.....	19
Figura 3: Menu principal do <i>Wavelet Toolbox</i>	19
Figura 4: Selecionando a imagem para processamento.....	20
Figura 5: Detalhamento do menu lateral.....	21
Figura 6: Tela de primeira análise e apresentação da imagem decomposta.....	21
Figura 7: Etapa de de-noise da imagem escolhida.....	22
Figura 8: Apresentação da imagem processada e processo para salvá-la no computador.....	22
Figura 9: Comparativo da imagem original com as processadas usando <i>wavelets</i> . a) Imagem mamográfica original do padrão 1; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.....	25
Figura 10: Comparativo da imagem original com as processadas usando <i>wavelets</i> . a) Imagem mamográfica original do padrão 2; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.....	26
Figura 11: Comparativo da imagem original com as processadas usando <i>wavelets</i> . a) Imagem mamográfica original do padrão 3; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.....	26
Figura 12: Comparativo da imagem original com as processadas usando <i>wavelets</i> . a) Imagem mamográfica original do padrão 4; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.....	27
Figura 13: Funções wavelet e escalada de coiflet 3.....	30

LISTA DE TABELAS

Tabela 1: Famílias <i>wavelet</i> disponíveis no Toolbox do MatLab™	18
Tabela 2: Resultados do primeiro teste para quatro mamografias.....	23
Tabela 3: Famílias de <i>wavelets</i> e respectivos níveis escolhidos para <i>denoising</i>	24
Tabela 4: Média SNR e PSNR de famílias <i>wavelet</i> para as mamas de padrão 1.....	27
Tabela 5: Média SNR e PSNR de famílias <i>wavelet</i> para as mamas de padrão 2.....	27
Tabela 6: Média SNR e PSNR de famílias <i>wavelet</i> para as mamas de padrão 3.....	28
Tabela 7: Média SNR e PSNR de famílias <i>wavelet</i> para as mamas de padrão 4.....	28
Tabela 8: Melhores resultados SNR e PSNR para cada família <i>wavelet</i> , separado por padrão de mama.....	29

LISTA DE ABREVIATURAS E SIGLAS

ICA - Análise de Componentes Independentes

GUI - Graphical User Interface

TWD - Transformada Wavelet Discreta

SNR - Signal-Noise Ratio

PSNR - Peak Signal-Noise Ratio

MSE - Mean Squared Error

ER - Energia Retida

DICOM - Digital Imaging and Communications in Medicine

TWDE - Transformada Wavelet Discreta Estacionária

Haar - Haar

Db - Daubechies

Sym - Symlets

Coif - Coiflets

Bior - BiorSplines

Rbio - ReverseBior

Meyr - Meyer

Dmey - DMeyer

Gaus - Gaussian

Mexh - Mexican hat

Morl - Morlet

Cgau - Complex Gaussian

Shan - Shannon

Fbsp - Frequency B-Spline

cmor - Complex Morlet

SUMÁRIO

ABSTRACT	5
LISTA DE ILUSTRAÇÕES	6
LISTA DE TABELAS	7
LISTA DE ABREVIATURAS E SIGLAS	8
1 INTRODUÇÃO	10
2 REVISÃO BIBLIOGRÁFICA	12
3 ESTADO DA ARTE	15
4 METODOLOGIA	16
5 RESULTADOS E DISCUSSÕES	25
6 CONCLUSÕES	30
REFERÊNCIAS	30

1 INTRODUÇÃO

Em 2013, a Organização Mundial da Saúde reportou o câncer de mama como uma das principais causas de morte em mulheres. Um dos métodos mais eficazes de rastreamento e detecção de tumores na mama é a mamografia, um exame que utiliza raios X de baixa energia e a menor dose possível para gerar uma imagem final de qualidade. O diagnóstico e tratamento dessa e muitas outras doenças, desde as que estão em estágios iniciais de identificação aos estágios mais complexos dependem profundamente das imagens médicas.

No Brasil, o câncer de mama é a doença com maior grau de mortalidade em mulheres, na maioria dos casos, devido ao estágio avançado da doença quando descoberta, decorrente do diagnóstico tardio. Foram estimados mais de 57 mil novos casos de câncer de mama no ano passado, com um risco estimado de 56.09 casos para cada 100 mil mulheres brasileiras. No entanto, mediante a realização de diagnósticos e tratamentos no início da doença, o prognóstico pode ser bem melhor. O sintoma mais comum de câncer de mama é o aparecimento de nódulo, geralmente indolor, duro e irregular, mas há tumores de consistência branda, globosos e bem definidos (INCA, 2015).

O baixo contraste entre os pequenos tumores e os tecidos glandulares adjacentes, que na maioria das vezes possui absorções semelhantes, dificulta a visualização de lesões. Além disso, durante o processo de aquisição, transmissão e recepção, as imagens digitais são corrompidas por ruído. Por isso, para aprimorar a mamografia é necessária a etapa de pré-processamento para reduzir o nível de ruído da imagem preservando as estruturas mamográficas para se obter a melhor estimativa da imagem original a partir da corrompida (BHATNAGAR, JAIN, 2016).

A mamografia digital foi desenvolvida em parte para abordar algumas das limitações da mamografia de filme, separar a aquisição e exibição de imagens, permitindo a otimização de ambas. O processamento de imagens digitais permite que o grau de contraste na imagem seja manipulado, podendo ser aumentado nas áreas densas da mama com o menor contraste (PISANO et al., 2005).

Dentre algumas técnicas de *denoising*, as mais utilizadas são o filtro adaptável de Wiener, desmembramento por análise de componentes independentes e *wavelets*. O filtro adaptável de Wiener realiza pouca suavização em regiões onde a variância é grande e mais suavização para pequenas variâncias. Esta abordagem geralmente produz melhores resultados do que a filtragem linear. O filtro adaptativo é mais seletivo do que um linear comparável, preservando bordas e outras partes de alta frequência de uma imagem. O objetivo do filtro Wiener é calcular uma estimativa estatística de um sinal desconhecido usando um sinal relacionado como entrada e filtrando um sinal conhecido para produzir uma estimativa como saída (MAYO, RODENAS, VERDU, 2004).

A análise de componentes independentes (ICA) é um método para representar um conjunto de vetores de dados multidimensionais em uma base onde os componentes são tão independentes quanto possível. Muitas vezes isso significa que é preciso encontrar uma transformação que forneça um vetor cujos componentes são dispersos, ou seja, a probabilidade de um componente ser significativamente diferente de zero é muito baixa. Os métodos de *denoising* ICA dependem do fato de que os componentes transformados possuem distribuições esparsas (super gaussianas), de modo que as técnicas de *denoising* tentam reduzir o ruído gaussiano por *shrinkage* (*soft thresholding*) dessas componentes (MAYO, RODENAS, VERDU, 2004).

As *wavelets* são funções matemáticas capazes de separar o dado em diferentes componentes de frequência e estudam cada componente com uma resolução combinada a sua escala. Existem semelhanças entre as transformadas *wavelet* e de Fourier, a matriz da transformada inversa de ambas é a transposta da original e podem ser consideradas como uma rotação do espaço de funções para um outro domínio (frequência). Neste novo domínio as funções bases da transformada de Fourier são cossenos e senos e da transformada *wavelet* são as *wavelets* mãe. A função *wavelet* mãe produz através de escalonamentos as famílias de funções base. As bases *wavelets* devem satisfazer alguns critérios como alta velocidade de processamento, regularidade, suavidade, boa localização (no tempo, espaço ou frequência) e boa descrição de sinais (SILVA, 2007).

Uma propriedade essencial é que elas estão bem localizadas no tempo (ou no caso da imagem, no espaço) e na frequência. No domínio *wavelet*, o ruído é distribuído

uniformemente entre todos os coeficientes, enquanto o sinal é bastante disperso, sendo concentrado em um pequeno número de coeficientes. Este método consiste na decomposição *wavelet* de uma imagem, que é a versão bidimensional de uma decomposição de sinal, e sua representação em uma função de base (MAYO, RODENAS, VERDU, 2004).

Diferente das funções seno e cosseno de Fourier, as funções individuais *wavelet* estão localizadas no espaço. Tal característica combinada com as *wavelets* no domínio da frequência resulta em várias aplicações, no campo gráfico permite a detecção de características, compressão de dados e remoção de ruído (SILVA, 2007).

O objetivo geral deste trabalho consiste em avaliar o GUI *Wavelet* do MatLab™ para ser utilizado no processo de *denoising* de imagens mamográficas, investigando os parâmetros e as famílias disponíveis na ferramenta.

2 REVISÃO BIBLIOGRÁFICA

Imagem mamográfica digital

O mamógrafo é um equipamento de raios X que obtém a imagem a partir de diferenças de coeficiente de atenuação linear dos diversos tipos de tecidos, detectando estruturas de diferentes densidades e formas. Os raios X são produzidos no tubo de raios X, composto por dois eletrodos submetidos à vácuo e à tensão, gerando assim um campo elétrico. O cátodo, eletrodo negativo, é formado por um filamento de ponto de fusão alto com um elemento resistivo que libera elétrons a partir da passagem de corrente elétrica. Tais elétrons são acelerados até colidirem com o eletrodo positivo. No anodo está o alvo, revestido geralmente por molibdênio, ródio, ou tungstênio, no qual possui uma pequena área denominada ponto focal, onde efetivamente é produzida a radiação (WOLBARST, 1993).

O filtro é responsável por impedir e retirar fótons de baixa energia do feixe de raios X, que nada contribuiriam para a formação da imagem, evitando assim que a paciente receba desnecessariamente uma maior dose de radiação. Durante o exame ocorre a compressão da mama, importante etapa para diminuir sua espessura e,

dessa forma, diminuir o borramento da imagem causado pela radiação espalhada. Outra finalidade da compressão é imobilizar a mama e possibilitar que a região de interesse fique mais próxima do detector da imagem, evitando a sobreposição de tecidos (POULOS, 2003).

A imagem digital é definida como uma função bidimensional $f(x,y)$, em que x e y correspondem à coordenadas espaciais e a amplitude do sinal corresponde ao nível de cinza da imagem na coordenada. Quando os valores de x , y e f são números finitos (discretos), a imagem é classificada digital (GONZALEZ, 2010). Imagens são frequentemente corrompidas em sua aquisição e transmissão por vários tipos de ruídos e diversas técnicas estão sendo estudadas no campo de processamento de imagens para resolver essa questão.

Wavelets

Wavelets foram encontradas inicialmente na literatura nos trabalhos de Haar, Grossman e Morlet. Elas são capazes de separar o dado em componentes de frequência e, assim, possibilita estudá-las com uma resolução combinada à sua escala. Assim como a transformada de Fourier, a Transformada *Wavelet* Discreta (TWD) possui a função base no domínio da frequência, denominada *wavelet* mãe (SILVA, 2007). Uma *wavelet* mãe deve possuir um valor médio igual a zero, o que requer que o núcleo de transformação da transformada *wavelet* suporte à localização no tempo, oferecendo assim o potencial para capturar os picos que ocorrem instantaneamente em um curto período de tempo.

Wavelets são obtidas através da decomposição do sinal em duas funções ψ e ϕ , a primeira denominada “*wavelet* mãe” e a segunda, “função de escalamento”. Ambas descrevem o banco de filtros passa-faixa a ser empregados na análise. Na Equação 1 temos a descrição da forma geral de uma *wavelet* mãe, onde a é o parâmetro de escala e b é o parâmetro de mudança ou conversão, na Equação 2 apresenta a condição que deve obedecer. E na Equação 3 temos a fórmula da transformada *wavelet* para sinais discretos.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (1)$$

$$\int \psi(t)dt = 0 \quad (2)$$

$$F_{m,n}(a,b) = a_0^{-m/2} \int f(t)\psi(a_0^{-m}t - nb_0)dt \quad (3)$$

As propriedades e vantagens de uma família de *wavelets* dependem das características da *wavelet* mãe (ERGHEN, 2012). O produto da TWD é um conjunto de coeficientes organizados da maneira que permite não apenas análises de espectro do sinal, mas também o comportamento espectral do sinal no tempo (GUPTA, CHOUBEY, 2015).

Um sinal ou uma imagem infelizmente é corrompido por vários fatores como ruído durante a aquisição ou transmissão. Esses efeitos ruidosos diminuem o desempenho da análise visual e computadorizada. O processo de *denoising* pode ser descrito como remover o ruído, mantendo e não distorcendo a qualidade do sinal processado ou imagem (CHEN, 2003).

O modo tradicional de *denoising* para remover o ruído de um sinal ou imagem é usar um filtro passa-baixa ou passa-banda com frequências de corte. As técnicas tradicionais de filtragem são capazes de remover um relevante do ruído, porém são ineficazes se o ruído estiver na banda do sinal a ser analisado. Portanto, muitas técnicas de *denoising* são propostas para superar este problema (ERGHEN, 2012).

A transformada *wavelet* aparece nesse cenário como uma técnica emergente de processamento de sinais e imagens, podendo ser utilizada para representar sinais não-estacionários reais com alta eficiência e também na análise de sinais com descontinuidades e picos. A transformada ganhou ampla aceitação no processamento de sinais, *denoising* e compressão de imagens devido à sua natureza inerente à multiresolução. Os esquemas de codificação *wavelet* são especialmente adequados

para aplicações em que a escalabilidade e a degradação tolerável são importantes (GUPTA, CHOUBEY, 2015).

Após o processo de *denoising*, o desempenho pode ser medido comparando o sinal processado e o sinal original. Muitos métodos têm sido propostos para medir o desempenho de algoritmos de *denoising*, entre eles a relação sinal-ruído (SNR) e a relação pico-ruído de sinal (PSNR) geralmente são aceitas como medida de qualidade do sinal e das imagens, respectivamente (ERGHEN, 2012).

3 ESTADO DA ARTE

No trabalho de Erghen publicado em 2012, foram utilizadas *wavelets* como técnica de *denoise* e, segundo o autor, elas oferecem alta qualidade e flexibilidade para o problema de ruído de sinais e imagem. Os desempenhos dos métodos de *denoising* para muitas variações incluindo regras de limiar e o tipo de *wavelet* foram examinados nos exemplos de modo a apresentar os resultados de *denoising* adequados. As comparações foram feitas para três métodos de estimar limiar, tipos de *wavelets* e tipos de limiar. Os resultados mostraram que o fator mais importante na redução da *wavelet* é o nível de decomposição, e não o tipo de onda, o tipo de limiar ou a estimativa do valor do limiar. No entanto, não se notou diferenças nos métodos do nível um ao nível seis e constatou que o tipo *wavelet* não é muito importante se o número de oscilação não for muito baixo, o nível de decomposição é absolutamente dependente da banda de frequência do sinal a ser analisado e da sua frequência de amostragem (ERGHEN, 2012).

Gupta em 2015 publicou um trabalho em que utiliza *wavelets* para compactação de imagem. Nesse artigo apresenta a comparação do desempenho de *wavelets* discretas como a *Wavelet* de Haar e a *Wavelet* Daubechies para implementação em um sistema de compressão de imagens estáticas. Os desempenhos dessas transformações são comparados em termos de Erro Quadrático Médio (MSE) e Energia Retida (ER). O objetivo principal é investigar a compressão de imagem estática de uma imagem em escala de cinza usando a teoria de *wavelets* com MATLAB *Wavelet* Toolbox e a técnica 2D-TWD. Idealmente, durante a compressão, o número de zeros e a retenção de energia serão os mais altos possíveis significando a permanência de dados. Em Daubechies, a energia retida da *wavelet* é

de 98,03% e na transformada de haar é de 97,70%. Assim, a partir dos resultados, concluiu-se que a transformada *wavelet* Daubechies apresenta os melhores resultados em termos de retenção de energia, comparando com a transformada *wavelet* de Haar (GUPTA, CHOUBEY, 2015).

Em 2014 Jaiswal publicou um trabalho no qual fez *denoising* de ruídos dos tipos sal-pimenta e gaussiano em quatro etapas: filtragem da imagem; *denoising* da imagem por *wavelet* usando *thresholding*; limiar *hard* e filtragem aplicados simultaneamente em imagem ruidosa. A filtragem da imagem ruidosa foi feita com filtro mediana e Wiener estabelecendo um limiar *soft* e *hard* até encontrar o melhor limiar e finalmente combinar o resultado do método de filtragem e o limiar *hard*. O *denoising* foi executado com base em medidas de desempenho como PSNR (Relação Sinal / Ruído de Pico), MSE (Erro Quadrático Médio) e na qualidade visual da imagem. Os resultados de PSNR e MSE foram calculados comparando todos os casos. Durante a análise de várias técnicas, descobriram que a técnica de filtragem e limiarização de *wavelet* dá uma melhor concordância entre PSNR e MSE do que de outros (JAISWAL, UPADHYAY, SOMKUWAR, 2014).

Jumah em 2012 aplicou técnicas de limiares não-lineares no domínio da *wavelet*, tais como limites hard e soft, encolhimentos de *wavelet* como *Visu-shrink* e *SURE* (não-adaptativos), *Bayes* e *Normal Shrink* (adaptativos), usando Transformada *Wavelet* Discreta Estacionária (TWDE) para diferentes *wavelets*, em diferentes níveis, para *denoise* de uma imagem e determinar a melhor delas. O desempenho do algoritmo de *denoising* foi medido usando SNR e MSE como medidas de desempenho quantitativas para várias técnicas de limiarização (AL JUMAH, 2013).

4 METODOLOGIA

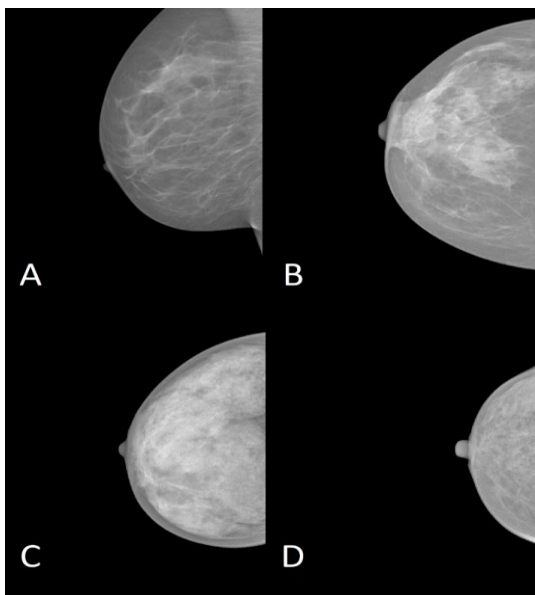
Para realização deste trabalho foram utilizadas quarenta imagens mamográficas digitais extraídas do banco virtual INbreast, provenientes do Centro Hospitalar de São João em Porto, Portugal (MOREIRA et al., 2012). As imagens foram obtidas do mesmo equipamento: mamógrafo MammoNovation FFDM com 12 bits de resolução de contraste, marca Siemens. As mamografias possuem tamanho 3328 x 4084 ou 2560 x 3328 pixels, dependendo da compressão da mama (de acordo com o tamanho). As imagens foram inicialmente salvas no padrão DICOM (*Digital Imaging*

and Communications in Medicine), mas para processamento em MatLab™ foram convertidas em extensão .bmp.

Os exames deste banco foram laudados por dois radiologistas experientes, usando a padronização proposta na quarta edição do BIRADS™ para caracterizar a densidade mamária de cada exame. As quatro categorias do BI-RADS™ usadas no banco INbreast (MOREIRA et al., 2012) são:

- Mamas predominantemente adiposas (25% do componente fibroglandular) (Figura 1 A);
- Mamas parcialmente gordurosas (com densidade de tecido fibroglandular em 26% a 50% do volume da mama) (Figura 1 B);
- Mamas com padrão denso e heterogêneo (51% a 75% de tecido fibroglandular) (Figura 1 C);
- Mamas muito densas, apresentando mais de 75% de tecido fibroglandular (Figura 1 D).

Figura 1: Exemplos de cada padrão de mama.



Fonte: MOREIRA et al. (2012).

A ferramenta principal utilizada foi o *Wavelet Toolbox* do MatLab™, composta por várias opções de processamento e aplicações para analisar e sintetizar sinais, imagens e dados que possuem comportamento regular pontuado com mudanças abruptas. A caixa de ferramentas inclui algoritmos para a transformada *wavelet* contínua (CWT) e coerência *wavelet*. Também contém algoritmos e visualizações de

wavelets discretas, análise do comportamento do conteúdo de frequência dos sinais ao longo do tempo e revela padrões variáveis no tempo comuns em múltiplos sinais. As famílias de *wavelets* disponíveis no *Wavelet Toolbox* do MatLab™ estão dispostos na Tabela 1.

Tabela 1: Famílias *wavelet* disponíveis no Toolbox do MatLab™.

Família <i>wavelet</i>	Sigla
Haar	haar
Daubechies	Db
Symlets	Sym
Coiflets	Coif
BiorSplines	Bior
ReverseBior	Rbio
Meyer	Meyr
DMeyer	Dmey
Gaussian	Gaus
Mexican hat	Mexh
Morlet	Morl
Complex Gaussian	Cgau
Shannon	Shan
Frequency B-Spline	Fbsp
Complex Morlet	cmor

O aplicativo GUI para processamento utilizando *wavelets* é acessado a partir do comando *wavemenu* digitado no *Command Window* do MatLab™, conforme Figura 2. Na janela principal existem os menus: *One-Dimensional*, *Two-Dimensional*, *Three-Dimensional*, *Multiple1-D*, *Wavelet Design*, *Specialized Tools 1-D*, *Specialized Tools 2-D*, *Display* e *Extension*, conforme apresentado na Figura 3.

Figura 2: Janela inicial do *Matlab*, comando *wavemenu* na *Command Window*.

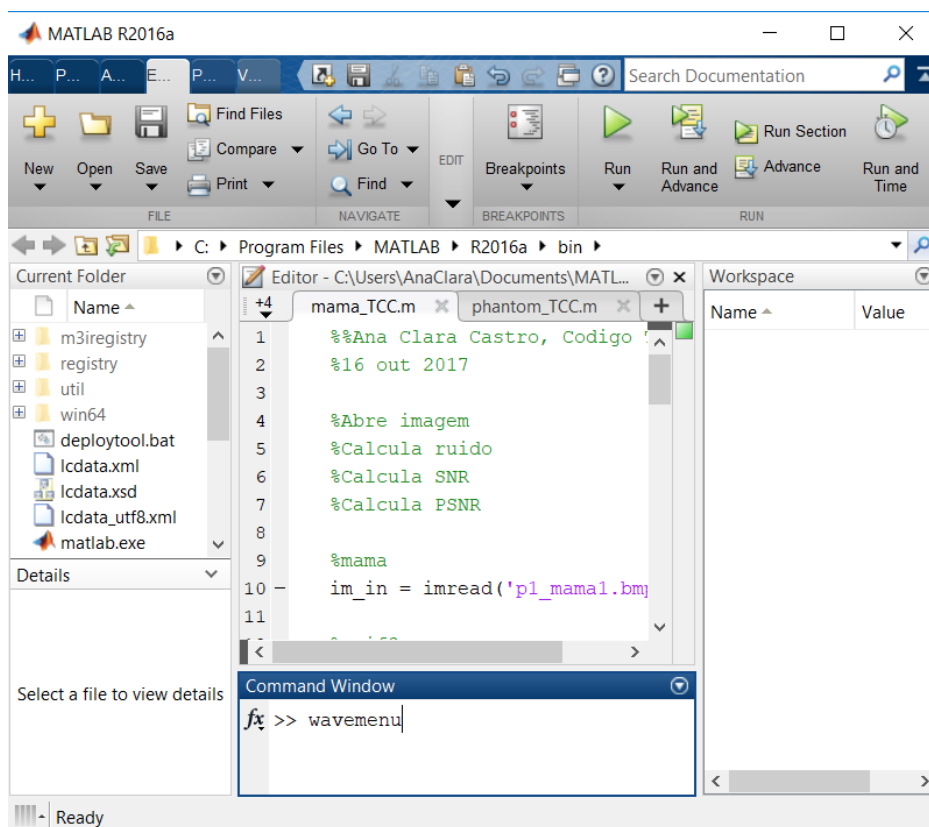
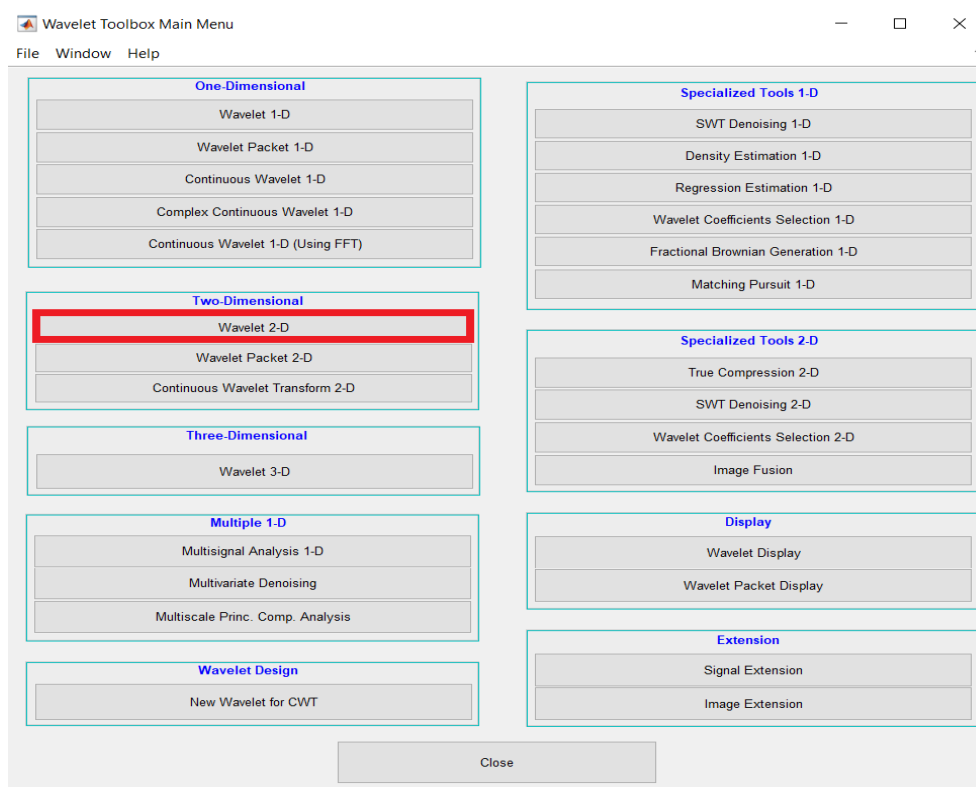


Figura 3: Menu principal do *Wavelet Toolbox*.



Para o processamento de *denoising* utilizando *wavelet* é preciso acessar o menu *Wavelet* 2D no grupo *Two-Dimensional* e assim a janela desta opção será inicializada. Nesta interface primeiramente deve-se escolher a imagem a ser processada através da opção *File, Load, Image* no menu superior (Figura 4). E então segue-se os seguintes passos:

- 1 - No menu lateral escolhe-se *gray* em *colormap* (Figura 5).
- 2 - No menu lateral, escolhe-se a família de *wavelet* em *Wavelet* e o nível de decomposição em *Level* (Figura 5).
- 3 - Ao clicar em *Analyze* as imagens processadas aparecem em estrutura árvore (padrão) (Figura 6).
- 4 - Deve-se clicar em *De-noise*, no menu lateral (Figura 5).
- 5 - Uma nova janela aparecerá, nela clica-se em *De-noise* novamente (Figura 7).
- 6 - A imagem processada aparecerá ao lado da original e para salvá-la deve-se clicar em *File, Save, De-noised Image* no menu superior (Figura 8).

Figura 4: Selecionando a imagem para processamento.

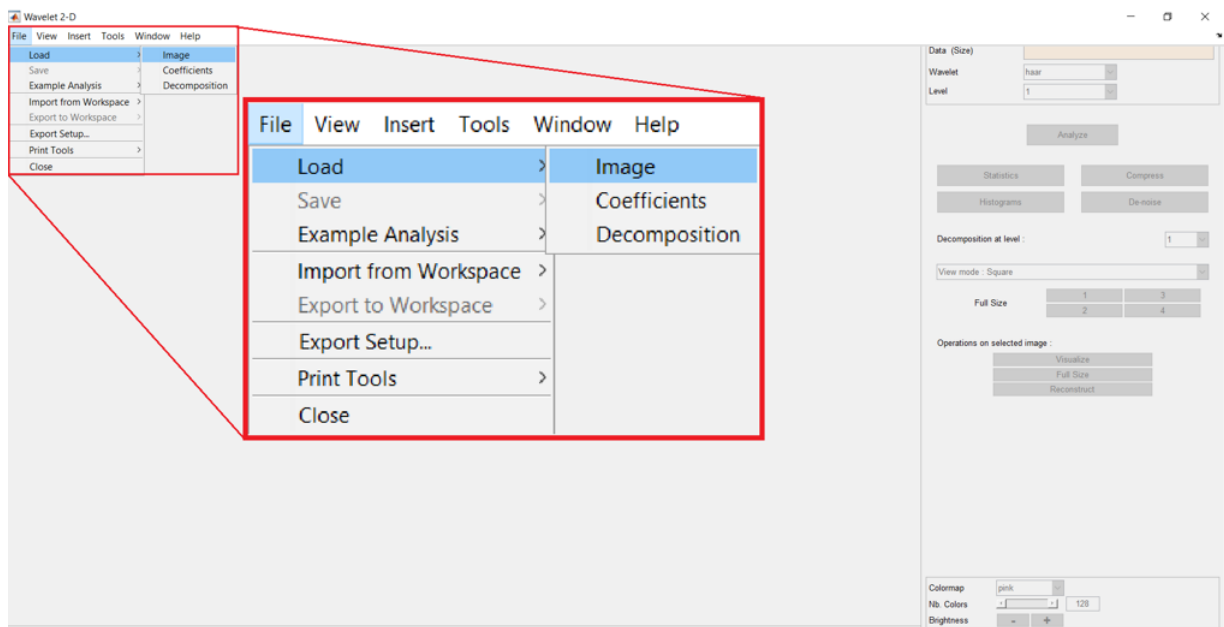


Figura 5: Detalhamento do menu lateral.

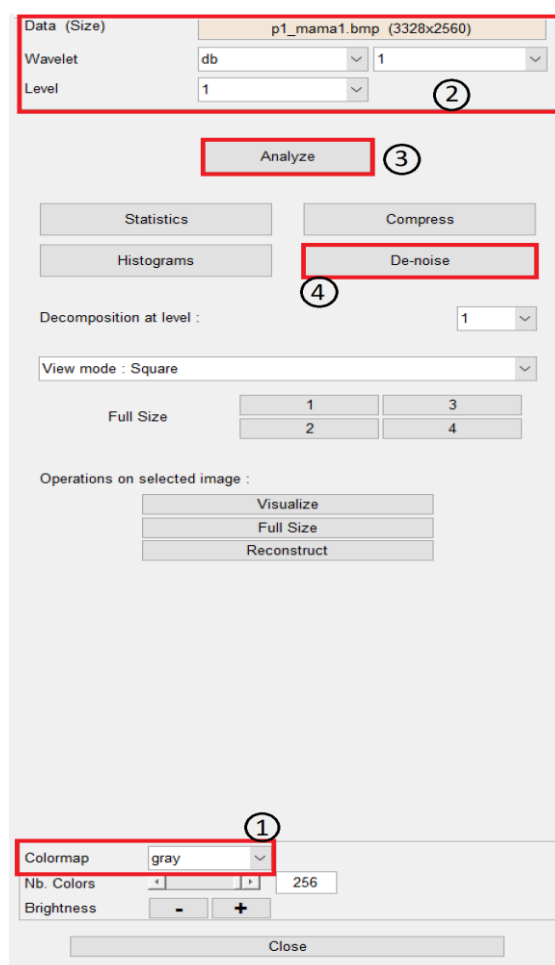


Figura 6: Tela de primeira análise e apresentação da imagem decomposta.

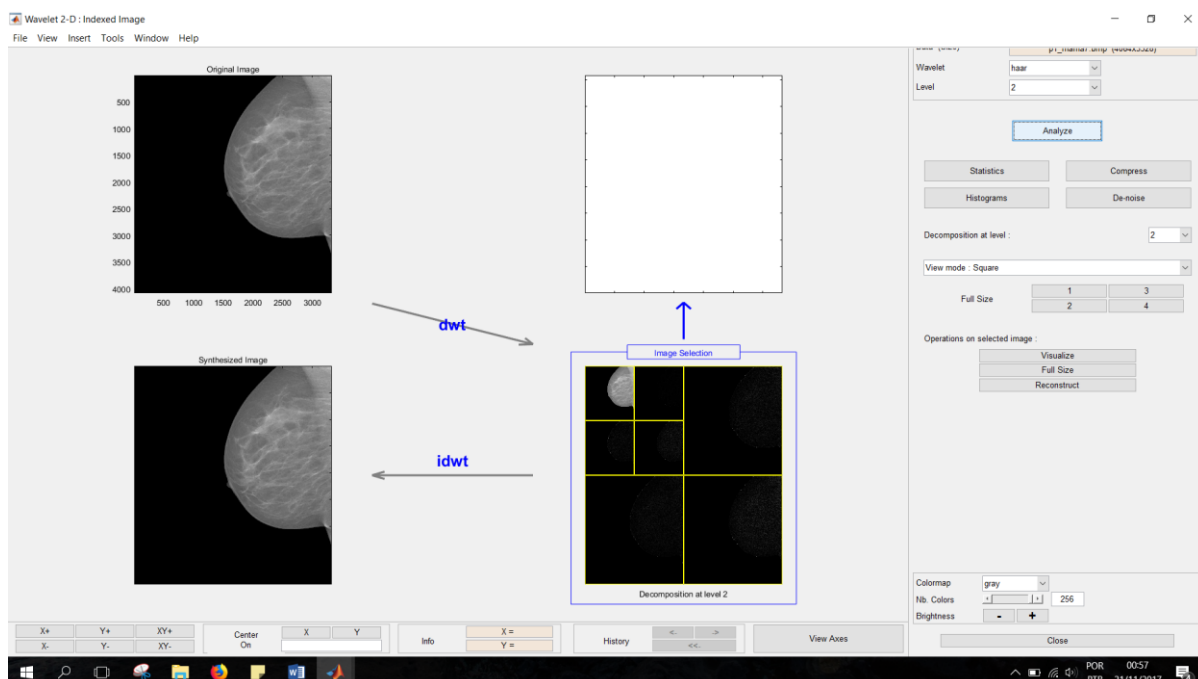


Figura 7: Etapa de *de-noise* da imagem escolhida.

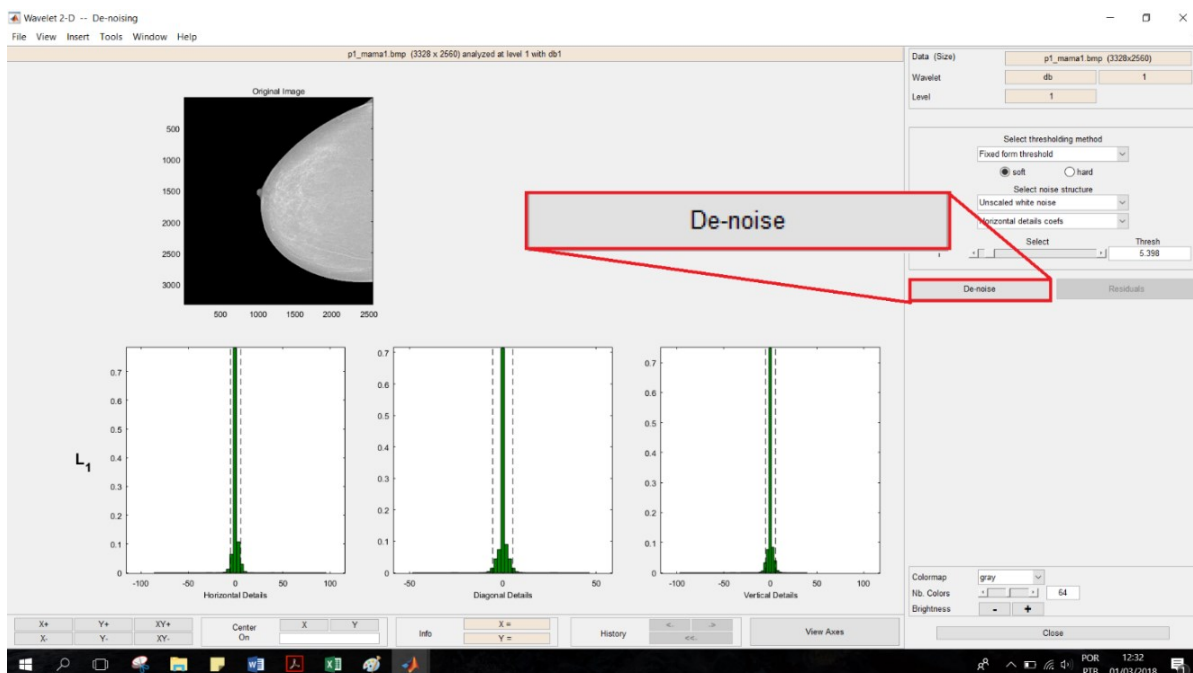
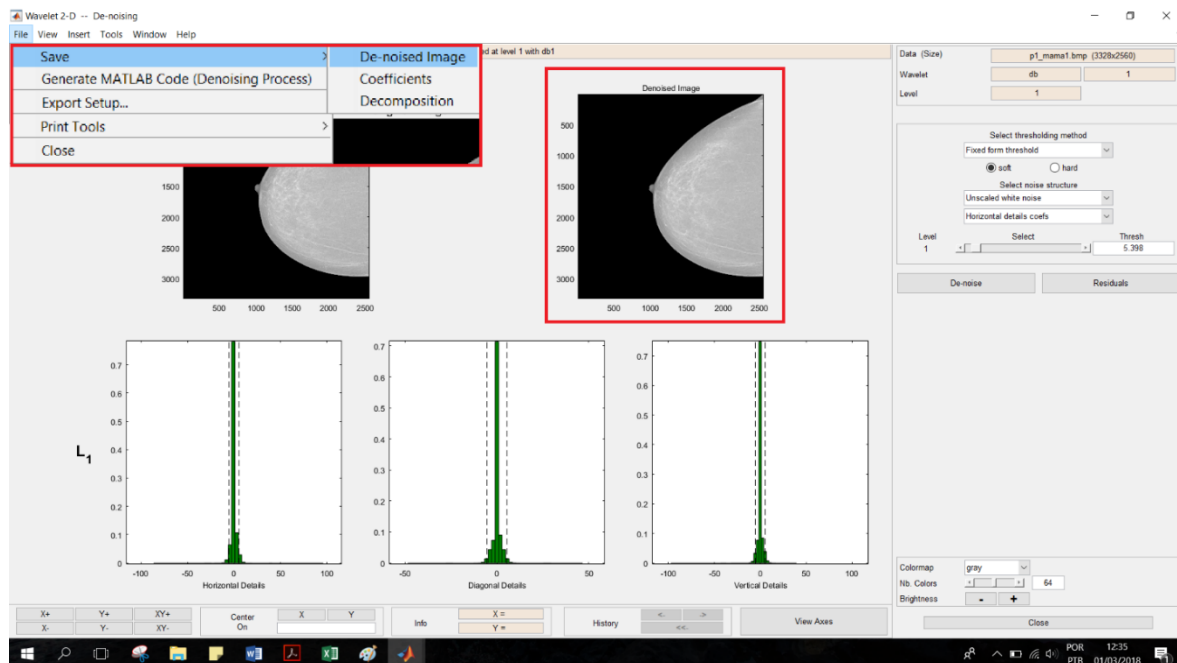


Figura 8: Apresentação da imagem processada e processo para salvá-la no computador.



Para encontrar o ruído é necessário subtrair a imagem original da imagem processada. Com esses três parâmetros é possível utilizar as funções *snr* e *psnr* do MatLab™ que retornam, respectivamente, valores de relação sinal-ruído (Equação 4) e relação sinal-ruído de pico (Equação 5). Esses quantificadores foram escolhidos

para determinar quais famílias possuem melhor desempenho no processo de *denoising* de imagens mamográficas.

$$SNR = \left(\frac{A_{\text{signal}}}{A_{\text{ruído}}} \right)^2 = \left(\frac{\sqrt{\frac{1}{n} \sum_{k=1}^n s_k^2}}{\sqrt{\frac{1}{n} \sum_{k=1}^n r_k^2}} \right)^2 \quad (4)$$

Na Equação 4, A é o valor quadrático médio da amplitude, s é o vetor que representa o sinal original, r representa o vetor de ruído e n o comprimento do vetor do sinal.

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (5)$$

Na Equação 5 MAX indica o valor máximo de pixel possível em uma imagem, MSE o erro quadrático médio (Equação 6). Abaixo, M e N indicam o número de linhas e colunas.

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} ||I(i, j) - K(i, j)||^2 \quad 6 \quad (6)$$

Utilizando o processo descrito acima, foram realizados 26 testes na mama 1 de cada padrão (Tabela 2). Os resultados foram coloridos de forma que os menores valores foram destacados em vermelho e os maiores em verde, os intermediários variam entre essas duas cores.

Tabela 2: Resultados do primeiro teste para quatro imagens mamográficas.

Famílias	Mama 1 - Padrão 1		Mama 1 - Padrão 2		Mama 1 - Padrão 3		Mama 1 - Padrão 4	
	SNR	PSNR	SNR	PSNR	SNR	PSNR	SNR	PSNR
coif1 - 1	16,6644	43,79	18,9945	47,4286	17,7823	44,9037	20,4717	51,8755
coif1 - 2	15,5674	42,5177	17,6687	45,8698	16,6038	43,5309	18,5319	49,6243
coif1 - 3	17,4162	42,6112	20,0309	46,1178	18,6773	43,6845	21,0755	49,8833
coif2 - 1	16,7694	43,9163	19,1158	47,5733	17,9054	45,0522	20,8504	52,3207
coif2 - 2	15,6556	42,6145	17,7887	46,0069	16,7142	43,656	18,7671	49,8916

coif2 - 3	15,352	42,2567	17,3236	45,4593	16,335	43,2128	18,1132	49,1355
coif3 - 1	16,801	43,9559	19,1432	47,6082	17,9383	45,0947	20,9113	52,4016
coif3 - 2	15,6803	42,6411	17,8174	46,0384	16,7411	43,6857	18,796	49,9319
coif3 - 3	15,3717	42,278	17,3476	45,4857	16,3604	43,2411	18,1435	49,168
db1 - 1	16,5717	43,5703	18,8388	47,0637	17,7002	44,6634	19,9547	51,1199
db1 - 2	15,3393	42,2464	17,217	45,3208	16,2805	43,1389	17,7743	48,7339
db1 - 3	15,0359	41,9142	16,761	44,8248	15,9042	42,7273	17,2231	48,1249
db2 - 1	16,7138	43,8696	18,8803	47,298	17,8451	44,9958	20,3884	51,8032
db2 - 2	15,6051	42,5618	17,6231	45,8144	16,642	43,5762	18,4715	49,5683
db2 - 3	15,2952	42,2039	17,1626	45,2783	16,259	43,1339	17,8711	48,8717
db3 - 1	16,7846	43,9481	18,9831	47,4173	17,9209	45,0819	20,6171	52,0672
db3 - 2	15,6557	42,6199	17,7198	45,9232	16,7041	43,6457	18,6615	49,7749
db3 - 3	15,3478	42,2609	17,2558	45,3816	16,3223	43,2036	18,0355	49,0397
sym2 - 1	16,7138	43,8696	18,8803	47,298	17,8451	44,9958	20,3884	51,8032
sym2 - 2	15,6051	42,5618	17,6231	45,8144	16,642	43,5762	18,4715	49,5683
sym2 - 3	17,4545	42,6563	19,9618	46,0534	18,7171	43,73	20,9895	49,8243
sym3 - 1	16,7846	43,9481	18,9831	47,4173	17,9209	45,0819	20,6171	52,0672
sym3 - 2	15,6557	42,6199	17,7198	45,9232	16,7041	43,6457	18,6615	49,7749
sym3 - 3	15,3478	42,2609	17,2558	45,3816	16,3223	43,2036	18,0355	49,0397
haar1	16,5717	43,5703	18,8388	47,0637	17,7002	44,6634	19,9547	51,1199
haar2	15,3393	42,2464	17,217	45,3208	16,2805	43,1389	17,7743	48,7339

Observa-se uma semelhança nos resultados que independe do padrão de mama, algumas famílias de *wavelet* obtiveram um desempenho igualmente ruim para todos. Por exemplo, Daubechies 1 nível 3, obteve o pior resultado para todos. A partir dos resultados obtidos, foram determinadas quais as famílias e seus respectivos níveis a serem utilizadas no processamento das demais imagens (Tabela 3). Escolheu-se as oito famílias com melhores resultados, visto que db2 nível 1 e sym2 nível 1 são equivalentes assim como db3 nível 1 e sym3 nível 1.

Com as oito famílias *wavelets* e seus respectivos níveis escolhidas no primeiro teste, foi realizado o processamento das demais imagens. As imagens mamográficas estavam agrupadas por tipo de padrão mamário: mamas predominantemente adiposas (Padrão 1), mamas parcialmente gordurosas (Padrão 2), mamas com padrão denso e heterogêneo (Padrão 3) e mamas muito densas (Padrão 4).

Tabela 3: Famílias de *wavelets* e respectivos níveis escolhidos para *denoising*.

<i>Wavelet</i>	Nível
Coiflets 1	1
	3
Coiflets 2	1
Coiflets 3	1
Daubechies 1	1
Daubechies 3	1
Symlets 2	1
	3

5 RESULTADOS E DISCUSSÕES

Exemplos dos resultados dos processamentos das imagens seguem nas Figuras de 9 a 12. Foram escolhidos os casos com o maior valor de desvio padrão de SNR e PSNR entre famílias *wavelet*, coincidentemente nessas imagens o melhor desempenho foi da *wavelet* coiflet 3 nível 1 (Letra e das Figuras).

Figura 9: Comparativo da imagem original com as processadas usando *wavelets*. a) Imagem mamográfica original do padrão 1; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.

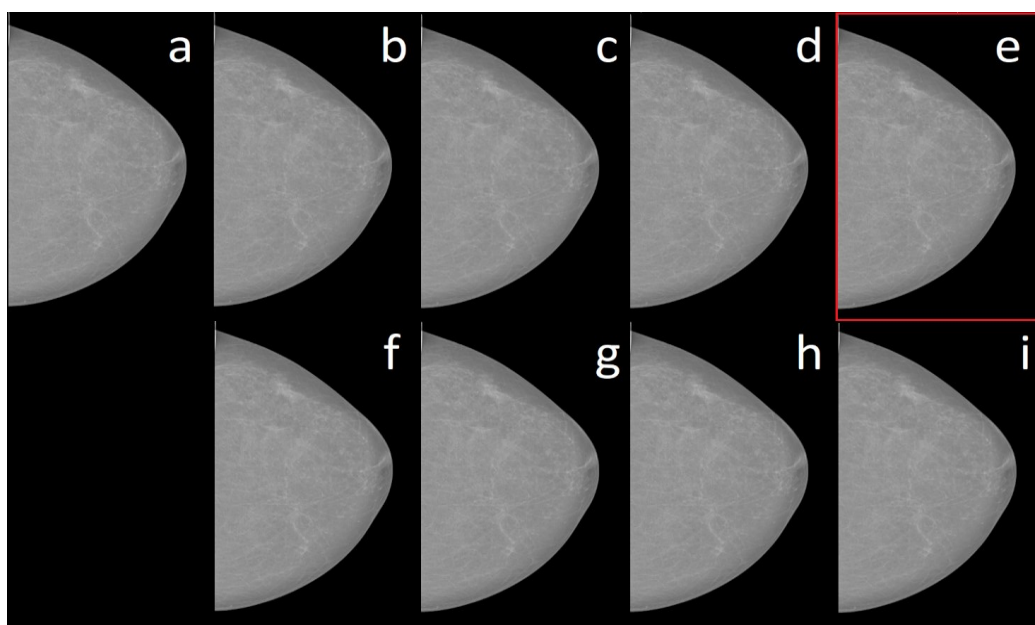


Figura 10: Comparativo da imagem original com as processadas usando *wavelets*. a) Imagem mamográfica original do padrão 2; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.

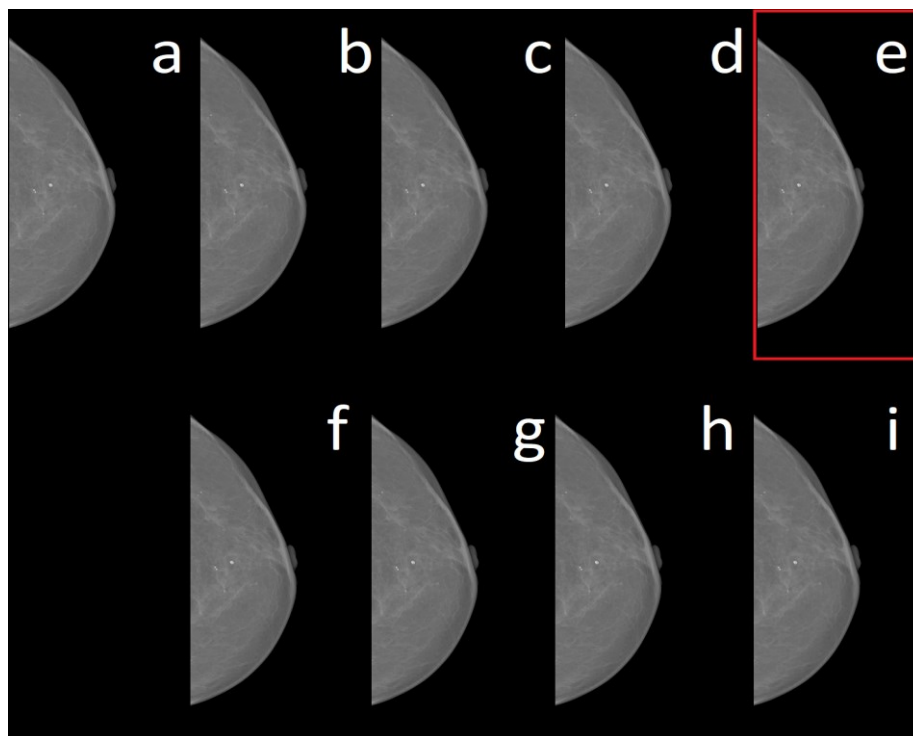


Figura 11: Comparativo da imagem original com as processadas usando *wavelets*. a) Imagem mamográfica original do padrão 3; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.

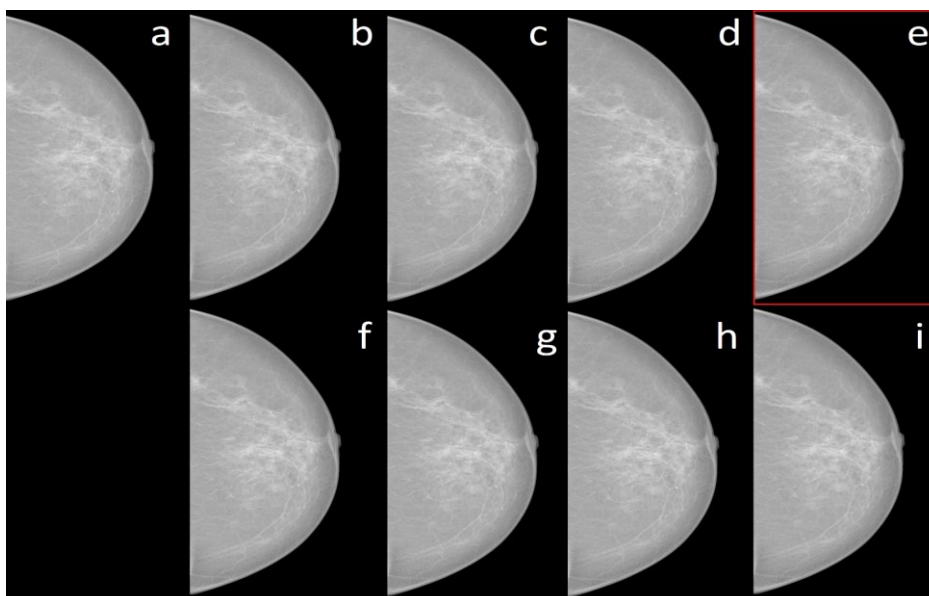
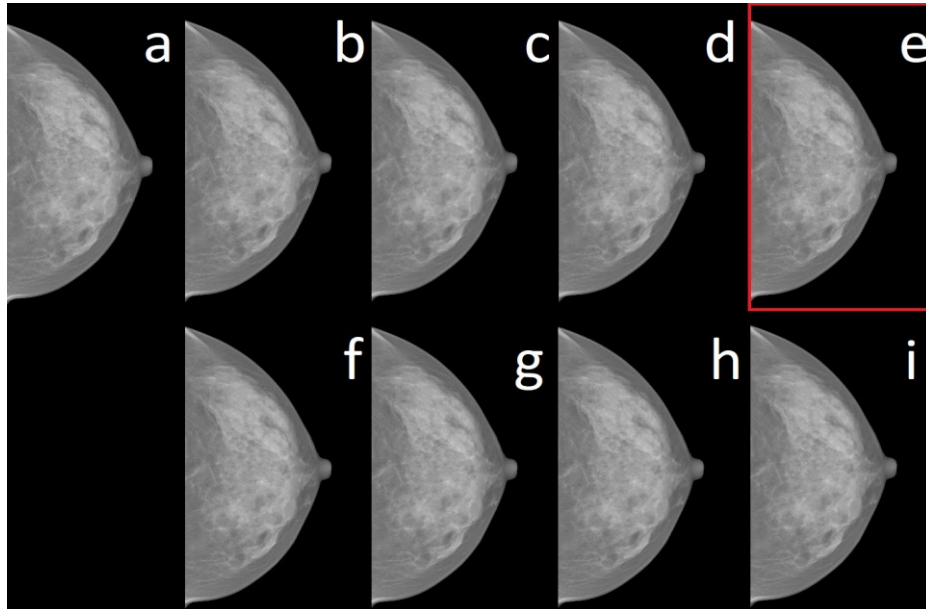


Figura 12: Comparativo da imagem original com as processadas usando *wavelets*. a) Imagem mamográfica original do padrão 4; b) Coif1 nível 1; c) Coif1 nível 3; d) Coif2 nível 1; e) Coif3 nível 1; f) Db1 nível 1; g) Db3 nível 1; h) Sym2 nível 1; i) Sym2 nível 3.



Visualmente os resultados podem não ser muito perceptíveis, por isso realizou-se a análise quantitativa usando a relação sinal-ruído (SNR) e pico da relação sinal-ruído (PSNR) como medidas. O SNR é resultado da razão entre a imagem final (pós-processamento) e a imagem original (ruidosa), quanto maior essa relação, menor é o efeito do ruído na imagem. O PSNR, relaciona a máxima energia de um sinal e o ruído que afeta, sua análise de resultado é semelhante ao SNR.

As tabelas seguintes apresentam as médias de SNR e PSNR para cada padrão de mama. Para o padrão 1 (mamas predominantemente adiposas), nota-se um melhor desempenho na medida SNR da família *coiflet1* nível 1 e PSNR da *coiflet3* nível 1. O pior resultado nas duas medidas foi a wavelet *coiflet1* nível 3 (Tabela 4).

Tabela 4: Média SNR e PSNR de famílias *wavelet* para as mamas de padrão 1.

Padrão 1	Média SNR	Média PSNR
coif1 - 1	18,89399	46,01139
coif1 - 3	17,06392	43,91108
coif2 - 1	18,77505	46,12845
coif3 - 1	18,81817	46,18209
db1 - 1	18,80021	45,72735
db3 - 1	18,68112	46,00221

sym2 - 1	18,90762	46,03633
sym2 - 3	17,06931	43,91838

Para os padrões de mama 2 (parcialmente gordurosas), 3 (denso e heterogêneo) e 4 (densas) tem-se um melhor resultado SNR e PSNR para coiflet 3 nível 1 e pior desempenho da wavelet coiflet 1 nível 3 (Tabelas 5, 6 e 7).

Tabela 5: Média SNR e PSNR de famílias *wavelet* para as mamas de padrão 2.

Padrão 2	Média SNR	Média PSNR
coif1 - 1	18,59815	46,97271
coif1 - 3	17,15055	45,02968
coif2 - 1	18,73336	47,13465
coif3 - 1	18,7623	47,17132
db1 - 1	18,46592	46,65912
db3 - 1	18,62241	46,99317
sym2 - 1	18,54444	46,91884
sym2 - 3	17,12684	45,00639

Tabela 6: Média SNR e PSNR de famílias *wavelet* para as mamas de padrão 3.

Padrão 3	Média SNR	Média PSNR
coif1 - 1	17,69584	45,48314
coif1 - 3	16,25687	43,57867
coif2 - 1	17,86548	45,6868
coif3 - 1	17,91112	45,74245
db1 - 1	17,46119	45,07584
db3 - 1	17,78128	45,57654
sym2 - 1	17,69484	45,4908
sym2 - 3	16,25714	43,57944

Tabela 7: Média SNR e PSNR de famílias *wavelet* para as mamas de padrão 4.

Padrão 4	Média SNR	Média PSNR
coif1 - 1	17,29954	46,97666
coif1 - 3	16,05512	45,26314
coif2 - 1	17,42049	47,11862
coif3 - 1	17,44813	47,15215

db1 - 1	17,1223	46,65531
db3 - 1	17,32286	46,98848
sym2 - 1	17,25993	46,9345
sym2 - 3	16,25712	45,41793

Os melhores resultados SNR e PSNR foram organizados na Tabela 8 separados por padrão de mama. Houve a necessidade de apresentá-los separadamente porque, em alguns casos, os melhores resultados não eram da mesma *wavelet*.

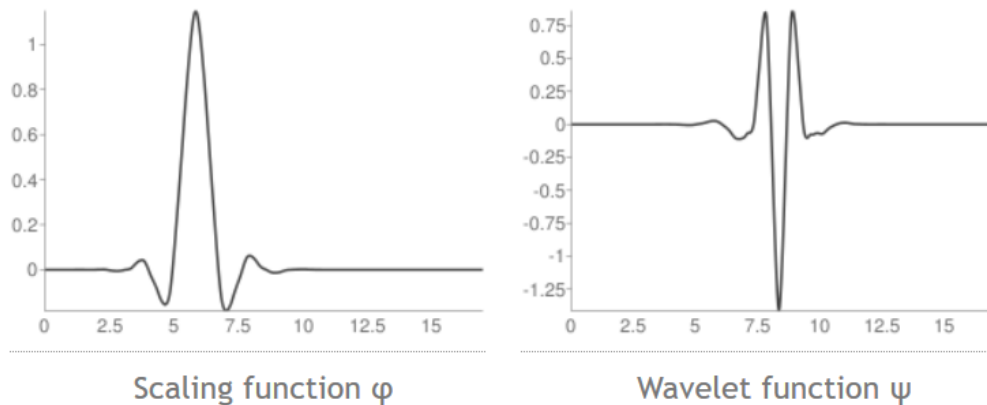
Tabela 8: Melhores resultados SNR e PSNR para cada família *wavelet* por padrão de mama.

Wavelet	Padrão 1		Padrão 2		Padrão 3		Padrão 4		Total	
	SNR	PSNR	SNR	PSNR	SNR	PSNR	SNR	PSNR	SNR	PSNR
Coif1 - 3	-	-	1	-	-	-	1	-	2	-
Coif2 - 1	-	-	-	1	-	-	-	-	-	1
Coif3 - 1	7	8	9	9	9	10	8	10	33	37
Sym2 - 1	1	1	-	-	-	-	-	-	1	1
Sym 2 - 3	1	-	-	-	1	-	1	-	3	-
Db3 - 1	1	1	-	-	-	-	-	-	1	1

É possível aferir que, independentemente do padrão de mama, os melhores resultados foram obtidos pela *wavelet* coiflet 3 no nível 1: melhor SNR em 82,5% das imagens e melhor PSNR em 92,5%.

As *wavelets* Coiflet foram desenvolvidas por I. Daubechies (apresentadas na Figura 13) e, diferentemente de suas outras *wavelets*, foi construída com momentos de desaparecimento não apenas para a função wavelet, mas também para a função de escala. Neste modo de design, a função de escalonamento do Coiflet exibe características de interpolação, podendo aproximar funções polinomiais em diferentes resoluções, aumentando assim a eficiência computacional da análise. Possibilitam ótimos resultados de processamento, análise e síntese também devido à sua regularidade (se comparada às outras famílias).

Figura 13: Funções wavelet e escalada de coiflet 3.



6 CONCLUSÕES

O principal problema do GUI *Wavelet Toolbox* é não aceitar imagens em padrão DICOM, padrão utilizado para imagens médicas, necessitando assim converter o banco de imagens para outro formato (neste caso *bitmap*). Porém a ferramenta possui uma interface intuitiva, apresenta várias opções para usuário e vem sendo amplamente aplicada em trabalhos de processamento de imagens.

Como visto em trabalhos sobre esse campo, pesquisa-se também sobre qual melhor valor de *threshold* para cada família *wavelet* e formas de estimá-lo, porém concluem que o mais importante é o nível de decomposição. Portanto, neste trabalho a configuração do *threshold* foi *default* para todo o processamento.

Conseguiu-se realizar a análise do desempenho de oito configurações de famílias e níveis *wavelets* no processo de *denoising*, de dez mamografias de cada padrão de densidade, e obteve-se um resultado significativo para a *wavelets* coiflet 3 nível de decomposição 1, o que reafirma o que foi visto na literatura (SILVA, 2015; ERGHEN, 2012; GUPTA, CHOUBEY, 2015), onde são destacados os bons resultados dados pelas famílias daubechies e coiflets.

REFERÊNCIAS

(AL JUMAH, 2013)

AL JUMAH, A. **Denoising of an image using discrete stationary wavelet transform and various thresholding techniques**. Journal of Signal and Information Processing, v. 4, n. 01, p. 33, 2013.

(BHATNAGAR, 2016)

BHATNAGAR, S.; JAIN, R. C. **Application of Discrete Multi-wavelet Transform in Denoising of Mammographic Images**. Indian Journal of Science and Technology, v. 9, n. 48, 2016.

(CHEN, 2003)

CHEN, G.; BUI, T. **Multiwavelets denoising using neighboring coefficients**. Signal Processing Letters, IEEE 10(7): 211-214, 2003.

(ERGHEN, 2012)

ERGEN, B. **Signal and image denoising using wavelet transform**. In: Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology. InTech, 2012.

(GONZALEZ, 2010)

GONZALEZ, R. C.; WOODS R. E. **Processamento digital de imagens**. 3ª Ed. Pearson Prentice Hall, São Paulo, 2010.

(GUPTA, 2015)

GUPTA, D.; CHOUBEY, S. **Discrete wavelet transform for image processing**. International Journal of Emerging Technology and Advanced Engineering, v. 4, n. 3, p. 598-602, 2015.

(INCA, 2015)

INSTITUTO NACIONAL DE CÂNCER (Brasil). **Estimativa 2016. Incidência do Câncer no Brasil**. Rio de Janeiro: INCA, 2015.

(JAISWAL, UPADHYAY, SOMKUWAR, 2014)

JAISWAL, A.; UPADHYAY, J.; SOMKUWAR, A. **Image denoising and quality measurements by using filtering and wavelet based techniques**. AEU-International Journal of Electronics and Communications, v. 68, n. 8, p. 699-705, 2014.

(MAYO, 2004)

MAYO, P.; RODENAS, F.; VERDU, G. **Comparing methods to denoise mammographic images**. In: Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE. IEEE, 2004. p. 247-250.

(MOREIRA, 2012)

MOREIRA, I. C. et al. INbreast: toward a full-field digital mammographic database. Academic Radiology, v. 19, n. 2, p. 236-248, 2012.

(PISANO, 2005)

PISANO, E. D. et al. **Diagnostic performance of digital versus film mammography for breast-cancer screening**. New England Journal of Medicine, v. 353, n. 17, p. 1773-1783, 2005.

(POULOS, 2003)

POULOS, A. et al. **Breast compression in mammography: how much is enough?**. Australasian radiology, v. 47, n. 2, p. 121-126, 2003.

(SILVA, 2007)

SILVA, C. Y. V. W. **Extração de características de imagens médicas utilizando wavelets para mineração de imagens e auxílio ao diagnóstico**. Tese de Doutorado. Universidade de São Paulo. 2007.

(WOLBARST, 1993)

WOLBARST, A. B. **Physics of radiology**. McGraw-Hill/Appleton & Lange, 1993.