



Pós-Graduação em Ciência da Computação

WASHINGTON WAGNER AZEVEDO DA SILVA

**MÁQUINAS MORFOLÓGICAS DE APRENDIZADO EXTREMO
APLICADAS À DETECÇÃO E CLASSIFICAÇÃO DE LESÕES EM
MAMOGRAFIAS**



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

RECIFE
2017

WASHINGTON WAGNER AZEVEDO DA SILVA

**MÁQUINAS MORFOLÓGICAS DE APRENDIZADO EXTREMO
APLICADAS À DETECÇÃO E CLASSIFICAÇÃO DE LESÕES EM
MAMOGRAFIAS**

Tese apresentada ao Programa de Pós-graduação em Ciência da Computação, do Centro de Informática, da Universidade Federal de Pernambuco - UFPE/CIn, como requisito parcial para obtenção do grau de Doutor em Ciência da Computação.

Orientador: Prof. Dr. Abel Guilhermino da Silva Filho

Coorientador: Prof. Dr. Wellington Pinheiro dos Santos

RECIFE
2017

Catálogo na fonte
Bibliotecário Jefferson Luiz Alves Nazareno, CRB4-1758

S586m Silva, Washington Wagner Azevedo da.
Máquinas morfológicas de aprendizado extremo aplicadas à detecção e
classificação de lesões em mamografias / Washington Wagner Azevedo da
Silva. – 2017.
160 f.: il., fig., tab.

Orientador: Abel Guilhermino da Silva Filho.
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da
Computação, Recife, 2017.
Inclui referências e anexo.

1. Inteligência artificial. 2. Mamografia digital 3. Câncer de mama. I. Silva
Filho, Abel Guilhermino (orientador). II. Título.

006.3

CDD (23. ed.)

UFPE- MEI 2017-198

Washington Wagner Azevedo da Silva

**Máquinas Morfológicas de Aprendizado Extremo Aplicadas à Detecção e
Classificação de Lesões em Mamografias**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação

Aprovado em: 15/03/2017.

Orientador: Prof. Dr. Abel Guilhermino da Silva Filho

BANCA EXAMINADORA

Prof. Dr. Ricardo Bastos Cavalcante Prudêncio
Centro de Informática / UFPE

Prof. Dr. Sérgio Ricardo de Melo Queiroz
Centro de Informática / UFPE

Prof. Dr. Leandro Maciel Almeida
Centro de Informática / UFPE

Prof. Dr. Filipe Rolçim Cordeiro
Departamento de Estatística e Informática / UFRPE

Prof. Dr. Herman Martins Gomes
Departamento de Sistemas e Computação / UFCG

*Dedico este trabalho à minha mãe Selma Dário,
e aos meus irmãos Wellinton, Wilma e Jaqueline.*

Agradecimentos

A Deus, por ter me ajudado nos momentos difíceis e ter me dado saúde e força para chegar até aqui.

À minha adorada mãe Selma Dário, pelo amor, carinho, apoio e exemplo de fé. ESSA É PARA VOCÊ, MINHA MÃE! Sem você, esse sonho não teria se realizado! EU TE AMO!

Aos meus irmãos Wellinton, Wilma e Jaqueline (em especial); a minha cunhada Priscila e as minhas sobrinhas (Eduarda e Sofia) e sobrinho (Rafael); pelo apoio, momentos de descontração e felicidade. Jaqueline, sem você, esse caminho teria sido muito mais árduo! OBRIGADO!

A minha tia/mãe, Maria de Lourdes Azevedo, e ao meu tio/pai, Juarez, por todo apoio, carinho e dedicação.

A minha primeira professora Maria de Lourdes Oliveira, que me ensinou as primeiras letras do alfabeto e abriu-me as portas para o mundo maravilhoso do conhecimento.

Aos bons professores da Escola André Cordeiro, em especial as professoras Erundina Barros (in memoriam), Irene Sousa (in memoriam), Margarida Oliveira, Hécia Amaral e Mercêis Vieira.

Aos professores da Universidade Salgado de Oliveira (UNIVERSO), em especial aos professores Carla Taciana, Mirele Moutinho e Byron Leite.

Aos "meus pais da capital" Reginaldo Monteiro e Elizete Felix, pelo apoio e ajuda em momentos especiais.

Aos grandes amigos que fiz durante o doutorado, tais como, Luciano Soares, Rebecca Linhares, Sidney Lima, Marcus Vinicius Duarte e Paulo Abadie.

A todos os amigos que me incentivaram nessa jornada, em especial àqueles que estiveram mais próximos: Fábila Wanderley, Sávio Wanderley, Kátia Alessandra, Alexandre Rodrigues.

A minha psicóloga, Lúcia Lira, pelos grandes conselhos, orientações e ajuda em um dos momentos mais turbulentos da minha vida.

Ao meu orientador Abel Guilhermino, pela oportunidade de ser seu orientando, além de todas as suas dicas e conselhos. Em especial ao meu coorientador Wellington Pinheiro, pela amizade, atenção, paciência e compreensão.

À Universidade Federal de Pernambuco (UFPE) por permitir tornar possível a realização do meu trabalho e aos professores e funcionários do Centro de Informática (CIn), pelo apoio constante e informação adquiridos durante os meus estudos.

A CAPES e REUNI, por financiar este trabalho.

A todos que colaboraram, direta ou indiretamente, com o meu trabalho, e que estarei sempre em falta.

*"Eu sou um cidadão do planeta
De raízes bem singelas até a alta percepção
Sou protegido pelos anjos
E o meu corpo mostra a direção a seguir*

*Meu passatempo favorito é superar os meus limites
Apaixonado pela minha condição humana
E por esses ideais que nascem dentro de mim!"*

Citizen Of The Planet (Alanis Morissette).

Resumo

De acordo com a Organização Mundial de Saúde, o câncer de mama é a forma mais comum de câncer entre as mulheres adultas em todo o mundo, sendo um dos tipos de câncer mais fatal. Estudos mostram que a disponibilização de diagnóstico precoce pode contribuir para a redução das taxas de mortalidade e aumentar as opções de tratamento. Diagnósticos assistidos por computador (*Computer Aided Detection/Diagnosis*) estão sendo propostos e usados visando auxiliar os profissionais da área de saúde. Morfologia Matemática é uma teoria de processamento não linear amplamente utilizada no processamento de imagens digitais. É baseada na teoria matemática da intersecção e união de conjuntos. Máquinas de Aprendizado Extremo (*Extreme Learning Machines*, ELMs) são máquinas de aprendizado compostas por pelo menos uma camada escondida, com nodos com *kernels* configuráveis e pesos aleatórios, e uma camada de saída, composta por nodos com *kernels* lineares cujos pesos são ajustados de forma não iterativa, por meio da pseudoinversa de Moore-Penrose. Neste trabalho são propostas as Máquinas Morfológicas de Aprendizado Extremo (*Morphological Extreme Learning Machines*, mELMs), ou seja, ELMs com neurônios na camada escondida baseados nos operadores morfológicos não lineares básicos de erosão e dilatação. O método proposto foi avaliado utilizando 2.796 imagens de mamografias da base de dados IRMA (*Image Retrieval in Medical Applications*). Através dos experimentos realizados foi verificado que as mELMs de erosão e dilatação, no que concerne à acurácia e ao índice kappa, apresentaram desempenho de classificação equivalente ou superior, quando comparados com os resultados obtidos pelos classificadores clássicos do estado da arte. A base de dados IRMA é dividida em quatro tipos de tecidos (adiposo, fibroso, heterogeneamente denso e extremamente denso). Foram realizados experimentos com cada tipo de tecido e com toda a base de dados. Os atributos foram extraídos usando momentos de *Haralick* e *Wavelets*.

Palavras-chave: Câncer de mama. Mamografia digital. Extreme Learning Machines (ELM). Morfologia Matemática. Máquinas Morfológicas de Aprendizado Extremo (mELM).

Abstract

According to the World Health Organization, breast cancer is the most common form of cancer among adult women worldwide, being one of the most fatal types of cancer. Studies show that providing early diagnosis can contribute to a reduction in mortality rates and increase as treatment options. Computer Assisted Diagnostics are being proposed and used in the health area. Mathematical Morphology is a nonlinear theory widely used in the processing of digital images. It is based on the mathematical theory of intersection and union of sets. Extreme Learning Machines (ELMs) are learning machines composed of at least one hidden layer, with nodes with configurable kernels and random weights, and an output layer composed of nodes with linear kernels whose weights are adjusted in a non-iterative way, by means of the Moore-Penrose's pseudoinverse. In this work are proposed as Morphological Machines of Extreme Learning (mELMs), that is, ELMs with neurons in the hidden layer based on the basic non-linear morphological operators of erosion and dilation. The proposed method was evaluated using 2,796 mammography images from the IRMA database (Image Recovery in Medical Applications). Through experiments carried out with verified that as mELMs of erosion and dilatation, we presented performance of equivalent or superior classification, when compared with the results obtained by classifiers classic of the state of the art. An IRMA database is divided into four tissue types (adipose, fibrous, heterogeneously dense and dense extremity). Experiments were performed with each type of tissue and with all database. The attributes were extracted using moments from Haralick and Wavelets.

Keywords: Breast Cancer. Digital Mammography. Extreme Learning Machines (ELM). Mathematical Morphology. Morphological Extreme Learning Machines (mELM).

Lista de Figuras

Figura 1-1 Fluxo de Desenvolvimento do Projeto.	17
Figura 2-1 Processo de metástase realizado por um tumor maligno. Fonte: (PINHEIRO 2014).	22
Figura 2-2 Exame Clínico.	23
Figura 2-3 Exame de mamografia.	24
Figura 2-4 Obtenção da imagem de mamografia digital e análise do diâmetro máximo do tumor.	25
Figura 2-5 Distorção de avaliação entre dois profissionais em um exame de mamografia digital (LIMA 2016).	25
Figura 2-6 Projeções nas quais o tumor pode ser visto.	26
Figura 2-7 Resultado do exame mamográfico em diferentes visões (incidências). (RSNA - Radiological Society of North America, 2013).	27
Figura 2-8 Classificação da mama.	28
Figura 2-9 Classificação do tumor quanto ao contorno.	29
Figura 2-10 Classificação do tumor quanto à densidade.	29
Figura 3-1 Atuação da operação morfológica de erosão.	33
Figura 3-2 Atuação da operação morfológica de dilatação.	34
Figura 3-3 Classificação supervisionada.	36
Figura 3-4 Classificação não-supervisionada.	36
Figura 3-5 Neurônio artificial generalizado.	39
Figura 3-6 Arquitetura do perceptron morfológico	41
Figura 3-7 Diagrama de uma célula nervosa	42
Figura 3-8 Modelo do perceptron morfológico com dendritos.	45
Figura 3-9 Arquitetura da rede neural FLNN.	49
Figura 3-10 Estrutura da l-ésima camada de uma rede MRL.	52
Figura 3-11 (a) Estrutura da rede MMNN usando dilatações morfológicas, (b) Estrutura da rede MMNN usando erosões morfológicas.	55
Figura 3-12 (a) Arquitetura usada para a decomposição de Banon e Barrera por sup-geradores, (b) Arquitetura usada para a decomposição de Banon e Barrera por inf-geradores.	56
Figura 3-13 Algoritmo genético simples.	58
Figura 3-14 Imagem histopatológica (IKONEN et al., 2015).	59
Figura 3-15 Tecido com características de malignidade (THIRAN et al., 1996).	61
Figura 3-16 Imagem original TIFF do tecido canceroso (IKONEN et al., 2015).	63
Figura 3-17 Resultados obtidos através da equação $B-(R + G)/2$ (IKONEN et al., 2015).	64
Figura 3-18 Resultados obtidos através da equação $R-(G + B)/2$ (IKONEN et al., 2015).	64
Figura 3-19 Imagem resultante da transformada Top-Hat após a manipulação RGB (IKONEN et al., 2015).	65
Figura 3-20 Imagem resultante da transformada Top-Hat sem a manipulação RGB (IKONEN et al., 2015).	66
Figura 3-21 Algoritmo de decomposição de <i>Wavelets</i> bi-dimensionais, para um nível de resolução.	78
Figura 3-22 Ilustração da decomposição das <i>wavelets</i> , em 2 níveis, baseada no algoritmo <i>Daubechies 8</i>	79
Figura 4-1 Exemplo do uso da aproximação proposta de erosão em uma imagem binária.	88
Figura 4-2 Exemplo do uso da aproximação proposta de dilatação em uma imagem binária.	89
Figura 4-3 Diagrama da metodologia proposta.	91
Figura 4-4 Extração de características (descritores de Haralick).	92
Figura 4-5 Exemplo do processo de classificação utilizado o método proposto mELM (erosão).	93
Figura 4-6 Apresentando o vetor de entradas ao mELM (erosão) e iniciando o processo de treinamento.	94
Figura 4-7 Cálculo do produtório.	95
Figura 4-8 Resultado da saída do primeiro neurônio erosão da camada intermediária.	95
Figura 4-9 Mamografias de diferentes tecidos da mama: (a) tecido adiposo, (b) tecido fibroso, (c) tecido heterogeneamente denso e (d) tecido extremamente denso.	96
Figura 4-10 Exemplo de falta de cuidado ao digitalizar as imagens de mamografia por parte dos especialistas.	98
Figura 4-11 Exemplo de regiões de interesse de imagens de mamografia da base de dados IRMA.	98
Figura 5-1 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido adiposo utilizando Atributos de <i>Haralick</i> (HA).	103
Figura 5-2 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido fibroso utilizando Atributos de <i>Haralick</i> (HA).	104
Figura 5-3 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido heterogeneamente denso utilizando Atributos de <i>Haralick</i> (HA).	104

Figura 5-4 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido extremamente denso utilizando Atributos de <i>Haralick</i> (HA).	105
Figura 5-5 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com todos os tipos de tecidos utilizando Atributos de <i>Haralick</i> (HA).	105
Figura 5-6 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido adiposo utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	110
Figura 5-7 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido fibroso utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	110
Figura 5-8 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido heterogeneamente denso utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	111
Figura 5-9 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido extremamente denso utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	111
Figura 5-10 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com todos os tipos de tecidos utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	112
Figura 8-1 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido adiposo utilizando Atributos de <i>Haralick</i> (HA).	134
Figura 8-2 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido fibroso utilizando Atributos de <i>Haralick</i> (HA).	135
Figura 8-3 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Heterogeneamente denso utilizando Atributos de <i>Haralick</i> (HA).	136
Figura 8-4 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Extremamente denso utilizando Atributos de <i>Haralick</i> (HA).	137
Figura 8-5 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com todos os tipos de tecidos utilizando Atributos de <i>Haralick</i> (HA).	138
Figura 8-6 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido adiposo utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	143
Figura 8-7 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido fibroso utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	144
Figura 8-8 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Heterogeneamente denso utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	145
Figura 8-9 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Extremamente denso utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	146
Figura 8-10 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com todos os tipos de tecidos utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).	147

Lista de Tabelas

Tabela 3-1 Dilatação, erosão e convolução de imagens.	72
Tabela 3-2 Matriz de Confusão	80
Tabela 3-3 Índice Kappa e a performance correspondente ao classificador	81
Tabela 3-4 Principais Redes Neurais utilizadas no estado da arte para a detecção de câncer de mama em imagens de mamografia.	83
Tabela 4-1 Operações de mínimo e máximo associadas às propriedades da álgebra de Boole, representados em (a) e (b), respectivamente.	85
Tabela 5-1 Resultados obtidos pelos dois melhores e os dois piores classificadores utilizando Atributos de <i>Haralick</i> (HA).....	102
Tabela 5-2 Teste de <i>Wilcoxon</i> não-paramétrico comparando os dois melhores e os dois piores classificadores utilizando Atributos de <i>Haralick</i>	106
Tabela 5-3 Resultados obtidos pelos dois melhores e os dois piores classificadores utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).....	109
Tabela 5-4 Teste de <i>Wilcoxon</i> não-paramétrico comparando os dois melhores e os dois piores classificadores utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i>	112
Tabela 6-1 Publicações.....	118
Tabela 8-1 Resultados expandidos, referentes à Tabela 5-1. Todos os resultados obtidos com os métodos propostos e os principais classificadores do estado da arte utilizando Atributos de <i>Haralick</i> (HA).	130
Tabela 8-2 Resultados expandidos, referentes à Tabela 5-3. Todos os resultados obtidos com os métodos propostos e os principais classificadores do estado da arte utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i> (WA-HA).....	139
Tabela 8-3 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de <i>Wilcoxon</i> não-paramétrico comparando o método proposto, mELM (dilatação), com os principais classificadores do estado da arte. Nos experimentos foram utilizando Atributos de <i>Haralick</i>	148
Tabela 8-4 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de <i>Wilcoxon</i> não-paramétrico comparando o método proposto, mELM (dilatação), com os principais classificadores do estado da arte. Nos experimentos foram utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i>	150
Tabela 8-5 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de <i>Wilcoxon</i> não-paramétrico comparando o método proposto, mELM (erosão), com os principais classificadores do estado da arte. Nos experimentos foram utilizando Atributos de <i>Haralick</i>	152
Tabela 8-6 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de <i>Wilcoxon</i> não-paramétrico comparando o método proposto, mELM (erosão), com os principais classificadores do estado da arte. Nos experimentos foram utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i>	154
Tabela 8-7 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de <i>Wilcoxon</i> não-paramétrico comparando o classificador SVM <i>kernel</i> linear com os métodos propostos e com os principais classificadores do estado da arte. Nos experimentos foram utilizando Atributos de <i>Haralick</i>	156
Tabela 8-8 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de <i>Wilcoxon</i> não-paramétrico comparando o classificador SVM <i>kernel</i> linear com os métodos propostos e com os principais classificadores do estado da arte. Nos experimentos foram utilizando <i>Wavelets</i> combinados com Atributos de <i>Haralick</i>	158

Principais Abreviações

Abreviações	Significado
<i>BI-RADS</i>	<i>Breast Image Reporting and Data System</i>
<i>CAD</i>	<i>Computer Aided Diagnosis</i>
<i>CL</i>	<i>Competitive Learning Network</i>
<i>CNN</i>	<i>Cellular Neural Network</i>
<i>DELP</i>	<i>Dilation-Erosion-Linear Perceptron</i>
<i>ELM</i>	<i>Extreme Learning Machine</i>
<i>fmELM</i>	<i>Fuzzy Morphological Extreme Learning Machines</i>
<i>IRMA</i>	<i>Image Retrieval in Medical Applications</i>
<i>LVQ</i>	<i>Learning Vector Quantization</i>
<i>mELM</i>	<i>Morphological Extreme Learning Machine</i>
<i>OMS</i>	<i>Organização Mundial de Saúde</i>
<i>PNN</i>	<i>Probabilistic Neural Networks</i>
<i>RBF</i>	<i>Radial Basis Function</i>
<i>RNA</i>	<i>Redes Neurais Artificiais</i>
<i>SOM</i>	<i>Self-Organizing Map</i>
<i>FLNN</i>	<i>Fuzzy Lattice Neural Network</i>

Sumário

1	INTRODUÇÃO	14
1.1	Motivação	15
1.2	Fluxo de Desenvolvimento de Projeto	16
1.3	Pergunta de pesquisa	18
1.4	Objetivos	18
1.4.1	<i>Objetivo Geral</i>	<i>18</i>
1.4.2	<i>Objetivos Específicos</i>	<i>18</i>
1.5	Estrutura do Documento	19
2	CÂNCER DE MAMA	21
2.1	Introdução	22
2.2	Métodos Clássicos para Detecção de Lesões de Mama	23
a)	Exame Clínico	23
b)	Mamografia	24
c)	Ultra-sonografia e Mamografia Digital	24
2.3	Comprometimento da Pele quanto à Orientação do Tumor	26
2.4	Classificação da Mama e Tumores em Mamografias	27
2.5	Considerações do Capítulo	30
3	REVISÃO BIBLIOGRÁFICA	31
3.1	Morfologia Matemática	32
3.2	Reconhecimento de Padrões	35
3.3	Redes Neurais Morfológicas	38
3.4	Extreme Learning Machines	66
3.5	Operadores Morfológicos	68
3.6	Descritores de Imagem	73
3.7	Métricas	80
3.8	Trabalhos Relacionados	81
3.9	Considerações do Capítulo	83
4	METODOLOGIA PROPOSTA	84
4.1	Aproximações Morfológicas em Imagens Binárias	85
4.2	Proposta: Neurônios Morfológicos	90
4.3	Detecção e Classificação de Lesões em Mamografias	91
4.4	Base de Dados IRMA	96
4.5	Considerações do Capítulo	99
5	RESULTADOS	100
5.1	Introdução	101
5.2	Resultados e Análise Experimental	102
5.3	Considerações do Capítulo	115
6	CONCLUSÕES	116
6.1	Considerações Finais	117
6.2	Contribuições	118
6.3	Trabalhos Futuros	119
	REFERÊNCIAS	120
	ANEXO A - Resultados Expandidos do Capítulo 5	130

1 INTRODUÇÃO

Este capítulo relata as principais motivações para realização deste trabalho, apresentando o contexto sobre a necessidade de um diagnóstico precoce e preciso na identificação de lesões em mama. São também apresentados neste capítulo os objetivos gerais e específicos do trabalho, o escopo da tese e, por fim, é descrito como está estruturado o restante da tese.

1.1 Motivação

O câncer de mama é o mais comum em mulheres em todo o mundo: a Organização Mundial de Saúde (OMS) estima a ocorrência de 1,7 milhões de novos casos por ano (DESANTIS et al., 2014). As taxas de sobrevivência para o câncer de mama podem variar de 80%, em países de alta renda, abaixo de 40% em países de baixa renda (COLEMAN et al., 2008). A baixa taxa de sobrevivência em alguns países está relacionada com a falta de programas de detecção precoce, com um impacto importante para o sucesso do tratamento do câncer, uma vez que o tratamento médico se torna mais difícil nos últimos estágios. Um dos métodos mais eficazes para a detecção de possíveis tumores de mama é a mamografia digital (MAITRA et al., 2011). No entanto, a compreensão visual e análise da mamografia pode ser uma tarefa difícil, mesmo para um especialista, uma vez que tal procedimento pode ser afetado por aspectos de qualidade de imagem, a experiência do radiologista, e forma do tumor.

O diagnóstico por imagem é uma tarefa complexa, devido à grande variabilidade de casos clínicos. Muitos casos observados na prática clínica não correspondem com imagens clássicas e descrições precisamente (JUHL et al., 2000). Por estas razões, *Mammography Computer Aided Diagnosis* (CAD) vem desempenhar um papel de importância para ajudar os radiologistas a melhorar a precisão do diagnóstico. Consequentemente, as técnicas tradicionais de processamento e análise de imagens foram aplicadas no campo médico para fazer o diagnóstico menos suscetível a erros através da identificação precisa das anomalias anatômicas (DA-XI et al., 2010), (YE et al., 2010).

Extreme Learning Machines (ELMs) são máquinas de aprendizado baseadas em *kernels* flexíveis cujas principais características são o treinamento rápido, baseado em pesos aleatórios nos nodos das camadas escondidas e um ótimo desempenho de classificação (HUANG et al., 2012). Basicamente, uma ELM é composta por pelo menos uma camada escondida, composta por neurônios artificiais com *kernels* configuráveis, e uma camada de saída, composta por nodos com *kernels* lineares (HUANG et al., 2012), (CAMBRIA et al., 2013). Dados o número de camadas escondidas e neurônios por camadas, os pesos destes nodos são selecionados aleatoriamente, enquanto que os pesos sinápticos dos neurônios da camada de

saída são determinados de uma forma não iterativa, por meio da pseudoinversa de Moore-Penrose (HUANG et al., 2012).

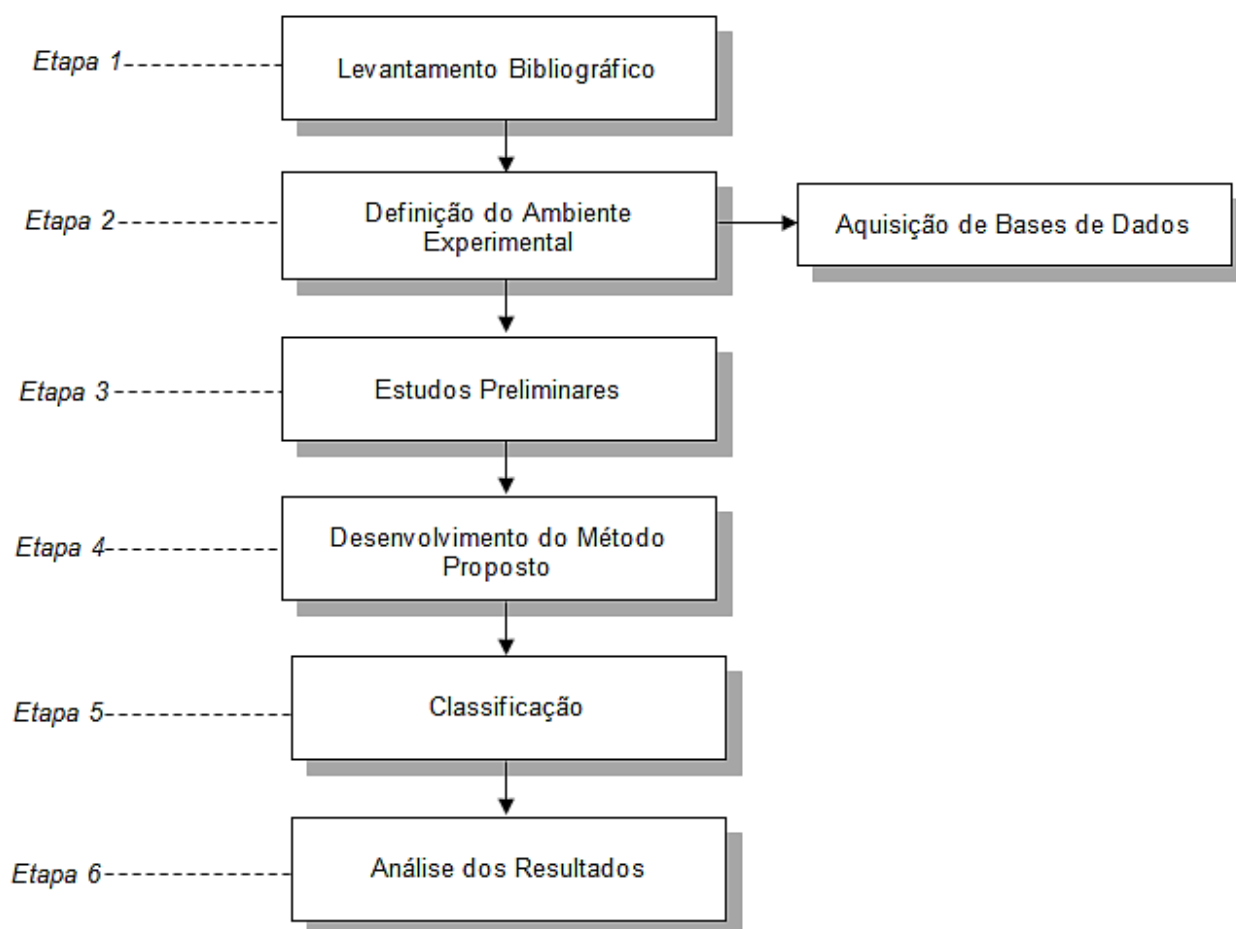
As redes neurais morfológicas emergiram da álgebra de imagens e de suas similaridades com as redes neurais artificiais (SANTOS et al. 2012). Nas redes neurais morfológicas, os neurônios são baseados em operações da Morfologia Matemática, como a erosão e a dilatação, ou de combinações destas, como no perceptron morfológico (SANTOS et al. 2012).

Uma característica interessante das redes neurais morfológicas é sua capacidade para construir fronteiras de decisão não lineares e relativamente complexas, que podem ajudar na resolução de problemas difíceis de classificação (SANTOS et al. 2012). Contudo, a principal limitação dessas redes é a dificuldade de implementação de algoritmos de treinamento e muitas vezes a sua ausência, o que faz com que, por vezes, sejam utilizados algoritmos de otimização para treinamento (SANTOS et al. 2012).

Neste trabalho são propostas as Máquinas Morfológicas de Aprendizado Extremo (*Morphological Extreme Learning Machines*, mELMs), ELMs com *kernels* na camada escondida baseados em operadores morfológicos não lineares básicos de erosão e dilatação. O objetivo foi buscar unir a capacidade das redes neurais morfológicas de construir fronteiras de decisão não lineares complexas com a simplicidade dos algoritmos de treinamento das ELMs. Como estudo de caso, as mELMs foram aplicadas à tarefa de detectar e classificar lesões em regiões de interesse previamente selecionadas em mamografias. Os resultados são comparados com classificadores clássicos do estado da arte e avaliados através de métricas de classificação amplamente utilizadas.

1.2 Fluxo de Desenvolvimento de Projeto

O fluxo de desenvolvimento do projeto mostra todas as etapas envolvidas na implementação da Tese. A Figura 1-1 mostra o fluxo de projeto para o desenvolvimento do trabalho.

Figura 1-1 Fluxo de Desenvolvimento do Projeto.

Inicialmente, houve o levantamento bibliográfico. O objetivo foi estudar os trabalhos relevantes quanto à detecção e classificação de lesões em mamografias.

Após o levantamento do estado da arte, na segunda etapa foi observado quais bases de dados são utilizadas no estado da arte e foi feita a definição do ambiente experimental, onde foram selecionadas as bases de dados de mamografia digital e definido o conjunto de imagens a ser trabalhado. Na terceira etapa, foram realizados estudos preliminares com os principais algoritmos de classificação apontados no estado da arte.

Na quarta etapa foi realizado a implementação da metodologia proposta, em seguida, na quinta etapa foram realizados experimentos com o método proposto e com outros classificadores apresentados no estado da arte. Por fim, na sexta etapa foi realizado a análise dos resultados obtidos pelas duas novas Redes Nerais propostas, mELM de Dilatação e mELM de erosão, realizando um comparativo entre os resultados alcançados pelos novos classificadores e os principais classificadores citados no levantamento bibliográfico. A descrição

detalhada da quarta, quinta e sexta etapas pode ser encontrada no quarto e quinto capítulo dessa dissertação.

1.3 Pergunta de pesquisa

É possível obter uma técnica para detecção e classificação de lesões através de imagens de mamografia com baixo custo computacional, e que seja capaz de construir fronteiras de decisão não lineares complexas?

- **Hipótese:** Usar redes neurais baseados em pesos aleatórios nas camadas escondidas e baseadas em funções de ativação complexas, não lineares, baseadas na morfologia matemática para modelar fronteiras de precisão complexas, buscando resolver o problema de apoio a diagnóstico do câncer de mama utilizando imagens de mamografia digitais.
- **Proposta:** Desenvolvimento de redes neurais de pesos aleatórios com funções de ativação baseadas na morfologia matemática para construir classificadores para apoio à diagnósticos por imagem usando mamografia digitais.

1.4 Objetivos

Baseado no problema apresentado e na hipótese descrita, este trabalho possui um objetivo geral pelo qual são direcionados todos os esforços, os quais são definidos por alguns objetivos específicos que precisam ser progressivamente alcançados. As próximas subseções abordam cada um deles.

1.4.1 Objetivo Geral

O trabalho proposto tem como objetivo principal, a aplicação e análise de Máquinas Morfológicas de Aprendizado Extremo (mELM), à detecção de lesões de mama em imagens de mamografia digital, a fim de permitir um diagnóstico mais preciso das alterações encontradas, além de agilizar o processo de diagnóstico e fornecer uma ferramenta de suporte à decisão para o profissional médico.

1.4.2 Objetivos Específicos

O objetivo geral pode ser decomposto nos seguintes objetivos específicos:

1. Estudar criticamente os principais classificadores utilizados na tarefa de detecção de lesões em mamografia, referenciados na literatura;
2. Levantar o estado da arte do uso de ELM como classificador de lesões em imagens de mamografia;
3. Desenvolvimento e análise de uma arquitetura de rede ELM morfológica (mELM), utilizando *kernels* de dilatação e erosão na camada intermediária;
4. Realizar experimentos utilizando os principais classificadores referenciados na literatura, aplicados a tarefa de classificação de tumores em mamografia;
5. Realizar experimentos com a rede mELM, utilizando *kernels* de dilatação e erosão, na classificação de tumores em imagens de mamografia;
6. Análise dos classificadores clássicos, apontados na literatura, e das mELMs, através de métricas de precisão em relação à taxa de classificação;
7. Análise da melhora quanto à taxa de classificação obtida pelas redes mELMs, confrontando os resultados obtidos com as técnicas clássicas apontadas no estado da arte.

1.5 Estrutura do Documento

Este documento, que contém a Monografia de Qualificação, encontra-se estruturado da seguinte forma:

- **Capítulo 1: Introdução**

O primeiro capítulo contextualizou este trabalho através do relato das principais motivações envolvidas, listando os objetivos de pesquisa almejados e por fim, discriminando como está estruturado o restante do documento.

- **Capítulo 2: Câncer de mama**

Esse capítulo apresenta uma breve descrição sobre câncer de mama, os principais métodos utilizados para obtenção de imagens digitais da mama e as dificuldades encontradas pelos profissionais de saúde na realização de diagnósticos. Também é apresentado a descrição dos diferentes tipos de tumor de mama quanto ao contorno.

- **Capítulo 3: Revisão Bibliográfica**

Mostra as principais contribuições produzidas nos últimos anos no contexto de detecção de lesões em imagens de mamografia. Também, serão evidenciados os principais trabalhos que serviram como base para o desenvolvimento da pesquisa proposta.

- **Capítulo 4: Metodologia Proposta**

Apresentar a metodologia utilizada ao longo do trabalho de pesquisa. Também, é descrita em maiores detalhes, a base de dados IRMA, que foi utilizada durante os experimentos com o método proposto.

- **Capítulo 5: Resultados**

Apresenta os resultados alcançados utilizando o método proposto (mELM) e as técnicas do estado da arte, utilizando a base de imagens de mamografia IRMA.

- **Capítulo 6: Conclusões**

Apresenta algumas considerações finais sobre o trabalho proposto, incluindo as contribuições alcançadas e indicações para trabalhos futuros.

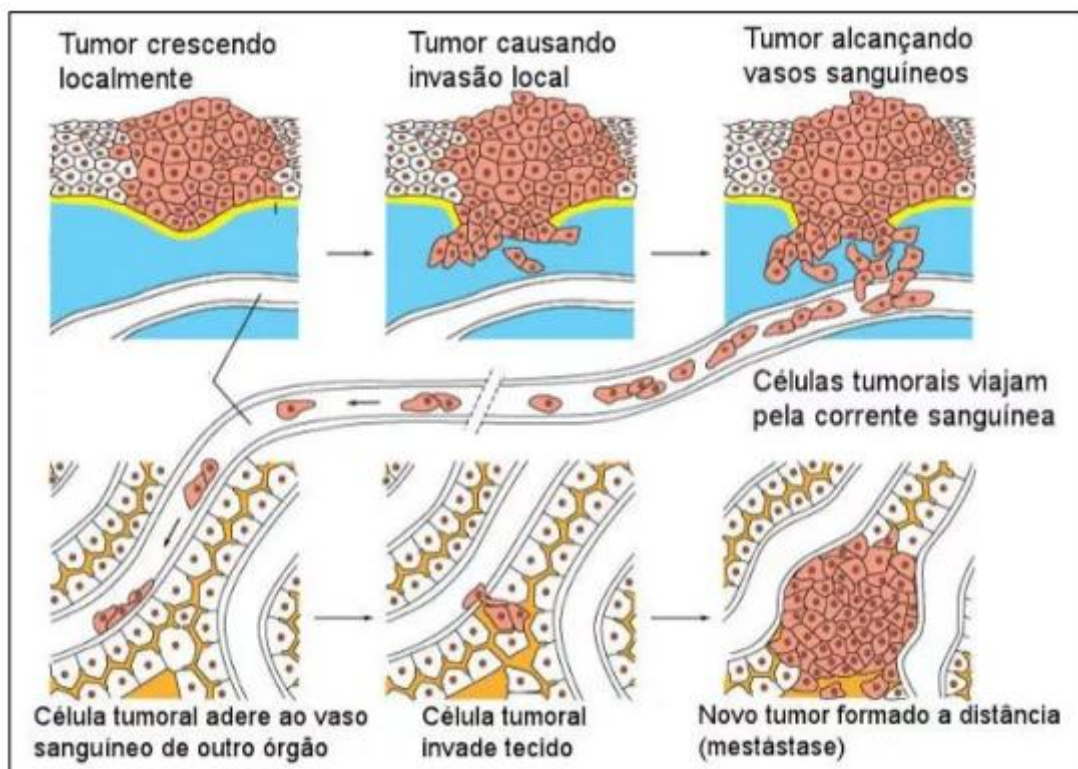
2 CÂNCER DE MAMA

Esse capítulo apresenta uma breve descrição sobre câncer de mama, os principais métodos utilizados para obtenção de imagens digitais da mama e as dificuldades encontradas pelos profissionais de saúde na realização de diagnósticos. Também é apresentado a descrição dos diferentes tipos de tumor de mama quanto ao contorno.

2.1 Introdução

(PINHEIRO 2014) afirma que um câncer inicia-se a partir de uma célula com informações genéticas incorretas que a incapacitam no cumprimento das funções para as quais foi designada. Quando esta célula reproduz novas células com a mesma construção genética incorreta, um tumor começará a ser formado, sendo composto por um aglomerado dessas células imperfeitas. Nos casos em que a doença se agrava, células malignas se separam do tumor original podendo espalhar-se no local e/ou à distância em relação ao tumor primário, disseminando a doença em outras partes do corpo, formando as metástases. O processo de metástase é ilustrado na Figura 2-1.

Figura 2-1 Processo de metástase realizado por um tumor maligno. Fonte: (PINHEIRO 2014).



A melhor forma de tratamento contra o câncer é a detecção precoce do tumor. Quanto mais cedo for identificada a lesão, mais fácil e eficiente será o tratamento da doença. No entanto, identificar o tumor em estágios iniciais é uma tarefa difícil e, por consequência, existe uma maior dependência de equipamentos de obtenção de imagem e técnicas de segmentação mais eficientes, além de uma maior experiência do profissional de radiologia.

2.2 Métodos Clássicos para Detecção de Lesões de Mama

De acordo com o documento de consenso do Ministério da Saúde, os métodos de detecção e medição de lesões em imagens médicas de mama são três: exame clínico, ultrassonografia e mamografia (COSTA et al., 2004).

a) Exame Clínico

O exame clínico de toque é um procedimento completamente manual. Este procedimento é realizado com a paciente sentada ou deitada e com os braços relaxados. O examinador apalpa a mama da paciente, tendo com objetivo, encontrar mudanças nos mamilos, dobras ou retrações na pele.

Figura 2-2Exame Clínico.



Quando o profissional de saúde detecta uma lesão, ela é fixada entre os dedos e mensurada através de uma régua milimetrada flexível (MONTEIRO et al., 2008). Esse método é muito dependente da experiência do examinador. O exame clínico não deve ser a única forma de diagnóstico adotada, pois o tumor leva cerca de 8 anos para se tornar palpável (DE ABREU et al., 2002). Desta forma, quando ele se torna palpável o mesmo pode se encontrar em um estágio desenvolvido.

b) Mamografia

A mamografia ou mastografia é um exame de diagnóstico por imagem, que tem como objetivo estudar o tecido mamário. Esse tipo de procedimento pode detectar um nódulo, mesmo que este ainda não seja palpável. Para tanto é utilizado um equipamento que usa uma fonte de raios-x, para obtenção de imagens radiográfica do tecido mamário. A Figura 2-3 demonstra um exame típico de mamografia.

Figura 2-3 Exame de mamografia.



De acordo com o Ministério da Saúde, mulheres com idade acima de 40 anos devem realizar o exame de mamografia anualmente, enquanto que as mulheres pertencentes ao grupo de risco elevado devem realizar o exame a cada seis meses, a partir dos 35 anos. Para as mulheres que se enquadram na faixa etária de 50 a 69 anos é recomendado o procedimento de mamografia a cada dois anos.

A dimensão da lesão é aferida na mastografia através de uma régua milimetrada aplicada ao filme resultante do exame. Por se tratar de um procedimento manual existe uma morosidade associada além de aumentar a possibilidade de erros de medição.

c) Ultra-sonografia e Mamografia Digital

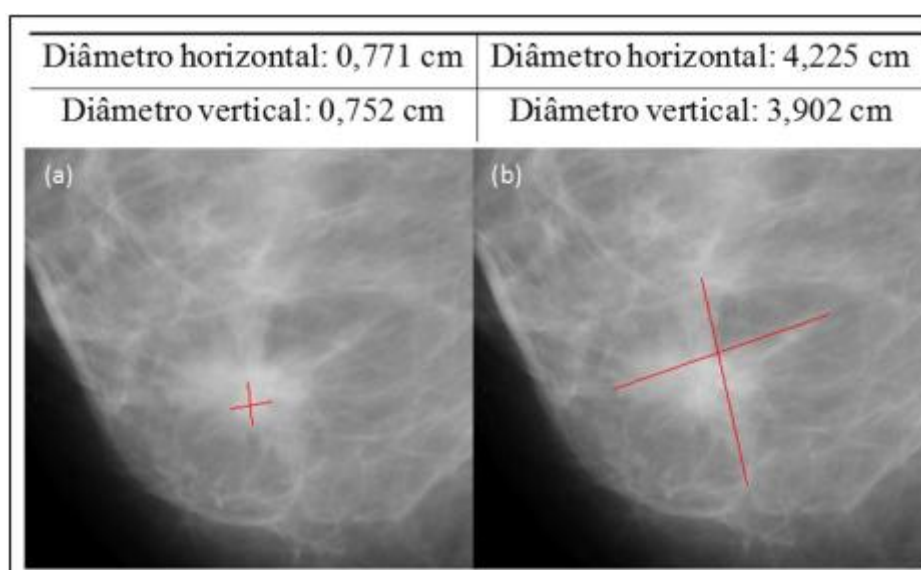
A Figura 2-4 demonstra o processo de obtenção e análise de lesão através do exame de mamografia digital. Inicialmente a paciente é submetida ao equipamento para obtenção da imagem da mama, onde são capturadas imagens digitais do seio, chamadas de mamogramas. Ao obter a imagem de mamografia digital, o radiologista realiza uma análise com a finalidade

de detectar alguma região de lesão. Caso alguma lesão seja encontrada, o radiologista irá analisar essa lesão quanto à forma e ao seu diâmetro máximo para poder fornecer um diagnóstico. No entanto essa análise ainda é bastante dependente da experiência do profissional em radiologia. Como exemplo, a imagem da Figura 2-5, analisada na tese de (LIMA 2016), mostra a diferença de análise de uma mesma imagem feita por dois radiologista da cidade de Recife.

Figura 2-4Obtenção da imagem de mamografia digital e análise do diâmetro máximo do tumor.



Figura 2-5Distorção de avaliação entre dois profissionais em um exame de mamografia digital(LIMA 2016).



Na Figura 2-5, pode-se observar que existe uma discrepância entre as avaliações realizadas pelos dois radiologistas. Nessas avaliações, que são utilizadas na prática clínica local, o médico tenta identificar os diâmetros máximos horizontal e vertical. Caso o tamanho do diâmetro máximo do tumor ultrapasse 3 cm, na maioria dos casos é indicado a realização de uma cirurgia de remoção de mama. Porém, existe uma grande inconsistência nesse método, pois a lesão completa é discretizada em apenas 4 pontos, que são eles, dois verticais e dois

horizontais, e todas as suas demais informações são desconsideradas. Outro equívoco se dá pelo fato dos pontos formadores do diâmetro, serem inseridos manualmente, através do uso de um mouse. No caso da Figura 2-5 (a), por exemplo, um radiologista identificou a região da lesão como sendo pequena, com diâmetro menor que 1 cm. Já no caso da Figura 2-5 (b), o outro profissional analisou a mesma imagem e indicou que o diâmetro máximo era superior a 3 cm. A disparidade de diagnóstico, no entanto, interfere diretamente na qualidade de vida da paciente e tem grande impacto psicológico, onde o diagnóstico errado pode causar a remoção desnecessária da mama, quanto no tratamento errado de um tumor em estágio avançado.

A ultra-sonografia, por sua vez, disponibiliza as mesmas ferramentas que a mamografia digital para as imagens de ultra-som. Portanto, as mesmas imprecisões descritas para a mamografia digital podem ser geradas nas análises das imagens de ultra-som. Apesar de haverem vários estudos sobre qual dos métodos, vistos na seção, é o melhor na investigação do tamanho da lesão na mama (HERRADA et al., 1997), (FIORENTINO et al., 2001), (FORNAGE et al., 1987), (MEDEN et al., 1995), (PIERIE et al., 1998), (BOSCH et al., 2003), (SHOMA et al., 2003), não há um consenso sobre qual procedimento é o mais eficaz. Todos eles são dependentes demasiadamente da experiência do profissional examinador. Logo, a mensuração da lesão, através desses exames, se torna bastante subjetiva e sujeita a diversas interpretações.

2.3 Comprometimento da Pele quanto à Orientação do Tumor

O diâmetro máximo do tumor é de extrema importância no estudo que se refere ao comprometimento da pele do paciente. Um tumor deve ser visto nas projeções crânio-caudal e médio-lateral, como pode ser mostrado na Figura 2-6 e Figura 2-7. Se o tumor em potencial for visto em apenas uma projeção, ele é denominado assimetria.

Figura 2-6 Projeções nas quais o tumor pode ser visto.

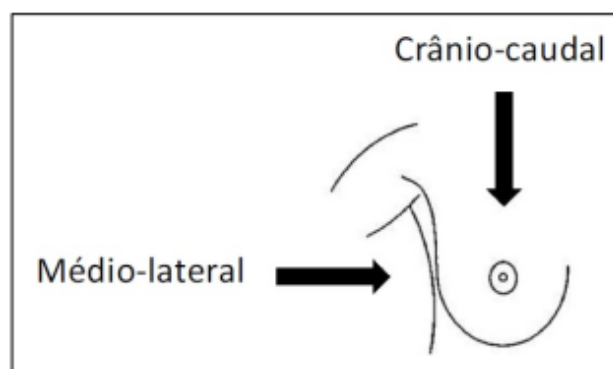
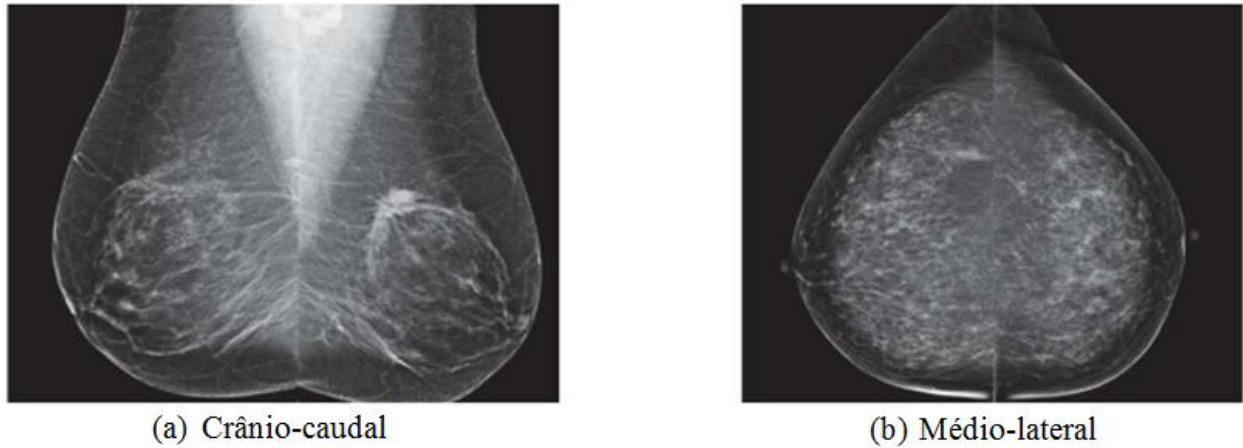


Figura 2-7 Resultado do exame mamográfico em diferentes visões (incidências). (RSNA - Radiological Society of North America, 2013).



Um tumor tem característica benigna quando o diâmetro máximo está paralelo à pele. Isso quer dizer que na projeção crânio-caudal o diâmetro máximo deve ter orientação horizontal, ou na projeção médio-lateral, o máximo diâmetro deve ter orientação vertical.

Um tumor tem característica maligna quando o diâmetro máximo está perpendicular à pele, chamado popularmente de "câncer enraizado". Então na projeção médio-lateral, o máximo diâmetro deve ter orientação horizontal.

Isso mostra como é importante a definição exata do contorno do tumor, obtendo seu diâmetro máximo, para identificar o diagnóstico correto.

2.4 Classificação da Mama e Tumores em Mamografias

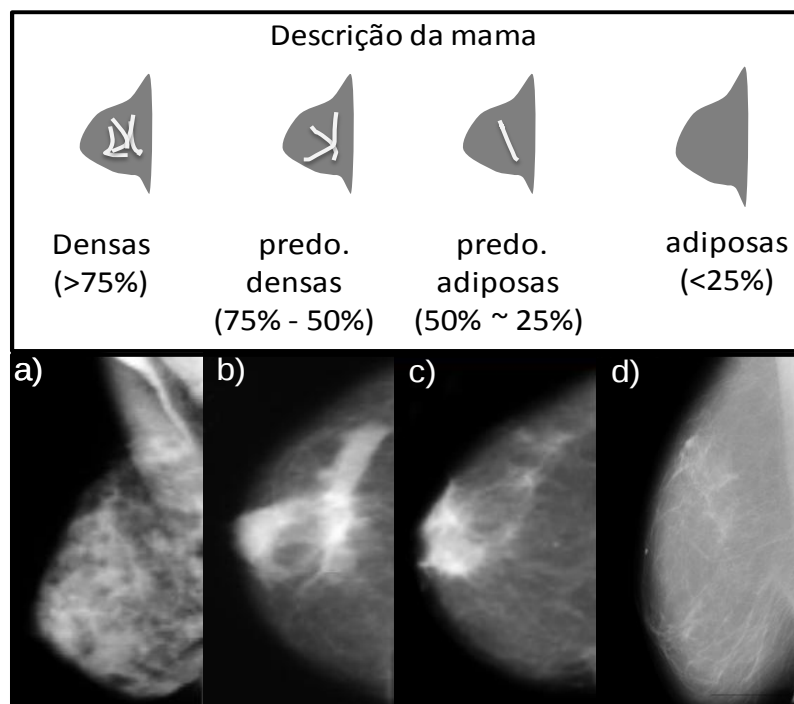
A classificação proposta pelo Colégio Americano de Radiologia, desenvolvida pelo Registro de Imagem da Mama e Sistema de Dados (Breast Imaging Reporting and Data System, BI-RADS) (LIBERMAN et al., 2002), reconhecida pelo Colégio Brasileiro de Radiologia por Imagem (CBR), através da Comissão Nacional de Mamografia, é a que tem a maior aceitação por apresentar simplicidade, abrangência e fácil entendimento (TEIXEIRA 2011). Essa classificação considera quatro padrões de densidade do tecido mamário:

- Tipo I: Mamas predominantemente adiposas (25% do componente fibroglandular);
- Tipo II: Mamas parcialmente gordurosas (com densidades de tecido fibroglandular ocupando de 26% a 50% do volume da mama);

- Tipo III: Mamas com padrão denso e heterogêneo (51% a 75% de tecido fibroglandular, dificultando a visualização de nódulos);
- Tipo IV: Mamas muito densas, apresentando mais de 75% de tecido fibroglandular, diminuindo a sensibilidade da mamografia.

A Figura 2-8 mostra imagens de mamografia para cada tipo de tecido, segundo a classificação do BI-RADS.

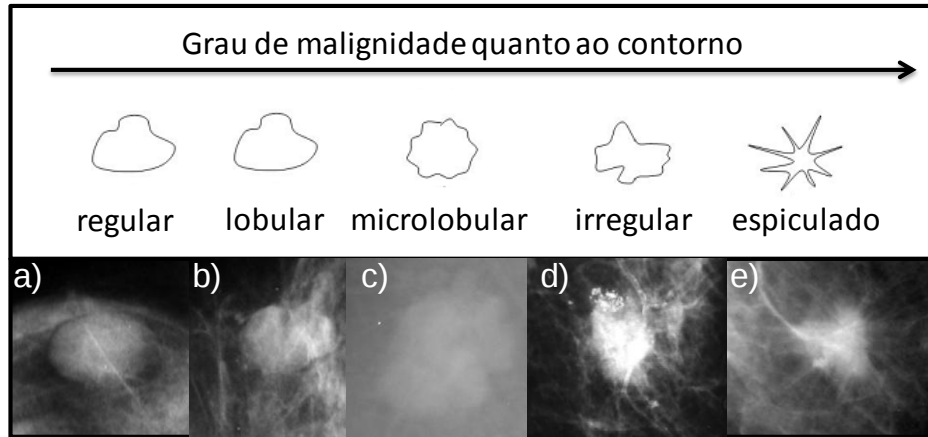
Figura 2-8 Classificação da mama.



Nas mulheres jovens, comumente existe pouco tecido adiposo, apresentando mamas mais densas. Os músculos e demais tecidos mamários, chamados tecnicamente de parênquima, ocupam maior parte da mama. Nas mamas com características predominantemente densas, o parênquima mamário ocupa entre 75% e 50% da mama. Em uma mama predominantemente adiposa, o parênquima mamário ocupa entre 50% e 25% da mama. Já em uma mama adiposa, existe pouco parênquima, menos de 25% e muito tecido. Comumente as pacientes idosas apresentam mamas com características adiposas. À medida que ocorre o envelhecimento da mulher, existe uma tendência dos músculos e demais tecidos mamários serem substituídos por tecidos adiposos.

O estudo do formato da lesão é de extrema importância para a escolha do tratamento apropriado. O tumor pode ser classificado em cinco grupos: regular, lobular, microlobular, irregular e espiculado, como podem ser observados na Figura 2-9. Um tumor regular tem características benignas, enquanto uma lesão espiculada costuma ser maligna.

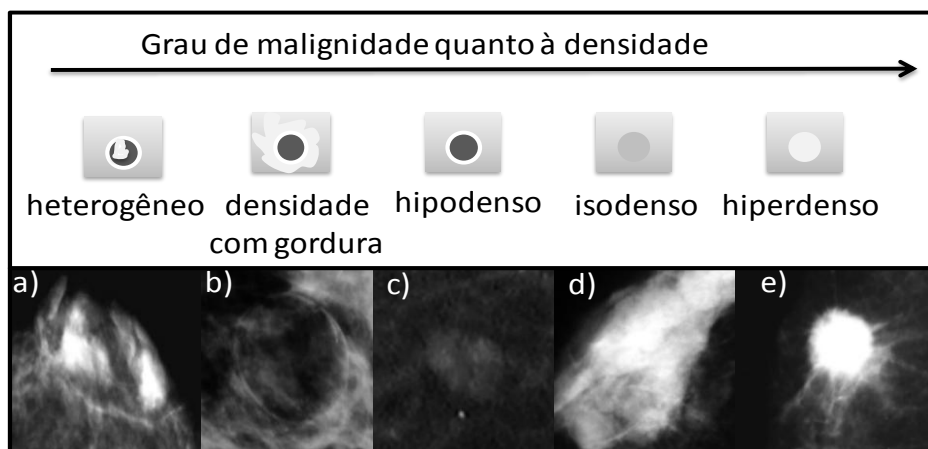
Figura 2-9 Classificação do tumor quanto ao contorno.



Um tumor regular apresenta uma borda pouco rugosa, com suaves alterações. Um tumor lobular tem um contorno ondulado. Um tumor microlobular possui uma borda com pequenas ondulações. Um tumor espiculado possui linhas radiantes nas margens do tumor. Por fim, um tumor irregular não se encaixa na definição de nenhum outro agrupamento.

Quanto à densidade, o INCA classifica o tumor em 5 grupos: heterogêneo, densidade com gordura, hipodenso, isodenso e hiperdenso, exibidas Figura 2-10.

Figura 2-10 Classificação do tumor quanto à densidade.



Um tumor hiperdenso apresenta densidade mais elevada do que à pele. Tumores isodenso tem densidade igual ao tecido mamário. Tumores hipodensos possuem densidade inferior à pele. Tumores com densidade de gordura são hipodensos e contém gorduras nas suas redondezas. Por último, tumores heterogêneos são hipodensos e parcialmente ocupados por parênquima mamário.

Teoricamente, tumores malignos apresentam densidade superior em relação à mama. Já tumores com densidade inferior à mama, geralmente, são benignos. O BI-RADS, por exemplo, cita que tumores de densidade com gordura são benignos, quase que na sua totalidade (BI-RADS 2003).

2.5 Considerações do Capítulo

Neste capítulo foram apresentados alguns dos principais métodos e aparelhos de detecção de lesões em mama. Ficou evidente o quanto esse tipo de procedimento ainda é dependente da experiência do profissional da área de saúde, e o quanto um diagnóstico incorreto pode causar impactos na vida de uma pessoa. Finalmente, é possível verificar como a utilização de técnicas computacionais, através de ferramentas que ajudam no diagnóstico, podem trazer benefícios na prevenção do câncer e na redução do percentual de óbitos causados pela doença.

3 REVISÃO BIBLIOGRÁFICA

O objetivo deste capítulo é mostrar as principais contribuições produzidas nos últimos anos no contexto de detecção de lesões em imagens de mamografia. Também, serão evidenciados os principais trabalhos que serviram como base para o desenvolvimento da pesquisa proposta.

3.1 Morfologia Matemática

Morfologia Matemática é uma teoria completa de processamento não linear amplamente utilizado no processamento de imagens digitais. Várias aplicações específicas são construídas a partir da Morfologia Matemática como detecção e segmentação de objetos, extração de características dentre outras (SOILLE 2004), (HÄNNI et al., 1999), (THANH et al., 2007) e (SANTOS et al., 2012). A morfologia é baseada nas transformações de formas preservando as relações de inclusão dos objetos. Há duas operações morfológicas fundamentais: erosão e dilatação (SOILLE 2004), (HÄNNI et al., 1999), (THANH et al., 2007) e (SANTOS et al., 2012). A Morfologia Matemática pode ser considerada uma teoria construtiva porque todas as operações são construídas tendo com base erosões e dilatações. Matematicamente, erosão e dilatação são formalizados de acordo com as Equação (3-1) e a Equação (3-2), respectivamente (SOILLE 2004), (HÄNNI et al., 1999), (THANH et al., 2007) e (SANTOS et al., 2012):

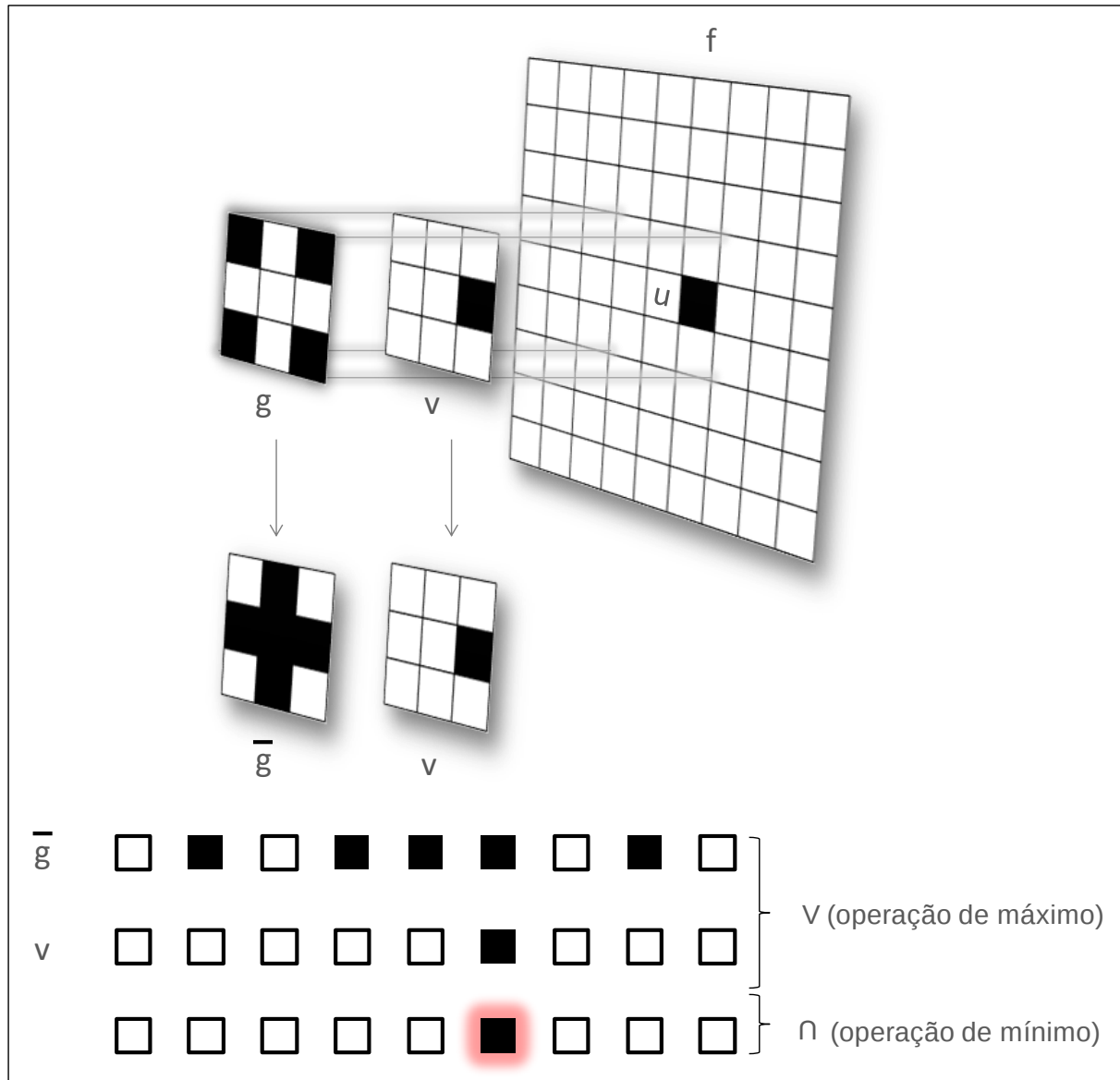
$$\varepsilon_g(f)(u) = \bigcap_{v \in S} f(v) \vee \bar{g}(u - v) \quad (3-1)$$

$$\delta_g(f)(u) = \bigcup_{v \in S} f(v) \wedge g(u - v), \quad (3-2)$$

onde $f: S \rightarrow [0,1]$ e $g: S \rightarrow [0,1]$ são imagens normalizadas em forma de matriz nomeada de formato S , onde $S \in \mathbb{N}^2$. O pixel é definido através do par cartesiano $(u, f(u))$, onde u é a posição associada ao valor $f(u)$. v é a matriz de $f(u)$, abrangida por g . Os operadores $\cup$ e $\vee$ estão associados à operação de máximo, enquanto $\cap$ e $\wedge$ estão associados à operação de mínimo. (g) é o elemento estruturante tanto para erosão quanto para dilatação (SOILLE 2004), (HÄNNI et al., 1999), (THANH et al., 2007) e (SANTOS et al., 2012). $(\bar{g})$ é a negação de (g) .

A Figura 3-1exibe a atuação da operação morfológica de erosão, descrita na Equação (3-1). Por precedência matemática, inicialmente ocorre a negação do elemento estruturante $\bar{g}$. Logo, acontece a operação de máximo $\vee$ denotada por $f(v) \vee \bar{g}(u - v)$, onde $f(v)$ diz respeito à matriz da imagem original f abrangida (casada) por $\bar{g}$. Por fim, o valor $\varepsilon_g(f)(u)$, na posição u , da imagem erodida recebe o mínimo valor entre os máximos, através do operador $\cap$. No exemplo exibido na Figura 3-1, $\varepsilon_g(f)(u)$, obtém o valor 0 associado ao preto absoluto.

Figura 3-1Atuação da operação morfológica de erosão.



A erosão sobrepõe ($\bar{g}$) à imagem original (f). O objetivo é que as áreas similares ao ($\bar{g}$) se expandam (SANTOS et al., 2012). Ao se associar o 1's ao branco absoluto e 0's ao preto absoluto, a erosão aumenta as áreas mais escuras e elimina as regiões com maior intensidade (SANTOS et al., 2012).

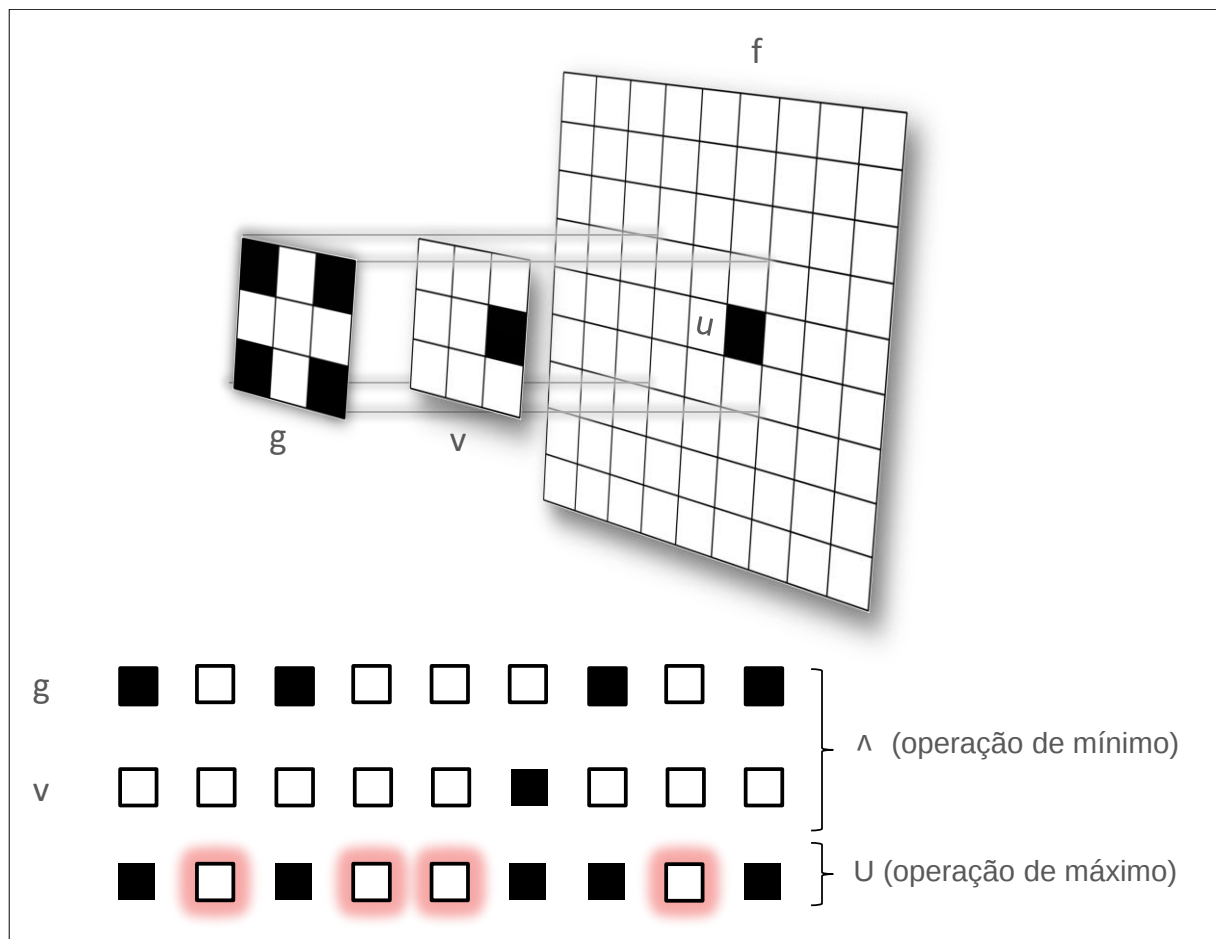
A Figura 3-2 exibe a atuação da operação morfológica de dilatação, descrita na Equação (3-2). Por precedência matemática, ocorre a operação de mínimo $\wedge$ denotada por $f(v) \wedge g(u - v)$, onde $f(v)$ diz respeito à matriz da imagem original f abrangida (casada) por g . Logo, o valor $\delta_g(f)(u)$, na posição u , da imagem dilatada recebe o máximo valor entre os mínimos, através

do operador U . No exemplo exibido na Figura 3-2, $\delta_g(f)(u)$ obtém o valor 1 associado ao branco absoluto.

A dilatação sobrepõe o elemento estruturante (g) à imagem original (f). O objetivo é que as áreas similares ao (g) se expandam (SANTOS et al., 2012). Ao se associar o 1's ao branco absoluto e 0's ao preto absoluto, a dilatação aumenta as áreas com tonalidade mais intensas e elimina as regiões escuras (SANTOS et al., 2012).

A execução de uma dilatação em uma imagem (f) seguida por uma erosão, com o mesmo elemento estruturante (g), é conhecida de fechamento $\phi_g(f)$. A aplicação de uma erosão seguida por uma dilatação, também com um mesmo elemento estruturante, é chamada de abertura $\gamma_g(f)$ (SANTOS et al., 2012).

Figura 3-2 Atuação da operação morfológica de dilatação.



Ao analisar as atuações das operações morfológicas de erosão e dilatação, exibidas nas Figura 3-1e aFigura 3-2respectivamente, nota-se que se trata de operações não-lineares. A

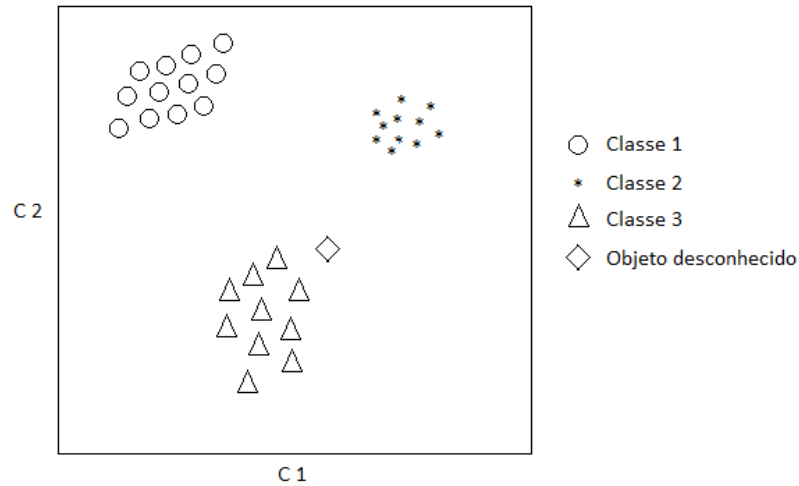
razão é que os operadores de mínimos $\cap$ e máximos $\cup$, formalizados matematicamente na Equação (3-1) e na Equação (3-2) respectivamente, apenas ocorrem após a conclusão das operações de $\vee$ e $\wedge$, respeitando a precedência matemática. Logo, as operações morfológicas são realizadas em duas etapas e não obedecem qualquer Equação linear.

3.2 Reconhecimento de Padrões

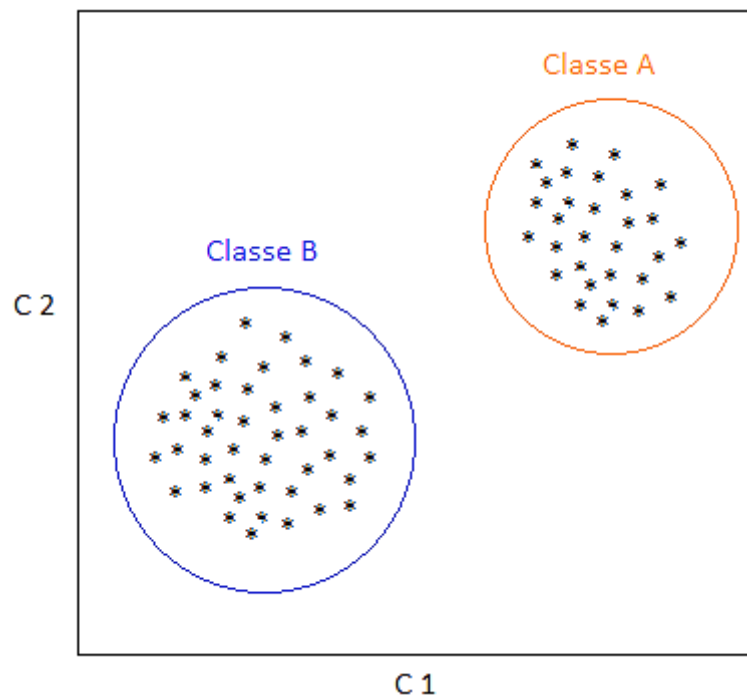
O reconhecimento de padrões é uma área da ciência que visa à classificação de objetos em um número de categorias ou classes com base na observação de suas características (Theodoridis et al., 2003). Os objetos a serem classificados variam de acordo com cada aplicação, podendo ser imagens, sinais em forma de ondas (voz, luz ou rádio) ou qualquer outro tipo de medida que tenha a necessidade de ser classificada.

Um problema de reconhecimento de padrões forma-se de uma tarefa de categorização ou classificação, em que as classes são estabelecidas pelo projetista do sistema (denominada classificação supervisionada) ou são "aprendidas" em conformidade com a similaridade dos padrões (classificação não-supervisionada).

Na Figura 3-3, é apresentado um conjunto de objetos demonstrados em um espaço de características bidimensionais. Como estão explícitos na legenda, estes objetos são conhecidos e foram antecipadamente rotulados (classe 1, classe 2 e classe 3). Dessa maneira, a partir desta base de conhecimento, pode ser atribuído um rótulo a um objeto desconhecido através de um processo de classificação supervisionado. No exemplo mostrado na Figura 3-3, o objeto desconhecido é rotulado como pertinente à classe 3. O parâmetro utilizado para essa decisão é a proximidade, no espaço de características, aos objetos desta classe.

Figura 3-3 Classificação supervisionada.

A Figura 3-4 exibe um exemplo de classificação não-supervisionada. Neste tipo de classificação, um outro conjunto de objetos está demonstrado no espaço de características bidimensional, porém, agora os padrões dos objetos não são conhecidos. Um processo de classificação não-supervisionada busca por objetos que tenham similaridades e os agrupam em classes. No exemplo demonstrado na Figura 3-4, os objetos foram divididos entre duas classes diferentes (classe A e classe B).

Figura 3-4 Classificação não-supervisionada.

- **Classificador**

A função de um classificador consiste em realizar o cálculo das medidas de similaridades existentes entre os objetos no espaço de características. Durante a classificação supervisionada, um objeto considerado desconhecido, é atribuído a uma determinada classe com base na similaridade deste objeto com os outros objetos representantes da classe no conjunto de treinamento. Existe uma grande variedade de classificadores e a escolha ou projeto do mais adequado depende especificamente de cada caso. Em geral, os classificadores podem ser divididos em dois tipos: Estatísticos (paramétrico e não-paramétricos) e os conexionistas (redes neurais).

Em se tratando dos classificadores estatísticos, os objetos contidos no conjunto de treinamento são utilizados para estimar certos parâmetros estatísticos de cada classe. Além de que, os classificadores estatísticos paramétricos necessitam estimar uma determinada distribuição para cada classe.

Os classificadores mais simples e intuitivos são os classificadores estatísticos não-paramétricos. Eles empregam uma função de distância como medida de similaridade, dando, a um objeto desconhecido, a classe que menos se distânci dele no espaço de características. Dentre os classificadores não-paramétricos mais utilizados, podemos citar o de distância euclidiana e o de distância de Mahalanobis.

Os denominados classificadores conexionistas são as redes neurais artificiais. As redes neurais compõem os classificadores mais complexos, e são inspirados na estrutura neural de organismos inteligentes, que aprendem por experiência.

A tarefa de escolher qual o melhor classificador utilizar depende especificamente do problema a ser tratado. No caso dos classificadores estatísticos, eles não mais indicados quando os objetos existentes nas várias classes estão distribuídos no espaço de características de acordo com topologias simples, e de preferencia conhecidas. Em contrapartida, os clasificadores conexionistas apresentam capacidade de se ajustar a qualquer topologia de classe. Dessa maneira, as redes neurais são mais indicadas para tarefas onde existem classes com distribuições mal-comportadas (GONZALEZ et al., 2007).

3.3 Redes Neurais Morfológicas

A. Introdução ao modelo neuronal morfológico

A definição de redes neurais morfológicas nasceu à medida que a álgebra de imagens ia se desenvolvendo. Demonstrou-se que uma sub-álgebra da álgebra de imagens (RITTER et al., 1990) inclui as formulações matemáticas de modelos de redes neurais clássicas. A partir de então alguns artigos envolvendo redes neurais morfológicas surgiram. No trabalho de (DAVIDSON et al., 1990) foram implementadas redes neurais morfológicas tendo como objetivo solucionar o problema de identificação de template e problemas de classificação de alvo. Em (ARAUJO et al., 1992) C.P. Suarez Araujo usou redes neurais morfológicas para calcular invariâncias visuais. Nos artigos (RITTER et al., 1996) (RITTER et al., 1997) Ritter e Sussner proporcionaram um melhor entendimento e uma base mais rigorosa para computação com redes neurais morfológicas. Dentre os distintos modelos morfológicos que vêm sendo desenvolvidos, podemos citar o perceptron morfológico (RITTER et al., 1997) (SUSSNER et al., 1998), o perceptron morfológico dentríco (RITTER et al., 2003), as memórias associativas morfológicas (RITTER et al., 1998) (SUSSNER 2000) (SUSSNER 2003A) (SUSSNER 2003B) (SUSSNER et al., 2006A) (SUSSNER et al., 2006B), redes neurais morfológicas modulares (ARAUJO et al., 2006A) (ARAUJO et al., 2006B) (DE SOUSA et al., 2000), redes neurais morfológicas de pesos compartilhados (KHABOU et al., 2000) (KHABOU et al., 2000B), redes neurais morfológicas de regularização (GADER et al., 2000), redes neurais morfológica /posto/lineares híbridas (PESSOA et al., 2000). As redes neurais morfológicas e morfológicas/lineares híbridas dispõem de muitas aplicações voltadas as áreas de reconhecimento de padrões (KABURLASOS et al., 2000) (KHABOU et al., 2000B) (PESSOA et al., 2000), processamento de imagens e visão computacional (GRAÑA et al., 2003) (RADUCANU et al., 2003), controle e previsão (ARAUJO et al., 2006B) (BEALE et al., 1997) (VALLE et al., 2004).

As redes neurais morfológicas compõem uma classe de RNAs que realiza uma operação elementar da morfologia matemática (erosão, anti-erosão, dilatação ou anti-dilatação), em um neurônio ou em uma camada, de um reticulado completo $\mathbb{L}$ para outro reticulado completo $\mathbb{M}$ seguida da aplicação de uma função de ativação. Em (RITTER et al., 1990) J.L. Davidson e G.X. Ritter propuseram o primeiro modelo neural morfológico.

Um ponto central da morfologia matemática é a decomposição de transformações entre reticulados completos em termos das operações elementares. No trabalho de Banon e Barrera foi proposto um teorema mostrando, que qualquer transformação $\psi : \mathbb{L} \rightarrow \mathbb{M}$ pode ser decomposta como o supremo de ínfimos entre erosões e anti-dilatações (BANON et al., 1993). Observe que o operador ínfimo pode ser apresentado como uma erosão, dado que este operador comuta com o ínfimo de uma coleção de elementos de um retículo $\mathbb{L}$, do mesmo modo o operador supremo pode ser apresentado como uma dilatação.

Colocaremos aqui uma definição importante, a respeito da explicação de um neurônio artificial generalizado. Um neurônio artificial generalizado pode ser definido como aquele que possui uma função de agregação que combina as entradas e os pesos sinápticos e posteriormente aplica uma função de ativação. A Figura 3-3 mostra um neurônio artificial generalizado. O neurônio morfológico é um caso especial do neurônio artificial generalizado.

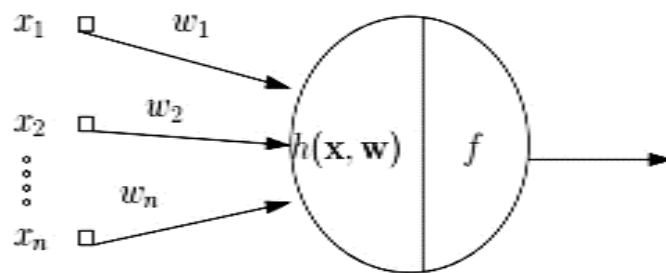


Figura 3-5 Neurônio artificial generalizado.

Trabalhos recentes na área de neurociência revelaram que um neurônio pode de fato executar uma operação de máximo (ANGELA et al., 2002). Vamos definir uma unidade de processamento como sendo um neurônio artificial generalizado. Nas redes neurais alimentadas a diante a equação $x_k^{(l)} = h(y^{l-1}, w_k^{(l)})$ $k = 1, 2, \dots, N_l$ diz simplesmente que o k -ésimo neurônio da camada l calcula uma operação $h = (y^{l-1}, w_k)$ que depende do vetor de entradas e do vetor de pesos associados ao k -ésimo neurônio. Se este nodo for um neurônio morfológico, esta operação deve ser uma erosão, anti-erosão ou uma dilatação, anti-dilatação no sentido da teoria de reticulados. Se o k -ésimo neurônio de uma camada l executa uma erosão, este neurônio é denominado como um neurônio erosivo. Igualmente, definimos um neurônio dilatativo, anti-dilatativo e anti-erosivo. Quando uma camada é formada apenas por neurônios erosivos, dizemos que essa camada realiza uma erosão e, por tanto, ela é denominada como

uma camada erosiva. O mesmo ocorre ao definirmos uma camada dilatativa, anti-dilatativa e anti-erosiva.

B. Perceptron Morfológico (MP)

O perceptron (MINSKY et al., 1969)(ROSENBLATT 1958)(SUSSNER 1998) é um modelo de rede neural *feed forward* com unidades de decisão binária que representam o comportamento das células nervosas do cérebro humano. Seu objetivo é agir como um classificador ou, simplesmente, reconhecer padrões. O modelo do perceptron convencional é naturalmente descrito pelo par de equações $h_k(x, w) = \sum_{j=1}^n w_{kj} x_j$ e $y_k = g(h_k(x, w))$, com a função limiar g , definida pela equação $g(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$, como função de ativação. Dado um vetor de entradas $x \in \mathbb{G}^n$ e um vetor de pesos sinápticos $w \in \mathbb{G}^n$, definimos o perceptron morfológico como uma RNM *feed forward*, tal que em cada neurônio calcula-se uma das quatro operações seguintes:

$$\varepsilon_w(x) = w^\top \wedge x, \quad (3-3)$$

$$\delta_w(x) = w^\top \vee x, \quad (3-4)$$

$$\bar{\varepsilon}_w(x) = \delta_w \circ v_*(x), \quad (3-5)$$

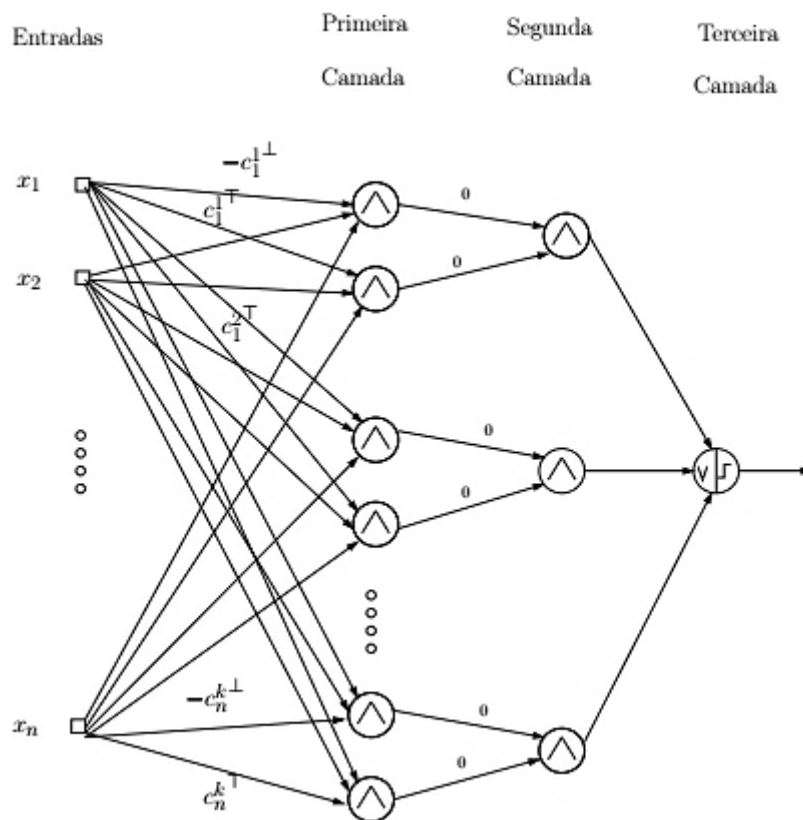
$$\bar{\delta}_w(x) = \varepsilon_w \circ v_*(x), \quad (3-6)$$

Em seguida aplica-se uma certa função de ativação g . Sobre a notação utilizada, temos que x_j denota o valor da j -ésima entrada, w_j denota o peso sináptico associado à j -ésima entrada, e f é uma certa função de ativação. Se g corresponder à função limiar, $g : \mathbb{G} \rightarrow \{0, 1\}$ é definida de acordo com a equação $g(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$.

A equação $y = g(\varepsilon_w(x)) = 0$, por exemplo, determina uma superfície, de dimensão $(n - 1)$ que divide o espaço euclidiano n -dimensional em duas regiões. Cada região é denominada como fronteira de decisão ou superfície de decisão e cada uma representa uma classe de padrões associados. Antes do perceptron atuar como um classificador, os valores dos seus pesos sinápticos devem ser determinados. Em situações nas quais a superfície de decisão é conhecida a priori, os pesos podem ser obtidos analiticamente, caso isto não ocorra, os pesos são obtidos em um estágio de treinamento, no qual o perceptron determinará sua própria

superfície de decisão, através de um algoritmo de aprendizagem. A Figura 3-6 representa a arquitetura de um perceptron morfológico.

Figura 3-6 Arquitetura do perceptron morfológico



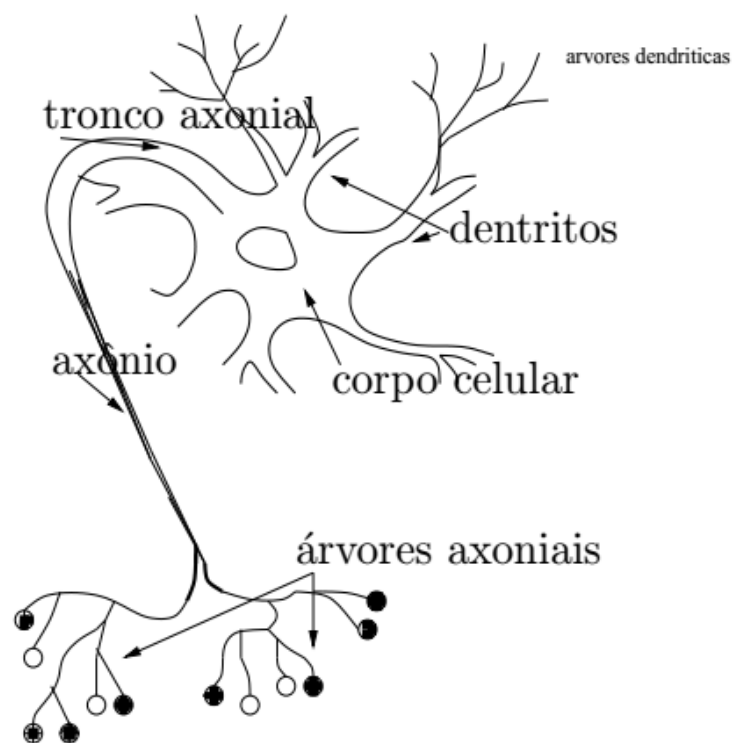
C. Perceptrons Morfológicos com Dendritos (MPD)

- **Motivação Biológica**

Estudos recentes na área de neurocomputação têm dado ênfase à importância de estruturas dendríticas em uma célula nervosa. Estas unidades são agora vistas como unidades computacionais primárias básicas do neurônio, que são capazes de realizar operações lógicas. Sobre esses novos modelos neurais, Ritter e Urcid desenvolveram um novo paradigma para a computação envolvendo neurônios morfológicos, onde o processo é realizado no dendrito (RITTER et al., 2003), produzindo assim um modelo mais fiel ao de um neurônio em redes neurais artificiais. Unidades computacionais e a topologia do modelo imitam neurônios biológicos do córtex cerebral. O neurônio, que é a unidade estrutural do sistema nervoso, é formado por um corpo celular, denominado soma, e alguns processos. Os processos são de dois

tipos, chamados dendritos e axônios, na devida ordem. Os dendritos transmitem impulsos de entrada para o corpo da célula. Via de regra, os dendritos apresentam ramificações que criam árvores longas e complicadas. O axônio que normalmente fica do lado contrário da célula em um ponto denominado tronco axonal, constitui-se de uma fibra longa cujos ramos formam a árvore axonal. Ocasionalmente, em alguns neurônios, o axônio pode ter ramos em intervalos ao longo do seu comprimento, além de sua terminação arborizada. O axônio é considerado o principal ramo fibral do neurônio com a finalidade de transmissão de sinais a outros neurônios. A Figura 3-7 descreve uma célula nervosa.

Figura 3-7 Diagrama de uma célula nervosa



Um impulso desloca-se pelo axônio, ao longo do tronco axonal, do neurônio para terminações nervosas. Essas terminações fazem contato com o corpo da célula e muitos dendritos de outras células nervosas. As regiões de contato entre axônio e dendrito são chamadas de regiões sinápticas porque é o local onde as sinapses ocorrem. A sinapse é apenas uma estrutura especializada com a tarefa de realizar a comunicação entre neurônios. Na prática o mecanismo exato de como a sinapse é realizada ainda é considerado desconhecido, embora o que é bem conhecido é que existem dois tipos de sinapses: as sinapses excitantes, as quais

tendem a despolarizar a membrana pós-sináptica facilitando a ativação do neurônio, e as sinápses inibitórias que tem como objetivo tentar prevenir a ativação final do neurônio.

A maioria das sinápses ocorrem nos dendritos, sendo de responsabilidade deles ocupar uma parcela significativa de área e volume do cérebro. Estudos recentes mostraram que a informação transmitida é processada nos dendritos. Neste sentido, parece bem razoável considerar a atuação dos dendritos na neurocomputação.

• Descrição detalhada do modelo

No perceptron morfológico com dendritos, consideramos que as informações serão processadas nos dendritos, ou seja, um neurônio morfológico N possui regiões pós-sinápticas que recebem entrada de terminações axoniais de outros neurônios. Se N recebe entradas de n neurônios $N_1, N_2, \dots, N_n$, então cada N_i possui ramos axoniais que entram em várias regiões sinápticas de N . Estas regiões são distribuídas ao longo de um número finito de dendritos $D_1, D_2, \dots, D_k$. Levando em consideração um neurônio N que possui K dendritos supondo que cada dendrito recebe o vetor de entrada $x = (x_1, x_2, \dots, x_n)^\top$. O peso sináptico correspondente à entrada x_i do dendrito k é indicado por w_{ki}^l , onde $l \in \{0, 1\}$, e o índice l simplesmente diferencia as sinápses excitantes das sinápses inibidoras. Ou seja, quando denotamos w_{ki}^1 , quer dizer que a entrada x_i produz uma excitação e w_{ki}^0 indica uma inibição induzida pela entrada x_i .

A entrada x_i produz um efeito final no k -ésimo dendrito dado por $+(w_{ki}^1 + x_i)$ para as sinápses excitantes e $-(w_{ki}^0 + x_i)$ para as sinápses inibidoras. Observamos que, quanto maior o valor de x_i , maior será o efeito de uma sinapse excitante e menor será o efeito de uma sinapse inibidora. Neste modelo, em uma entrada é permitido a produção de excitação e inibição no dendrito. O que ocorre por fim é que uma competição entre o efeito de todas as conexões sinápticas é realizada e apenas a sinapse de menor valor é considerada como ativação do dendrito em questão.

A computação realizada pelo k -ésimo dendrito para uma entrada $x \in \mathbb{R}^n$ é fornecida pela seguinte expressão:

$$\tau_k(x) = p_k \bigwedge_{i=1}^n \bigwedge_{l \in L} (-1)^{l+1} (x_i + w_{ki}^l), \quad (3-7)$$

onde $L = \{0, 1\}$ e $p_k = \{-1, 1\}$ denota uma resposta excitante ou inibidora do k -ésimo dendrito. Evidenciamos que $w_{ki}^1 = \infty$ e $w_{ki}^0 = -\infty$ respectivamente, representam ausência de conexão sináptica excitante ou inibitória.

O valor $\tau_k(x)$ é emitido para o corpo celular do neurônio e o estado neurônio N é uma função das entradas recebidas de todos os dendritos:

$$\tau_k(x) = \bigwedge_{k=1}^K \tau_k(x), \quad (3-8)$$

na qual K indica o número de dendritos. Quer dizer, a ativação do neurônio é considerada o menor valor dentre as excitações de todos os dendritos. O próximo estado do neurônio N , isto é, a saída, é determinada por uma função de ativação aplicada à ativação do neurônio N , que adotaremos como sendo a função limiar definida na equação $g(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$.

Por conseguinte, o neurônio será ativado se, e somente se, todos os dendritos forem saídas positivas, isto é, se $\tau_k(x) \geq 0$ para todo $k = 1, \dots, K$. A saída de um neurônio morfológico com dendritos é dada por:

$$y(x) = f \left(\bigwedge_{k=1}^K [p_k \bigwedge_{i=1}^n \bigwedge_{l \in L} (-1)^{l+1} (w_{ik}^l + x_i)] \right), \quad (3-9)$$

Vejamos como contar as unidades computacionais de um MPD. Segundo a equação $\tau_k(x) = p_k \bigwedge_{i=1}^n \bigwedge_{l \in L} (-1)^{l+1} (x_i + w_{ki}^l)$, temos que um k -ésimo dendrito calcula a seguinte operação:

$$\tau_k(x) = p_k \bigwedge_{i=1}^n \bigwedge_{l \in \{0,1\}} (-1)^{l+1} (x_i + w_i^l), \quad (3-10)$$

$$= p_k \bigwedge_{i=1}^n \{(-1)^{0+1} (x_i + w_i^0) \wedge (-1)^{1+1} (x_i + w_i^1)\}, \quad (3-11)$$

$$= p_k \{ \{(-1) \cdot (x_1 + w_1^0) \wedge (x_1 + w_1^1)\} \wedge \dots \wedge \{(-1) \cdot (x_n + w_n^0) \wedge (x_n + w_n^1)\} \}, \quad (3-12)$$

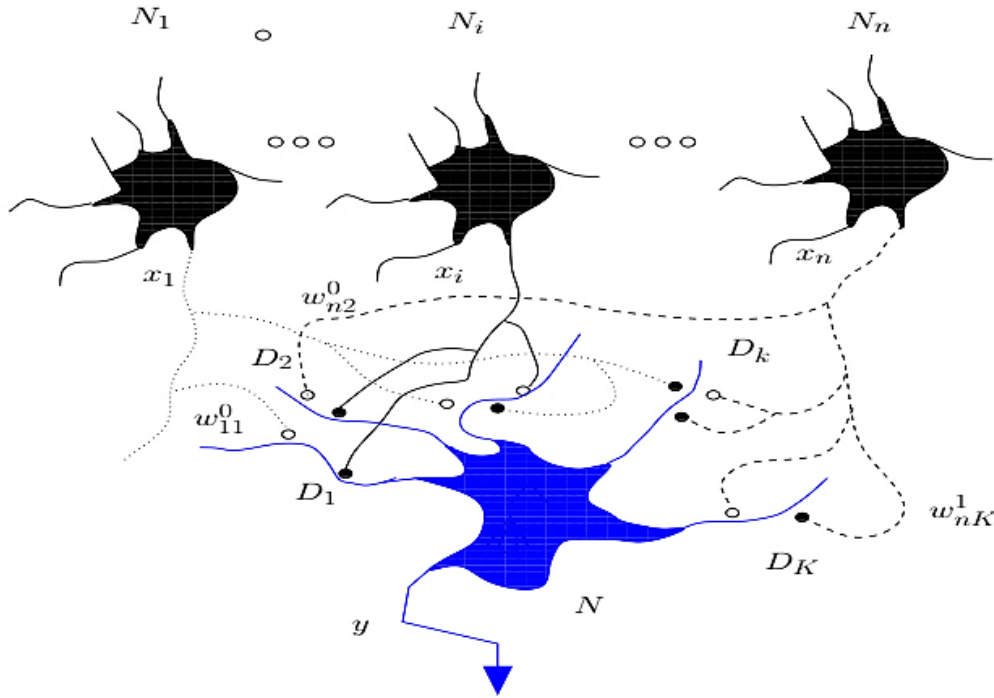
Como o reticulado completo é distributivo, rearranjamos os termos e obtemos:

$$\tau_k(x) = p_k \{ \bigwedge_{i=1}^n (-1)(x_i + w_i^0) \} \wedge \{ \bigwedge_{i=1}^n (x_i + w_i^1) \}, \quad (3-13)$$

Observamos que existem três unidades computacionais no k -ésimo dendrito; uma calcula a expressão $\bigwedge_{i=1}^n (-1)(x_i + w_i^0)$, outra calcula a expressão $\bigwedge_{i=1}^n (x_i + w_i^1)$ por fim a expressão $\{ \bigwedge_{i=1}^n (-1)(x_i + w_i^0) \} \wedge \{ \bigwedge_{i=1}^n (x_i + w_i^1) \}$.

Pela equação: $y(x) = f\left(\bigwedge_{k=1}^K [p_k \bigwedge_{i=1}^n \bigwedge_{l \in L} (-1)^{l+1} (w_{ik}^l + x_i)]\right)$ temos que, por fim, um neurônio realiza uma competição entre as ativações produzidas pelos K dendritos. Assim, teríamos um total de $3K + 1$ unidades computacionais. A Figura 3-8 representa o modelo de um perceptron morfológico com dendritos.

Figura 3-8 Modelo do perceptron morfológico com dendritos.



D. Fuzzy Lattice Neural Network-FLNN

Os fundamentos teóricos da rede neural FLNN são constituídos de uma combinação bem sucedida de conjuntos *fuzzy* (KLIR et al., 1995) e teoria de reticulados (BIRKHOFF et al., 1993) somado a teoria de ressonância adaptativa (ART) (CARPENTER et al., 1987). A rede FLNN implementa dois paradigmas de aprendizagem, sendo um esquema de clusterização e outro para classificação. No processo, conjuntos de elementos de reticulados são especificados por conjuntos de intervalos de retículos.

- **Reticulado dos intervalos generalizados**

Seja $\mathbb{L}$ um reticulado completo e $O_{\mathbb{L}}$ e $I_{\mathbb{L}}$, o menor elemento de $\mathbb{L}$, nesta ordem. O reticulado completo $\mathbb{P}\mathbb{L}$, denominado de reticulado dos intervalos generalizados de $\mathbb{L}$, é dado por

$$\mathbb{PL} = \{[a, b] : a, b \in \mathbb{L}\}, \quad (3-14)$$

Com a relação de ordem parcial dada por $[a, b] \leq [c, d] \Leftrightarrow c \leq a \text{ e } b \leq d$. Seja x^i uma coleção de elementos de $\mathbb{PL}$. Então o ínfimo e o supremo de elementos deste reticulado $\mathbb{PL}$ são dados naturalmente, na devida ordem, por:

$$\bigwedge_{i \in I} x^i = \bigwedge_{i \in I} [a^i, b^i] = [\bigvee_{i \in I} a^i, \bigwedge_{i \in I} b^i], \quad (3-15)$$

$$\bigvee_{i \in I} x^i = \bigvee_{i \in I} [a^i, b^i] = [\bigwedge_{i \in I} a^i, \bigvee_{i \in I} b^i], \quad (3-16)$$

Um elemento de $\mathbb{PL}$ é chamado de intervalo generalizado. O reticulado completo $\mathbb{V}_{\mathbb{L}}$ foi definido como um conjunto de todos os intervalos fechados de elementos $\mathbb{L}$ aumentado pelo elemento mínimo $[I_{\mathbb{L}}, O_{\mathbb{L}}]$ de $\mathbb{PL}$:

$$\mathbb{V}_{\mathbb{L}} = \{[a, b] : a, b \in \mathbb{L} \text{ e } a \leq b\} \cup \{[I_{\mathbb{L}}, O_{\mathbb{L}}]\}, \quad (3-17)$$

Podemos definir as operações de ínfimo e supremo no reticulado completo $\mathbb{V}_{\mathbb{L}}^N = (\mathbb{V}_{\mathbb{L}})^N$ da seguinte maneira: Seja $x^i = (x_1^i, x_2^i, \dots, x_N^i)$ uma coleção de elementos de $\mathbb{V}_{\mathbb{L}}^N$. Então o ínfimo e o supremo de x^i em $\mathbb{V}_{\mathbb{L}}^N$ são dados, respectivamente, por:

$$\bigwedge_{i \in I} x^i = (\bigwedge_{i \in I} x_1^i, \bigwedge_{i \in I} x_2^i, \dots, \bigwedge_{i \in I} x_N^i), \quad (3-18)$$

$$\bigvee_{i \in I} x^i = (\bigvee_{i \in I} x_1^i, \bigvee_{i \in I} x_2^i, \dots, \bigvee_{i \in I} x_N^i), \quad (3-19)$$

• Conjuntos Fuzzy

Um conjunto *fuzzy* **a** é determinado por meio de uma função, denominada função de pertinência, de um conjunto clássico X para o intervalo $[0, 1]$:

$$a(x) : X \rightarrow [0, 1], \quad (3-20)$$

Na verdade esta função caracteriza o grau de pertinência de **x** em **a**. Todo conjunto *fuzzy* fica unicamente determinado pela sua função de pertinência **a(x)**. A classe dos conjuntos *fuzzy* é dada por:

$$\mathcal{F}(X) = \{a : X \rightarrow [0, 1]\}, \quad (3-21)$$

Um conjunto clássico $A \in \mathcal{P}(X)$ se relaciona com um conjunto nebuloso $a \in \mathcal{F}(X)$ da forma a seguir:

$$a(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \notin A \end{cases} \quad (3-22)$$

• Reticulados *Fuzzy*

O conceito de reticulado *fuzzy* foi introduzido para estender a relação de ordem parcial a todos pares $(x, y) \in \mathbb{L} \times \mathbb{L}$ de um reticulado convencional $\mathbb{L}$. Um reticulado *fuzzy* é um par $(\mathbb{L}, p)$, onde $\mathbb{L}$ é um reticulado convencional e $p : X \rightarrow [0, 1]$ é uma função de ordem parcial nebulosa no universo de discurso $X = \{(x, y) : x, y \in \mathbb{L}\}$, tal que:

$$p(x, y) = 1 \Leftrightarrow x \leq y \text{ em } \mathbb{L}, \quad (3-23)$$

Uma função $h : \mathbb{L} \rightarrow \mathbb{R}$ é dita ser uma função de avaliação positiva se satisfaz as seguintes propriedades:

$$(P1) \quad h(0) = 0, \quad \text{onde } 0 \text{ é o menor elemento em } \mathbb{L}.$$

$$(P2) \quad u < w \Rightarrow h(u) < h(w), \quad u, w \in \mathbb{L}.$$

$$(P3) \quad u \leq w \Rightarrow h(x \vee w) - h(x \vee u) \leq h(w) - h(u), \quad x, u, w \in \mathbb{L}.$$

Uma função de avaliação positiva não necessariamente existe em um reticulado $\mathbb{L}$. Todavia, quando esta função existe pode-se mostrar que $p(x, y) = h(y)/(h \vee y)$, $x, y \in \mathbb{L}$, sendo que $h(\cdot)$ é uma função de avaliação positiva. Esta função determina uma ordem parcial nebulosa em $\mathbb{L}$, pois $p(x, y) = \frac{h(y)}{h \vee y} = 1 \Leftrightarrow h(y) = h(x \vee y) \Leftrightarrow y = x \vee y \Leftrightarrow x \leq y$ (PETRIDIS et al., 1998). Desse modo, $(\mathbb{L}, p(x, y))$ é um reticulado *fuzzy*. Assim sendo, a função p satisfaz as seguintes propriedades:

$$(C1) \quad p(u, 0) = 0, \quad u \neq 0;$$

$$(C2) \quad p(u, u) = 1, \quad \forall u \in \mathbb{L};$$

$$(C3) \quad u \leq w \Rightarrow p(x, u) \leq p(x, w), \quad x, u, w \in \mathbb{L}.$$

A função $p(x, y)$ será usada como função de agregação dos neurônios da rede neural FLNN. Na verdade, para calcularmos a função de ordem parcial nebulosa $p(x, u)$ de um intervalo de reticulados x em outro intervalo de reticulados u , primeiro mapeamos $x, u \in \mathbb{V}_{\mathbb{L}}^N \subseteq \mathbb{P}\mathbb{L}^N$ a seus isomorfos $x', u' \in \mathbb{L}^N$.

O lema a seguir garante a existência de uma função de avaliação positiva no produto reticulado *fuzzy*, uma vez que todos os reticulados *fuzzy* constituintes possuam uma função de avaliação positiva.

Lema 1: Sejam $\mathbb{L} = \mathbb{L}_1 \times \mathbb{L}_2 \times \dots \times \mathbb{L}_N$ o produto de N reticulados completos, cada qual com as suas respectivas funções de avaliação positiva $h_1, h_2, \dots, h_n$. Logo a função $h((x_1, \dots, x_N)) = h_1(x_1) + \dots + h_N(x_N)$ define uma função positiva ao produto de reticulados $\mathbb{L}$ (De Oliveira et al., 2010).

Uma função de avaliação positiva qualquer em $\mathbb{V}_{\mathbb{L}}$ define uma função de ordem parcial em $\mathbb{V}_{\mathbb{L}}$. Logo, para se obter uma função de ordem parcial nebulosa em $\mathbb{V}_{\mathbb{L}}$ é preciso primeiramente estabelecer um isomorfismo dual entre $\mathbb{P}\mathbb{L}$ e $\mathbb{L}^2$ sendo que para construir este isomorfismo é suficiente que exista um automorfismo dual $\theta: \mathbb{L}^2 \rightarrow \mathbb{L}^2$. Considere o isoformismo:

$$\varphi_{\theta}: \mathbb{P}\mathbb{L} \rightarrow \mathbb{L}^2, \quad (3-24)$$

$$[a, b] \mapsto (\theta(a), b), \quad (3-25)$$

Deste modo, existe um isomorfismo entre $\mathbb{P}\mathbb{L}$ e $\mathbb{L}^2$. Uma função de ordem parcial nebulosa pode ser definida no reticulado $\mathbb{V}_{\mathbb{L}}$ da seguinte forma:

1. Uma função de avaliação positiva h no reticulado $\mathbb{L}$ pode definir uma função de avaliação positiva no reticulado $\mathbb{L}^2$ de acordo com o **Lema 1**.
2. Uma função de avaliação positiva h em $\mathbb{L}^2$ define uma função de ordem parcial nebulosa $p_h(x, y) = \frac{h(y)}{h(x \vee y)}$ em $\mathbb{L}^2$.
3. Um automorfismo dual $\theta(\cdot)$ em $\mathbb{L}$ pode definir um isoformismo φ_{θ} entre $\mathbb{P}\mathbb{L}$ e $\mathbb{L}^2$.
4. Uma função de ordem parcial nebulosa entre dois intervalos $[a, b], [c, d] \in \mathbb{V}_{\mathbb{L}}$ é dada por:

$$p([a, b], [c, d]) = p_h(\varphi_{\theta}[a, b], \varphi_{\theta}[c, d]), \quad (3-26)$$

$$= p_h((\theta(a), b), (\theta(c), d)), \quad (3-27)$$

onde $(\theta(a), b), (\theta(c), d) \in \mathbb{L}^2$ e p_h é a função de ordem parcial nebulosa em $\mathbb{L}^2$ dada por:

$$p_h(x, y) = \frac{h(y)}{h(xvy)}, \quad (3-28)$$

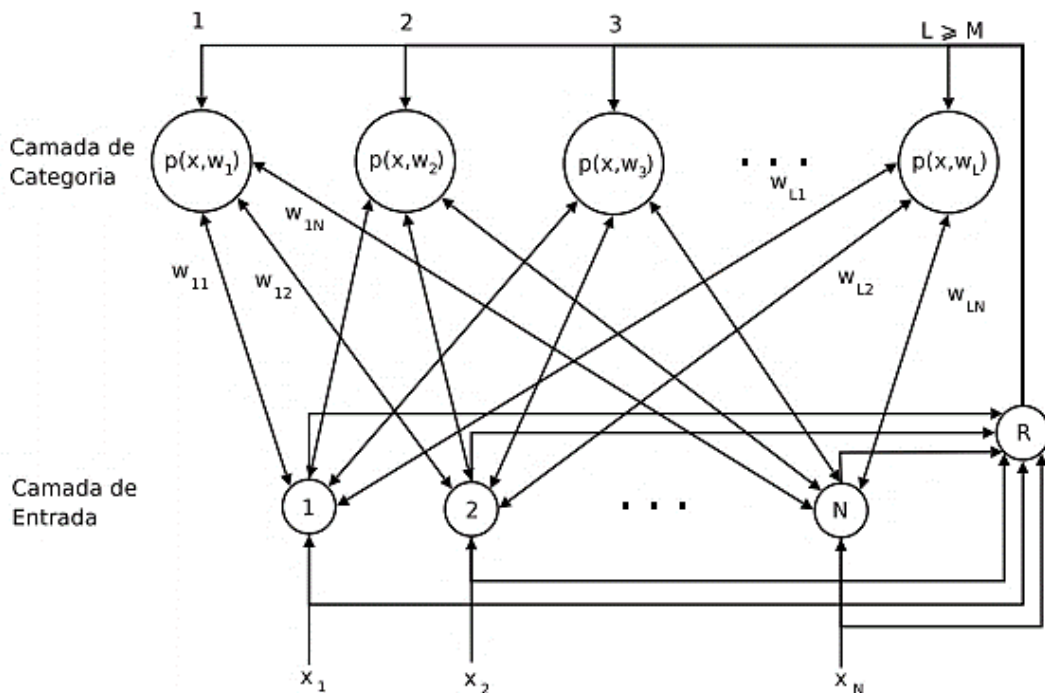
Então sempre podemos associar um elemento $x \in \mathbb{L}^2$ com um intervalo $[a, b]$ de $\mathbb{V}_{\mathbb{L}}$ pelo fato da existência deste automorfismo θ e do isomorfismo φ_θ entre $\mathbb{P}\mathbb{L}$ e $\mathbb{L}$. A função de ordem parcial *fuzzyp* : $\mathbb{V}_{\mathbb{L}} \rightarrow [0, 1]$ pode representar uma anti-erosão do reticulado completo $\mathbb{V}_{\mathbb{L}}$ para o reticulado completo $[0, 1]$ ou pode representar também uma anti-dilatação do reticulado completo $\mathbb{V}_{\mathbb{L}}$ para o reticulado completo $[0, 1]$.

Uma ordem parcial nebulosa p em $\mathbb{V}_{\mathbb{L}}^N$ é obtida por meio do isomorfismo $\varphi_\theta^N : (\mathbb{P}\mathbb{L})^N \rightarrow (\mathbb{L}^2)^N$. Ou seja, para mapear um elemento $a \in \mathbb{V}_{\mathbb{L}}^N$ para $(\mathbb{L}^2)^N$, basta aplicar $\varphi_\theta^N(a) = \varphi_\theta^N(a_1, \dots, a_N) = (\varphi_\theta(a_1), \dots, \varphi_\theta(a_N))$, onde $(a_1, \dots, a_N)$.

• Arquitetura da rede neural FLNN

A rede neural FLNN funciona de forma semelhante a rede neural ART (CARPENTER et al., 1987), quer dizer, a rede FLNN constitui-se de uma arquitetura recorrente em duas camadas: a camada de entrada e a camada de categoria, de acordo com a Figura 3-9.

Figura 3-9 Arquitetura da rede neural FLNN.



A camada de entrada dispõe de N neurônios artificiais que são utilizados para armazenar e comparar os dados de entrada. A camada de categoria contém L neurônios artificiais, que representam M classes. Cada nodo da camada de categoria pode representar no máximo uma classe, ou seja, $L \geq M$.

As duas camadas são totalmente conectadas pelo vetor de pesos sinápticos $\mathbf{w}$, que tem como objetivo filtrar os dados tanto de baixo para cima (pesos *bottom-up*), quanto de cima para baixo (pesos *top-down*). Também existe um nodo utilizado para desabilitar as classes que opera de modo semelhante à rede neural ART. A entrada da rede FLNN deve ser uma hipercaixa $\mathbf{x}$ pertencente a $(\mathbb{V}_{\mathbb{L}})^N$, ao invés de vetores de $\mathbb{R}^N$, isto é, cada componente x_j , para $j = \{1, \dots, N\}$, representa um valor $\mathbb{V}_{\mathbb{L}}$. O tamanho de uma hipercaixa arbitrária $\mathbf{x} = ([a_1, b_1], \dots, [a_N, b_N])$, denotado por $Z(\mathbf{x})$, é definido por $Z(\mathbf{x}) = \prod_{i=1}^N (h(b_i) - h(a_i))$, onde $h_1, \dots, h_N$ são funções de avaliação positiva definidas em $\mathbb{L}$. A função de agregação utilizada pelos neurônios da camada superior da rede FLNN, é uma função de ordem parcial *fuzzy*, denotada por p , que calcula o valor de pertinência da entrada $\mathbf{x}$ no intervalo de pesos aprendidos $\mathbf{w}_i$, os quais ligam os nodos da camada de entrada ao i -ésimo nodo da camada de categoria, onde $i = \{1, \dots, N\}$. A rede neural FLNN lida com certas famílias de intervalos de reticulados, denotadas por $\{\mathbf{w}_i\}$ onde $\mathbf{w}_i \in \mathbb{V}_{\mathbb{L}}^N$ e $i \in I$, I finito. A finalidade da rede FLNN é identificar conjuntos c de elementos de reticulados, chamados de classe que podem ser representados por uma união de um número finito de intervalos $(\mathbb{V}_{\mathbb{L}})^N$, isto é:

$$c = \cup_i \mathbf{w}_i. \quad (3-29)$$

A aplicação da rede FLNN depende fundamentalmente da existência de uma função de ordem parcial fuzzy em $(\mathbb{V}_{\mathbb{L}})^N$. Se $\mathbf{x} \in (\mathbb{V}_{\mathbb{L}})^N$ e $c = \cup_i \mathbf{w}_i$ então o grau de inclusão de $\mathbf{x}$ na classe c é definido como

$$p(\mathbf{x}, c) := \max(p(\mathbf{x}, \mathbf{w}_i)). \quad (3-30)$$

A tomada de decisão na rede neural FLNN é alcançada através do grau de verdade da proposição " $\mathbf{x} \leq c$ " como expressado por $p(\mathbf{x}, c)$, onde $\mathbf{x}$ é um um intervalo de reticulado de entrada que excita o sistema e c representa a classe aprendida guardada na memória do sistema. Quando $p(\mathbf{x}, c) = 1$ nós escrevemos $\mathbf{x} \leq c$ e diremos $\mathbf{x}$ está dentro da classe c .

E. Rede Neural Linear/Posto/Morfológica híbrida

A rede neural Linear/Posto/Morfológica Híbrida (MRL) (PESSOA et al., 2010) pertence à classe das redes neurais *feed forward*, e contém propriedades dos perceptrons de múltiplas camadas (MLPs) (SUSSNER 1999) e das redes neurais posto morfológicas (MRNNs). Ou seja, cada neurônio realiza uma operação que consiste em uma combinação híbrida de operações lineares e não-lineares seguida da aplicação de uma função de ativação. Do ponto de vista matemático, podemos descrever a rede MRL por meio das equações:

$$\mathbf{y}^l = G(\mathbf{x}^l) = (g(x_1^{(l)}), g(x_2^{(l)}), \dots, g(x_{N_l}^{(l)})), \quad (3-31)$$

$$l = 1, 2, \dots, L$$

$$\mathbf{x}_k^{(l)} = h(\mathbf{y}^{l-1}, \mathbf{w}_k^{(l)}), \quad (3-32)$$

$$k = 1, 2, \dots, N_l$$

tais que:

$$x_k^{(l)} = \lambda_k^{(l)} a_k^{(l)} + (1 - \lambda_k^{(l)}) \beta_k^{(l)}, \quad (3-33)$$

$$a_k^{(l)} = \mathcal{R}_{r_k^{(l)}}(\mathbf{y}^{(l-1)} + \mathbf{a}_k^{(l)}), \quad (3-34)$$

$$\beta_k^{(l)} = \mathbf{y}^{(l-1)\top} \mathbf{b}_k^{(l)} + \theta_k^{(l)}, \quad (3-35)$$

onde $r_k^{(l)} \in \mathbb{Z}$, $\lambda_k^{(l)}, \theta_k^{(l)} \in \mathbb{R}$; $\mathbf{a}_k^{(l)}, \mathbf{b}_k^{(l)} \in \mathbb{R}^{N_{l-1}}$. O símbolo $\mathcal{R}_r$ representa a função posto (PESSOA et al., 1998), de ordem r , que determina o r -ésimo elemento da sequência obtida colocando-se os componentes do vetor $\mathbf{t} = (t_1, \dots, t_n) \in \mathbb{R}^n$ dado em ordem decrescente, ou seja, considere a sequência em ordem decrescente:

$$t_{(1)} \geq t_{(2)} \geq \dots \geq t_{(n)},$$

Tomando o r -ésimo elemento desta sequência, definimos a r -ésima função posto de $\mathbf{t}$ por

$$\mathcal{R}_r(\mathbf{t}) = t_{(r)}, \text{ onde } r = 1, 2, \dots, k. \quad (3-36)$$

O vetor peso associado a cada nodo é definido por

$$\mathbf{w}_k^{(l)} = (\mathbf{a}_k^{(l)}, \rho_k^{(l)}, \mathbf{b}_k^{(l)}, \theta_k^{(l)}, \lambda_k^{(l)}), \quad (3-37)$$

em

$$\mathbb{R}^{2N_{l-1}+3}. \quad (3-38)$$

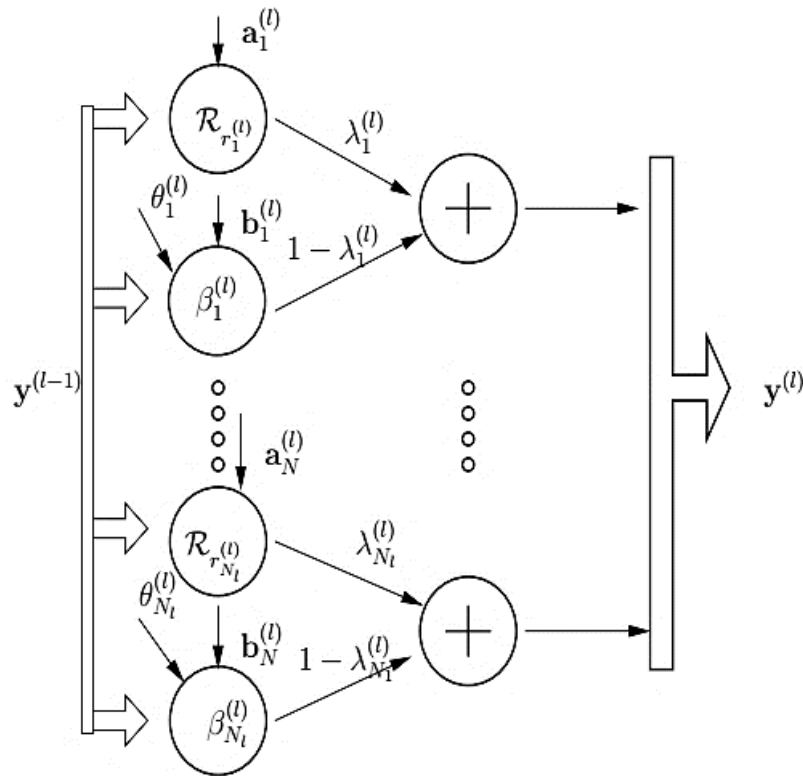
Usa-se uma variável real $\rho_k^{(l)}$ ao invés de uma variável inteira $r_k^{(l)}$, devido ao fato de se precisar avaliar derivadas posto (PESSOA et al., 1998) durante o algoritmo de treinamento da rede MRL.

A relação entre $\rho_k^{(l)}$ e $r_k^{(l)}$ é dada por:

$$r_k^{(l)} = \left\lfloor N_{l-1} - \frac{N_{l-1}-1}{1+\exp(\rho_k^{(l)})} + 0.5 \right\rfloor, \quad (3-39)$$

onde $\lfloor \cdot \rfloor$ indica a função maior inteiro. A Figura 3-10 representa um módulo da rede MRL

Figura 3-10 Estrutura da l-ésima camada de uma rede MRL.



Observamos que o cálculo realizado pela equação $x_k^{(l)} = \lambda_k^{(l)} \alpha_k^{(l)} + (1 - \lambda_k^{(l)}) \beta_k^{(l)}$ não consegue ser realizado por um único neurônio. Isto ocorre devido ao fato da variável $r_k^{(l)}$ sempre variar. Neste seguimento em um neurônio é responsável pela soma vetorial $\mathbf{y}^{(l-1)} + \mathbf{a}_k^{(l)}$, e outro responsável por tomar a $r_k^{(l)}$ posição do vetor resultante da soma. Assim, em um módulo da

rede MRL, representado pela Figura 3-10, temos que existem quatro unidades computacionais ao invés de três (impressão equivocada que a Figura 3-8 pode passar). Dois casos especiais da rede MRL são obtidos quando a função de ativação g , empregada pela rede MRL de acordo com as equações $\mathbf{y}^l = G(\mathbf{x}^l) = (g(\mathbf{x}_1^{(l)}), g(\mathbf{x}_2^{(l)}), \dots, g(\mathbf{x}_{N_l}^{(l)}))$, $l = 1, 2, \dots, L$ e $\mathbf{x}_k^{(l)} = h(\mathbf{y}^{l-1}, \mathbf{w}_k^{(l)})$ $k = 1, 2, \dots, N_l$, é identidade, isto é, $g(\mathbf{x}_1^{(l)}) = \mathbf{x}_k^{(l)}$, e quando a função g é uma não linearidade do tipo logística, isto é, $g(\mathbf{x}_1^{(l)}) = (1 + \exp[-\eta \mathbf{x}_k^{(l)}])^{-1}$, $\eta \geq 1$, de acordo com o trabalho de (PESSOA et al., 2000). Previamente, antes de definirmos a derivada da função posto, iremos definir a função impulso. A função impulso é definida por:

$$Q(\mathbf{v}) = [q(v_1), \dots, q(v_1)], \quad (3-40)$$

tal que $\mathbf{v} = (v_1, \dots, v_n)$,

Onde

$$q(v) = \begin{cases} 1, & \text{se } v = 0 \\ 0, & \text{caso contrário} \end{cases} \quad (3-41)$$

Definição 1: Dado um vetor $\mathbf{t} = (t_1, t_2, \dots, t_n) \in \mathbb{R}^n$, e um posto $r \in \{1, 2, \dots, n\}$, o r -ésimo vetor indicador posto $\mathbf{c}$ de $\mathbf{t}$ é definido por:

$$\mathbf{c}(\mathbf{t}, r) = \frac{Q(z\mathbf{1} - \mathbf{t})}{Q(z\mathbf{1} - \mathbf{t}) \cdot \mathbf{1}^\top}, \quad (3-42)$$

onde $z = \mathcal{R}_r(\mathbf{t})$ e $\mathbf{1} = (1, 1, \dots, 1)$.

O vetor indicador posto tem a função de marcar as localizações em $\mathbf{t}$ onde $z = \mathcal{R}_r(\mathbf{t})$ ocorrem. Para evitar mudanças abruptas e conseguir robustez numérica, a função $q(v)$ é, normalmente (PESSOA et al., 1998) $q_\sigma(v)$, $\sigma \geq 0$ como, por exemplo, $\exp[-\frac{1}{2}(v/\sigma)^2]$ ou $\text{sech}^2(v/\sigma)$.

Definição 2: Seja $\mathbf{t} \in \mathbb{R}^n$, $r \in \{1, 2, \dots, n\}$ e $\mathbf{c} = \mathbf{c}(\mathbf{t}, r)$. Então,

- $\mathbf{c} \cdot \mathbf{1}^\top = 1$
- $\mathbf{c} \cdot \mathbf{t}^\top = \mathcal{R}_r(\mathbf{t})$
- Se r é fixo, então é uma função constante por partes de $\mathbf{t}$ com exatamente $2^n - 1$ possíveis diferentes valores. Para todos os pontos $\mathbf{t}_0 \in \mathbb{R}^n$ com componentes diferentes,

existe uma vizinhança em torno deles, $\|\mathbf{t} - \mathbf{t}_0\|_\infty < \frac{1}{2} \min_{i \neq j} |t_{0_i} - t_{0_j}|$, tal que dentro dela $\mathbf{c}$ é constante, isto é, $\mathbf{c}(\mathbf{t}, r) = \mathbf{c}(\mathbf{t}_0, r)$.

Segundo a definição e a proposição (PESSOA et al., 1998) apresentadas a cima, se considerarmos que, para r fixo, $\mathbf{c} = \mathbf{c}(\mathbf{t}, r)$ seja constante em uma certa vizinhança de $\mathbf{t}_0$ então a função posto $\mathcal{R}_r(\mathbf{t})$ é diferenciável em $\mathbf{t}_0$ e sua derivada é dada por:

$$\nabla \mathcal{R}_r(\mathbf{t})|_{\mathbf{t}=\mathbf{t}_0} = \mathbf{c}(\mathbf{t}_0, r). \quad (3-43)$$

Nos pontos em cuja vizinhança $\mathbf{c}$ não seja constante, a função posto não é diferenciável. De fato, se $\mathbf{t}$ satisfaz as hipóteses da definição 2(c), temos que $\mathbf{t}$ e $\mathbf{t}_0$ apresentam a mesma ordem decrescente de suas componentes, consequentemente, como o posto r é fixo, o vetor indicador posto $\mathbf{c}$ de $\mathbf{t}$ e $\mathbf{t}_0$ serão os mesmos, pois o vetor indicador posto simplesmente marca as localizações em $\mathbf{t}$ onde $z = \mathcal{R}_r(\mathbf{t})$ ocorrem. Ou seja, teremos que $\mathbf{c}(\mathbf{t}, r) = \mathbf{c}(\mathbf{t}_0, r)$. Assim, considerando $\mathbf{c}(\mathbf{t}, r)$ constante em uma vizinhança de $\mathbf{t}_0$, e usando a representação de produto interno para a função posto $\mathcal{R}_r(\mathbf{t})$, (Definição 2(c)), para todo $\mathbf{t}$ dentro da vizinhança, temos que:

$$\nabla \mathcal{R}_r(\mathbf{t})|_{\mathbf{t}=\mathbf{t}_0} = \nabla(\mathbf{c}(\mathbf{t}_0, r) \cdot \mathbf{t}^\top)|_{\mathbf{t}=\mathbf{t}_0} = \mathbf{c}(\mathbf{t}_0, r). \quad (3-44)$$

F. Redes Neurais Morfológicas Modulares (MMNN)

A rede neural morfológica modular (MMNN), idealizada por R Sousa et al. (DE SOUSA et al., 2000), é uma rede neural *feed forward* e sua arquitetura é baseada nas decomposições de Banon e Barrera (BANON et al., 1993), que trata o que diz respeito à decomposição de um operador $\psi : \mathbb{L} \rightarrow \mathbb{M}$ em termos de um supremo de ínfimo de erosões e anti-dilatações ou, semelhantemente, de um ínfimo de supremos de dilatações e anti-erosões. A expressão decomposição de Matheron refere-se à decomposição de um operador ψ que é crescente e invariante com respeito à translações em termos de um supremo de erosões ou, da mesma forma, em termos de um ínfimo de dilatações, onde estes operadores são definidos sobre reticulados da forma $\mathcal{P}(\mathbb{R}^d)$. Podemos citar como exemplo onde estas decomposições podem ser utilizadas: detecção de bordas, restauração de imagens e reconhecimento de padrões. A MMNN, pode ser utilizada para desenvolver operadores invariantes por translação binária ou

em tons de cinza. Ao desenvolver operadores em tons de cinza, os valores de sinais são normalizados no conjunto $[0, 1]$. A arquitetura geral da MMNN é capaz de aprender tanto operadores invariantes por translação em tons de cinza quanto binários. Para o treinamento da MMNN, são usadas ideias do algoritmo *backpropagation* (BP) (SUSSNER 1999) e a metodologia proposta por Pessoa e Maragos (PESSOA et al., 1998) para superar o problema da não-diferenciabilidade das funções posto, bem como algoritmos genéticos adaptativos (AG). A estrutura da rede MMNN pode ser vista como um caso especial da rede MRL, mas com arquiteturas e regras de treinamento específico.

A Figura 3-11 apresenta as possíveis arquiteturas de uma rede MMNN de acordo com a decomposição de mapeamento crescente invariante à translação em termos de um supremo de erosões e em termos de um ínfimo de dilatações. A Figura 3-12 exibe as possíveis arquiteturas da rede MMNN segundo o teorema de Banon e Barrera (BANON et al., 1994). As duas estruturas representam decomposição por alguns módulos independentes.

Figura 3-11(a) Estrutura da rede MMNN usando dilatações morfológicas, (b) Estrutura da rede MMNN usando erosões morfológicas.

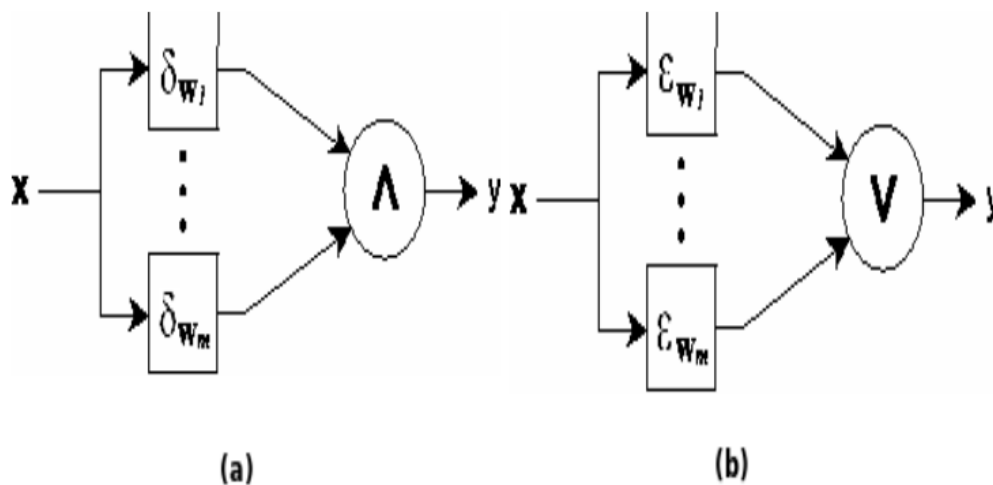
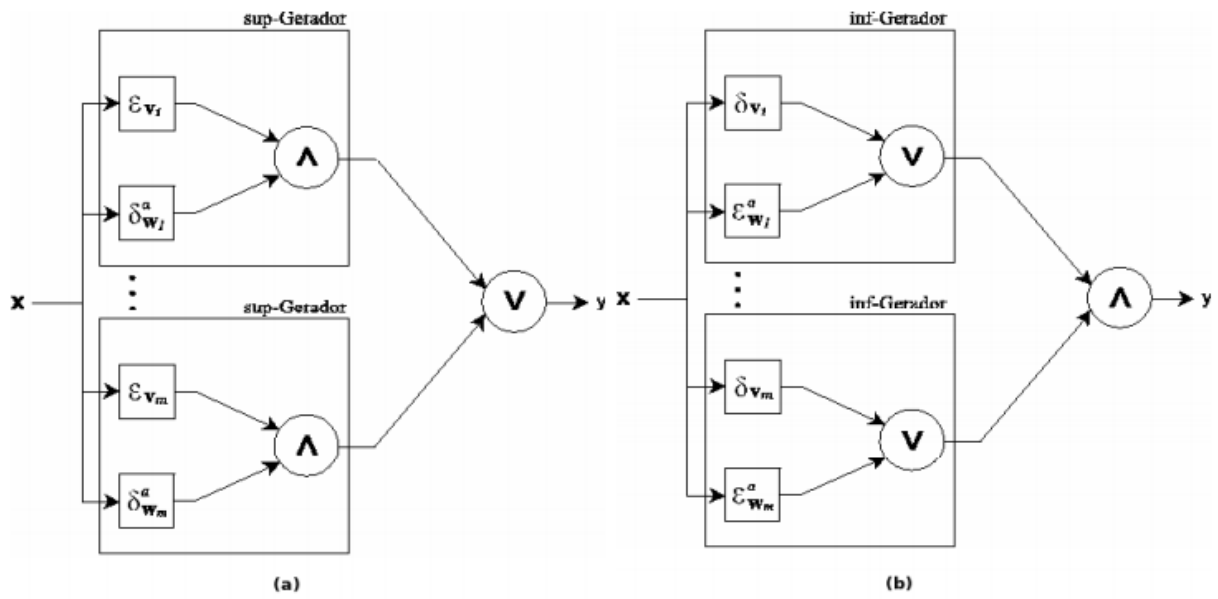


Figura 3-12(a) Arquitetura usada para a decomposição de Banon e Barrera por sup-geradores, (b) Arquitetura usada para a decomposição de Banon e Barrera por inf-geradores.



Definimos a arquitetura ilustrada na Figura 3-12 de acordo com a decomposição por erosões de matheron (MATHERON 1975):

$$\varepsilon_k = v_k = \bigwedge_{i=1}^N (x_i - a_{ki}), \quad (3-45)$$

$$y = V_{i=1}^N \{v_i\}, \quad \text{onde } \mathbf{v} = (v_1, v_2, \dots, v_N), \quad (3-46)$$

onde $\mathbf{x}$ é o sinal de entrada para a MMNN, $\mathbf{a}_k \in \mathbb{R}^N$ representam os elementos de estruturação das erosões v_k , $k = 1, 2, \dots, N$.

A matriz de pesos sinápticos, A , da MMNN é definida por:

$$A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N), \quad (3-47)$$

De uma maneira dual, a arquitetura para a decomposição de Matheron (MATHERON 1975) por dilatações, de acordo com a Figura 3-12(b), é definida substituindo erosões por dilatações e o símbolo $\wedge$ por $\vee$.

As equações apresentadas a seguir representam a arquitetura da Figura 3-12(a), de acordo com a representação de Banon e Barrera (BANON et al., 1994) através de sup-geradores.

$$\varepsilon_k = u_{k1} = \bigwedge_{i=1}^N (x_i - a_{ki}), \quad (3-48)$$

$$\delta_k^a = u_{k2} = 1 - V_{i=1}^N (x_i + b_{ki}), \quad (3-49)$$

$$v_k = \bigwedge_{i=1}^N u_{ki}, \quad (3-50)$$

$$\mathbf{u}_k = (u_{k1}, u_{k2}), \quad (3-51)$$

$$k = 1, 2, \dots, N$$

$$y = V_{i=1}^N v_i, \quad (3-52)$$

onde, $\mathbf{v} = (v_1, v_2, \dots, v_N)$.

Usando a equação $v_*(x) = (x^*)^\top, \forall x \in \mathbb{G}^n$, observamos que:

$$\delta_w^a(x) = v_*(\mathbf{w} \wedge x) = [(\mathbf{w} \wedge x)^*]^\top, \quad (3-53)$$

$$= (\mathbf{w} \wedge x)^* = x^* \vee \mathbf{w}^* = (\mathbf{w}^*)^\top \wedge (x^*)^\top, \quad (3-54)$$

$$= \delta_{(\mathbf{w}^*)^\top}((x^*)^\top) = \delta_{(\mathbf{w}^*)^\top}(v_*(x)) = \bar{\delta}_{(\mathbf{w}^*)^\top}(x), \quad (3-55)$$

Isso mostra, essencialmente, que as MMNNs pertencem à classe de perceptrons morfológicos. Similarmente, é possível mostrar que $\varepsilon_w^a = \bar{\varepsilon}_{(\mathbf{w}^*)^\top}(x)$.

As matrizes de pesos sinápticos das redes MMNN são definidos por:

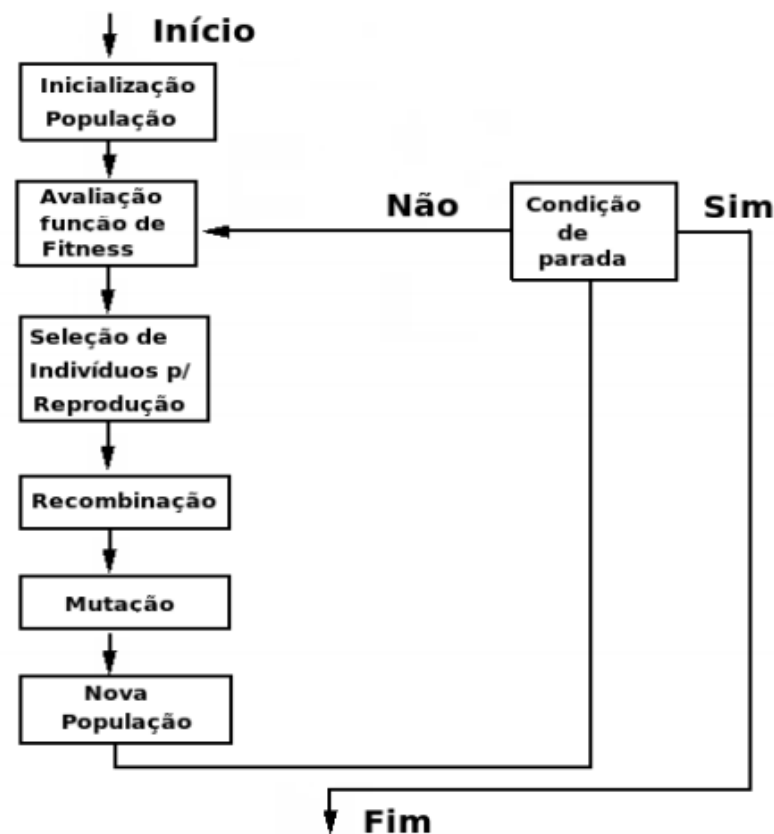
$$A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N),$$

$$B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N),$$

onde $\mathbf{a}_k, \mathbf{b}_k \in \mathbb{R}^N$ representam os elementos de estruturação e são considerados parâmetros livres da rede neural MMNN.

De maneira dual, a arquitetura para a decomposição de Banon e Barrera (BANON et al., 1994) através de infgeradores, Figura 3-12(b), é definida substituindo dilatações por erosões, anti-dilatações por anti-erosões, o símbolo $\wedge$ por $\vee$ nas unidades sub-integradoras, e o símbolo $\vee$ por $\wedge$ na unidade integradora geral.

Figura 3-13 Algoritmo genético simples.



G. Morfologia Matemática e novas aplicações

No trabalho de (IKONEN et al., 2015) o núcleo de células cancerosas foi estudado do ponto de vista morfológico com foco na análise de: Anisonucleose, deformidade nuclear, hipercromasia e núcleo citoplasmático com relação a soluções CAD (*Computer Assisted Diagnosis*). Morfologia Matemática (MM), mais especificamente a transformada *Top-Hat* foi usada junto com a manipulação dos canais de cores RGB para processar imagens histopatológicas para detectar e extrair núcleos para análise posterior. Alguns resultados preliminares são apresentados e eles parecem ser bastante promissores na área de extração núcleos. A separação de núcleos a partir de outras seções de tecidos é o componente chave na análise de tecidos cancerosos. Além disso, novas ideias são apresentadas e que serão utilizadas em novos trabalhos futuros dos autores (IKONEN et al., 2015).

- Transformada clássica *Top-hat*

A transformada *Top-Hat* é uma operação morfológica importante na detecção de pequenos alvos. A análise da transformada de dois conjuntos, de imagens originais e elementos estruturantes, calculando uma abertura numa imagem e subtraindo o resultado a partir da imagem original. A transformada *Top-Hat* é capaz de detectar características da imagem que correspondem aos elementos estruturantes usados. A transformada White Top-Hat (WTH) é definida como a diferença entre a função e sua abertura morfológica e as regiões brilhantes extraídas de uma imagem, de acordo com elemento estruturante usado (BAI et al., 2012):

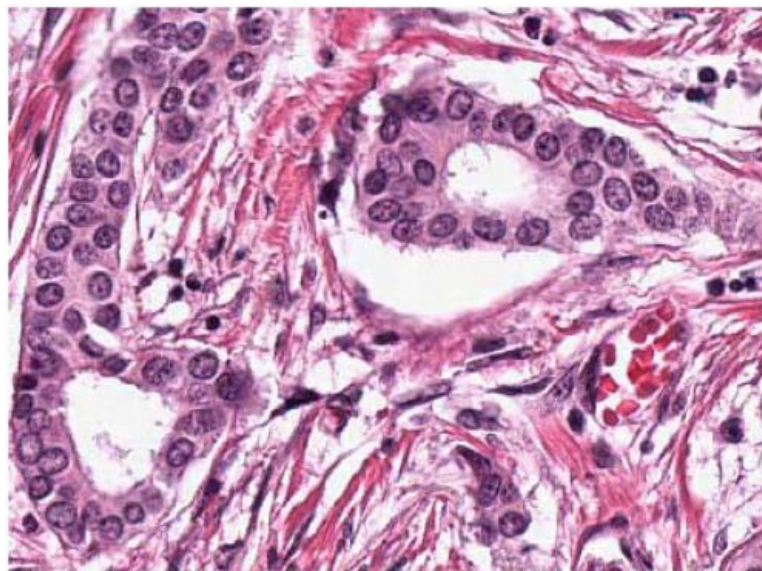
$$WTH(x, y) = f(x, y) - f \circ B(x, y), \quad (3-56)$$

A transformada *Black Top-Hat* (BTH) usa o fechamento para melhorar as regiões escuras (BAI et al., 2012).

$$BTH(x, y) = f \bullet B(x, y) - f(x, y), \quad (3-57)$$

O núcleo numa imagem aparece pequeno e com manchas escuras dentro de uma célula, como pode ser observado na Figura 3-14. Portanto, o uso da transformada *Black Top-Hat* é justificado.

Figura 3-14 Imagem histopatológica (IKONEN et al., 2015).



- **Modificações para a transformada Top-Hat**

A transformada *Top-Hat* clássica usa apenas um elemento estruturante, o que torna a utilização de informações diferentes entre a região de interesse (ROI) e as regiões circunvizinhas

difícil. Várias abordagens para a melhoria da transformada Top-Hat foram propostas (BAI et al., 2012):

- 1) *Pré-processamento ou pós-processamento*: Não melhora o desempenho da transformada.
- 2) *Suprimir certos pixels*: Tradicionalmente a detecção de bordas é realizada de acordo com a propriedade e o limiar do gradiente, o que elimina detalhes finos nas regiões escuras. Como uma melhoria, Chen et al. propuseram um algoritmo que detecta as bordas nas regiões escuras e também preserva as características das bordas nas regiões brilhantes (CHEN et al., 2002).
- 3) *Aplicar a técnica de multi-escala*: O uso de apenas um elemento estruturante pode extrair características que correspondem ao tamanho do elemento estruturante utilizado. Com a finalidade de extrair todas as características desejadas, elementos estruturantes multi-escala com diferentes tamanhos devem ser abordados. No entanto, o tempo de cálculo excederia a um, por exemplo, wavelet pirâmide e algoritmos médios diretos (BAI et al., 2011).
- 4) *Aplicar elementos multi estruturantes*: A construção de elementos multi estruturantes têm sido utilizados para melhorar o desempenho de operações morfológicas na detecção de características (Huang et al., 2012). Uma característica é considerada como um conjunto que é dividido em subconjuntos e elementos estruturantes são construídos correspondentes a esses diferentes subconjuntos (BAI et al., 2010).

• Morfologia do núcleo

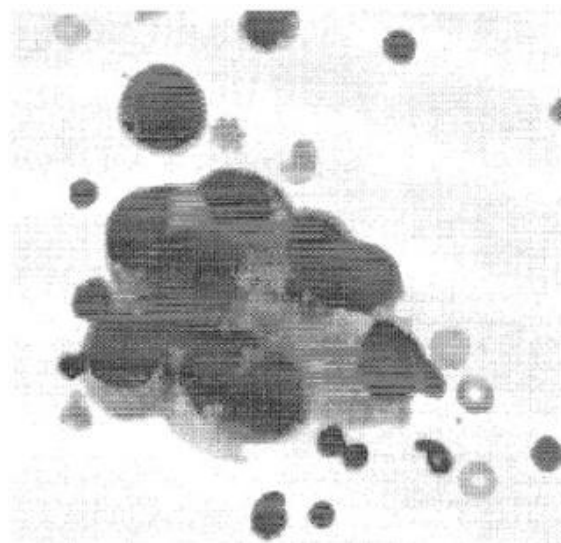
As células malignas possuem morfologia irregular que é mais visível no núcleo da célula, bem como no citoplasma. As células malignas possuem morfologia irregular que é mais visível no núcleo da célula, bem como no citoplasma. Os núcleos malignos são maiores e variam em tamanho e forma, comparado aos núcleos benignos. A variação no tamanho é chamada de anisonucleose e a variação na forma nuclear é chamada de deformidade. Uma área nuclear maligna exibe hipercromasia, o que significa a distribuição irregular de cromatina que aparece como aglomeração e condensação (THIRAN et al., 1996). Nas células malignas o citoplasma é escasso, o núcleo alargado, mas o citoplasma não é.

A malignidade de um tecido é baseada nos critérios definidos a seguir:

- Relação núcleo citoplasmático
- Anisonucleose
- Deformidade nuclear
- Hiperchromasia

A Figura 3-15 representa uma imagem de tecido onde os quatro critérios malignidade mencionados são visíveis.

Figura 3-15 Tecido com características de malignidade (THIRAN et al., 1996).



- **Segmentação e extração de núcleos cancerosos**

A segmentação dos núcleos é importante devido as características malignas das células estarem sendo relacionadas com o próprio núcleo. Métodos de segmentação que não levam em conta a forma entre outros, são: Detecção de bordas, limiarização hierárquica e histograma limiar 2D (THIRAN et al., 1996). (YU et al., 2004) apresentou uma extensão da morfologia matemática que usa a correlação entre os três canais de cores para processar uma imagem colorida de forma eficiente. A abordagem morfológica de cores pode ser utilizada, entre outras, na extração de características (YU et al., 2004). O trabalho de (YU et al., 2004) utilizou a manipulação do canal de cores RGB como uma etapa de pré-processamento antes de usar métodos de morfologia matemática. A análise de formas é a chave e morfologia matemática fornece uma abordagem para fazê-la. Thiran et al. propôs um método onde a morfologia matemática é desenvolvida para remover o ruído de fundo de uma imagem e o segmento dos

núcleos, após, a forma e o tamanho do núcleo é analisada e a textura avaliada (THIRAN et al., 1996).

a) Relação núcleo citoplasmático

Relação núcleo citoplasmático é avaliado pela primeira contagem de todos os pixels não brancos nas imagens de núcleos, que vai para a primeira parte da relação. Após a extração do citoplasma da imagem pela operação de uma diferença entre a imagem original e a imagem do núcleo. Também um duplo limiar é executado para remover o ruído de fundo e as manchas. Com um pequeno elemento estruturante, a imagem resultante é, então, fechada para uniformizar a região citoplasmática. Como resultado, é obtida uma imagem onde apenas o citoplasma está presente e os pixels não brancos são calculados. A relação núcleo citoplasmático é então obtida pela comparando a área nuclear e a área citoplasmática (THIRAN et al., 1996).

b) Anisonucleose e deformidade nuclear

Anisonucleose é avaliada primeiramente encontrando o centro de cada núcleo e, em seguida, calculando a distância do centro para a borda do núcleo em várias direções. Em busca do centro de gravidade, um grupo de erosões são realizados para separar diferentes partes da imagem uma da outra. Quando as distâncias das características têm sido obtidas, a anisonucleose é avaliada por variações do valor da média das distâncias das características. A deformidade nuclear é calculada da mesma forma (THIRAN et al., 1996).

c) Hiper Cromasia

Hiper Cromasia é extraída de uma imagem pela transformada Top-Hat. A textura granular do núcleo é eliminada por fechamento com um pequeno elemento estruturante que após o fechamento a imagem é subtraída da imagem original. Hiper Cromasia é avaliada através da análise do valor médio das pontas sobre a área nuclear (THIRAN et al., 1996).

- **Resultados**

O trabalho de pesquisa (IKONEN et al., 2015) é dividido em duas etapas que formam *pipeline* do estudo apresentado. O primeiro passo é realizar a manipulação dos canais RGB usando a modificação apresentada anteriormente no trabalho de (ALI et al., 2013). O segundo passo é a aplicação da transformada Top-Hat na imagem resultante e analisar os resultados.

a) Manipulação RGB

No trabalho feito anteriormente por (ALI et al., 2013), foi sugerido o uso de uma imagem de referência formada a partir da média dos canais R e G. Em seguida, essa média foi subtraída do canal B de acordo com a equação (ALI et al., 2013):

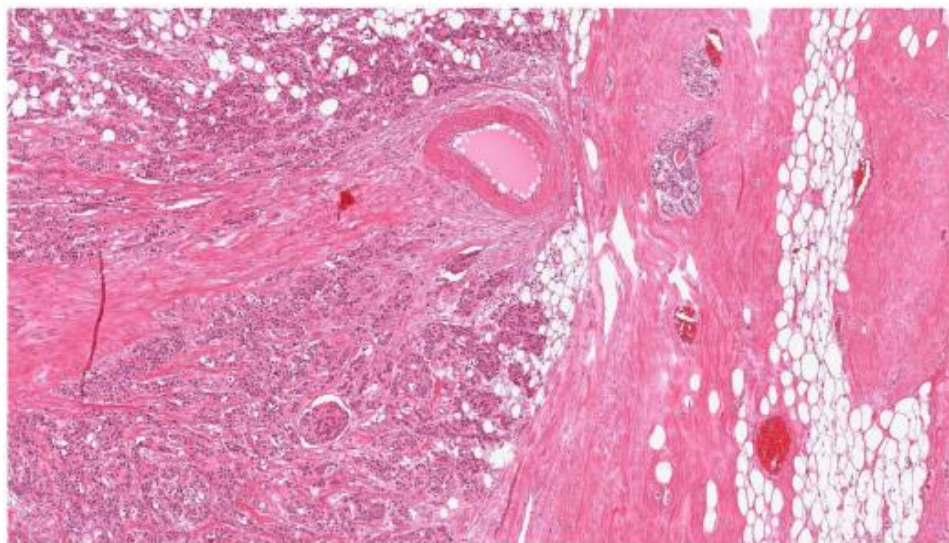
$$B - (R + G)/2, \quad (3-58)$$

Para melhorar a imagem histopatológica, a manipulação dos canais RGB foram aplicados de acordo com a fórmula proposta em (IKONEN et al., 2015):

$$R - (G + B)/2, \quad (3-59)$$

A análise é apresentada na Figura 3-16, que representa um caso de carcinoma ductal in situ. Este tipo de câncer é o mais comum não invasivo câncer de mama, que é localizado nos dutos de leite da mama.

Figura 3-16 Imagem original TIFF do tecido canceroso (IKONEN et al., 2015).



A imagem resultante da equação $B - (R + G)/2$ é demonstrada na Figura 3-17 e os resultados da imagem resultantes da equação $R - (G + B)/2$ são apresentados na Figura 3-18. Na imagem obtida através da fórmula proposta em (IKONEN et al., 2015) mostra a área do núcleo mais visível, isto é, a área de gordura tem-se disposta a partir da imagem excluindo as bordas. Principalmente apenas a área núcleos é destacada como um tom mais escuro de cinza, que é o ponto de partida fundamental para analisar os quatro critérios de malignidade.

Figura 3-17 Resultados obtidos através da equação $B - (R + G)/2$ (IKONEN et al., 2015).

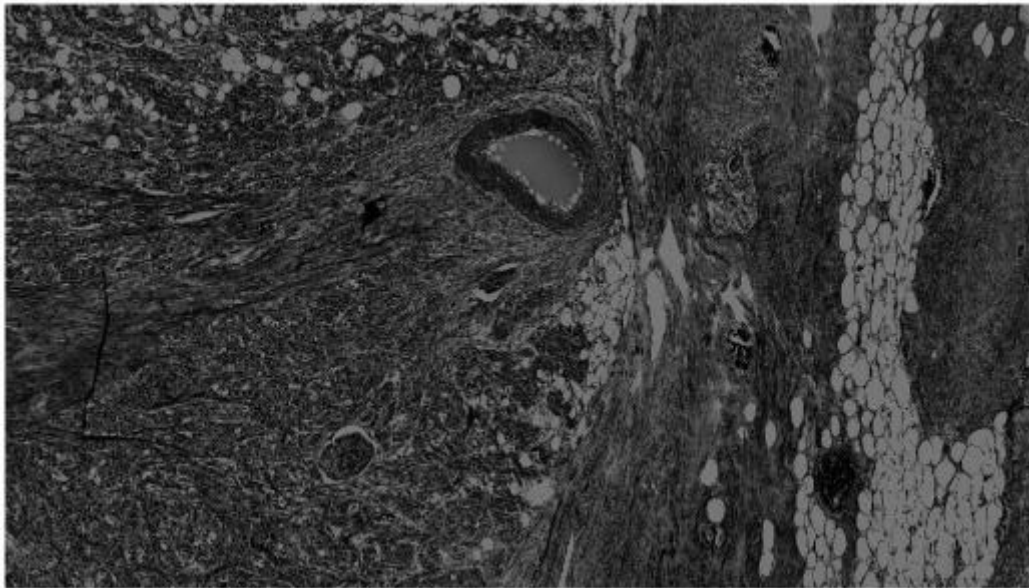
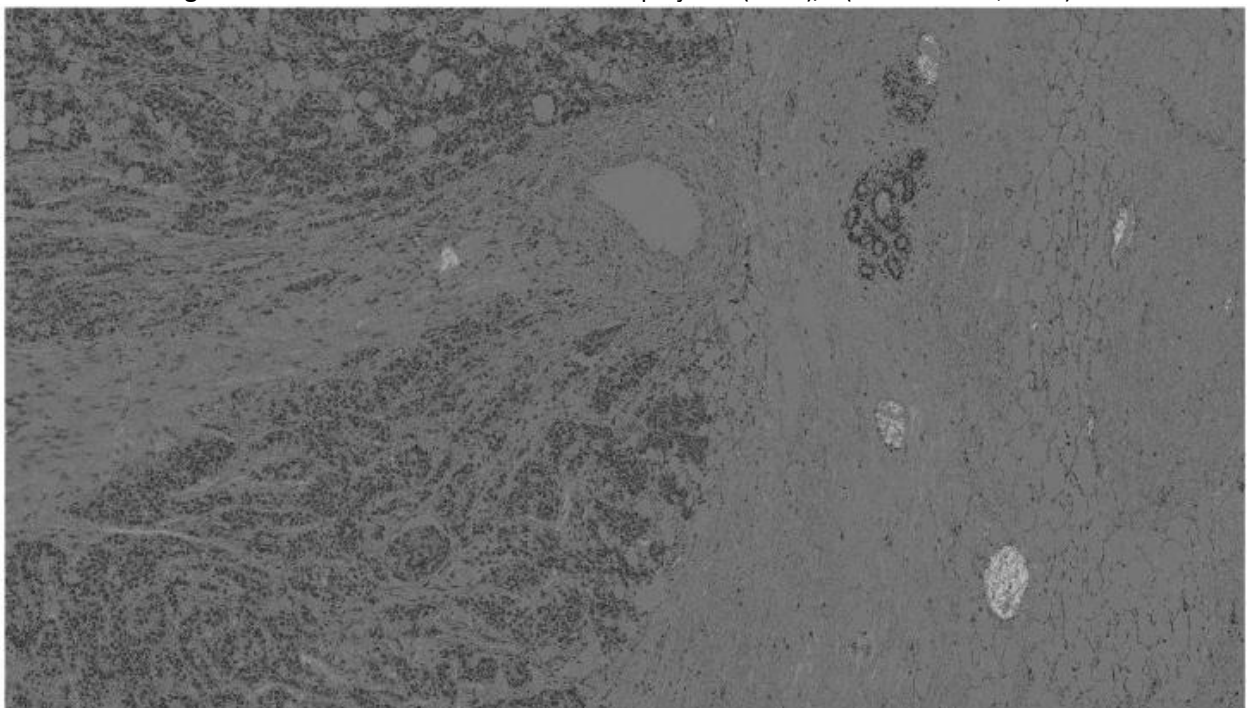


Figura 3-18 Resultados obtidos através da equação $R - (G + B)/2$ (IKONEN et al., 2015).



b) *Transformada Top-Hat*

Após a manipulação dos canais de cores RGB, a transformada *Top-Hat* é aplicada com um elemento estruturante em forma de disco com um raio de 3.O resultado pode ser visto na Figura 3-19. A imagem mostra uma diferença de contraste entre a área nuclear e outras seções de tecidos. Para demonstrar a utilidade da manipulação RGB, a imagem contendo os resultados obtidos através da transformada *Top-Hat* sem a manipulação RGB pode ser observado na Figura 3-20. Como proposta para trabalhos futuros, os autores do trabalho (IKONEN et al., 2015) pretendem aplicar melhorias a transformada *Top-Hat* tendo como objetivo alcançar resultados ainda melhores na difetença de intensidade.

Figura 3-19 Imagem resultante da transformada Top-Hat após a manipulação RGB (IKONEN et al., 2015).

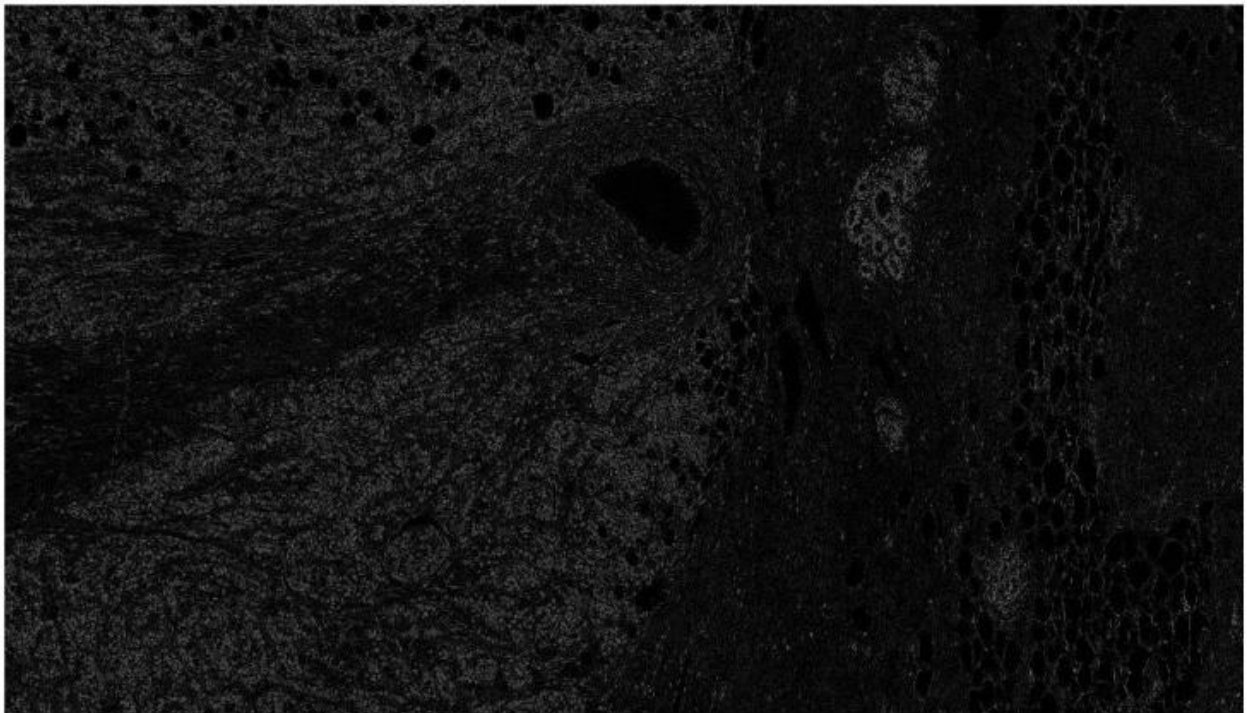
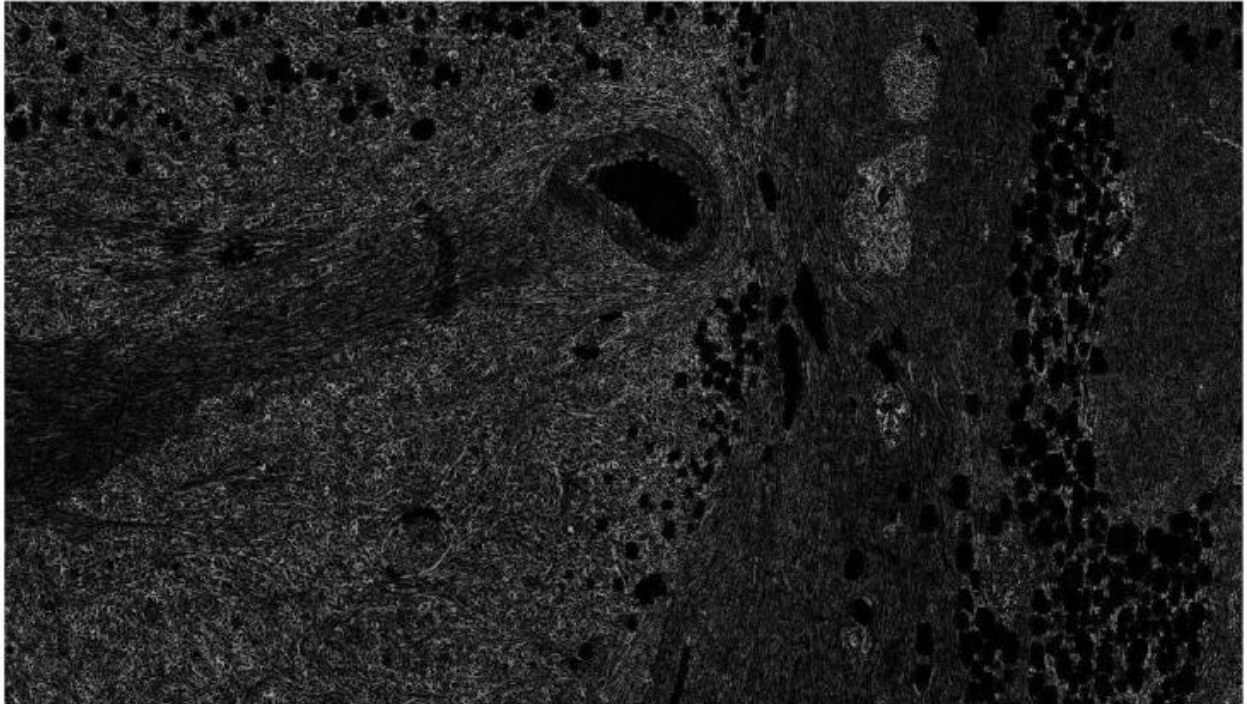


Figura 3-20 Imagem resultante da transformada Top-Hat sem a manipulação RGB (IKONEN et al., 2015).



3.4 Extreme Learning Machines

Redes Neurais Artificiais é uma das técnicas mais bem sucedidas utilizadas em problemas de classificação (JEATRAKUL et al., 2009). Na maioria dos tipos de RNAs, como Perceptron multicamadas (KRUSE et al., 2013), o conhecimento sobre os parâmetros da rede é necessário para obter um desempenho ótimo na solução do problema de classificação. Uma preocupação comum deste tipo de rede é evitar ser preso em mínimos locais, tornando necessário adicionar métodos para controlar a rede, a fim de escapar dessas regiões. Outra característica comum deste tipo de rede é o elevado tempo de treinamento necessário para tornar a rede capaz de executar classificações corretamente.

Extreme Learning Machine (HUANG et al., 2012) tem como principal característica a velocidade de treinamento e predição de dados comparada com outros métodos não lineares. As redes ELM geralmente são formadas por uma única camada intermediária, não recorrente, baseada num método analítico para estimar os pesos de saída da rede, em qualquer inicialização aleatória de pesos de entrada (CAMBRIA et al., 2013), (HUANG et al., 2012). Como os pesos das saídas são determinados de forma não iterativa, a rede é obtida através de alguns passos e baixo custo computacional (HUANG et al., 2012). Diferente da rede MLP, na rede ELM

não é necessário definir um critério de parada para a fase de treinamento, nem o uso de validação cruzada, como a rede apresenta uma única iteração (HUANG et al., 2012).

O processo de aprendizagem da rede ELM baseia-se na função inversa generalizada de Moore-Penrose (pseudo-inversa) (NASHED 2014), onde os pesos da saída são calculados. O processo de aprendizagem ocorre em lote, com dados apresentados à rede antes do ajuste de pesos em uma única iteração, tornando o treinamento mais rápido do que as abordagens convencionais (HUANG et al., 2012), (HUANG et al., 2011). A rede ELM não se baseia no gradiente descendente. Assim, a rede é livre do problema de mínimos locais. Além disso, não é necessária a definição de um parâmetro de taxa de aprendizagem nem o máximo número de iterações, uma vez que o algoritmo não é recursivo (MICHE et al., 2010), (HUANG et al., 2008), (HUANG et al., 2006). O resultado é uma camada de saída formada por neurônios de *kernel* linear semelhante à máquina de aprendizagem linear e ao perceptron de uma única camada.

Vamos considerar uma camada intermediária de uma rede ELM com m_h neurônios intermediários e n entradas, a saída será dada por:

$$y_{h,i} = K_h(x, w_{h,i}), \quad (3-60)$$

Onde $y_{h,i}$ é a saída e $w_{h,i} \in \mathbb{R}^n$ são os vetores de pesos do i -ésimo neurônio da camada escondida; $x \in \mathbb{R}^n$ é o vetor das entradas, enquanto $K_h : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ é o kernel dos neurônios da camada escondida, para $i = 1, 2, \dots, m_h$.

Como exemplo dos *kernels* mais usados na camada oculta, podemos citar o *kernel* linear

$$K_h(x, w_{h,i}) = b_i + x \bullet w_{h,i}, \quad (3-61)$$

onde b_i é o *bias* e $\bullet$ indica o produto interno entre dois vetores; e a função gaussiana de base radial do *kernel*

$$K_h(x, w_{h,i}) = \exp\left(-\frac{\|x - w_i\|^2}{2\sigma_i^2}\right), \quad (3-62)$$

onde σ_i controla o raio da função gaussiana, para $i = 1, 2, \dots, m_h$.

Considerando uma rede ELM com uma camada intermediária com m saídas, n entradas, e m_h neurônios na camada escondida, a saída é dada por:

$$y_i = \sum_{j=1}^{m_h} w_{o,i,j} K_h(x, w_{h,j}), \quad (3-63)$$

Onde $w_{o,i,j}$ é o j -ésimo peso sináptico da saída i -ésimo do neurônio, para $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, m_h$.

Vamos considerar o conjunto de treinamento $\Omega = \{(x_i, y_i)\}_{i=1}^M$, a matriz H é dada por:

$$H = \begin{pmatrix} K_h(x_1, w_{h,1}) & \cdots & K_h(x_1, w_{h,m_h}) \\ \vdots & \ddots & \vdots \\ K_h(x_M, w_{h,1}) & \cdots & K_h(x_M, w_{h,m_h}) \end{pmatrix}. \quad (3-64)$$

As saídas da rede ELM são calculadas pela seguinte relação básica:

$$HW_o = Y, \quad (3-65)$$

Onde

$$W_o = (w_{o,1}^T, w_{o,2}^T, \dots, w_{o,m}^T)^T$$

e

$$Y = (y_1^T, y_2^T, \dots, y_m^T)^T.$$

O processo de treinamento em lote consiste em calcular W_o pela relação inversa:

$$W_o = \check{H}Y, \quad (3-66)$$

Onde $\check{H}$ é a inversa generalizada de Moore-Penrose de H .

3.5 Operadores Morfológicos

Morfologia Matemática é uma teoria completa de processamento não linear extremamente usada no processamento digital de imagem. Baseia-se nas transformações de formas preservando as relações de inclusão dos objetos. Seus operadores fundamentais são chamados de erosão e dilatação (SANTOS 2012), (SOILLE 2004). Além disso, a Morfologia Matemática pode ser considerada uma teoria construtiva, porque todos os operadores são construídos usando erosões e dilatações como bases. É possível a construção de vários

operadores destinados a aplicações específicas, como filtragem de imagens, extração de características e outros (SANTOS 2012), (SOILLE 2004).

É possível afirmar que as operações fundamentais da Morfologia Matemática consistem da investigação de imagens usando elementos da estrutura (SANTOS 2012), (HÄNNI et al., 1999), (SOILLE 2004). Elementos da estrutura são tão importantes para Morfologia Matemática como a máscara de convolução é para processamento de imagem linear. Na verdade, os elementos de estrutura são geralmente imagens com as dimensões menores do que as dimensões da imagem a ser morfológicamente transformadas. As várias transformações morfológicas são baseadas em operações lógicas entre o elemento estrutural e regiões da imagem sob o processamento.

Enquanto a máscara de convolução é usada para gerar médias ponderadas por regiões de filtragem linear, a estrutura do elemento tem uma função similar na Morfologia Matemática: ela é usada na erosão para maximizar regiões, gerando imagens mais claras para o caso de adotar os valores maiores de pixel para os *pixels* mais brilhantes. Da mesma forma, na erosão a estrutura do elemento é usada para minimizar regiões, gerando imagens mais escuras. Estas operações de minimização e maximização estão relacionadas com as operações lógicas de interseção e união. Isto leva-nos a pensar em estrutura de elementos como unidades básicas para reconstruir as regiões de imagem.

Uma *imagem* $f : S \rightarrow K$ é definida como uma função para mapear um vetor bidimensional $u \in S$ no ponto $f(u) \in K$ onde S é a *grade* ou imagem de *suporte*, u é o *pixel* da grade S ; $f(u) \in K$ é o valor do *pixel* $u \in S$ (Santos, 2012).

Mais rigorosamente, o *pixel* é definido como o par cartesiano $(u, f(u))$, isto é, o vetor formado pela posição u na grade S e seu valor associado $f(u)$. K é o conjunto de possíveis valores de $f(u)$. A notação K^S representa o conjunto de todas as possíveis imagens da grade S em K (SANTOS 2012), (SOILLE 2004).

Quando a grade S é discreta, isto é, $S \subseteq \mathbb{Z}^2$, dizemos que $f : S \rightarrow K$ é uma *imagem discreta*. K é definido por $k = V^p$, onde $V \in \mathbb{R}$ é o *conjunto de níveis de cinza* e $p \in \mathbb{N}^*$ é o *número de faixas* presentes na imagem f (SANTOS 2012).

A imagem f com p faixas, $f : S \rightarrow V^p$, pode também ser representado pela seguinte notação (SANTOS 2012):

$$f(u) = \{f_1(u), f_2(u), \dots, f_p(u)\}, \quad u \in S,$$

onde $f_i(u)$ é a i -ésima faixa de $f : S \rightarrow V^p$, para $f_i : S \rightarrow V$ e $1 \leq i \leq p$.

Uma imagem $f : S \rightarrow K$ é uma *imagem digital* quando é discreta e *quantizada*, isto é, o *pixel* assume apenas valores discretos. Matematicamente, $f : S \rightarrow K$ é uma imagem digital quando $K = \{0, 1, \dots, K\}^p$ e $S \in \mathbb{Z}$, onde $k, p \in \mathbb{N}^*$.

Usando uma imagem digital é representada utilizando N bits, $k = 2^n - 1$.

Para facilitar a notação, é relativamente comum para representar imagens digitais como imagens discreta normalizadas, mudando $K = \{0, 1, \dots, K\}^p$ por $\bar{K} = [0, 1]^p$. Usando esta notação não é necessário dizer quantos bits foram usados para representar os valores possíveis dos *pixels*. Portanto, uma imagem $f : S \rightarrow \{0, 1, \dots, k\}^p$ pode também ser representado por $\tilde{f} : S \rightarrow [0, 1]^p$.

As operações básicas entre as imagens na Matemática Morfológica são: comparação, subtração, interseção e união. É importante afirmar que as operações básicas só podem ser aplicadas a imagens que respeitam a relação de ordem entre os *pixels*. Apenas nível de cinza e imagens binárias mantêm a relação de ordem. Para lidar com imagens multiespectrais é necessário trabalhar com cada faixa separadamente ou adotar uma abordagem determinada para manter a relação de ordem. A relação de ordem é à base da Morfologia Matemática (SANTOS 2012), (SOILLE 2004). Dadas duas imagens $f_1 : S \rightarrow [0, 1]$ e $f_2 : S \rightarrow [0, 1]$ a *comparação* $\leq$ é a operação de $[0, 1]^S \times [0, 1]^S$ para $[0, 1]^S$ definido como se segue (SANTOS 2012), (SOILLE 2004):

$$(f_1 \leq f_2)(u) = \begin{cases} 1, & f_1(u) \leq f_2(u) \\ 0, & \text{c.c.} \end{cases}, \quad \forall_u \in S. \quad (3-67)$$

Dada à imagem $f : S \rightarrow [0, 1]$, o *negativo* de f , é denotada por $\bar{f}$ ou $\sim f$, é a operação de $[0, 1]^S$ para $[0, 1]^S$ definido como se segue (SANTOS 2012), (BANON et al., 1994):

$$\bar{f}(u) = 1 - f(u), \quad \forall_u \in S. \quad (3-68)$$

A *união* entre $f_1 : S \rightarrow [0, 1]$ e $f_2 : S \rightarrow [0, 1]$ é a operação de $[0, 1]^S \times [0, 1]^S$ para $[0, 1]^S$, denotado por $f_1 \vee f_2$, definido como se segue (SANTOS 2012), (SOILLE 2004):

$$(f_1 \vee f_2)(u) = \max\{f_1(u), f_2(u)\}, \quad \forall u \in S. \quad (3-69)$$

A *interseção* entre $f_1 : S \rightarrow [0, 1]$ e $f_2 : S \rightarrow [0, 1]$ é a operação de $[0, 1]^S \times [0, 1]^S$ para $[0, 1]^S$, denotado por $f_1 \wedge f_2$, definido como se segue (SANTOS 2012), (SOILLE 2004):

$$(f_1 \wedge f_2)(u) = \min\{f_1(u), f_2(u)\}, \quad \forall u \in S. \quad (3-70)$$

É importante notar que as operações de união e interseção entre as imagens são semelhantes às respectivas operações entre conjuntos *fuzzy* (SANTOS 2012). As semelhanças entre Morfologia Matemática e operações de Lógica Fuzzy aumentam com a adoção da notação de imagens normalizadas (SANTOS 2012). Consequentemente, também é possível pensar em imagens normalizadas como conjuntos *fuzzy*.

A *subtração* ou a *diferença* entre $f_1 : S \rightarrow [0, 1]$ e $f_2 : S \rightarrow [0, 1]$ é a operação de $[0, 1]^S \times [0, 1]^S$ para $[0, 1]^S$ dado por (SANTOS 2012), (BANON et al., 1994):

$$(f_1 \sim f_2)(u) = (f_1 \wedge f_2)(u), \quad \forall u \in S. \quad (3-71)$$

É possível afirmar que Morfologia Matemática é uma teoria construtiva, dado o fato de que os operadores mais complexos podem ser construídos pelo mais simples. Os operadores fundamentais, base para outros operadores, são a dilatação e erosão. Todos os outros operadores são construídos como combinações destes dois operadores fundamentais (SANTOS 2012), (SOILLE 2004).

A *dilatação* de $f : S \rightarrow [0, 1]$ por $g : S \rightarrow [0, 1]$, denotado por $\delta_g(f)$ ou $f \oplus g$ (Soma de Minkowski, em casos binários (SANTOS 2012), (GONZALEZ et al., 2007), (SOILLE 2004)), é um operador de $[0, 1]^S \times [0, 1]^S$ para $[0, 1]^S$ definido como se segue:

$$\delta_g(f)(u) = (f \oplus g)(u) := \bigvee_{v \in S} f(v) \wedge g(u - v), \quad \forall u \in S. \quad (3-72)$$

onde g é o elemento estrutural. A dilatação transforma a imagem original de modo a fazer com que o elemento estrutural se encaixe na imagem original, isto é, a imagem original f será modificada para tornar as áreas semelhantes ao aumento g , no caso g .

Se associarmos os 1's para o branco absoluto e os 0's para o preto absoluto, vamos perceber os resultados de dilatação no *aumento das áreas mais brilhantes* e a *eliminação das áreas mais escuras*.

A erosão de $f : S \rightarrow [0, 1]$ por $g : S \rightarrow [0, 1]$, denotado por $\epsilon_g(f)$ ou $f \ominus g$ (Subtração de Minkowski, em casos binários (SANTOS 2012), (GONZALEZ et al., 2007), (SOILLE 2004)), é um operador de $[0, 1]^S \times [0, 1]^S$ para $[0, 1]^S$ definido como se segue:

$$\epsilon_g(f)(u) = (f \ominus g)(u) := \bigwedge_{v \in S} f(v) \vee \bar{g}(v - u), \quad \forall u \in S. \quad (3-73)$$

A erosão transforma a imagem original de modo a fazer o negativo do elemento estrutural no caso de a imagem original, isto é, a imagem original f será modificada para tornar as áreas semelhantes ao aumento $\bar{g}$, no caso $\bar{g}$.

Se associarmos os 1's para o branco absoluto e os 0's ao preto absoluto, vamos perceber que os resultados de erosão no *aumento das áreas mais escuras* e a *eliminação das áreas mais brilhantes*.

Na Tabela 3-1, mostramos as expressões matemáticas de dilatação, erosão e convolução. Ela ilustra algumas semelhanças entre estas operações.

Tabela3-1 Dilatação, erosão e convolução de imagens.

Operação	Expressão
Dilatação	$(f \oplus g)(u) = \bigvee_{v \in S} f(v) \wedge g(u - v)$
Erosão	$(f \ominus g)(u) = \bigwedge_{v \in S} f(v) \vee \bar{g}(v - u)$
Convolução	$(f * g)(u) = \sum_{v \in S} f(v) g(u - v)$

Um resultado interessante poderia ser obtido se nós aplicarmos o Teorema de De Morgan a expressão 3-73:

$$\delta_g(f)(u) = \overline{\bigwedge_{v \in S} \bar{f}(v) \vee \bar{g}(u - v)}, \quad \forall u \in S. \quad (3-74)$$

Se o elemento estrutural g é simétrico, isto é, $g(u) = g(-u)$, $\forall u \in S$, a expressão 3-74 torna-se:

$$\overline{\delta_g(f)}(x) = \bigwedge_{v \in S} \bar{f}(v) \vee \bar{g}(v - u), \quad \forall_u \in S. \quad (3-75)$$

A partir da definição da erosão, mudando f por $\bar{f}$, temos uma relação importante:

$$\epsilon_g(f)(u) = \overline{\delta_g(\bar{f})(u)}, \quad \forall_u \in S. \quad (3-76)$$

Para todo $g : S \rightarrow [0, 1]$ simétrico. Esta relação é de fundamental importância, porque quantitativamente expressa a *dualidade* entre dilatação e erosão.

A n -dilatação de $f : S \rightarrow [0, 1]$ de $g : S \rightarrow [0, 1]$, denotado por $\delta_g^n(f)$, é dada pela seguinte expressão, para $n > 1$ (SANTOS 2012):

$$\delta_g^n(f)(u) := \underbrace{\delta_g \delta_g \dots \delta_g}_n(f)(u), \quad \forall_u \in S. \quad (3-77)$$

Do mesmo modo, a n -erosão de $f : S \rightarrow [0, 1]$ de $g : S \rightarrow [0, 1]$, denotada por $\epsilon_g^n(f)$, é dada como se segue, para $n > 1$ (SANTOS 2012):

$$\epsilon_g^n(f)(u) := \underbrace{\epsilon_g \epsilon_g \dots \epsilon_g}_n(f)(u), \quad \forall_u \in S. \quad (3-78)$$

Banon (1994) (BANON et al., 1994) demonstrou que cada elemento da estrutura, cujas dimensões são maiores do que 3×3 pode ser decomposto em vários elementos de estruturais 3×3 , isto é, uma dilatação por um determinado elemento estrutural $h \times m \times n$, desde que $m > 3$ e $n > 3$, pode ser executado através da utilização de um determinado número de dilatações por elementos estruturais de b_k , com dimensões de 3×3 , onde b_k é um elemento estrutural compondo h [36].

3.6 Descritores de Imagem

A partir de uma imagem digital podem ser extraídas características da imagem, tendo como finalidade identificar alguma propriedade ou elemento contido na imagem. As características de uma imagem são constituídas por informações numéricas, oriundas de análise da decomposição dos pixels, que podem ajudar na classificação de um objeto na imagem, tal como um tumor em uma imagem de mamografia.

A extração desses atributos pode ser aplicados à imagem toda, para extrair informação de textura, por exemplo, ou para regiões de objetos, tais como área, perímetro e forma de uma região segmentada. Geralmente essa extração é realizada por descritores de imagem, os quais podem extrair características quanto à forma, ou textura, por exemplos. Normalmente esses descritores apresentam propriedades invariantes à rotação e escala, o que permite que mesmo se o objeto estiver em uma escala menor ou rotacionado, o valor do descritor será o mesmo.

1) Atributos de Textura de Haralick: Os descritores de Haralick (HARALICK et al., 1973) são descritores de imagem bastante utilizados na literatura e que também têm sido aplicados para detecção de tumores em imagens de mamografia (JENIFER et al., 2014). Os descritores de Haralick descrevem características baseados em texturas realizado a partir de cálculos estatísticos entre pixels vizinhos, que definem características obtidas através de uma matriz de co-ocorrência. A matriz de co-ocorrência possui os valores de ocorrência de cores em uma determinada imagem. Essas características são utilizadas para diferenciar texturas que não seguem um determinado padrão de repetição. As matrizes de co-ocorrência representam a distribuição espacial e a dependência dos níveis de cinza dentro de uma área local. Cada $(i, j)^{th}$ de uma matriz é a possibilidade de passar de um *pixel* com nível de cinza i para outro com nível de cinza j sob um determinado ângulo e distância. De acordo com a literatura, os ângulos utilizados como padrão são: 0° , 45° , 90° e 135° . Este trabalho fixou a distância d como 32 características de textura, que são utilizadas como entradas para a rede neural. Conjuntos de vetores de características são calculados a partir dessas matrizes para fazer diferentes modelos de textura. Os atributos de Haralick (HUANG et al., 2008) são medidas estatísticas derivadas de cada matriz. O cálculo dos atributos de Haralick sobre a matriz de co-ocorrência $P(i, j)$ para um certo ângulo e uma certa distância d é representado pela equação 3-79 (HARALICK et al., 1973).

$$P(i, j, d, 0^\circ) = \#\{(k, l), (m, n) \in (L_y \times L_x) \times (L_y \times L_x) | k - m = 0, |l - n| = d, I(k, l) = i, I(m, n) = j\}$$

$$\begin{aligned}
P(i, j, d, 45^\circ) &= \#\{(k, l), (m, n)\} \\
&\in (L_y \times L_x) \times (L_y \\
&\times L_x) | (k - m = d, l - n = -d) \text{ ou } (k - m = -d, l - n = d), I(k, l) \\
&= i, I(m, n) = j\}
\end{aligned}$$

$$\begin{aligned}
P(i, j, d, 90^\circ) &= \#\{(k, l), (m, n)\} \in (L_y \times L_x) \times (L_y \times L_x) | |k - m| = d, l - n = 0, I(k, l) \\
&= i, I(m, n) = j\}
\end{aligned}$$

$$\begin{aligned}
P(i, j, d, 135^\circ) &= \#\{(k, l), (m, n)\} \in (L_y \times L_x) \times (L_y \times L_x) | (k - m = d, l - n = \\
&\text{dou } (k - m = -d, l - n = -d), I(k, l) = i, I(m, n) = j\} \quad (3-79)
\end{aligned}$$

onde # significa o número de elementos no conjunto.

Os descritores de Haralick são compostos por 13 descritores relacionados às características da imagem conforme descrito abaixo (HARALICK et al., 1973):

1. Energia ou Uniformidade:

$$f_1 = \sum_i \sum_j \{p(i, j)\}^2.$$

2. Contraste:

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}.$$

3. Correlação:

$$f_3 = \frac{\sum_i \sum_j (i, j) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}.$$

onde μ_x, μ_y, σ_x e σ_y são as médias e o desvio padrão de p_x e p_y .

4. Soma de Variância:

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j).$$

5. Momento da Diferença Inversa:

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j).$$

6. Média da Soma:

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i).$$

7. Variância da Soma:

$$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i).$$

8. Entropia da Soma:

$$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\}.$$

9. Entropia:

$$f_9 = - \sum_i \sum_j p(i, j) \log(p(i, j)).$$

10. Variância da Diferença:

$$f_{10} = \text{variância de } p_{x-y}.$$

11. Entropia da Diferença:

$$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\}.$$

12. Medida de Informação de Correlação 1:

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}.$$

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

13. Medida de Informação de Correlação 2:

$$f_{13} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2}.$$

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

Onde, nos descritores 12 e 13, HX e HY são as entropias de p_x e p_y , e

$$HXY1 = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(j)\}$$

$$HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}.$$

14. Matriz de Coeficiente de Correlação:

$$f_{14} = (\text{segundo maior valor próprio de } Q)^{1/2}$$

onde

$$Q(i, j) = \sum_k \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)}.$$

2) Decomposição em Série de Wavelets: A *wavelet* $\psi(t)$ é uma função em forma de onda que obedece às seguintes condições:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0, \quad (3-80)$$

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 dt < +\infty. \quad (3-81)$$

A *wavelet* $\psi(t)$ é chamada de *wavelet* mãe com as *wavelets* filhas $\psi_{a,b}(t)$ versões transladas e escalonadas da *wavelet* mãe, dada como:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad (3-82)$$

onde a e b são os fatores de escalonamento e transladação, respectivamente. Dada um sinal no tempo contínuo, a sua transformada *wavelet* contínua é dada por:

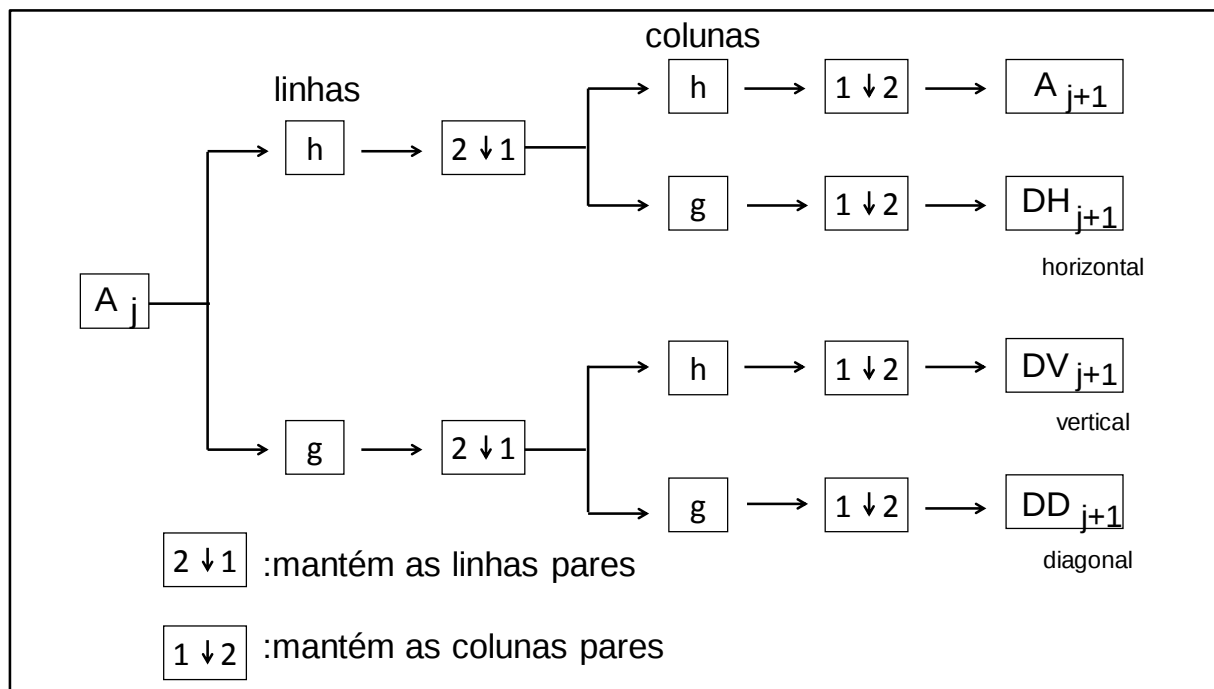
$$X(a, b) = \int_{-\infty}^{+\infty} x(t) \psi_{a,b}^*(t) dt, \quad (3-83)$$

onde $*$ significa o conjugado.

Wavelets são ferramentas muito eficazes para a representação das imagens em multi-resolução. A transformada de *wavelet* relativa a processamento de imagens pode ser implementada de maneira bidimensional. Mallat propôs uma Transformada de *Wavelet* Discreta de um sinal através da decomposição de uma imagem original em uma série de imagens geradas através de filtros discretos passa-alta e passa-baixa (MALLAT 1999).

A Figural 3-21 exibe o algoritmo de decomposição de *wavelets* bi-dimensional para um nível de resolução, proposto por (MALLAT 1999). A imagem A_j é convoluída pelos filtros $h(n)$ e $g(n)$. As duas imagens, resultantes da convolução, são decimadas por um fator de 2, onde as linhas ímpares são retiradas. Em seguida, as colunas das imagens decimadas são convoluídas pelos filtros $h(n)$ e $g(n)$. Logo as colunas ímpares são excluídas no processo de decimação. Em cada nível, são geradas uma imagem de aproximação e três imagens de detalhes.

Figura 3-21 Algoritmo de decomposição de *Wavelets* bi-dimensionais, para um nível de resolução.

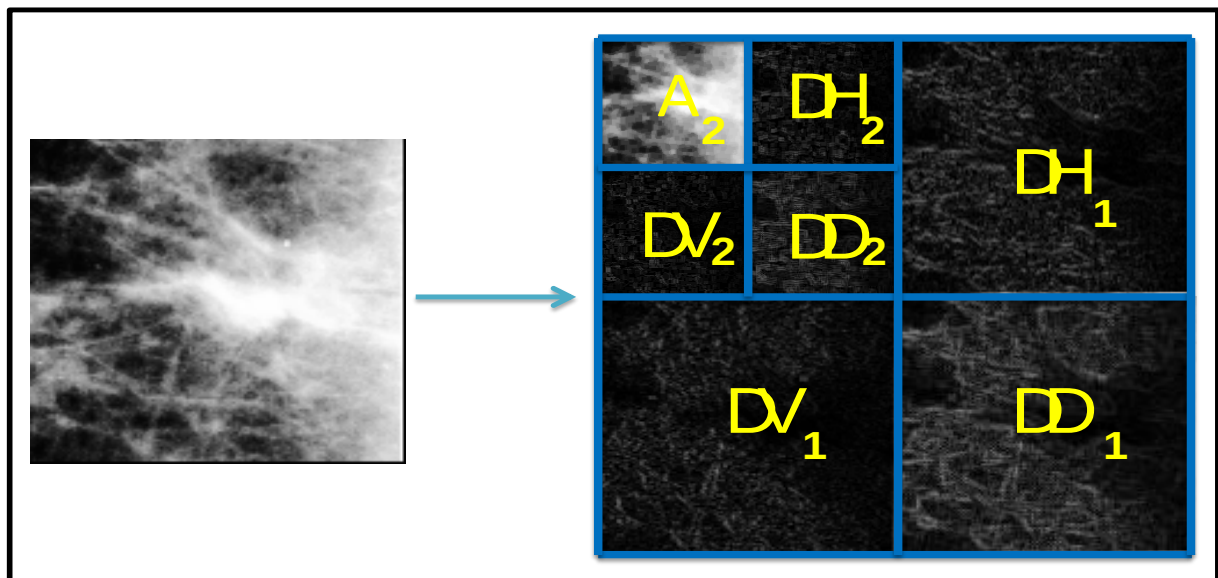


Uma imagem de aproximação A diz respeito a uma representação da imagem original sem detalhes, resultante da aplicação de um FPB (Filtro Passa-Baixa). O efeito visual é o embaçamento da imagem original, através da redução do número de níveis de cinza. Enquanto uma imagem de detalhes tem como objetivo representar a imagem original, em alta frequência. Um FPA (Filtro Passa-Alta) intensifica as bordas dos elementos presentes na imagem. Em cada nível das *wavelets*, o algoritmo de Mallat gera três imagens de detalhes. Uma para DH (Detalhes Horizontais), outra para DV (Detalhes Verticais) e por fim uma para DD

(Detalhes Diagonais). O processo de decomposição do nível seguinte emprega, como fonte, a imagem de aproximação do nível imediatamente anterior.

A Figura 3-22 representa o processo de decomposição das *wavelets*, em 2 níveis, baseada no algoritmo *Daubechies* 8. O objetivo dessa técnica é decompor a imagem original em uma imagem de aproximação e em três imagens de detalhes, para cada nível.

Figura 3-22 Ilustração da decomposição das *wavelets*, em 2 níveis, baseada no algoritmo *Daubechies* 8.



Os filtros dependem de cada algoritmo empregado pelas *wavelets*. Os algoritmos podem também ser chamados de família. Dentro de cada família, os filtros podem possuir comprimentos diferentes. Quanto maior for o comprimento dos filtros mais suave se tornam as imagens de aproximação. A suavidade implica na redução de detalhes da imagem. A família *Daubechies*, por exemplo, é formada por vários filtros de comprimentos diferentes. *Daubechies* 8 é capaz de suavizar a imagem com maior intensidade em relação à *Daubechies* 4. Isso ocorre porque os filtros das funções *Daubechies* 8 e *Daubechies* 4 possuem comprimento de 16 e 8, respectivamente.

3.7 Métricas

1) *Índice Kappa*: O índice Kappa (THOMAS et al., 2012) é uma métrica que avalia a precisão do classificador. O índice Kappa é utilizado como um método para avaliar a qualidade global de um processo de classificação de uma métrica única, envolvendo informações de falsos positivos, falsos negativos, verdadeiros positivos e verdadeiros negativos, que são obtidos a partir da matriz de confusão. A Tabela 3-2 mostra o cálculo feito na matriz de confusão.

Tabela 3-2 Matriz de Confusão

Classe	Predição C_1	Predição C_2	Precisão da Classe	Precisão Global
C_1	Verdadeiro Positivo V_p	Falso Negativo F_n	$\frac{V_p}{V_p + F_n}$	$\frac{V_p + V_n}{n}$
C_2	Falso Positivo F_p	Verdadeiro Negativo V_n	$\frac{V_n}{F_p + V_n}$	

A partir da Tabela 3-2, V_p , V_n e n são as quantidades de verdadeiras classificações positivas, verdadeiras negativas e a quantidade de dados, respectivamente. Classificações positivas verdadeiras são feitas quando a amostra da classe C_1 é prevista como sendo de classe C_1 . Classificação negativa verdadeira representa quando uma amostra de classe C_2 é previsto como sendo de classe C_2 . Falso negativo é obtido quando uma amostra de classe C_1 é previsto como classe C_2 , e Falso Positivo é quando uma amostra de classe C_2 é prevista como sendo de classe C_1 . O índice Kappa é calculado a partir da matriz de confusão através da seguinte equação:

$$k = \frac{n \sum_{i=1}^c x_{ii} - \sum_{i=1}^c x_{i+} x_{+i}}{n^2 - \sum_{i=1}^c x_{i+} x_{+i}}, \quad (3-84)$$

No qual x_{ii} é o valor da linha i e coluna i ; x_{i+} é a soma da linha i e x_{+i} é a soma da coluna i da matriz de confusão; n é o número total de amostras e c é o número total de classes. As taxas do Índice Kappa têm valores entre 0 e 1, com a classificação descrita na Tabela 3-3, de acordo com o índice (THOMAS et al., 2012).

Tabela 3-3 Índice Kappa e a performance correspondente ao classificador

Índice Kappa	Performance
$k < 0$	Muito ruim
$0 < k \leq 0.2$	Ruim
$0.2 < k \leq 0.4$	Regular
$0.4 < k \leq 0.6$	Bom
$0.6 < k \leq 0.8$	Muito Bom
$0.8 < k \leq 1.0$	Excelente

2) *Taxa de Classificação*: O método proposto é avaliado através da taxa média de classificação usando validação cruzada com 10 *folds*. As amostras são divididas em três classes de lesão: benignas, malignas e normais.

3.8 Trabalhos Relacionados

O uso de Redes Neurais Artificiais (RNAs) na detecção de câncer de mama é muito difundido, porque não é necessário fornecer um algoritmo específico sobre como identificar a doença (BELCIUG et al., 2010). As redes *Multiple Layer Perceptron* (MLP) geralmente conseguem um bom desempenho em termos de taxa de reconhecimento na classificação de lesões em mamografias. Infelizmente, há limites quando se utiliza MLPs em tarefas de classificação:

- Em primeiro lugar, não existe uma relação teórica entre a estrutura de uma rede MLP (ex: número de camadas escondidas e número de neurônios por camada) e a tarefa de classificação;
- O algoritmo *Backpropagation* é muito lento em superfícies complexas;
- Não é raro o algoritmo convergir para mínimos locais. Mínimos locais são pontos na superfície de erro que apresentam uma solução estável, embora não seja a saída correta. Algumas técnicas são utilizadas tanto para acelerar o algoritmo *backpropagation* quanto para reduzir a incidência de mínimos locais:
 - Utilizar taxa de aprendizado decrescente;

- Adicionar nós intermediários;
- Utilizar um termo *momentum*;
- Adicionar ruídos aos dados;
- *Overfitting*: Depois de um certo ponto do treinamento, a rede piora ao invés de melhorar. A rede passa a memorizar os padrões de treinamento, incluindo suas peculiaridades e assim, piorando a generalização.

No trabalho de(CORDEIRO et al., 2012) foi proposto o uso da rede ELM (HUANG et al., 2004) tendo como tarefa segmentar imagens de mamografia. *Extreme Learning Machines* são redes com tempo de aprendizagem rápido e de alta precisão, o que faz com que ELMs sejam adequadas para resolverem problemas de segmentação. Os resultados da rede ELM são comparados com a rede MLP (JEATRAKUL et al., 2009).

Em (BELCIUG et al., 2010) foram comparados dois métodos da aprendizagem em redes neurais: **supervisionado**: redes neurais *feed forward* clássicas: *Multiple Layer Perceptron* (MLP), *Radial Basis Function* (RBF) e *Probabilistic Neural Networks* (PNN); e **não-supervisionado**: *Self-Organizing Feature Maps* (SOFM), ou *Kohonen Map*; tendo como principal objetivo avaliar os seus desempenhos na classificação de câncer de mama em imagens de mamografias.

Em (ZRIBI et al., 2012) experimentos foram realizados usando *Self-Organizing Maps of Kohonen* (SOM) para classificar lesões de mama (benignas e malignas).

Em (SINGH et al., 2011) os autores propuseram um método de detecção automático, classificação, *scoring* e *grading* de câncer de mama utilizando *Feed Forward Backpropagation Neural Network*. Foram calculadas a sensibilidade, precisão e especificidade que se encontravam iguais. Através dos resultados obtidos, concluíram que as redes *Feed Forward Backpropagation* apresentaram resultados satisfatórios quanto ao tempo de treinamento e taxa de classificação.

Em (JANGHEL et al., 2010) foram implementados quatro modelos de redes neurais que foram: uma rede MLP com algoritmo *Backpropagation*, uma *Radial Basis Function Networks* (RBFN), um *Learning Vector Quantization* (LVQ) e uma *Competitive Learning Network* (CL). Após

os resultados experimentais, concluíram que o LVQ apresentou melhores resultados quanto à classificação de tumores de câncer de mama.

Tabela 3-4 Principais Redes Neurais utilizadas no estado da arte para a detecção de câncer de mama em imagens de mamografia.

Rede Neural	Desvantagens
MLP	• <i>Backpropagation</i> é muito lento em superfícies complexas; • Mínimos locais: solução estável que não fornece saída correta; • <i>Overfitting</i>: Depois de um certo ponto do treinamento, a rede piora ao invés de melhorar. Memoriza padrões de treinamento, incluindo suas peculiaridades (piora a generalização).
RBF	• Geralmente precisam de ao menos dez vezes mais dados de treinamento para atingir a acuracidade das redes de retropropagação; • Exige muitos neurônios na camada oculta; • Processamento em tempo real lento. Isto se deve ao fato da utilização das funções radiais de base da camada escondida.
PNN	• Todos os vetores do treinamento devem ser armazenados e usados para classificar os vetores novos, requerendo uma memória grande; • O teste pode ser mais lento do que para o <i>backpropagation</i> para a execução do software porque o momento da computação para uma classificação é proporcional ao tamanho do treinamento ajustado.
SOM	• Definição de vizinhança na fase de ordenação; • Definição das fronteiras entre um e outro <i>cluster</i>; • Identificação da geometria ótima de mapa; • Pode falhar em representar estatisticamente o espaço de atributos, assinalando regiões pequenas do mapa para classes com o maior número de amostras de treinamento.
LVQ	• Precisa ser capaz de gerar medidas úteis à distância para todos os atributos (a distância euclidiana é normalmente utilizada para atributos numéricos); • A acurácia do modelo é altamente dependente da inicialização do modelo, bem como os parâmetros de aprendizagem utilizados (taxa de aprendizagem, iterações de treinamento, etc.); • A acurácia também é dependente da distribuição das classes no conjunto de treinamento, uma boa distribuição de amostras é necessária para a construção de modelos úteis; • É difícil determinar um bom número de vetores da tabela de codificação para um determinado problema.
CL	• Tempo de treinamento longo; • Possibilidade de encontrar e ficar preso a soluções sub-ótimas; • Instabilidade diante mudanças no ambiente.

3.9 Considerações do Capítulo

Neste capítulo foram apresentados os principais trabalhos do estado da arte no contexto de detecção de lesões em imagens de mamogramas. Também foram colocados em evidência os trabalhos que serviram como inspiração para o desenvolvimento do modelo proposto.

4 METODOLOGIA PROPOSTA

Este capítulo tem como objetivo apresentar a metodologia utilizada ao longo do trabalho de pesquisa. Também, é descrita em maiores detalhes, a base de dados IRMA, que foi utilizada durante os experimentos com o método proposto.

4.1 Aproximações Morfológicas em Imagens Binárias

A aproximação aritmética dos operadores morfológicos emprega operações aritméticas e também álgebra booleana. As operações de conjuntos empregadas pelas aproximações morfológicas propostas, podem ser representadas através da álgebra booleana, vista na Tabela 4-1, onde $f_1: S \rightarrow [0,1]$ e $f_2: S \rightarrow [0,1]$, onde S está em formato de matriz bidimensional.

Tabela 4-1 Operações de mínimo e máximo associadas às propriedades da álgebra de Boole, representados em (a) e (b), respectivamente.

(a)

f_1	f_2	$\min(f_1, f_2) = \wedge(f_1, f_2)$
0	0	0
0	1	0
1	0	0
1	1	1

(b)

f_1	f_2	$\max(f_1, f_2) = \vee(f_1, f_2)$
0	0	0
0	1	1
1	0	1
1	1	1

A operação booleana AND, graficamente representada por $\wedge$, é apresentada na Tabela 4-1 (a). Representando a operação de *mínimo*. A equação 4-1 exhibe a aproximação da operação $\wedge$ através dos operadores aritméticos. O operador "." representa uma multiplicação.

$$\min\{f_1(u), f_2(u)\} = \wedge\{f_1(u), f_2(u)\} = f_1(u) \cdot f_2(u), \forall u \in S. \quad (4-1)$$

Ao se extrapolar para n -funções booleanas, a operação aproximada da teoria de conjuntos de mínimos se dá entre o produtório $\prod$ dessas n -funções, como exhibe a Equação (4-2), onde $f_i: S \rightarrow [0,1]$.

$$\begin{aligned} \min\{f_1(u), f_2(u), f_3(u), \dots, f_n(u)\} &= \wedge\{f_1(u), f_2(u), f_3(u), \dots, f_n(u)\} = \\ f_1(u) \cdot f_2(u) \cdot f_3(u) \cdot \dots \cdot f_n(u) &= \prod_{i=1}^n f_i(u), \forall u \in S. \end{aligned} \quad (4-2)$$

Analisando a Tabela 4-1 (b), é possível observar que a operação booleana OR, representada de forma gráfica por $\vee$, equivale à operação de máximo, como pode ser observado na equação 4-3. O último termo da Equação está de acordo com a propriedade da álgebra de Boole chamada de Teorema da Involução: $A = \overline{\overline{A}}$, $A \rightarrow [0,1]$.

$$\max \{f_1(u), f_2(u)\} = \vee \{f_1(u), f_2(u)\} = \overline{\overline{f_1(u)} \wedge \overline{f_2(u)}}, \forall u \in S. \quad (4-3)$$

Entre a equação 4-4 e a equação 4-7, a operação de conjunto *máximo* é desenvolvida sempre a partir dos desdobramentos da Equação imediatamente anterior. A equação 4-4 segue o Teorema de De Morgan, onde $\overline{\overline{A} \vee \overline{B}} = \overline{\overline{A}} \wedge \overline{\overline{B}}, \{A, B\} \rightarrow [0,1]$.

$$\max\{f_1(u), f_2(u)\} = \overline{\overline{f_1(u)} \wedge \overline{f_2(u)}}, \forall u \in S \quad (4-4)$$

A equação 4-5 utiliza o complemento padrão das operações de conjunto: $\bar{A} = 1 - A, A \rightarrow [0,1]$ (KLIR, et al., 1995).

$$\max\{f_1(u), f_2(u)\} = 1 - (\bar{f}_1(u) \wedge \bar{f}_2(u)), \forall u \in S \quad (4-5)$$

A equação 4-6 emprega novamente o complemento padrão das operações de conjunto, dessa vez para as funções $\bar{f}_1(u)$ e $\bar{f}_2(u)$.

$$\max \{f_1(u), f_2(u)\} = 1 - (1 - f_1(u)) \wedge (1 - f_2(u)), \forall u \in S \quad (4-6)$$

A equação 4-7 corresponde à aproximação aritmética da operação booleana de $\wedge$, como foi demonstrado na equação 4-1.

$$\max \{f_1(u), f_2(u)\} = 1 - (1 - f_1(u)) \cdot (1 - f_2(u)), \forall u \in S. \quad (4-7)$$

Logo, ao se extrapolar para *n-funções* booleanas, a operação de *máximo* está de acordo com a equação 4-8, $f_i: S \rightarrow [0,1]$.

$$\max\{f_1(u), f_2(u), \dots, f_n(u)\} = 1 - \prod_{i=1}^n 1 - f_i(u), \forall u \in S. \quad (4-8)$$

Assim, as operações da teoria de conjunto *máximo* e *mínimo* podem ser implementadas, de forma aproximada, através de operadores aritméticos. Portanto, é possível gerar aproximações morfológicas, tomando como base noções de conjuntos, por meio de operadores aritméticos. Depois dos cálculos das aproximações de *mínimo* e *máximo* implementadas, de modo aproximado, através de operações aritméticas, existe o objetivo de modificar as formulações clássicas de *mínimo* e *máximo* presentes nas operações de erosão e dilatação vistas na equação 4-9 e equação 4-10, respectivamente.

$$\varepsilon_g(f)(u) = \bigcap_{v \in S} f(v) \vee \bar{g}(u - v) \quad (4-9)$$

$$\delta_g(f)(u) = \bigcup_{v \in S} f(v) \wedge g(u - v), \quad (4-10)$$

Na equação 4-11, a aproximação da erosão $\tilde{\epsilon}_g$ inicialmente transforma *aerosão clássica*, exposta na equação 4-9. A operação de *máximo* $\vee$, entre f e $\bar{g}$, passa a ser implementada, de forma aproximada, através de operadores aritméticos, conforme equação 4-7.

$$\tilde{\epsilon}_g(f)(u) = \bigcap_{v \in S} 1 - (1 - f(v)) \cdot (1 - \bar{g}(u - v)). \quad (4-11)$$

A equação 4-12 utiliza o complemento padrão das operações de conjunto para o termo $\bar{g}$ (KLIR, et al., 1995).

$$\tilde{\epsilon}_g(f)(u) = \bigcap_{v \in S} 1 - (1 - f(v)) \cdot (1 - (1 - g(u - v))). \quad (4-12)$$

A equação 4-13 simplifica a equação apresentada anteriormente.

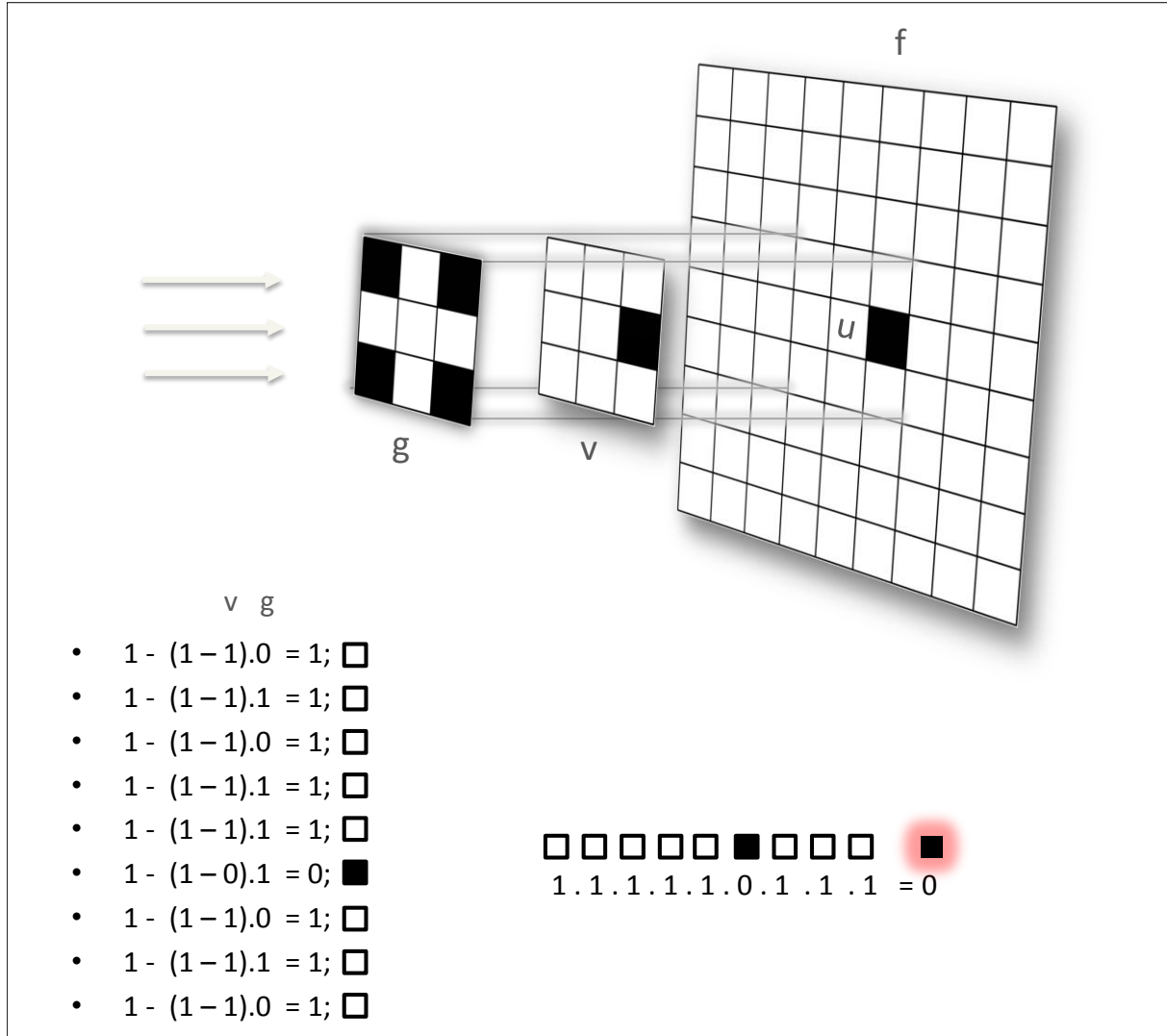
$$\tilde{\epsilon}_g(f)(u) = \bigcap_{v \in S} 1 - (1 - f(v)) \cdot g(u - v) \quad (4-13)$$

Na equação 4-14 a aproximação da erosão $\tilde{\epsilon}_g$ utiliza a equação 4-2 para a aproximação da operação de *mínimo* $\cap$, através de operações aritméticas.

$$\tilde{\epsilon}_g(f)(u) = \prod_{v \in S} 1 - (1 - f(v)) \cdot g(u - v) \quad (4-14)$$

A Figura 4-1 descreve a atuação da aproximação da erosão, exposta na equação 4-14. A princípio, existe o cálculo da expressão $1 - (1 - f(v)) \cdot g(u - v)$ entre a região v da imagem original f abrangida (casada) por g . Os valores 1's estão associados ao branco absoluto e 0's ao preto absoluto. Posteriormente, ocorre o cálculo do produtório $\prod$ entre os n -resultados da expressão. No exemplo apresentado na Figura 4 1, o valor $\tilde{\epsilon}_g(f)(u)$, na posição u da imagem erodida, recebe o valor 0 associado ao preto absoluto.

Figura 4-1 Exemplo do uso da aproximação proposta de erosão em uma imagem binária.



Entre a equação 4-15 e a equação 4-16, é originada a aproximação da *dilatação* $\tilde{\delta}_g$. A operação clássica de dilatação, exibida na equação 4-10, é alterada. A operação de *mínimo* $\wedge$, entre f e g , passa a ser implementada, de forma aproximada, por meio de operadores aritméticos, conforme a equação 4-1.

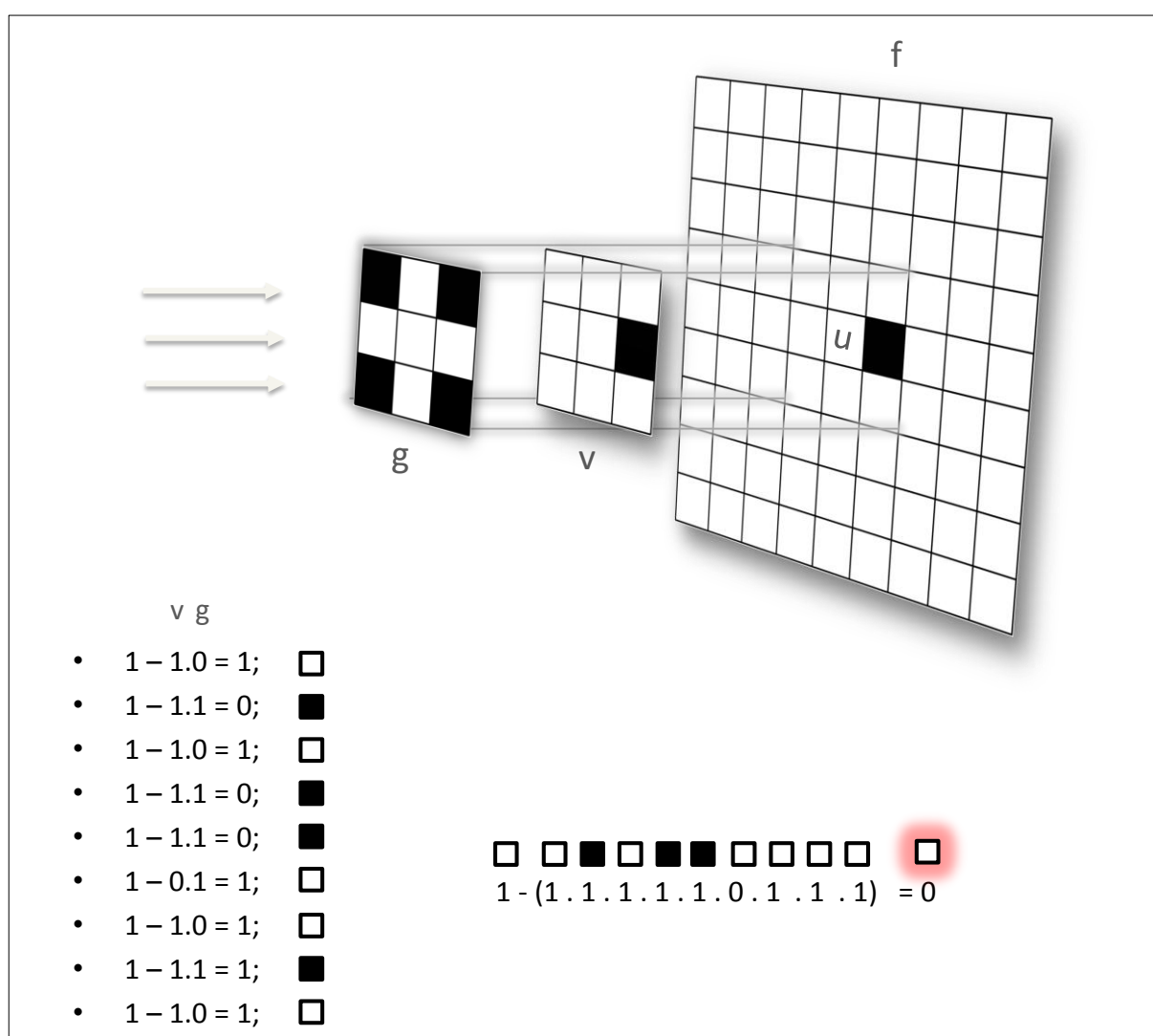
$$\tilde{\delta}_g(f)(u) = \bigcup_{v \in S} f(v).g(u - v) \quad (4-15)$$

Na equação 4-16, a aproximação da *dilatação* $\tilde{\delta}_g$ emprega a equação 4-8 a aproximação da operação de *máximo* $\cup$ com operações aritméticas.

$$\tilde{\delta}_g(f)(u) = 1 - \prod_{v \in S} 1 - f(v).g(u - v) \quad (4-16)$$

A Figura 4-2 mostra a representação da operação de dilatação, exposta na equação 4-16. Primeiramente, é realizado o cálculo da expressão $1 - f(v) \cdot g(u - v)$ entre a região v da imagem original f e a máscara g . Os valores 1's estão associados ao branco absoluto e 0's ao preto absoluto. Em seguida, é realizado o cálculo do produto $\prod$ entre os n -resultados da expressão. Finalmente, o valor 1 é subtraído pelo resultado do produto $\prod$. No exemplo apresentado na Figura 4-2, o valor $\tilde{\delta}_g(f)(u)$, na posição u da imagem erodida, recebe o valor 0 associado ao preto absoluto.

Figura 4-2 Exemplo do uso da aproximação proposta de dilatação em uma imagem binária.



Nas imagens binárias, a aproximação da erosão $\tilde{\epsilon}_g$ e a aproximação da dilatação $\tilde{\delta}_g$ produzem resultados iguais à erosão e dilatação clássicas, respectivamente.

4.2 Proposta: Neurônios Morfológicos

Com base nas definições morfológicas de dilatação e erosão, e nas aproximações aritméticas dos operadores morfológicos apresentadas na seção anterior, definimos dois neurônios não lineares: dilatação e erosão. Um ELM cujos nodos escondidos são baseados nesses neurônios pode ser considerado um ELM Morfológico (mELM). Dada a expressão de dilatação morfológica para uma imagem monoespectral $f : S \rightarrow [0, 1]$ e um elemento estrutural $g : S \rightarrow [0, 1]$:

$$\delta_g(f)(u) = \bigvee_{v \in S} f(v) \wedge g(u - v), \forall u \in S. \quad (4-17)$$

podemos definir a operação de um neurônio de dilatação da seguinte forma:

$$K_\delta(x, w) = \bigvee_{i=1}^n (x_i \wedge w_i). \quad (4-18)$$

onde $x, w \in [0, 1]^n$.

Da mesma forma, tendo em conta a expressão da erosão morfológica para uma imagem monoespectral $f : S \rightarrow [0, 1]$ e um elemento estrutural $g : S \rightarrow [0, 1]$:

$$\epsilon_g(f)(u) = (f \ominus g)(u) := \bigwedge_{v \in S} f(v) \vee \bar{g}(v - u), \quad \forall u \in S. \quad (4-19)$$

podemos definir a operação de um neurônio de erosão da seguinte forma:

$$K_\epsilon(x, w) = \bigwedge_{i=1}^n (x_i \vee \bar{w}_i). \quad (4-20)$$

A fim de obter expressões eficientes computacionalmente, podemos usar a relação de De Morgan e aproximações de Lógica Fuzzy para operações de máximo e mínimo, o que resulta nas seguintes expressões aproximadas (aproximações aritméticas):

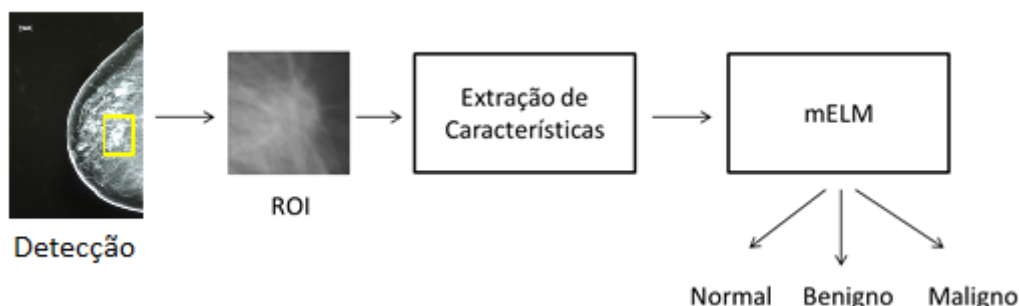
$$K_\delta(x, w) = \overline{\bigwedge_{i=1}^n \bar{x}_i \wedge \bar{w}_i} = 1 - \prod_{i=1}^n (1 - x_i w_i). \quad (4-21)$$

$$K_\epsilon(x, w) = \bigwedge_{i=1}^n (\bar{x}_i \wedge \bar{w}_i) = \prod_{i=1}^n (1 - (1 - x_i) \cdot w_i). \quad (4-22)$$

4.3 Detecção e Classificação de Lesões em Mamografias

O trabalho proposto utiliza a detecção e classificação de lesões em mamografias como estudo de caso para a validação de dois classificadores, o mELM de Dilatação e o mELM de Erosão. A Figura 4-3exibe o diagrama da metodologia proposta em diagrama de blocos.

Figura 4-3Diagrama da metodologia proposta.



Inicialmente, é realizada a **Detecção** da existência, ou não, de alguma lesão na imagem digital de mamografia. Ao detectar a presença de uma lesão, será selecionada a região de interesse na imagem (**ROI**). Através da imagem da região de interesse, dá-se início a fase de **Extração de Características**. Esses dados servirão como atributos de entrada para a tarefa de **Classificação**, realizada pelo **mELM**. As classes são definidas de acordo com os critérios da *American College of Radiology*, expressos na escala (BI-RADS 2003). Finalmente, os casos são classificados em normais (sem lesão), benignos e malignos.

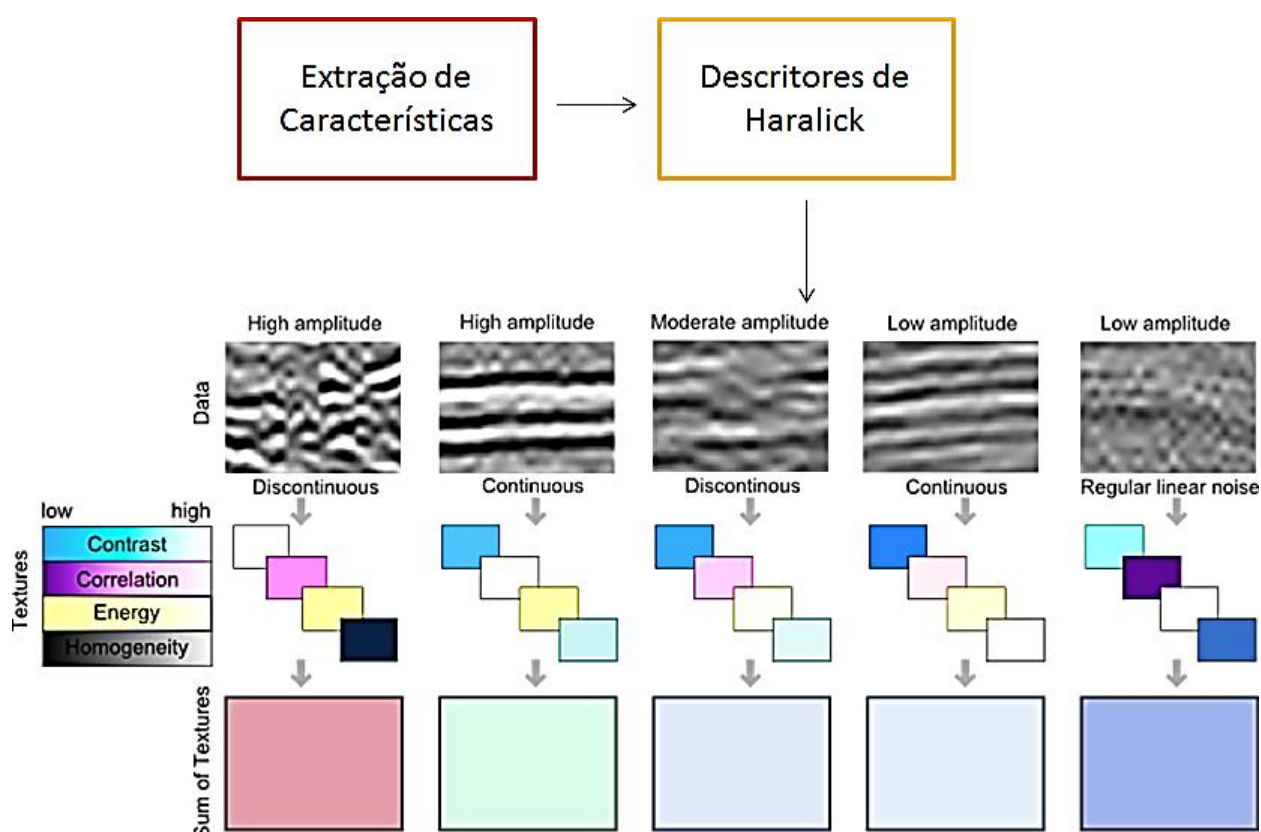
- **Extração de Características da Imagem**

A Figura 4-4mostra em diagrama de blocos, a extração de características da imagem. A região de interesse da imagem de mamografia é primeiramente pré-processada. Ocorre uma equalização baseada nas informações do histograma. O objetivo é tornar os pixels mais claros mais próximos do branco absoluto e os pixels mais escuros mais próximo preto absoluto. Isoladamente, a equalização baseada no histograma não é capaz de segregar a lesão dos demais elementos da imagem. A razão é que os músculos e demais tecidos mamários, chamados tecnicamente de parênquima, geralmente apresentam uma textura análoga à lesão

e podem se encontrar sobrepostos a ela. Logo, a equalização baseada nas informações do histograma, ao atuar no contraste da imagem, acaba por realçar não somente a lesão como também os músculos e parênquima mamário.

Posteriormente à equalização baseada nas informações do histograma, a imagem é decomposta, em multi-resolução, através da Decomposição Fuzzy-Morfológica. Após isso, são extraídos vetores de características das imagens decompostas.

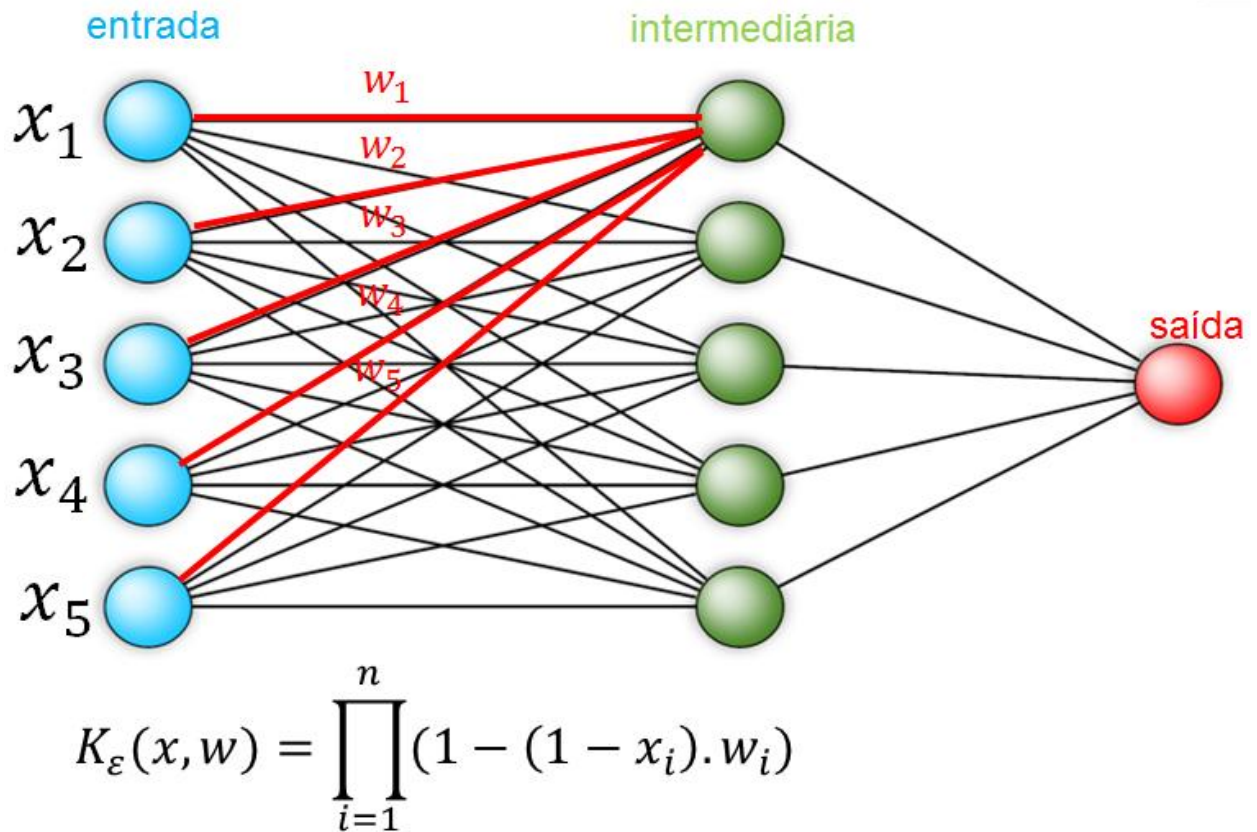
Figura 4-4 Extração de características (descritores de Haralick).



• Classificação utilizando o método proposto (mELM)

A etapa de classificação utiliza como atributo de entrada os vetores de características obtidos na fase de extração de características, de forma concatenada. A Figura 4-5 apresenta o método proposto mELM com neurônios do tipo erosão na camada intermediária.

Figura 4-5 Exemplo do processo de classificação utilizado o método proposto mELM (erosão).



Suponha o seguinte vetor de entrada $x = 1, 1, 0, 1, 1$ e o seguinte conjunto de pesos $w = 1, 0, 1, 0, 1$, como mostra a Figura 4-6. Primeiramente, será realizado o cálculo da expressão $(1 - (1 - x_i) \cdot w_i)$, utilizando o vetor de entrada e os pesos sinápticos entre os nós da camada de entrada e o primeiro nó da camada intermediária. Assim, ocorrerá o cálculo do produto $\prod$ entre os n -resultados da expressão, como pode ser visto na Figura 4-7. No exemplo da Figura 4-5, o valor de saída do primeiro neurônio erosão da camada intermediária, será zero. Como ilustra a Figura 4-8.

Figura 4-6 Apresentando o vetor de entradas ao mELM (erosão) e iniciando o processo de treinamento.

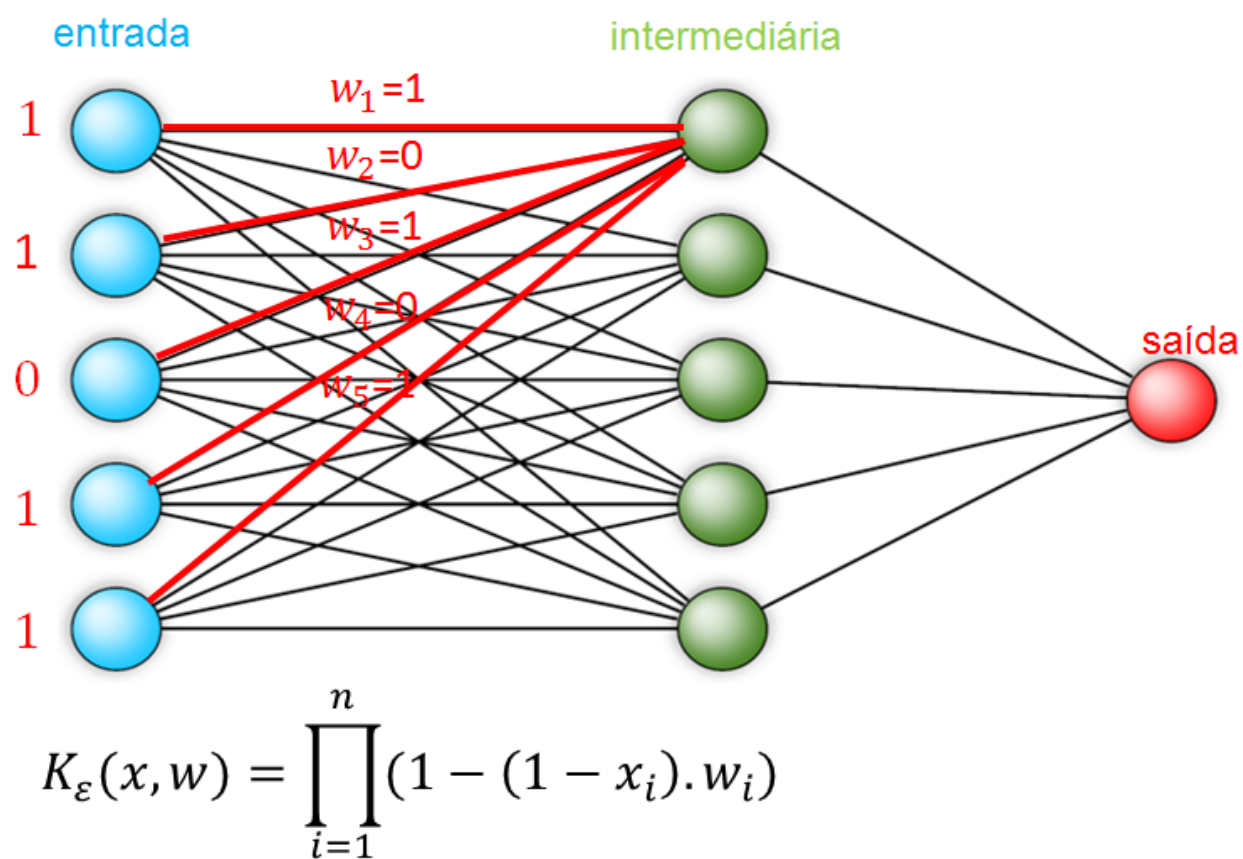


Figura 4-7 Cálculo do produto.

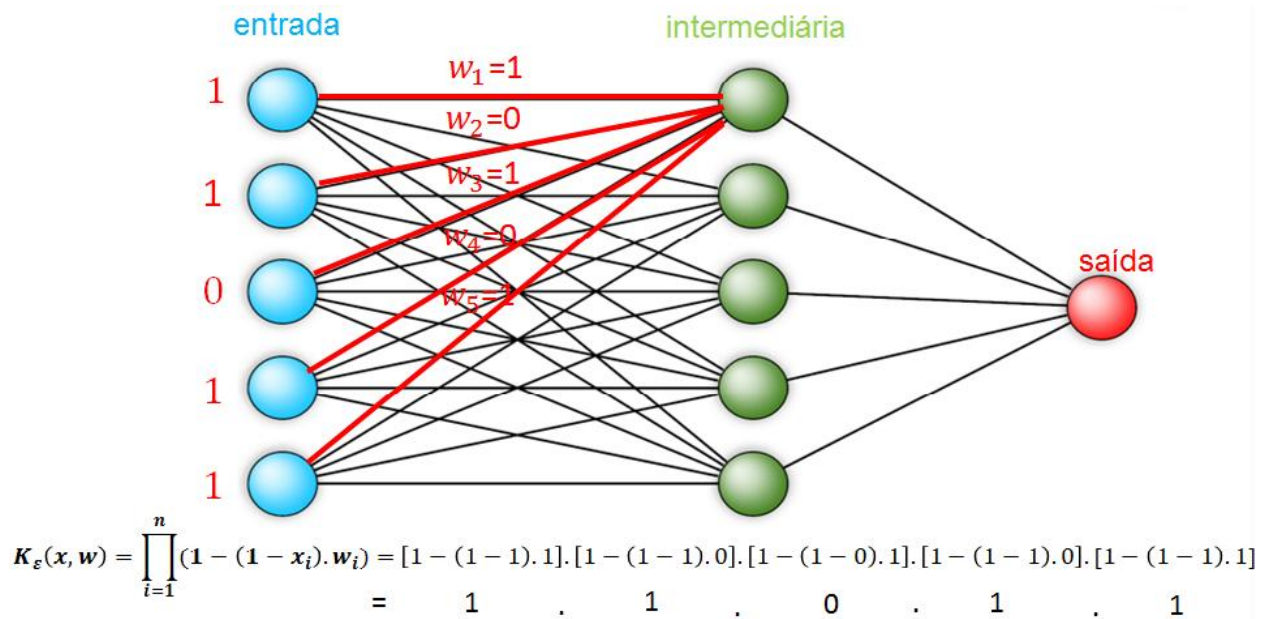
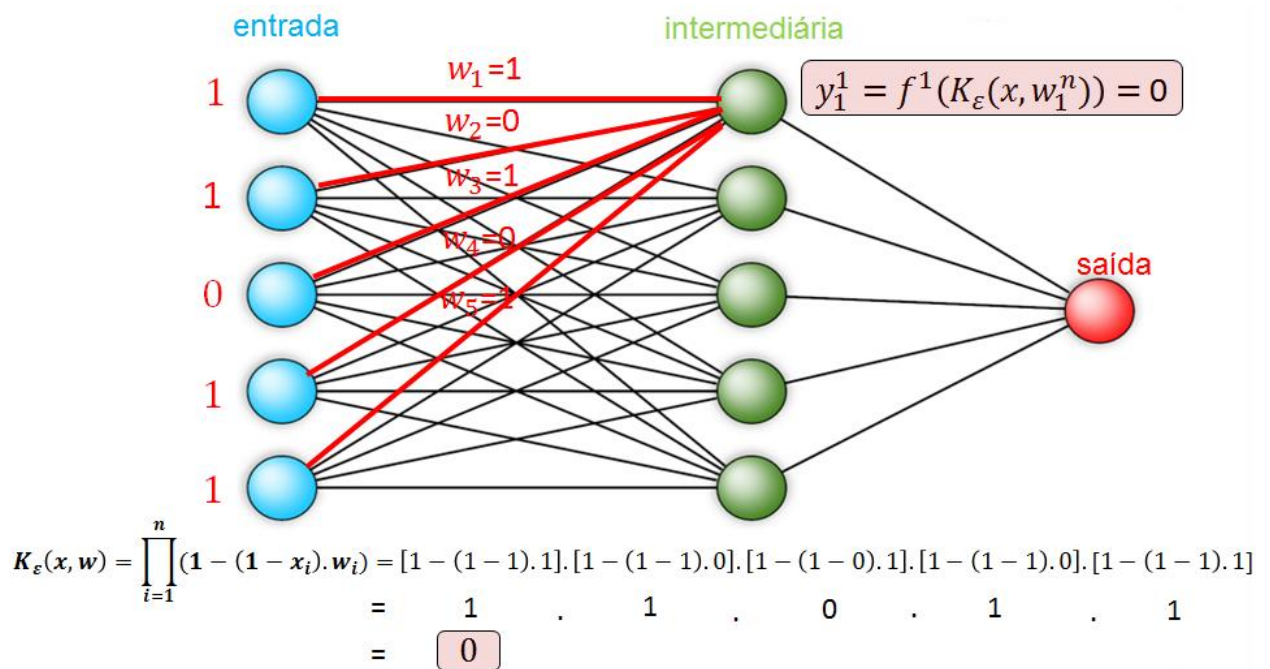


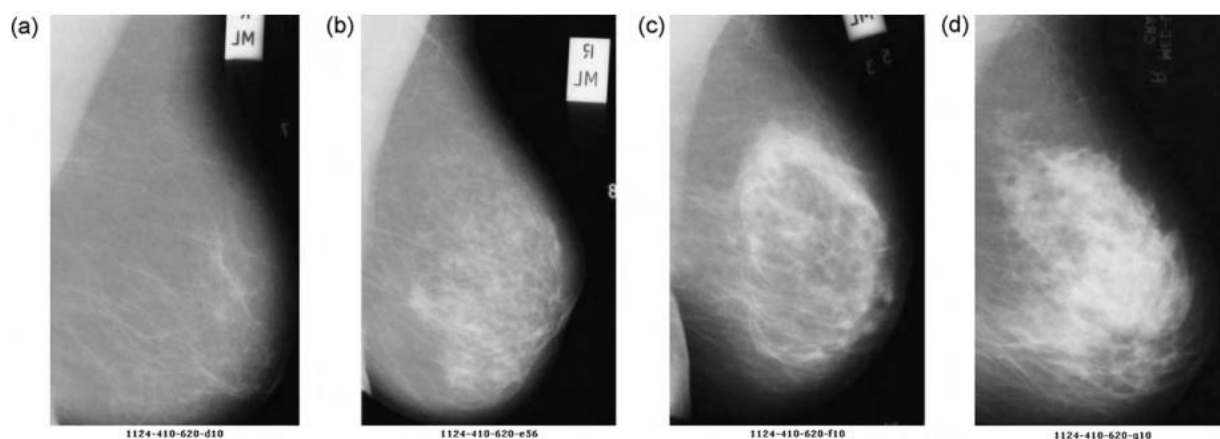
Figura 4-8 Resultado da saída do primeiro neurônio erosão da camada intermediária.



4.4 Base de Dados IRMA

Nesse trabalho foram realizados experimentos utilizando a Base de Dados IRMA (*Image Retrieval in Medical Applications - Redistribuidor de Imagens em Aplicações Médicas*) (DESERNO et al., 2011) (DE OLIVEIRA et al., 2010) (DESERNO et al., 2012), que foi desenvolvida a partir de um projeto realizado pela Universidade de Aachen (RWTH Aachen), na Alemanha, que reúne imagens de outras bases, disponíveis para consulta na internet. A base de dados é composta por regiões de interesse das mamografias, que foram classificadas por radiologistas e redimensionadas para 128x128 *pixels*. A base de dados é composta por 2.796 imagens de mamografias de quatro repositórios: 150 imagens da base de dados Mini-MIAS (SUCKLING et al., 1994), 2.576 imagens do DDSM (*Digital Database for Screening Mammography*) (HEATH et al., 2000), 1 da base de dados LLN (*Lawrence Livermore National Laboratory*), e 69 a partir do Departamento de Radiologia da *Aachen University of Technology* (RWTH), Alemanha. As imagens da base de dados IRMA tem quatro tipos de densidade de tecido, que são classificados em quatro tipos, de acordo com a classificação do BI-RADS (D'Orsi, 1998): tecido adiposo (Tipo I), tecido fibroso (Tipo II), tecido heterogeneamente denso (Tipo III) e tecido extremamente denso (Tipo IV); como pode ser visto na Figura 4-9.

Figura 4-9 Mamografias de diferentes tecidos da mama: (a) tecido adiposo, (b) tecido fibroso, (c) tecido heterogeneamente denso e (d) tecido extremamente denso.



O conjunto de dados IRMA, detalha todo o procedimento de digitalização de exame de imagem da mama. Além disso, apresenta a direção do exame. As direções são médio-lateral

e crânio-caudal. Na base de dados IRMA, é especificado se a imagem se refere a mama direita ou a esquerda.

Para este trabalho foram utilizadas todas as imagens do banco de dados do IRMA, considerando todos os tipos de massas, totalizando 2.796 imagens. O método proposto identifica e classifica cada imagem como benigna, maligna ou normal.

Pode ser observada a falta de cuidado ao digitalizar as imagens de mamografia por parte dos especialistas. Em muitos casos, as imagens apresentam alguns elementos que sobrepõem a mama. Entre eles podemos citar: esparadrapo, espátulas e anotações (manuscritas ou digitais). Também podem ser encontradas, ranhuras e marcas de impressões digitais. Assim, o estudo da região mamária ocupada por alguns desses tipos de elementos externos, pode se tornar inviável; como pode ser visto na Figura 4-10. Para contornar essa situação, o IRMA utiliza apenas a região de interesse de cada caso; como pode ser visto na Figura 4-11. A base garante que existe apenas uma única lesão por região de interesse.

A base de dados IRMA foi obtida através de um acordo estabelecido entre a UFPE (Universidade Federal de Pernambuco) e o DPI (Department of Medical Informatics - Universidade de Tecnologia da Aachen - Alemanha). O acordo veta o uso comercial da base, tanto de forma parcial quanto integral.

Figura 4-10Exemplo de falta de cuidado ao digitalizar as imagens de mamografia por parte dos especialistas.

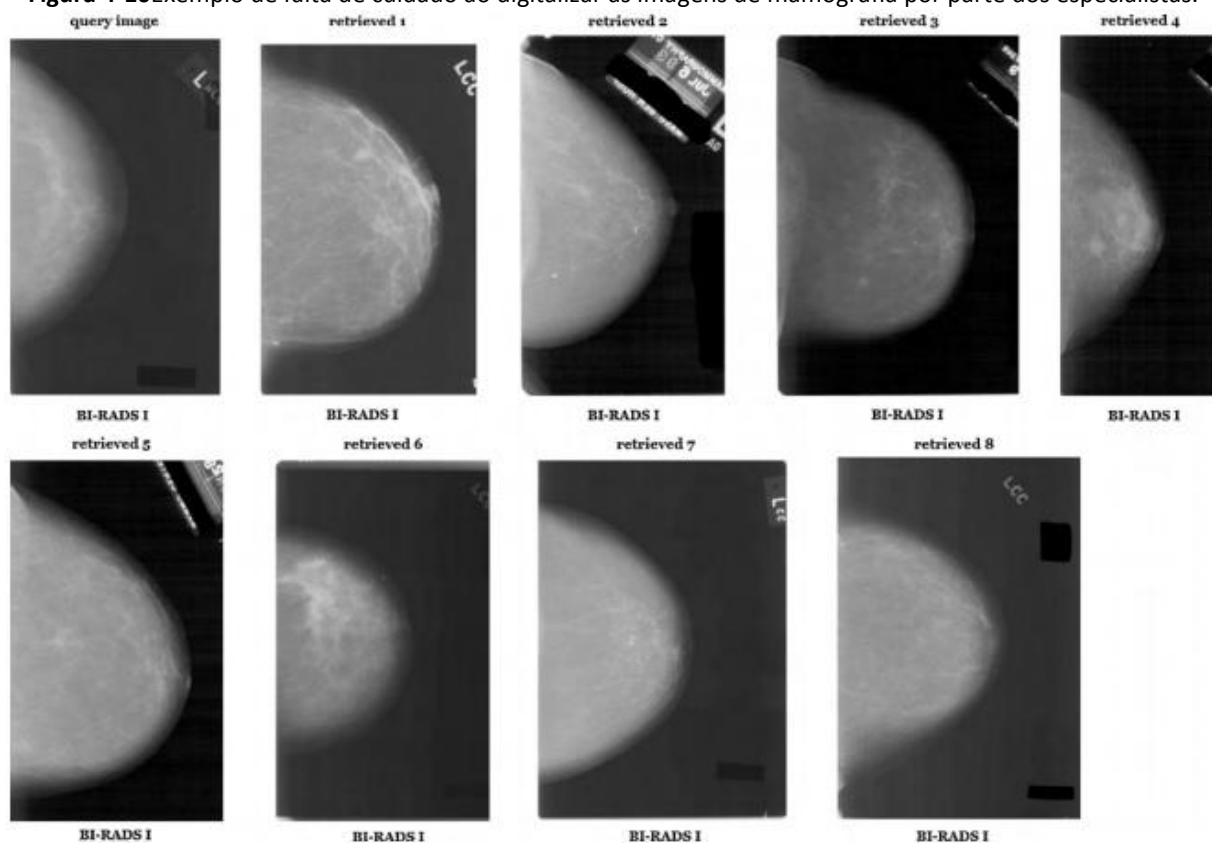
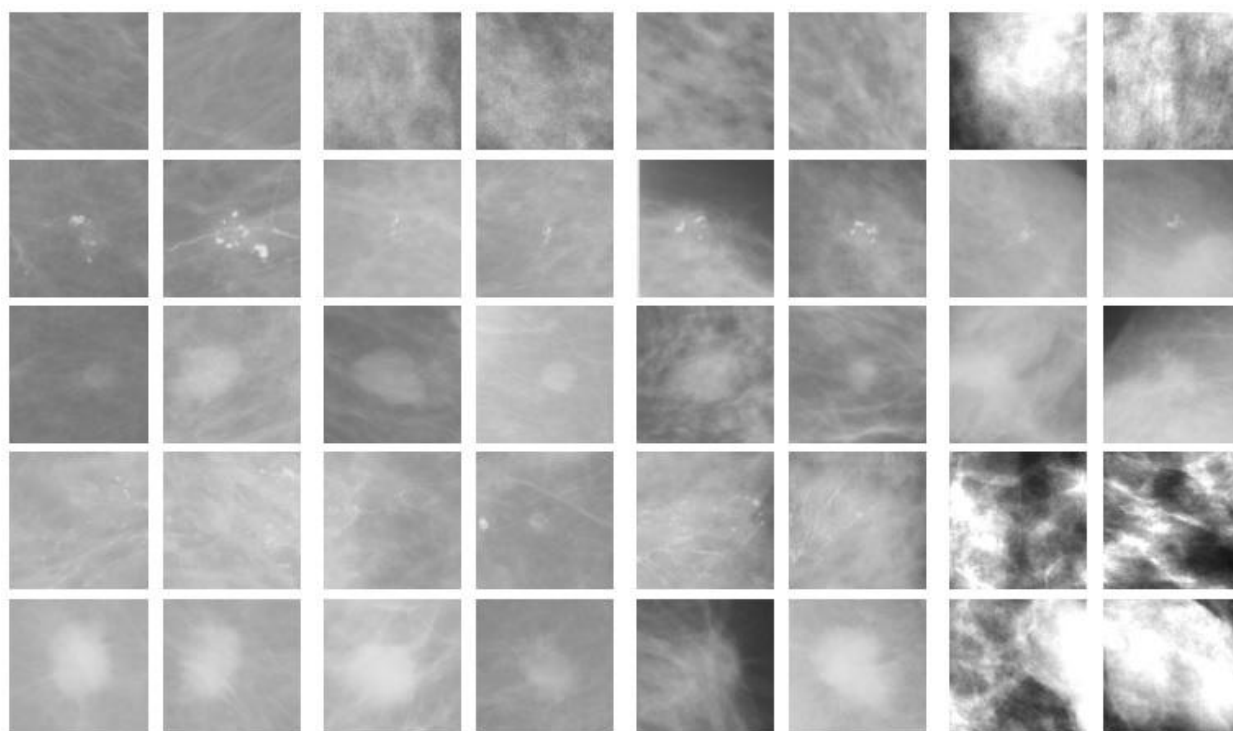


Figura 4-11Exemplo de regiões de interesse de imagens de mamografia da base de dados IRMA.



4.5 Considerações do Capítulo

Neste capítulo foi apresentado o modelo proposto, onde foi descrito um novo modelo denominado Máquinas Morfológicas de Aprendizado Extremo (mELMs), ou seja, ELMs com neurônios morfológicos de dilatação e erosão na camada intermediária.

5 RESULTADOS

Este capítulo tem como objetivo apresentar os resultados obtidos com o modelo proposto (mELM) aplicado ao problema de detecção de lesões em imagens de mamografia. Também, foram realizados experimentos com outros classificadores clássicos comuns no estado da arte. A base de dados IRMA foi utilizada para todos os experimentos realizados.

5.1 Introdução

A seção a seguir apresenta os resultados obtidos com o método proposto, mELM, e com os seguintes classificadores comuns no estado da arte: ELM *kernel* linear, ELM *kernel* RBF, ELM com função de ativação RBF, ELM com função de ativação sigmóide, Redes *Bayesianas*, Redes *Naive Bayes*, Redes MLP, Redes RBF, SVM *kernel* linear, SVM *kernel* polinomial grau 2, SVM *kernel* polinomial grau 3, SVM *kernel* polinomial grau 4, SVM *kernel* polinomial grau 5, SVM *kernel* RBF, kNN ($k = 1$), kNN ($k = 3$), kNN ($k = 5$), kNN ($k = 7$), kNN ($k = 9$), J48, *Random Tree* e *Random Forest*.

Foram utilizadas 2.796 imagens de mamografias da base de dados IRMA. As imagens contidas nessa base de dados é composta por quatro tipos de tecido, de acordo com a classificação do BI-RADS (D'Orsi, 1998): Tecido adiposo (Tipo I), tecido fibroso (Tipo II), tecido heterogeneamente denso (Tipo III) e tecido extremamente denso (Tipo IV). Os resultados foram divididos mostrando a taxa de classificação e o índice Kappa, para cada tipo de tecido e para toda a base de dados. Todos os classificadores foram avaliados utilizando dois diferentes tipos de descritores de imagens, que serviram como entrada para cada técnica. Os dois descritores de imagens utilizados durante os experimentos foram: Atributos de *Haralick* (HA) e *Wavelets* combinadas com atributos de *Haralick* (WA-HA). As redes mELM, ELM, RBF e MLP foram configuradas com 100 neurônios na camada escondida. Nos SVMs foram utilizados os seguintes kernels: Linear, RBF, polinomial de grau 2, polinomial de grau 3, polinomial de grau 4 e polinomial de grau 5. O classificador *Random Forest* foi configurado com 20 árvores. Com o kNN foram realizados experimentos variando o valor de k da seguinte forma: $k = 1$, $k = 3$, $k = 5$, $k = 7$ e $k = 9$. Os experimentos realizados com todos os classificadores identificam cada imagem como benigna, maligna ou normal. Cada análise foi realizada usando validação cruzada com 10 *fold* e o processo foi repetido 30 vezes para obter os resultados de média e desvio padrão.

O índice Kappa foi avaliado para cada classificador, para todos os experimentos realizados com cada tipo específico de tecido e toda a base de dados. O índice Kappa tem um intervalo entre 0 e 1, em que 1 corresponde a um excelente desempenho do classificador e 0 quando muito ruim (Tabela 3-3). O índice Kappa para todos os métodos é mostrado na Figura 4-4. Os resultados da Figura 4-4 validam os resultados de taxa de classificação de Figura 4-1.

5.2 Resultados e Análise Experimental

Nas tabelas e *boxplots* seguintes, há apenas resultados obtidos pelos dois melhores e os dois piores casos, nessa ordem, para cada classificador. Os resultados expandidos, para todas as técnicas, estão disponíveis no Anexo A.

A Tabela 5-2 e a Tabela 5-4 exibem os testes de Wilcoxon não-paramétricos entre os modelos propostos e os principais classificadores do estado da arte. As amostras são relativas às taxas de acerto durante a fase de teste. Ainda sobre a Tabela 5-2 e a Tabela 5-4, *op-value* é o valor escalar [0, 1]. *p* é a probabilidade do teste estatístico apresentar o valor observado ou algo mais extremo. Se *p* apresentar um valor baixo, o resultado observado é estatisticamente relevante.

Tabela 5-1 Resultados obtidos pelos dois melhores e os dois piores classificadores utilizando Atributos de *Haralick* (HA).

<i>Tecido</i>	<i>Descritor de imagem</i>	<i>Classificador</i>	<i>Acerto teste (%)</i>	<i>Kappa</i>
I (Adiposo)	HA	mELM (erosão)	92,13 (0,00)	0,9113 (0,00)
	HA	mELM (dilatação)	91,56 (0,00)	0,9048 (0,00)
	HA	SVM <i>kernel</i> polinomial grau 5	73,61 (4,58)	0,60 (0,07)
	HA	ELM <i>kernel</i> RBF	62,65 (3,39)	0,5779 (0,04)
II (Fibroso)	HA	mELM (dilatação)	91,41 (0,00)	0,9034 (0,00)
	HA	mELM (erosão)	91,13 (0,00)	0,8998 (0,00)
	HA	SVM <i>kernel</i> polinomial grau 5	70,74 (4,44)	0,56 (0,07)
	HA	ELM <i>kernel</i> RBF	39,26 (4,62)	0,3140 (0,05)
III (Heterogeneamente denso)	HA	mELM (erosão)	94,13 (0,00)	0,9337 (0,00)
	HA	ELM função ativação sigmóide	93,35 (0,83)	0,9243 (0,01)
	HA	SVM <i>kernel</i> polinomial grau 5	73,94 (4,55)	0,61 (0,07)
	HA	ELM <i>kernel</i> RBF	47,35 (6,34)	0,4102 (0,06)
IV (Extremamente denso)	HA	mELM (erosão)	91,27 (0,00)	0,9015 (0,00)
	HA	mELM (dilatação)	90,13 (0,00)	0,8873 (0,00)
	HA	Redes <i>Naive Bayes</i>	62,02 (5,66)	0,43 (0,08)
	HA	ELM <i>kernel</i> RBF	47,71	0,4506

			(3,43)	(0,04)
Todos os tecidos	HA	SVM <i>kernel</i> linear	65,27 (2,50)	0,62 (0,03)
	HA	mELM (erosão)	64,66 (0,00)	0,6443 (0,00)
	HA	ELM <i>kernel</i> linear	30,87 (0,53)	0,3079 (0,01)
	HA	ELM <i>kernel</i> RBF	10,51 (0,53)	0,1024 (0,01)

Figura 5-1 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido adiposo utilizando Atributos de *Haralick* (HA).

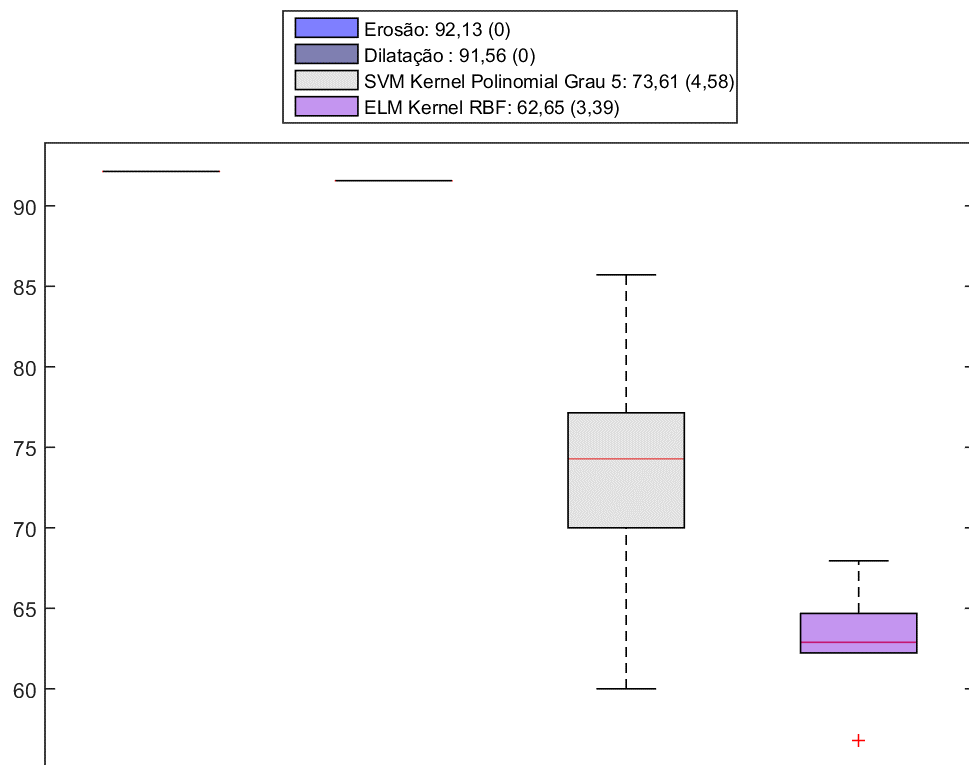


Figura 5-2 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido fibroso utilizando Atributos de *Haralick* (HA).

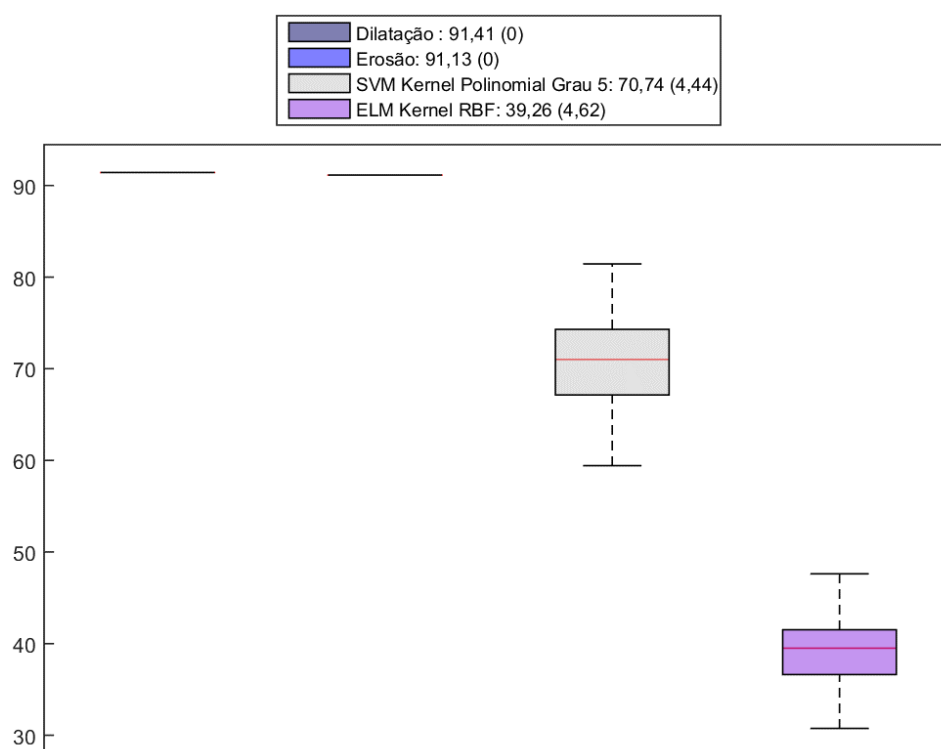


Figura 5-3 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido heterogeneamente denso utilizando Atributos de *Haralick* (HA).

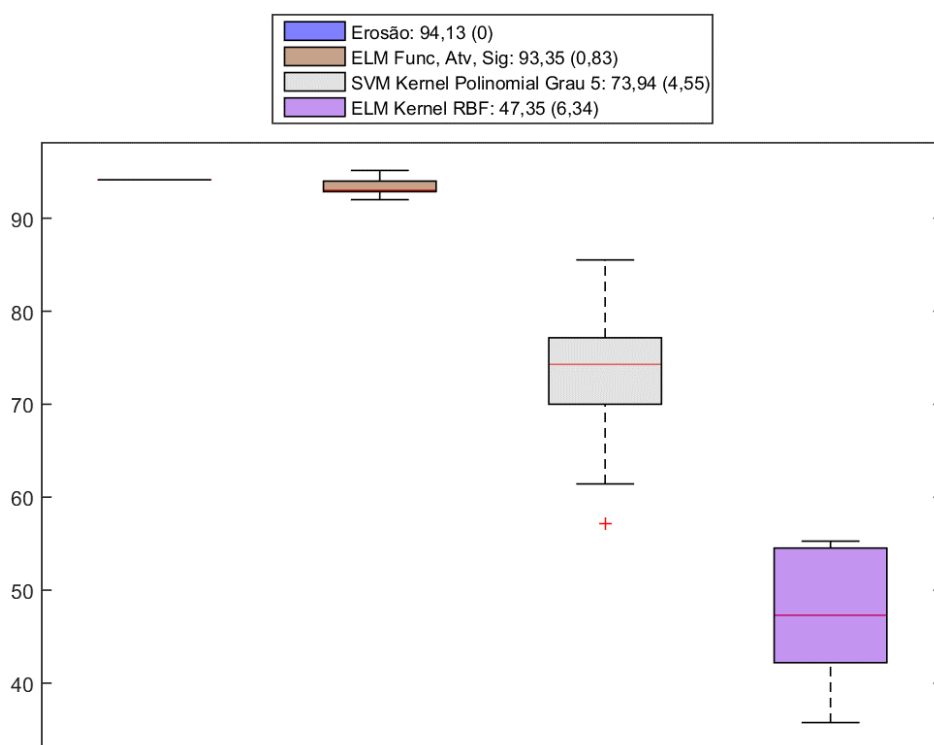


Figura 5-4 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido extremamente denso utilizando Atributos de *Haralick* (HA).

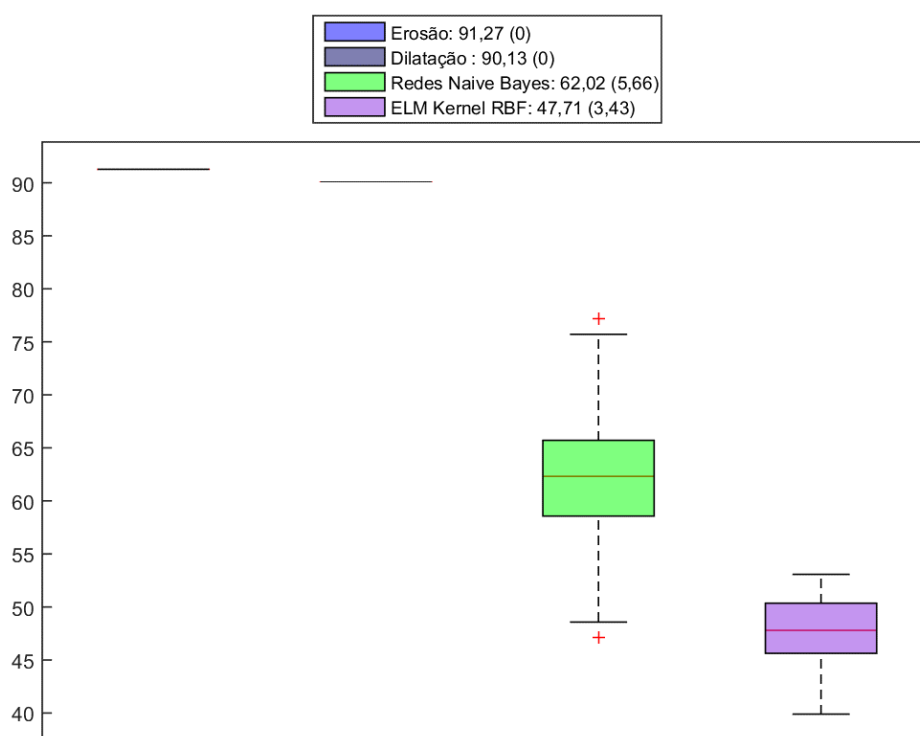


Figura 5-5 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com todos os tipos de tecidos utilizando Atributos de *Haralick* (HA).

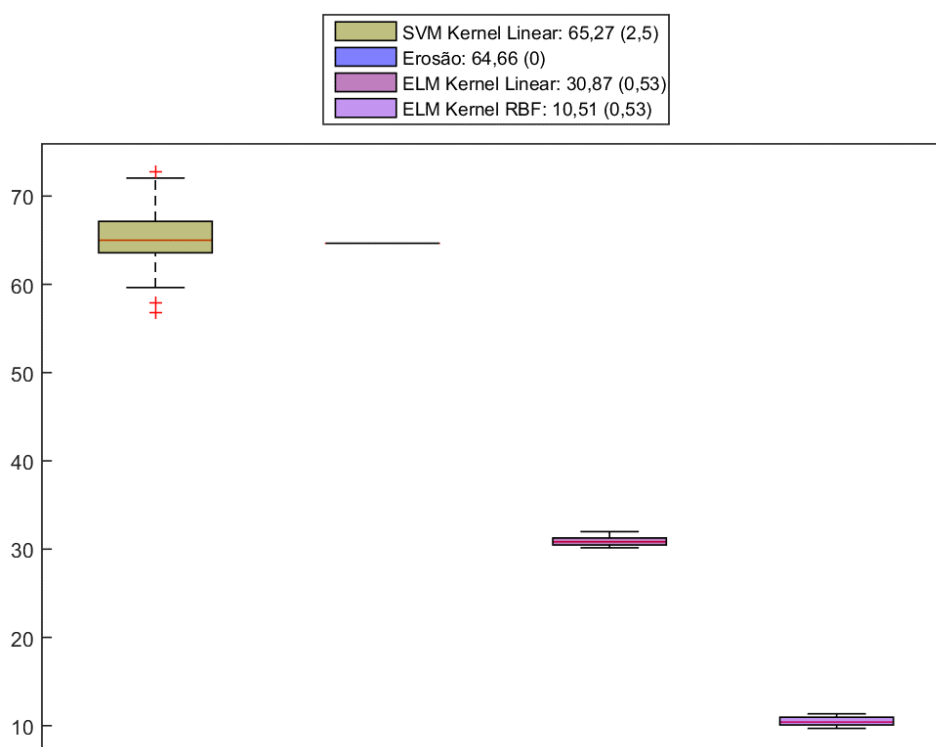


Tabela 5-2 Teste de *Wilcoxon* não-paramétrico comparando os dois melhores e os dois piores classificadores utilizando Atributos de *Haralick*.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	mELM (erosão)vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão)vs SVM (kernel linear)	1	1.20796e-38
	mELM (erosão)vs SVM (kernel polinomial grau 5)	1	8.83377e-114
	mELM (erosão)vs ELM (kernel RBF)	1	6.35092e-114
II (Fibroso)	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação)vs ELM (funç. de ativ. sigmoidal)	1	3.79431e-77
	mELM (dilatação)vs SVM (kernel polinomial grau 5)	1	8.80274e-114
	mELM (dilatação)vs ELM (kernel RBF)	1	6.35092e-114
III (Heterogeneamente denso)	mELM (erosão)vs ELM (funç. de ativ. sigmoidal)	1	9.36891e-47
	mELM (erosão)vs SVM (kernel linear)	1	3.53302e-12
	mELM (erosão)vs SVM (kernel polinomial grau 5)	1	9.02723e-114
	mELM (erosão)vs ELM (kernel RBF)	1	6.35092e-114
IV (Extremamente denso)	ELM M. (erosão)vs mELM (dilatação)	1	2.78101e-132
	ELM M. (erosão)vs SVM (kernel linear)	1	0.023354
	ELM M. (erosão)vs NaiveBayes	1	9.43394e-114
	ELM M. (erosão)vs ELM (kernel RBF)	1	6.35092e-114
Todos os tecidos	SVM (kernel linear)vs ELM M. (erosão)	1	0.00410445
	SVM (kernel linear)vs mELM (dilatação)	1	2.05879e-23
	SVM (kernel linear)vs ELM (kernel linear)	1	7.81526e-100
	SVM (kernel linear)vs ELM (kernel RBF)	1	6.59731e-100

Ao observar a Tabela 5-1, nos resultados realizados com o tecido adiposo, pode ser observado que o modelo proposto, mELM (erosão), apresentou melhor taxa de classificação com o percentual de acerto de 92,13%. O índice kappa para esse método avalia o modelo como excelente. Os piores resultados ficaram por conta dos classificadores: ELM *kernel* RBF e SVM *kernel* polinomial grau 5. Apresentando uma acurácia de 62,65% e 73,61% respectivamente. Ainda sobre os piores resultados, de acordo com o valor kappa, ambas as técnicas foram avaliadas como boas. A Figura 5-1 apresenta o *boxplot* para os experimentos realizados com o tecido adiposo. Nesse gráfico, podemos observar que os modelos propostos apresentaram desvio padrão muito próximo de zero. Nas outras duas técnicas clássicas, podemos observar um maior valor do desvio padrão e a presença de *outlier*, especificamente para o ELM *kernel* RBF. A Tabela 5-2, exibe o teste de *Wilcoxon* não-paramétrico. Ao observar essa tabela, para os resultados obtidos com o tecido adiposo, é possível concluir que os resultados obtidos pelo modelo proposto são amostras estatisticamente distintas daqueles obtidos pelos classificadores clássicos, visto que o teste de hipótese é sempre 1 e o valor *p* é sempre extremamente baixo.

Nos experimentos realizados com o tecido fibroso, na Tabela 5-1, o modelo proposto mELM (dilatação) apresentou a melhor acurácia, tendo um valor de 91,41%. O método proposto recebeu avaliação quanto ao índice kappa como excelente. Os piores

resultados ficaram por conta das técnicas clássicas: ELM *kernel* RBF e SVM *kernel* polinomial grau 5. Com valores de acurácia 39,26% e 70,74% respectivamente. De acordo com índice kappa, as técnicas foram avaliadas como regular e muito boa, respectivamente. A Figura 5-2 apresenta o *boxplot* para os experimentos realizados com o tecido fibroso. No gráfico citado, pode ser observado que os modelos propostos apresentaram desvio padrão muito próximo de zero. Nas outras duas técnicas clássicas, podemos observar um maior valor do desvio padrão. A Tabela 5-2, apresenta o teste de *Wilcoxon* não-paramétrico. Ao observar o teste de *Wilcoxon*, para os resultados obtidos com o tecido fibroso, pode ser concluído que os resultados obtidos pelo modelo proposto são estatisticamente distintos em comparação aos classificadores clássicos, observando que o teste de hipótese é sempre 1 e o valor p é baixo.

Analizando os experimentos feitos com o tecido heterogeneamente denso, presentes na Tabela 5-1, pode ser comprovado que o modelo proposto mELM (erosão) apresentou melhor taxa de classificação, com um valor de 94,13%. O classificador proposto foi avaliado pelo índice kappa como excelente. Quanto aos piores resultados, pode ser notado que as seguintes técnicas clássicas obtiveram os resultados menos favorecidos: ELM *kernel* RBF e SVM *kernel* polinomial grau 5. Com valores de 47,35% e 73,94% respectivamente. De acordo com o valor de kappa, as técnicas foram avaliadas como boa e muito boa, respectivamente. A Figura 5-3 apresenta o *boxplot* para os experimentos realizados com o tecido heterogeneamente denso. No gráfico mencionado, pode ser observado que o modelo proposto apresenta desvio padrão muito próximo de zero. Nas outras três técnicas clássicas, podemos observar um maior valor do desvio padrão e a presença de *outlier*, especificamente para o SVM com função de ativação sigmóide. A Tabela 5-2, apresenta o teste de *Wilcoxon* não-paramétrico. Analisando o teste de *Wilcoxon*, para os resultados obtidos com o tecido heterogeneamente denso, pode ser notado que os resultados alcançados pelo modelo proposto são estatisticamente diferentes em comparação aos classificadores clássicos, visto que o teste de hipótese é sempre 1 e o valor p é baixo.

Observando os experimentos feitos com o tecido extremamente denso, exibidos na Tabela 5-1 podemos comprovar que o modelo proposto mELM (erosão), apresentou melhor resultado quanto a taxa de classificação, com uma acurácia de 91,27%. O modelo proposto foi avaliado pelo índice kappa como excelente. Além disso, podemos observar que o mELM, nos resultados obtidos com o tecido extremamente denso, considerado de difícil análise clínica,

alcançou uma acurácia muito satisfatória. Os resultados menos favorecidos ficaram por conta das seguintes técnicas clássicas: ELM *kernel* RBF e Redes *Naive Bayes*. Com acurácia de 47,71% e 62,02% respectivamente. O valor de kappa avaliou essas duas técnicas como boa e muito boa, respectivamente. A Figura 5-4 exibe o *boxplot* para os experimentos realizados com o tecido extremamente denso. No *boxplot* citado, pode ser notado que os modelos propostos apresentam desvio padrão muito próximo de zero. As outras duas técnicas clássicas, apresentam um desvio padrão com maior valor e a presença de *outliers*, especificamente para as Redes *Naive Bayes*. Na Tabela 5-2, é apresentado o teste de *Wilcoxon* não-paramétrico. Observando a tabela citada, para os resultados com o tecido extremamente denso, pode ser concluído que os resultados obtidos pelo modelo proposto são estatisticamente diferentes em comparação aos classificadores clássicos, visto que o teste de hipótese é sempre 1 e o valor p é baixo.

Nos experimentos realizados com toda a base de dados (Tabela 5-1), isto é, englobando todos os tipos de tecidos, pode ser analisado que o classificador SVM *kernel* linear apresentou o melhor resultado, quanto a acurácia, com um percentual de acerto de 65,27%. O método proposto mELM (erosão) ficou em segundo lugar, em termos de taxa de classificação, com acurácia de 64,66%. O índice kappa classificou tanto o SVM *kernel* linear quanto o método proposto mELM (erosão) como muito bom. Os piores resultados foram alcançados pelas seguintes técnicas: ELM *kernel* RBF e ELM *kernel* linear. Apresentando uma acurácia de 10,51% e 30,87% respectivamente. Ainda sobre os piores resultados, o índice kappa avaliou os classificadores como ruim e regular respectivamente. A Figura 5-5 apresenta o *boxplot* para os experimentos realizados com todos os tipos de tecidos. Nesse gráfico, pode ser observado que o modelo proposto apresenta desvio padrão muito próximo de zero. Nas outras três técnicas clássicas, pode ser notado um maior valor de desvio padrão e a presença de *outliers*, especificamente para a técnica vencedora, o SVM *kernel* linear. Na Tabela 5-2, é exibido o teste de *Wilcoxon* não-paramétrico. Ao analisar essa tabela, para os resultados obtidos com todos os tipos de tecido, é possível observar que os resultados alcançados pelo SVM *kernel* Linear são amostras estatisticamente distintas daqueles obtidos pelos demais classificadores, visto que o teste de hipótese é sempre 1 e o valor p é baixo para maioria dos casos.

Tabela 5-3 Resultados obtidos pelos dois melhores e os dois piores classificadores utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

<i>Tecido</i>	<i>Descritor de imagem</i>	<i>Classificador</i>	<i>Acerto teste (%)</i>	<i>Kappa</i>
I (Adiposo)	WA-HA	mELM (erosão)	93,28 (0,00)	0,9241 (0,00)
	WA-HA	SVM <i>kernel</i> linear	92,29 (3,05)	0,88 (0,05)
	WA-HA	Redes <i>Naive Bayes</i>	61,66 (5,82)	0,42 (0,09)
	WA-HA	ELM <i>kernel</i> RBF	54,24 (1,14)	0,4889 (0,01)
II (Fibroso)	WA-HA	mELM (dilatação)	91,56 (0,00)	0,9050 (0,00)
	WA-HA	SVM <i>kernel</i> linear	91,41 (3,46)	0,87 (0,05)
	WA-HA	Redes <i>Bayesianas</i>	57,62 (4,63)	0,36 (0,07)
	WA-HA	ELM <i>kernel</i> RBF	51,02 (1,81)	0,4561 (0,02)
III (Heterogeneamente denso)	WA-HA	SVM <i>kernel</i> linear	96,57 (2,17)	0,95 (0,03)
	WA-HA	mELM (erosão)	96,42 (0,00)	0,9598 (0,00)
	WA-HA	ELM <i>kernel</i> RBF	64,18 (1,28)	0,5886 (0,01)
	WA-HA	Redes <i>Naive Bayes</i>	62,15 (4,83)	0,43 (0,07)
IV (Extremamente denso)	WA-HA	SVM <i>kernel</i> linear	94,09 (2,68)	0,91 (0,04)
	WA-HA	mELM (dilatação)	93,71 (0,00)	0,9285 (0,00)
	WA-HA	Redes <i>Naive Bayes</i>	57,45 (4,33)	0,36 (0,06)
	WA-HA	ELM <i>kernel</i> RBF	56,79 (0,50)	0,5375 (0,01)
Todos os tecidos	WA-HA	SVM <i>kernel</i> linear	73,79 (2,47)	0,71 (0,03)
	WA-HA	mELM (erosão)	66,42 (0,00)	0,6618 (0,00)
	WA-HA	ELM <i>kernel</i> linear	24,58 (0,21)	0,2429 (0,00)
	WA-HA	ELM <i>kernel</i> RBF	15,79 (0,66)	0,1537 (0,01)

Figura 5-6 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido adiposo utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

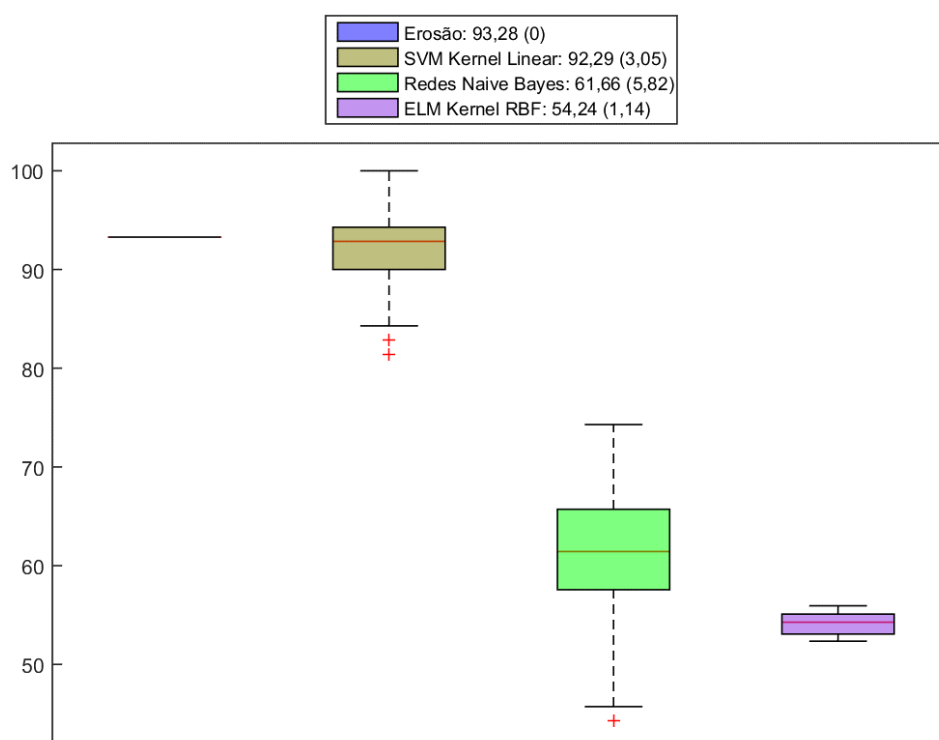


Figura 5-7 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido fibroso utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

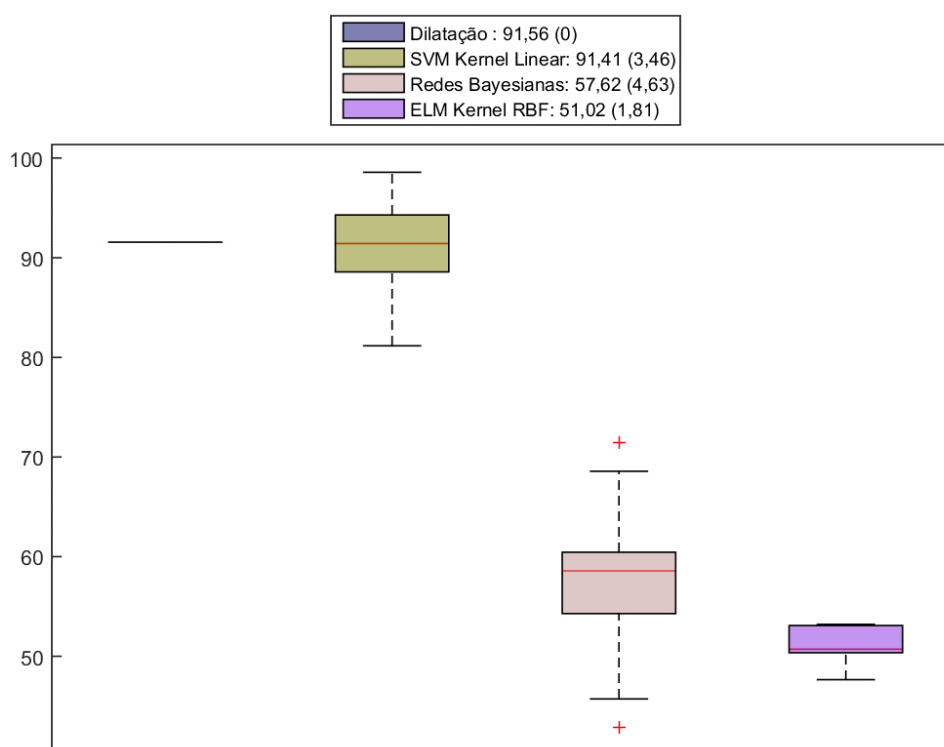


Figura 5-8 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido heterogeneamente denso utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

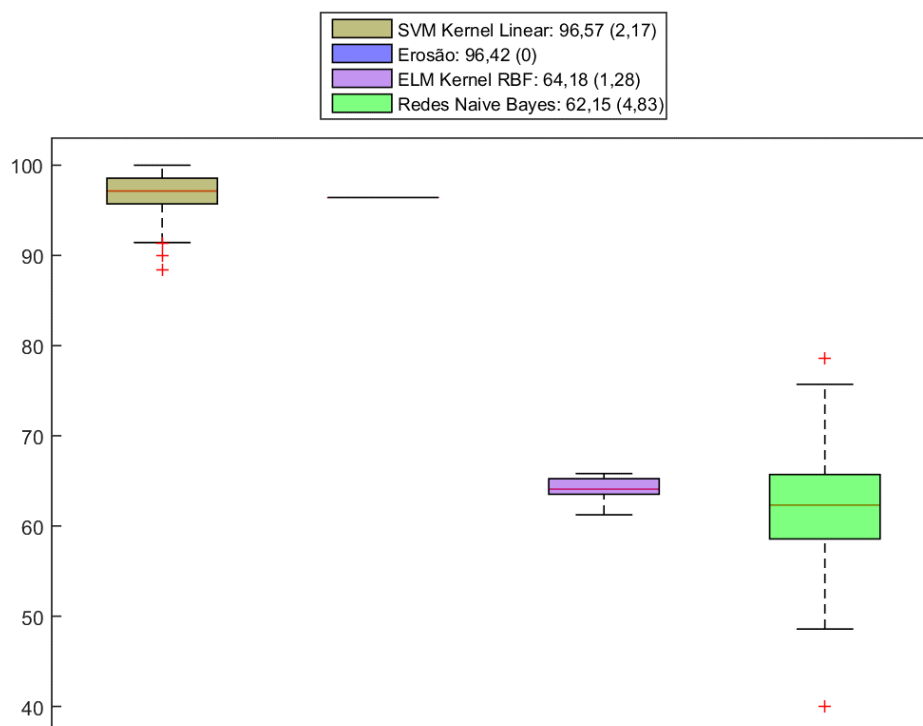


Figura 5-9 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com o tecido extremamente denso utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

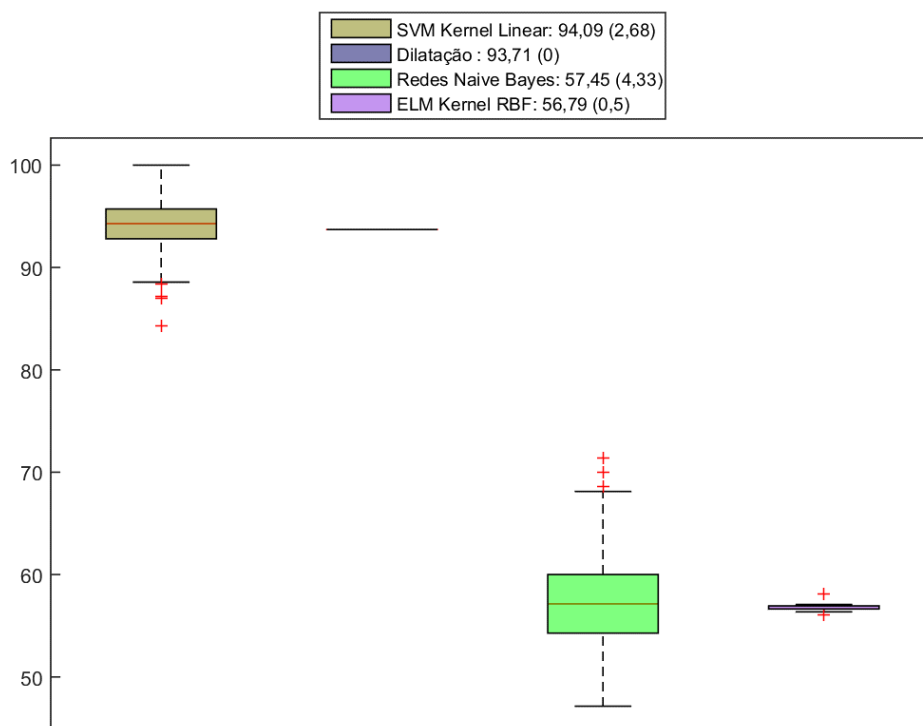


Figura 5-10 Boxplots das taxas de classificações obtidas pelos dois melhores e os dois piores classificadores. Testes realizados com todos os tipos de tecidos utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

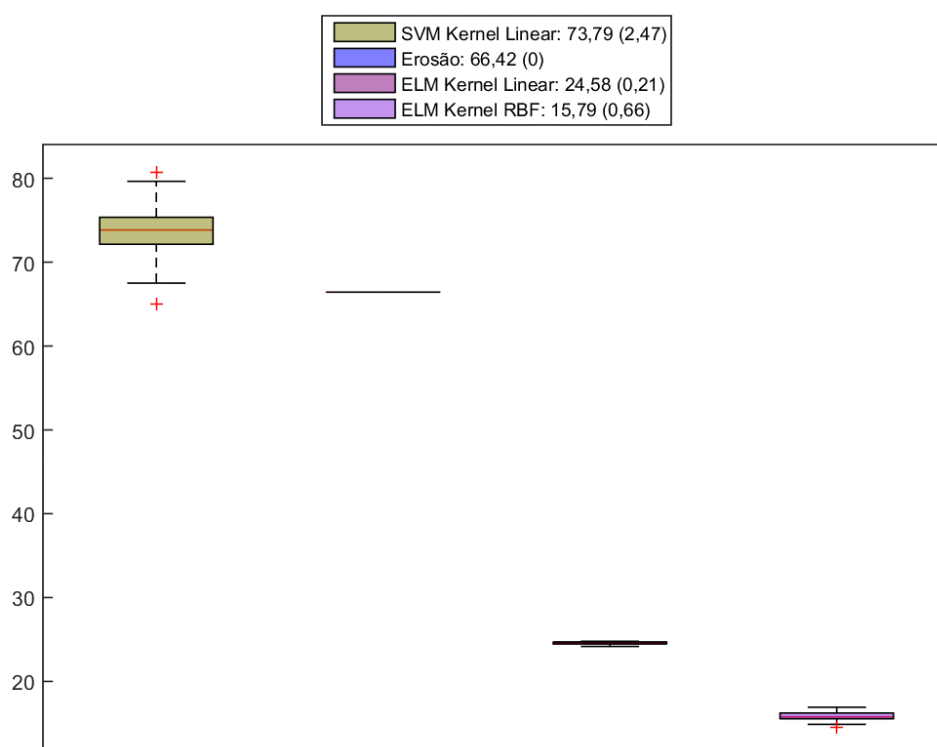


Tabela 5-4 Teste de *Wilcoxon* não-paramétrico comparando os dois melhores e os dois piores classificadores utilizando *Wavelets* combinados com Atributos de *Haralick*.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	mELM (erosão)vs SVM (kernel linear)	1	3.44658e-12
	mELM (erosão)vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão)vs NaiveBayes	1	9.70242e-114
	mELM (erosão)vs ELM (kernel RBF)	1	6.35092e-114
II (Fibroso)	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação)vs ELM (funç. de ativ. sigmoidal)	1	7.12282e-114
	mELM (dilatação)vs Bayesianas	1	8.9078e-114
	mELM (dilatação)vs ELM (kernel RBF)	1	6.35092e-114
III (Heterogeneamente denso)	SVM (kernel linear) vsELM (erosão)	1	0.0490823
	SVM (kernel linear)vs mELM (dilatação)	1	1.28758e-39
	SVM (kernel linear)vs ELM (kernel RBF)	1	2.79694e-100
	SVM (kernel linear)vs NaiveBayes	1	3.66829e-100
IV (Extremamente denso)	SVM (kernel linear)vs mELM (dilatação)	1	5.97642e-07
	SVM (kernel linear) vsELM (erosão)	1	3.9234e-32
	SVM (kernel linear) vs NaiveBayes	1	4.73226e-100
	SVM (kernel linear)vs ELM (kernel RBF)	1	3.77904e-100
Todos os tecidos	SVM (kernel linear)vs ELM M. (erosão)	1	3.41118e-112
	SVM (kernel linear)vs mELM (dilatação)	1	3.41118e-112
	SVM (kernel linear)vs ELM (kernel linear)	1	7.81884e-100
	SVM (kernel linear)vs ELM (kernel RBF)	1	6.60034e-100

Observando a Tabela 5-3, nos resultados realizados com o tecido adiposo, pode ser observado que o modelo proposto, mELM (erosão), apresentou melhor taxa de classificação com o percentual de acerto de 93,28%. O índice kappa para esse método avalia o modelo proposto como excelente. Os piores resultados ficaram por conta dos classificadores: ELM *kernel* RBF e as Redes *Naive Bayes*. Apresentando uma acurácia de 54,24% e 61,66% respectivamente. Ainda sobre os piores resultados, de acordo com o valor kappa, ambas as técnicas foram avaliadas como boas. A Figura 5-6 apresenta o *boxplot* para os experimentos realizados com o tecido adiposo. Nesse gráfico, podemos observar que o modelo proposto apresenta desvio padrão muito próximo de zero. Nas outras três técnicas clássicas, podemos observar um maior valor do desvio padrão e a presença de outlier, especificamente para o SVM *kernel* linear e para as Redes *Naive Bayes*. A Tabela 5-4, exibe o teste de *Wilcoxon* não-paramétrico. Ao observar essa tabela, para os resultados obtidos com o tecido adiposo, é possível concluir que os resultados obtidos pelo modelo proposto são amostras estatisticamente distintas daqueles obtidos pelos classificadores clássicos, visto que o teste de hipótese é sempre 1 e o valor p é sempre baixo.

Nos experimentos realizados com o tecido fibroso, na Tabela 5-3, o modelo proposto mELM (dilatação) apresentou a melhor acurácia, tendo um valor de 91,56%. O método proposto recebeu avaliação quanto ao índice kappa como excelente. Os piores resultados ficaram por conta das técnicas clássicas: ELM *kernel* RBF e Redes *Bayesianas*. Com valores de acurácia 51,02% e 57,62% respectivamente. De acordo com o índice kappa, as técnicas foram avaliadas como boa e regular, respectivamente. A Figura 5-7 apresenta o *boxplot* para os experimentos realizados com o tecido fibroso. No gráfico citado, pode ser observado que o modelo proposto apresenta desvio padrão muito próximo de zero. Nas outras três técnicas clássicas, podemos observar um maior valor do desvio padrão e a presença de *outliers*, especificamente nas Redes *Bayesianas*. A Tabela 5-4, apresenta o teste de *Wilcoxon* não-paramétrico. Ao observar o teste de *Wilcoxon*, para os resultados obtidos com o tecido fibroso, pode ser concluído que os resultados obtidos pelo modelo proposto são estatisticamente diferentes em comparação aos classificadores clássicos, observando que o teste de hipótese é sempre 1 e o valor p é baixo.

Ao analisar os experimentos feitos com o tecido heterogeneamente denso, presentes na Tabela 5-3, pode ser comprovado que o classificador SVM *kernel* linear

apresentou melhor taxa de classificação, com um valor de 96,57%. O classificador proposto mELM (erosão) ficou em segundo lugar, quanto a acurácia, com um percentual de acerto de 96,42. O SVM *kernel* linear e o mELM (erosão) foram avaliados pelo índice kappa como excelentes. Quanto aos piores resultados, pode ser notado que as seguintes técnicas clássicas obtiveram os resultados menos favorecidos: Redes *Naive Bayes* e o ELM *kernel* RBF. Com valores de 62,15% e 64,18% respectivamente. De acordo com o valor de kappa, as técnicas foram avaliadas como boas, respectivamente. A Figura 5-8 apresenta o *boxplot* para os experimentos realizados com o tecido heterogeneamente denso. No gráfico mencionado, pode ser observado que o mELM(erosão) apresenta desvio padrão muito próximo de zero. Nas outras três técnicas clássicas, podemos observar um maior valor do desvio padrão e a presença de *outliers*, especificamente para o SVM *kernel* linear, técnica com maior acurácia, e para as Redes *Naive Bayes*. A Tabela 5-4, apresenta o teste de *Wilcoxon* não-paramétrico. Analisando o teste de *Wilcoxon*, para os resultados obtidos com o tecido heterogeneamente denso, pode ser notado que os resultados alcançados pelo SVM *kernel* linear são estatisticamente diferentes em comparação aos outros classificadores, visto que no teste de hipótese é sempre 1 e o valor *p* é baixo para maioria dos casos.

Analisando os experimentos feitos com o tecido extremamente denso, exibidos na Tabela 5-3 podemos comprovar que o SVM *kernel* linear, apresentou melhor resultado quanto a taxa de classificação, com uma acurácia de 94,09%. O modelo proposto, mELM (dilatação), ficou em segundo lugar, com um percentual de acerto de 93,71%. Ainda sobre o mELM, podemos observar que, o tecido extremamente denso, considerado de difícil análise clínica, o método proposto alcançou uma acurácia muito satisfatória. O SVM *kernel* linear e o mELM (dilatação) foram avaliados pelo índice kappa como excelentes. Os resultados menos favorecidos ficaram por conta das seguintes técnicas clássicas: ELM *kernel* RBF e Redes *Naive Bayes*. Com acurácia de 56,79% e 57,45% respectivamente. O valor de kappa avaliou essas duas técnicas como boa e regular, respectivamente. A Figura 5-9 exibe o *boxplot* para os experimentos realizados com o tecido extremamente denso. No *boxplot* citado, pode ser notado que o modelo proposto apresenta desvio padrão muito próximo de zero. As outras três técnicas clássicas, apresentam um desvio padrão com maior valor e a presença de *outliers*, especificamente para o SVM *kernel* linear (técnica com maior acurácia), Redes *Naive Bayes* e o ELM *kernel* RBF. Na Tabela 5-4, é apresentado o teste de *Wilcoxon* não-paramétrico. Observando a tabela citada, para os resultados com o tecido extremamente denso, pode ser

concluído que os resultados obtidos pelo SVM *kernel* linear são estatisticamente diferentes em comparação aos outros classificadores, visto que no teste de hipótese é sempre 1 e o valor p é baixo.

Nos experimentos realizados com toda a base de dados (Tabela 5-3), isto é, com todos os tipos de tecidos, pode ser analisado que o classificador SVM *kernel* linear apresentou o melhor resultado, quanto a acurácia, com um percentual de acerto de 73,79%. O método proposto mELM (erosão) ficou em segundo lugar, em termos de taxa de classificação, com acurácia de 66,42%. O índice kappa classificou tanto o SVM *kernel* linear quanto o método proposto mELM (erosão) como muito bom. Os piores resultados foram alcançados pelas seguintes técnicas: ELM *kernel* RBF e ELM *kernel* linear. Apresentando uma acurácia de 15,79% e 24,58% respectivamente. Ainda sobre os piores resultados, o índice kappa avaliou os classificadores como ruim e regular respectivamente. A Figura 5-10 apresenta o boxplot para os experimentos realizados com todos os tipos de tecidos. Nesse gráfico, pode ser observado que o modelo proposto e o ELM *kernel* linear apresentam desvio padrão muito próximo de zero. Nas outras duas técnicas clássicas, pode ser notado um maior valor de desvio padrão (principalmente na técnica com maior acurácia, o SVM *kernel* linear) e a presença de outliers, especificamente para a técnica vencedora, o SVM *kernel* linear. Na Tabela 5-2, é exibido o teste de Wilcoxon não-paramétrico. Ao analisar essa tabela, para os resultados obtidos com todos os tipos de tecido, é possível observar que os resultados alcançados pelo SVM *kernel* Linear são amostras estatisticamente distintas daqueles obtidos pelos demais classificadores, visto que o teste de hipótese é sempre 1 e o valor p é muito baixo.

5.3 Considerações do Capítulo

Neste capítulo foram apresentados os resultados experimentais obtidos com os modelos propostos mELM (dilatação) e o mELM (erosão) aplicados a detecção e classificação de lesões em mamografias. O desempenho do método proposto foi comparado com os principais classificadores citados no estado da arte. A partir dos resultados mostrados, conclui-se que o modelo proposto apresentou melhor desempenho, para a maioria dos casos, quando comparados aos classificadores clássicos.

6 CONCLUSÕES

Este capítulo tem como objetivo apresentar algumas considerações finais sobre o trabalho proposto, incluindo as contribuições alcançadas e indicações para trabalhos futuros.

6.1 Considerações Finais

Neste trabalho foram propostas as Máquinas Morfológicas de Aprendizado Extremo (mELM). O método proposto utiliza redes neurais com camadas escondidas de pesos aleatórios, com nodos baseados em operadores morfológicos não lineares de erosão e dilatação. A principal motivação foi construir novos métodos de apoio ao diagnóstico por imagem para detecção e classificação de lesões de mama em imagens mamográficas. O principal objetivo do método proposto foi aliar a capacidade das redes neurais morfológicas, de construir superfícies de decisão complexas com a simplicidade dos algoritmos de treinamento de redes neurais de camadas escondidas com pesos aleatórios, como as máquinas de aprendizado extremo (ELM).

Nos experimentos realizados com a base de imagens de mamografia IRMA, foram utilizadas 2.796 imagens, organizadas em quatro tipos de tecidos: adiposo, fibroso, heterogeneamente denso e extremamente denso. A análise foi realizada para cada tipo de tecido separadamente e para toda a base de dados. Para a análise, foram comparados os resultados obtidos usando classificadores clássicos do estado da arte, neurônio de Dilatação e Erosão, com dois tipos diferentes de entradas para as técnicas: atributos de Haralick e *Wavelets* combinadas com Haralick. Os resultados mostraram que a utilização de neurônios morfológicos melhorou a taxa de classificação para a maioria dos cenários. Em todas as configurações, o método proposto alcançou acurácia média superior a 80% e tendo seu desempenho de classificação considerado como excelente, na maioria dos casos, de acordo com o índice kappa.

Esses resultados mostram que as Máquinas Morfológicas de Aprendizado Extremo são arquiteturas de redes neurais competitivas com relação a abordagens com *kernels* e funções de ativação usuais. Os neurônios de dilatação e de erosão propostos neste trabalho, contudo, podem ser utilizados para construir outras arquiteturas de redes, sejam morfológicas ou híbridas, a serem aplicadas a outros problemas de classificação tão complexos quanto a detecção e classificação de lesões de mama em imagens mamográficas obtidas para diversos tipos de tecido mamário.

6.2 Contribuições

Como principais contribuições do trabalho proposto, podemos citar, um estudo sobre classificação de lesões de mama em imagens de mamografias, focado no uso de ELM com neurônios morfológicos na camada escondida. Com isso, foi proposto o ELM morfológico, denominado mELM, contendo neurônios de Dilatação e de Erosão na camada intermediária.

De acordo com os experimentos realizados, verificamos que, o modelo proposto, mELM, apresenta melhorias significativas em relação à taxa de reconhecimento, comparados com os principais classificadores apresentados pelo estado da arte.

Também podemos mencionar que foram realizados experimentos com quatro tipos de tecidos (adiposo, fibroso, heterogeneamente denso e extremamente denso). Ao analisarmos os resultados obtidos com o tecido extremamente denso, o método proposto, chegou a atingir uma acurácia de 93,71%. O tecido extremamente denso é considerado clinicamente de difícil análise, e o método proposto conseguiu apresentar resultados satisfatórios para esse tipo de tecido.

Os resultados preliminares do trabalho proposto contemplam as seguintes publicações:

Tabela 6-1Publicações

Conferência	Qualis
AZEVEDO, W. ; LIMA, S. M. L. ; FERNANDES, I. M. M. ; ROCHA, A. D. D. ; CORDEIRO, F. R. ; SILVA FILHO, A. G. ; SANTOS, W. P. . Morphological Extreme Learning Machines Applied to Detect and Classify Masses in Mammograms. In: The International Joint Conference on Neural Networks - IJCNN 2015, 2015, Killarney, Irlanda. Proceedings of IJCNN 2015, 2015.	A2
AZEVEDO, W. ; LIMA, S. M. L. ; FERNANDES, I. M. M. ; ROCHA, A. D. D. ; CORDEIRO, F. R. ; SILVA FILHO, A. G. ; SANTOS, W. P. . Fuzzy Morphological Extreme Learning Machines to Detect and Classify Masses in Mammograms. In: IEEE International Conference on Fuzzy Systems - FUZZY-IEEE 2015, 2015, Istanbul. Proceedings of FUZZ-IEEE 2015, 2015.	A2
LIMA, S. M. L. ; AZEVEDO, W. ; CORDEIRO, F. R. ; SILVA FILHO, A. G. ; SANTOS, W. P. . Feature Extraction Employing Fuzzy-Morphological Decomposition for Detection and Classification of Mass on Mammograms. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - EMBC 2015, 2015, Milão. Proceedings of IEEE EMBC 2015, 2015.	B1

Foi submetido um artigo para o periódico Applied Soft Computing, Qualis A1.

6.3 Trabalhos Futuros

Além das contribuições geradas a partir desse trabalho, algumas linhas de pesquisas podem ser seguidas a partir do método proposto. Trabalhos futuros para a pesquisa desenvolvida envolvem, construir novos métodos de apoio ao diagnóstico de lesões em imagens de mamografia. Também podemos mencionar a necessidade de realizar novos experimentos com diferentes bases de dados de mamogramas, visando avaliar se os resultados obtidos com a base IRMA podem ser generalizados a outras bases.

Além disso, temos a pretensão de utilizar os neurônios morfológicos propostos (dilatação e erosão) para criar e avaliar o desempenho de outras arquiteturas de redes neurais, por exemplo, SVMs morfológicos, redes MLPs morfológicas, entre outras.

Por fim, pretendemos realizar experimentos com o método proposto em diferentes domínios de aplicação, tendo como principal objetivo, verificar se o desempenho do método proposto consegue generalizar na resolução de outros problemas.

REFERÊNCIAS

ALI, H. R. et al. Astronomical algorithms for automated analysis of tissue protein expression in breast cancer. **British journal of cancer**, v. 108, n. 3, p. 602-612, 2013.

ANDERSON, B. D. O.; MOORE, J. B. Optimal filtering, ser. **Information and system science series**. Englewood Cliffs, NJ, USA: Prentice Hall, 1979.

ANGELA, J. Yu; GIESE, Martin A.; POGGIO, Tomaso A. Biophysiological plausible implementations of the maximum operation. **Neural Computation**, v. 14, n. 12, p. 2857-2881, 2002.

ARAUJO, Carmen P. Suarez; RITTER, Gerhard X. Morphological neural networks and image algebra in artificial perception systems. In: **San Diego'92**. International Society for Optics and Photonics, p. 128-142, 1992.

ARAUJO, R. de A. et al. An evolutionary morphological approach for financial time series forecasting. In: **Evolutionary Computation, 2006. CEC 2006. IEEE Congress**, p. 2467-2474, 2006B.

ARAUJO, R. de A. et al. Modular morphological neural network training via adaptive genetic algorithm for designing translation invariant operators. In: **Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference**, p. II-II, 2006A.

ARAÚJO, Ricardo de A.; OLIVEIRA, Adriano LI; MEIRA, Silvio RL. **A dilation-erosion-linear perceptron for bovespa index prediction**. In: International Conference on Intelligent Data Engineering and Automated Learning. Springer Berlin Heidelberg, 2012. p. 407-415.

AZEVEDO, Washington W. et al. Morphological Extreme Learning Machines Applied to Detect and Classify Masses in Mammograms. In: **The International Joint Conference on Neural Networks - IJCNN 2015**, 2015, Killarney, Irlanda. Proceedings of IJCNN 2015, 2015A.

AZEVEDO, Washington W. et al. Fuzzy Morphological Extreme Learning Machines to detect and classify masses in mammograms. In: **Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on**. IEEE, 2015. p. 1-8, 2015B.

BAI, Xiangzhi; ZHOU, Fugen. Multi structuring element top-hat transform to detect linear features. In: **Signal Processing (ICSP), 2010 IEEE 10th International Conference on**. IEEE, 2010. p. 877-880, 2010.

BAI, Xiangzhi; ZHOU, Fugen; XUE, Bindang. Edge preserved image fusion based on multiscale toggle contrast operator. **Image and Vision Computing**, v. 29, n. 12, p. 829-839, 2011.

BAI, Xiangzhi; ZHOU, Fugen; XUE, Bindang. Multiple linear feature detection through top-hat transform by using multi linear structuring elements. **Optik-International Journal for Light and Electron Optics**, v. 123, n. 22, p. 2043-2049, 2012.

BANON, Gerald Jean Francis; BARRERA, Junior. Decomposition of mappings between complete lattices by mathematical morphology, Part I. General lattices. **Signal Processing**, v. 30, n. 3, p. 299-327, 1993.

BANON, Gerald Jean Francis; BARRERA, Junior. **Bases da Morfologia Matemática para a análise de imagens binárias**. UFPE-DI, 1994.

BEALE, R.; FIESLER, E., Eds. **Handbook of Neural Computation**. Institute of Physics Publishing and Oxford University Press, 1997.

BELCIUG, Smaranda et al. **Assessing performances of unsupervised and supervised neural networks in breast cancer detection**. In: Informatics and Systems (INFOS), 2010 The 7th International Conference on. IEEE, 2010. p. 1-8.

BERTSEKAS, Dimitri P. **Dynamic Programming and Stochastic Control (Mathematics in science and engineering; v. 125)**. Elsevier, 1976.

BI-RADS. **BI-RADSTM(Breast Imaging Reporting and Data System - Sistema de Diagnóstico e Relatório para Imagens de Mama)**. American College of Radiology. Quarta edição, 2003.

BOSCH, Anne M. et al. Preoperative estimation of the pathological breast tumour size by physical examination, mammography and ultrasound: a prospective study on 105 invasive tumours. **European journal of radiology**, v. 48, n. 3, p. 285-292, 2003.

BRAGA, A.; CARVALHO, A.; LUDERMIR, T. **Redes Neurais Artificiais: Teoria e Aplicações**. 1ª edição, Rio de Janeiro: LTC, 2000.

CAMBRIA, Erik et al. **Extreme learning machines [trends & controversies]**. IEEE Intelligent Systems, v. 28, n. 6, p. 30-59, 2013.

CARPENTER, Gail A.; GROSSBERG, Stephen. A massively parallel architecture for a self-organizing neural pattern recognition machine. **Computer vision, graphics, and image processing**, v. 37, n. 1, p. 54-115, 1987.

CHEN, Ting et al. A pseudo top-hat mathematical morphological approach to edge detection in dark regions. **Pattern Recognition**, v. 35, n. 1, p. 199-210, 2002.

COLEMAN, Michel P. et al. **Cancer survival in five continents: a worldwide population-based study (CONCORD)**. The lancet oncology, v. 9, n. 8, p. 730-756, 2008.

CORDEIRO, Filipe R. et al. **Segmentation of mammography by applying extreme learning machine in tumor detection**. In: International Conference on Intelligent Data Engineering and Automated Learning. Springer Berlin Heidelberg, 2012. p. 92-100.

COSTA, Humberto; SOLLA, Jorge; TEMPORÃO, José Gomes. Ministério da Saúde. **Instituto Nacional do Câncer. Controle do Câncer de Mama: Documento de Consenso**. Rio de Janeiro, 2004.

D'Orsi, C. J. **Breast Imaging Reporting and Data System (BI-RADS)**. American College of Radiology, 1998.

DAVIDSON, J.; RITTER, G. A theory of morphological neural networks. In **Digital Optical Computing II**, vol. 1215 of Proceedings of SPIE, p. 378–388, 1990.

DA-XI, Wang; YUAN, Fang; SHENG, Huang. **An algorithm for medical imaging identification based on edge detection and seed filling**. In: Computer Application and System Modeling (ICCA SM), 2010 International Conference on. IEEE, 2010. p. V15-547-V15-548.

DE ABREU, Evaldo; KOIFMAN, Sérgio. Fatores prognósticos no câncer da mama feminina. **Revista Brasileira de Cancerologia**, v. 48, n. 1, p. 113-31, 2002.

DE OLIVEIRA, Julia EE et al. MammoSys: A content-based image retrieval system using breast density patterns. **Computer methods and programs in biomedicine**, v. 99, n. 3, p. 289-297, 2010.

DESANTIS, C. E. et al. Cancer treatment and survivorship statistics, 2014. **CA: a cancer journal for clinicians**, [S.l.], v.64, n.4, p.252–71, 2014.

DE SOUSA, Robson P. et al. Designing translation invariant operations via neural network training. In: **Image Processing, 2000. Proceedings. 2000 International Conference on**. IEEE, 2000. p. 908-911, 2000.

DE SOUZA, Flávio Joaquim. **Modelos neuro-fuzzy hierárquicos**. 1999. Tese de Doutorado. Pontifícia Universidade Católica do Rio de Janeiro.

DE SOUZA, Flavio Joaquim; VELLASCO, Marley MR; PACHECO, Marco AC. Hierarchical neuro-fuzzy quadtree models. **Fuzzy sets and systems**, v. 130, n. 2, p. 189-205, 2002.

DESERNO, Thomas et al. Towards computer-aided diagnostics of screening mammography using content-based image retrieval. In: **Graphics, Patterns and Images (Sibgrapi), 2011 24th SIBGRAPI Conference on**. IEEE, 2011. p. 211-219.

DESERNO, Thomas M. et al. Computer-aided diagnostics of screening mammography using content-based image retrieval. In: **SPIE medical imaging**. International Society for Optics and Photonics, 2012. p. 831527-831527-9.

DOYLE, John; STEIN, Gunter. Multivariable feedback design: Concepts for a classical/modern synthesis. **IEEE transactions on Automatic Control**, v. 26, n. 1, p. 4-16, 1981.

FERNANDES, I.; SANTOS, W. dos. Classificação de Mamografias Utilizando Extração de Atributos de Textura e Redes Neurais Artificiais. Congresso Brasileiro de Engenharia Biomédica, [S.l.], p.8, 2014.

FIORENTINO, Carla et al. Accuracy of mammography and echography versus clinical palpation in the assessment of response to primary chemotherapy in breast cancer patients with operable disease. **Breast cancer research and treatment**, v. 69, n. 2, p. 143-151, 2001.

FORNAGE, Bruno D.; TOUBAS, Olivier; MOREL, Michel. Clinical, mammographic, and sonographic determination of preoperative breast cancer size. **Cancer**, v. 60, n. 4, p. 765-771, 1987.

FRANK, Andrew; ASUNCION, Arthur. UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California. **School of Information and Computer Science**, v. 213, 2010.

FUKUNAGA, Reinosuke. **Statistical pattern recognition**. Academic Press., 1990.

GADER, Paul D.; KHABOU, Mohamed A.; KOLDOBSKY, Alexander. Morphological regularization neural networks. **Pattern Recognition**, v. 33, n. 6, p. 935-944, 2000.

GONZALES, R. C.; WOODS, R. E. **Digital Image Processing**. New York: Prentice Hall, 2007.

GRAÑA, Manuel et al. On the application of associative morphological memories to hyperspectral image analysis. In: **International Work-Conference on Artificial Neural Networks**. Springer Berlin Heidelberg, 2003. p. 567-574, 2003.

GUPTA, Madan; JIN, Liang; HOMMA, Noriyasu. **Static and dynamic neural networks: from fundamentals to advanced theory**. John Wiley & Sons, 2004.

GUPTA, S. L.; KASHYAP, A. K.; SINGH, A. P. Responses of blue-green algae to cooper: a comparative study. **Indian journal of environmental health**, v. 27, n. 3, p. 224-9, 1985.

HÄNNI, Mari et al. Quantitation of atherosclerosis by magnetic resonance imaging and 3-D morphology operators. **Magnetic resonance imaging**, v. 17, n. 4, p. 585-591, 1999.

HARALICK, Robert M. et al. Textural features for image classification. **IEEE Transactions on systems, man, and cybernetics**, v. 3, n. 6, p. 610-621, 1973.

HAYKIN, S. **Redes Neurais: Princípios e Prática**. 2ª ed. Porto Alegre: Bookman, 2001.

HEATH, Michael et al. The digital database for screening mammography. In: **Proceedings of the 5th international workshop on digital mammography**. Medical Physics Publishing, 2000. p. 212-218.

HERRADA, Juan et al. Relative value of physical examination, mammography, and breast sonography in evaluating the size of the primary tumor and regional lymph node metastases in women receiving neoadjuvant chemotherapy for locally advanced breast carcinoma. **Clinical Cancer Research**, v. 3, n. 9, p. 1565-1569, 1997.

HOLMBLAD, L. P.; OSTERGAARD, J. J. Fuzzy logic control--operator experience applied to automatic process control. **ZKG, Zement-Kalk-Gips, Edition A**, v. 34, p. 127-33, 1981.

HOTZMAN, J. M. *Nonlinear System Theory*. 1970.

HUANG, Guang-Bin et al. **Extreme learning machine for regression and multiclass classification**. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, v. 42, n. 2, p. 513-529, 2012.

HUANG, Guang-Bin; CHEN, Lei. **Enhanced random search based incremental extreme learning machine**. *Neurocomputing*, v. 71, n. 16, p. 3460-3468, 2008.

HUANG, Guang-Bin; CHEN, Yan-Qiu; BABRI, Haroon A. Classification ability of single hidden layer feedforward neural networks. *IEEE Transactions on Neural Networks*, v. 11, n. 3, p. 799-801, 2000.

HUANG, Guang-Bin; ZHU, Qin-Yu; SIEW, Chee-Kheong. **Extreme learning machine: a new learning scheme of feedforward neural networks**. In: *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*. IEEE, 2004. p. 985-990, 2004.

HUANG, Guang-Bin; ZHU, Qin-Yu; SIEW, Chee-Kheong. Extreme learning machine: theory and applications. *Neurocomputing*, v. 70, n. 1, p. 489-501, 2006.

HWANG, Sun-Kyoo; KIM, Whoi-Yul. A novel approach to the fast computation of Zernike moments. *Pattern Recognition*, v. 39, n. 11, p. 2065-2076, 2006.

IKONEN, Tiia et al. Morphological extraction of cancerous nucleus in the diagnostics of breast cancer. In: **Intelligent Systems Design and Applications (ISDA), 2015 15th International Conference on**. IEEE, 2015. p. 148-153, 2015.

JANGHEL, R. R. et al. **Breast cancer diagnosis using artificial neural network models**. In: *Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on*. IEEE, 2010. p. 89-94, 2010.

JEATRAKUL, P.; WONG, K. W. **Comparing the performance of different neural networks for binary classification problems**. In: *Natural Language Processing, 2009. SNLP'09. Eighth International Symposium on*. IEEE, 2009. p. 111-115, 2009.

JENIFER, Sheeba; PARASURAMAN, S.; KADIRVEL, Amudha. An efficient biomedical imaging technique for automatic detection of abnormalities in digital mammograms. *Journal of Medical Imaging and Health Informatics*, v. 4, n. 2, p. 291-296, 2014.

JUHL, John; CRUMMY, Andrew B.; KUHLMAN, Janet E. Paul & Juhl. **Interpretação radiológica**. In: *Paul & Juhl interpretação radiológica*. Guanabara Koogan, 2000.

KABURLASOS, Vassilis G.; PETRIDIS, Vassilios. Fuzzy lattice neurocomputing (FLN) models. *Neural Networks*, v. 13, n. 10, p. 1145-1170, 2000.

KHABOU, Mohamed A.; GADER, Paul D. Automatic target detection using entropy optimized shared-weight neural networks. **IEEE transactions on neural networks**, v. 11, n. 1, p. 186-193, 2000A.

KHABOU, Mohamed A.; GADER, Paul D.; KELLER, James M. LADAR target detection using morphological shared-weight neural networks. **Machine Vision and Applications**, v. 11, n. 6, p. 300-305, 2000B.

KLIR, George; YUAN, Bo. **Fuzzy sets and fuzzy logic**. New Jersey: Prentice hall, 1995.

KOHONEN, Teuvo. Self-organizing maps, volume 30 of Springer Series in Information Sciences. 1995.

KRUSE, R.; BORGELT, C.; KLAWONN, F.; MOEWES, C.; STEINBRECHER, M.; HELD, P. **Multi-layer perceptrons**, In: Computational Intelligence. Springer, 2013, p. 47–81.

LEE, Chuen-Chien. Fuzzy logic in control systems: fuzzy logic controller. I. **IEEE Transactions on systems, man, and cybernetics**, v. 20, n. 2, p. 404-418, 1990.

LEMBESSIS, Evangelos. **Dynamic learning behaviour of a rule-based self-organising controller**. 1984. Tese de Doutorado. Queen Mary, University of London.

LIBERMAN, Laura; MENELL, Jennifer H. Breast imaging reporting and data system (BI-RADS). **Radiologic Clinics of North America**, v. 40, n. 3, p. 409-430, 2002.

MAITRA, Indra Kanta; NAG, Sanjay; BANDYOPADHYAY, Samir Kumar. **Identification of abnormal masses in digital mammography images**. International Journal of Computer Graphics, v. 2, n. 1, p. 17-30, 2011.

MALLAT, S. **A Wavelet Tour Of Signal Processing**. Segunda edição, Editora: New york: Academic, 1999.

MAMDANI, Ebrahim H. Application of fuzzy algorithms for control of simple dynamic plant. In: **Proceedings of the Institution of Electrical Engineers**. IET, 1974. p. 1585-1588.

MAMDANI, Ebrahim H.; ASSILIAN, Sedrak. An experiment in linguistic synthesis with a fuzzy logic controller. **International journal of man-machine studies**, v. 7, n. 1, p. 1-13, 1975.

MATHERON, Georges et al. Random sets and integral geometry. 1975.

MATHERS, Colin; FAT, Doris Ma; BOERMA, Jan Ties. **The global burden of disease: 2004 update**. World Health Organization, 2008.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115-133, 1943.

MEDEN, H.; NEUES, K. P.; RÖBEN-KÄMPKEN, S.; KUHN, W. A. Clinical, **Mamographic, Sonographic and Histologic Evaluation of Breast Cancer. International Journal of Gynecology & Obstetrics. Volume 48, número 2, p. 193-199, 1995.**

MICHE, Yoan et al. OP-ELM: optimally pruned extreme learning machine. **IEEE Transactions on Neural Networks**, v. 21, n. 1, p. 158-162, 2010.

MINSKY, Marvin; PAPERT, Seymour. Perceptron (expanded edition). 1969.

MONTEIRO, FERNANDA; SIQUEIRA, PAULA. Correlação entre o exame clínico, a mamografia e a ultra-sonografia com o exame anatomopatológico na determinação do tamanho tumoral no câncer de mama. **Rev Bras Ginecol Obstet**, v. 30, n. 3, p. 107-12, 2008.

NASHED, M. Zuhair (Ed.). **Generalized Inverses and Applications: Proceedings of an Advanced Seminar Sponsored by the Mathematics Research Center**, the University of Wisconsin—Madison, October 8-10, 1973. Elsevier, 2014.

NGUYEN, Duc Thanh et al. A robust document skew estimation algorithm using mathematical morphology. In: **Tools with Artificial Intelligence, 2007. ICTAI 2007. 19th IEEE International Conference on**. IEEE, 2007. p. 496-503.

PAWAR, Punam S.; PATIL, Dharmaraj R. **Breast cancer detection using neural network models**. In: Communication Systems and Network Technologies (CSNT), 2013 International Conference on. IEEE, 2013. p. 568-572.

PESSOA, Lúcio FC; MARAGOS, Petros. MRL-filters: A general class of nonlinear systems and their optimal design for image processing. **IEEE Transactions on Image Processing**, v. 7, n. 7, p. 966-978, 1998.

PESSOA, Lucio FC; MARAGOS, Petros. Neural networks with hybrid morphological/rank/linear nodes: a unifying framework with applications to handwritten character recognition. **Pattern Recognition**, v. 33, n. 6, p. 945-960, 2000.

PETRIDIS, Vassilios; KABURLASOS, Vassilis G. Fuzzy lattice neural network (FLNN): a hybrid model for learning. **IEEE Transactions on Neural Networks**, v. 9, n. 5, p. 877-890, 1998.

PIERIE, J. P. E. N. et al. Clinical assessment, mammography and ultrasonography as methods of measuring the size of breast cancer: a comparison. **The Breast**, v. 7, n. 5, p. 247-250, 1998.

PINHEIRO, P. **O que é Câncer?** Disponível em <http://www.mdsaude.com/2009/02/cancer-cancro-sintomas.html>. Último acesso: 07/02/16.

QUINTANILLA-DOMINGUEZ, Joel et al. Microcalcification detection applying artificial neural networks and mathematical morphology in digital mammograms. In: **World Automation Congress (WAC)**, 2010. IEEE, 2010. p. 1-6, 2010.

RADUCANU, Bogdan; GRAÑA, Manuel; ALBIZURI, F. Xabier. Morphological scale spaces and associative morphological memories: Results on robustness and practical applications. **Journal of Mathematical Imaging and Vision**, v. 19, n. 2, p. 113-131, 2003.

RESCHER, N. **Many-valued logic**, McGraw-Hill, New York, 1969.

REZENDE, Solange Oliveira. **Sistemas fuzzy, In Sistemas inteligentes: fundamentos e aplicações**. Editora Manole Ltda, 2002.

RITTER, G. X.; SUSSNER, P. Morphological neural networks. In **Intelligent Systems: A Semiotic Perspective; Proceedings of the 1996 International Multidisciplinary Conference**, p. 221–226, 1996.

RITTER, G. X.; SUSSNER, P. Morphological perceptrons. In **ISAS'97, Intelligent Systems and Semiotics**, 1997.

RITTER, G. X.; SUSSNER, P.; DE LEON, J. L. D. Morphological associative memories. **IEEE Transactions on Neural Networks**, p. 281–293, 1998.

RITTER, Gerhard X.; URCID, Gonzalo. Lattice algebra approach to single-neuron computation. **IEEE Transactions on Neural Networks**, v. 14, n. 2, p. 282-295, 2003.

RITTER, Gerhard X.; WILSON, Joseph N.; DAVIDSON, Jennifer L. Image algebra: An overview. **Computer Vision, Graphics, and Image Processing**, v. 49, n. 3, p. 297-331, 1990.

ROSENBLATT, Frank. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychological review**, v. 65, n. 6, p. 386, 1958.

ROUHI, Rahimeh et al. Benign and malignant breast tumors classification based on region growing and CNN segmentation. **Expert Systems with Applications**, v. 42, n. 3, p. 990-1002, 2015.

RSNA - Radiological Society of North America. **Mammography Screening Intervals Mays Affect Breast Cancer Prognosis**. 2013. [Acessado em: 17/06/2017]. Disponível em: http://press.rsna.org/timssnet/media/pressreleases/pr_target.cfm?ID=70

SAGE, Andrew P.; WHITE, Chelsea C. **Optimum systems control**. Englewood Cliffs, NJ: Prentice-Hall, 1977.

SANTOS, W. P. **Mathematical Morphology**, in Digital Document Analysis and Processing, MELO, C. A. B.; OLIVEIRA, A. L. I.; SANTOS, W. P. Eds. New York: Nova Science, 2012, pp. 159–192.

SHOMA, Ashraf et al. Ultrasound for accurate measurement of invasive breast cancer tumor size. **The breast journal**, v. 12, n. 3, p. 252-256, 2006.

SIMONTON, Oscar Carl; SIMONTON, Stephanie; CREIGHTON, James L. **Com a vida de novo: uma abordagem de auto-ajuda para pacientes com câncer**. Summus, 1987.

SINGH, S.; Gupta, P.R. Breast cancer detection and classification using neural network, (**IJAEST International Journal of Advanced Engineering Sciences And Technologies**, Vol No. 6, Issue No. 1, p. 04-09, 2011).

SOILLE, P. **Morphological Image Analysis Principles and Applications**. Springer-Verlag, 2004.

SOILLE, Pierre. **Morphological image analysis: principles and applications**. Springer Science & Business Media, 2013.

STORK, David G. et al. Pattern classification. **A Wiley-Interscience Publication**, 2001.

SUCKLING, John et al. The mammographic image analysis society digital mammogram database. In: **Exerpta Medica. International Congress Series**, p. 375-378, 1994.

SUSSNER, P. Morphological perceptron learning. In **Proceedings of IEE ISIC/CIRA/ISAS Joint Conference**, p. 447-482, 1998.

SUSSNER, P. Perceptrons. In **Wiley Encyclopedia of Electrical and Electronics Engineering**, J. G. Webster, Ed. John Wiley and Sons, Inc., New York, 1999, p. 44-59, 1999.

SUSSNER, Peter. Associative morphological memories based on variations of the kernel and dual kernel methods. **Neural Networks**, v. 16, n. 5, p. 625-632, 2003A.

SUSSNER, Peter. Generalizing operations of binary autoassociative morphological memories using fuzzy set theory. **Journal of mathematical imaging and vision**, v. 19, n. 2, p. 81-93, 2003B.

SUSSNER, Peter. Observations on morphological associative memories and the kernel method. **Neurocomputing**, v. 31, n. 1, p. 167-183, 2000.

SUSSNER, Peter; VALLE, Marcos Eduardo. Gray-scale morphological associative memories. **IEEE Transactions on Neural Networks**, v. 17, n. 3, p. 559-570, 2006A.

SUSSNER, Peter; VALLE, Marcos Eduardo. Implicative fuzzy associative memories. **IEEE Transactions on Fuzzy Systems**, v. 14, n. 6, p. 793-807, 2006B.

TAHMASBI, Amir; SAKI, Fatemeh; SHOKOUHI, Shahriar B. Classification of benign and malignant masses based on Zernike moments. **Computers in biology and medicine**, v. 41, n. 8, p. 726-735, 2011.

TEIXEIRA, Marta de Betânia Rabelo. Avaliação dos achados mamográficos classificados na categoria 4 do sistema BI_RADS® e sua correlação histopatológica. 2011.

THEODORIDIS, Sergios. L. koutroumbas. Pattern recognition, v. 4, 2003.

THIRAN, J.-P.; MACQ, Benoit. Morphological feature extraction for the classification of digital images of cancerous tissues. **IEEE Transactions on biomedical engineering**, v. 43, n. 10, p. 1011-1020, 1996.

THOMAS, Simon J.; COWLEY, Ian R. A comparison of four indices for combining distance and dose differences. **International Journal of Radiation Oncology* Biology* Physics**, v. 82, n. 5, p. e717-e723, 2012.

VALLE, Marcos Eduardo; SUSSNER, Peter; GOMIDE, Fernando. Introduction to implicative fuzzy associative memories. In: **Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on**. IEEE, 2004. p. 925-930, 2004.

VEGA-CORONA, Antonio et al. Contextual and Non-Contextual Features Extraction and a Selection Method for Microcalcifications Detection. In: **Automation Congress, 2006. WAC'06. World**. IEEE, 2006. p. 1-6, 2006.

WANG, Weizhuo; MOTTERSHEAD, John E.; MARES, Cristinel. Mode-shape recognition and finite element model updating using the Zernike moment descriptor. **Mechanical Systems and Signal Processing**, v. 23, n. 7, p. 2088-2112, 2009.

WANG, Weizhuo; MOTTERSHEAD, John E.; MARES, Cristinel. Mode-shape recognition and finite element model updating using the Zernike moment descriptor. **Mechanical Systems and Signal Processing**, v. 23, n. 7, p. 2088-2112, 2009.

WANG, Xi-Zhao. **International journal of machine learning and cybernetics**. 2010.

XIANG, Cheng; DING, Shenqiang Q.; LEE, Tong Heng. Geometrical interpretation and architecture selection of MLP. **IEEE Transactions on Neural Networks**, v. 16, n. 1, p. 84-96, 2005.

YE, Shuzhi; ZHENG, Sheng; HAO, Wei. **Medical image edge detection method based on adaptive facet model**. In: Computer application and system modeling (ICCASM), 2010 international conference on. IEEE, 2010. p. V3-574-V3-578.

YU, Mei et al. New morphological operators for color image processing. In: **TENCON 2004. 2004 IEEE Region 10 Conference**. IEEE, 2004. p. 443-446, 2004.

ZADEH, Lotfi A. Fuzzy sets. **Information and control**, v. 8, n. 3, p. 338-353, 1965.

ZADEH, Lotfi A. Outline of a new approach to analysis of complex systems and decision processes. **IEEE Transactions on Systems Man and Cybernetics**, n. 1, p. 28-44, 1973.

ZADEH, Lotfi Asker. Fuzzy logic. **Computer**, v. 21, n. 4, p. 83-93, 1988.

ZRIBI, Manel et al. **The self-organizing maps of Kohonen in the medical classification**. In: Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2012 6th International Conference on. IEEE, 2012. p. 852-856.

ANEXO A-Resultados Expandidos do

Capítulo 5

Tabela 0-1 Resultados expandidos, referentes à Tabela 5-1. Todos os resultados obtidos com os métodos propostos e os principais classificadores do estado da arte utilizando Atributos de *Haralick* (HA).

<i>Tecido</i>	<i>Descritor de imagem</i>	<i>Classificador</i>	<i>Acerto teste (%)</i>	<i>Kappa</i>
I (Adiposo)	HA	mELM (dilatação)	91,56 (0,00)	0,9048 (0,00)
	HA	mELM (erosão)	92,13 (0,00)	0,9113 (0,00)
	HA	ELM <i>kernel</i> linear	75,31 (0,77)	0,7118 (0,01)
	HA	ELM <i>kernel</i> RBF	62,65 (3,39)	0,5779 (0,04)
	HA	ELM com função de ativação RBF	85,67 (0,55)	0,8383 (0,01)
	HA	ELM com função de ativação sigmóide	88,62 (0,80)	0,8713 (0,01)
	HA	Redes <i>Bayesianas</i>	79,34 (4,40)	0,69 (0,07)
	HA	Redes <i>Naive Bayes</i>	74,00 (4,76)	0,61 (0,07)
	HA	Redes MLP	85,04 (4,16)	0,78 (0,06)
	HA	Redes RBF	81,91 (4,52)	0,73 (0,07)
	HA	SVM <i>kernel</i> linear	89,32 (3,55)	0,84 (0,05)
	HA	SVM <i>kernel</i> polinomial grau 2	84,27 (4,02)	0,76 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 3	83,41 (3,67)	0,75 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 4	74,18 (4,49)	0,61 (0,07)
	HA	SVM <i>kernel</i> polinomial grau 5	73,61 (4,58)	0,60 (0,07)
	HA	SVM <i>kernel</i> RBF	84,66 (3,91)	0,77 (0,06)
	HA	kNN (k = 1)	84,16 (4,06)	0,76 (0,06)
	HA	kNN (k = 3)	83,83 (4,03)	0,76 (0,06)
	HA	kNN (k = 5)	82,86 (3,99)	0,74 (0,06)
	HA	kNN (k = 7)	83,35 (4,12)	0,75 (0,06)
	HA	kNN (k = 9)	83,12 (4,11)	0,75 (0,06)
	HA	J48	80,70 (4,18)	0,71 (0,06)
	HA	<i>Random Tree</i>	78,24	0,67

			(4,97)	(0,07)
	HA	<i>Random Forest</i>	85,15 (4,02)	0,78 (0,06)
II (Fibroso)	HA	mELM (dilatação)	91,41 (0,00)	0,9034 (0,00)
	HA	mELM (erosão)	91,13 (0,00)	0,8998 (0,00)
	HA	ELM <i>kernel</i> linear	76,01 (1,24)	0,7217 (0,02)
	HA	ELM <i>kernel</i> RBF	39,26 (4,62)	0,3140 (0,05)
	HA	ELM com função de ativação RBF	89,06 (0,56)	0,8765 (0,01)
	HA	ELM com função de ativação sigmóide	90,74 (0,46)	0,8951 (0,01)
	HA	Redes <i>Bayesianas</i>	82,41 (4,10)	0,74 (0,06)
	HA	Redes <i>Naive Bayes</i>	80,63 (4,32)	0,71 (0,06)
	HA	Redes MLP	85,04 (4,16)	0,78 (0,06)
	HA	Redes RBF	81,91 (4,52)	0,73 (0,07)
	HA	SVM <i>kernel</i> linear	90,41 (3,26)	0,86 (0,05)
	HA	SVM <i>kernel</i> polinomial grau 2	85,76 (3,99)	0,79 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 3	83,20 (3,86)	0,75 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 4	77,98 (4,21)	0,67 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 5	70,74 (4,44)	0,56 (0,07)
	HA	SVM <i>kernel</i> RBF	87,15 (3,75)	0,81 (0,06)
	HA	kNN (k = 1)	83,37 (3,75)	0,75 (0,06)
	HA	kNN (k = 3)	85,01 (3,93)	0,78 (0,06)
	HA	kNN (k = 5)	85,03 (3,86)	0,78 (0,06)
	HA	kNN (k = 7)	85,08 (3,91)	0,78 (0,06)
	HA	kNN (k = 9)	85,29 (3,79)	0,78 (0,06)
	HA	J48	84,54 (4,07)	0,77 (0,06)
	HA	<i>Random Tree</i>	82,67 (4,28)	0,74 (0,06)
	HA	<i>Random Forest</i>	88,14 (3,68)	0,82 (0,06)
III (Heterogeneamente denso)	HA	mELM (dilatação)	78,97 (0,00)	0,7599 (0,00)
	HA	mELM (erosão)	94,13 (0,00)	0,9337 (0,00)
	HA	ELM <i>kernel</i> linear	74,58 (0,29)	0,6870 (0,00)
	HA	ELM <i>kernel</i> RBF	47,35 (6,34)	0,4102 (0,06)
	HA	ELM com função de	89,70	88,22

		ativação RBF	(0,73)	(0,01)
	HA	ELM com função de ativação sigmóide	93,35 (0,83)	0,9243 (0,01)
	HA	Redes <i>Bayesianas</i>	82,99 (4,55)	0,74 (0,07)
	HA	Redes <i>Naive Bayes</i>	74,76 (3,35)	0,62 (0,05)
	HA	Redes MLP	86,48 (3,67)	0,80 (0,05)
	HA	Redes RBF	81,21 (4,55)	0,72 (0,07)
	HA	SVM <i>kernel</i> linear	92,01 (3,18)	0,88 (0,05)
	HA	SVM <i>kernel</i> polinomial grau 2	84,47 (3,92)	0,77 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 3	83,39 (4,05)	0,75 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 4	80,50 (3,61)	0,71 (0,05)
	HA	SVM <i>kernel</i> polinomial grau 5	73,94 (4,55)	0,61 (0,07)
	HA	SVM <i>kernel</i> RBF	85,11 (3,93)	0,78 (0,06)
	HA	kNN (k = 1)	85,39 (3,94)	0,78 (0,06)
	HA	kNN (k = 3)	85,91 (3,98)	0,79 (0,06)
	HA	kNN (k = 5)	85,07 (3,84)	0,78 (0,06)
	HA	kNN (k = 7)	84,70 (3,82)	0,77 (0,06)
	HA	kNN (k = 9)	84,58 (3,86)	0,77 (0,06)
	HA	J48	86,30 (4,05)	0,79 (0,06)
	HA	<i>Random Tree</i>	82,72 (4,18)	0,74 (0,06)
	HA	<i>Random Forest</i>	89,21 (3,56)	0,84 (0,05)
IV (Extremamente denso)	HA	mELM (dilatação)	90,13 (0,00)	0,8873 (0,00)
	HA	mELM (erosão)	91,27 (0,00)	0,9015 (0,00)
	HA	ELM <i>kernel</i> linear	63,50 (0,44)	0,6261 (0,01)
	HA	ELM <i>kernel</i> RBF	47,71 (3,43)	0,4506 (0,04)
	HA	ELM com função de ativação RBF	85,12 (1,08)	0,8344 (0,01)
	HA	ELM com função de ativação sigmóide	89,77 (0,67)	0,8848 (0,01)
	HA	Redes <i>Bayesianas</i>	78,15 (4,53)	0,67 (0,07)
	HA	Redes <i>Naive Bayes</i>	62,02 (5,66)	0,43 (0,08)
	HA	Redes MLP	85,38 (4,23)	0,78 (0,06)
	HA	Redes RBF	72,70 (4,90)	0,59 (0,07)
	HA	SVM <i>kernel</i> linear	90,03	0,85

			(3,61)	(0,05)
	HA	SVM <i>kernel</i> polinomial grau 2	79,60 (4,29)	0,69 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 3	77,13 (4,28)	0,66 (0,06)
	HA	SVM <i>kernel</i> polinomial grau 4	67,82 (5,60)	0,52 (0,08)
	HA	SVM <i>kernel</i> polinomial grau 5	63,85 (5,88)	0,46 (0,09)
	HA	SVM <i>kernel</i> RBF	82,28 (4,09)	0,73 (0,06)
	HA	kNN (k = 1)	79,91 (4,25)	0,70 (0,06)
	HA	kNN (k = 3)	80,72 (3,95)	0,71 (0,06)
	HA	kNN (k = 5)	80,20 (4,20)	0,70 (0,06)
	HA	kNN (k = 7)	80,07 (4,28)	0,70 (0,06)
	HA	kNN (k = 9)	80,33 (4,26)	0,70 (0,06)
	HA	J48	83,68 (4,07)	0,76 (0,06)
	HA	<i>Random Tree</i>	77,65 (5,07)	0,66 (0,08)
	HA	<i>Random Forest</i>	85,17 (3,95)	0,78 (0,06)
Todos os tecidos	HA	mELM (dilatação)	63,91 (0,00)	0,6367 (0,00)
	HA	mELM (erosão)	64,66 (0,00)	0,6443 (0,00)
	HA	ELM <i>kernel</i> linear	30,87 (0,53)	0,3079 (0,01)
	HA	ELM <i>kernel</i> RBF	10,51 (0,53)	0,1024 (0,01)
	HA	ELM com função de ativação RBF	55,01 (0,64)	54,75 (0,01)
	HA	ELM com função de ativação sigmóide	63,65 (0,53)	0,6340 (0,01)
	HA	Redes <i>Bayesianas</i>	47,31 (2,80)	0,43 (0,03)
	HA	Redes <i>Naive Bayes</i>	38,87 (2,44)	0,33 (0,03)
	HA	Redes MLP	54,32 (2,99)	0,50 (0,03)
	HA	Redes RBF	35,61 (2,19)	0,30 (0,02)
	HA	SVM <i>kernel</i> linear	65,27 (2,50)	0,62 (0,03)
	HA	SVM <i>kernel</i> polinomial grau 2	54,69 (2,72)	0,51 (0,03)
	HA	SVM <i>kernel</i> polinomial grau 3	52,77 (2,69)	0,48 (0,03)
	HA	SVM <i>kernel</i> polinomial grau 4	46,52 (2,69)	0,42 (0,03)
	HA	SVM <i>kernel</i> polinomial grau 5	42,34 (2,53)	0,37 (0,03)
	HA	SVM <i>kernel</i> RBF	56,20 (2,64)	0,52 (0,03)
	HA	kNN (k = 1)	49,51	0,45

			(2,57)	(0,03)
	HA	kNN (k = 3)	51,28 (2,69)	0,47 (0,03)
	HA	kNN (k = 5)	52,54 (2,62)	0,48 (0,03)
	HA	kNN (k = 7)	53,67 (2,59)	0,49 (0,03)
	HA	kNN (k = 9)	53,55 (2,77)	0,49 (0,03)
	HA	J48	52,19 (3,17)	0,48 (0,03)
	HA	<i>Random Tree</i>	46,47 (3,01)	0,42 (0,03)
	HA	<i>Random Forest</i>	58,46 (2,89)	0,55 (0,03)

Figura 0-1 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido adiposo utilizando Atributos de *Haralick* (HA).

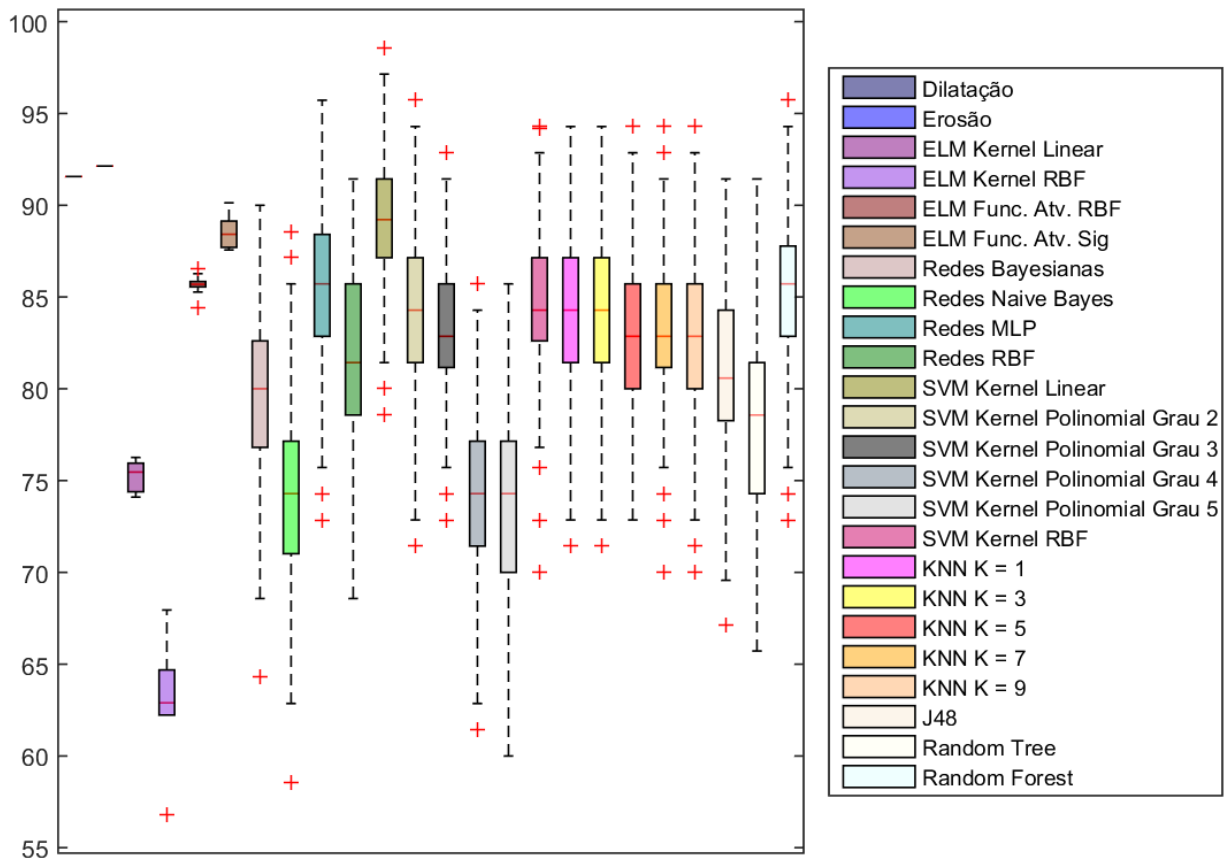


Figura 0-2 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido fibroso utilizando Atributos de *Haralick* (HA).

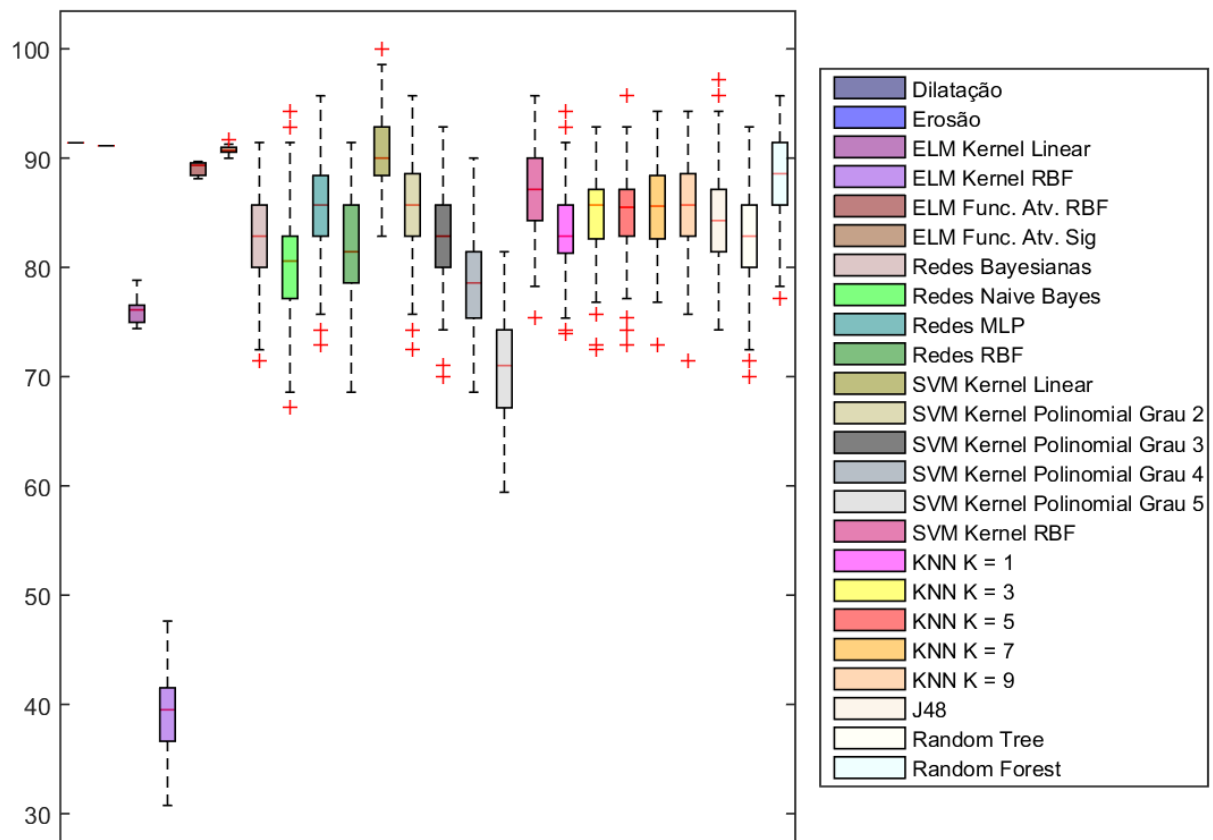


Figura 0-3 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Heterogeneamente denso utilizando Atributos de *Haralick* (HA).

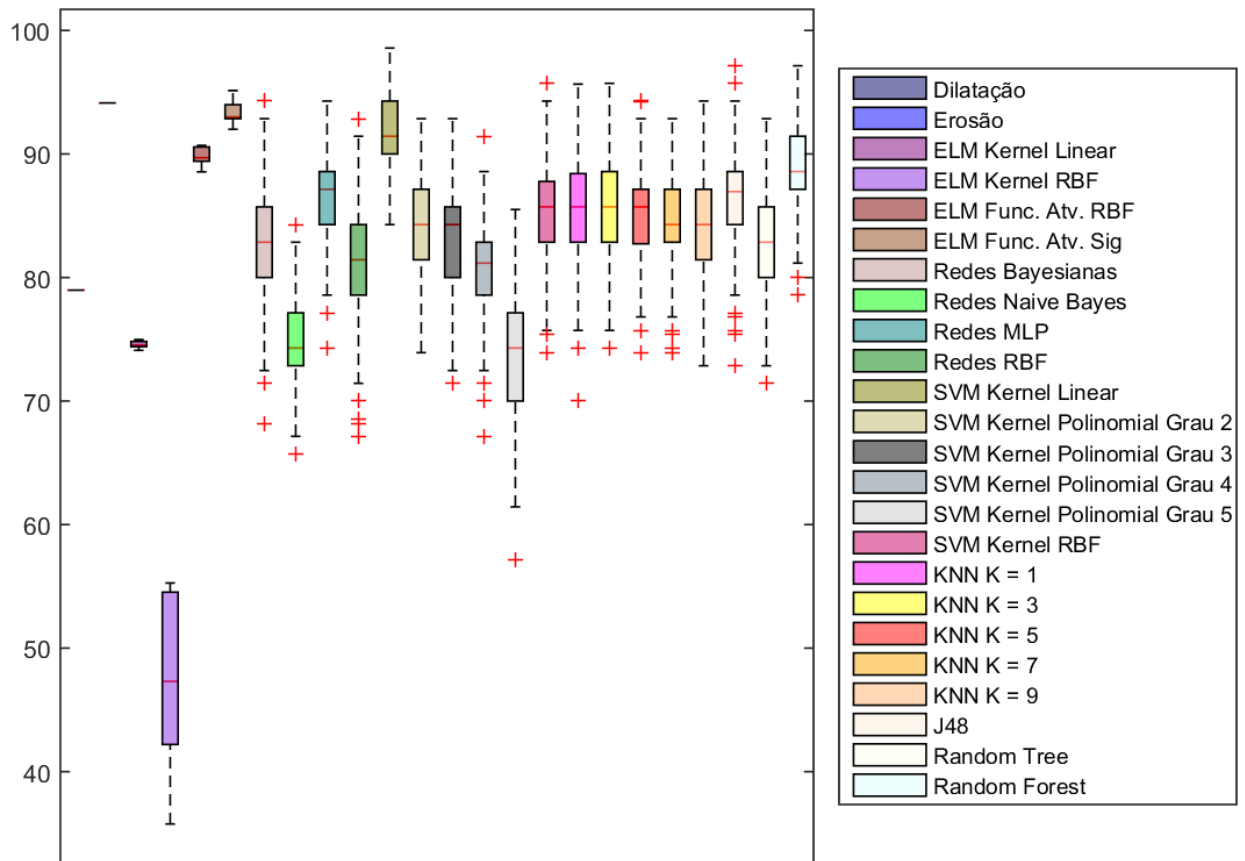


Figura 0-4 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Extremamente denso utilizando Atributos de *Haralick* (HA).

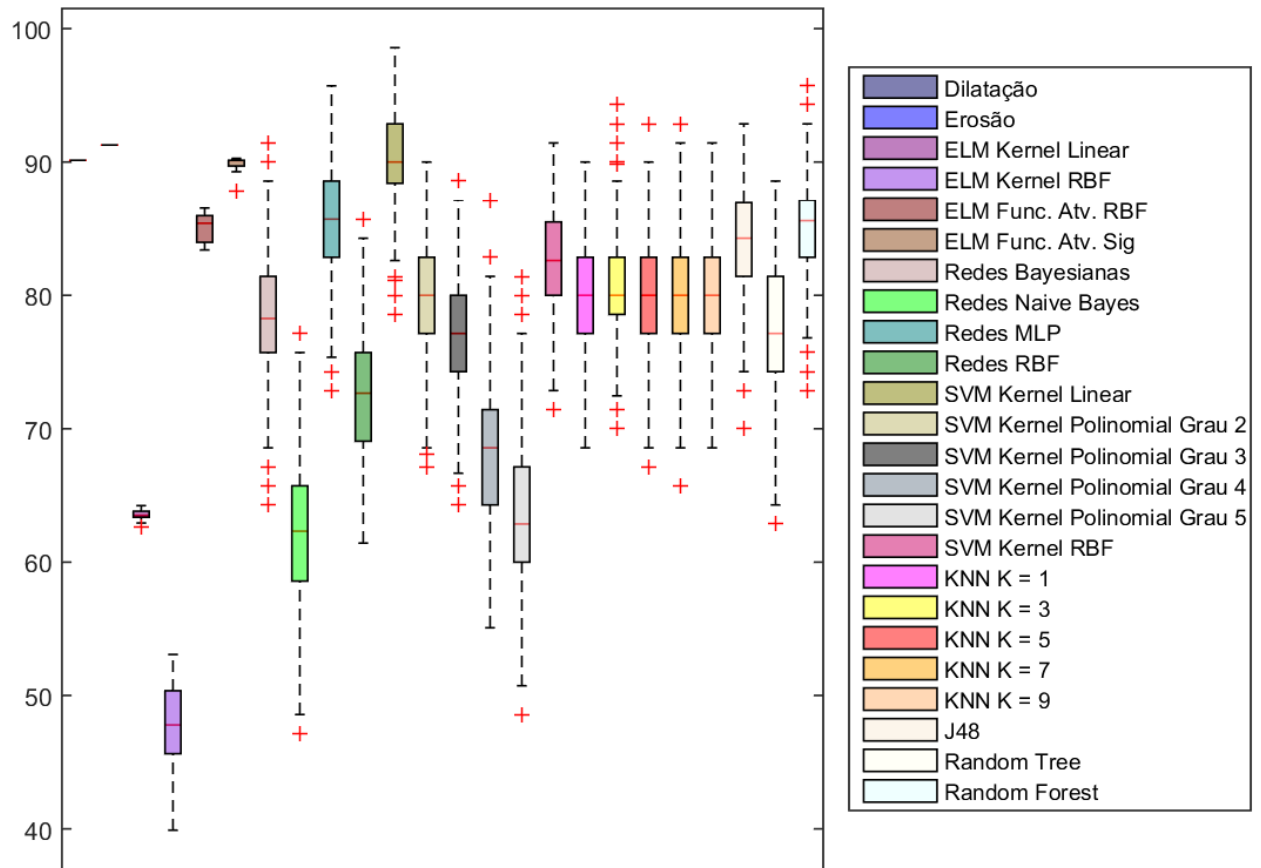


Figura 0-5 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com todos os tipos de tecidos utilizando Atributos de *Haralick* (HA).

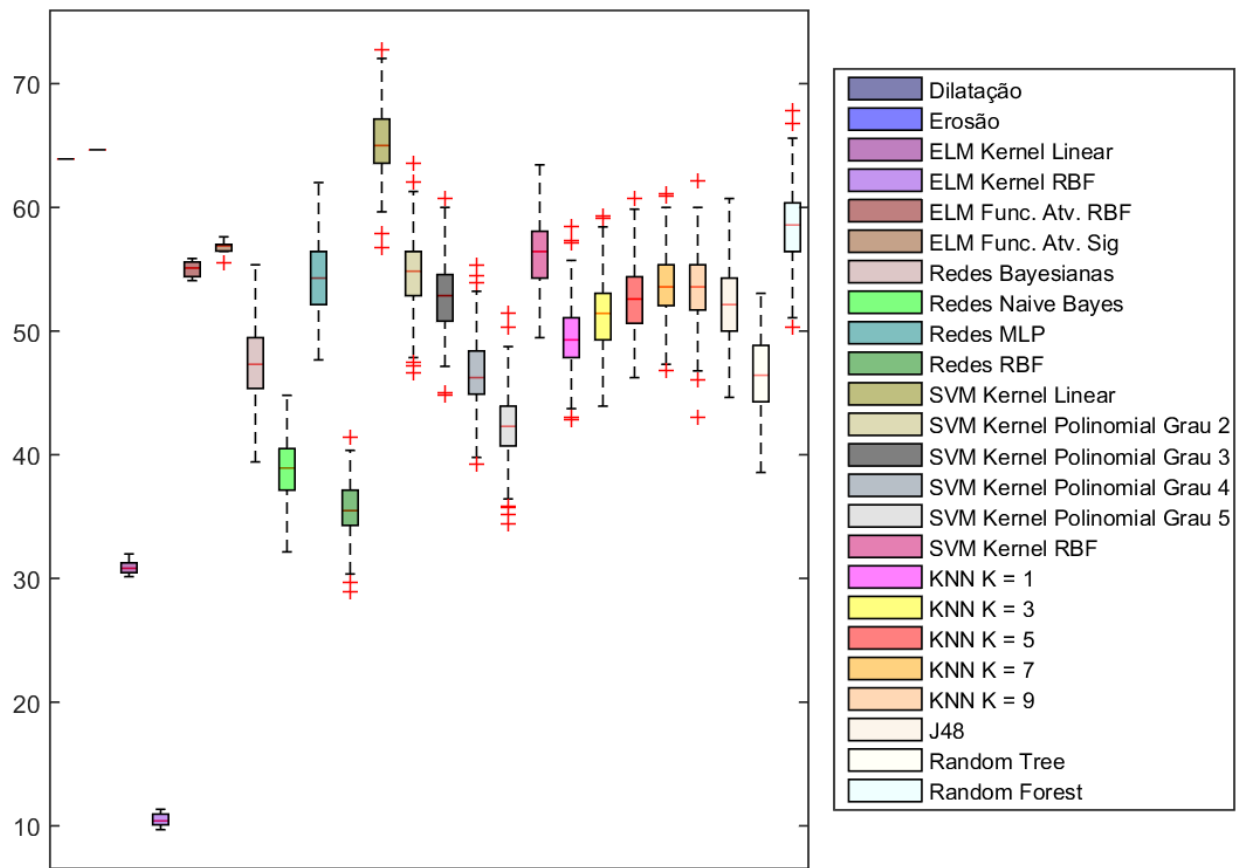


Tabela 0-2 Resultados expandidos, referentes à Tabela 5-3. Todos os resultados obtidos com os métodos propostos e os principais classificadores do estado da arte utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

<i>Tecido</i>	<i>Descritor de imagem</i>	<i>Classificador</i>	<i>Acerto teste (%)</i>	<i>Kappa</i>
I (Adiposo)	WA-HA	mELM (dilatação)	92,13 (0,00)	0,9108 (0,00)
	WA-HA	mELM (erosão)	93,28 (0,00)	0,9241 (0,00)
	WA-HA	ELM <i>kernel</i> linear	66,98 (0,51)	0,6221 (0,01)
	WA-HA	ELM <i>kernel</i> RBF	54,24 (1,14)	0,4889 (0,01)
	WA-HA	ELM com função de ativação RBF	83,41 (0,97)	0,8119 (0,01)
	WA-HA	ELM com função de ativação sigmóide	91,86 (0,77)	0,9078 (0,01)
	WA-HA	Redes <i>Bayesianas</i>	64,08 (5,12)	0,46 (0,08)
	WA-HA	Redes <i>Naive Bayes</i>	61,66 (5,82)	0,42 (0,09)
	WA-HA	Redes MLP	84,20 (5,04)	0,76 (0,08)
	WA-HA	Redes RBF	67,15 (4,85)	0,51 (0,07)
	WA-HA	SVM <i>kernel</i> linear	92,29 (3,05)	0,88 (0,05)
	WA-HA	SVM <i>kernel</i> polinomial grau 2	78,27 (5,02)	0,67 (0,08)
	WA-HA	SVM <i>kernel</i> polinomial grau 3	77,83 (4,88)	0,67 (0,07)
	WA-HA	SVM <i>kernel</i> polinomial grau 4	67,32 (5,59)	0,51 (0,08)
	WA-HA	SVM <i>kernel</i> polinomial grau 5	66,63 (5,37)	0,50 (0,08)
	WA-HA	SVM <i>kernel</i> RBF	82,41 (3,95)	0,74 (0,06)
	WA-HA	kNN (k = 1)	73,56 (4,87)	0,60 (0,07)
	WA-HA	kNN (k = 3)	76,26 (4,46)	0,64 (0,07)
	WA-HA	kNN (k = 5)	77,43 (4,68)	0,66 (0,07)
	WA-HA	kNN (k = 7)	78,75 (4,22)	0,68 (0,06)
	WA-HA	kNN (k = 9)	77,99 (4,54)	0,67 (0,07)
	WA-HA	J48	78,45 (4,89)	0,68 (0,07)
	WA-HA	<i>Random Tree</i>	71,11 (5,61)	0,57 (0,08)
	WA-HA	<i>Random Forest</i>	81,46 (4,47)	0,72 (0,07)
II (Fibroso)	WA-HA	mELM (dilatação)	91,56 (0,00)	0,9050 (0,00)
	WA-HA	mELM (erosão)	90,13 (0,00)	0,8886 (0,00)
	WA-HA	ELM <i>kernel</i> linear	60,26 (0,49)	0,5234 (0,01)
	WA-HA	ELM <i>kernel</i> RBF	51,02	0,4561

			(1,81)	(0,02)
	WA-HA	ELM com função de ativação RBF	78,89 (1,34)	0,7592 (0,02)
	WA-HA	ELM com função de ativação sigmóide	88,48 (0,89)	0,8689 (0,01)
	WA-HA	Redes <i>Bayesianas</i>	57,62 (4,63)	0,36 (0,07)
	WA-HA	Redes <i>Naive Bayes</i>	59,46 (4,38)	0,39 (0,07)
	WA-HA	Redes MLP	74,06 (5,52)	0,61 (0,08)
	WA-HA	Redes RBF	63,45 (3,99)	0,45 (0,06)
	WA-HA	SVM <i>kernel</i> linear	91,41 (3,46)	0,87 (0,05)
	WA-HA	SVM <i>kernel</i> polinomial grau 2	70,70 (5,01)	0,56 (0,08)
	WA-HA	SVM <i>kernel</i> polinomial grau 3	69,65 (5,06)	0,54 (0,08)
	WA-HA	SVM <i>kernel</i> polinomial grau 4	60,82 (6,04)	0,41 (0,09)
	WA-HA	SVM <i>kernel</i> polinomial grau 5	60,83 (5,97)	0,41 (0,09)
	WA-HA	SVM <i>kernel</i> RBF	75,89 (4,96)	0,64 (0,07)
	WA-HA	kNN (k = 1)	63,71 (5,10)	0,46 (0,08)
	WA-HA	kNN (k = 3)	64,99 (5,20)	0,47 (0,08)
	WA-HA	kNN (k = 5)	65,15 (5,66)	0,48 (0,08)
	WA-HA	kNN (k = 7)	66,29 (5,65)	0,49 (0,08)
	WA-HA	kNN (k = 9)	66,46 (5,42)	0,50 (0,08)
	WA-HA	J48	70,22 (5,11)	0,55 (0,08)
	WA-HA	<i>Random Tree</i>	64,67 (5,65)	0,47 (0,08)
	WA-HA	<i>Random Forest</i>	74,48 (5,03)	0,62 (0,08)
III (Heterogeneamente denso)	WA-HA	mELM (dilatação)	94,71 (0,00)	0,9400 (0,00)
	WA-HA	mELM (erosão)	96,42 (0,00)	0,9598 (0,00)
	WA-HA	ELM <i>kernel</i> linear	73,78 (0,45)	0,6944 (0,01)
	WA-HA	ELM <i>kernel</i> RBF	64,18 (1,28)	0,5886 (0,01)
	WA-HA	ELM com função de ativação RBF	81,23 (1,20)	0,7844 (0,01)
	WA-HA	ELM com função de ativação sigmóide	93,57 (0,59)	0,9268 (0,01)
	WA-HA	Redes <i>Bayesianas</i>	64,58 (5,37)	0,47 (0,08)
	WA-HA	Redes <i>Naive Bayes</i>	62,15 (4,83)	0,43 (0,07)
	WA-HA	Redes MLP	83,30 (7,14)	0,75 (0,11)
	WA-HA	Redes RBF	76,16	0,64

			(4,82)	(0,07)
	WA-HA	SVM <i>kernel</i> linear	96,57 (2,17)	0,95 (0,03)
	WA-HA	SVM <i>kernel</i> polinomial grau 2	78,26 (4,71)	0,67 (0,07)
	WA-HA	SVM <i>kernel</i> polinomial grau 3	75,67 (4,79)	0,64 (0,07)
	WA-HA	SVM <i>kernel</i> polinomial grau 4	72,22 (5,15)	0,58 (0,08)
	WA-HA	SVM <i>kernel</i> polinomial grau 5	69,05 (5,07)	0,54 (0,08)
	WA-HA	SVM <i>kernel</i> RBF	86,19 (4,13)	0,79 (0,06)
	WA-HA	kNN (k = 1)	76,35 (4,56)	0,65 (0,07)
	WA-HA	kNN (k = 3)	77,58 (4,72)	0,66 (0,07)
	WA-HA	kNN (k = 5)	78,05 (4,69)	0,67 (0,07)
	WA-HA	kNN (k = 7)	76,92 (4,93)	0,65 (0,07)
	WA-HA	kNN (k = 9)	75,63 (4,86)	0,63 (0,07)
	WA-HA	J48	78,06 (4,90)	0,67 (0,07)
	WA-HA	<i>Random Tree</i>	71,99 (5,35)	0,58 (0,08)
	WA-HA	<i>Random Forest</i>	82,42 (4,36)	0,74 (0,07)
IV (Extremamente denso)	WA-HA	mELM (dilatação)	93,71 (0,00)	0,9285 (0,00)
	WA-HA	mELM (erosão)	92,70 (0,00)	0,9172 (0,00)
	WA-HA	ELM <i>kernel</i> linear	65,67 (0,39)	0,6187 (0,00)
	WA-HA	ELM <i>kernel</i> RBF	56,79 (0,50)	0,5375 (0,01)
	WA-HA	ELM com função de ativação RBF	83,59 (1,00)	0,8140 (0,01)
	WA-HA	ELM com função de ativação sigmóide	91,47 (0,94)	0,9032 (0,01)
	WA-HA	Redes <i>Bayesianas</i>	62,78 (5,42)	0,44 (0,08)
	WA-HA	Redes <i>Naive Bayes</i>	57,45 (4,33)	0,36 (0,06)
	WA-HA	Redes MLP	71,91 (7,85)	0,58 (0,12)
	WA-HA	Redes RBF	70,49 (5,06)	0,56 (0,08)
	WA-HA	SVM <i>kernel</i> linear	94,09 (2,68)	0,91 (0,04)
	WA-HA	SVM <i>kernel</i> polinomial grau 2	71,27 (4,86)	0,57 (0,07)
	WA-HA	SVM <i>kernel</i> polinomial grau 3	68,27 (5,15)	0,52 (0,08)
	WA-HA	SVM <i>kernel</i> polinomial grau 4	61,88 (5,31)	0,43 (0,08)
	WA-HA	SVM <i>kernel</i> polinomial grau 5	64,56 (5,27)	0,47 (0,08)
	WA-HA	SVM <i>kernel</i> RBF	82,31	0,73

			(4,17)	(0,06)
	WA-HA	kNN (k = 1)	73,18 (4,97)	0,60 (0,07)
	WA-HA	kNN (k = 3)	73,55 (5,36)	0,60 (0,08)
	WA-HA	kNN (k = 5)	71,69 (5,10)	0,58 (0,08)
	WA-HA	kNN (k = 7)	71,79 (5,33)	0,58 (0,08)
	WA-HA	kNN (k = 9)	71,93 (5,09)	0,58 (0,08)
	WA-HA	J48	75,42 (5,23)	0,63 (0,08)
	WA-HA	<i>Random Tree</i>	70,25 (5,41)	0,55 (0,08)
	WA-HA	<i>Random Forest</i>	80,29 (4,37)	0,70 (0,07)
Todos os tecidos	WA-HA	mELM (dilatação)	65,95 (0,00)	0,6572 (0,00)
	WA-HA	mELM (erosão)	66,42 (0,00)	0,6618 (0,00)
	WA-HA	ELM <i>kernel</i> linear	24,58 (0,21)	0,2429 (0,00)
	WA-HA	ELM <i>kernel</i> RBF	15,79 (0,66)	0,1537 (0,01)
	WA-HA	ELM com função de ativação RBF	50,56 (0,71)	0,5031 (0,01)
	WA-HA	ELM com função de ativação sigmóide	63,99 (0,58)	0,6374 (0,01)
	WA-HA	Redes <i>Bayesianas</i>	33,57 (2,21)	0,28 (0,02)
	WA-HA	Redes <i>Naive Bayes</i>	29,06 (1,98)	0,23 (0,02)
	WA-HA	Redes MLP	36,02 (4,78)	0,30 (0,05)
	WA-HA	Redes RBF	27,00 (2,22)	0,20 (0,02)
	WA-HA	SVM <i>kernel</i> linear	73,79 (2,47)	0,71 (0,03)
	WA-HA	SVM <i>kernel</i> polinomial grau 2	50,69 (2,70)	0,46 (0,03)
	WA-HA	SVM <i>kernel</i> polinomial grau 3	48,27 (2,67)	0,44 (0,03)
	WA-HA	SVM <i>kernel</i> polinomial grau 4	39,94 (2,44)	0,34 (0,03)
	WA-HA	SVM <i>kernel</i> polinomial grau 5	39,60 (2,55)	0,34 (0,03)
	WA-HA	SVM <i>kernel</i> RBF	55,35 (2,61)	0,51 (0,03)
	WA-HA	kNN (k = 1)	42,32 (2,58)	0,37 (0,03)
	WA-HA	kNN (k = 3)	40,99 (2,52)	0,36 (0,03)
	WA-HA	kNN (k = 5)	42,73 (2,75)	0,38 (0,03)
	WA-HA	kNN (k = 7)	43,04 (2,69)	0,38 (0,03)
	WA-HA	kNN (k = 9)	42,56 (2,60)	0,37 (0,03)
	WA-HA	J48	45,24	0,40

			(3,02)	(0,03)
	WA-HA	<i>Random Tree</i>	37,17	0,31
			(2,87)	(0,03)
	WA-HA	<i>Random Forest</i>	52,26	0,48
			(2,72)	(0,03)

Figura 0-6Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido adiposo utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

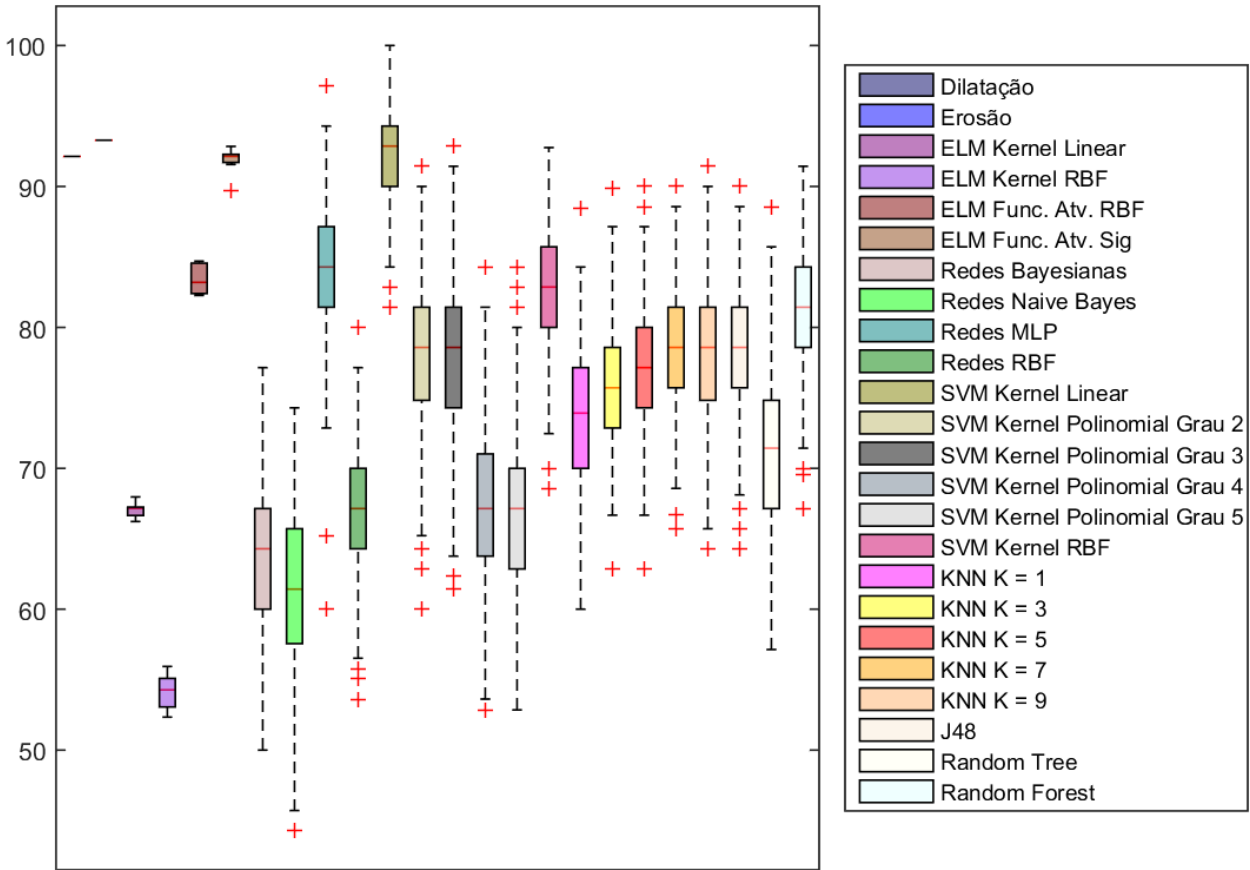


Figura 0-7 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido fibroso utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

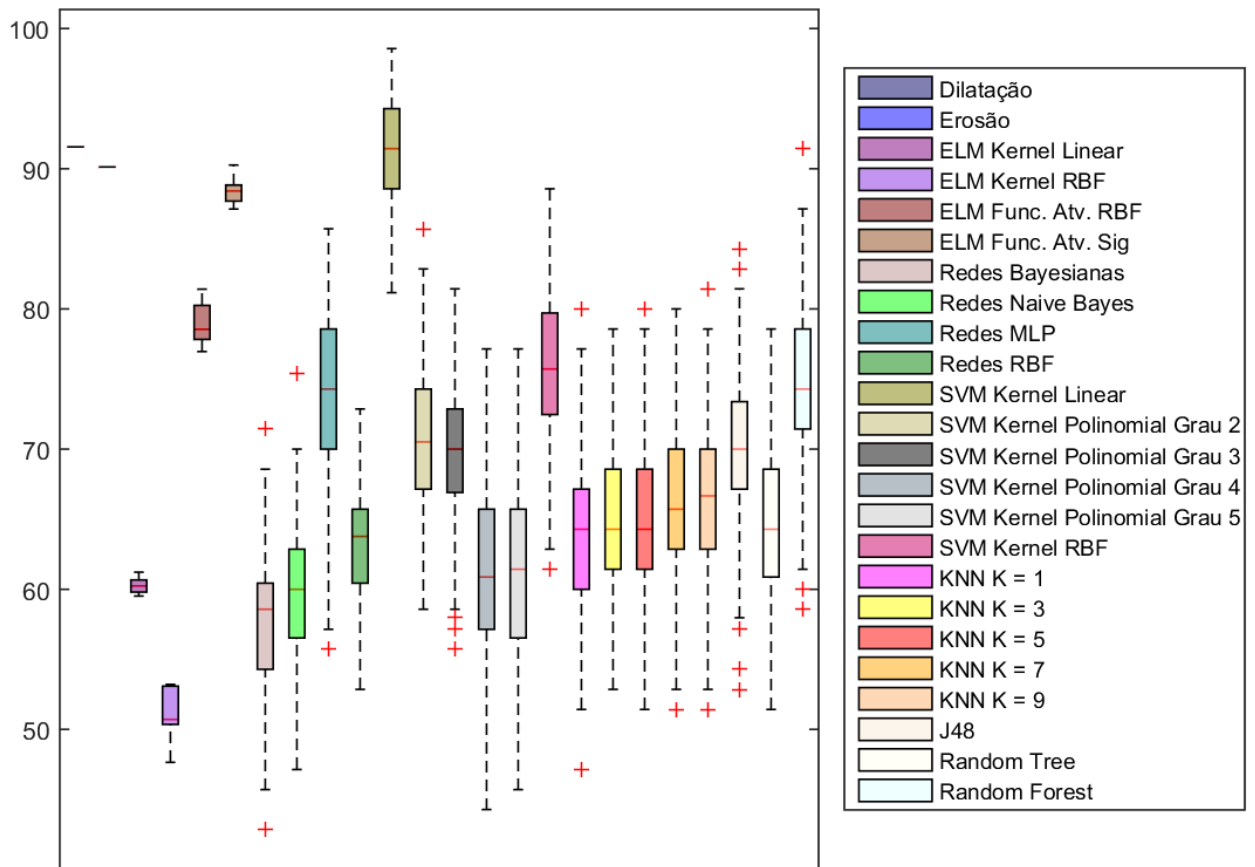


Figura 0-8 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Heterogeneamente denso utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

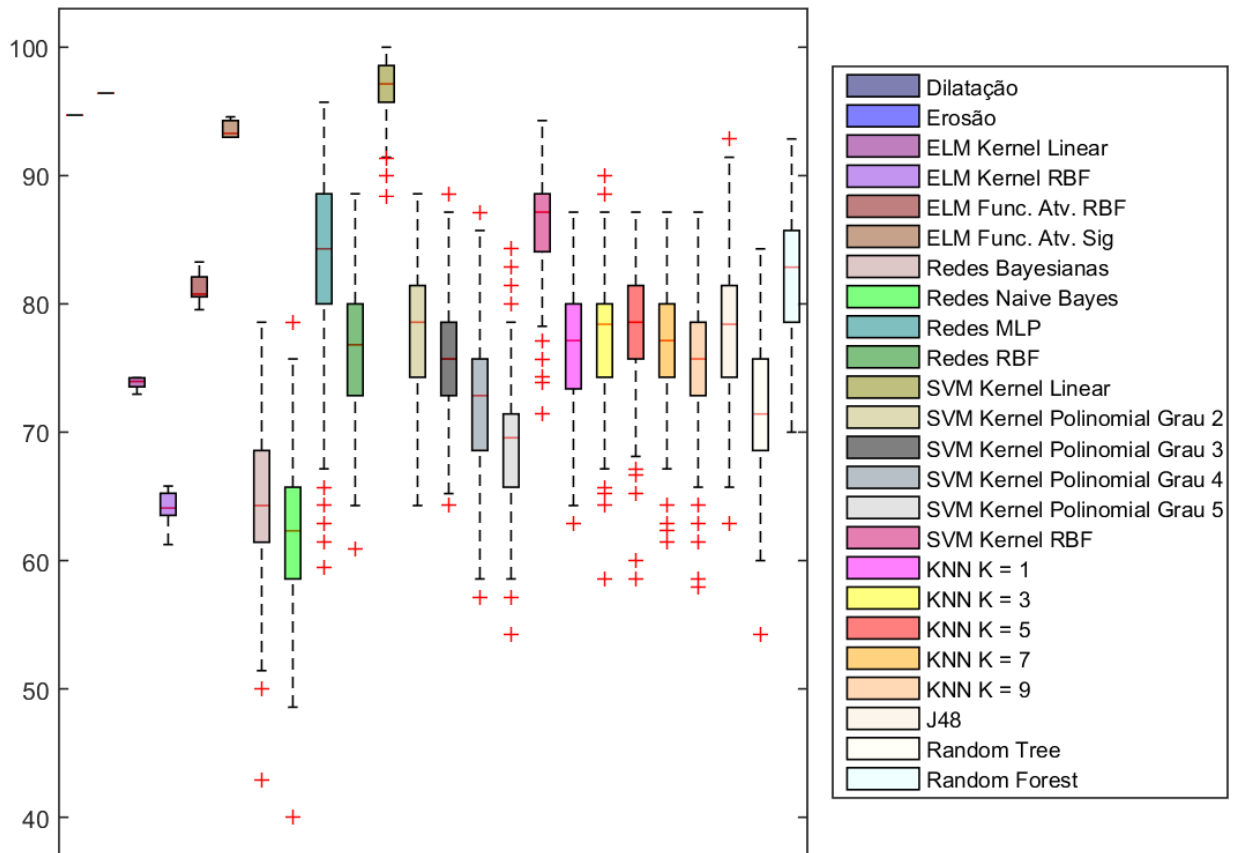


Figura 0-9 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com o tecido Extremamente denso utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

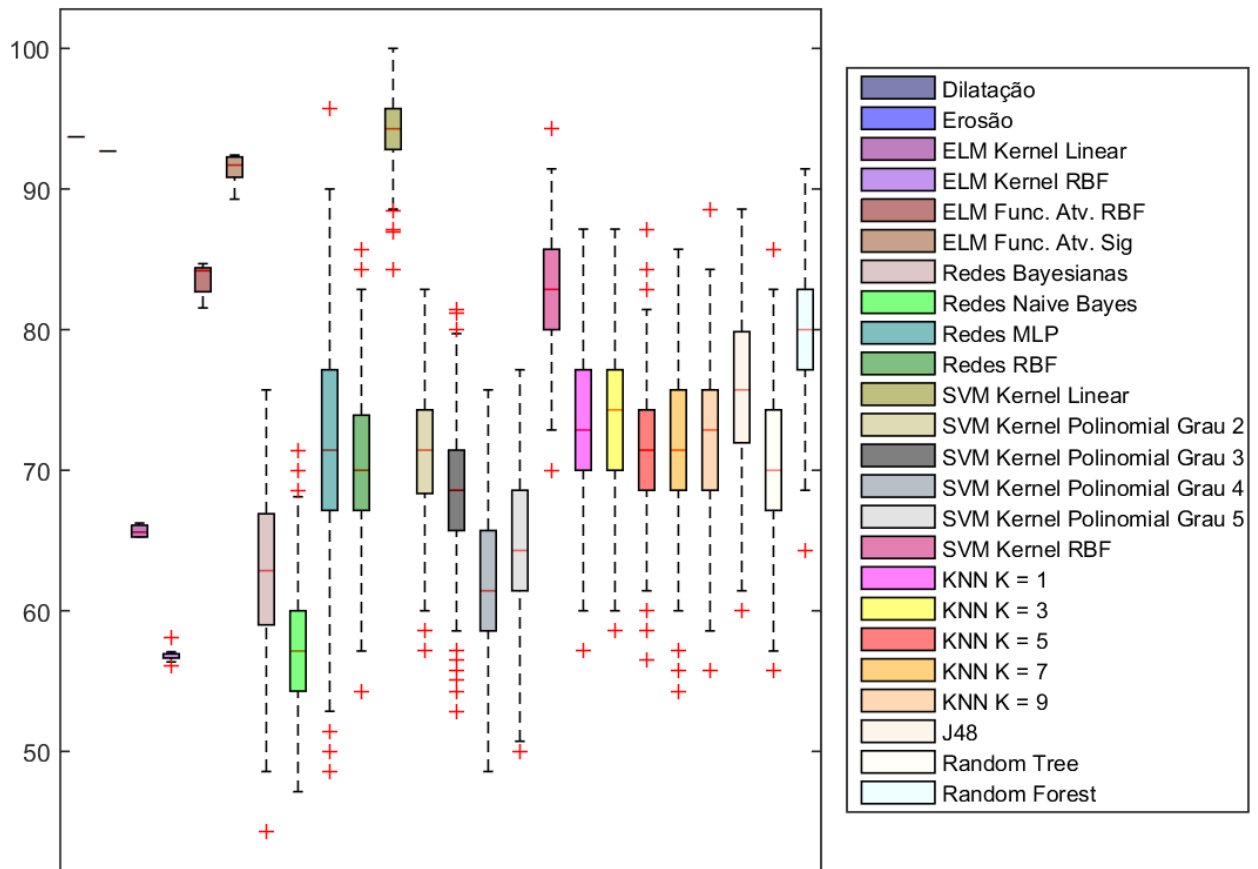


Figura 0-10 Boxplots das taxas de classificações obtidas pelos métodos propostos e os principais classificadores do estado da arte. Testes realizados com todos os tipos de tecidos utilizando *Wavelets* combinados com Atributos de *Haralick* (WA-HA).

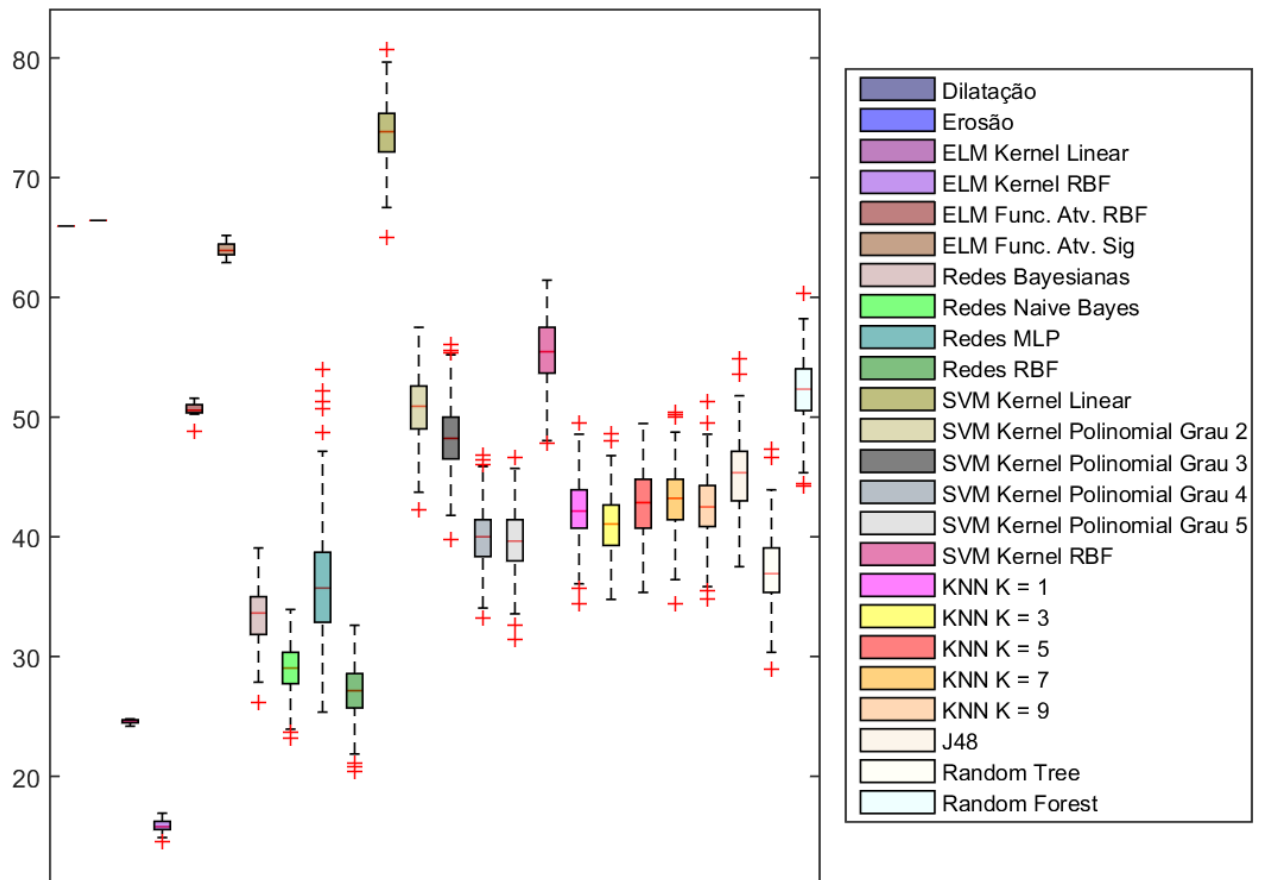


Tabela 0-3 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de Wilcoxon não-paramétrico comparando o método proposto, mELM (dilatação), com os principais classificadores do estado da arte. Nos experimentos foram utilizados Atributos de Haralick.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	mELM (dilatação) vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	7.92389e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	8.255e-114
	mELM (dilatação) vsBayesianas	1	8.76752e-114
	mELM (dilatação) vsNaiveBayes	1	9.08902e-114
	mELM (dilatação) vsMLP	1	9.05226e-101
	mELM (dilatação) vsRBF	1	9.01874e-114
	mELM (dilatação) vsSVM (kernel linear)	1	1.20796e-38
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	4.8099e-105
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	2.47514e-112
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	8.47755e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	8.83377e-114
	mELM (dilatação) vs SVM (kernel RBF)	1	1.29146e-103
	mELM (dilatação) vskNN (k = 1)	1	4.78005e-105
	mELM (dilatação) vs kNN (k = 3)	1	6.02932e-108
	mELM (dilatação) vs kNN (k = 5)	1	7.63911e-111
	mELM (dilatação) vs kNN (k = 7)	1	1.76959e-106
	mELM (dilatação) vs kNN (k = 9)	1	4.91012e-105
	mELM (dilatação) vs J48	1	8.36917e-114
	mELM (dilatação) vsRandomTree	1	9.24778e-114
	mELM (dilatação) vs RandomForest	1	8.75089e-101
II (Fibroso)	mELM (dilatação) vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	6.35092e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.48185e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	3.79431e-77
	mELM (dilatação) vsBayesianas	1	7.73232e-111
	mELM (dilatação) vsNaiveBayes	1	2.25526e-109
	mELM (dilatação) vsMLP	1	2.21265e-83
	mELM (dilatação) vsRBF	1	5.26739e-105
	mELM (dilatação) vsSVM (kernel linear)	0	0.130625
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	1.01274e-74
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	1.27916e-103
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	8.57931e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	8.80274e-114
	mELM (dilatação) vs SVM (kernel RBF)	1	6.91849e-45
	mELM (dilatação) vskNN (k = 1)	1	1.25334e-103
	mELM (dilatação) vs kNN (k = 3)	1	3.93286e-82
	mELM (dilatação) vs kNN (k = 5)	1	2.13045e-78
	mELM (dilatação) vs kNN (k = 7)	1	2.93673e-87
	mELM (dilatação) vs kNN (k = 9)	1	1.40602e-88
	mELM (dilatação) vs J48	1	1.46084e-88
	mELM (dilatação) vsRandomTree	1	7.89984e-111
	mELM (dilatação) vs RandomForest	1	4.76012e-29
III (Heterogeneamente denso)	mELM (dilatação) vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	7.92389e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	8.445e-114
	mELM (dilatação) vsBayesianas	1	4.86123e-43
	mELM (dilatação) vsNaiveBayes	1	1.39032e-73
	mELM (dilatação) vsMLP	1	3.14153e-102

	mELM (dilatação) vsRBF	1	4.78121e-19
	mELM (dilatação) vsSVM (kernel linear)	1	7.22879e-114
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	3.67537e-77
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	9.82269e-57
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	7.98799e-11
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	1.17688e-67
	mELM (dilatação) vs SVM (kernel RBF)	1	2.13345e-78
	mELM (dilatação) vskNN (k = 1)	1	2.79248e-87
	mELM (dilatação) vs kNN (k = 3)	1	1.50436e-92
	mELM (dilatação) vs kNN (k = 5)	1	1.07635e-84
	mELM (dilatação) vs kNN (k = 7)	1	1.56299e-73
	mELM (dilatação) vs kNN (k = 9)	1	2.1761e-78
	mELM (dilatação) vs J48	1	3.25008e-91
	mELM (dilatação) vsRandomTree	1	5.88825e-44
	mELM (dilatação) vs RandomForest	1	2.35391e-112
IV (Extremamente denso)	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	7.92389e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	5.08826e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	5.68561e-10
	mELM (dilatação) vsBayesianas	1	2.49764e-112
	mELM (dilatação) vsNaiveBayes	1	9.43394e-114
	mELM (dilatação) vsMLP	1	9.92397e-75
	mELM (dilatação) vsRBF	1	8.96035e-114
	mELM (dilatação) vsSVM (kernel linear)	1	0.023354
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	8.505e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	8.15893e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	9.60377e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	9.6649e-114
	mELM (dilatação) vs SVM (kernel RBF)	1	1.80938e-106
	mELM (dilatação) vskNN (k = 1)	1	7.97479e-114
	mELM (dilatação) vs kNN (k = 3)	1	5.81734e-108
	mELM (dilatação) vs kNN (k = 5)	1	2.51822e-112
	mELM (dilatação) vs kNN (k = 7)	1	7.63783e-111
	mELM (dilatação) vs kNN (k = 9)	1	2.25361e-109
	mELM (dilatação) vs J48	1	5.22797e-98
	mELM (dilatação) vsRandomTree	1	9.47654e-114
	mELM (dilatação) vs RandomForest	1	2.14569e-78
Todos os tecidos	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	7.92389e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	6.35092e-114
	mELM (dilatação) vsBayesianas	1	1.12072e-113
	mELM (dilatação) vsNaiveBayes	1	1.10871e-113
	mELM (dilatação) vsMLP	1	1.12383e-113
	mELM (dilatação) vsRBF	1	1.10905e-113
	mELM (dilatação) vsSVM (kernel linear)	1	2.05879e-23
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	1.11674e-113
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	1.12198e-113
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	1.11774e-113
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	1.1154e-113
	mELM (dilatação) vs SVM (kernel RBF)	1	1.11947e-113
	mELM (dilatação) vskNN (k = 1)	1	1.11547e-113
	mELM (dilatação) vs kNN (k = 3)	1	1.11747e-113
	mELM (dilatação) vs kNN (k = 5)	1	1.11652e-113
	mELM (dilatação) vs kNN (k = 7)	1	1.11588e-113
	mELM (dilatação) vs kNN (k = 9)	1	1.12234e-113
	mELM (dilatação) vs J48	1	1.12832e-113

	mELM (dilatação) vsRandomTree	1	1.12231e-113
	mELM (dilatação) vs RandomForest	1	1.13074e-100

Tabela 0-4Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de Wilcoxon não-paramétrico comparando o método proposto, mELM (dilatação), com os principais classificadores do estado da arte. Nos experimentos foram utilizandoWaveletscombinados com Atributos de Haralick.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	7.92389e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	6.35092e-114
	mELM (dilatação) vsBayesianas	1	9.34242e-114
	mELM (dilatação) vsNaiveBayes	1	9.70242e-114
	mELM (dilatação) vsMLP	1	3.4692e-91
	mELM (dilatação) vsRBF	1	9.14398e-114
	mELM (dilatação) vsSVM (kernel linear)	1	4.42567e-05
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	9.18529e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	2.716e-112
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	9.40251e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	9.51474e-114
	mELM (dilatação) vs SVM (kernel RBF)	1	7.15621e-111
	mELM (dilatação) vskNN (k = 1)	1	9.16502e-114
	mELM (dilatação) vs kNN (k = 3)	1	8.65249e-114
	mELM (dilatação) vs kNN (k = 5)	1	8.79936e-114
	mELM (dilatação) vs kNN (k = 7)	1	8.27523e-114
	mELM (dilatação) vs kNN (k = 9)	1	8.67894e-114
	mELM (dilatação) vs J48	1	9.00908e-114
	mELM (dilatação) vsRandomTree	1	9.37909e-114
	mELM (dilatação) vs RandomForest	1	8.70697e-114
II (Fibroso)	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	6.35092e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	7.12282e-114
	mELM (dilatação) vsBayesianas	1	8.9078e-114
	mELM (dilatação) vsNaiveBayes	1	8.90474e-114
	mELM (dilatação) vsMLP	1	9.58635e-114
	mELM (dilatação) vsRBF	1	8.38583e-114
	mELM (dilatação) vsSVM (kernel linear)	1	0.00249821
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	9.48336e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	9.17027e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	9.93406e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	9.8073e-114
	mELM (dilatação) vs SVM (kernel RBF)	1	9.36744e-114
	mELM (dilatação) vskNN (k = 1)	1	9.15978e-114
	mELM (dilatação) vs kNN (k = 3)	1	9.57429e-114
	mELM (dilatação) vs kNN (k = 5)	1	9.70869e-114
	mELM (dilatação) vs kNN (k = 7)	1	9.82391e-114
	mELM (dilatação) vs kNN (k = 9)	1	9.48018e-114
	mELM (dilatação) vs J48	1	9.34815e-114
	mELM (dilatação) vsRandomTree	1	9.74862e-114
	mELM (dilatação) vs RandomForest	1	9.21772e-114
III (Heterogeneamente denso)	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	6.35092e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	5.08826e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	4.42632e-114

	mELM (dilatação) vsBayesianas	1	9.57617e-114
	mELM (dilatação) vsNaiveBayes	1	9.02619e-114
	mELM (dilatação) vsMLP	1	2.57576e-109
	mELM (dilatação) vsRBF	1	8.93868e-114
	mELM (dilatação) vsSVM (kernel linear)	1	1.28758e-39
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	9.02294e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	8.9e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	9.26542e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	9.06578e-114
	mELM (dilatação) vs SVM (kernel RBF)	1	8.18502e-114
	mELM (dilatação) vskNN (k = 1)	1	8.83963e-114
	mELM (dilatação) vs kNN (k = 3)	1	8.65625e-114
	mELM (dilatação) vs kNN (k = 5)	1	8.85896e-114
	mELM (dilatação) vs kNN (k = 7)	1	9.18153e-114
	mELM (dilatação) vs kNN (k = 9)	1	9.05007e-114
	mELM (dilatação) vs J48	1	9.11195e-114
	mELM (dilatação) vsRandomTree	1	9.68317e-114
	mELM (dilatação) vs RandomForest	1	8.7007e-114
IV (Extremamente denso)	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	7.92389e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	8.445e-114
	mELM (dilatação) vsBayesianas	1	9.68586e-114
	mELM (dilatação) vsNaiveBayes	1	8.50743e-114
	mELM (dilatação) vsMLP	1	9.42991e-111
	mELM (dilatação) vsRBF	1	9.17207e-114
	mELM (dilatação) vsSVM (kernel linear)	1	5.97642e-07
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	9.10927e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	8.80217e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	9.23704e-114
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	9.34395e-114
	mELM (dilatação) vs SVM (kernel RBF)	1	2.62965e-112
	mELM (dilatação) vskNN (k = 1)	1	9.01689e-114
	mELM (dilatação) vs kNN (k = 3)	1	9.58612e-114
	mELM (dilatação) vs kNN (k = 5)	1	9.3081e-114
	mELM (dilatação) vs kNN (k = 7)	1	9.40813e-114
	mELM (dilatação) vs kNN (k = 9)	1	9.35672e-114
	mELM (dilatação) vs J48	1	9.17402e-114
	mELM (dilatação) vsRandomTree	1	9.49919e-114
	mELM (dilatação) vs RandomForest	1	8.73537e-114
Todos os tecidos	mELM (dilatação)vs mELM (erosão)	1	2.78101e-132
	mELM (dilatação) vsELM (kernel linear)	1	7.92389e-114
	mELM (dilatação) vsELM (kernel RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (dilatação) vsELM (Função atv. sigmoidal)	1	8.445e-114
	mELM (dilatação) vsBayesianas	1	1.10594e-113
	mELM (dilatação) vsNaiveBayes	1	1.0978e-113
	mELM (dilatação) vsMLP	1	1.13273e-113
	mELM (dilatação) vsRBF	1	1.10723e-113
	mELM (dilatação) vsSVM (kernel linear)	1	3.41118e-112
	mELM (dilatação) vsSVM (kernel polinomial grau 2)	1	1.11852e-113
	mELM (dilatação) vs SVM (kernel polinomial grau 3)	1	1.11656e-113
	mELM (dilatação) vs SVM (kernel polinomial grau 4)	1	1.11195e-113
	mELM (dilatação) vs SVM (kernel polinomial grau 5)	1	1.11211e-113
	mELM (dilatação) vs SVM (kernel RBF)	1	1.11589e-113
	mELM (dilatação) vskNN (k = 1)	1	1.11122e-113
	mELM (dilatação) vs kNN (k = 3)	1	1.11367e-113
	mELM (dilatação) vs kNN (k = 5)	1	1.11843e-113

	mELM (dilatação) vs kNN (k = 7)	1	1.11492e-113
	mELM (dilatação) vs kNN (k = 9)	1	1.1134e-113
	mELM (dilatação) vs J48	1	1.12388e-113
	mELM (dilatação) vs RandomTree	1	1.12161e-113
	mELM (dilatação) vs RandomForest	1	1.11821e-113

Tabela 0-5 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de Wilcoxon não-paramétrico comparando o método proposto, mELM (erosão), com os principais classificadores do estado da arte. Nos experimentos foram utilizados Atributos de Haralick.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	mELM (erosão) vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vs ELM (kernel linear)	1	7.92389e-114
	mELM (erosão) vs ELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vs ELM (Função atv. RBF)	1	6.35092e-114
	mELM (erosão) vs ELM (Função atv. sigmoidal)	1	8.255e-114
	mELM (erosão) vs Bayesianas	1	8.76752e-114
	mELM (erosão) vs NaiveBayes	1	9.08902e-114
	mELM (erosão) vs MLP	1	9.05226e-101
	mELM (erosão) vs RBF	1	9.01874e-114
	mELM (erosão) vs SVM (kernel linear)	1	1.20796e-38
	mELM (erosão) vs SVM (kernel polinomial grau 2)	1	4.8099e-105
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	2.47514e-112
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	8.47755e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	8.83377e-114
	mELM (erosão) vs SVM (kernel RBF)	1	1.29146e-103
	mELM (erosão) vs kNN (k = 1)	1	4.78005e-105
	mELM (erosão) vs kNN (k = 3)	1	6.02932e-108
	mELM (erosão) vs kNN (k = 5)	1	7.63911e-111
	mELM (erosão) vs kNN (k = 7)	1	1.76959e-106
	mELM (erosão) vs kNN (k = 9)	1	4.91012e-105
	mELM (erosão) vs J48	1	8.36917e-114
	mELM (erosão) vs RandomTree	1	9.24778e-114
	mELM (erosão) vs RandomForest	1	8.75089e-101
II (Fibroso)	mELM (erosão) vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vs ELM (kernel linear)	1	6.35092e-114
	mELM (erosão) vs ELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vs ELM (Função atv. RBF)	1	6.48185e-114
	mELM (erosão) vs ELM (Função atv. sigmoidal)	1	1.0463e-47
	mELM (erosão) vs Bayesianas	1	7.73232e-111
	mELM (erosão) vs NaiveBayes	1	2.25526e-109
	mELM (erosão) vs MLP	1	2.28379e-78
	mELM (erosão) vs RBF	1	5.26739e-105
	mELM (erosão) vs SVM (kernel linear)	0	0.545508
	mELM (erosão) vs SVM (kernel polinomial grau 2)	1	1.01274e-74
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	1.27916e-103
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	8.57931e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	8.80274e-114
	mELM (erosão) vs SVM (kernel RBF)	1	3.7845e-42
	mELM (erosão) vs kNN (k = 1)	1	1.25334e-103
	mELM (erosão) vs kNN (k = 3)	1	7.13477e-81
	mELM (erosão) vs kNN (k = 5)	1	2.13045e-78
	mELM (erosão) vs kNN (k = 7)	1	1.13313e-84
	mELM (erosão) vs kNN (k = 9)	1	1.10406e-84
	mELM (erosão) vs J48	1	2.96882e-87
	mELM (erosão) vs RandomTree	1	7.89984e-111
	mELM (erosão) vs RandomForest	1	7.14245e-27
III (Heterogeneamente)	mELM (erosão) vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vs ELM (kernel linear)	1	7.92389e-114

denso)	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	9.36891e-47
	mELM (erosão) vsBayesianas	1	2.67896e-112
	mELM (erosão) vsNaiveBayes	1	6.72986e-114
	mELM (erosão) vsMLP	1	5.85853e-108
	mELM (erosão) vsRBF	1	8.60209e-114
	mELM (erosão) vsSVM (kernel linear)	1	3.53302e-12
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	8.05466e-114
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	7.88531e-114
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	7.3991e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	9.02723e-114
	mELM (erosão) vs SVM (kernel RBF)	1	7.22616e-111
	mELM (erosão) vskNN (k = 1)	1	2.39865e-112
	mELM (erosão) vs kNN (k = 3)	1	2.25095e-109
	mELM (erosão) vs kNN (k = 5)	1	7.16535e-111
	mELM (erosão) vs kNN (k = 7)	1	8.07909e-114
	mELM (erosão) vs kNN (k = 9)	1	7.43263e-111
	mELM (erosão) vs J48	1	1.3046e-103
	mELM (erosão) vsRandomTree	1	8.52011e-114
	mELM (erosão) vs RandomForest	1	5.23696e-70
IV (Extremamente denso)	mELM (erosão)vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vsELM (kernel linear)	1	7.92389e-114
	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	5.08826e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	7.12282e-114
	mELM (erosão) vsBayesianas	1	2.49764e-112
	mELM (erosão) vsNaiveBayes	1	9.43394e-114
	mELM (erosão) vsMLP	1	9.92397e-75
	mELM (erosão) vsRBF	1	8.96035e-114
	mELM (erosão) vsSVM (kernel linear)	1	0.023354
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	8.505e-114
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	8.15893e-114
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	9.60377e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	9.6649e-114
	mELM (erosão) vs SVM (kernel RBF)	1	1.80938e-106
	mELM (erosão) vskNN (k = 1)	1	7.97479e-114
	mELM (erosão) vs kNN (k = 3)	1	5.81734e-108
	mELM (erosão) vs kNN (k = 5)	1	2.51822e-112
	mELM (erosão) vs kNN (k = 7)	1	7.63783e-111
	mELM (erosão) vs kNN (k = 9)	1	2.25361e-109
	mELM (erosão) vs J48	1	5.22797e-98
	mELM (erosão) vsRandomTree	1	9.47654e-114
	mELM (erosão) vs RandomForest	1	2.14569e-78
Todos os tecidos	mELM (erosão)vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vsELM (kernel linear)	1	7.92389e-114
	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	6.35092e-114
	mELM (erosão) vsBayesianas	1	1.12072e-113
	mELM (erosão) vsNaiveBayes	1	1.10871e-113
	mELM (erosão) vsMLP	1	1.12383e-113
	mELM (erosão) vsRBF	1	1.10905e-113
	mELM (erosão) vsSVM (kernel linear)	1	0.00410445
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	1.11674e-113
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	1.12198e-113
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	1.11774e-113
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	1.1154e-113
	mELM (erosão) vs SVM (kernel RBF)	1	1.11947e-113

	mELM (erosão) vskNN (k = 1)	1	1.11547e-113
	mELM (erosão) vs kNN (k = 3)	1	1.11747e-113
	mELM (erosão) vs kNN (k = 5)	1	1.11652e-113
	mELM (erosão) vs kNN (k = 7)	1	1.11588e-113
	mELM (erosão) vs kNN (k = 9)	1	1.12234e-113
	mELM (erosão) vs J48	1	1.12832e-113
	mELM (erosão) vsRandomTree	1	1.12231e-113
	mELM (erosão) vs RandomForest	1	2.35596e-106

Tabela 0-6 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de Wilcoxon não-paramétrico comparando o método proposto, mELM (erosão), com os principais classificadores do estado da arte. Nos experimentos foram utilizados *Wavelets* combinados com Atributos de Haralick.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	mELM (erosão)vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vsELM (kernel linear)	1	7.92389e-114
	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	6.35092e-114
	mELM (erosão) vsBayesianas	1	9.34242e-114
	mELM (erosão) vsNaiveBayes	1	9.70242e-114
	mELM (erosão) vsMLP	1	6.89832e-108
	mELM (erosão) vsRBF	1	9.14398e-114
	mELM (erosão) vsSVM (kernel linear)	1	3.44658e-12
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	9.18529e-114
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	8.87179e-114
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	9.40251e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	9.51474e-114
	mELM (erosão) vs SVM (kernel RBF)	1	7.78477e-114
	mELM (erosão) vskNN (k = 1)	1	9.16502e-114
	mELM (erosão) vs kNN (k = 3)	1	8.65249e-114
	mELM (erosão) vs kNN (k = 5)	1	8.79936e-114
	mELM (erosão) vs kNN (k = 7)	1	8.27523e-114
	mELM (erosão) vs kNN (k = 9)	1	8.67894e-114
II (Fibroso)	mELM (erosão) vs J48	1	9.00908e-114
	mELM (erosão) vsRandomTree	1	9.37909e-114
	mELM (erosão) vs RandomForest	1	8.70697e-114
	mELM (erosão)vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vsELM (kernel linear)	1	6.35092e-114
	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	3.38349e-77
	mELM (erosão) vsBayesianas	1	8.9078e-114
	mELM (erosão) vsNaiveBayes	1	8.90474e-114
	mELM (erosão) vsMLP	1	9.58635e-114
	mELM (erosão) vsRBF	1	8.38583e-114
	mELM (erosão) vsSVM (kernel linear)	1	2.22158e-08
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	9.48336e-114
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	9.17027e-114
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	9.93406e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	9.8073e-114
	mELM (erosão) vs SVM (kernel RBF)	1	9.36744e-114
	mELM (erosão) vskNN (k = 1)	1	9.15978e-114
	mELM (erosão) vs kNN (k = 3)	1	9.57429e-114
	mELM (erosão) vs kNN (k = 5)	1	9.70869e-114
	mELM (erosão) vs kNN (k = 7)	1	9.82391e-114
	mELM (erosão) vs kNN (k = 9)	1	9.48018e-114
	mELM (erosão) vs J48	1	9.34815e-114
	mELM (erosão) vsRandomTree	1	9.74862e-114

	mELM (erosão) vs RandomForest	1	2.82047e-112
III (Heterogeneamente denso)	mELM (erosão) vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vsELM (kernel linear)	1	6.35092e-114
	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	5.08826e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	4.42632e-114
	mELM (erosão) vsBayesianas	1	9.57617e-114
	mELM (erosão) vsNaiveBayes	1	9.02619e-114
	mELM (erosão) vsMLP	1	9.64461e-114
	mELM (erosão) vsRBF	1	8.93868e-114
	mELM (erosão) vsSVM (kernel linear)	1	0.0490823
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	9.02294e-114
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	8.9e-114
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	9.26542e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	9.06578e-114
	mELM (erosão) vs SVM (kernel RBF)	1	8.18502e-114
	mELM (erosão) vskNN (k = 1)	1	8.83963e-114
	mELM (erosão) vs kNN (k = 3)	1	8.65625e-114
	mELM (erosão) vs kNN (k = 5)	1	8.85896e-114
	mELM (erosão) vs kNN (k = 7)	1	9.18153e-114
	mELM (erosão) vs kNN (k = 9)	1	9.05007e-114
	mELM (erosão) vs J48	1	9.11195e-114
	mELM (erosão) vsRandomTree	1	9.68317e-114
	mELM (erosão) vs RandomForest	1	8.7007e-114
IV (Extremamente denso)	mELM (erosão) vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vsELM (kernel linear)	1	7.92389e-114
	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	8.445e-114
	mELM (erosão) vsBayesianas	1	9.68586e-114
	mELM (erosão) vsNaiveBayes	1	8.50743e-114
	mELM (erosão) vsMLP	1	9.42991e-111
	mELM (erosão) vsRBF	1	9.17207e-114
	mELM (erosão) vsSVM (kernel linear)	1	3.9234e-32
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	9.10927e-114
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	8.80217e-114
	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	9.23704e-114
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	9.34395e-114
	mELM (erosão) vs SVM (kernel RBF)	1	2.62965e-112
	mELM (erosão) vskNN (k = 1)	1	9.01689e-114
	mELM (erosão) vs kNN (k = 3)	1	9.58612e-114
	mELM (erosão) vs kNN (k = 5)	1	9.3081e-114
	mELM (erosão) vs kNN (k = 7)	1	9.40813e-114
	mELM (erosão) vs kNN (k = 9)	1	9.35672e-114
	mELM (erosão) vs J48	1	9.17402e-114
	mELM (erosão) vsRandomTree	1	9.49919e-114
	mELM (erosão) vs RandomForest	1	8.73537e-114
Todos os tecidos	mELM (erosão) vs mELM (dilatação)	1	2.78101e-132
	mELM (erosão) vsELM (kernel linear)	1	7.92389e-114
	mELM (erosão) vsELM (kernel RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. RBF)	1	6.35092e-114
	mELM (erosão) vsELM (Função atv. sigmoidal)	1	8.445e-114
	mELM (erosão) vsBayesianas	1	1.10594e-113
	mELM (erosão) vsNaiveBayes	1	1.0978e-113
	mELM (erosão) vsMLP	1	1.13273e-113
	mELM (erosão) vsRBF	1	1.10723e-113
	mELM (erosão) vsSVM (kernel linear)	1	3.41118e-112
	mELM (erosão) vsSVM (kernel polinomial grau 2)	1	1.11852e-113
	mELM (erosão) vs SVM (kernel polinomial grau 3)	1	1.11656e-113

	mELM (erosão) vs SVM (kernel polinomial grau 4)	1	1.11195e-113
	mELM (erosão) vs SVM (kernel polinomial grau 5)	1	1.11211e-113
	mELM (erosão) vs SVM (kernel RBF)	1	1.11589e-113
	mELM (erosão) vs kNN (k = 1)	1	1.11122e-113
	mELM (erosão) vs kNN (k = 3)	1	1.11367e-113
	mELM (erosão) vs kNN (k = 5)	1	1.11843e-113
	mELM (erosão) vs kNN (k = 7)	1	1.11492e-113
	mELM (erosão) vs kNN (k = 9)	1	1.1134e-113
	mELM (erosão) vs J48	1	1.12388e-113
	mELM (erosão) vs RandomTree	1	1.12161e-113
	mELM (erosão) vs RandomForest	1	1.11821e-113

Tabela 0-7 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de Wilcoxon não-paramétrico comparando o classificador SVM *kernel* linear com os métodos propostos e com os principais classificadores do estado da arte. Nos experimentos foram utilizando Atributos de Haralick.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	SVM (kernel linear) vs mELM (dilatação)	1	1.20796e-38
	SVM (kernel linear) vs mELM (erosão)	1	1.20796e-38
	SVM (kernel linear) vs ELM (kernel linear)	1	5.8606e-100
	SVM (kernel linear) vs ELM (kernel RBF)	1	4.94513e-100
	SVM (kernel linear) vs ELM (Função atv. RBF)	1	3.46703e-52
	SVM (kernel linear) vs ELM (Função atv. sigmoidal)	1	0.000673966
	SVM (kernel linear) vs Bayesianas	1	3.1974e-85
	SVM (kernel linear) vs NaiveBayes	1	9.32838e-98
	SVM (kernel linear) vs MLP	1	6.1837e-33
	SVM (kernel linear) vs RBF	1	2.00734e-64
	SVM (kernel linear) vs SVM (kernel polinomial grau 2)	1	4.97339e-44
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	2.08018e-57
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	1.91372e-98
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	6.05856e-99
	SVM (kernel linear) vs SVM (kernel RBF)	1	4.88727e-40
	SVM (kernel linear) vs kNN (k = 1)	1	2.53883e-45
	SVM (kernel linear) vs kNN (k = 3)	1	1.26774e-49
	SVM (kernel linear) vs kNN (k = 5)	1	4.35564e-61
	SVM (kernel linear) vs kNN (k = 7)	1	3.92851e-54
	SVM (kernel linear) vs kNN (k = 9)	1	2.19084e-56
II (Fibroso)	SVM (kernel linear) vs J48	1	2.18698e-79
	SVM (kernel linear) vs RandomTree	1	5.42596e-87
	SVM (kernel linear) vs RandomForest	1	6.68774e-33
	SVM (kernel linear) vs mELM (dilatação)	0	0.130625
	SVM (kernel linear) vs mELM (erosão)	0	0.545508
	SVM (kernel linear) vs ELM (kernel linear)	1	4.78759e-100
	SVM (kernel linear) vs ELM (kernel RBF)	1	4.78759e-100
	SVM (kernel linear) vs ELM (Função atv. RBF)	1	1.34355e-15
	SVM (kernel linear) vs ELM (Função atv. sigmoidal)	0	0.520811
	SVM (kernel linear) vs Bayesianas	1	9.9709e-78
	SVM (kernel linear) vs NaiveBayes	1	1.86053e-85
	SVM (kernel linear) vs MLP	1	7.64874e-48
	SVM (kernel linear) vs RBF	1	3.32157e-76
	SVM (kernel linear) vs SVM (kernel polinomial grau 2)	1	5.06216e-40
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	1.32672e-71
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	6.74487e-97
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	6.15149e-100
	SVM (kernel linear) vs SVM (kernel RBF)	1	1.12873e-23
	SVM (kernel linear) vs kNN (k = 1)	1	2.15774e-70
	SVM (kernel linear) vs kNN (k = 3)	1	5.69692e-50
	SVM (kernel linear) vs kNN (k = 5)	1	1.4278e-50
	SVM (kernel linear) vs kNN (k = 7)	1	1.00957e-48

	SVM (kernel linear) vs kNN (k = 9)	1	6.58336e-48
	SVM (kernel linear) vs J48	1	2.26362e-54
	SVM (kernel linear) vs RandomTree	1	2.13171e-73
	SVM (kernel linear) vs RandomForest	1	5.12183e-13
III (Heterogeneamente denso)	SVM (kernel linear) vs mELM (dilatação)	1	7.22879e-114
	SVM (kernel linear) vs mELM (erosão)	1	3.53302e-12
	SVM (kernel linear) vsELM (kernel linear)	1	5.59601e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	4.72154e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	1.76533e-28
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	1	2.15737e-08
	SVM (kernel linear) vs Bayesianas	1	5.98161e-83
	SVM (kernel linear) vs NaiveBayes	1	5.08684e-100
	SVM (kernel linear) vs MLP	1	6.45258e-56
	SVM (kernel linear) vs RBF	1	1.31455e-92
	SVM (kernel linear) vs SVM (kernel polinomial grau 2)	1	2.46544e-75
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	4.35973e-83
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	1.57293e-97
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	6.53537e-100
	SVM (kernel linear) vs SVM (kernel RBF)	1	3.42977e-69
	SVM (kernel linear) vs kNN (k = 1)	1	6.21008e-67
	SVM (kernel linear) vs kNN (k = 3)	1	1.48467e-58
	SVM (kernel linear) vs kNN (k = 5)	1	8.5584e-70
	SVM (kernel linear) vs kNN (k = 7)	1	4.08083e-75
	SVM (kernel linear) vs kNN (k = 9)	1	2.1596e-74
	SVM (kernel linear) vs J48	1	1.93255e-54
	SVM (kernel linear) vs RandomTree	1	1.63611e-86
	SVM (kernel linear) vs RandomForest	1	2.38435e-19
IV (Extremamente denso)	SVM (kernel linear) vs mELM (dilatação)	1	0.023354
	SVM (kernel linear) vs mELM (erosão)	1	0.023354
	SVM (kernel linear) vsELM (kernel linear)	1	5.66936e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	4.78353e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	2.90685e-68
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	0	0.291397
	SVM (kernel linear) vs Bayesianas	1	2.07849e-91
	SVM (kernel linear) vs NaiveBayes	1	6.48199e-100
	SVM (kernel linear) vs MLP	1	3.24848e-37
	SVM (kernel linear) vs RBF	1	1.4894e-98
	SVM (kernel linear) vs SVM (kernel polinomial grau 2)	1	5.45118e-87
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	7.35121e-95
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	1.58221e-99
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	7.52166e-100
	SVM (kernel linear) vs SVM (kernel RBF)	1	1.07609e-70
	SVM (kernel linear) vs kNN (k = 1)	1	2.62015e-86
	SVM (kernel linear) vs kNN (k = 3)	1	3.56348e-82
	SVM (kernel linear) vs kNN (k = 5)	1	1.97338e-84
	SVM (kernel linear) vs kNN (k = 7)	1	2.16378e-84
	SVM (kernel linear) vs kNN (k = 9)	1	1.03359e-82
	SVM (kernel linear) vs J48	1	1.3205e-58
	SVM (kernel linear) vs RandomTree	1	1.89551e-91
	SVM (kernel linear) vs RandomForest	1	1.4167e-41
Todos os tecidos	SVM (kernel linear) vs mELM (dilatação)	1	2.05879e-23
	SVM (kernel linear) vs mELM (erosão)	1	0.00410445
	SVM (kernel linear) vsELM (kernel linear)	1	7.81526e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	6.59731e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	6.59731e-100
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	1	3.99562e-99
	SVM (kernel linear) vs Bayesianas	1	1.01916e-99
	SVM (kernel linear) vs NaiveBayes	1	1.01079e-99
	SVM (kernel linear) vs MLP	1	7.08302e-99

	SVM (kernel linear) vsRBF	1	1.01102e-99
	SVM (kernel linear) vsSVM (kernel polinomial grau 2)	1	1.09735e-98
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	1.49152e-99
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	1.01709e-99
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	1.01545e-99
	SVM (kernel linear) vs SVM (kernel RBF)	1	2.95781e-97
	SVM (kernel linear) vskNN (k = 1)	1	1.05698e-99
	SVM (kernel linear) vs kNN (k = 3)	1	1.14661e-99
	SVM (kernel linear) vs kNN (k = 5)	1	1.41364e-99
	SVM (kernel linear) vs kNN (k = 7)	1	2.1387e-99
	SVM (kernel linear) vs kNN (k = 9)	1	2.4344e-99
	SVM (kernel linear) vs J48	1	1.74049e-99
	SVM (kernel linear) vsRandomTree	1	1.02027e-99
	SVM (kernel linear) vs RandomForest	1	1.18018e-85

Tabela 0-8 Resultados expandidos, referentes à Tabela 5-2 e a Tabela 5-4. Teste de Wilcoxon não-paramétrico comparando o classificador SVM *kernel* linear com os métodos propostos e com os principais classificadores do estado da arte. Nos experimentos foram utilizados *Wavelets* combinados com Atributos de Haralick.

<i>Tecido</i>	<i>Comparação</i>	<i>Hipótese</i>	<i>Valor p</i>
I (Adiposo)	SVM (kernel linear) vs mELM (dilatação)	1	4.42567e-05
	SVM (kernel linear) vs mELM (erosão)	1	3.44658e-12
	SVM (kernel linear) vsELM (kernel linear)	1	4.58823e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	3.87009e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	2.80105e-97
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	1	3.87009e-100
	SVM (kernel linear) vsBayesianas	1	5.208e-100
	SVM (kernel linear) vsNaiveBayes	1	5.36175e-100
	SVM (kernel linear) vsMLP	1	8.38221e-72
	SVM (kernel linear) vsRBF	1	5.12268e-100
	SVM (kernel linear) vsSVM (kernel polinomial grau 2)	1	5.08142e-97
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	3.46676e-97
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	5.61427e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	5.55338e-100
	SVM (kernel linear) vs SVM (kernel RBF)	1	1.02985e-90
	SVM (kernel linear) vskNN (k = 1)	1	8.6825e-100
	SVM (kernel linear) vs kNN (k = 3)	1	3.92535e-99
	SVM (kernel linear) vs kNN (k = 5)	1	8.51192e-98
	SVM (kernel linear) vs kNN (k = 7)	1	7.57461e-98
	SVM (kernel linear) vs kNN (k = 9)	1	1.25052e-97
	SVM (kernel linear) vs J48	1	1.29777e-97
II (Fibroso)	SVM (kernel linear) vsRandomTree	1	8.95522e-100
	SVM (kernel linear) vs RandomForest	1	1.88934e-92
	SVM (kernel linear) vs mELM (dilatação)	1	0.00249821
	SVM (kernel linear) vs mELM (erosão)	1	2.22158e-08
	SVM (kernel linear) vsELM (kernel linear)	1	4.77393e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	4.77393e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	6.45132e-100
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	1	5.49878e-35
	SVM (kernel linear) vsBayesianas	1	6.19011e-100
	SVM (kernel linear) vsNaiveBayes	1	6.18847e-100
	SVM (kernel linear) vsMLP	1	2.08615e-99
	SVM (kernel linear) vsRBF	1	5.90962e-100
	SVM (kernel linear) vsSVM (kernel polinomial grau 2)	1	8.59195e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	6.45794e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	6.73083e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	6.66476e-100
	SVM (kernel linear) vs SVM (kernel RBF)	1	2.98789e-98
	SVM (kernel linear) vskNN (k = 1)	1	6.32415e-100

	SVM (kernel linear) vs kNN (k = 3)	1	6.54281e-100
	SVM (kernel linear) vs kNN (k = 5)	1	6.61324e-100
	SVM (kernel linear) vs kNN (k = 7)	1	6.67343e-100
	SVM (kernel linear) vs kNN (k = 9)	1	6.5588e-100
	SVM (kernel linear) vs J48	1	7.9674e-100
	SVM (kernel linear) vs RandomTree	1	6.63412e-100
	SVM (kernel linear) vs RandomForest	1	1.20361e-98
III (Heterogeneamente denso)	SVM (kernel linear) vs mELM (dilatação)	1	1.28758e-39
	SVM (kernel linear) vs mELM (erosão)	1	0.0490823
	SVM (kernel linear) vsELM (kernel linear)	1	2.79694e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	2.79694e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	2.35741e-100
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	1	1.51582e-58
	SVM (kernel linear) vs Bayesianas	1	3.83959e-100
	SVM (kernel linear) vs NaiveBayes	1	3.66829e-100
	SVM (kernel linear) vs MLP	1	6.49456e-97
	SVM (kernel linear) vs RBF	1	3.6776e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 2)	1	3.89518e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	3.66532e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	3.74309e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	3.6807e-100
	SVM (kernel linear) vs SVM (kernel RBF)	1	2.60019e-97
	SVM (kernel linear) vs kNN (k = 1)	1	3.60965e-100
	SVM (kernel linear) vs kNN (k = 3)	1	3.77237e-100
	SVM (kernel linear) vs kNN (k = 5)	1	3.61574e-100
	SVM (kernel linear) vs kNN (k = 7)	1	3.71691e-100
	SVM (kernel linear) vs kNN (k = 9)	1	3.67578e-100
	SVM (kernel linear) vs J48	1	4.8647e-100
	SVM (kernel linear) vs RandomTree	1	3.87265e-100
	SVM (kernel linear) vs RandomForest	1	8.78293e-100
IV (Extremamente denso)	SVM (kernel linear) vs mELM (dilatação)	1	5.97642e-07
	SVM (kernel linear) vs mELM (erosão)	1	3.9234e-32
	SVM (kernel linear) vsELM (kernel linear)	1	4.48045e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	3.77904e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	9.32922e-100
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	1	2.11752e-43
	SVM (kernel linear) vs Bayesianas	1	5.22909e-100
	SVM (kernel linear) vs NaiveBayes	1	4.73226e-100
	SVM (kernel linear) vs MLP	1	7.8361e-98
	SVM (kernel linear) vs RBF	1	5.09032e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 2)	1	4.98784e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	4.85793e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	5.0416e-100
	SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	5.08644e-100
	SVM (kernel linear) vs SVM (kernel RBF)	1	6.98716e-97
	SVM (kernel linear) vs kNN (k = 1)	1	5.20354e-100
	SVM (kernel linear) vs kNN (k = 3)	1	5.56413e-100
	SVM (kernel linear) vs kNN (k = 5)	1	5.22639e-100
	SVM (kernel linear) vs kNN (k = 7)	1	5.21699e-100
	SVM (kernel linear) vs kNN (k = 9)	1	5.6555e-100
	SVM (kernel linear) vs J48	1	6.03559e-100
	SVM (kernel linear) vs RandomTree	1	5.20332e-100
	SVM (kernel linear) vs RandomForest	1	8.87738e-99
Todos os tecidos	SVM (kernel linear) vs mELM (dilatação)	1	3.41118e-112
	SVM (kernel linear) vs mELM (erosão)	1	3.41118e-112
	SVM (kernel linear) vsELM (kernel linear)	1	7.81884e-100
	SVM (kernel linear) vsELM (kernel RBF)	1	6.60034e-100
	SVM (kernel linear) vsELM (Função atv. RBF)	1	6.60034e-100
	SVM (kernel linear) vsELM (Função atv. sigmoidal)	1	1.07582e-99

SVM (kernel linear) vsBayesianas	1	1.00931e-99
SVM (kernel linear) vsNaiveBayes	1	1.00362e-99
SVM (kernel linear) vsMLP	1	1.02799e-99
SVM (kernel linear) vsRBF	1	1.01021e-99
SVM (kernel linear) vsSVM (kernel polinomial grau 2)	1	1.01809e-99
SVM (kernel linear) vs SVM (kernel polinomial grau 3)	1	1.01673e-99
SVM (kernel linear) vs SVM (kernel polinomial grau 4)	1	1.01351e-99
SVM (kernel linear) vs SVM (kernel polinomial grau 5)	1	1.01363e-99
SVM (kernel linear) vs SVM (kernel RBF)	1	1.01626e-99
SVM (kernel linear) vskNN (k = 1)	1	1.01301e-99
SVM (kernel linear) vs kNN (k = 3)	1	1.01471e-99
SVM (kernel linear) vs kNN (k = 5)	1	1.01803e-99
SVM (kernel linear) vs kNN (k = 7)	1	1.01559e-99
SVM (kernel linear) vs kNN (k = 9)	1	1.01453e-99
SVM (kernel linear) vs J48	1	1.02183e-99
SVM (kernel linear) vsRandomTree	1	1.02025e-99
SVM (kernel linear) vs RandomForest	1	1.01788e-99