



Universidade Federal de Uberlândia
Faculdade de Engenharia Elétrica
Programa de Pós-Graduação em Engenharia Biomédica

BRUNO AUGUSTO MACHADO GOES

Avaliação da segmentação de microcalcificações em mamografias com operações morfológicas.

Uberlândia

2024

BRUNO AUGUSTO MACHADO GOES

Avaliação da segmentação de microcalcificações em mamografias com operações morfológicas.

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Biomédica da Universidade Federal de Uberlândia como parte dos requisitos para obtenção do título de Mestre em Ciências

Orientadora: Prof.^a Dra. Ana Cláudia Patrocínio

Uberlândia

2024

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

G598
2024

Goes, Bruno Augusto Machado, 1998-
Avaliação da segmentação de microcalcificações em
mamografias com operações morfológicas [recurso
eletrônico] / Bruno Augusto Machado Goes. - 2024.

Orientadora: Ana Claudia Patrocinio.
Dissertação (Mestrado) - Universidade Federal de
Uberlândia, Pós-graduação em Engenharia Biomédica.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2024.635>
Inclui bibliografia.
Inclui ilustrações.

1. Engenharia biomédica. I. Patrocinio, Ana Claudia,
1976-, (Orient.). II. Universidade Federal de
Uberlândia. Pós-graduação em Engenharia Biomédica. III.
Título.

CDU: 62:61

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Biomédica				
Defesa de:	Dissertação de Mestrado Acadêmico, 111, PPGEB				
Data:	vinte e nove de agosto de dois mil e vinte e quatro	Hora de início:	09:00	Hora de encerramento:	11:20
Matrícula do Discente:	12122EBI001				
Nome do Discente:	Bruno Augusto Machado Goes				
Título do Trabalho:	Avaliação da segmentação de microcalcificações em mamografias com operações morfológicas.				
Área de concentração:	Engenharia Biomédica				
Linha de pesquisa:	Tecnologias em Radiações, Imagens Médicas e Biológicas				
Projeto de Pesquisa de vinculação:	Fatores que influenciam na detecção precoce do Câncer de Mama				

Reuniu-se no Anfiteatro 1E, Campus Santa Mônica, da Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Biomédica, assim composta: Professores Doutores: Diego Merigue da Cunha - INFIS/UFU; Homero Schiabel -EESC USP; Ana Claudia Patrocinio - PPGEB/UFU orientador(a) do candidato.

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Ana Claudia Patrocinio, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

[A]provado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Ana Claudia Patrocínio, Professor(a) do Magistério Superior**, em 29/08/2024, às 11:16, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Homero Schiabel, Usuário Externo**, em 29/08/2024, às 11:16, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Diego Merigue da Cunha, Professor(a) do Magistério Superior**, em 29/08/2024, às 11:16, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5653095** e o código CRC **213E244A**.

RESUMO

A mamografia é o exame radiográfico recomendado para detecção precoce do câncer de mama, o segundo tipo de câncer mais diagnosticado no mundo, pois permite a identificação de achados mamográficos, como nódulos e microcalcificações, em estágios iniciais da doença. As microcalcificações são pequenos depósitos de cálcio que podem estar associados ao câncer de mama, sendo importante a detecção dessas estruturas para avaliação correta da mamografia. Nesse sentido, pesquisas em processamento digital de imagens mamográficas procuram desenvolver técnicas de detecção e segmentação de microcalcificações para facilitar a identificação desses objetos na imagem. Este trabalho busca realizar a segmentação de microcalcificações utilizando a segmentação por *watershed* com marcadores definidos através das operações morfológicas *h-maxima* e *top-hat*. Para isso, foram realizados testes com diferentes tamanhos de elemento estruturante, valores de diferença mínima de intensidade entre pixels, métodos de limiarização e processamento de redução de ruído, como o filtro de Wiener e a transformada *wavelet*. Além disso, foram desenvolvidos dois modelos de redes neurais artificiais para classificação das regiões dos tipos verdadeiro e falso positivo obtidas na segmentação, com o objetivo de reduzir o número de regiões do tipo falso positivo. A segmentação por *watershed* usando os marcadores definidos com *h-maxima* e *top-hat* foi bem-sucedida na identificação de regiões com microcalcificações, resultando em valores altos para a métrica sensibilidade, porém, em baixos valores para especificidade e acurácia, devido a segmentação de muitos falsos positivos. Os resultados da classificação indicam que os modelos produzidos foram efetivos na redução do número de falso positivos, com aumento das métricas acurácia e especificidade, porém, com redução da sensibilidade do processamento.

Palavras-chave: Câncer de mama. Mamografia. Microcalcificação. Segmentação. *Watershed*.

ABSTRACT

Mammography is an x-ray imaging method recommended for breast cancer early detection, the second most diagnosed cancer type worldwide, because allows to visualize suspicious findings, such as masses and microcalcifications, in early stages of disease. Microcalcifications are small calcium deposits that could be associate to breast cancer, being important the detection of these structures to correct image evaluation. In this scenario, research in digital mammography image processing aims to develop methods to detect and segmentate microcalcifications to facilitate the searching for these objects in image. This work intents to perform microcalcification segmentation using watershed segmentation with markers defined by the morphological operators *h-maxima* and *top-hat*. To achieve this, tests were carried out with different structuring element size, minimum intensity difference between pixels, thresholding and denoising techniques, such as the Wiener filter and wavelet transform. Besides that, two artificial neural networks were developed for classification of true and false positive regions, aiming to reduce the number of false positive regions. The marker-controlled watershed segmentation using markers from *h-maxima* and *top-hat* operations was well succeeded in identifying regions with microcalcifications, resulting in high sensitivity value, but with low specificity and accuracy values, due to high false positive numbers. The results of the classification step indicate that the developed models were effective to reduce the number of false positive regions, with increase in accuracy and specificity, but decrease in sensitivity.

Keywords: Breast cancer. Mammography. Microcalcification. Segmentation. Watershed

LISTA DE FIGURAS

Figura 1 - Microcalcificações em imagens mamográficas.	18
Figura 2 - Microcalcificações em visão topográfica.	29
Figura 3 - Configuração de uma rede neural artificial multicamadas.	34
Figura 4 - Padronização do <i>ground truth</i> . (a) Imagem original; (b) <i>Ground truth</i> original; (c) <i>Ground truth</i> produzido; (d) Imagem original; (e) <i>Ground truth</i> original; (f) <i>Ground truth</i> produzido.	41
Figura 5- Etapas desenvolvidas para o processamento das imagens.	42
Figura 6 - Resultados do pré-processamento. (a) Imagem original; (b) Filtro de Wiener; (c) Transformada <i>wavelet</i> ; (d) Imagem original; (e) Filtro de Wiener; (f) Transformada <i>wavelet</i>	50
Figura 7 – Resultados dos testes de detecção com <i>h-maxima</i> e máximo de 1000 objetos detectados: (a) SE=3, h=0,02; (b) SE=3, h=0,04; (c) SE=3, h=0,08; (d) SE=11, h=0,02; (e) SE=11, h=0,04; e (f) SE=11, h=0,08.	53
Figura 8 - Resultados dos testes de detecção com <i>h-maxima</i> com máximo de 1000 objetos detectados: (a) SE=3, h=0,02; (b) SE=3, h=0,04; (c) SE=3, h=0,08; (d) SE=11, h=0,02; (e) SE=11, h=0,04; e (f) SE=11, h=0,08.	53
Figura 9 - Resultados dos testes de detecção com <i>h-maxima</i> variando o número máximo objetos detectados: (a) 500 objetos; (b) 750 objetos; (c) 1000 objetos; (d) 500 objetos; (e) 750 objetos; (f) 1000 objetos.	54
Figura 10 - Resultados dos testes de detecção com <i>top-hat</i> : (a) SE=3, 500 objetos; (b) SE=3, 750 objetos; (c) SE=3, 1000 objetos; (d) SE=11, 500 objetos; (e) SE=11, 750 objetos; (f) SE=11, 1000 objetos.	56
Figura 11 - Resultados dos testes de detecção com <i>top-hat</i> : (a) SE=3, 500 objetos; (b) SE=3, 750 objetos; (c) SE=3, 1000 objetos; (d) SE=11, 500 objetos; (e) SE=11, 750 objetos; (f) SE=11, 1000 objetos.	56
Figura 12 – Combinação dos marcadores internos e externos na detecção com <i>h-maxima</i> : (a) Filtro de Wiener; (b) Marcador interno; (c) Marcador final; (d) Filtro de Wiener; (e) Marcador interno; (f) Marcador final,	58
Figura 13 - Combinação dos marcadores internos e externos na detecção com <i>top-hat</i> : (a) Imagem original; (b) Marcador interno; (c) Marcador final; (d) Imagem original; (e) Marcador interno; (f) Marcador final,	58

Figura 14 - Resultados da segmentação: (a) Transformada *wavelet*; (b) Detecção por *h-maxima*; (c) Detecção por *top-hat*; (d) Filtro de Wiener; (e) Detecção por *h-maxima*; (f) Detecção por *top-hat*. 63

Figura 15 - Resultados da segmentação: (a) Transformada *wavelet*; (b) Detecção por *h-maxima*; (c) Detecção por *top-hat*; (d) Filtro de Wiener; (e) Detecção por *h-maxima*; (f) Detecção por *top-hat*. 64

Figura 16 - Distribuição do módulo dos valores de correlação entre as características extraídas e a saída esperada: (a) Dados da detecção com *h-maxima*; (b) Dados da detecção com *top-hat*. 66

Figura 17 - Comparação entre os resultados da segmentação e classificação: (a) Segmentação - *h-maxima* (b) Classificação - *h-maxima* (c) Segmentação - *top-hat* (d) Classificação - *top-hat*. 70

Figura 18 - Comparação entre os resultados da segmentação e classificação: (a) Segmentação *h-maxima* (b) Classificação *h-maxima* (c) Segmentação *top-hat* (d) Classificação *top-hat*. 71

LISTA DE TABELAS

Tabela 1 - Categorias BI-RADS® de acordo com achados mamográficos.....	18
Tabela 2 - Matriz de confusão.	31
Tabela 3 - Imagens selecionadas conforme a classificação BI-RADS®	41
Tabela 4 - Parâmetros de configuração para detecção dos marcadores.....	44
Tabela 5 – Características extraídas das regiões analisadas.	46
Tabela 6 - Divisão dos dados em conjuntos de treinamento e teste.	47
Tabela 7 - Parâmetros de configuração das redes neurais artificiais.	48
Tabela 8 - Média e desvio padrão do PSNR das imagens processadas.	50
Tabela 9 - Resultados da detecção de objetos com <i>h-maxima</i>	52
Tabela 10 - Resultados da detecção de objetos com <i>top-hat</i>	55
Tabela 11 - Média e desvio padrão das métricas de avaliação da segmentação das imagens com detecção por <i>h-maxima</i>	59
Tabela 12 - Média e desvio padrão das métricas de avaliação da segmentação das imagens com detecção por <i>top-hat</i>	61
Tabela 13 - Atributos selecionados e medidas calculadas nos conjuntos de treinamento.	66
Tabela 14 - Resultados das métricas calculadas para segmentação e classificação das imagens do grupo de teste.	68

SUMÁRIO

1. INTRODUÇÃO.....	13
1.1. Objetivos.....	14
2. MAMOGRAFIA	15
3. PROCESSAMENTO DIGITAL DE IMAGENS	20
3.1. Restauração de Imagens.....	21
3.2. Operadores Morfológicos	24
3.3. Segmentação	28
4. REDES NEURAIS	34
5. ESTADO DA ARTE	37
6. METODOLOGIA.....	40
6.1. Base de Imagens	40
6.2. Processamento.....	42
7. RESULTADOS E DISCUSSÃO	50
8. CONCLUSÃO.....	73
REFERÊNCIAS	75
APÊNDICE A – Resultados da segmentação.....	78
APÊNDICE B – Resultados da classificação	80

1. INTRODUÇÃO

O câncer de mama é o segundo tipo de câncer mais diagnosticado no mundo, sendo que em 2022 foram registradas cerca de 2.3 milhões de novos casos e 665 mil mortes (BRAY et al., 2024). No Brasil, as estimativas do Instituto Nacional do Câncer (INCA) indicam cerca 73.610 casos em 2023, o que corresponde a um risco estimado de 66,54 casos a cada 100 mil mulheres. Quanto a mortalidade, os dados mais recentes indicam 17.825 óbitos em 2020, o equivalente a 16,47 mortes a cada 100 mil mulheres brasileiras (INSTITUTO NACIONAL DE CÂNCER, 2022).

Considerando a importância do controle do câncer de mama para garantir o diagnóstico e o tratamento adequado para mulheres brasileiras, o Ministério da Saúde recomenda que mulheres entre 50 e 69 anos realizem mamografia de rastreamento a cada dois anos. A mamografia é um exame radiológico realizado em um mamógrafo e é capaz de detectar lesões nas mamas antes do surgimento de sintomas relacionados ao câncer, permitindo que o tratamento comece mais cedo e seja menos agressivo com a paciente. O rastreamento por mamografia é a principal ação utilizada para detecção precoce do câncer de mama, sendo essencial para avaliação diagnóstica de mulheres assintomáticas (INSTITUTO NACIONAL DE CÂNCER JOSÉ DE ALENCAR GOMES DA SILVA, 2015).

Um dos desafios no diagnóstico de câncer de mama é avaliar corretamente variedade de tipos, formas, tamanhos e densidades dos achados mamográficos, os quais podem aparecer como nódulos, calcificações, assimetrias e distorções. Uma das estruturas suspeitas mais comuns em mamografias são as microcalcificações, pequenos depósitos de cálcio menores que 1 mm com padrão variável. As microcalcificações podem representar um sinal precoce de câncer, sendo imprescindível a correta identificação dessas estruturas caso presentes em uma mamografia (DE PAREDES, 2007).

Apesar de possuir alta qualidade, a imagem da mamografia, assim como outras imagens radiográficas, possui diversas características indesejáveis causadas pela interação da radiação com o tecido mamário e que atrapalham a visualização da imagem produzida. O espalhamento dos fótons de raios-X causa ruído com características semelhantes as microcalcificações e dificultam a diferenciação das estruturas na imagem. Outro problema é o baixo contraste da imagem, característica resultante da baixa energia utilizada na mamografia e que limita a distribuição dos níveis de cinza na imagem (BUSHBERG, 2012).

Nesse sentido, o processamento digital das imagens pode ser uma ferramenta para amenizar as características indesejáveis nas mamografias e facilitar a identificação de estruturas

e o diagnóstico do câncer de mama. Um exemplo é a segmentação de objetos em imagens, um tipo de processamento capaz identificar e delimitar automaticamente as regiões em uma imagem e que pode ser aplicado para detecção e avaliação de achados mamográficos, inclusive microcalcificações (LAURITZEN et al., 2024; REZAEI, 2021; SAMALA et al., 2024).

1.1. Objetivos

Dada a importância da detecção precoce do câncer de mama e o desafio na identificação de microcalcificações em imagens mamográficas, este trabalho tem por objetivo avaliar a segmentação de microcalcificações em mamografias utilizando a técnica *watershed* associada a diferentes métodos de detecção de marcadores. Para isso, pretende-se:

- Examinar o impacto de métodos de redução de ruído na segmentação de microcalcificações;
- Avaliar a capacidade de detecção de objetos de interesse com operadores morfológicos;
- Desenvolver modelos de aprendizado de máquina com redes neurais para redução de falsos positivos.

2. MAMOGRAFIA

A mamografia é um exame radiográfico, ou seja, utiliza radiação e a diferença de atenuação entre os tecidos para formar a imagem. Para a realização do exame é necessário utilizar um mamógrafo, um equipamento que gera, direciona e detecta fótons de raios X que, ao entrarem em contato a mama, perdem energia conforme as diferentes características das estruturas do tecido mamário. Quando comparado com um aparelho de raios X convencional, o mamógrafo possui aspectos específicos que buscam permitir a diferenciação entre tecidos normais e cancerígenos, como por exemplo a configuração do tubo de raios X e seus filtros e a operação com baixa diferença de potencial no tubo (BUSHBERG, 2012).

A formação dos raios X se dá pela conversão da energia cinética dos elétrons em radiação eletromagnética quando os elétrons são desacelerados pela interação com um material. Em um tubo de raios X, a fonte dos elétrons é o cátodo, geralmente de tungstênio, que libera elétrons quando aquecido no processo de excitação térmica. Quando uma diferença de potencial entre 15 e 35 kV, faixa de tensão utilizada na mamografia, é aplicada no tubo de raios X, os elétrons são acelerados e ganham energia cinética máxima equivalente a tensão aplicada no tubo, sendo essa energia caracterizada pela unidade de medida Elétron-volt (eV) (BUSHBERG, 2012).

As partículas aceleradas pela diferença de potencial são direcionadas ao ânodo, o alvo, que produz raios X com o impacto dos elétrons. Na produção dos raios X, apenas cerca de 1% da energia cinética total é convertida em radiação, sendo o restante liberado como calor, devido as diferentes possibilidades de interação entre os elétrons acelerados e os átomos do alvo. Para garantir que os elétrons não percam a energia cinética durante o deslocamento, o tubo de raios X é submetido a vácuo. Além disso, para aprimorar a dissipação do calor produzido no processo e permitir uso de correntes e tempos de exposição maiores, é comum a utilização de ânodos rotativos (BUSHBERG, 2012).

A região do alvo na qual ocorre a produção dos raios X é o ponto focal, cujas características são importantes na qualidade do feixe e para evitar o borramento geométrico da imagem. Na mamografia, o ponto focal tem tamanho de 0,3 mm, porém, algumas aplicações, como a magnificação da imagem, utilizam ponto focal com tamanho 0,1 mm. Tamanhos menores para o ponto focal aumentam a resolução espacial, porém limitam a corrente máxima do tubo e, conseqüentemente, aumentam o tempo de exposição (BUSHBERG, 2012)

Considerando que a energia dos fótons de raios X produzidos depende da distância da interação do elétron com o núcleo do alvo, sendo que distâncias menores resultam em maiores

energias, o resultado da formação de raios X consiste em fótons com energia entre zero e a energia cinética máxima produzida, também chamados de radiação de freamento. Portanto, como são produzidos fótons com diferentes energias e alguns níveis são conhecidamente irrelevantes para a formação da imagem, pode-se selecionar os níveis de energia de um feixe através da filtragem dos fótons de raios X (BUSHBERG, 2012; OKUNO; YOSHIMURA, 2010).

Além da pouca contribuição na formação da imagem, os fótons de baixa energia também são facilmente absorvidos pelos tecidos mamários, ou seja, é possível evitar doses impróprias de radiação nos tecidos pela filtragem do feixe. Para isso filtragem dos níveis de energia dos fótons, o feixe passa por filtros constituídos de materiais capazes de absorver e selecionar os níveis energia nas faixas desejadas. No caso da mamografia, a filtragem é realizada pela janela de berílio e por filtros compostos de finas chapas de molibdênio, ródio ou prata, a depender da configuração do mamógrafo. (BUSHBERG, 2012).

No processo de produção de raios X, os principais parâmetros configurados pelo operador do equipamento são a tensão aplicada no tubo, dada em kV, e a corrente utilizada para aquecer o cátodo, caracterizada por mA. A tensão altera diretamente a energia dos fótons, de forma que quanto maior a tensão, maior a capacidade de penetração do feixe nos tecidos. Já a corrente modifica o número de elétrons produzidos no cátodo e, conseqüentemente, o número de fótons de raios X emitidos pelo ânodo. Quando a tensão aplicada no tubo é alterada, é necessário certa compensação no valor de corrente utilizado para manter a exposição desejada (BUSHBERG, 2012).

Durante um exame radiográfico, o feixe de raios X atravessa os tecidos humanos e é detectado em um detector capaz de converter a energia dos fótons em luz visível ou sinal elétrico. Essa interação entre os fótons de raios X e a matéria pode ocorrer de duas formas, sendo elas a absorção fotoelétrica e o espalhamento Compton. Na absorção fotoelétrica, a energia do fóton é absorvida, causando a liberação de um elétron da camada mais interna do átomo e atenuação do feixe, sendo esse o tipo de interação responsável pela formação da imagem. Já no espalhamento Compton, apenas parte da energia é absorvida na ejeção do elétron e a trajetória do fóton é alterada, diminuindo a qualidade da imagem pela adição de ruído e borramento (BUSHBERG, 2012).

Para atenuar os efeitos do espalhamento dos fótons na mamografia, pode-se utilizar uma grade antiespalhamento entre a mama e o detector. Essa grade é composta de chumbo combinado com alumínio ou carbono e permite que apenas fótons com trajetória paralela a

parede da grade atinjam o detector, impedindo que as partículas com trajetórias alteradas pelo efeito Compton aumentem o ruído na imagem (BUSHBERG, 2012).

Na mamografia, para aquisição de uma imagem de boa qualidade, é necessário realizar a compressão da mama através de manobras adequadas para uniformizar a exposição ao feixe de raios X. Esse procedimento permite reduzir a sobreposição dos tecidos mamários, evitando artefatos de movimento, diminuindo o espalhamento dos fótons e, conseqüentemente, reduzindo o ruído e o borramento da imagem (BUSHBERG, 2012).

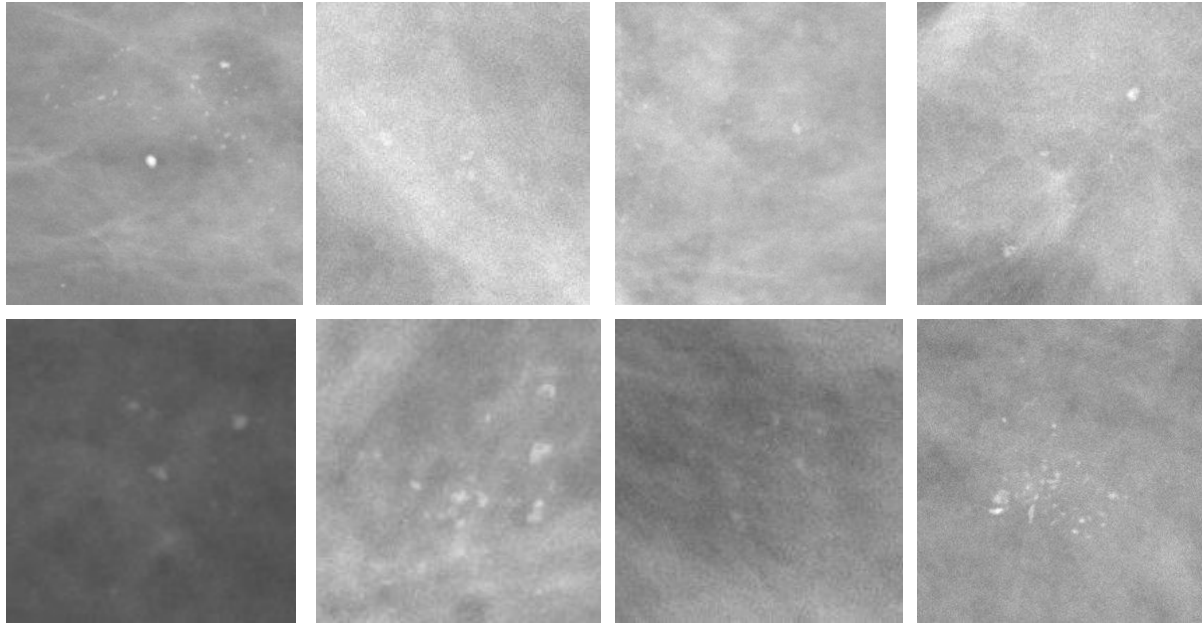
O exame completo é realizado pela aquisição duas imagens de cada mama, cada uma com um posicionamento do tubo. Na incidência médio lateral oblíqua (MLO) é possível visualizar o tecido mamário e a região axilar posicionando o tubo em uma angulação entre 30° e 60°. A segunda incidência, chamada craniocaudal (CC), é obtida quando o tubo está localizado perpendicularmente à mama e expõe as partes lateral e medial da mama, além do músculo grande peitoral (BUSHBERG, 2012).

Através da mamografia é possível detectar estruturas suspeitas que podem ser indicativos de câncer de mama ou não, entre elas nódulos, microcalcificações, assimetrias, distorções focais e dilatações ductais. Essas lesões apresentam muitas formas, tamanhos, contornos e densidades diferentes, características importantes para avaliação da malignidade dos achados. A detecção e classificação correta das estruturas encontradas em uma mamografia é importante para evitar resultados incorretos causadores de sobrediagnóstico e sobretratamento (DE PAREDES, 2007; INSTITUTO NACIONAL DE CÂNCER, 2007).

Um dos achados mais comuns em mamografias são as microcalcificações, objetos de difícil detecção e avaliação, principalmente pelas diferentes formas e tamanhos que podem apresentar. As microcalcificações possuem tamanho menor que 1 mm e podem assumir diversos formatos como arredondadas, puntiformes, irregulares e ramificadas. Além disso, essas estruturas podem estar distribuídas pela mama de forma agrupada, em *clusters* com mais de cinco microcalcificações em 1 cm², em segmento ocupando parte de um quadrante ou em um trajeto ductal (DE PAREDES, 2007; INSTITUTO NACIONAL DE CÂNCER, 2007; SICKLES et al., 2013).

Devido a composição das microcalcificações, geralmente fosfato de cálcio ou oxalato de cálcio, os fótons de raios X possuem grande probabilidade de atenuação quando em contato com esses objetos. Dessa maneira, as microcalcificações tendem a aparecer em mamografias como pequenos componentes conectados com alta intensidade cercados por regiões de menor intensidade (TOUIL et al., 2021). A Figura 1 traz exemplos de mamografias com diferentes densidades de mama contendo microcalcificações.

Figura 1 - Microcalcificações em imagens mamográficas.



Fonte: Elaborado pelo autor.

Os achados mamográficos podem ser classificados de acordo com a classificação BI-RADS® conforme a suspeita de malignidade e o risco de câncer. Conforme a categoria na qual o achado foi classificado, diferentes condutas deverão ser tomadas para o diagnóstico e tratamento correto do caso. Essas condutas incluem a biópsia dos achados, repetição do exame usando diferentes incidências ou após determinado período e exames complementares como ultrassonografia e a tomossíntese (INSTITUTO NACIONAL DE CÂNCER JOSÉ DE ALENCAR GOMES DA SILVA, 2021; SICKLES et al., 2013).

Tabela 1 - Categorias BI-RADS® de acordo com achados mamográficos.

Categoria BI-RADS®	Achados mamográficos	Risco de câncer
1 – Negativo	Sem achados	< 0,05%
2 – Benigno	Achados benignos	< 0,05%
3 - Provavelmente benigno	Achados provavelmente benignos	< 2%
4 – Suspeito (baixa, média e alta suspeição)	Achados suspeitos de malignidade	Entre 2% e 95%
5 – Altamente suspeito	Achados altamente suspeitos de malignidade	> 95%
6 – Achados já com diagnóstico de câncer	Diagnóstico de câncer comprovado histologicamente	100%
0 – Indefinido	Necessidade de avaliação adicional	---

Fonte: Adaptado de Instituto Nacional de Câncer José Alencar Gomes da Silva, 2021

Além disso, a classificação BI-RADS® distribui as mamografias em quatro categorias conforme a composição da mama, a qual passa por substituição do parênquima mamário por tecido adiposo com o avanço da idade. Desse modo, as mamas podem ser classificadas em:

densas (sem substituição adiposa), predominantemente densas, predominantemente adiposas e adiposas (composta apenas de tecido adiposo) (SICKLES et al., 2013).

3. PROCESSAMENTO DIGITAL DE IMAGENS

Devido ao processo de formação da imagem radiográfica, algumas características podem afetar a visualização de estruturas e, conseqüentemente, dificultar a interpretação do exame e a detecção de achados mamográficos. A presença de ruído e o baixo contraste entre os objetos na imagem são alguns dos problemas que podem impactar a qualidade da imagem e, inclusive, o desenvolvimento de algoritmos para processamento digital de imagens (GONZALEZ; WOODS, 2010).

Dito isso, é comum a utilização de métodos de filtragem e aumento de contraste para melhorar a qualidade da imagem adquirida, além de algoritmos de segmentação para auxílio na detecção e análise de objetos de interesse. Esses processamentos podem envolver a suavização de ruído através de filtros espaciais e no domínio da frequência, o aumento de contraste pela manipulação do histograma da imagem e a segmentação pela diferenciação das regiões da imagem segundo características dos pixels e seus vizinhos (GONZALEZ; WOODS, 2010; REZAEI, 2021).

No processamento digital de imagens, a imagem é representada como uma matriz com M linhas e N colunas na qual cada elemento corresponde a um pixel, um elemento de imagem caracterizado por coordenadas espaciais e um valor de intensidade ou nível de cinza. Considerando que cada elemento da matriz é definido por $f(x, y)$, sendo x e y as coordenadas espaciais, podemos representar a matriz $M \times N$ da imagem como (GONZALEZ; WOODS, 2010):

$$f(x, y) = \begin{bmatrix} f(0, 0) & \cdots & f(0, N - 1) \\ \vdots & \ddots & \vdots \\ f(M - 1, 0) & \cdots & f(M - 1, N - 1) \end{bmatrix} \quad (1)$$

O número de níveis de cinza presentes em uma imagem está diretamente relacionado ao número de *bits* necessários para armazenar cada pixel da imagem. Esse número é expresso no intervalo $[0, 2^n - 1]$, sendo n um número inteiro referente ao número de *bits*, também chamado de resolução de intensidade (GONZALEZ; WOODS, 2010).

As relações entre os pixels são importantes no processamento de imagens, principalmente em técnicas que levam em consideração as coordenadas de posicionamento dos pixels. Considerando p um pixel localizado em (x, y) , representamos a vizinhança-4 de p pelos vizinhos horizontais e verticais, ou seja, os pixels nas coordenadas $(x + 1, y)$, $(x - 1, y)$, $(x, y + 1)$ e $(x, y - 1)$. No mesmo sentido, a vizinhança-8 é composta dos pixels da vizinhança-4 e os vizinhos diagonais localizados nas coordenadas $(x + 1, y + 1)$, $(x + 1, y - 1)$, $(x - 1, y + 1)$ e $(x - 1, y - 1)$ (GONZALEZ; WOODS, 2010).

Além da vizinhança, outra relação entre pixels importante é o caminho, também chamado curva digital. Considerando os pixels p e q de coordenadas (x, y) e (s, t) , respectivamente, o caminho entre p e q é definido pela sequência $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ de pixels distintos que obedecem às seguintes restrições: *i)* $(x_0, y_0) = (x, y), (x_n, y_n) = (s, t)$; *ii)* Os pixels (x_i, y_i) e (x_{i-1}, y_{i-1}) são vizinhos em todo o caminho; e *iii)* se $(x_0, y_0) = (x_n, y_n)$, tem-se um caminho fechado. Um caminho pode ser do tipo caminho-4 ou caminho-8 se o tipo de vizinhança analisada for alterado (GONZALEZ; WOODS, 2010).

Com os conceitos de vizinhanças e caminhos, chamamos de região (R) um conjunto conexo, ou seja, um subconjunto de pixels conectados no qual para todos os pixels no conjunto existe um caminho conectando p e q . Se uma imagem possuir várias regiões, a união de todas as regiões pode ser expressa por R_u e seu complemento por $(R_u)^c$. Os pontos presentes em R_u são a frente (*foreground*) e os pontos em $(R_u)^c$ são o fundo da imagem (*background*) (GONZALEZ; WOODS, 2010).

Considerando uma região como um objeto de interesse, é possível extrair as fronteiras interna e externa das regiões, definidas como um conjunto de pixels com relações com o objeto e o fundo da imagem. A fronteira interna de uma região, também chamada de contorno interno, corresponde ao conjunto de pixels da região que possuem pelo menos um vizinho no fundo da imagem. Já a fronteira externa, inclui os pixels do fundo com pelo menos um vizinho no objeto analisado (GONZALEZ; WOODS, 2010).

3.1. Restauração de Imagens

A restauração de imagens digitais corresponde ao uso de técnicas para estimar a função de degradação da imagem e recuperar a imagem original. Esses processamentos podem ser realizados nos domínios espacial e da frequência e tem por objetivo o *denoising* da imagem, ou seja, reduzir o ruído da imagem original (GONZALEZ; WOODS, 2010).

O filtro de mínimo erro quadrático médio, também conhecido por filtro de Wiener, é uma das técnicas utilizadas para suavização de ruído e restauração de imagens. O objetivo é encontrar \hat{f} da imagem não corrompida f minimizando o erro quadrático médio entre as variáveis para que a relação sinal ruído da imagem seja maximizada. A medida desse erro é calculada pela Equação (2) (GONZALEZ; WOODS, 2010):

$$e^2 = E\{(f(x, y) - \hat{f}(x, y))^2\} \quad (2)$$

Para encontrar o valor mínimo da função de erro, é preciso considerar que: *i)* O ruído e a imagem não são correlacionados e possuem média igual a zero; e *ii)* A intensidade dos pixels da imagem estimada é uma função linear dos níveis encontrados na imagem degradada. Além

disso, é necessário conhecimento prévio da função de degradação do ruído e dos espectros de potência do ruído e da imagem não degradada. Dessa forma, o mínimo valor da função de erro é calculado no domínio da frequência através expressão (GONZALEZ; WOODS, 2010):

$$\begin{aligned}\hat{F}(u, v) &= \left[\frac{H^*(u, v)S_f(u, v)}{S_f(u, v)|H(u, v)|^2 + S_f(u, v)} \right] G(u, v) \\ &= \left[\frac{H^*(u, v)}{|H(u, v)|^2 + S_\eta(u, v)/S_f(u, v)} \right] G(u, v) \\ &= \left[\frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + S_\eta(u, v)/S_f(u, v)} \right] G(u, v) \quad (3)\end{aligned}$$

Os termos da Equação (3) são:

$H(u, v)$ = função de degradação do ruído;

$H^*(u, v)$ = conjugado complexo de $H(u, v)$;

$|H(u, v)|^2 = H^*(u, v)H(u, v)$, ou seja, o produto de um número complexo pelo seu conjugado corresponde à sua magnitude ao quadrado;

$S_\eta(u, v) = |N(u, v)|^2$ = espectro de potência do ruído;

$S_f(u, v) = |F(u, v)|^2$ = espectro da imagem não degradada.

Dado que o resultado da filtragem é uma função no domínio da frequência, é preciso realizar a transformada inversa de Fourier da estimativa $\hat{F}(u, v)$ para obter a imagem restaurada (GONZALEZ; WOODS, 2010).

Outra técnica utilizada para *denoising* é a limiarização de coeficientes *wavelet* da imagem original. A transformada *wavelet* é uma técnica utilizada para decompor a imagem em diferentes escalas no domínio da frequência e do tempo, resultando em um determinado número de subimagens a depender do número de escalas utilizado. Para a construção de uma *wavelet*, a onda que representa o sinal da transformada, é preciso definir uma *wavelet-mãe* ($\psi(t)$) que corresponde à função oscilatória utilizada para gerar as bases a partir de dilatações (a) e deslocamentos (b). A Transformada Discreta de *Wavelet* é definida pela Equação (4) (CARNEIRO, 2019):

$$TWD(m, n) = \frac{1}{\sqrt{a_0^m}} \int_{-\infty}^{\infty} f(x) \psi * (a_0^{-m}x - k) dx \quad (4)$$

Sendo $a = 2^j$ e $b = k$ números inteiros que representam a escala e a translação, respectivamente, e que permitem a representação do sinal de forma hierárquica, permitindo a análise multiresolução da imagem. Se m e n são número inteiros que caracterizam a escala e o

deslocamento, respectivamente, podemos discretizar o parâmetro a como $a = (a_0^m)$ e b é como $b = b_0 n_0 a_0^m$, sendo que a_0 define os passos discretos de escalonamento e b_0 o deslocamento.

Utilizando a análise multiresolução desenvolvida por Mallat, é possível obter aproximações e detalhes de um sinal através de um algoritmo piramidal. Nesse tipo de representação, considerando uma sequência crescente de resoluções espaciais $(r_j)_{j \in \mathbb{Z}}$, a diferença de informação entre a aproximação (r_j) e a aproximação com resolução mais baixa (r_{j-1}) é definida como o coeficiente de detalhes na resolução r_j . Dessa forma, o sinal $f(x)$ pode ser decomposto nos dois termos indicados na Equação (5) (CARNEIRO, 2019):

$$f_{2^{j+1}} = f_{2^j}(x) + D_{2^j}f(x) \quad (5)$$

Na Equação (5), o termo $f_{2^j}(x)$ representa o coeficiente de aproximação que contém as baixas frequências do sinal. O coeficiente de detalhes, indicado pelo termo $D_{2^j}f(x)$, é subdividido em três direções: vertical, horizontal e diagonal. Portanto, em cada nível de resolução da imagem obtém-se quatro subimagens: aproximação, vertical, horizontal e diagonal (CARNEIRO, 2019).

Os coeficientes obtidos em cada resolução na transformada *wavelet* podem ser manipulados para obter diversos resultados, como a redução de ruído e o aumento de contraste nas imagens. É o caso das pesquisas desenvolvidas por Carneiro e Costa Júnior (CARNEIRO, 2019; COSTA JÚNIOR, 2019) que realizaram a limiarização dos coeficientes de detalhe para diminuir o ruído presente na imagem. A limiarização, também chamada de *thresholding*, é utilizada para zerar os elementos com valores abaixo de um limiar pré-estabelecido e pode ser definida pela Equação (6) (GONZALEZ; WOODS, 2010):

$$thr_T^{\square}(D_{j,k}) = \begin{cases} 0, & se |D_{j,k}| \leq T \\ (D_{j,k}), & se |D_{j,k}| > T \end{cases} \quad (6)$$

Nesses casos, o valor T do limiar utilizado para limiarizar os coeficientes de detalhes como método de redução de ruído foi calculado através da Equação (7). Nessa equação, o termo σ_{im}^2 corresponde a variância da imagem original e k é uma constante definida manualmente (CARNEIRO, 2019).

$$T = \sigma_{im}^2 \sqrt{\frac{\pi}{2}} + k \sqrt{\frac{4 - \pi}{2}} (\sigma_{im}^2)^2 \quad (7)$$

Para avaliar a eficiência de um processamento de restauração de imagens, é possível utilizar métricas que comparam a imagem restaurada e a imagem original. É o caso da relação sinal-ruído de pico, conhecida por PSNR (*Peak Signal-to-Noise Ratio*), definida como relação

entre a potência máxima do sinal e do ruído. O PSNR é dado em decibéis (dB) e pode ser calculado com a Equação (8) (CARNEIRO, 2019).

$$PSNR = 10 \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (8)$$

Sendo que n representa o número de bits da imagem, ou seja, 2^n resulta no número de níveis de cinza presentes na imagem. Já valor do erro quadrático médio entre as imagens original e restaurada, conhecido pela sigla MSE (*Mean Square Error*), é calculado através da Equação (9) (GONZALEZ; WOODS, 2010).

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_A(i, j) - I_B(i, j))^2 \quad (9)$$

3.2. Operadores Morfológicos

O processamento morfológico de imagens envolve a utilização elementos estruturantes para extrair representações da forma dos objetos em uma imagem. No geral, os objetos da imagem são tratados como conjuntos e podem ser manipulados de diversas maneiras para expansão, afinamento, rotação e detecção de padrões nas imagens. O elemento estruturante é uma pequena subimagem, com forma escolhida de acordo com a finalidade, utilizada para investigar propriedades de interesse em uma imagem (GONZALEZ; WOODS, 2010).

Dentre as operações morfológicas, duas se destacam como elementares: a erosão e a dilatação. A erosão é uma operação utilizada para identificar e eliminar objetos menores que o elemento estruturante. Considerando A o conjunto que representa a imagem binária $f(x, y)$ e B o conjunto do elemento estruturante com coordenadas (s, t) , podemos definir a erosão de A por B com a Equação (10) (GONZALEZ; WOODS, 2010; SOILLE, 2004).

$$A \ominus B = \{z | (B)_z \cap A^c = \emptyset\} \quad (10)$$

O termo $(B)_z$ representa a translação do elemento estruturante, realizada pela substituição das coordenadas (s, t) de B por $(s + z_1, t + z_2)$. Além disso, a operação A^c indica o complemento da imagem, que se refere aos pixels que não pertencem ao conjunto A (DOUGHERTY, 2009; GONZALEZ; WOODS, 2010).

Na erosão de uma imagem binária, se pelo menos um pixel da vizinhança delimitada pelo elemento estruturante pertencer ao fundo da imagem, o pixel analisado é considerado como parte do fundo, ou seja, a vizinhança deve coincidir perfeitamente com elemento estruturante. No caso de imagens em níveis de cinza, a erosão corresponde ao valor mínimo da imagem na região que coincide com o elemento estruturante B e pode ser representada pela Equação (11) (DOUGHERTY, 2009; GONZALEZ; WOODS, 2010).

$$[f \ominus b](x, y) = \min_{(s,t) \in b} \{f(x + s, y + t)\} \quad (11)$$

Já na operação morfológica dilatação, os objetos da imagem binária são expandidos e pequenas lacunas são preenchidas. Na dilatação, um pixel é considerado parte de uma região se pelo menos um pixel da vizinhança definida pelo elemento estruturante coincide com a máscara definida. A dilatação pode ser representada pela Equação (12), na qual a transformação \hat{B} refere-se a reflexão do elemento estruturante, ou seja, as coordenadas (s, t) são substituídas por $(-s, -t)$ (DOUGHERTY, 2009; GONZALEZ; WOODS, 2010).

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (12)$$

Em imagens em níveis de cinza, a dilatação corresponde ao valor máximo da região sobreposta pela reflexão do elemento estruturante \hat{b} e pode ser descrita pela Equação (13) (DOUGHERTY, 2009; GONZALEZ; WOODS, 2010).

$$[f \oplus b](x, y) = \max_{(s,t) \in b} \{f(x - s, y - t)\} \quad (13)$$

É importante ressaltar que a erosão e a dilatação em imagens com níveis de cinza são equivalentes a filtragem da imagem com os filtros espaciais não-lineares de mínimo e máximo, respectivamente. Além disso, a erosão e a dilatação possuem a propriedade dualidade, ou seja, o complemento de uma erosão é a dilatação do complemento da imagem original pelo reflexo do elemento estruturante. A propriedade dualidade das operações erosão e dilatação podem ser descritas pelas Equações (14) e (15), respectivamente. (GONZALEZ; WOODS, 2010).

$$(A \ominus B)^c = A^c \oplus \hat{B} \quad (14)$$

$$(A \oplus B)^c = A^c \ominus \hat{B} \quad (15)$$

Outras duas operações morfológicas importantes são a abertura e fechamento, as quais são implementadas utilizando a erosão e a dilatação da imagem. A abertura de uma imagem é utilizada para suavizar as fronteiras dos objetos, eliminar estruturas finas e pequenos detalhes claros. Já o fechamento é útil na restauração da conectividade entre regiões, eliminação de pequenos buracos no contorno dos objetos e supressão pequenos detalhes escuros (GONZALEZ; WOODS, 2010; SOILLE, 2004).

A abertura morfológica preserva as regiões com formato semelhante ao do elemento estruturante. Nesse processamento, as regiões com valor de intensidade maior que as regiões vizinhas e não contém o elemento estruturante são eliminadas. Formalmente, a abertura da imagem A pelo elemento estruturante B é dada pela erosão de A por B e a dilatação do resultado por B e pode ser descrita pela Equação (16) (DOUGHERTY, 2009; GONZALEZ; WOODS, 2010; SOILLE, 2004).

$$A \circ B = (A \ominus B) \oplus B \quad (16)$$

Analogamente, o fechamento da imagem A pelo elemento estruturante B é dado pela dilatação de A por B e a erosão do resultado por B , como apresentado na Equação (17) (GONZALEZ; WOODS, 2010).

$$A \bullet B = (A \oplus B) \ominus B \quad (17)$$

Assim como a erosão e a dilatação, a abertura e o fechamento são duais entre si. Além disso, uma característica importante nesses processamentos é que múltiplas aberturas ou fechamentos de uma imagem usando o mesmo elemento estruturante não produz efeito na imagem (GONZALEZ; WOODS, 2010).

As operações morfológicas apresentadas também podem ser aplicadas na reconstrução morfológica, utilizando duas imagens e um elemento estruturante. A reconstrução morfológica parte dos pontos definidos na imagem marcadora e restringe a transformação através da máscara. Na reconstrução morfológica por dilatação, a dilatação geodésica do marcador F em relação à máscara G deve ser iteragida até a estabilidade, como estabelecido pela Equação (18) (GONZALEZ; WOODS, 2010).

$$R_G^D(F) = D_G^{(k)}(F) \text{ com } k \text{ tal que } D_G^{(k)}(F) = D_G^{(k+1)}(F) \quad (18)$$

Sendo que a dilatação geodésica, representada por $D_G^{(k)}$, corresponde à intersecção da máscara G e o resultado da dilatação de F pelo elemento estruturante B . A dilatação geodésica pode ser calculada pela Equação (19), considerando $D_G^{(1)}(F)$ como o resultado da primeira iteração, dado pela Equação (20) (GONZALEZ; WOODS, 2010).

$$D_G^{(n)}(F) = D_G^{(1)}[D_G^{(n-1)}(F)] \quad (19)$$

$$D_G^{(1)}(F) = (F \oplus B) \cap G \quad (20)$$

Da mesma forma, a reconstrução morfológica por erosão é definida pela erosão geodésica do marcador F em relação à máscara G iteragida até a estabilidade. Nesse caso, a erosão geodésica é tratada pela união da máscara G e o resultado da erosão de F pelo elemento estruturante B . A reconstrução morfológica por erosão pode ser representada pela Equação (21), sendo que a erosão geodésica é definida com auxílio das Equações (22) e (23) (GONZALEZ; WOODS, 2010).

$$R_G^E(F) = E_G^{(k)}(F) \text{ com } k \text{ tal que } E_G^{(k)}(F) = E_G^{(k+1)}(F) \quad (21)$$

$$E_G^{(n)}(F) = E_G^{(1)}[E_G^{(n-1)}(F)] \quad (22)$$

$$E_G^{(1)}(F) = (F \ominus B) \cup G \quad (23)$$

A operação de reconstrução também pode ser aplicada na abertura morfológica permitindo restaurar a forma dos objetos que permanecem após a erosão. A abertura por reconstrução morfológica é dada pela reconstrução por dilatação da imagem resultante da erosão de tamanho n da imagem F e pode ser descrita pela Equação (24) (GONZALEZ; WOODS, 2010).

$$O_R^{(n)}(F) = R_F^D[F \ominus nB] \quad (24)$$

Analogamente, o fechamento por reconstrução é obtido pela reconstrução morfológica por erosão da imagem resultante da dilatação de tamanho n da imagem F e é dado pela Equação (25) (GONZALEZ; WOODS, 2010).

$$C_R^{(n)}(F) = R_F^E[F \oplus nB] \quad (25)$$

Uma das aplicações das operações morfológicas é no cálculo do gradiente morfológico da imagem, processamento obtido pela diferença entre os resultados da dilatação e da erosão da imagem utilizando o mesmo elemento estruturante, como indicado na Equação (26). O gradiente morfológico representa as fronteiras dos objetos da imagem, com base no conjunto de pixels cuja intensidade muda abruptamente, também chamado de bordas da imagem. O resultado do gradiente morfológico contém as bordas internas (g^-) e externas (g^+) dos objetos, duas componentes que podem ser obtidas separadamente pela subtração da imagem original e os resultados da erosão e dilatação, respectivamente. (GONZALEZ; WOODS, 2010; SOILLE, 2004).

$$g = (f \oplus b) - (f \ominus b) \quad (26)$$

Utilizando a subtração da imagem original e o resultado de aberturas e fechamentos é possível obter as transformadas *top-hat* e *bottom-hat*, operações cujo resultado é a uma imagem contendo objetos removidos da imagem original. Esses objetos são removidos da imagem de acordo com a forma e tamanho do elemento estruturante definido, sendo necessário adaptá-lo ao tipo de região que deverá ser removida. A operação *top-hat* é mais utilizada para objetos claros sobre fundos escuros, enquanto *bottom-hat* é preferível para remover objetos mais escuros que o fundo (GONZALEZ; WOODS, 2010; SOILLE, 2004).

O *top-hat* é a diferença entre a imagem original e o resultado da abertura pelo elemento estruturante b , enquanto o *bottom-hat* corresponde à diferença entre o resultado do fechamento da imagem pelo elemento estruturante b e a imagem original. Esses processamentos também podem ser aplicados utilizando a abertura e o fechamento por reconstrução morfológica na subtração com a imagem original. As Equações (27) e (28) descrevem o cálculo das operações *top-hat* e *bottom-hat* (GONZALEZ; WOODS, 2010; SOILLE, 2004).

$$T_{hat}(f) = f - (f \circ b) \quad (27)$$

$$B_{hat}(f) = (f \bullet b) - f \quad (28)$$

Quando tratamos de regiões de uma imagem f que possuem um nível de intensidade t , podemos definir dois tipos regiões quanto a sua intensidade e relacionamento com as regiões vizinha. Um máximo regional de uma imagem a um nível de intensidade t é um componente de pixels conectados cujo valor de intensidade da borda externa do objeto possui um valor menor que t . Seguindo o mesmo princípio, um mínimo regional é dado por um componente de pixels conectados cujo valor de intensidade dos pixels da borda externa possuem valor maior que t (SOILLE, 2004).

Nesse sentido, visando identificar e selecionar os máximos regionais, a operação morfológica *h-maxima* pode ser utilizada para eliminar os máximos regionais com profundidade menor que h através da dilatação por reconstrução de f por $(f - h)$. Esse resultado pode ser subtraído da imagem original para obter a imagem *h-convex*, a qual pode ser limiarizada para obter objetos mais claros que as regiões vizinhas. O cálculo para produção das imagens *h-maxima* e *h-convex* estão descritos nas Equações (29) e (30) (SOILLE, 2004).

$$HMAX = R_f^D(f - h) \quad (29)$$

$$HCONVEX = f - HMAX \quad (30)$$

3.3. Segmentação

A segmentação é um tipo de processamento que subdivide uma imagem em regiões distintas de acordo com características semelhantes entre pixels e sua vizinhança. Uma das utilidades a segmentação é o aplicação do resultado em procedimentos de análise computadorizada que envolvem classificação e extração de características das regiões (GONZALEZ; WOODS, 2010).

As técnicas de segmentação baseiam-se em propriedades básicas dos valores de intensidade da imagem, como a descontinuidade e a similaridade. A descontinuidade consiste na divisão da imagem em regiões com limites definidos conforme mudanças locais de intensidade. A similaridade permite a divisão da imagem em regiões com características semelhantes de acordo com critérios pré-estabelecidos (GONZALEZ; WOODS, 2010).

Um dos métodos para segmentação de imagens é a limiarização de intensidade, a qual consiste na divisão do histograma da imagem em duas partes, o objeto e o fundo, a partir de um limiar pré-estabelecido. A limiarização global é realizada quando o mesmo limiar é utilizado em toda imagem, sendo possível a variação do limiar conforme as coordenadas espaciais (x, y) e as características da vizinhança na limiarização variável, local ou adaptativa. A limiarização

global utilizando um único limiar pode ser descrita pela Equação (31) (GONZALEZ; WOODS, 2010).

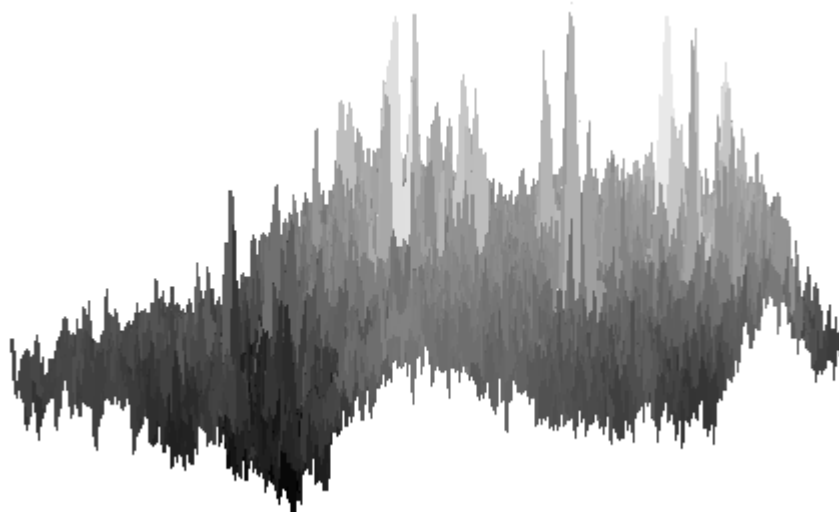
$$g(x, y) = \begin{cases} 1, & \text{se } f(x, y) > T \\ 0, & \text{se } f(x, y) \leq T \end{cases} \quad (31)$$

Além disso, o conceito de limiarização pode ser estendido para uso com múltiplos limiares para divisão da imagem em várias regiões através da atribuição de valores de intensidade diferentes a cada um dos limiares. A definição do limiar de intensidade utilizado na divisão histograma é dependente do ruído e contraste presentes na imagem e deve ser definido conforme as características dos objetos que serão segmentados (GONZALEZ; WOODS, 2010).

Uma das formas de seleção do limiar utilizado no *thresholding* global é o método de Otsu que busca selecionar um limiar pela divisão do histograma da imagem em duas classes de forma a minimizar o erro médio na atribuição dos pixels em cada grupo. Isso é feito visando maximizar a variância entre as classes do histograma normalizado, partindo do cálculo da probabilidade do pixel em pertencer a cada grupo, as intensidades médias e variâncias globais e de cada classe (GONZALEZ; WOODS, 2010).

Outro método de segmentação é a detecção de *watersheds* morfológicas, um tipo de processamento baseado em operadores morfológicos utilizado para extrair a fronteira dos objetos da imagem e que tem como principal vantagem resultar em contornos contínuos entre as regiões. Para aplicá-la, é preciso projetar a imagem em três dimensões, em uma visão topográfica composta por duas coordenadas espaciais e um eixo que associa cada pixel com uma elevação proporcional ao nível de cinza. A Figura 2 exemplifica um conjunto de microcalcificações em visão topográfica (DOUGHERTY, 2009; GONZALEZ; WOODS, 2010).

Figura 2 - Microcalcificações em visão topográfica.



Fonte: Elaborado pelo autor.

Utilizando a abordagem topográfica, o processamento identifica e expande os mínimos regionais da imagem utilizando sucessivas dilatações limitadas pelas fronteiras dos objetos pertencentes a cada mínimo regional. Os mínimos regionais são representados por pixels conectados com intensidade constante e cercados por pixels de maior valor de intensidade. Na segmentação por *watershed* os mínimos regionais separam um conjunto de pixels em pixels cuja variação de intensidade tendem a um mínimo simples, cujo conjunto forma as *watershed* e os pontos que a variação tem a mesma probabilidade de tender a mais de um mínimo regional, formando as linhas de *watershed*. As linhas de *watershed* correspondem as fronteiras dos objetos, ou seja, são os pontos que devem ser extraídos na segmentação (DOUGHERTY, 2009; GONZALEZ; WOODS, 2010).

Para construir as *watersheds*, é possível utilizar a dilatação morfológica do conjunto de mínimos regionais da imagem. Quando a dilatação dos pixels de um mínimo regional causa a mistura de pixels com outros conjuntos dilatados, os pixels em comum são marcados como parte das linhas de watershed que correspondem as fronteiras dos objetos encontrados (GONZALEZ; WOODS, 2010).

Considerando um conjunto de mínimos regionais representados por M_1, M_2, \dots, M_i , sendo cada elemento um conjunto conexo de pontos pertencentes a uma *watershed* $C(M_i)$. Além disso, sendo $T[n]$ o conjunto de pontos com intensidade menor que o valor de intensidade n , pode-se encontrar os pontos da *watershed* $C_n(M_i)$ relacionadas ao mínimo M_i abaixo do plano definido por n através da Equação (32). (GONZALEZ; WOODS, 2010):

$$C_n(M_i) = C(M_i) \cap T[n] \quad (32)$$

Podemos definir então $C[n]$ como a união das *watersheds* associadas ao nível de intensidade n , sendo esse valor variável entre os níveis mínimo e máximo de intensidade da imagem, como representado pela Equação (33). Quando o valor de n é maior que o valor máximo de intensidade, o conjunto $C[n_{m\acute{a}x} + 1]$ representa a união de todas as *watersheds* presentes na imagem (GONZALEZ; WOODS, 2010).

$$C[n] = \bigcup_{i=1}^R C_n(M_i) \quad (33)$$

Para obter $C[n]$ a partir de $C[n - 1]$ analisa-se cada componente compactado q que pertence ao conjunto de componentes conectados em $T[n]$. Se $q \cap C[n - 1]$ está vazia um novo mínimo foi encontrado e o componente deve ser incorporado a $C[n - 1]$ para formar $C[n]$. Se $q \cap C[n - 1]$ contém apenas um componente conectado em $C[n - 1]$ o componente q fica

dentro da *watershed* associada ao mínimo regional e é adicionado para compor $C[n]$. Por fim, caso $q \cap C[n - 1]$ possua mais de um componente conectado em $C[n - 1]$ é preciso separar q em regiões pois um pixel de fronteira foi encontrado. As linhas de *watershed* que dividem q podem ser construídas utilizando a dilatação $q \cap C[n - 1]$ com um elemento estruturante 3×3 e restringindo a dilatação a q (GONZALEZ; WOODS, 2010).

O principal desafio na aplicação da segmentação por *watershed* é evitar a supersegmentação da imagem, ou seja, a detecção de muitos objetos em sua maioria irrelevantes. Esse problema ocorre, principalmente, devido ao ruído presente na imagem processada, causa da detecção exagerada de mínimos potenciais que leva a formação muitas *watershed* (GONZALEZ; WOODS, 2010).

Como solução, é possível realizar o pré-processamento da imagem, aplicando filtros de suavização como forma de diminuir o ruído existente. Alguns filtros utilizados para esse fim são os filtros de média, mediana, Wiener e processamentos utilizando a transformada *wavelet*. Além disso, pode-se limitar os números de mínimos regionais permitidos e diminuir a quantidade de regiões nas quais a *watershed* é aplicada, utilizando marcadores definidos com base nos objetos de interesse e no fundo da imagem (GONZALEZ; WOODS, 2010).

Os marcadores consistem em componentes conectados pertencentes a imagem que identificam as regiões de interesse e restringem a segmentação por *watershed*. Um marcador interno é composto de componentes associados ao objeto que será segmentado. O marcador externo é relacionado ao fundo da imagem e delimita a expansão dos marcadores internos (GONZALEZ; WOODS, 2010).

A avaliação dos resultados da segmentação de uma imagem tem como princípio a comparação entre as regiões detectadas e as regiões originais definidas por um *ground truth*. Um dos métodos para avaliação é verificar a eficiência do algoritmo quanto a detecção ou não de um determinado objeto utilizando uma matriz de confusão, também chamada de tabela de contingência. Considerando um objeto p a ser detectado e o fundo da imagem como n , uma matriz de confusão pode ser descrita como na Tabela 2, na qual é possível identificar quatro tipos de resultados quanto a detecção do objeto (FAWCETT, 2006).

Tabela 2 - Matriz de confusão.

		Acontecimento	
		Positivo (p)	Negativo (n)
Resposta	Sim (S)	VP (verdadeiro positivo)	FP (falso positivo)

	Não (N)	FN (falso negativo)	VN (verdadeiro negativo)
--	-------------	--------------------------	-------------------------------

Fonte: Adaptado de Fawcett, 2006.

Dessa forma, os resultados de uma segmentação podem ser categorizados em (FAWCETT, 2006):

1. VP : Verdadeiro positivo, referente à detecção do objeto quando ele existe;
2. VN : Verdadeiro negativo, relativo à não detecção do objeto quando não existe;
3. FP : Falso positivo, correspondente à detecção do objeto quando ele não existe;
4. FN : Falso negativo, equivalente à não detecção do objeto quando ele existe.

Através da classificação dos resultados da segmentação nessas categorias podemos definir métricas capazes de descrever de maneira ampla a eficiência do algoritmo avaliado. Dentre essas métricas está a sensibilidade, definida como a capacidade de detecção do objeto, ou seja, ao número de verdadeiros positivos com relação ao número total de detecções corretas esperadas. Analogamente, a especificidade corresponde à capacidade de identificar a ausência de objeto na segmentação, ou seja, corresponde ao número de verdadeiro negativos em relação ao número de detecções negativas esperadas. Matematicamente, podemos definir a sensibilidade e a especificidade pelo cálculo das Equações (34) e (35) (FAWCETT, 2006).

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (34)$$

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (35)$$

Analogamente, a taxa de falsos positivos (fp_{rate}) é estimada pela razão entre o número de objetos detectados incorretamente e o número total de objetos negativos no *ground truth*, como descrito pela Equação (36) (FAWCETT, 2006).

$$fp_{rate} = 1 - \text{especificidade} = \frac{FP}{FP + VN} \quad (36)$$

Outra medida de avaliação é a precisão, capaz de identificar o número de objetos classificados corretamente dentre o número total de objetos segmentados. A precisão pode ser calculada através da Equação (37) (FAWCETT, 2006).

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (37)$$

A matriz de confusão também permite o cálculo da acurácia, indicadora da proporção de acertos da segmentação, ou seja, o número de verdadeiros positivos e negativos e comparado com o total de itens analisados. A Equação (38) demonstra o cálculo da acurácia a partir dos valores encontrados na matriz de confusão (FAWCETT, 2006).

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (38)$$

Além disso, a partir das métricas apresentadas, é possível estimar a curva ROC (*Receiver Operating Characteristics*), um tipo de gráfico utilizado para avaliar a performance de um algoritmo. A curva ROC é desenvolvida posicionando a taxa de falsos positivos do resultado na coordenada x e a sensibilidade na coordenada y . Através da curva ROC determina-se uma importante métrica para avaliação da performance de um algoritmo de segmentação, sendo essa medida obtida pelo cálculo da área sob a curva e conhecida pela sigla AUC (*Area Under Curve*) (FAWCETT, 2006).

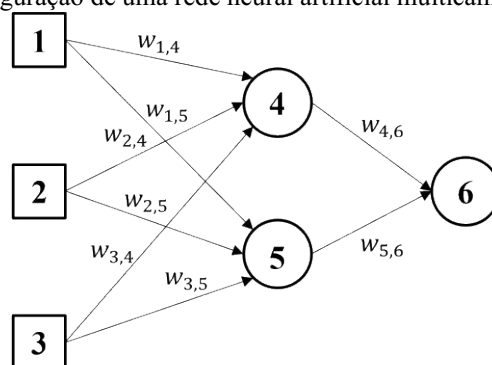
4. REDES NEURAIS

Os problemas de aprendizado de máquina tratam da capacidade de um algoritmo em prever a ocorrência de um evento através dos dados disponíveis e as relações entre eles. No caso da aprendizagem supervisionada, o objetivo é encontrar uma função h que se aproxime de uma função desconhecida $y = f(x)$ a partir de um conjunto de treinamento com N pares de exemplos com entradas e saídas conhecidas e forma (x_n, y_n) . Considerando que a função h é uma hipótese para a solução do problema, é possível medir a precisão dessa hipótese utilizando um conjunto de teste, formado por exemplos distintos do conjunto de treinamento. Em problemas de classificação, nos quais a saída y assume um conjunto de valores finitos, o objetivo do aprendizado é determinar a distribuição de probabilidade condicional $P(Y|x)$ (RUSSELL; NORVIG, 2013).

O aprendizado supervisionado pode ser aplicado através de algoritmos clássicos como as regressões linear e logística, máquinas de vetores de suporte, redes neurais artificiais e árvores de decisão. No caso das Redes Neurais Artificiais do tipo Perceptron Multicamadas (RNA-MLP), problemas com funções não lineares podem ser solucionados através da combinação de variáveis de entrada, aplicação de funções de ativação e propagação de ativação através da ligação entre unidades (RUSSELL; NORVIG, 2013).

As redes neurais artificiais são modelos matemáticos inspirados nos neurônios biológicos e são compostas por nós, também chamados de unidades, conectados por ligações para formação de camadas de entrada, intermediárias e de saída. As unidades calculam a soma ponderada de suas entradas e aplicam uma função de ativação para obter uma saída. A Figura 3 ilustra a configuração de uma RNA com ligações entre três camadas: uma camada de entrada com três unidades; uma camada intermediária com dois nós; e uma camada de saída única, comumente utilizada em problemas de classificação binária (RUSSELL; NORVIG, 2013).

Figura 3 - Configuração de uma rede neural artificial multicamadas.



Fonte: Elaborado pelo autor.

Considerando a função de ativação g e dois nós i e j conectados por uma ligação com um peso numérico $w_{i,j}$ associado, pode-se calcular a saída da unidade j através da Equação (39). Nesse caso, a variável a_i representa a saída da unidade i (RUSSELL; NORVIG, 2013).

$$a_j = g \left(\sum_{i=0}^n w_{i,j} a_i \right) \quad (39)$$

A função de ativação é utilizada para transformar os valores calculados nas unidades em um nível de saída, de acordo com o intervalo estabelecido pela função. No caso da função sigmoide, os valores são restritos ao intervalo $[0, 1]$ sendo possível calculá-los através da Equação (40). Outras funções muito utilizadas são a tangente hiperbólica, que produz valores entre $[-1, 1]$, a *softmax* e a ReLU (*Rectified Linear Unit*) (LITJENS et al., 2017; SOUSA, 2022).

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (40)$$

Dadas as características das redes neurais artificiais, o resultado de um nó de saída pode ser descrito como uma função não linear das entradas da unidade, sendo essas, por sua vez, formadas pelos resultados das ativações das camadas anteriores. Dessa forma, é possível associar uma função de custo, também chamada de erro, às respostas produzidas pelos neurônios de saída em relação aos resultados esperados e aplicar métodos para minimizar o erro e obter respostas mais precisas (RUSSELL; NORVIG, 2013; SOUSA, 2022).

Uma das formas de aprimorar a performance das RNA-MLP é a utilização de redes recorrentes, ou *backpropagation*, que pelo cálculo do gradiente da função de erro da camada de saída pode ser aplicado recursivamente em métodos de atualização dos pesos das ligações entre as unidades da rede. Nesse sentido, otimizadores capazes de percorrer o gradiente do erro em relação aos pesos podem ser aplicados como forma de aprimorar o algoritmo quanto a velocidade de processamento, consumo de memória e atualização dos pesos, visto que a minimização do erro depende da convergência do gradiente do erro. Alguns otimizadores amplamente utilizados em problemas de RNA-MLP são o SGD (*Stochastic Gradient Descent*), AdaGrad (*Adaptive Gradient Algorithm*), RMSProp (*Root Mean Square Propagation*) e ADAM (*Adaptive Moment Estimation*) (SOUSA, 2022).

O otimizador ADAM utiliza as estimativas do primeiro (m_t) e segundo (v_t) momentos (média e variância descentralizada, respectivamente) na manipulação dos pesos a cada iteração t de atualização de pesos. Considerando ∇_t o gradiente em uma iteração e β_1 e β_2 duas constantes pré-definidas no intervalo $[0, 1]$, pode-se estimar a média e a variância descentralizada através das Equações (41) e (42) (KINGMA; BA, 2015; SOUSA, 2022).

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla_t \quad (41)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \nabla_t^2 \quad (42)$$

Com o auxílio das Equações (43) e (44), o otimizador realiza a atualização através da Equação (45), na qual ϵ corresponde a uma constante para evitar a divisão por zero (KINGMA; BA, 2015; SOUSA, 2022).

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (43)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (44)$$

$$w_{t+1} = w_t - \frac{\alpha \hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (45)$$

Na Equação (45), a variável α representa a taxa de aprendizagem do modelo, a qual controla o passo com que o otimizador percorre o gradiente em direção ao valor mínimo da função. As variáveis taxa de aprendizagem, β_1 , β_2 e ϵ são normalmente chamadas de hiperparâmetros e devem ser definidas antes do treinamento do modelo. Normalmente, são realizados vários testes com valores aleatórios para os hiperparâmetros, de forma que o modelo com melhor performance seja selecionado pela comparação entre o desempenho de cada teste (SOUSA, 2022).

Uma das formas de avaliar a performance de um modelo RNA-MLP durante o treinamento é pela validação cruzada, também chamada *k-fold cross validation*. A validação cruzada com k -repetições divide o conjunto em k subconjuntos iguais e realiza k rodadas de aprendizagem, cada uma utilizando conjuntos de treinamento e validação distintos. Em cada etapa de aprendizagem, um subconjunto é selecionado para validação enquanto os $k - 1$ subconjuntos restantes serão utilizados para treinamento. Através da pontuação média obtida com os subconjuntos de validação, é possível avaliar e comparar a performance de diferentes modelos para seleção da hipótese com maior capacidade de classificação correta dos dados. Normalmente, o número k de subconjuntos é definido como cinco ou dez, dependendo do volume de dados disponíveis para treinamento e teste dos modelos (RUSSELL; NORVIG, 2013).

5. ESTADO DA ARTE

Considerando o desafio na detecção e segmentação de microcalcificações nas imagens de mamografia devido a diversidade na morfologia e distribuição dessas estruturas, inúmeras pesquisas foram realizadas com o objetivo de facilitar a detecção do câncer de mama com mamografia. Os métodos de detecção e segmentação de microcalcificações mais utilizados envolvem detectores de bordas, limiarização, operadores morfológicos e consideram as principais características desses objetos: componentes pequenos com intensidade maior que a intensidade das regiões vizinhas. Além disso, é comum a aplicação de estratégias de classificação para reduzir o número de falsos positivos do resultado da segmentação e o desenvolvimento de modelos de redes neurais para encontrar microcalcificações em mamografias (REZAEI, 2021).

Os métodos que utilizam detectores de borda são baseados em mudanças locais abruptas de intensidade para encontrar os pixels de borda dos objetos (GONZALEZ; WOODS, 2010). A tese de Silva Júnior (SILVA JÚNIOR, 2009) implementa a combinação de dois filtros detectores de borda: um filtro diferencial e o filtro de Sobel. Os resultados, que contêm o gradiente das imagens, são limiarizados e combinados para identificar objetos de interesse e identificar falsos positivos com análise fractal.

Outro detector de borda utilizado na detecção de estruturas que se assemelham a microcalcificações é o laplaciano da gaussiana (LoG) da imagem. Uma das formas de aplicação deste método é adaptá-lo ao tamanho dos componentes de interesse, ou seja, aos tamanhos mínimo e máximo das microcalcificações, como no caso de Cascio *et al.* (CASCIO *et al.*, 2018), que realizou a detecção com o filtro LoG adaptado e agrupou os objetos usando um método de clusterização. Para minimizar o número de falsos positivos, os clusters foram classificados com redes neurais, utilizando atributos geométricos e de intensidade dos agrupamentos e dos objetos.

As operações morfológicas são empregadas em problemas que requerem atenção a forma e características das regiões do pixel. Utilizando elementos estruturantes em diferentes, como disco, polígonos em geral, cruz, asterisco e até mesmo tridimensionais, é possível extrair elementos relevantes da imagem conforme a forma dos objetos (SOILLE, 2004). O trabalho de Ciecholewski (CIECHOLEWSKI, 2017) utiliza a técnica pirâmide morfológica, ou seja, a combinação de redução de escala com as operações abertura e fechamento para detectar as microcalcificações em várias escalas e extrair marcadores para serem utilizados na segmentação por *watershed*.

Além dos marcadores representando as microcalcificações, Ciecholewski (CIECHOLEWSKI, 2017) também definiu marcadores externos para caracterizar o fundo dos objetos. Esses marcadores foram obtidos com linhas de *watersheds* resultantes da segmentação por *watershed* da imagem original após a filtragem por fechamento e abertura e inversão dos níveis de cinza.

Outro exemplo de operação morfológica usada na segmentação de microcalcificações é a reconstrução morfológica por dilatação, utilizada na transformada *h-maxima*. A pesquisa de Halkiotis, Botsis e Rangoussi (HALKIOTIS; BOTSIS; RANGOUSI, 2007) utiliza a transformada *h-maxima* para detectar e extrair os máximos regionais da imagem de acordo com um limiar pré-estabelecido e segmenta os objetos com transformada máximos estendidos. As regiões detectadas foram classificadas para redução de falsos positivos utilizando redes neurais e atributos de intensidade dos objetos segmentados. Com essa mesma premissa, o trabalho de Karale *et al.* (KARALE *et al.*, 2019) extraiu características morfológicas, de intensidade e textura dos objetos e utilizou um modelo de máquina de vetor de suporte (SVM) na classificação dos objetos segmentados.

O processamento *top-hat* é dado pela subtração da imagem original e o resultado da abertura morfológica e é aplicado em várias pesquisas envolvendo a segmentação de microcalcificações. O trabalho de Wang *et al.* (WANG *et al.*, 2020) combina a imagem *top-hat* e o coeficiente de detalhes do terceiro nível de decomposição da transformada *wavelet* para obter as sementes que utilizadas na segmentação por crescimento de regiões. Nesse caso, a detecção e segmentação dos objetos só foi realizada nas subimagens obtidas pela classificação das imagens originais quanto a presença de microcalcificações utilizando redes neurais.

Nesse mesmo sentido, Gómez, Echeverry-Correa e Gutiérrez (GÓMEZ; ECHEVERRY-CORREA; GUTIÉRREZ, 2021) também aplicam a etapa de detecção apenas nas imagens classificadas positivamente em modelos de redes neurais convolucionais. Neste trabalho, a segmentação da imagem foi feita pela limiarização por Otsu do resultado do *top-hat* com aumento de contraste pela transformada *wavelet*.

A pesquisa de Cai *et al.* (CAI *et al.*, 2019) também realiza a limiarização por Otsu da imagem *top-hat*, porém, o elemento estruturante utilizado na operação morfológica foi do tipo bola com raio de 8 pixels e altura de 100 pixels, ou seja, uma máscara tridimensional. Utilizando atributos morfológicos, de intensidade e extraídas por redes neurais convolucionais, os objetos detectados foram classificados utilizando aprendizado supervisionado com máquina de vetor de suporte (SVM).

Outra pesquisa que trata da segmentação de microcalcificações é a de Touil *et al.* (TOUIL et al., 2021) que utilizou um algoritmo de crescimento de regiões para segmentação os objetos. As sementes utilizadas como ponto de partida da segmentação foram obtidas pela detecção dos máximos regionais das imagens e selecionadas com base em superpixels relacionados ao fundo da imagem. Um dos métodos para obtenção dos superpixels foi a seleção das linhas de *watershed* resultantes da segmentação por *watershed* aplicada ao complemento da imagem original.

6. METODOLOGIA

6.1. Base de Imagens

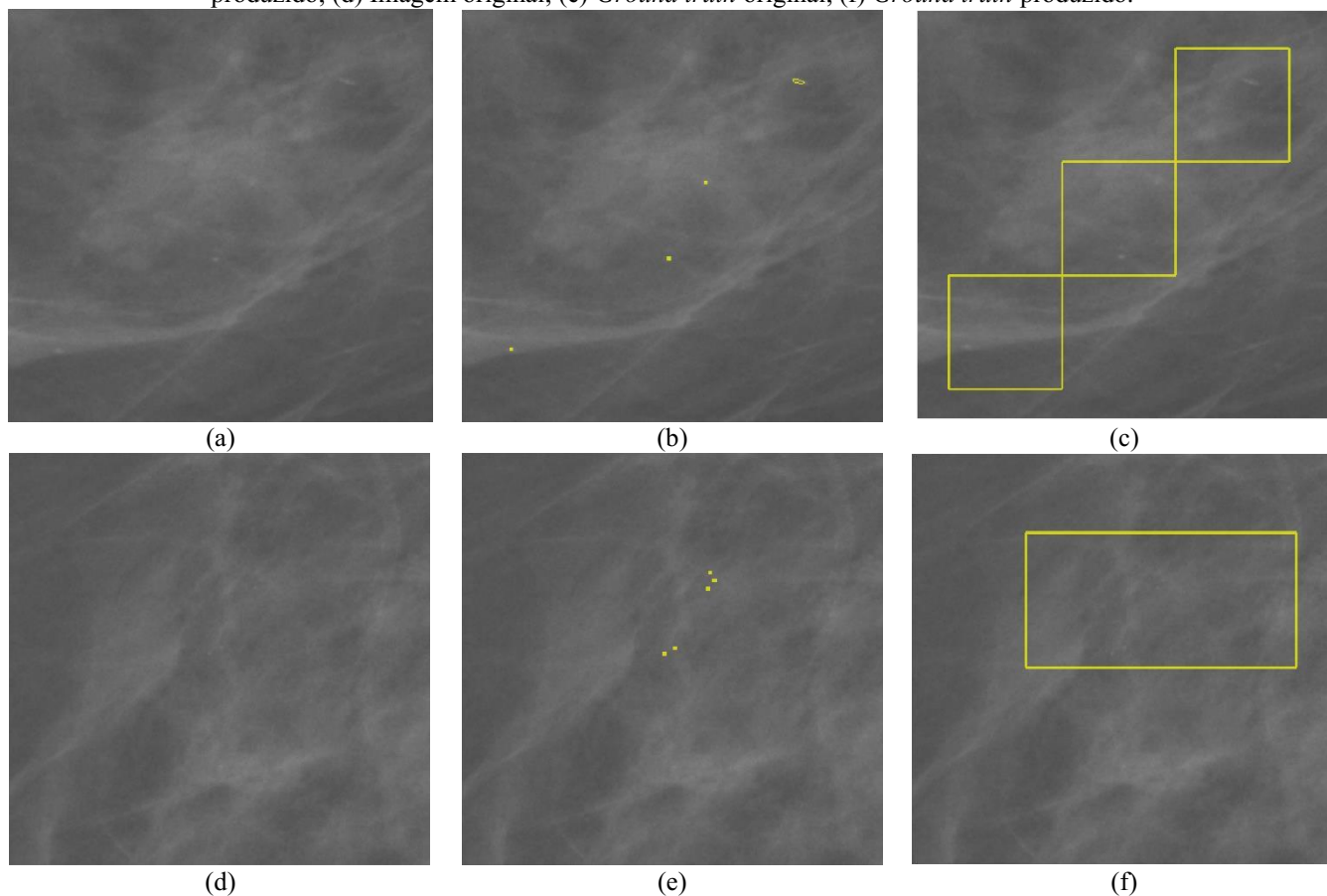
As imagens de mamografia utilizadas nesse trabalho são provenientes da base de imagens INbreast no formato DICOM (Digital Imaging and Communications in Medicine). As imagens foram adquiridas no Centro Hospitalar de São João na cidade de Porto, Portugal entre abril de 2008 e julho de 2010 utilizando um mamógrafo digital FFDM (Full Field Digital Mammography) MammoNovation da fabricante Siemens contendo um detector de selênio amorfo e possuem resolução de 12 bits e pixels com tamanho 70 μm . As imagens variam de tamanho, entre 3328 \times 4084 e 2560 \times 3328 pixels, devido aos diferentes tamanhos de mama e compressão utilizados (MOREIRA et al., 2012).

Os achados mamográficos são acompanhados de anotações contendo a localização dos pixels da borda dos objetos, além da descrição do tipo de estrutura detectada, como massas, distorções, calcificações e *clusters* de microcalcificações. As anotações estão disponíveis no formato XML (Extensible Markup Language) e, através da reconstrução dos pontos, foi possível produzir uma imagem contendo o *ground truth* dos objetos em cada imagem, ou seja, a representação a localização real dos objetos conforme laudo radiográfico (MOREIRA et al., 2012).

Para produzir as imagens do *ground truth*, os pontos indicados como pixels de borda foram utilizados para construção da envoltória convexa do conjunto, resultando em uma imagem binária com componentes conectados. A envoltória convexa produz o menor polígono convexo que contém todos os pontos do conjunto e mostrou-se apropriado para reconstrução de objetos a partir de poucos pontos, como no caso das anotações de microcalcificações do INbreast.

No geral, as anotações do INbreast foram registradas de diferentes formas, por exemplo, alguns objetos tiveram as bordas delimitadas individualmente, outros foram marcados apenas como pontos e algumas imagens apenas as regiões com *clusters* foram identificadas. Por esse motivo, visando a utilização do maior número de imagens contendo microcalcificações relevantes possível, as imagens de *ground truth* foram padronizadas. As imagens foram divididas em regiões de 100 \times 100 pixels e as regiões contendo ao menos um objeto foram consideradas como verdadeiras. A Figura 4 exemplifica os pontos encontrados no *ground truth* do INbreast e o *ground truth* produzido para utilização nesse trabalho.

Figura 4 - Padronização do ground truth. (a) Imagem original; (b) *Ground truth* original; (c) *Ground truth* produzido; (d) Imagem original; (e) *Ground truth* original; (f) *Ground truth* produzido.



Fonte: Elaborado pelo autor.

O INbreast é composto por 410 imagens, sendo que 301 contêm calcificações e 21 imagens apresentam *clusters* de microcalcificações (MOREIRA et al., 2012). Dentre as imagens com *ground truth*, foram selecionadas 178 imagens com calcificações ou *clusters*, considerando apenas imagens com no mínimo 4 objetos com área menor que 200 pixels. Em média, as imagens foram divididas em 691,94 regiões de tamanho 100×100 pixels, com média de 32,46 regiões verdadeiras por imagem. A Tabela 3 traz a distribuição das imagens selecionadas conforme a classificação BI-RADS[®] para achados mamográficos e densidade mamária.

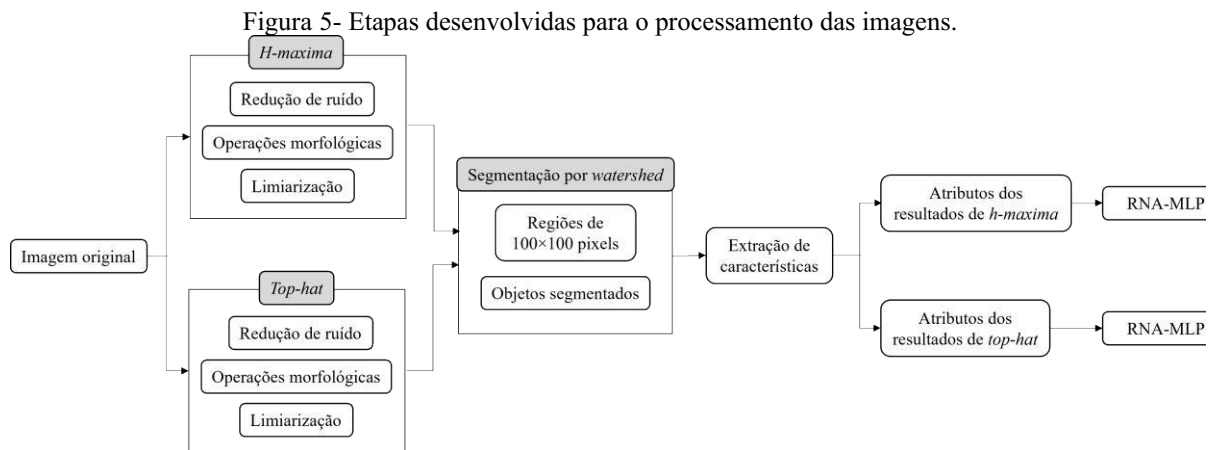
Tabela 3 - Imagens selecionadas conforme a classificação BI-RADS[®].

	Dens. A	Dens. B	Dens. C	Dens. D	Total
Categoria 2	40	39	33	14	126
Categoria 3	6	2	0	0	8
Categoria 4	0	4	5	2	11
Categoria 5	6	21	6	0	33
Total	52	66	44	16	178

Fonte: Elaborado pelo autor.

6.2. Processamento

Para o processamento digital das imagens, foram desenvolvidos algoritmos em linguagem Python utilizando bibliotecas para manipulação de imagens e matrizes como *numpy*, *scikit-image*, *open-cv*, *scipy* e *tensorflow*. O processamento foi realizado em etapas para realização de testes com diferentes métodos de pré-processamento para redução de ruído, variações na configuração das técnicas de detecção de objetos, segmentação de objetos com *watershed*, divisão das imagens em regiões de 100×100 pixels, extração de características e classificação dos resultados com redes neurais artificiais visando a diminuição no número de resultados do tipo falso positivo. A Figura 5 ilustra as etapas desenvolvidas neste trabalho para o processamento das imagens.



Fonte: Elaborado pelo autor.

Considerando que as microcalcificações aparecem como pequenos componentes de alta frequência, característica muito semelhante ao ruído aleatório, é importante que as imagens sejam tratadas de forma a reduzir o ruído antes da detecção de máximos regionais. Neste trabalho, foram utilizados dois métodos de redução de ruído, ou *denoising*, com o objetivo de avaliar como essas técnicas podem interferir na segmentação de microcalcificações, sendo eles o filtro de Wiener e a transformada *wavelet*.

Primeiramente, a imagem original foi normalizada, ou seja, os níveis de intensidade foram redistribuídos em uma escala entre 0 e 1. Após isso, o primeiro o filtro para redução de ruído aplicado foi o filtro de Wiener utilizando uma máscara de forma 3×3 pixels, tamanho escolhido seguindo os resultados de Costa Júnior (COSTA JÚNIOR, 2019) para imagens do INbreast. O filtro de Wiener busca minimizar o erro quadrático médio entre a imagem original e a imagem restaurada, considerando no cálculo os espectros de potência do ruído e da imagem (GONZALEZ; WOODS, 2010).

A transformada *wavelet*, utilizada como segundo método de *denoising*, é empregada na extração de coeficientes em diversos níveis de decomposição da imagem de acordo com a

família de *wavelets* definida. No processamento realizado, a imagem foi decomposta em um nível de decomposição resultando em quatro coeficientes, sendo três de detalhes nas direções horizontal, vertical e diagonal e um de aproximação utilizando a função *Coiflets5*. Para redução de ruído, os coeficientes de detalhes passaram por limiarização utilizando um limiar calculado com variância da imagem original e a distribuição Rayleigh. O resultado do processamento foi obtido pela transformada *wavelet* inversa utilizando os coeficientes limiarizados e o componente de aproximação (CARNEIRO, 2019; COSTA JÚNIOR, 2019).

Para avaliar a efetividade dos processamentos de *denoising*, a métrica relação sinal-ruído de pico, PSNR, foi calculada para cada método empregado. O PSNR calcula a relação entre as energias máximas do sinal e do ruído comparando a imagem original e a imagem processada. Valores maiores para o PSNR sugerem menor presença de ruído nas imagens (COSTA JÚNIOR, 2019).

As imagens processadas com os métodos de redução de ruído foram submetidas aos métodos de detecção de microcalcificações para obter imagens binárias contendo os marcadores internos utilizados na segmentação por *watershed*. As técnicas foram escolhidas considerando a necessidade de detecção de pequenos objetos com intensidade maior que a intensidade das regiões vizinhas, ou seja, com aspecto semelhante aos máximos regionais da imagem (TOUIL et al., 2021).

O primeiro método escolhido para detecção dos marcadores internos foi a operação morfológica conhecida como *h-maxima*. Esse processamento é dado pela reconstrução morfológica por dilatação da imagem f por $f-h$. O valor de h corresponde a diferença mínima de intensidade entre o pixel e seus vizinhos (SOILLE, 2004). Nesse trabalho, após testes preliminares, optou-se por realizar testes com elementos estruturantes do tipo disco com raios de tamanhos 3 e 11 pixels e valor de h igual a 0,02, 0,04 e 0,08, o que corresponde a uma diferença mínima de intensidade entre os pixels de aproximadamente 80, 160 e 320 níveis de cinza, respectivamente.

O próximo passo nesse processamento foi a subtração entre a imagem resultante da operação *h-maxima* e a imagem original, que resulta na imagem *h-convex*, composta de máximos regionais com h maior que o valor estabelecido (SOILLE, 2004). A imagem *h-convex* é limiarizada para obter uma versão binária que será utilizada como marcador interno na segmentação por *watershed*.

Utilizar a limiarização para obter uma imagem binária a partir de uma imagem em níveis de cinza representa um desafio, visto a necessidade da escolha de um limiar ideal, ou seja, encontrar um valor tal que as estruturas sejam binarizadas corretamente sem detecção de

objetos irrelevantes. Após testes com diversos limiares, entre eles o calculado pelo método de Otsu, optou-se pelo processamento usando a limiarização iterativa.

Na limiarização iterativa, a imagem é binarizada com diferentes limiares e, comparando a quantidade de componentes conectados obtidos na limiarização com cada valor de teste, é possível escolher o limiar capaz de produzir um número pré-estabelecido de objetos na imagem binária (KARALE et al., 2019). Na limiarização das imagens *h-convex*, optou-se por testar vários limiares começando pelo máximo valor de intensidade da imagem e seguindo em sequência decrescente com razão -0,003, o correspondente a uma variação de intensidade de aproximadamente 12 níveis de cinza. O algoritmo de teste foi interrompido quando o limiar testado produziu mais objetos que o máximo determinado, de forma que o valor escolhido como limiar para binarização da imagem correspondeu ao utilizado no teste anterior a interrupção. Neste trabalho foram testados três limiares como parâmetro de interrupção: 500, 750 e 1000 componentes conectados na imagem binária.

O segundo processamento escolhido para detecção de microcalcificações consistiu na operação *top-hat* com abertura por reconstrução morfológica. A imagem *top-hat* é obtida pela subtração entre a imagem original e o resultado da abertura por reconstrução morfológica (SOILLE, 2004). Nesse trabalho, foram testados processamentos com elementos estruturantes do tipo disco com raios de tamanho 3 e 11 pixels para produzir a imagem *top-hat*. Após isso, a imagem passou por limiarização iterativa para produção de imagens binárias com no máximo 500, 750 e 1.000 componentes conectados.

Ao todo, 48 combinações de processamentos foram realizadas para avaliar o papel de cada parâmetro na detecção dos marcadores para utilização na segmentação por *watershed*. Essas combinações foram obtidas variando os parâmetros descritos na Tabela 4.

Tabela 4 - Parâmetros de configuração para detecção dos marcadores.

Parâmetro	Descrição	Intervalo de teste
Pré-processamento	Processamento para redução de ruído ou aumento de contraste da imagem.	Filtro de Wiener e transformada <i>wavelet</i>
Deteção	Método de detecção de marcadores.	<i>H-maxima</i> e <i>top-hat</i>
Valor de <i>h</i>	Mínima diferença de intensidade na operação <i>h-maxima</i> .	0,02, 0,04 e 0,08
Elemento estruturante	Forma e tamanho do elemento estruturante utilizado nas operações morfológicas.	Tipo disco de raio 3 e 11 pixels
Limiarização	Máximo de componentes conectados permitidos na imagem binária.	500, 750 e 1000 objetos

Fonte: Elaborado pelo autor.

As imagens binárias produzidas foram processadas para eliminar os componentes detectados que poderiam causar supersegmentação no resultado. Considerando as limitações

quanto ao tamanho dos pixels e das microcalcificações, os objetos detectados com áreas menores que 4 pixels e maiores que 250 pixels foram eliminados das imagens binárias.

Após a detecção e seleção dos marcadores internos, os marcadores externos foram obtidos com as linhas de *watershed* do complemento da imagem original filtrada com fechamento e abertura morfológica, operações realizadas com um elemento estruturante de tamanho 3×3 pixels. Além disso, a segmentação por *watershed* da imagem filtrada foi realizada utilizando os marcadores internos detectados como mínimos regionais no processamento para obtenção do marcador externo. Por fim, para obter a versão final do marcador utilizado na segmentação por *watershed*, o marcador interno, que localiza os objetos, foi adicionado ao marcador externo, referente ao fundo da imagem (CIECHOLEWSKI, 2017).

Utilizando os marcadores detectados com as operações *h-maxima* e *top-hat*, a segmentação por *watershed* foi aplicada no gradiente morfológico da imagem para obter uma imagem contendo os componentes conectados e as linhas de *watershed* correspondentes as fronteiras desses objetos. Novamente, considerando as características das microcalcificações, as regiões segmentadas com área maior que 250 pixels foram eliminadas.

Para avaliação dos resultados da segmentação, a imagem com o resultado da segmentação por *watershed* foi dividida em regiões quadradas com tamanho 100×100 pixels, seguindo os mesmos procedimentos da padronização do *ground truth*. Cada região foi comparada com a região correspondente no *ground truth* e categorizada conforme a detecção ou não de objetos naquela região. Neste trabalho, a análise se deu da seguinte forma:

- Verdadeiro positivo: a região contém ao menos um objeto segmentado e o *ground truth* mostra a presença de microcalcificação na região;
- Verdadeiro negativo: não foram encontrados objetos na região segmentada e no *ground truth*;
- Falso positivo: a região possui ao menos um objeto segmentado, contudo, o *ground truth* não indica presença de microcalcificação na região;
- Falso negativo: não foram encontrados objetos na região segmentada e o *ground truth* aponta a existência de microcalcificação na região.

A partir dessa análise, foi possível montar uma matriz de confusão para cada imagem segmentada e extrair métricas de avaliação como a sensibilidade, especificidade, precisão, acurácia e área sob a curva ROC para cada exemplo. A avaliação da eficiência dos processamentos testados foi realizada analisando a média e o desvio padrão das métricas calculadas em cada caso.

Considerando a possibilidade de detecção de muitos falsos positivos, as regiões de 100×100 pixels categorizadas como resultados verdadeiro positivo e falso positivo foram utilizadas para treinamento de modelos de Redes Neurais Artificiais do tipo Perceptron Multicamadas (RNA-MLP) capazes de classificá-las e reduzir o número de regiões do tipo falso positivo nas imagens. Para isso, essas regiões passaram por uma etapa de extração de características, cujo objetivo principal é a obtenção de atributos que possam descrever as propriedades morfológicas, de intensidade e de textura das regiões de 100×100 pixels e dos objetos segmentados. As 113 características calculadas estão indicadas na Tabela 5, sendo que 68 características foram extraídas das regiões de 100×100 pixels e 45 atributos foram extraídos dos objetos segmentados em cada região.

Tabela 5 – Características extraídas das regiões analisadas.

Características das regiões de 100×100 pixels	Características dos objetos
1. Número de objetos	69. Área
2. Intensidade máxima	70. Área da <i>bounding box</i>
3. Intensidade mínima	71. Maior eixo da elipse
4. Intensidade média	72. Menor eixo da elipse
5. Desvio padrão da intensidade	73. Excentricidade
6. Diferença entre intensidades média e mínima	74. Área da elipse com diâmetro equivalente
7. Diferença entre intensidades média e máxima	75. Número de Euler
8. Diferença entre intensidades mínima e máxima	76. Extensão
9. <i>Skewness</i>	77. Diâmetro máximo de Feret
10. <i>Kurtosis</i>	78. Intensidade máxima
11. Moda	79. Intensidade mínima
12. N° de pixels com intensidade igual a moda	80. Intensidade média
13. N° de pixels com intensidade maior que a moda	81. Orientação
14. N° de pixels com intensidade menor que a moda	82. Perímetro
15-18. Segundo Momento Angular - ASM em quatro direções	83. Perímetro de Crofton
19-22. Contraste em quatro direções	84. Solidez
23-26. Correlação em quatro direções	85-91. 7 momentos de Hu
27-30. Soma da variância do quadrado em quatro direções	92-107. 16 momentos centrais
31-34. Momento de Diferença Inverso (IDM) em quatro direções	108. Razão entre a intensidade máxima do objeto e a intensidade máxima da região 100×100 pixels
35-38. Soma da média em quatro direções	109. Razão entre a intensidade média do objeto e a intensidade máxima da região 100×100 pixels
39-42. Soma da variância em quatro direções	110. Razão entre a intensidade mínima do objeto e a intensidade máxima da região 100×100 pixels
43-46. Soma da entropia em quatro direções	111. Diferença entre a intensidade máxima do objeto e a intensidade máxima da região 100×100 pixels
47-50. Entropia em quatro direções	112. Diferença entre a intensidade média do objeto e a intensidade máxima da região 100×100 pixels
51-54. Diferença da variância em quatro direções	113. Diferença entre a intensidade mínima do objeto e a intensidade máxima da região 100×100 pixels
55-58. Diferença da entropia	
61-64. Medida de Informação de Correlação 1 (IMC 1) em quatro direções	
65-68. Medida de Informação de Correlação 2 (IMC 2) em quatro direções	

Fonte: Elaborado pelo autor.

Os descritores de textura de Haralick (HARALICK, 1979) indicados nos itens 15 a 68 da Tabela 5 foram calculados utilizando a matriz de coocorrência de níveis de cinza, conhecida pela sigla GLCM (*Gray Level Co-occurrence Matrix*), considerando a distância de um pixel e os ângulos 0° , 45° , 90° e 180° (COELHO, 2013). Além disso, foram obtidas as médias de cada característica de textura utilizando os valores calculados em cada direção.

As características dos objetos segmentados foram calculadas com auxílio da biblioteca *scikit-image* e incluem descritores de forma e intensidade de componentes conectados, além de momentos centrais e momentos invariantes de Hu. Também foram obtidas as médias de cada atributo dos objetos para cada região de 100×100 pixels. A extração de características ocorreu em dois grupos de dados, cada um contendo os resultados do melhor teste para os processamentos de detecção *h-maxima* e *top-hat*.

Após essa etapa, as 178 imagens segmentadas foram divididas em grupos de treinamento e teste com tamanhos 124 e 54 imagens, respectivamente. As características extraídas do conjunto de imagens de treinamento representam os dados utilizados para o treinamento e validação dos modelos de redes neurais. Já o conjunto de imagens de teste contém as novas amostras que serão aplicadas para testar e avaliar os modelos finais. A Tabela 6 traz o número de imagens e regiões de 100×100 pixels conforme a divisão dos dados em conjuntos de treinamento e teste. Esses dados são referentes aos resultados obtidos com pré-processamento utilizando a transformada *wavelet* com as seguintes configurações de detecção: *h-maxima* com valor de h igual a 0,04, elemento estruturante com raio de 3 pixels e máximo de 1000 objetos na imagem binária; e *top-hat*, também com elemento estruturante de raio igual a 3 pixels e máximo de 1000 objetos na imagem binária

Tabela 6 - Divisão dos dados em conjuntos de treinamento e teste.

Item	<i>H-maxima</i>		<i>Top-hat</i>	
	Treinamento	Teste	Treinamento	Teste
Nº de imagens	124	54	124	54
Nº de regiões	30.983	13.867	27.710	12.181
Nº de regiões VP	3.483	1.519	3.438	1.474
Nº de regiões FP	27.500	12.348	24.272	10.707

Fonte: Elaborado pelo autor.

Com a divisão dos conjuntos, os dados de treinamento foram analisados para seleção dos atributos que serão utilizados como entrada dos modelos de RNA-MLP. Considerando a presença de mais regiões da classe falso positivo (com alvo igual a zero) do que da classe verdadeiro positivo (com alvo igual a um), foi necessário balancear os conjuntos de treinamento. Para isso, no caso das imagens processadas com *h-maxima*, foram selecionadas aleatoriamente 3.483 regiões da classe falso positivo de forma a igualar a quantidade de dados

ao número total de regiões da classe verdadeiro positivo. Já para as imagens processadas com *top-hat*, o processo de balanceamento foi feito pela seleção de 3.438 regiões da classe falso positivo. Portanto, o número total de regiões nos conjuntos de treinamento dos dados foi igual a 6.966 regiões no caso dos dados obtidos com detecção por *h-maxima* e 6.876 regiões para as imagens processadas com *top-hat*.

A seleção de atributos foi realizada com o conjunto de treinamento balanceado através da análise de medidas estatísticas descritivas, dos histogramas e da matriz de correlação das características extraídas. Além disso, transformações de variáveis através de métodos matemáticos, como a normalização e a transformação logarítmica, foram realizadas em alguns casos para identificar e aprimorar possíveis relações relevantes entre os dados. Ao final do processo, foram selecionados 16 atributos com base nos maiores valores absolutos de correlação entre os atributos e a saída esperada, ou seja, zero para regiões do tipo falso positivo e um para regiões verdadeiro positivo. É importante ressaltar que os atributos selecionados para o treinamento também são utilizados como características de entrada durante os testes, sendo necessário garantir a adequação dos dados de teste as configurações de entrada definidas no treinamento dos modelos.

Considerando as duas fontes de dados, as regiões dos tipos verdadeiro positivo e falso positivo resultantes da segmentação com detecção *h-maxima* e *top-hat*, foram construídos dois modelos de RNA-MLP, utilizando os mesmos atributos como entrada e diferentes configurações dos parâmetros de construção das redes. A saída das redes foi definida como a probabilidade de uma região classificada ser um verdadeiro positivo em todos os modelos testados. Outros parâmetros de configuração dos modelos, como o número de neurônios nas camadas ocultas e a taxa de aprendizagem, foram testados com diferentes valores aleatórios até a seleção de um modelo satisfatório. Os parâmetros de configuração dos modelos com melhor performance estão indicados na Tabela 7.

Tabela 7 - Parâmetros de configuração das redes neurais artificiais.

Parâmetros	<i>H-maxima</i>	<i>Top-hat</i>
Nº de nós da camada de entrada	16	
Nº de camadas ocultas	3	5
Nº de nós das camadas ocultas	63	40
Nº de nós da camada de saída	1	
Função de ativação	Função sigmoide	
Taxa de aprendizagem	$5,174 \times 10^{-5}$	$6,237 \times 10^{-5}$
Otimizador	Adam	
β_1	0,127	0,214
β_2	0,206	0,013

Fonte: Elaborado pelo autor.

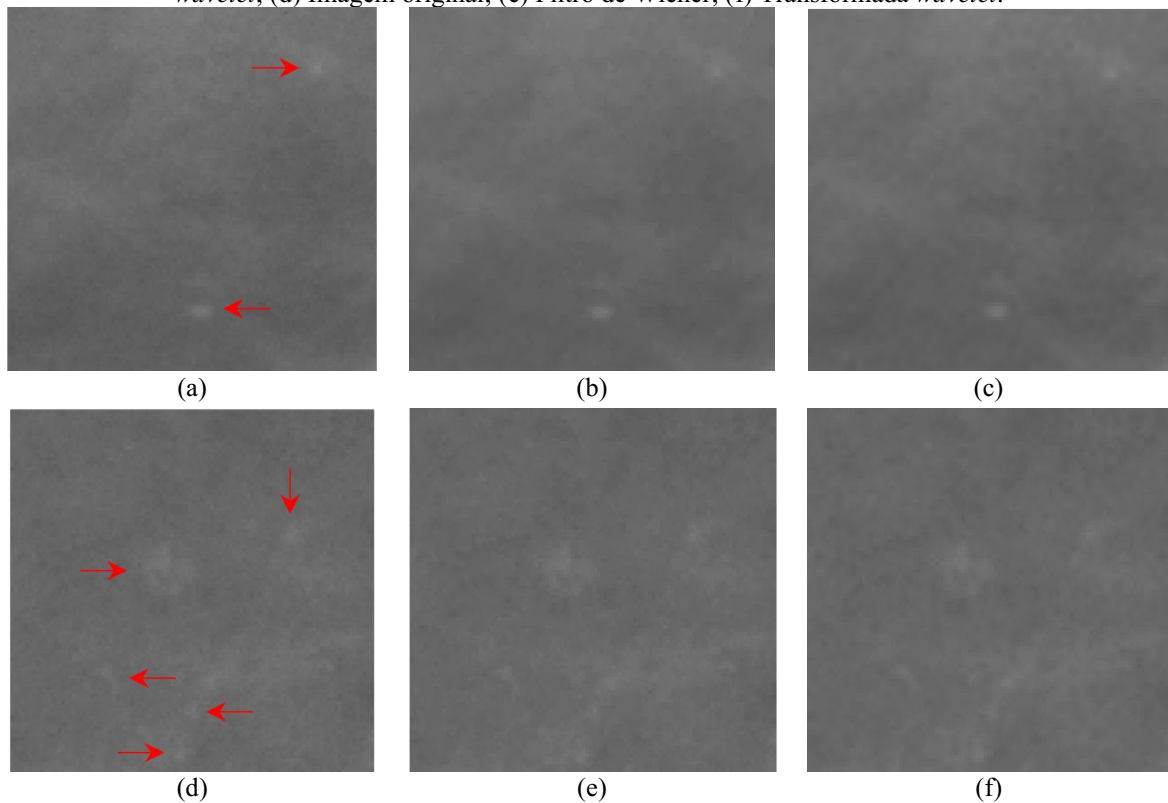
Uma das formas de definir se um modelo é satisfatório é através do uso de um conjunto de validação durante o treinamento da rede. Neste trabalho, optou-se pelo método de validação cruzada, também conhecido por *k-fold cross validation*, o qual consiste na reutilização de dados do conjunto de treinamento para compor subconjuntos de validação. O conjunto de dados de treinamento é dividido em k partes iguais com o objetivo de produzir k estimativas de performance, sendo que em cada estimativa $k - 1$ partes são utilizadas no treinamento e a parte restante é aplicada na validação do modelo (SANTOS et al., 2019).

Após os testes com parâmetros aleatórios e a escolha dos modelos finais, as imagens separadas para teste foram preparadas para passar por classificação. Nesse caso, os dados foram agrupados por imagem e cada grupo foi utilizado para testar os modelos, resultando em 54 testes para cada modelo. Apesar de não passarem pelo balanceamento no número de regiões categorizadas como falso positivo, os subconjuntos de teste foram adaptados às entradas dos modelos de acordo com os procedimentos realizados no treinamento das redes. Por fim, após os testes dos modelos finais, os resultados foram avaliados pela análise dos dados observados durante o treinamento e a validação das redes e com o cálculo de métricas como a sensibilidade, especificidade, acurácia, precisão e área da curva ROC.

7. RESULTADOS E DISCUSSÃO

As 178 imagens originais foram pré-processadas com os algoritmos de redução de ruído antes da detecção de marcadores internos com *h-maxima* e *top-hat*. Visto o interesse em verificar o efeito dos métodos de *denoising* em regiões com possíveis microcalcificações, a Figura 6 traz exemplos dos resultados do pré-processamento das imagens em regiões de 100×100 pixels.

Figura 6 - Resultados do pré-processamento. (a) Imagem original; (b) Filtro de Wiener; (c) Transformada *wavelet*; (d) Imagem original; (e) Filtro de Wiener; (f) Transformada *wavelet*.



Fonte: Elaborado pelo autor.

As setas em vermelho na Figura 6 indicam a localização das microcalcificações nos exemplos apresentados. É possível perceber pequenas mudanças no aspecto das imagens após a filtragem para redução de ruído, sendo que as imagens processadas com filtro de Wiener apresentaram aspecto com mais borramento do que as filtradas com a transformada *wavelet*. A métrica relação sinal-ruído de pico (PSNR) foi calculada para cada imagem processada e, para avaliar a capacidade de redução de ruído dos dois pré-processamentos, os valores de média e desvio padrão foram calculados e apresentados na Tabela 8.

Tabela 8 - Média e desvio padrão do PSNR das imagens processadas.

Pré-processamento	PSNR	
	Média	Desvio
<i>Wavelet</i>	123,31	2,71
Wiener	125,75	2,77

Fonte: Elaborado pelo autor.

Os valores na Tabela 8 indicam que o pré-processamento com o filtro de Wiener foram mais efetivos no *denoising* das imagens, visto que a média do PSNR das imagens tratadas com esse método é maior que a média para imagens processadas com a transformada *wavelet*. O efeito de borramento causado pelo filtro de Wiener na imagem pode ser aumentado conforme o tamanho da máscara utilizada, o que pode causar a diminuição do ruído na imagem e o aumento do PSNR. Entretanto, com o aumento do borramento, microcalcificações muito pequenas podem desaparecer, sendo importante equilibrar a capacidade de borramento dos métodos de redução de ruído e a perda de informação na imagem.

As imagens processadas com as técnicas de redução de ruído foram utilizadas na segmentação por *watershed* limitada por marcadores pré-definidos. Esses marcadores estabelecem pixels de referência pertencentes aos objetos de interesse e guiam o algoritmo de produção das *watershed*. O primeiro método utilizado para detecção de objetos de interesse para composição dos marcadores foi a operação morfológica *h-maxima*. Para aplicá-la foi necessário definir o valor de h , a diferença mínima entre uma região e os pixels vizinhos, e o elemento estruturante utilizado na reconstrução por dilatação morfológica (SOILLE, 2004).

Pela subtração entre a imagem original e o resultado da operação *h-maxima*, obtém-se a imagem *h-convex* que, quando limiarizada, representa objetos que correspondem aos máximos regionais da imagem (SOILLE, 2004). Para realizar a limiarização, deve-se escolher um valor de intensidade para binarizar a imagem e eliminar os elementos com intensidade abaixo do limiar definido. Neste trabalho, utilizou-se a limiarização iterativa para binarizar as imagens e limitar o número de componentes conectados na imagem binária. A Tabela 9 traz os resultados dos testes de detecção de objetos utilizando o método *h-maxima* com variação no valor de h , no tamanho do elemento estruturante (SE) e do número de objetos permitidos na limiarização. Para comparar os resultados, utilizou-se a média e o desvio padrão do número de marcadores detectados por imagem (MI) e o número de imagens nas quais não foram detectados objetos (N/A).

Os dados da Tabela 9 mostram que os testes realizados com valor de h igual a 0,02 (cerca de 80 níveis de cinza) resultaram em imagens sem objetos detectados. Isso ocorreu porque quanto menor o valor de h , mais máximos regionais são encontrados e a limiarização iterativa não foi bem-sucedida em encontrar um limiar cujo número de componentes conectados binarizados fosse menor que os limites estabelecidos. Já com o elemento estruturante de raio 11 pixels, os testes realizados com valor de h igual a 0,02 resultaram em objetos detectados. Isso se deve ao uso de um elemento estruturante maior, o que aumenta a região de análise durante a detecção e, no caso das imagens analisadas, causa a detecção de menos objetos,

permitindo a seleção de limiares que produzam o número máximo de objetos desejados na imagem binária (SOILLE, 2004).

Tabela 9 - Resultados da detecção de objetos com *h-maxima* considerando a média e desvio padrão do número de marcadores por imagem (MI) e resultados sem detecção (N/A).

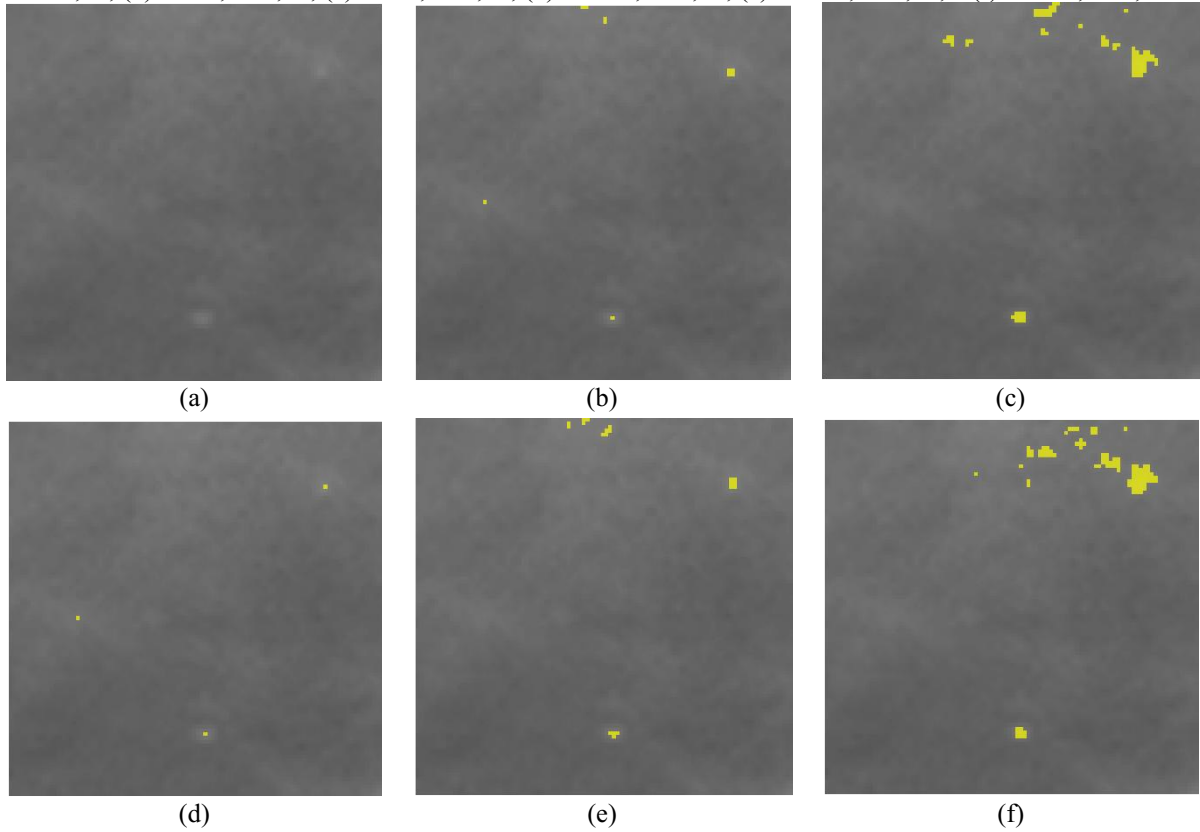
SE	Valor de <i>h</i>	Limiar	<i>Wavelet</i>			Wiener		
			MI - Média	MI - STD	N/A	MI - Média	MI - STD	N/A
3	0,02	500	13,72	80,97	173	16,16	86,82	172
3	0,02	750	48,46	180,81	166	61,21	202,37	163
3	0,02	1000	153,39	353,07	158	130,69	332,05	154
3	0,04	500	478,42	63,44	3	467,30	95,16	7
3	0,04	750	730,28	14,38	0	731,30	11,28	0
3	0,04	1000	975,38	15,27	0	973,88	14,64	0
3	0,08	500	481,18	8,27	0	483,83	10,20	0
3	0,08	750	719,11	11,56	0	723,76	14,64	0
3	0,08	1000	956,63	14,73	0	962,13	18,76	0
11	0,02	500	309,28	232,54	64	285,62	238,93	73
11	0,02	750	710,21	109,08	4	695,16	151,75	8
11	0,02	1000	970,78	21,26	0	966,07	74,80	1
11	0,04	500	488,47	7,58	0	488,10	7,51	0
11	0,04	750	732,63	10,78	0	730,64	12,02	0
11	0,04	1000	971,37	16,86	0	969,85	17,11	0
11	0,08	500	477,26	7,87	0	480,12	10,21	0
11	0,08	750	716,79	10,73	0	718,98	13,97	0
11	0,08	1000	955,26	14,40	0	958,48	17,84	0

Fonte: Elaborado pelo autor,

Quanto ao número de objetos, no caso dos testes com detecção bem-sucedida, observou-se que a média de objetos detectados por imagem tem valor próximo ao número máximo de objetos permitidos na limiarização. Além disso, os resultados na Tabela 9 indicam que os testes realizados com valor de *h* igual a 0,04 (aproximadamente 160 níveis de cinza) detectaram mais objetos do que os testes com valor igual a 0,08 (cerca de 320 níveis de cinza). Esses dados estão de acordo com os resultados esperados, visto que o aumento da mínima diferença de intensidade entre os pixels indica a seleção de objetos com maior destaque em relação ao fundo da imagem e eliminação de componentes com baixo contraste (SOILLE, 2004).

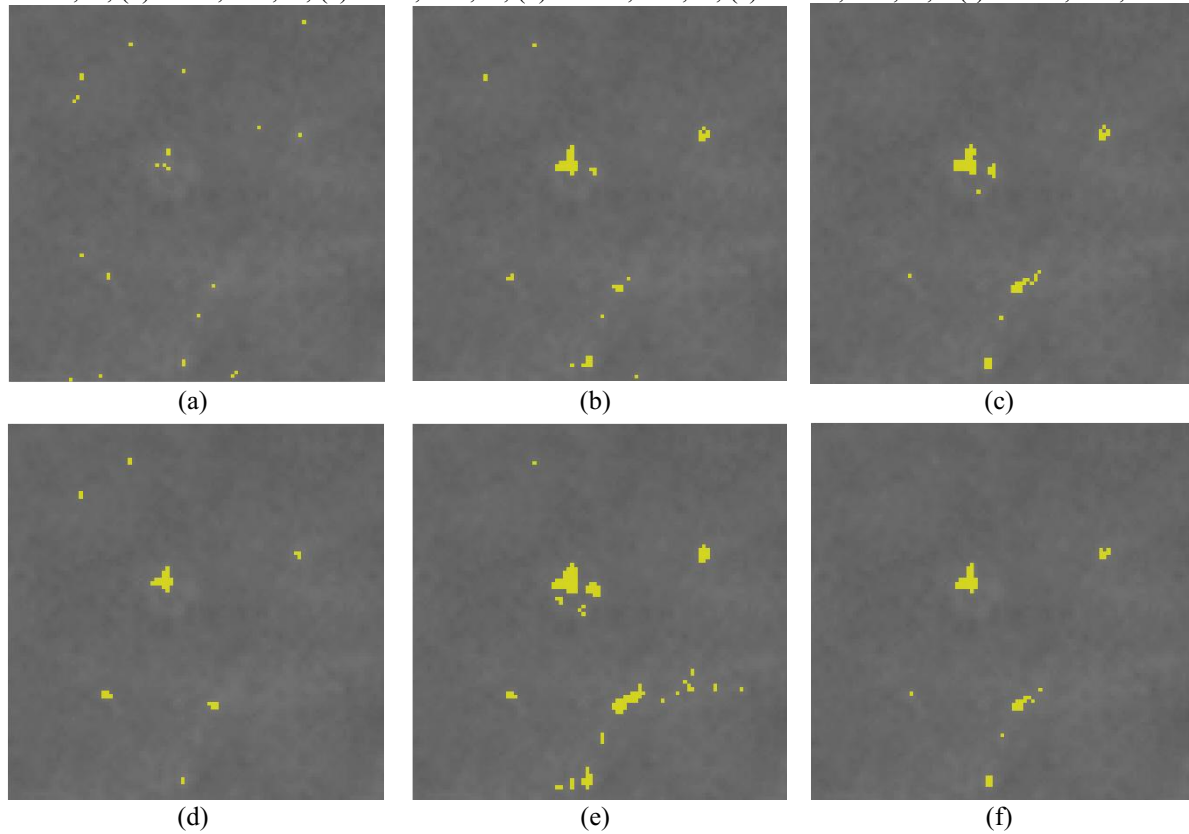
O elemento estruturante permite especificar a forma das regiões analisadas na detecção, ou seja, dependendo do formato e tamanho escolhidos diferentes objetos podem ser detectados. As Figuras 7 e 8 exemplificam a mudança na forma e no número de objetos detectados de acordo com os diferentes valores de *h* e elementos estruturantes (SE) testados. Neste exemplo, utilizou-se a transformada *wavelet* como pré-processamento e definiu-se 1000 objetos como máximo de componentes conectados na imagem binária.

Figura 7 – Resultados dos testes de detecção com *h-maxima* e máximo de 1000 objetos detectados: (a) SE=3, $h=0,02$; (b) SE=3, $h=0,04$; (c) SE=3, $h=0,08$; (d) SE=11, $h=0,02$; (e) SE=11, $h=0,04$; e (f) SE=11, $h=0,08$.



Fonte: Elaborado pelo autor.

Figura 8 - Resultados dos testes de detecção com *h-maxima* com máximo de 1000 objetos detectados: (a) SE=3, $h=0,02$; (b) SE=3, $h=0,04$; (c) SE=3, $h=0,08$; (d) SE=11, $h=0,02$; (e) SE=11, $h=0,04$; e (f) SE=11, $h=0,08$.

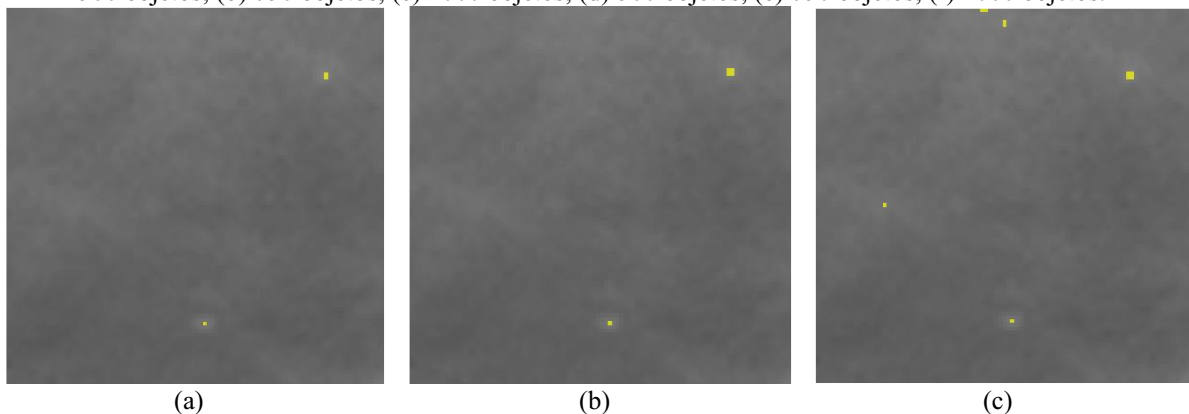


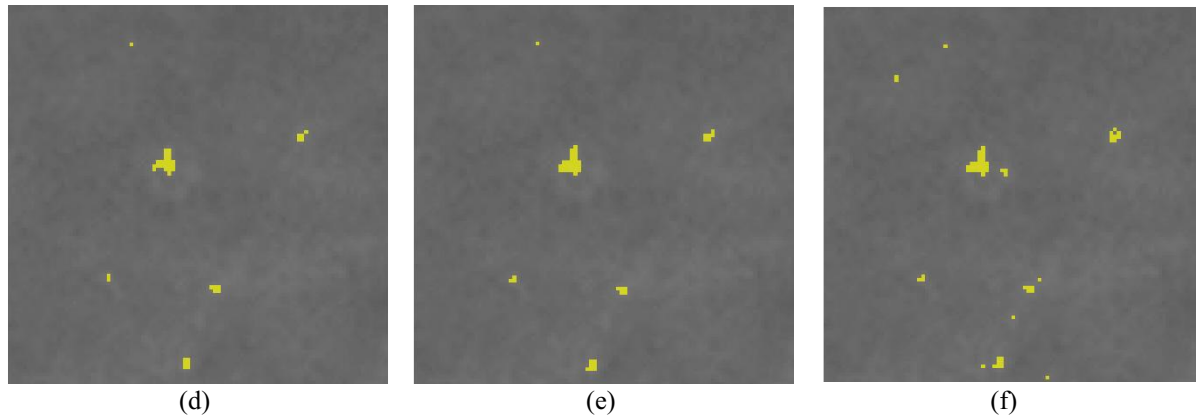
Fonte: Elaborado pelo autor.

Nos exemplos das Figuras 7 e 8 observa-se que o número de pixels dos objetos e de componentes conectados é alterado conforme a mudança no valor de h e no tamanho do elemento estruturante utilizados na detecção com h -maxima. Combinações com diferentes tamanhos de elemento estruturante e valores de h criam mecanismos distintos de seleção de pixels, permitindo que em cada teste as relações entre os objetos e as regiões vizinhas nas imagens seja analisada de uma forma diferente. O tamanho do elemento estruturante delimita a região que deve possuir diferença de intensidade mínima igual a h em relação aos pixels vizinhos, sendo que valores maiores podem auxiliar na detecção de objetos de forma mais uniforme e reduzir a detecção de objetos com apenas um pixel.

Além disso, o limiar utilizado para binarizar a imagem h -convex também pode causar diferenças no número de objetos detectado e na forma dessas regiões. Na limiarização iterativa, são testados vários limiares para binarização, de forma que a cada iteração diferentes pixels são eliminados, criando diferentes configurações de objetos na imagem binária. Dessa forma, quanto menor o nível de cinza estabelecido para limiarização, mais pixels são adicionados a imagem binária, provocando aumento no número de componentes conectados e expandindo regiões que já haviam sido detectadas. A Figura 9 exemplifica esse processo com os resultados obtidos nos testes com diferentes números de objetos permitidos no resultado da limiarização da imagem h -convex. Neste exemplo, os objetos foram detectados na imagem pré-processada com a transformada *wavelet* utilizando valor de h igual a 0,04 e elemento estruturante com tamanho de três pixels.

Figura 9 - Resultados dos testes de detecção com h -maxima variando o número máximo objetos detectados: (a) 500 objetos; (b) 750 objetos; (c) 1000 objetos; (d) 500 objetos; (e) 750 objetos; (f) 1000 objetos.





Fonte: Elaborado pelo autor.

O segundo processamento realizado para detecção de objetos de interesse foi a limiarização da imagem obtida pela operação morfológica *top-hat*. A imagem *top-hat* foi obtida pela subtração da imagem original e o resultado da abertura por reconstrução com diferentes tamanhos de elemento estruturante. Novamente, a limiarização iterativa foi utilizada para binarizar a imagem considerando diferentes valores como número máximo de componentes conectados na imagem binária. A Tabela 10 traz os resultados dos testes de detecção de objetos utilizando o método *top-hat* com variação no tamanho do elemento estruturante e do número de objetos permitidos na limiarização.

Tabela 10 - Resultados da detecção de objetos com *top-hat* considerando a média e desvio padrão do número de marcadores por imagem (MI) e resultados sem detecção (N/A).

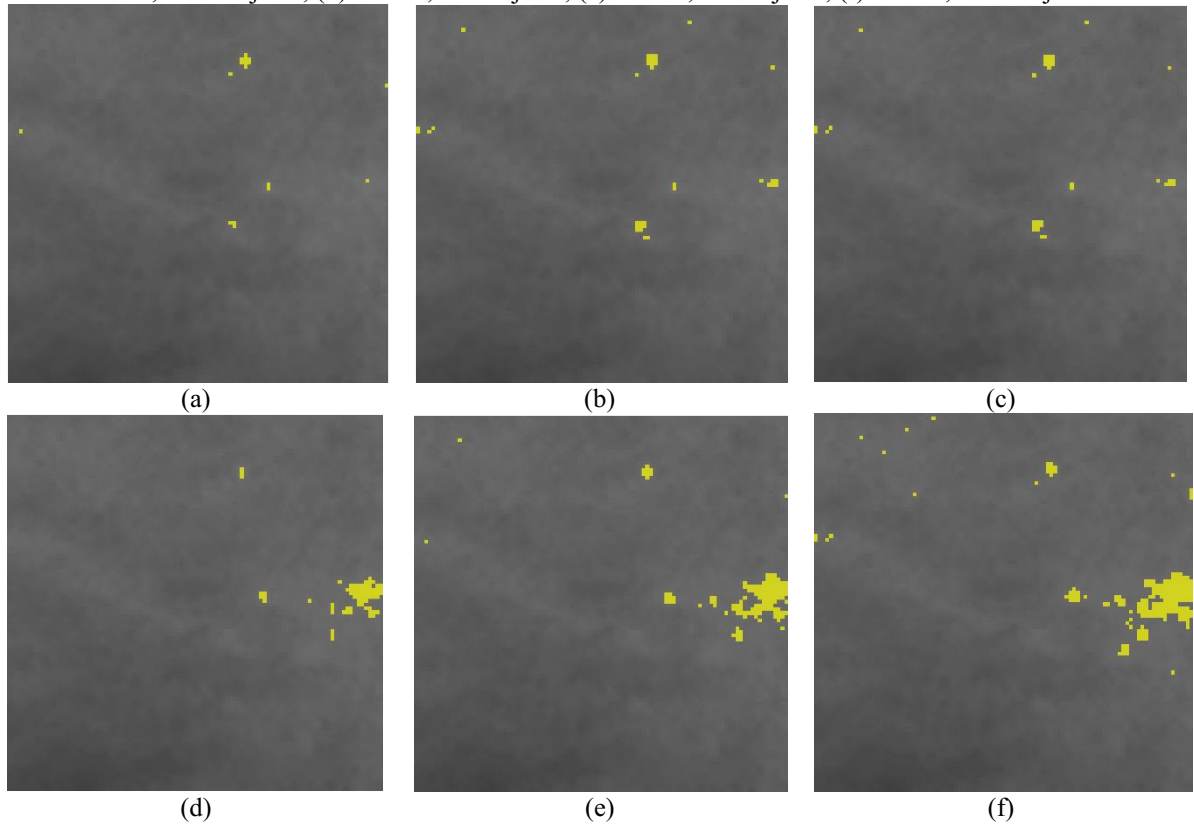
SE	Limiar	Wavelet			Wiener		
		Nº MI - Média	Nº MI - STD	N/A	Nº MI - Média	Nº MI - STD	N/A
3	500	373,18	67,85	0	389,27	56,72	0
3	750	572,63	91,99	0	593,19	94,36	0
3	1000	748,23	140,62	0	761,20	122,36	0
11	500	426,70	33,82	0	428,46	34,16	0
11	750	644,87	50,44	0	645,62	46,14	0
11	1000	856,49	66,20	0	859,79	61,19	0

Fonte: Elaborado pelo autor.

Os dados da Tabela 10 mostram que todos os testes de detecção com *top-hat* resultaram em imagens binárias com objetos. Além disso, dado que a operação *top-hat* analisa o formato dos objetos em relação ao elemento estruturante, com a mudança desse parâmetro diferentes regiões são detectadas, pois a tendência desse processamento é a eliminação de objetos com tamanho menor que o elemento estruturante (SOILLE, 2004).

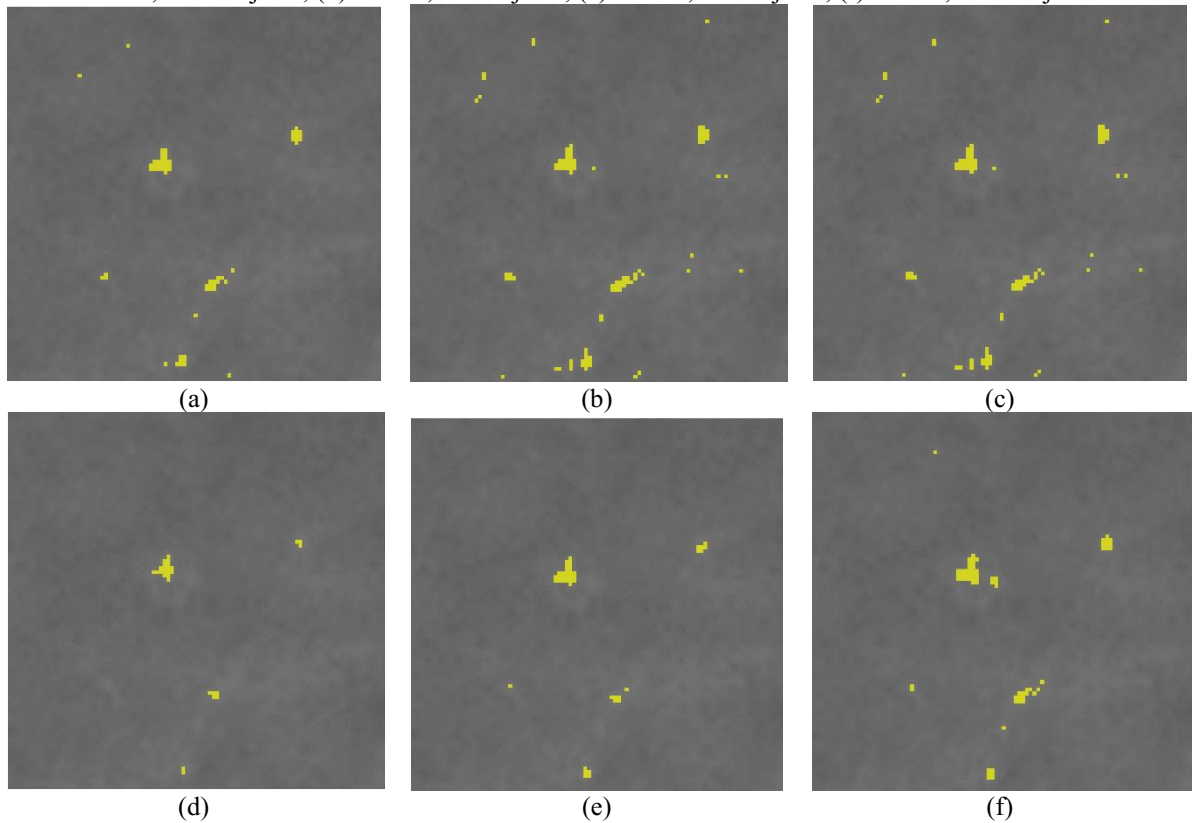
As Figuras 10 e 11 apresentam os resultados obtidos nos testes de detecção com diferentes números de objetos permitidos na limiarização da imagem *top-hat* e a mudança no tamanho do elemento estruturante (SE). Esses exemplos demonstram os testes utilizando o pré-processamento com transformada *wavelet*.

Figura 10 - Resultados dos testes de detecção com *top-hat*: (a) SE=3, 500 objetos; (b) SE=3, 750 objetos; (c) SE=3, 1000 objetos; (d) SE=11, 500 objetos; (e) SE=11, 750 objetos; (f) SE=11, 1000 objetos.



Fonte: Elaborado por autor.

Figura 11 - Resultados dos testes de detecção com *top-hat*: (a) SE=3, 500 objetos; (b) SE=3, 750 objetos; (c) SE=3, 1000 objetos; (d) SE=11, 500 objetos; (e) SE=11, 750 objetos; (f) SE=11, 1000 objetos.



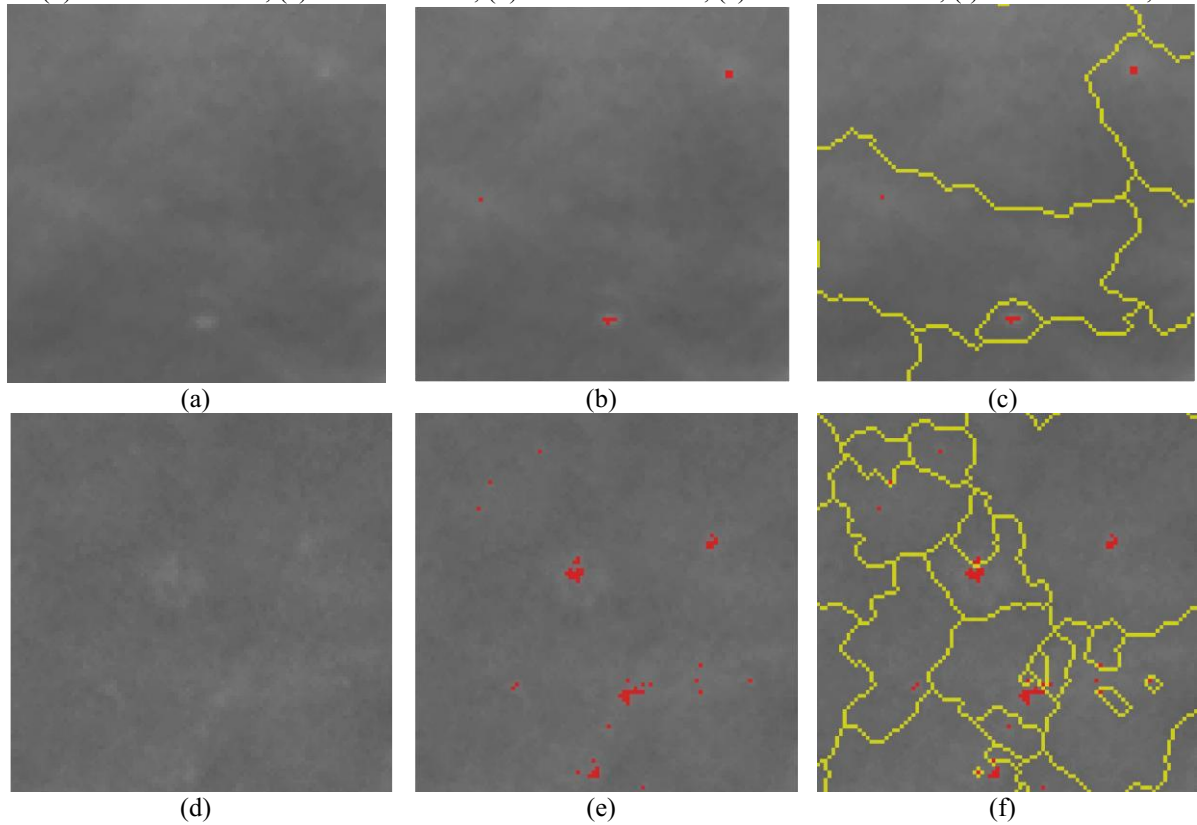
Fonte: Elaborado por autor.

A operação *top-hat* é baseada na abertura morfológica da imagem, a qual tem como primeira etapa a erosão das regiões, ou seja, o processamento primeiro atenua regiões menores que o elemento estruturante e depois ressalta os objetos restantes através da dilatação da imagem. Observa-se nas Figuras 10 e 11 que objetos pequenos foram eliminados com o aumento do elemento estruturante. Também é possível perceber que, com o aumento do número de objetos permitidos na imagem binária, mais pixels foram adicionados na imagem, mudando o formato das regiões e alterando o número de objetos detectados. Algumas regiões pequenas podem permanecer na imagem binária devido a limiarização iterativa, mesmo quando o elemento estruturante de tamanho 11 pixels é utilizado, sendo possível a expansão desses marcadores diminuindo o valor do limiar utilizado na binarização e, conseqüentemente, aumentando o número de objetos detectados na imagem binária.

Após a detecção dos marcadores internos, as imagens passaram por processamento para identificar marcadores externos relativos ao fundo da imagem. Os marcadores externos limitam a expansão das regiões iniciada a partir dos marcadores internos durante a segmentação por *watershed*. Neste trabalho, os marcadores externos foram obtidos com as linhas de *watershed* obtidas na segmentação do complemento da imagem filtrada por abertura e fechamento morfológico. A filtragem com abertura e fechamento tem por objetivo suavizar os máximos da imagem, de forma que o resultado represente as regiões uniformes do fundo dos objetos (CIECHOLEWSKI, 2017).

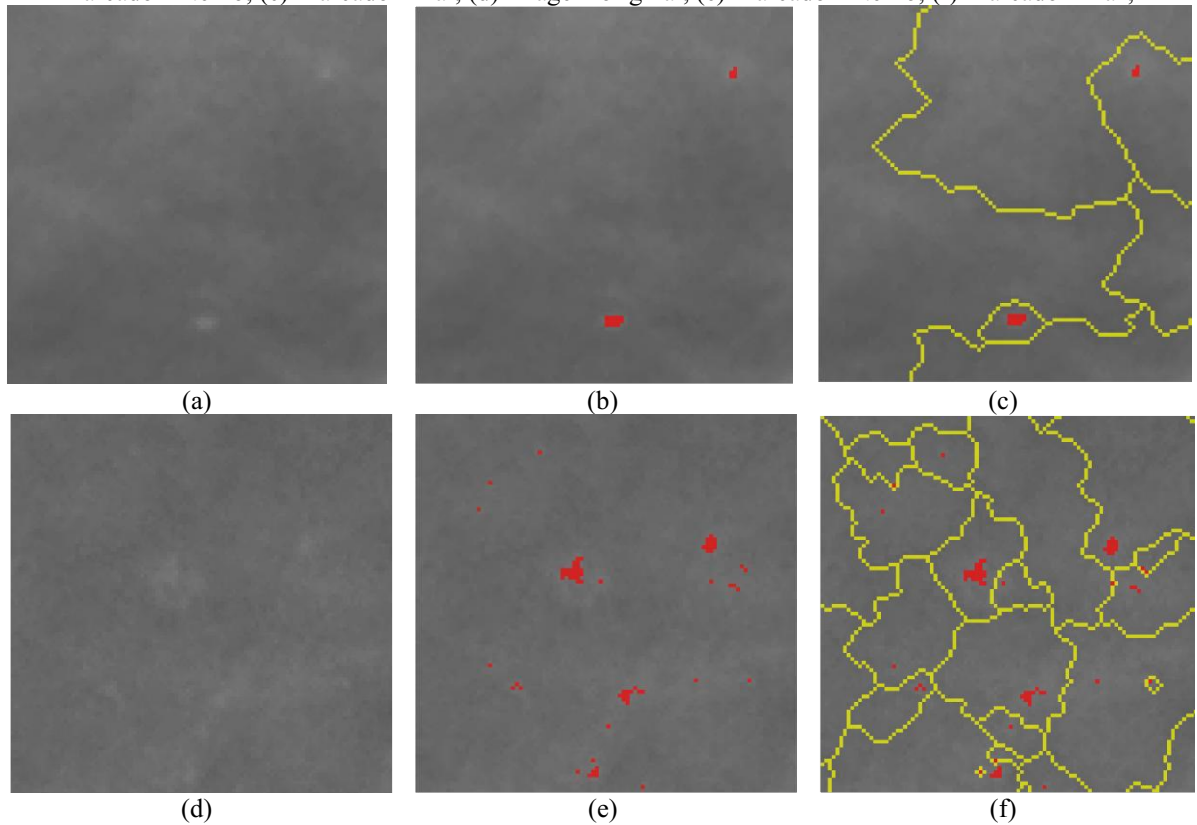
O resultado da combinação dos marcadores internos e externos pode ser verificado nas Figuras 12 e 13 nas quais os marcadores internos estão indicados em vermelho e os externos em amarelo. Na Figura 12 os exemplos foram pré-processados com o filtro de Wiener e a detecção de marcadores internos foi realizada com o método *h-maxima* com elemento estruturante de tamanho 3 pixels, valor de *h* igual a 0,04 e máximo de 1000 objetos na imagem binária. A Figura 13 apresenta marcadores detectados com a operação *top-hat* utilizando um elemento estruturante de tamanho 3 pixels, máximo de 1000 objetos no resultado da limiarização, além de pré-processamento com o filtro de Wiener.

Figura 12 – Combinação dos marcadores internos e externos na detecção com *h-maxima*: (a) Filtro de Wiener; (b) Marcador interno; (c) Marcador final; (d) Filtro de Wiener; (e) Marcador interno; (f) Marcador final,



Fonte: Elaborado pelo autor.

Figura 13 - Combinação dos marcadores internos e externos na detecção com *top-hat*: (a) Imagem original; (b) Marcador interno; (c) Marcador final; (d) Imagem original; (e) Marcador interno; (f) Marcador final,



Fonte: Elaborado pelo autor.

Utilizando os marcadores internos e externos, a segmentação por *watershed* foi realizada no gradiente morfológico das imagens pré-processadas e as métricas de avaliação foram calculadas para verificar a performance dos processamentos. O resultado da segmentação foi dividido em regiões de 100×100 pixels e comparado com o *ground truth* para categorização das regiões em verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo. A partir desses dados, foram calculadas a sensibilidade, especificidade, acurácia, precisão e a área da curva ROC (AUC) para cada imagem. A Tabela 11 traz a média (M) e o desvio padrão (STD) das métricas calculadas em cada teste realizado na detecção com *h-maxima*.

Tabela 11 - Média e desvio padrão das métricas de avaliação da segmentação das imagens com detecção por *h-maxima*.

Configuração			Sensibilidade		Especificidade		Acurácia		Precisão		AUC	
SE	Valor de <i>h</i>	Limiar	M	STD	M	STD	M	STD	M	STD	M	STD
Filtro de Wiener												
3	0,02	500	2,49	14,00	98,29	10,82	91,07	11,47	16,97	12,31	0,51	≈ 0,00
3	0,02	750	7,12	23,78	95,27	18,02	89,20	15,69	19,79	13,10	0,53	≈ 0,00
3	0,02	1000	11,94	30,66	92,53	22,42	87,47	18,39	19,30	11,98	0,55	≈ 0,00
3	0,04	500	62,40	25,04	77,73	16,08	76,84	14,77	16,30	11,74	0,71	≈ 0,00
3	0,04	750	72,61	19,56	69,06	19,12	70,11	16,96	14,20	10,60	0,74	≈ 0,00
3	0,04	1000	78,36	17,17	63,01	21,65	65,18	18,87	13,06	9,95	0,75	≈ 0,00
3	0,08	500	61,06	21,83	82,75	12,07	80,86	12,32	18,96	13,06	0,72	≈ 0,00
3	0,08	750	68,14	20,86	77,16	15,21	76,60	14,26	16,83	12,07	0,74	≈ 0,00
3	0,08	1000	73,66	19,42	72,29	17,70	72,80	15,91	15,52	11,35	0,75	≈ 0,00
11	0,02	500	38,33	35,65	82,85	21,22	80,56	18,32	15,57	10,65	0,63	≈ 0,00
11	0,02	750	64,97	23,96	67,74	22,97	68,80	19,82	12,87	9,32	0,71	≈ 0,00
11	0,02	1000	74,34	19,25	59,57	24,78	62,18	21,12	11,80	8,86	0,73	≈ 0,00
11	0,04	500	63,95	20,28	75,36	15,42	74,82	14,78	15,06	10,89	0,71	≈ 0,00
11	0,04	750	71,13	19,13	68,06	19,14	69,05	17,19	13,50	10,11	0,73	≈ 0,00
11	0,04	1000	77,47	17,15	62,37	21,80	64,55	18,98	12,67	9,58	0,74	≈ 0,00
11	0,08	500	57,63	22,05	83,49	11,87	81,35	12,22	18,84	13,05	0,71	≈ 0,00
11	0,08	750	65,51	20,76	77,83	15,10	77,07	14,09	16,66	11,62	0,73	≈ 0,00
11	0,08	1000	70,27	19,20	72,73	17,93	73,05	15,96	15,11	10,79	0,74	≈ 0,00
Transformada Wavelet												
3	0,02	500	2,63	15,53	98,92	8,37	91,48	10,03	20,31	16,14	0,51	≈ 0,00
3	0,02	750	6,62	24,69	96,70	15,38	90,06	13,68	18,42	12,64	0,52	≈ 0,00
3	0,02	1000	14,46	34,59	90,56	26,09	85,77	21,20	14,78	10,20	0,55	≈ 0,00
3	0,04	500	78,09	19,12	65,25	19,55	66,80	17,52	13,66	10,59	0,75	≈ 0,00
3	0,04	750	85,44	13,48	57,28	22,47	60,37	19,65	12,47	9,89	0,76	≈ 0,00
3	0,04	1000	89,27	11,20	52,35	24,65	56,39	21,17	11,84	9,42	0,76	≈ 0,00
3	0,08	500	67,47	20,55	78,35	13,0	77,34	12,94	17,04	12,35	0,73	≈ 0,00
3	0,08	750	74,59	18,19	71,60	16,67	72,13	15,40	15,21	11,35	0,75	≈ 0,00
3	0,08	1000	79,56	15,80	66,02	19,11	67,69	17,01	13,99	10,66	0,76	≈ 0,00
11	0,02	500	54,70	42,25	71,64	28,88	71,83	23,84	13,16	10,08	0,66	≈ 0,00
11	0,02	750	86,38	16,73	50,85	26,82	55,06	22,89	11,37	8,88	0,75	≈ 0,00
11	0,02	1000	91,43	9,50	44,86	28,03	50,05	23,73	10,78	8,67	0,75	≈ 0,00

11	0,04	500	74,64	17,31	67,92	17,76	68,87	16,40	13,82	10,33	0,74	≈ 0,00
11	0,04	750	81,20	14,92	60,27	21,43	62,75	18,88	12,56	9,61	0,75	≈ 0,00
11	0,04	1000	86,56	12,63	54,82	23,84	58,38	20,45	11,92	9,12	0,76	≈ 0,00
11	0,08	500	64,67	20,66	79,55	12,48	78,23	12,68	17,04	11,85	0,72	≈ 0,00
11	0,08	750	70,82	18,98	72,85	16,12	73,01	14,98	14,99	10,95	0,74	≈ 0,00
11	0,08	1000	76,23	17,31	67,06	18,92	68,47	16,80	13,79	10,27	0,75	≈ 0,00

Fonte: Elaborado pelo autor.

Os dados apresentados na Tabela 11 mostram que, no geral, os testes com maiores médias para sensibilidade e AUC também resultaram nas menores médias de especificidade, acurácia e precisão. Esses resultados apontam que apesar de altas taxas de acerto, grande parte das regiões de 100×100 pixels com objetos segmentados são, na verdade, falsos positivos.

Além disso, observa-se que numa mesma configuração de detecção os resultados de sensibilidade foram melhores nas imagens pré-processadas com a transformada *wavelet*. A capacidade de borrimento do pré-processamento com a transformada *wavelet* é menor que do filtro de Wiener, aumentando o número de objetos detectados, probabilidade de acerto e, conseqüentemente, a sensibilidade. Entretanto, esse aumento também ocorre no número de falsos positivos, causando a diminuição da especificidade, acurácia e precisão.

As detecções realizadas com valor de h igual a 0,02 e elemento estruturante com raio igual a 3 pixels obtiveram os piores resultados de segmentação quando analisamos a média da sensibilidade das imagens. Isso ocorreu devido ao número de imagens sem objetos detectados, conforme observado anteriormente na Tabela 9. Além disso, o valor da área sob a curva ROC próximo a 0,5 indica que a maior parte das regiões não foi categorizada corretamente. Os valores mais altos para especificidade, acurácia e precisão se justificam pelo alto número de verdadeiro negativos, visto que a maior parte das regiões de 100×100 pixels nesses testes não contém objetos segmentados.

Considerando a média da sensibilidade, os melhores resultados foram obtidos nas imagens pré-processadas com a transformada *wavelet* e valor de h igual a 0,04. Dentre esses, se destaca o teste realizado com elemento estruturante com raio 3 pixels, com sensibilidade média igual a 89,27% e AUC com média igual a 0,76. A Tabela 11 também indica que o desvio padrão é elevado para a maior parte das métricas calculadas, sendo essa característica justificada pela diferença no número de regiões 100×100 pixels das imagens, principalmente devido aos tamanhos desiguais de imagens e quantidade de regiões com microcalcificações em cada uma delas.

O resultado do teste realizado com a transformada *wavelet*, valor de h igual a 0,02, elemento estruturante de raio 11 pixels e máximo de 1000 objetos obteve sensibilidade média

igual a 91,46%, o maior valor registrado com o menor desvio padrão calculado (9,50%). Entretanto, observa-se que esse processamento obteve os menores valores de especificidade, acurácia e precisão dentre os resultados, indicando que apesar da alta sensibilidade, grande parte das regiões com objetos segmentados nesse teste são falsos positivos.

Os testes do segundo método de detecção, a limiarização da imagem *top-hat*, foram realizados com a mudança do tamanho do elemento estruturante e do número máximo de objetos obtidos na binarização da imagem. Novamente, o resultado da segmentação por *watershed* usando os marcadores obtidos com *top-hat* foi dividido em regiões de 100×100 pixels e comparado com o *ground truth*. Com a matriz de confusão de cada imagem, foram calculadas a sensibilidade, especificidade, acurácia, precisão e a área da curva ROC (AUC). A Tabela 12 traz os resultados desses testes comparando a média (M) e o desvio padrão (STD) das métricas calculadas em cada processamento realizado.

Tabela 12 - Média e desvio padrão das métricas de avaliação da segmentação das imagens com detecção por *top-hat*.

Configuração		Sensibilidade		Especificidade		Acurácia		Precisão		AUC	
SE	Limiar	M	STD	M	STD	M	STD	M	STD	M	STD
Filtro de Wiener											
3	500	66,74	24,33	84,07	13,05	82,51	12,53	22,33	15,46	0,75	≈ 0,00
3	750	72,85	21,67	77,86	16,51	77,66	14,88	18,91	13,63	0,76	≈ 0,00
3	1000	76,41	19,82	73,58	18,44	74,27	16,38	17,05	12,48	0,77	≈ 0,00
11	500	55,13	21,54	84,29	12,02	82,00	12,27	19,40	13,56	0,70	≈ 0,00
11	750	62,52	20,83	78,85	15,41	77,81	14,42	17,10	12,31	0,72	≈ 0,00
11	1000	67,71	19,24	74,12	18,09	74,11	16,17	15,63	11,45	0,73	≈ 0,00
Transformada Wavelet											
3	500	77,35	20,21	72,34	16,95	72,89	15,40	16,43	12,94	0,76	≈ 0,00
3	750	84,88	15,59	63,28	21,32	65,64	18,63	14,24	11,3	0,77	≈ 0,00
3	1000	88,57	13,23	57,21	23,65	60,69	20,31	12,85	10,11	0,78	≈ 0,00
11	500	63,70	19,83	79,14	12,86	77,93	12,88	16,96	12,33	0,72	≈ 0,00
11	750	70,48	17,80	72,32	16,47	72,58	15,25	14,94	11,18	0,73	≈ 0,00
11	1000	75,50	16,11	66,70	19,58	68,13	17,35	13,74	10,33	0,74	≈ 0,00

Fonte: Elaborado pelo autor,

A Tabela 12 indica que os testes com maior média de sensibilidade foram os realizados com o elemento estruturante de raio 3 pixels. Elementos estruturantes de tamanhos menores permitem a detecção de mais objetos, aumentando os valores de sensibilidade, porém, diminuindo as métricas especificidade, acurácia e precisão devido ao aumento do número de falsos positivos. Como esperado, para o mesmo tamanho de elemento estruturante e número máximo de objetos na imagem binária, a média da sensibilidade obtida nos testes com pré-processamento com a transformada *wavelet* foi maior do que nos testes com filtro de Wiener.

Esses resultados reforçam que o aumento de borramento na imagem reduz o número de objetos detectados, diminuindo a sensibilidade, porém, aumentando a especificidade,

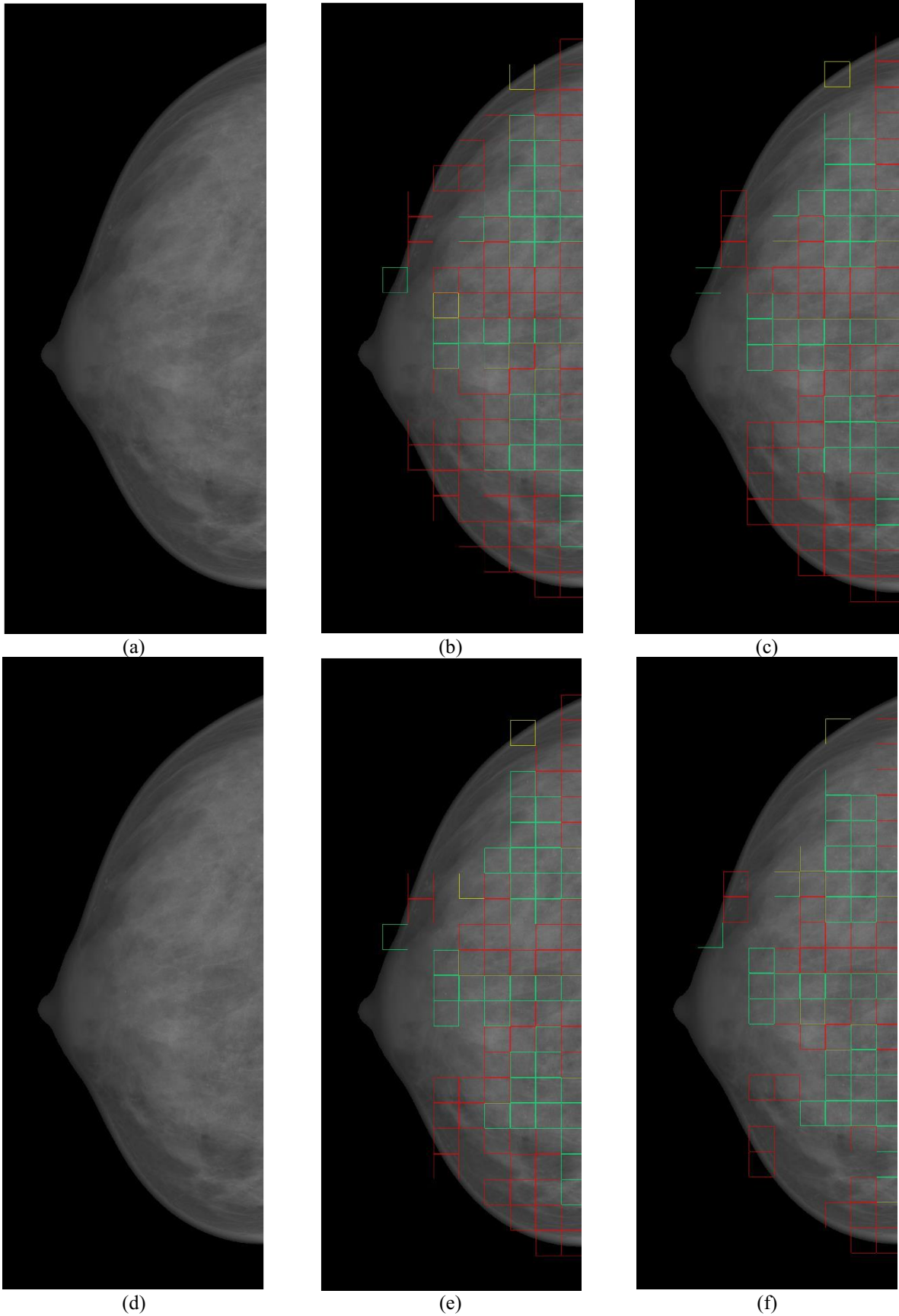
Dado que a operação *top-hat* é realizada através da abertura morfológica, a qual tem como primeiro passo a erosão da imagem, objetos com tamanho menor que o elemento estruturante utilizado tendem a ser eliminados. Os resultados corroboram essa premissa, visto que a sensibilidade média obtida nos testes realizados com o elemento estruturante de tamanho 11 pixels foram inferiores devido a detecção de menos objetos, o que diminui a probabilidade de acerto.

Considerando os dados das Tabelas 11 e 12, os resultados mais promissores para cada método de detecção foram escolhidos para extração das características e criação de modelos de redes neurais artificiais. Para detecção com *h-maxima*, a configuração selecionada utiliza a transformada *wavelet* como pré-processamento, elemento estruturante com raio de três pixels, valor de *h* igual a 0,04 e máximo de 1000 objetos no resultado da limiarização. Apesar desse processamento possuir a segunda maior sensibilidade dentre os processamentos com *h-maxima* (89,27%), esta configuração foi selecionada pois apresentou médias de especificidade, acurácia e precisão superiores aos valores encontrados com o processamento utilizando valor de *h* igual a 0,02, o qual obteve sensibilidade média de 91,46%. Dito isso, nesse caso, considerou-se que o pequeno ganho em sensibilidade não foi o suficiente para compensar as perdas em especificidade, acurácia e precisão.

Para o caso da detecção com *top-hat*, a configuração selecionada para extração de características foi a combinação do pré-processamento com a transformada *wavelet*, elemento estruturante de raio três pixels e máximo de 1000 objetos permitidos na imagem binária dos marcadores. Esse processamento resultou na maior média de sensibilidade (88,57%) quando comparado com os outros testes com *top-hat*, além da maior média para área da curva ROC (0,78) dentre todos os processamentos de segmentação realizados.

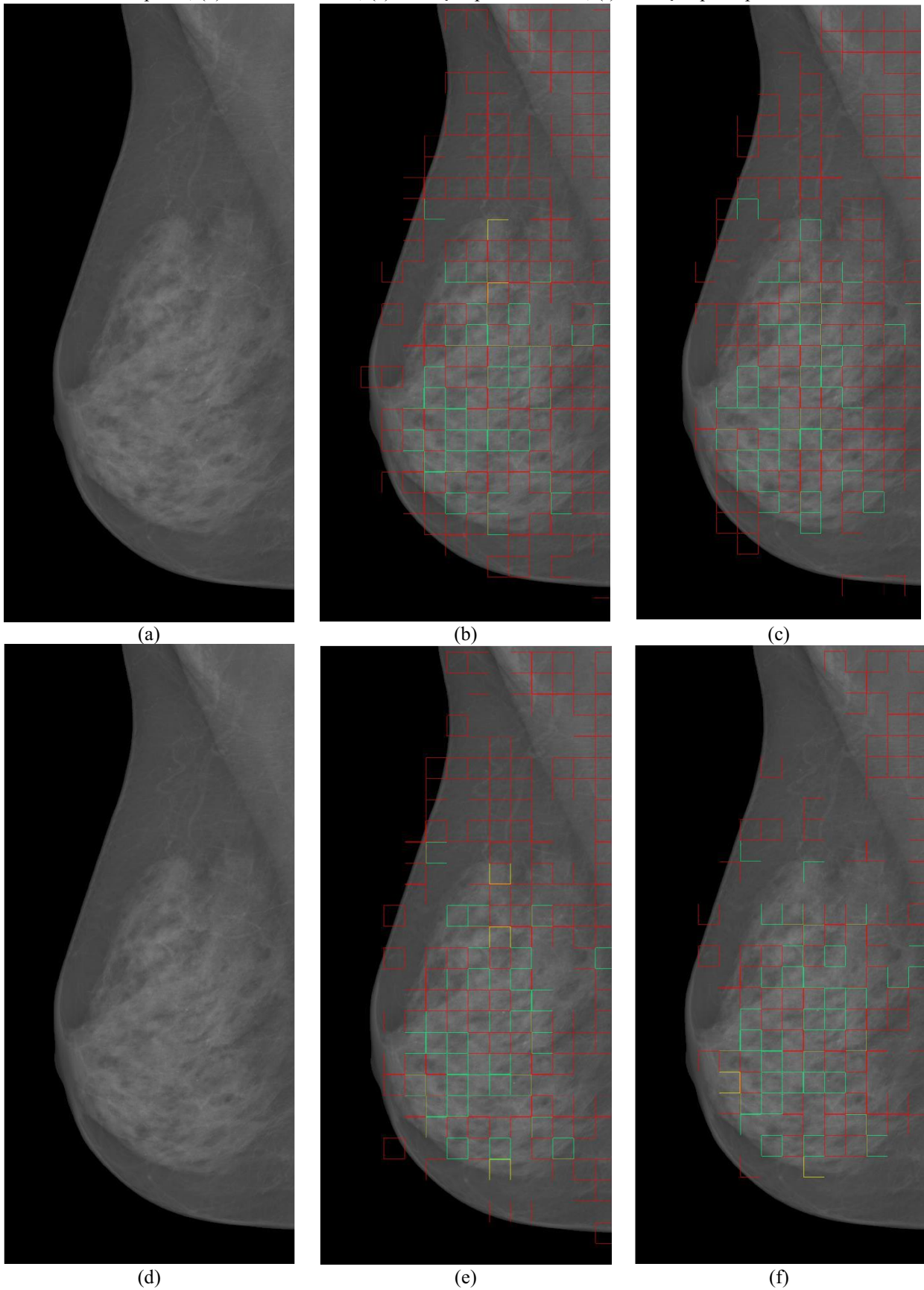
Os resultados obtidos com esses processamentos estão exemplificados nas Figuras 14 e 15, nas quais as regiões de 100×100 pixels categorizadas como verdadeiro positivas estão marcadas em verde; as regiões do tipo falso positivo em vermelho; e as regiões falso negativas em amarelo.

Figura 14 - Resultados da segmentação: (a) Transformada *wavelet*; (b) Detecção por *h-maxima*; (c) Detecção por *top-hat*; (d) Filtro de Wiener; (e) Detecção por *h-maxima*; (f) Detecção por *top-hat*.



Fonte: Elaborado pelo autor.

Figura 15 - Resultados da segmentação: (a) Transformada *wavelet*; (b) Detecção por *h-maxima*; (c) Detecção por *top-hat*; (d) Filtro de Wiener; (e) Detecção por *h-maxima*; (f) Detecção por *top-hat*.



Fonte: Elaborado pelo autor.

As imagens das Figuras 14 e 15 mostram que diferentes regiões são categorizadas como verdadeiro e falso positivo, conforme a segmentação ou não de objetos detectados pelos diferentes métodos de detecção utilizados. Percebe-se que, comparando os resultados com cada pré-processamento de redução de ruído utilizando a mesma técnica de detecção de objetos, os resultados obtidos com o filtro de Wiener apresentaram menos regiões do tipo falso positivo. Essa tendência também é observada nos valores obtidos nas Tabelas 11 e 12 para as médias de especificidade e acurácia desses resultados e se justifica pela maior capacidade de borramento do filtro de Wiener, que causa a suavização da imagem e eliminação de objetos que poderiam ser detectados.

Além disso, a comparação dos resultados obtidos em cada método de detecção utilizando a mesma técnica de redução de ruído também mostra mudanças nas regiões de 100×100 pixels categorizadas como verdadeiro e falso positivas. A presença de regiões 100×100 categorizadas corretamente (verdadeiro positivo ou verdadeiro negativo) nos resultados dos testes com um tipo de detecção e classificadas erroneamente (falso positivo ou falso negativo) com o outro já era esperada, visto que os processamentos *h-maxima* e *top-hat* são implementados utilizando operações morfológicas diferentes. Além disso, a operação *h-maxima* inclui a análise da mínima diferença de intensidade entre regiões vizinhas, ao contrário do método *top-hat*, o qual prioriza o tamanho e a forma da região analisada (SOILLE, 2004).

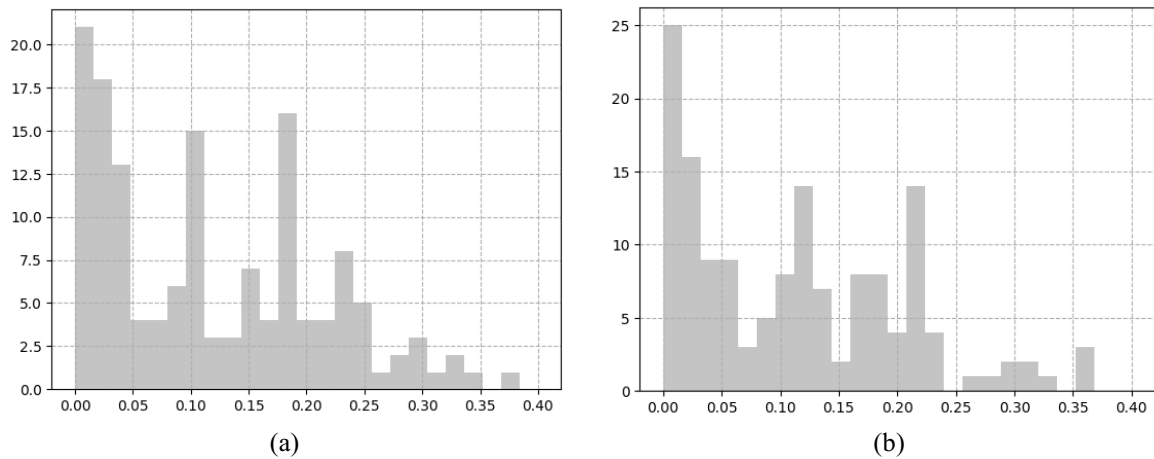
Apesar de alta sensibilidade média, os dois processamentos selecionados possuem valores baixos de média da acurácia, especificidade e precisão quando comparados aos resultados dos testes com outras configurações, o que justifica a necessidade de encontrar métodos para redução do número de regiões do tipo falso positivo. Dessa forma, buscou-se desenvolver dois modelos de redes neurais artificiais (RNA-MLP) capazes de diferenciar as regiões de 100×100 pixels obtidas na segmentação e classificá-las corretamente como verdadeiro positivo ou negativo. Com a redução do número de falsos positivos espera-se aumentar a especificidade, a acurácia e a precisão dos resultados obtidos na segmentação da imagem.

Os dados utilizados na produção dos modelos de RNA-MLP correspondem às características extraídas do conjunto de regiões 100×100 pixels categorizadas como verdadeiro e falso positivo e os objetos segmentados em cada uma delas. Essas características representam descritores dos níveis de intensidade, tipos de textura e morfologia das regiões 100×100 e dos objetos segmentados com *watershed*. Para seleção de atributos capazes de descrever as diferenças entre as classes verdadeiro positivo e falso positivo, foram realizadas avaliações

estatísticas dos conjuntos de dados através de do cálculo de medidas como a média, o desvio padrão e a matriz de correlação das características extraídas.

A matriz de correlação permite verificar se há relação linear entre variáveis e, neste trabalho, foi utilizada como critério para seleção das características de entrada dos modelos de RNA-MLP. Dessa forma, buscou-se selecionar as características com maior correlação com a saída esperada dos modelos, ou seja, zero para as regiões do tipo falso positivo e um para as do tipo verdadeiro positivo. A Figura 16 mostra a distribuição do módulo dos valores de correlação entre as características extraídas e a saída esperada para os conjuntos de resultados obtidos com na segmentação com marcadores detectados com *h-maxima* e *top-hat*.

Figura 16 - Distribuição do módulo dos valores de correlação entre as características extraídas e a saída esperada: (a) Dados da detecção com *h-maxima*; (b) Dados da detecção com *top-hat*.



Fonte: Elaborado pelo autor.

Observa-se pela Figura 16 que os atributos extraídos possuem baixa ou quase nenhuma correlação com a saída esperada, sinalizando relações não lineares, a complexidade dos dados e a dificuldade de distinção entre as classes verdadeiro e falso positivo. Apesar disso, foram selecionados 16 atributos como entrada dos modelos de RNA-MLP com base nos maiores valores absolutos de correlação entre as características analisadas e a saída esperada, os quais estão indicados na Tabela 13.

Tabela 13 - Atributos selecionados e medidas calculadas nos conjuntos de treinamento.

Atributo	Correlação	
	<i>H-maxima</i>	<i>Top-hat</i>
Intensidade máxima da região.	0,37	0,36
Média da diferença entre a intensidade média dos objetos e a intensidade máxima da região.	0,34	0,35
Média da diferença entre a intensidade mínima dos objetos e a intensidade máxima da região.	0,33	0,36
Diferença entre as intensidades média e máxima da região.	0,32	0,33
Média da intensidade máxima dos objetos.	0,32	0,29

Média da diferença entre a intensidade máxima dos objetos e a intensidade máxima da região.	0,30	0,29
Média da intensidade média dos objetos.	0,30	0,27
Média da razão entre a intensidade média dos objetos e a intensidade máxima da região.	0,29	0,31
Média da razão entre a intensidade mínima dos objetos e a intensidade máxima da região.	0,28	0,31
Média da intensidade mínima dos objetos.	0,28	0,24
Média da razão entre a intensidade máxima dos objetos e a intensidade máxima da região.	0,27	0,27
Média da diferença da variância em quatro direções.	0,25	0,24
Média do Segundo Momento Angular (ASM) em quatro direções.	0,24	0,22
Média da entropia em quatro direções.	0,24	0,22
Média da soma da média em quatro direções.	0,23	0,21
Intensidade média da região	0,23	0,21

Fonte: Elaborado pelo autor.

A Tabela 13 indica que a maior parte dos atributos selecionados são referentes a características que descrevem os níveis de intensidade das regiões de 100×100 pixels e seus objetos. Além disso, também foram selecionadas as características de textura diferença da variância, segundo momento angular, entropia e soma da média, sendo que os maiores valores de correlação foram obtidos com a média dos valores calculados nas quatro direções. É possível verificar que, apesar das mesmas características apresentarem os maiores valores correlação nos dois conjuntos de dados analisados, existe uma diferença nos valores calculados quando são utilizados dados de métodos diferentes de detecção de objetos.

É possível expandir a etapa de extração de características para incluir cálculo de outros atributos e relações entre os descritores das regiões de 100×100 pixels e dos objetos segmentados. Por exemplo, é utilizar modelos de regressão linear e logística e de árvores de decisão para combinar variáveis com alta correlação, além de algoritmos com redes neurais convolucionais e transformada *wavelet* (LITJENS et al., 2017; RUSSELL; NORVIG, 2010; SOUSA, 2022).

Como indicado na Tabela 6, as 178 imagens de cada conjunto de resultados foram divididas em dois grupos: 124 imagens cujos dados serão utilizados no treinamento e validação dos modelos e 54 imagens para o teste das soluções obtidas. Os dados das imagens de teste não devem ser utilizados durante o treinamento, de forma que durante o teste, apenas amostras inéditas sejam apresentadas aos modelos. Além disso, considerando a discrepância entre o

número de amostras do verdadeiras e falsas, foi necessário balancear os conjuntos de treinamento, eliminando aleatoriamente amostras da classe mais numerosa até que o número de amostras dos tipos verdadeiro e falso positivo fosse igualado. Dessa forma, o conjunto de dados de treinamento referente às regiões obtidas com detecção por *h-maxima* foi composto de 6.966 regiões, enquanto no grupo de imagens processadas com *top-hat* esse número foi igual a 6.876 regiões de 100×100 pixels.

Em posse dos dados de treinamento e com os atributos selecionados, os dois modelos de RNA-MLP foram produzidos com o objetivo de diminuir o número de regiões do tipo falso positivo nas imagens de teste. Os parâmetros para construção dos modelos foram testados com valores aleatórios e, pela comparação dos resultados da validação cruzada desses testes, foi possível selecionar as melhores configurações para composição das redes. Os parâmetros utilizados nos dois modelos produzidos, como número de camadas, número de nós, função de ativação e taxa de aprendizagem foram descritos na Tabela 7.

Após a seleção das melhores configurações de parâmetros para os modelos, o conjunto de imagens de teste foi utilizado para testar e avaliar a capacidade de classificação dos modelos de RNA-MLP. Para isso, as regiões de 100×100 pixels do tipo verdadeiro e falso positivo de cada imagem de teste foram classificadas e, para cada resultado, calculou-se a sensibilidade, especificidade, acurácia, precisão e área sob a curva ROC (AUC). A Tabela 14 permite comparar a média das métricas calculadas em cada teste de classificação, os resultados obtidos na segmentação das imagens do conjunto de teste e os valores obtidos com o novo arranjo de regiões dos tipos verdadeiro e falso positivo.

Tabela 14 - Resultados das métricas calculadas para segmentação e classificação das imagens do grupo de teste.

Métrica	Segmentação				Classificação				Final			
	<i>H-maxima</i>		<i>Top-hat</i>		<i>H-maxima</i>		<i>Top-hat</i>		<i>H-maxima</i>		<i>Top-hat</i>	
	M	STD	M	STD	M	STD	M	STD	M	STD	M	STD
Sensibilidade	90,66	9,82	88,64	12,07	60,72	17,55	55,98	20,06	55,84	19,09	50,64	20,78
Especificidade	55,12	24,78	60,35	23,20	74,96	5,51	77,22	5,48	88,73	6,99	90,90	6,05
Acurácia	58,74	21,60	63,25	20,34	72,33	4,77	73,50	5,43	85,87	8,67	87,39	8,38
Precisão	12,03	9,86	13,05	10,68	23,49	18,22	25,29	20,00	23,49	18,22	25,29	20,00
AUC	0,73	0,16	0,74	0,15	0,68	0,10	0,67	0,11	0,72	0,11	0,71	0,12

Fonte: Elaborado pelo autor.

É importante ressaltar que os valores indicados na Tabela 14 consideram apenas as métricas calculadas para as 54 imagens utilizadas nos testes dos modelos de RNA-MLP. Os resultados da classificação foram calculados com referência nas regiões categorizadas como verdadeiro e falso positivo na segmentação, ou seja, os valores de especificidade, acurácia e de AUC não consideram o total de regiões do tipo verdadeiro e falso negativo identificadas na segmentação. Por sua vez, os dados das colunas com index “Final” englobam todas essas

regiões, pois foram calculados com base no total de regiões de 100×100 pixels da imagem e os resultados da classificação pelas redes neurais artificiais.

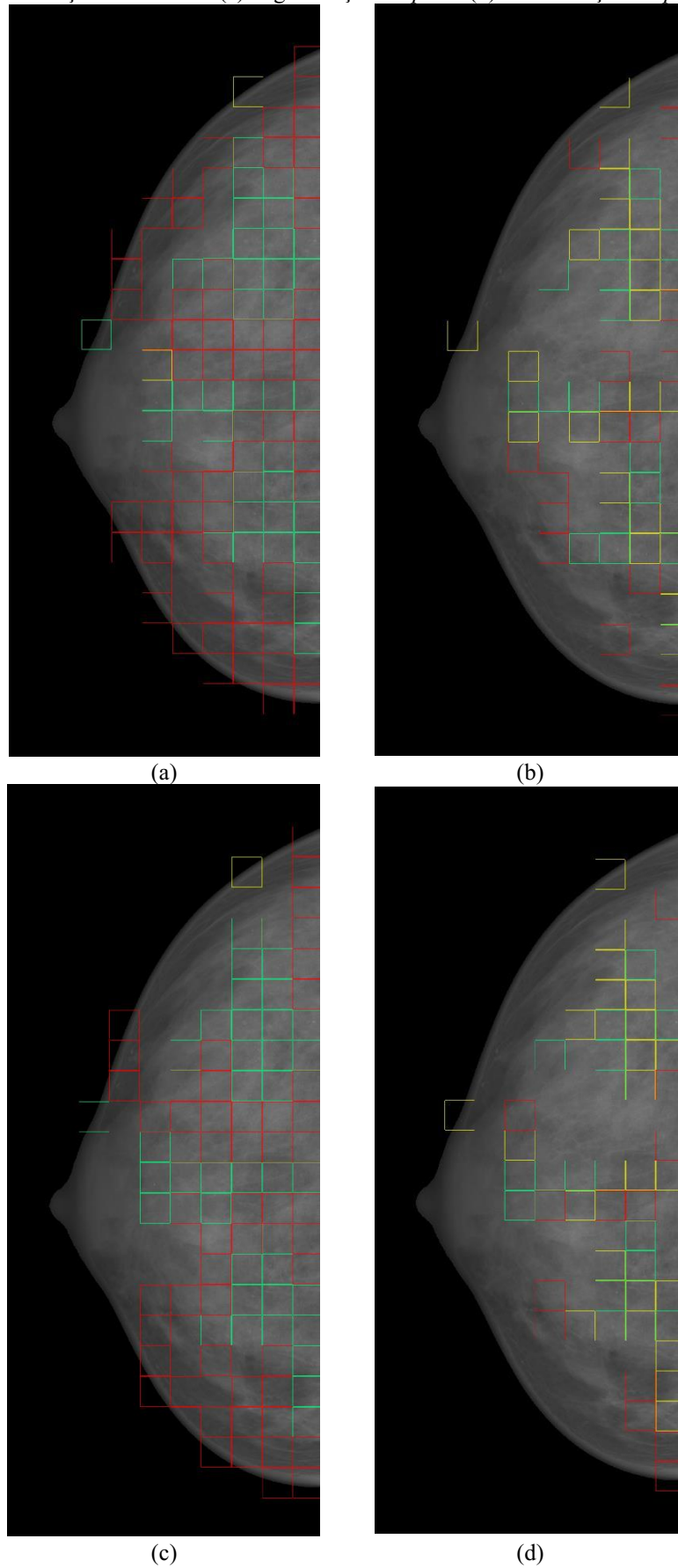
Os resultados na Tabela 14 mostram que, quando consideramos o total de regiões de 100×100 pixels na imagem, a classificação das regiões do tipo verdadeiro e falso positivo resultou em diminuição nas médias da sensibilidade e AUC. Por sua vez, o desvio padrão da sensibilidade aumentou, indicando há maior variação no número de regiões classificadas corretamente em cada imagem.

Além disso, observa-se que as médias da especificidade e da acurácia aumentaram após a classificação das regiões, indicando que o número de regiões do tipo falso positivo nas imagens de teste diminuiu. Apesar de pequeno aumento quando comparada ao valor obtido na segmentação, a média da precisão após a classificação das regiões permaneceu baixa, indicando que parte das regiões classificadas como verdadeiro positivo são, na verdade, falsos positivos. Esses dados mostram que os modelos de RNA-MLP são eficientes para diminuir o número de falsos positivos encontrados na segmentação e aumentar a especificidade e acurácia dos resultados, porém, isso ocorre em função da diminuição da sensibilidade.

Os resultados obtidos na classificação das regiões de 100×100 pixels podem ser aprimorados através de diferentes abordagens na extração de características, seleção de atributos e configurações nas RNA-MLP. Algumas soluções propostas na literatura envolvem o uso de algoritmos para detecção e eliminação de dados atípicos, técnicas para aproximar os dados de uma distribuição normal, diferentes disposições de camadas, nós e ligações e a utilização de redes neurais artificiais pré-treinadas. É importante ressaltar que o tamanho da classe falso positivo é muito maior que da classe verdadeiro positivo, sendo possível que a ampliação do número de dados disponível para treinamento melhore a performance dos modelos (LITJENS et al., 2017; RUSSELL; NORVIG, 2013).

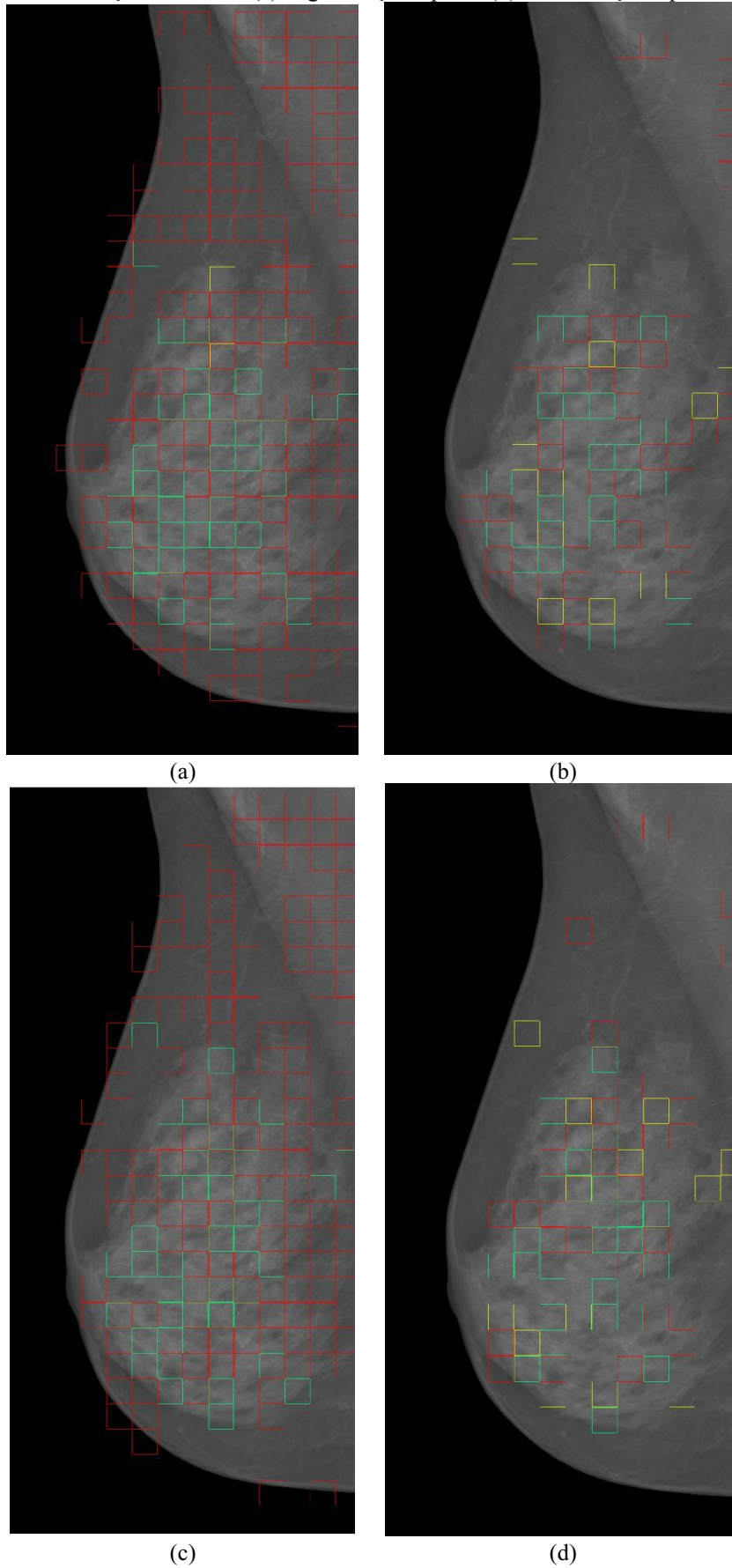
A seguir, as Figuras 17 e 18 trazem a comparação visual entre os resultados obtidos na segmentação e classificação das imagens separadas para teste dos modelos de RNA-MLP. Nessas imagens as regiões de 100×100 pixels categorizadas como verdadeiro positivas estão marcadas em verde; as regiões do tipo falso positivo em vermelho; e as regiões falso negativas em amarelo. As imagens são referentes aos resultados obtidos nas imagens pré-processadas com transformada *wavelet*.

Figura 17 - Comparação entre os resultados da segmentação e classificação: (a) Segmentação - *h-maxima* (b) Classificação - *h-maxima* (c) Segmentação - *top-hat* (d) Classificação - *top-hat*.



Fonte: Elaborado pelo autor.

Figura 18 - Comparação entre os resultados da segmentação e classificação: (a) Segmentação *h-maxima* (b) Classificação *h-maxima* (c) Segmentação *top-hat* (d) Classificação *top-hat*.



Fonte: Elaborado pelo autor.

As imagens das Figuras 17 e 18 indicam a presença de mais regiões do tipo falso negativo nos resultados da classificação, reforçando os dados da Tabela 14 que trazem a diminuição na sensibilidade nesses resultados. Também é possível notar a diminuição do número de regiões do tipo falso positivo após a classificação nas imagens, o que também corrobora o aumento na especificidade e acurácia dos resultados. O aumento da sensibilidade nos resultados da classificação seria possível ao custo de diminuição da precisão, ou seja, através do aumento no número de regiões do tipo falso positivo classificadas.

8. CONCLUSÃO

A detecção e segmentação de microcalcificações em imagens de mamografia através do processamento de digital de imagens pode ter papel relevante no auxílio ao diagnóstico do câncer de mama. As microcalcificações são o tipo de achado mamográfico mais comum em mamografias e a visualização e avaliação correta dessas estruturas pelo radiologista é imprescindível para o encaminhamento correto de cada caso. Neste sentido, este trabalho utilizou as operações morfológicas *h-maxima* e *top-hat* para detectar marcadores para restringir a segmentação de objetos por *watershed*.

Através da divisão dos resultados da segmentação em regiões de 100×100 pixels e comparando os resultados com o *ground truth* produzido, observou-se que os processamentos com maior sensibilidade também obtiveram valores menores para especificidade, acurácia e precisão. Isso significa que com o aumento de regiões categorizadas como verdadeiro positivo e, conseqüentemente, da sensibilidade, também houve aumento no número de regiões do tipo falso positivo, causando a diminuição da especificidade, acurácia e precisão.

Além disso, foi possível verificar que, devido a maior capacidade de borramento do filtro de Wiener, os testes cujas imagens foram pré-processadas com esse método obtiveram média de sensibilidade menor do que os testes com imagens filtradas com a transformada *wavelet*. Isso ocorre pois, com o aumento do borramento da imagem, objetos que poderiam ser detectados como microcalcificações desaparecem, ou seja, menos regiões são detectadas, o que diminui a probabilidade de acerto a sensibilidade da segmentação.

Os testes de detecção de máximos regionais com *h-maxima* indicaram que valores pequenos para a mínima diferença de nível de intensidade entre as regiões pode causar a detecção de muitos objetos, o que pode dificultar a limiarização iterativa e diminuir a especificidade e acurácia dos resultados. Já os testes com *top-hat* mostraram que utilizar elementos estruturantes maiores permite eliminar objetos muito pequenos, sendo uma alternativa para detecção de regiões quando se sabe a forma e o tamanho dos objetos de interesse.

A limiarização iterativa realizada com diferentes números de objetos permitidos na imagem binária demonstrou ser capaz de binarizar as imagens *h-convex* e *top-hat*. Observou-se que com mais objetos permitidos no resultado da binarização, menor o nível de cinza do limiar selecionado, fazendo com que mais pixels sejam adicionados a imagem binária e a quantidade de componentes conectados e o formato dos objetos sofram alteração.

Considerando o grande número de regiões categorizadas como falso positivo na segmentação, foram desenvolvidos dois modelos de Redes Neurais Artificiais do tipo Perceptron Multicamadas (RNA-MLP) com o objetivo de classificar as regiões de 100×100 pixels categorizadas como verdadeiro e falso positivo na segmentação. Esses modelos foram treinados e testados utilizando descritores de formato, intensidade e textura das regiões 100×100 e dos objetos segmentado nessas regiões. Verificou-se com a avaliação dos resultados que os modelos de RNA-MLP foram capazes de reduzir o número de regiões do tipo falso positivo e, conseqüentemente, aumentar a especificidade e acurácia dos resultados. Entretanto, essa melhora também causou a diminuição na sensibilidade do algoritmo, ou seja, mais regiões foram categorizadas como falso negativo.

Por fim, os resultados obtidos na segmentação das imagens podem ser aprimorados através de trabalhos futuros quanto ao método de limiarização utilizado para binarizar as imagens *h-convex* e *h-maxima*, técnicas de produção do marcador externo e seleção criteriosa dos marcadores internos. Já os modelos de RNA-MLP podem apresentar resultados melhores conforme a utilização de outras características de entrada, aumento no número de dados disponíveis para treinamento e reaproveitamento de modelos de aprendizado de máquina.

REFERÊNCIAS

- BRAY, F. *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, v. 74, n. 3, p. 229–263, 2024. DOI 10.3322/caac.21834
- BUSHBERG, J. T. **The Essential Physics of Medical Imaging**. 3. ed. Philadelphia: Lippincott Williams & Wilkins, 2012.
- CAI, H. *et al.* Breast Microcalcification Diagnosis Using Deep Convolutional Neural Network from Digital Mammograms. **Computational and Mathematical Methods in Medicine**, v. 2019, 2019. DOI 10.1155/2019/2717454
- CARNEIRO, P. **Proposta de metodologia para realce de contraste em imagens de mamas densas utilizando decomposição multiescala com transformada discreta wavelet**. 2019. 210 p. Tese (Doutorado), Universidade Federal de Uberlândia, Uberlândia, 2019. DOI 10.14393/ufu.te.2019.2593
- CASCIO, D. *et al.* **A Microcalcification Detection System in Mammograms based on ANN Clustering**. In: 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). 2018, Sydney. **Anais [...]. IEEE**, 2018. DOI 10.1109/NSSMIC.2018.8824729
- CIECHOLEWSKI, M. Microcalcification Segmentation from Mammograms: A Morphological Approach. **Journal of Digital Imaging**, v. 30, n. 2, p. 172–184, 2017. DOI 10.1007/s10278-016-9923-8
- COELHO, L. P. Mahotas: Open source software for scriptable computer vision. **Journal of Open Research Software**, v. 1, n. 3, 2013. DOI 10.48550/arXiv.1211.4907
- COSTA JÚNIOR, C. A. **Proposta de uma metodologia para suavização de ruído em imagens mamográficas de mamas densas**. 2019. 80 p. Dissertação (Mestrado em Engenharia Biomédica) - Universidade Federal de Uberlândia, Uberlândia, 2019. DOI 10.14393/ufu.di.2019.2036
- DE PAREDES, E. S. **Atlas of mammography**. 3. ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2007.
- DOUGHERTY, G. **Digital Image Processing for Medical Applications**. 1. ed. Cambridge: Cambridge University Press, 2009.
- FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, v. 27, n. 8, p. 861–874, 2006. DOI 10.1016/j.patrec.2005.10.010
- GÓMEZ, K. A. H.; ECHEVERRY-CORREA, J. D.; GUTIÉRREZ, Á. Á. O. Automatic Pectoral Muscle Removal and Microcalcification Localization in Digital Mammograms. **Healthcare Informatics Research**, v. 27, n. 3, p. 222–230, 2021. DOI 10.4258/hir.2021.27.3.222
- GONZALEZ, RAFAEL. C.; WOODS, R. E. **Processamento Digital de Imagens**. 3. ed. São Paulo: Pearson Prentice Hall, 2010.

HALKIOTIS, S.; BOTSIS, T.; RANGOUSI, M. Automatic detection of clustered microcalcifications in digital mammograms using mathematical morphology and neural networks. **Signal Processing**, v. 87, n. 7, p. 1559–1568, jul. 2007. DOI 10.1016/j.sigpro.2007.01.004

HARALICK, R. M. Statistical and structural approaches to texture. **Proceedings of the IEEE**, v. 67, n. 5, p. 786–804, 1979. DOI 10.1109/PROC.1979.11328

INSTITUTO NACIONAL DE CÂNCER. **Mamografia: da prática ao controle**. Rio de Janeiro, RJ: INCA, 2007.

INSTITUTO NACIONAL DE CÂNCER. **Estimativa 2023: incidência de câncer no Brasil**. Rio de Janeiro, RJ: INCA, 2022.

INSTITUTO NACIONAL DE CÂNCER JOSÉ DE ALENCAR GOMES DA SILVA. **Diretrizes para a detecção precoce do câncer de mama no Brasil**. Rio de Janeiro, RJ: INCA, 2015.

INSTITUTO NACIONAL DE CÂNCER JOSÉ DE ALENCAR GOMES DA SILVA. **Deteção precoce do câncer**. Rio de Janeiro, RJ: INCA, 2021.

KARALE, V. A. *et al.* A Screening CAD Tool for the Detection of Microcalcification Clusters in Mammograms. **Journal of Digital Imaging**, v. 32, n. 5, p. 728–745, 2019. DOI 10.1007/s10278-019-00249-5

KINGMA, D. P.; BA, J. **Adam: A Method for Stochastic Optimization**. *In*: 3rd International Conference for Learning Representations. 2015, San Diego. 2017. DOI 10.48550/arXiv.1412.6980

LAURITZEN, A. D. *et al.* Early Indicators of the Impact of Using AI in Mammography Screening for Breast Cancer. **Radiology**, v. 311, n. 3, 2024. DOI 10.1148/radiol.232479

LITJENS, G. *et al.* A survey on deep learning in medical image analysis. **Medical Image Analysis**, v. 42, p. 60–88, 2017. DOI 10.48550/arXiv.1702.05747

MOREIRA, I. C. *et al.* INbreast: Toward a Full-field Digital Mammographic Database. **Academic Radiology**, v. 19, n. 2, p. 236–248, 2012. DOI 10.1016/j.acra.2011.09.014

OKUNO, E.; YOSHIMURA, E. **Física das radiações**. 1. ed. São Paulo: Oficina de Textos, 2010.

REZAEI, Z. A review on image-based approaches for breast cancer detection, segmentation, and classification. **Expert Systems with Applications**, v. 182, nov. 2021. DOI 10.1016/j.eswa.2021.115204

RUSSELL, S.; NORVIG, P. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013.

SAMALA, R. K. *et al.* AI and machine learning in medical imaging: key points from development to translation. **BJR|Artificial Intelligence**, v. 1, n. 1, jan. 2024. DOI 10.1093/bjr/ai/ubae006

SICKLES, E. A. *et al.* ACR BI-RADS Mammography. Em: **ACR BI-RADS Atlas, Breast Imaging Reporting and Data System**. 5 ed. Reston, VA: American College of Radiology, 2013.

SILVA JÚNIOR, E. C. **Modelo de processamento de imagens mamográficas para detecção de agrupamentos de microcalcificações**. 2009. 209 p. Tese (Doutorado) - Universidade de São Paulo, São Carlos, 2009.

SOILLE, P. **Morphological Image Analysis**. 2. ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

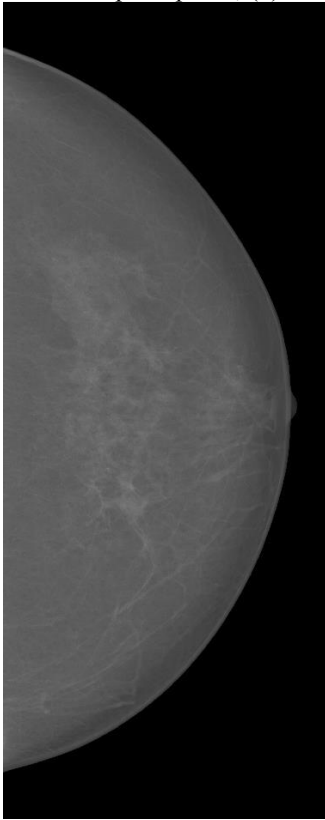
SOUSA, P. M. DE. **Estratégias de classificação de imagens radiológicas utilizando redes neurais convolucionais e transformada Wavelet**. 2022. 173 p. Tese (Doutorado) - Universidade Federal de Uberlândia, Uberlândia, jul. 2022. DOI 10.14393/ufu.te.2022.433

TOUIL, A. *et al.* A new conditional region growing approach for microcalcification delineation in mammograms. **Medical & Biological Engineering & Computing**, v. 59, n. 9, p. 1795–1814, set. 2021. DOI 10.1007/s11517-021-02379-x

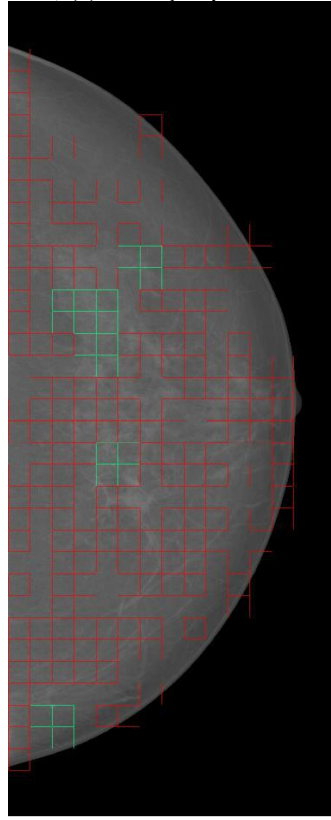
WANG, Z. *et al.* Breast microcalcifications detection based on fusing features with DTCWT. **Journal of X-Ray Science and Technology**, v. 28, n. 2, p. 197–218, 9 abr. 2020. DOI 10.3233/XST-190583

APÊNDICE A – Resultados da segmentação

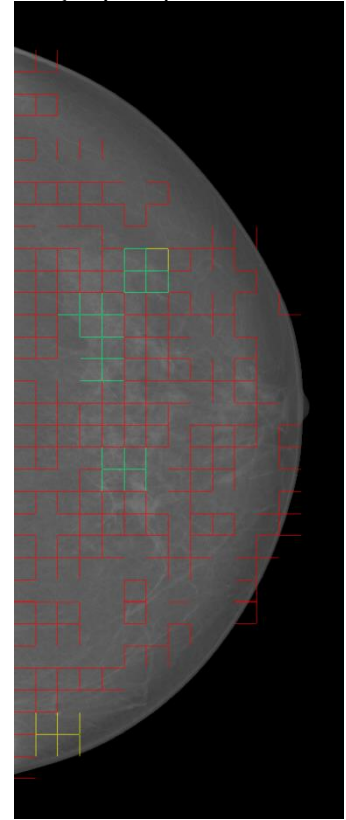
Resultados da segmentação por *watershed*: (a) Transformada *wavelet*; (b) Detecção por *h-maxima*; (c) Detecção por *top-hat*; (d) Filtro de Wiener; (e) Detecção por *h-maxima*; (f) Detecção por *top-hat*.



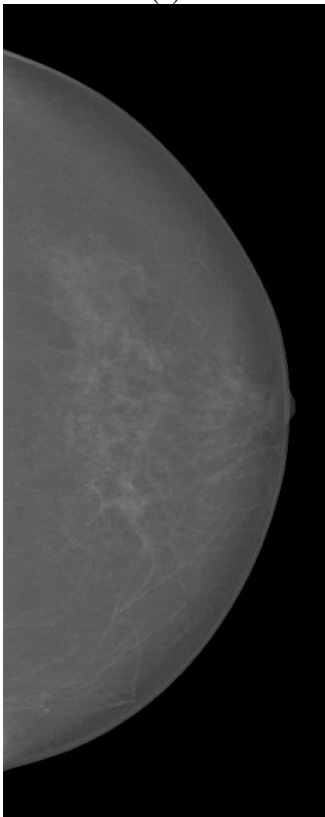
(a)



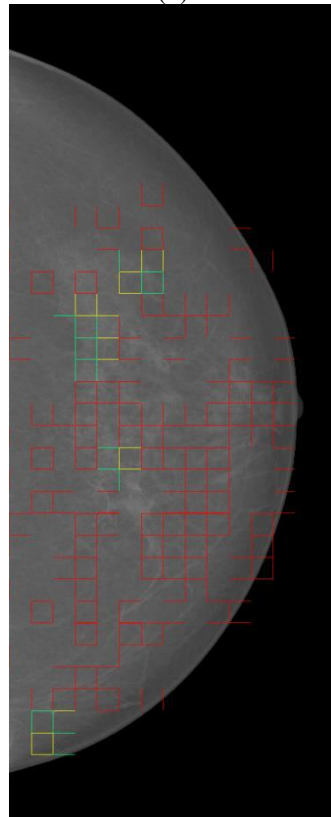
(b)



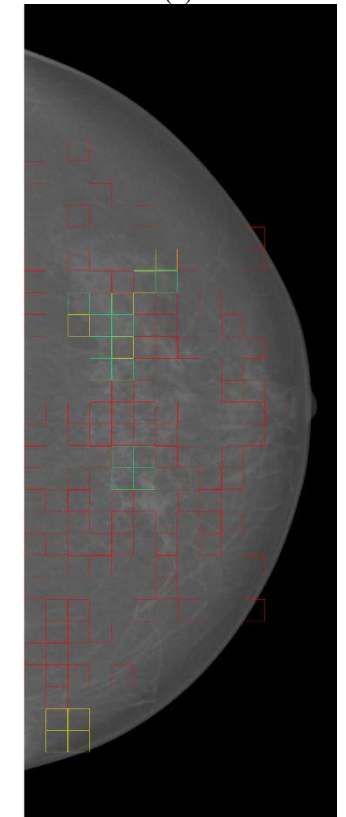
(c)



(d)



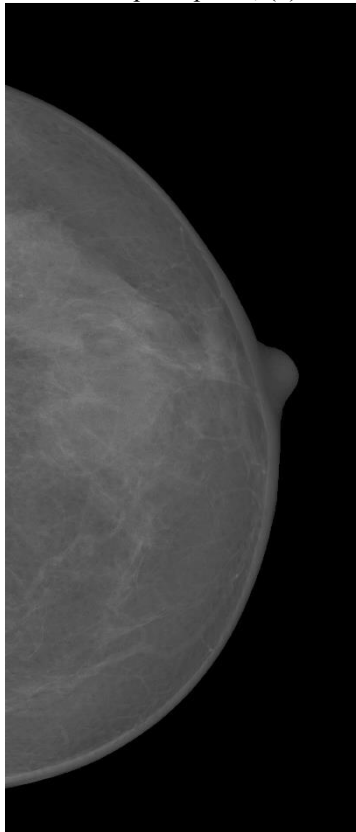
(e)



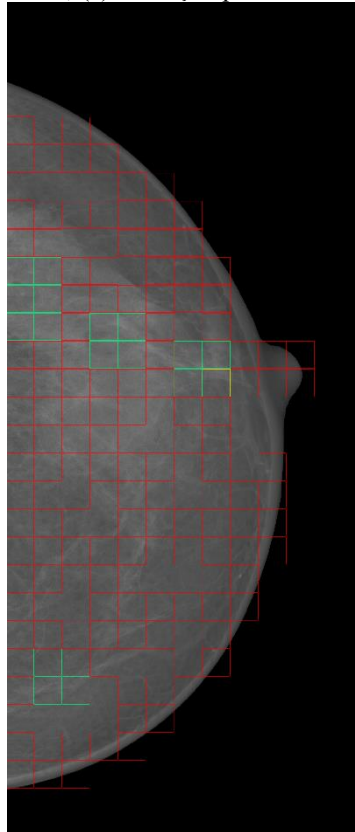
(f)

Fonte: Elaborado pelo autor.

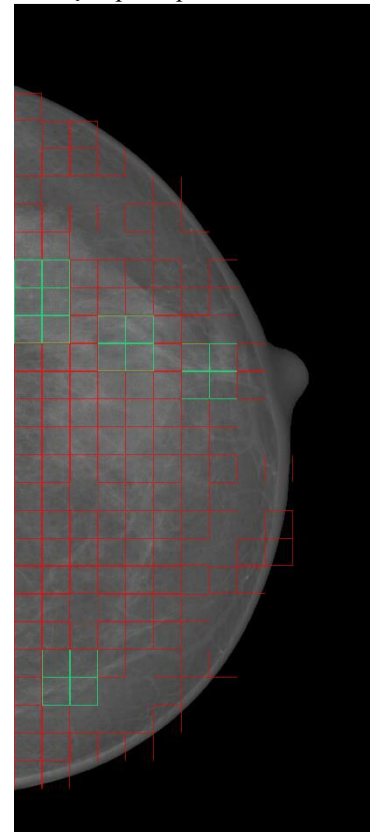
Resultados da segmentação por *watershed*: (a) Transformada *wavelet*; (b) Detecção por *h-maxima*; (c) Detecção por *top-hat*; (d) Filtro de Wiener; (e) Detecção por *h-maxima*; (f) Detecção por *top-hat*.



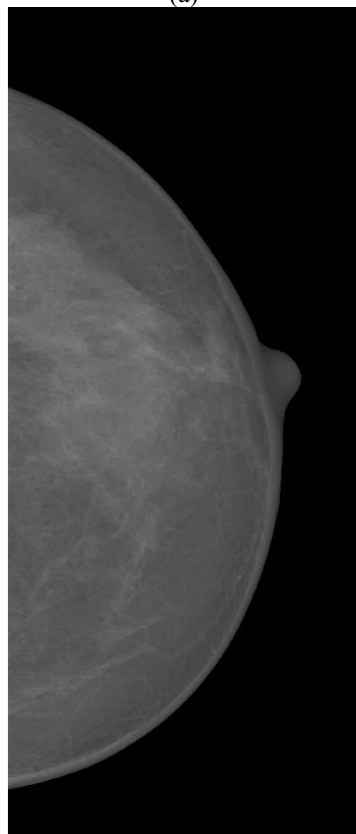
(a)



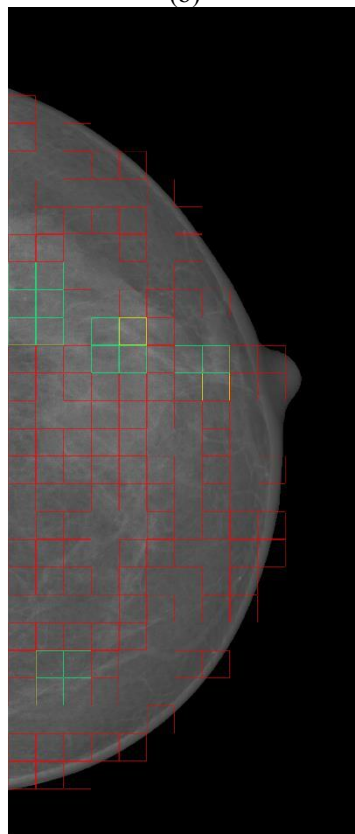
(b)



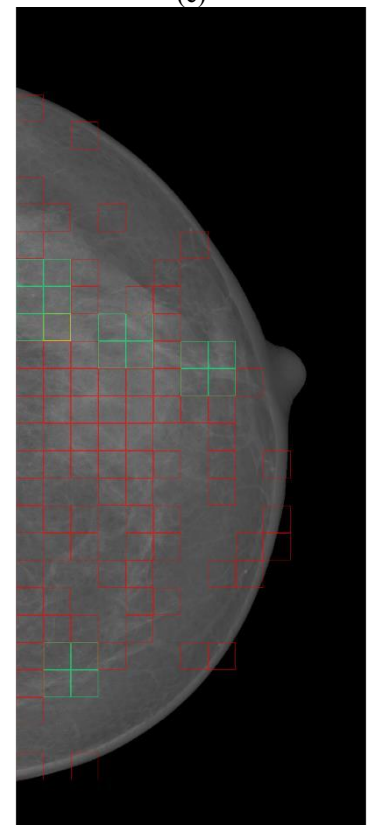
(c)



(d)



(e)

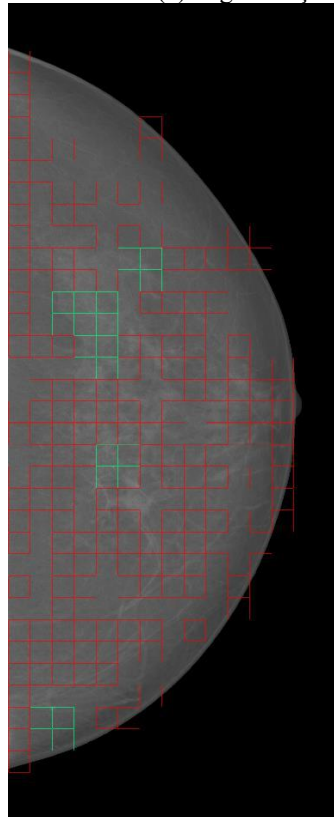


(f)

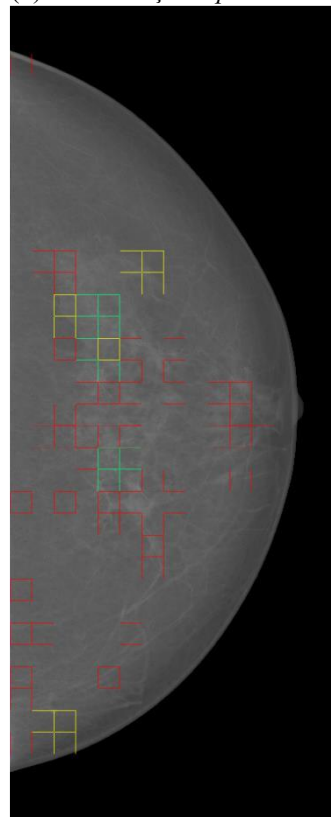
Fonte: Elaborado pelo autor.

APÊNDICE B – Resultados da classificação

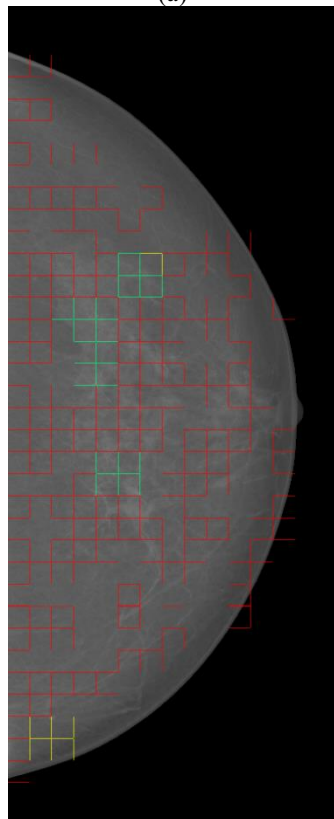
Comparação entre os resultados da segmentação e classificação: (a) Segmentação *h-maxima* (b) Classificação *h-maxima* (c) Segmentação *top-hat* (d) Classificação *top-hat*.



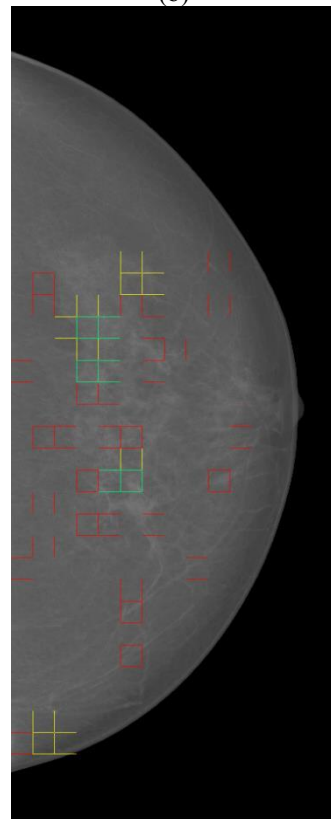
(a)



(b)



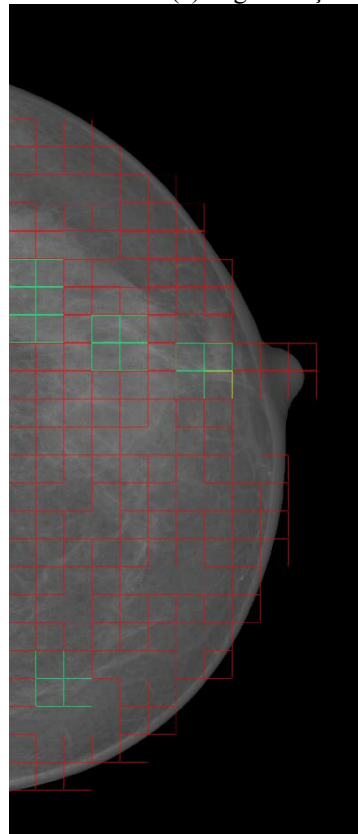
(c)



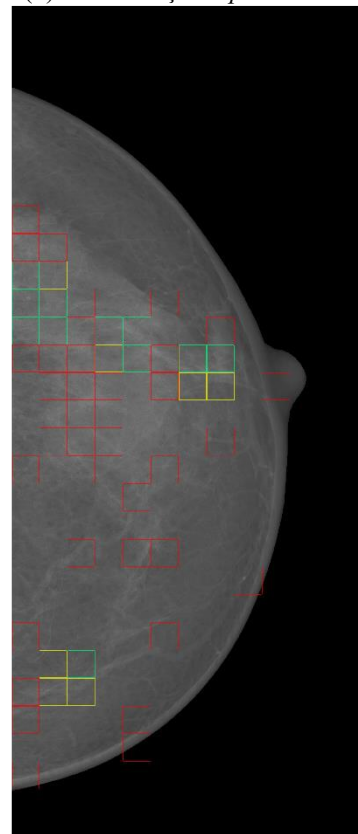
(d)

Fonte: Elaborado pelo autor.

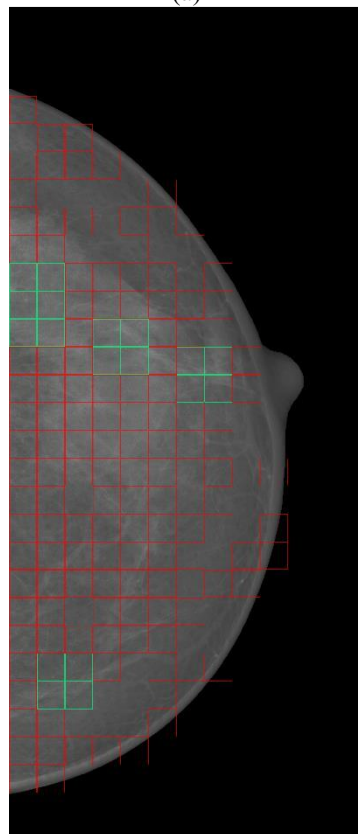
Comparação entre os resultados da segmentação e classificação: (a) Segmentação *h-maxima* (b) Classificação *h-maxima* (c) Segmentação *top-hat* (d) Classificação *top-hat*.



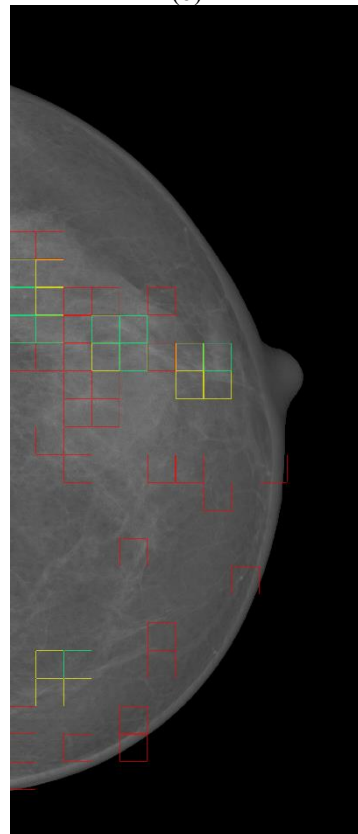
(a)



(b)



(c)



(d)

Fonte: Elaborado pelo autor.