

MARCELO DORNBUSCH LOPES

**CLASSIFICAÇÃO DE COLORAÇÃO IMUNO-HISTOQUÍMICA
EM IMAGENS MICROSCÓPICAS
COMBINANDO COR E TEXTURA COMO DESCRIPTOR DE
CARACTERÍSTICAS**

São José (SC), Fevereiro de 2015



UNIVERSIDADE DO VALE DO ITAJAÍ
CURSO DE MESTRADO ACADÊMICO EM
COMPUTAÇÃO APLICADA

**CLASSIFICAÇÃO DE COLORAÇÃO IMUNO-HISTOQUÍMICA
EM IMAGENS MICROSCÓPICAS**
EMPREGANDO ATRIBUTOS DE COR E TEXTURA COMO DESCRIPTOR DE
CARACTERÍSTICAS

por

Marcelo Dornbusch Lopes

Dissertação apresentada como requisito parcial à
obtenção do grau de Mestre em Computação
Aplicada.
Orientador: Eros Comunello, Dr. rer. nat.

São José (SC), Fevereiro de 2015

FOLHA DE APROVAÇÃO

Esta página é reservada para inclusão da folha de assinaturas, a ser disponibilizada pela Secretaria do Curso para coleta da assinatura no ato da defesa.

Dedico este trabalho aos meus pais, Rudi Pereira Lopes e Mara Maria Dornbusch Lopes, que me proporcionaram essa existência. Bem como aos demais familiares que me aturaram até aqui.

Página opcional reservada para uma epígrafe, a qual deve ser escrita em itálico, alinhada à direita e posicionada na base da página. Exclua esta página se não for incluir nenhuma epígrafe.

AGRADECIMENTOS

Gostaria de agradecer meu orientador Eros Comunello, e ao meu coorientador de fato (que só não foi por direito, devido a detalhes burocráticos) Antonio C. Sobieranski, ambos em diversos momentos enxergaram na minha pessoa, mais do que eu mesmo conseguia.

Gostaria de agradecer ao Luís Eduardo R. Carvalho, que gentilmente cedeu parte de seu trabalho, o código do classificador PMD otimizado, fato providencial para a conclusão desse trabalho.

Gostaria de agradecer aos amigos de programa, Marlon C. Domenech, Leonardo R. P. Rauta e Maykon Chagas, que por mais atribulados que se encontrassem, sempre despendiam tempo para ouvir minhas lamúrias, contribuindo com críticas construtivas.

Gostaria de agradecer ao amigo de laboratório Nicholas S. Meschke, que embarcou nesse projeto, decidindo empenhar parte de seu tempo no desenvolvimento da extensão dos experimentos, defendidos na dissertação.

CLASSIFICAÇÃO DE COLORAÇÃO IMUNO-HISTOQUÍMICA EM IMAGENS MICROSCÓPICAS

COMBINANDO COR E TEXTURA COMO DESCRIPTOR DE CARACTERÍSTICAS

Marcelo Dornbusch Lopes

Fevereiro / 2015

Orientador: Eros Comunello, Dr. rer. nat.

Área de Concentração: Computação Aplicada

Linha de Pesquisa: Inteligência Aplicada

Palavras-chave: Análise de Imagens, Reconhecimento de Padrões, Imuno-histoquímica.

Número de páginas: 128

RESUMO

No campo de visão computacional existem diversas tarefas de contínuo interesse e importância científica, tais como, o reconhecimento de padrões, análise de textura e descrição de características. Estas tarefas podem ser empregadas na análise imuno-histoquímica, guiada por computador, que consiste na avaliação da expressão de proteínas, a partir da relação antígeno/anticorpo, em imagens de células. Desta forma, o presente trabalho teve por objetivo melhorar o desempenho de classificação (considerando medidas de acerto e erro), no reconhecimento de coloração em imagens de imuno-histoquímica, utilizando a abordagem de aprendizado supervisionada que emprega a distância polinomial de Mahalanobis. A melhora foi atingida a partir da definição de um descritor de características híbrido, combinando cor e textura, baseado no método Local Binary Pattern. O descritor é formado por valores de cor no espaço cromático RGB, mais um grupo de medidas estatísticas tomadas em janelas de amostragem com dimensões distintas, proporcionando um descritor 12-dimensional para cada pixel. Os experimentos realizados compreenderam uma análise de componentes principais e avaliações no desempenho de classificação. Os testes de hipóteses realizados avaliaram quatro medidas de acerto, Sensibilidade, Especificidade, Precisão e Acurácia, bem como três medidas de erro, Taxa de Falso Positivo, Taxa de Falso Negativo e Taxa de Falsa Descoberta. Com os resultados encontrados, percebeu-se que a adição de informação de textura no descritor tornou o processo de classificação mais específico e restritivo (nas previsões positivas). Foi possível observar uma melhora de desempenho nas medidas de Especificidade, Precisão (relação das previsões positivas contra anotações positivas), Taxa de Falso Positivo e Taxa de Falsa Descoberta.

IMMUNOHISTOCHEMICAL STAINING CLASSIFICATION IN MICROSCOPIC IMAGES

BY COMBINING COLOR AND TEXTURE AS A FEATURE DESCRIPTOR

Marcelo Dornbusch Lopes

February / 2015

Advisor: Eros Comunello, Dr. rer. nat.

Area of Concentration: Applied Computer Science

Research Line: Applied Intelligence

Keywords: Image Analysis, Pattern Recognition, Immunohistochemical.

Number of pages: 128

ABSTRACT

In the field of computer vision, there are several scientific tasks of continuing interest and scientific importance, such as pattern recognition, texture analysis and feature description. These tasks can be used in computer-aided immunohistochemical analysis, which consists in evaluating the expression of proteins from the antigen-antibody complex, in cell or tissue sample images. The present research aims to improve the classification performance (considering hits and error measurement), based on a supervised learning approach that uses polynomial Mahalanobis distance, when applied to staining recognition in immunohistochemistry images. The improvement was achieved based on the definition of a descriptor with hybrid characteristics, combining color and texture based on the Local Binary Pattern method. The descriptor is formed by the RGB color values, plus a group of statistical measurements taken in sampling windows with different dimensions, providing a 12-dimensional descriptor for each pixel. The experiments included a principal component analysis and performance evaluations of the classification. The hypothesis tests performed evaluated four measurements of success: Sensitivity, Specificity, Precision and Accuracy, as well as three error measurements: False Positive Rate, False Negative Rate and False Discovery Rate. Based on the results, it was found that the addition of texture information in the descriptor made the classification process more specific and restrictive (in terms of predicting a pixel as positive). Improved performance was also observed in Specificity, Precision (ratio of positive predictions against positive observation), False Positive Rate and False Discovery Rate.

LISTA DE FIGURAS

Figura 1. Classificação da pesquisa realizada.....	26
Figura 2. (a) imagem com níveis de cinza; (b) notação de matriz de coocorrência e (c) direções angulares	33
Figura 3. (a) 25 combinações 1D que geram as máscaras de Laws e (b) exemplo da máscara L5E5	34
Figura 4. Um píxel central $x_{0,0}$ e sua vizinhança espaçada por um raio r	36
Figura 5. Ilustração do cálculo do valor $LBP_{(8,1)}$ para um píxel central de valor 5.....	36
Figura 6. Fluxo no processo de construção de classificadores	38
Figura 7. Comparação entre DE (a), DM (b), DPM ordem 2 (c) e DPM ordem 8 (d).....	41
Figura 8. Etapas básicas no processo de IHQ	45
Figura 9- Imagens de IHQ: a) Carcinoma de pulmão com DAB; b) Amostra de pele com AEC	46
Figura 10. (a) fluxo de processamento; (b) imagem original e (c) imagem segmentada.....	51
Figura 11. (a) máscaras, medidas estatísticas e desempenho e (b) imagem de núcleos reconhecidos	53
Figura 12. Visão geral do fluxo de processamento da abordagem proposta.....	54
Figura 13. (a) esboço da abordagem proposta e (b) plotagem após clusterização	56
Figura 14. (a) extração da linha de vale e (b) superfície topográfica da imagem	57
Figura 15. (a) estrutura da rede neural empregada e (b) resultados de classificação, imagem original, classe HLN, classe DAB e classe PRD (respectivamente)	59
Figura 16. (a) visão geral da abordagem proposta e (b) resultado de reconhecimento, imagem original e segmentada.....	60
Figura 17. Exemplos de janelas de amostragem: a) dimensão 3×3 ; e b) dimensão 9×9	65
Figura 18. Visão geral da etapa de treinamento, extração dos descritores (fase i-a)	67
Figura 19. Visão geral da fase de treinamento, calibração do modelo, etapa (i-b).....	67
Figura 20. Visão geral da etapa de classificação	68
Figura 21. Imagens resultantes da filtragem por textura: a) Imagem original; b) Textura LBP tradicional e c) Textura LBP baseado em variância	70
Figura 22. Diagrama de Classe simples das implementações efetuadas e utilizadas	71
Figura 23. Ferramenta para seleção do conjunto de treinamento	72
Figura 24. Resultados do processo de classificação e contingência	72
Figura 25. Vista aproximada do mapa temático	73
Figura 26. a) Imagem LCH-IHQ; e b) Imagem anotada respectiva.....	77
Figura 27.a) Imagem original; b) Imagem anotada e c) Resultado do filtro de textura LBP	88
Figura 28. Mapa temático resultante para cada descritor: a) \mathbf{d}_3 , b) \mathbf{d}_{23} e c) \mathbf{d}_{12}	88
Figura 29. Imagem original utilizada no experimento e sua respectiva imagem de anotação.....	95
Figura 30. Exemplo de imagem resultante do processo de classificação, para o descritor RGB e para o descritor LBP reduzido, respectivamente	95
Figura 31. Exemplo de imagem resultante da avaliação de contingência, para o descritor RGB e para o descritor LBP Reduzido, respectivamente	95
Figura 32. String de busca IEEE	119
Figura 33. String Science Direct	120
Figura 34. String ACM	121
Figura 35. Sequência das etapas de extração de características, treinamento e classificação	122
Figura 36. Comparação das etapas de treinamento e classificação da DPM vs. DE	123
Figura 37. Representação dos tecidos cancerosos no HPA	125

Figura 38. Lâminas de Microarranjo usadas no HPA	126
Figura 39. Metadados para as imagens na categoria linha da LCH, onde cada imagem diz respeito a um núcleo de TMAs.....	127
Figura 40. Detalhe de metadados para uma imagem	128

LISTA DE QUADROS

Quadro 1. Termos comumente utilizados no processo de classificação	38
Quadro 2. Identificação dos sete trabalhos relacionados	50
Quadro 3. Comparação entre os trabalhos relacionados e a proposta.....	63
Quadro 4. Medidas computadas para os experimentos de desempenho de classificação.....	82

LISTA DE TABELAS

Tabela 1. Tabela de contingência ou matriz de confusão	42
Tabela 2. Resultado de pesquisa e seleção de trabalhos para análise	47
Tabela 3 Reunião de técnicas e campos de aplicação encontrados na revisão sistemática	48
Tabela 4. Contribuição de cada variável para as 10 primeiras componentes (%)	81
Tabela 5. Resultados com o descritor d_3 (RGB), limiar 230.....	84
Tabela 6. Resultados com o descritor d_3 (RGB), limiar 240.....	84
Tabela 7. Resultados com o descritor d_3 (RGB), limiar 250.....	84
Tabela 8. Resultados com o descritor d_{23} (RGB-LBP completo), limiar 230.....	85
Tabela 9. Resultados com o descritor d_{23} (RGB-LBP completo), limiar 240.....	85
Tabela 10. Resultados com o descritor d_{23} (RGB-LBP completo), limiar 250.....	85
Tabela 11. Resultados com o descritor d_{12} (RGB-LBP reduzido), limiar 230.....	85
Tabela 12. Resultados com o descritor d_{12} (RGB-LBP reduzido), limiar 240.....	86
Tabela 13. Resultados com o descritor d_{12} (RGB-LBP reduzido), limiar 250.....	86
Tabela 14. Resultados de erro com o descritor d_{23} (RGB-LBP completo), limiares 230/250	87
Tabela 15. Resultados de erro com o descritor d_{12} (RGB-LBP reduzido), limiares 230/250.....	87
Tabela 16. Maiores valores de ESP, SEN, PRE e ACC entre os três descritores.....	87
Tabela 17. Menores valores de T-FP, T-FD e T-FN entre os três descritores.....	87
Tabela 18. Resultados do Experimento principal (média para 23 imagens).....	90
Tabela 19. Estatísticas descritivas do teste de hipótese para as medidas de acerto	97
Tabela 20. Resultados do teste de hipótese para as medidas de acerto.....	97
Tabela 21. Estatísticas descritivas do teste de hipótese para as medidas de erro	99
Tabela 22. Resultados do teste de hipótese para as medidas de erro	99

LISTA DE GRÁFICOS

Gráfico 1. Distribuição da porcentagem de marcações positivas nas imagens.....	78
Gráfico 2. Resultado do experimento e PCA.....	80
Gráfico 3. Correlação entre P e VP, descriptores d_3 e d_{23} , parâmetros GP 05/ CO -1.0 / LI 250	89
Gráfico 4. Correlação entre N e VN, descriptores d_3 e d_{23} , parâmetros GP 05 / CO -1.0 / LI 250	89
Gráfico 5. Gráfico de Caixa Sensibilidade.....	91
Gráfico 6. Gráfico de Caixa Especificidade.....	91
Gráfico 7. Gráfico de Caixa Precisão.....	91
Gráfico 8. Gráfico de Caixa Predição Negativa.....	91
Gráfico 9. Gráfico de Caixa F1	92
Gráfico 10. Gráfico de Caixa Acurácia.....	92
Gráfico 11. Gráfico de Caixa Correlação de Matthews	92
Gráfico 12. Gráfico de Caixa T-FN	92
Gráfico 13. Gráfico de Caixa T-FP	92
Gráfico 14. Gráfico de Caixa T-FD	93
Gráfico 15. Correlação entre anotações e predições positivas.....	93
Gráfico 16. Correlação entre anotações e predições negativas	94

LISTA DE ABREVIATURAS E SIGLAS

ACC	Acurácia
AEC	Amino-etil-carbazol
AUC	Area Under a Curve
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CB	Centroblastos
CE	Conjunto de Experimentos
CMYK	Cyan Magenta Yellow Black
CN	Contraste
CO	Corte
COb	Campo de Observação
COTB	Computer on The Beach
CP	Combinação de Parâmetros
CP	Combinação de Parâmetros
C-PI	Classificação Piloto
C-PR	Classificação Principal
CSV	Comma Separated Value
CT	Conjunto de Treinamento
DA	Distância Absoluta
DAB	Diaminobenzidina
DC	Descriptor de Características
DE	Distância Euclidiana
DH	Distância de Hausdoff
DM	Distância de Mahalanobis
DPM	Distância Polinomial de Mahalanobis
EC	Espaço de Características
ESP	Especificidade
F1	F1 Escore
FD	Função Discriminante
FN	Falso Negativo
FP	Falso Positivo
FR	Fisher Ratio
GM	Geometric Margin
GP	Grau Polinomial
GT	Ground Truth
HPA	Human Protein Atlas
HSV	Hue Saturation Value
IAD	Iteração de Avaliação de Desempenho
IF	Imagem Filtrada
IHQ	Imuno-histoquímica
JA	Janela de Amostragem
k-NN	k-Nearest Neighbor's
LBP	Local Binary Pattern
LCH	Linhas de Células Humanas
LF	Linfoma Folicular
LI	Limiar

MCC	Coeficiente de Correlação de Matthews
MPL	Multilayer Perceptron
N	Negativo
P	Positivo
PCA	Principal Component Analysis
PDI	Processamento Digital de Imagens
PRE	Precisão
RGB	Red Green Blue
ROC	Receiver Operating Characteristic
RP	Reconhecimento de Padrões
RSL	Revisão Sistemática da Literatura
SEM	Sensibilidade
SM	Dispersão
SVM	Support Vector Machine
T-FD	Taxa de Detecção Falsa
T-FN	Taxa de Falso Negativo
T-FP	Taxa de Falso Positivo
TMA	Tissue Micro Array
T-PN	Taxa de Predição Negativa
VC	Visão Computacional
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

SUMÁRIO

1 INTRODUÇÃO	18
1.1 PROBLEMA DE PESQUISA.....	21
1.1.1 Solução Proposta	22
1.1.2 Delimitação de Escopo	22
1.1.3 Justificativa.....	23
1.2 OBJETIVOS	24
1.2.1 Objetivo Geral	25
1.2.2 Objetivos Específicos	25
1.3 METODOLOGIA.....	25
1.3.1 Metodologia da Pesquisa	26
1.3.2 Procedimentos Metodológicos.....	27
1.4 ESTRUTURA DA DISSERTAÇÃO.....	28
2 FUNDAMENTAÇÃO TEÓRICA.....	29
2.1 DESCRIPTORES DE CARACTERÍTICAS	29
2.2 DESCRIPTORES DE TEXTURA	30
2.2.1 Matriz de coocorrência	32
2.2.2 Texturas de Laws	34
2.2.3 Padrão Binário Local.....	35
2.3 CLASSIFICAÇÃO DE PADRÕES	37
2.4 POLINOMIAL MAHALANOBIS.....	40
2.5 AVALIAÇÃO DE CLASSIFICADORES.....	41
2.6 IMUNO-HISTOQUÍMICA	44
3 REVISÃO BIBLIOGRÁFICA	47
3.1 REVISÃO SISTEMÁTICA DA LITERATURA	47
3.2 TRABALHOS RELACIONADOS	49
3.2.1 A Semi-automatic Image Analysis Tool For Biomarker Detection In Immunohistochemistry Analysis.....	50
3.2.2 Colour-Texture Based Image Analysis Method For Assessing The Hormone Receptors Status In Breast Tissue Sections.....	52
3.2.3 Classification And Immunohistochemical Scoring Of Breast Tissue Microarray Spots.....	54
3.2.4 Detection of Follicles From IHQ-Stained Slides Of Follicular Lymphoma Using Iterative Watershed	55
3.2.5 Supervised Learning-based Cell Image Segmentation For P53 Immunohistochemistry	56
3.2.6 Stain Unmixing In Bright Field Multiplexed Immunohistochemistry	58

3.2.7 Learning a Nonlinear Color Distance Metric For The Identification Of Skin Immunohistochemical Staining.....	59
3.3 ANÁLISE COMPARATIVA.....	61
3.4 CONSIDERAÇÕES	62
4 DESENVOLVIMENTO.....	64
4.1 PROPOSTA DO DESCRIPTOR.....	64
4.1.1 Descritor Baseado no Método LPB	64
4.2 MÉTODO DE CLASSIFICAÇÃO	66
4.2.1 Etapa i: Treinamento.....	66
4.2.2 Etapa ii: Classificação.....	68
4.3 DESCRIÇÃO DA IMPLEMENTAÇÃO DO PROTÓTIPO	69
4.4 DESCRIÇÃO DOS EXPERIMENTOS	73
4.4.1 Experimento de Dimensionalidade.....	74
4.4.2 Experimento de Desempenho.....	74
4.4.3 Experimento de Variância.....	74
4.5 DESCRIÇÃO DO CONJUNTO DE DADOS	75
4.5.1 Conjunto de Dados.....	77
4.6 CONSIDERAÇÕES	78
5 ANÁLISE EXPERIMENTAL.....	79
5.1 ANÁLISE DE DIMENSIONALIDADE.....	79
5.2 DESEMPENHOS DE CLASSIFICAÇÃO	81
5.2.1 Resultado do Experimento Piloto	83
5.2.2 Resultados do Experimento Principal.....	89
5.3 TESTES DE HIPÓTESES	96
5.3.1 Teste de Medidas Assertivas	96
5.3.2 Teste para Medidas de Erro.....	98
5.4 CONSIDERAÇÕES	100
6 CONCLUSÃO	102
6.1 CONTRIBUIÇÕES DA DISSERTAÇÃO	106
6.2 TRABALHOS FUTUROS	107
REFERÊNCIAS BIBLIOGRÁFICAS	108
GLOSSÁRIO.....	115
APÊNDICE A – PROTOCOLO DE BUSCA	116
APÊNDICE B – DIAGRAMA DE SEQUÊNCIA	122
ANEXO A – DPM VS. DISTÂNCIA EUCLIDIANA	123
ANEXO B – ILUSTRAÇÃO DE TMAS USADOS NO HPA.....	125
ANEXO C – METADADOS HPA LINHAS DE CÉLULAS.....	127

1 INTRODUÇÃO

O problema do reconhecimento de padrões (RP) concerne na descoberta de regularidades em conjunto de dados, empregando algoritmos computacionais. As regularidades identificadas são usadas na tomada de decisão, como a classificação de dados em diferentes categorias de classes. Uma abordagem consistente para este fim é adotar alguma técnica de aprendizado de máquina, na qual um número de amostras N $\{x_1, x_2 \dots x_N\}$ é definido como Conjunto de Treinamento (CT). O CT é então empregado no ajuste de um modelo adaptativo para a classificação, em abordagens supervisionadas (BISHOP, 2006).

Frequentemente, no campo de RP, o termo ‘característica’ é utilizado para denotar um descriptor. As classes agrupam padrões de uma mesma família que se correlacionam por alguma propriedade comum. Normalmente são representadas como $\omega_1, \omega_2, \dots, \omega_w$, em que W é o número de classes. Comumente, na literatura, três abordagens são empregadas para representação de padrões: (i) vetores, para descrições quantitativas; (ii) *strings* e (iii) árvores, para descrições estruturais (GONZALEZ; WOODS, 2010).

A identificação de células, em uma imagem histológica, como saudáveis ou anormais; o controle de qualidade em linhas de produção; a categorização de imagens de sensoriamento remoto são exemplos de classificação, uma atividade essencial na etapa inicial de processos de seleção ou tomada de decisão. No campo do processamento de imagens, a classificação consiste em identificar propriedades, padrões e estruturas características em uma imagem, a fim de categorizá-las (ou a própria imagem) em classes específicas. As técnicas de classificação podem ser divididas em duas abordagens: (i) supervisionadas e (ii) não supervisionadas (SOLOMON; BRECKON, 2013).

Na área médica, tipicamente o diagnóstico de câncer é feito pelo patologista, através de uma inspeção visual da morfologia e cromaticidade de amostras de tecidos. Uma atividade laboriosa e sujeita a erro, em que a efetividade do método depende diretamente da atenção e conhecimentos do patologista. O desenvolvimento contínuo em tecnologias relacionadas à bioimagem, especialmente na microscopia, vem estabelecendo técnicas guiadas por computador, para avaliação de imagens biológicas, como uma forma efetiva para extração de informações clínicas e funcionais, a partir de amostras de tecido. (MOUELHI *et al.*, 2013; CATALDO; FICARRA; MACII, 2012)

O desenvolvimento de tecnologias voltadas a sistemas para diagnóstico médico guiado por computador se tornou um dos principais temas de pesquisa em imagens médicas e radiodiagnóstico, uma vez que cresce o consenso, entre especialistas da área de análise de imagens e patologia, sobre a importância da análise quantitativa de imagens patológicas. Essas análises envolvem tanto o ponto de vista do diagnóstico clínico, quanto à área de pesquisa, proporcionando o entendimento dos mecanismos biológicos das doenças (GURCAN *et al.*, 2009).

Imuno-histoquímica (IHQ) é um procedimento aplicado no diagnóstico patológico, desde a década de setenta, amplamente utilizada como uma ferramenta para a detecção de proteínas em amostras de tecido, através da aplicação de anticorpos específicos. Permite o monitoramento e a quantificação da expressão de proteínas, com fins de diagnóstico e prognóstico. Os anticorpos são marcados com uma coloração específica, e a intensidade da referida expressão está relacionada com a intensidade da coloração alcançada no processo (LIU; WANG; FANG, 2006; WEMMERT *et al.* 2013; FICARRA; MACII, 2006).

O procedimento de imuno-histoquímica é utilizado na busca por抗ígenos em amostras de células ou tecidos, que vão dos aminoácidos e proteínas até agentes infecciosos e populações celulares específicas. O processo é conduzido em dois momentos: (i) preparação das lâminas e os estágios da reação (contraste de coloração empregando cromógeno); (ii) interpretação e quantificação da expressão de coloração obtida (PHAM *et al.*, 2007; MATOS *et al.*, 2010).

As abordagens tradicionais, para a análise nos padrões de coloração em imagens de histopatologia, tais como, observação direta por um grupo de especialistas ou empregando análise estereológica¹, constituem processos laboriosos e subjetivos. Por outro lado, soluções que empregam uso do computador, para a análise de imagens, proporcionam uma quantificação eficiente, objetiva e de fácil reprodução (PRASAD; PRABHU, 2011).

Abordagens computacionais especificamente projetadas para mensurar, de forma automatizada, a extensão e o grau de expressões de imagens de imuno-histoquímica abrangem um pequeno grupo de biomarcadores e tecidos. Métodos genéricos normalmente são baseados em

¹ Procedimento de sobrepor uma retícula sobre a lâmina a ser avaliada, contando visualmente a ocorrência do padrão alvo em cada intersecção da retícula, usando o microscópio.

algum procedimento de aprendizagem, no qual características dos biomarcadores são modeladas em um método semi-supervisionado. Segundo Kothari *et al.* (2011), abordagens supervisionadas têm mostrado melhor desempenho, se comparadas a métodos não supervisionados.

Sistemas para patologia que possam ser amplamente utilizados, tanto no campo clínico quanto na pesquisa, precisam fornecer soluções de aquisição, análise compreensiva e gerenciamento de imagens. Em Yagi (2007) é apresentada uma avaliação de nove soluções comerciais com este propósito. Em Prasad e Prabhu (2011) se encontra uma revisão de ferramentas computacionais, empregadas por pesquisadores, para avaliação de imagens de imuno-histoquímica.

Especificamente no processo de análise compreensiva, a tarefa de classificação é uma etapa desafiadora, na qual conjuntos de itens são agrupados com base em uma ou mais características inerentes (p. ex.² cor, textura, forma, filtragem no domínio de frequência, dentre outras), a partir do treinamento de um grupo de itens previamente anotados. A acurácia dos resultados está diretamente ligada ao tamanho do conjunto de treinamento, bem como da variabilidade requerida para a generalização do padrão alvo (ZHENG; XUE, 2009).

Este trabalho de pesquisa abrange conceitos na computação relativos a segmentação, descrição, classificação e mensuração de padrões. A segmentação consiste em organizar e dividir regiões da imagem com mesmo significado semântico na cena. A classificação, neste contexto, tem a finalidade de reconhecer a presença ou ausência do biomarcador em questão. Diversos métodos encontrados na literatura atendem a esta tarefa, como k-NN (k-Nearest Neighbor's) e classificadores Bayesianos, este último muito usado para atender demandas na área médica (RUSSELL; NORVING, 1994; SINGH; JINDAL, 2012).

Diante do contexto apresentado, esta pesquisa almeja proporcionar uma contribuição ao que tange a descrição e classificação de padrões IHQ, em imagens microscópicas de amostras biológicas de forma automatizada, aplicando para tanto técnicas de Processamento Digital de Imagens (PDI) e Visão Computacional (VC). Como descritores de características, a serem utilizados na classificação, serão empregados atributos de cor e textura.

² por exemplo (p. ex.)

1.1 PROBLEMA DE PESQUISA

No campo de PDI e VC, uma importante área de pesquisa está relacionada à segmentação de imagens. Devido a sua grande disseminação e variada gama de aplicações, a segmentação traz contribuições relevantes para áreas como medicina, sensoriamento remoto e recuperação de imagens por conteúdo. O principal objetivo da segmentação de imagens consiste em dividir a imagem em conjuntos disjuntos, agrupados por alguma característica como cor ou textura (ZUVA *et al.*, 2011).

Segundo Yang, Zhang e Wang (2012), a segmentação de imagens continua sendo um importante, entretanto complexo, problema de pesquisa. A crescente demanda em tarefas de análise de imagens, no tocante ao resultado da segmentação, introduz a necessidade da aplicação de diversas técnicas, para aumentar a acurácia deste resultado.

Os recentes avanços na microscopia digital são guiados, não somente pelas inovações nas áreas de engenharia, ciência da computação, ótica e da manufatura de precisão, mas também pelos avanços nas ciências químicas e biológicas. A teoria relacionada à microscopia é complexa, e o consenso entre as avaliações de dados é que não são perfeitos. O desafio consiste em sintetizar modelos teóricos e evidências empíricas a fim de resolver um problema específico, tendo em vista que sistemas genéricos, para análise de imagens, capazes de automaticamente aprender e identificar as informações desejadas não existem (BOVIK, 2009).

Sobieranski *et al.* (2009) apresentam uma abordagem, para identificação de coloração imuno-histoquímica em imagens microscópicas de amostra de pele, fundamentada em um processo de aprendizado definido como Distância Polinomial de Mahalanobis (DPM), com base na projeção polinomial da Distância de Mahalanobis (DM), para a construção de uma função de similaridade.

O método DPM permite a classificação de padrões não lineares de IHQ em imagens microscópicas de histopatologia. A abordagem apresentada por Sobieranski *et al.*(2009) emprega atributos de cor, no modelo cromático Vermelho, Verde e Azul (*red, green e blue – RGB*), como Descritores de Características (DC), para o aprendizado da métrica de discriminação.

Portanto, o problema de pesquisa abordado neste trabalho está relacionado à descrição e classificação automatizada da expressão de biomarcadores em imagens microscópicas de imuno-histoquímica. Um biomarcador consiste em qualquer entidade, passível de ser mensurada e indica

uma condição normal ou patológica, permitindo diagnóstico, prognóstico, planejamento no tratamento e pesquisa de doenças. A partir deste problema foi possível formular a seguinte pergunta de pesquisa:

- De que forma é possível aumentar o desempenho na classificação (considerando taxas de acerto e erro) da abordagem computacional DPM, que considera atributos de cor como descriptor de características, na análise de biomarcadores em imagens de imuno-histoquímica?

1.1.1 Solução Proposta

A solução proposta neste trabalho teve por objetivo aprimorar o desempenho de classificação do método DPM (considerando taxas de acerto e erro), apresentado por Sobieranski *et al.* (2009). O referido aprimoramento foi alcançado através da incorporação de atributos de textura ao descriptor de características, usado como entrada para o método de classificação. Diante da contextualização e definição do problema de pesquisa até aqui apresentados foi possível formular a seguinte hipótese de pesquisa:

- Combinando métricas de textura e um grupo de medidas estatísticas, com a finalidade de definir um descriptor de características híbrido formado por atributos de cor e textura, é possível aprimorar o desempenho de classificação do método DPM (considerando taxas de acerto e erro).

1.1.2 Delimitação de Escopo

Segundo Kayser *et al.* (2009), sistemas, para patologia virtual inteligente, requerem uma série de ferramentas e etapas padronizadas, a fim de serem interconectadas entre si, independentes de requisitos futuros. Estas ferramentas e etapas tratam desde a aquisição de imagens até o diagnóstico automatizado, e são enumeradas a seguir: (i) normalização das imagens adquiridas; (ii) pré-processamento e identificação de características; (iii) amostragem de imagens; (iv) algoritmos para análise de textura; (v) algoritmos para segmentação, identificação e mensuração de características; (vi) extração e análise de estruturas; (vii) análise estatística e sistemas de classificação; e (viii) monitoramento contínuo para o reajuste dos procedimentos de classificação.

O escopo do presente trabalho abordou essencialmente os aspectos relativos às definições (iv), (v) e (vii) (segmentação, classificação e análise estatística e de textura), definindo, implementando e avaliando uma abordagem supervisionada, em que uma etapa de treinamento é executada, para classificação e mensuração automatizada de imagens patológicas de imuno-histoquímica. O principal campo investigado foi o desempenho de classificação (taxas de acerto e erro), considerando a adição de componentes de textura, no descritor de características.

As etapas de implementação foram efetuadas utilizando a linguagem de programação C++ e a biblioteca para processamento de imagens, OpenCV³, apresentada por Bradski (2000) e Bradski e Kaehler (2008). As atividades de análises estatísticas foram efetuadas com o software de planilha eletrônica Excel⁴ e uma extensão XLSTAT⁵.

As demais etapas (i, ii, iii, vi e viii) não foram abordadas, considerando que não estava no escopo deste trabalho desenvolver um sistema completo. As atividades de modelagem e implementação contemplaram os requisitos mínimos, para permitir a execução dos testes planejados, a fim de comprovar a eficácia do descritor de características e do método de classificação empregados, quando aplicados na classificação de padrões de coloração IHQ. Não foi objetivo modelar e desenvolver uma ferramenta completa de diagnóstico clínico, com telas de usuário ou conduzir avaliações no sentido ergonômico e de usabilidade.

1.1.3 Justificativa

O termo telepatologia define a prática do diagnóstico à distância, empregando tecnologias de informação e comunicação. Desta forma, pacientes e centros médicos, mesmo de localidades remotas, têm acesso a laudos e diagnósticos de patologistas especializados (BARR *et al.*, 2004). Considerada um referencial da patologia moderna, a telepatologia alia conhecimento de informática médica e patologia digital a tecnologias de comunicação, proporcionando diversas modalidades da telemedicina: telediagnóstico, teleconsulta e tele-educação (BORGES *et al.*, 2005).

³ <http://opencv.org/>

⁴Excel - <http://office.microsoft.com/pt-br/excel/>

⁵XLSTAT - <http://www.xlstat.com/en/>

No campo da informática médica, diversas técnicas de processamento de imagens vêm sendo utilizadas na análise de microscopia. Processos laboriosos e subjetivos de diagnósticos podem ser aperfeiçoados, proporcionando a patologistas e biólogos a extração de informação quantitativa, a partir deste tipo de imagem, tanto no campo da citologia (estudo das células), quanto da histologia (estudo anatômico das estruturas dos tecidos) (LIU; WANG; FANG, 2006).

Outros fatores que contribuem para justificar a motivação e viabilidade da abordagem proposta são os resultados alcançados em por Sobieranski *et al.*(2011), onde nos quais a mesma métrica de aprendizado supervisionado foi utilizada para guiar o processo de segmentação por agrupamento de regiões, baseado na funcional de Mumford-Shah, e ainda, os resultados apresentados por Carvalho *et al.* (2014), em que a mesma métrica de aprendizado supervisionado é aplicada a um método de segmentação de imagens, baseado em uma abordagem versão que combina crescimento e fusão de regiões (Color Structure Code). Ambos os trabalhos fazem alusão à segmentação de propósito geral, diferindo ainda desta proposta que emprega somente atributos de cor, como descritor de características.

No trabalho de Aguiar (2012), outra abordagem semelhante é apresentada, empregando a mesma técnica de classificação supervisionada e aplicando atributos de cor e textura como descritor de características. No referido trabalho, o método para a geração da filtragem de textura foi baseado em filtros de Gabor, com janela de amostragem fixa (3×3). A área de aplicação foi a classificação de vegetação, em imagens de sensoriamento remoto, diferindo ainda da presente proposta, no método de filtragem de textura, no tamanho fixo da janela de amostragem e nas medidas estatísticas a serem computadas nestas janelas.

A rotulação de imagens histológicas, entre regiões cancerígenas ou não, é uma tarefa crítica, no diagnóstico de câncer. É importante também, o agrupamento destes resultados em classes distintas. Abordagens supervisionadas existentes requerem anotações manuais detalhadas a respeito dos padrões em questão (XU *et al.*, 2014).

1.2 OBJETIVOS

Diante da argumentação apresentada até este ponto, esta seção apresenta o objetivo geral do trabalho, bem como os cinco objetivos específicos propostos, a fim de concretizar a pesquisa.

1.2.1 Objetivo Geral

Demonstrar a melhora nas taxas de acerto e erro de classificação do método que emprega a distância polinomial de Mahalanobis, comparando o descritor de características que considera cor em relação ao descritor combinando cor e textura.

1.2.2 Objetivos Específicos

1. Definir o método, para gerar as respostas de textura;
2. Definir as medidas estatísticas, para compor o descritor de características;
3. Definir um descritor de características, combinando cor e textura;
4. Avaliar o desempenho de classificação do descritor de características composto com atributos de cor; e
5. Avaliar o desempenho de classificação do descritor de característica composto por cor e textura.

1.3 METODOLOGIA

Segundo Wazlawick (2009), as diferentes subáreas da computação empregam diversos estilos de pesquisa científica que podem ser caracterizados em três tipos básicos. O presente trabalho concentrou esforços na modalidade empírica. Os três tipos citados são descritos a seguir:

- Pesquisa formal: emprega lógica formal para elaboração de uma teoria e sua prova;
- Pesquisa Empírica: emprega métodos estatísticos para avaliação da proposta comparada a métodos do estado da arte; e
- Pesquisa Exploratória: emprega análises qualitativas, sem resultados estatísticos. Esta modalidade está diretamente relacionada à capacidade de argumentação do pesquisador.

O trabalho assumiu este estilo de classificação, no contexto da computação, quando conduziu a execução de experimentos quantitativos com uma análise estatística dos resultados, bem como a avaliação da abordagem proposta em comparação com resultados apresentados em trabalhos correlatos e aceitos pela comunidade científica.

1.3.1 Metodologia da Pesquisa

Formalmente a pesquisa conduzida neste trabalho se enquadrou no método hipotético-dedutivo. Tal classificação foi atingida partindo da definição do problema, passando pela pergunta, até a formulação da hipótese de pesquisa. Aliada a proposta de uma experimentação e avaliação quantitativa da ferramenta, na Figura 1 é apresentada a caracterização da pesquisa desenvolvida.

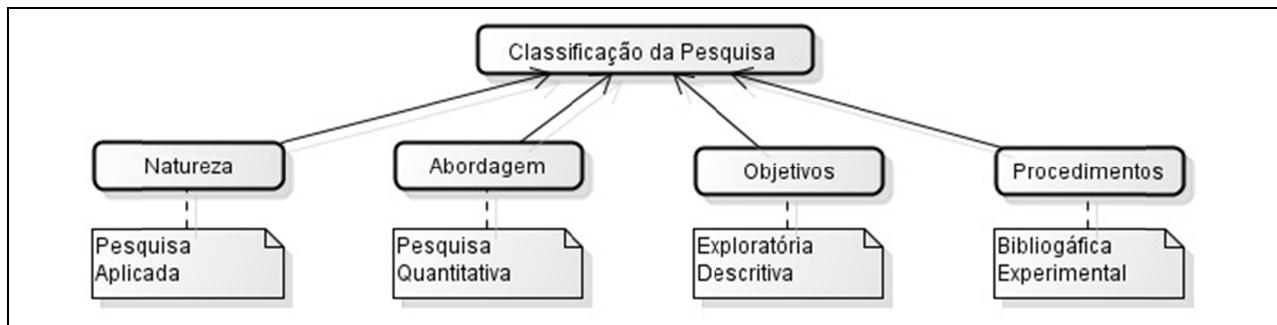


Figura 1. Classificação da pesquisa realizada

Fonte: Adaptado de Wazlawick (2009).

Como pode ser observado na Figura 1, o trabalho assumiu a característica a partir da sua natureza, como pesquisa aplicada, pelo fato de tratar um problema específico, relacionado à análise de imagens médicas de imuno-histoquímica; a abordagem quantitativa, ao executar experimentações e avaliações, empregando métodos estatísticos estabelecidos na literatura.

Ainda na Figura 1 é possível identificar a caracterização do trabalho, segundo seus objetivos, como uma pesquisa exploratória, uma vez que propõe o levantamento bibliográfico e apresenta ainda características de pesquisa descritiva, ao avaliar a relação de causa e efeito entre as variáveis do processo, tais como parâmetros dos algoritmos envolvidos, componentes dos DC e desempenho de classificação (taxas de acerto e erros).

Consequentemente, os procedimentos técnicos adotados para concretizar o trabalho proposto foram a pesquisa bibliográfica e a pesquisa experimental. Tais procedimentos incorporaram um grupo de atividades desempenhadas a fim de atingir os objetivos específicos estabelecidos na Subseção 1.2.2

A referida pesquisa bibliográfica foi realizada previamente, o que permitiu o refinamento da hipótese de pesquisa e adequada contextualização do problema, de acordo com o campo de

aplicação. A modalidade escolhida para efetuar a revisão bibliográfica e levantamento dos trabalhos correlatos foi a Revisão Sistemática da Literatura (RSL).

Segundo Sampaio e Mancini (2006), uma RSL é metódica, explícita e passível de reprodução por terceiros. Métodos sistematizados são atuais e mais adequados para eliminar viés e proporcionar uma análise objetiva do tema abordado.

1.3.2 Procedimentos Metodológicos

Com a finalidade de atender os objetivos específicos baseados nos procedimentos bibliográfico e experimental, o trabalho foi dividido em cinco atividades: Pesquisa Bibliográfica; Proposta; Modelagem; Desenvolvimento e Experimentos. As referidas atividades são apresentadas a seguir:

- **i – Pesquisa Bibliográfica:** esta atividade contemplou a execução da revisão sistemática da literatura sobre métodos de classificação e métricas de avaliação relacionadas a análise de imagens de imuno-histoquímica, bem como o levantamento bibliográfico sobre métodos, para descrição de textura. Esta atividade contribuiu para fornecer o embasamento teórico necessário à conclusão dos três primeiros objetivos específicos, os resultados são apresentados no Capítulo 2 e no Capítulo 3.
- **ii – Proposta:** esta atividade contemplou o atendimento dos objetivos específicos 1, 2 e 3. Foi apresentado o descritor de características proposto, descrevendo o método de filtragem de textura escolhido e as medidas estatísticas a serem tomadas sobre esta filtragem. A definição do método de classificação não linear, baseado em uma abordagem de aprendizado supervisionado, também foi executada nesta atividade. Descrições detalhadas destas definições são apresentadas nas Seções 4.1 e 4.2;
- **iii – Modelagem:** esta atividade contemplou o levantamento de requisitos, a modelagem das etapas de processamento envolvidas, bem como a criação do diagrama de classe. O objetivo foi gerar a documentação para guiar o processo de desenvolvimento do protótipo, que permitiu efetuar a atividade de experimentos; e
- **iv – Desenvolvimento:** esta atividade contemplou a implementação dos protótipos que permitiram a avaliação do descritor de características definido. Uma descrição detalhada desta atividade é apresentada na Seção 4 .

- **v – Experimentos:** esta atividade contemplou o atendimento dos objetivos específicos 4 e 5. Os experimentos compreenderam desde a análise de dimensionalidade (Seção 5.1) do descritor definido inicialmente, passando por dois experimentos de classificação (Seção 5.2), bem como a execução de um teste não paramétrico de hipóteses (Seção 5.3).

1.4 ESTRUTURA DA DISSERTAÇÃO

O documento está organizado em seis capítulos correlacionados. O Capítulo 1, Introdução, contextualiza o problema de pesquisa, a pergunta e a hipótese de pesquisa, bem como os objetivos, as justificativas, o escopo e o contra escopo da proposta de trabalho. Os demais capítulos são dispostos da seguinte forma:

- **Capítulo 2:** apresenta a fundamentação teórica, apontando tópicos considerados relevantes ao entendimento do trabalho, relacionando conceitos de descrição de textura, classificação supervisionada, avaliação de classificadores e o processo de IHQ;
- **Capítulo 3:** apresenta a revisão bibliográfica a respeito do tema de pesquisa, conduzida através de RSL e aponta o resultado da execução do protocolo de busca estabelecido, com a análise de sete trabalhos considerados relacionados e uma comparação com a presente pesquisa;
- **Capítulo 4:** descreve a solução para o problema de pesquisa, demonstrando o descritor de características, o método de classificação, a descrição da base de imagens e as anotações a serem utilizadas nos experimentos, além da descrição detalhada dos testes planejados e o resultado da implementação do protótipo;
- **Capítulo 5:** descreve a execução dos experimentos propostos, a discussão a respeito dos resultados alcançados e a execução e resultado do teste de hipóteses; e
- **Capítulo 6:** apresenta as conclusões extraídas com a finalização do trabalho, relacionadas às definições estabelecidas e ao atendimento dos objetivos, bem como uma seção de sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

O Capítulo 2 apresenta uma relação de tópicos diretamente relacionados a presente proposta. Na Seção 2.1 é apresentada a formalização do conceito de descritores de características e na Seção 2.2 são apresentadas quatro abordagens para descrição de características baseadas em textura. Nas seções 2.3 e 2.4 são apresentados conceitos a respeito da classificação de padrões, bem como a abordagem polinomial a ser empregada neste trabalho, respectivamente. As seções 2.5 e 2.6 apresentam os conceitos a respeito da avaliação de classificadores e da área de aplicação (desde a preparação das amostras até o diagnóstico), respectivamente.

2.1 DESCRIPTORES DE CARACTERÍTICAS

Vetores de características (*feature vector* ou *pattern vector*) são representados por letras minúsculas em negrito, como \mathbf{x} , assumindo a forma em (1):

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (1)$$

onde cada componente, x_i , representa a i -ésima medida que compõe o descritor e n corresponde ao número de medidas associadas ao descritor. Normalmente os vetores são expressos na forma de matrizes $nx1$, assim a forma apresentada em (1) pode usar a seguinte notação:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T \quad (2)$$

onde T indica a transposta. A natureza de cada medida que compõe um vetor de características \mathbf{x} depende da metodologia empregada para descrever o padrão físico, propriamente dito (GONZALEZ; WOODS, 2010).

A seleção de características discriminatórias define o espaço de características. Essencialmente é desejável que a seleção de características atinja duas propriedades: (i) dimensionalidade reduzida e (ii) alto grau de discriminação. Considerando o problema de classificação de duas imagens: leão e leopardo, uma abordagem seria empregar algum descritor de

forma para o problema, entretanto, os dois animais são felinos de grande porte e semelhantes. Neste caso específico, uma abordagem mais eficiente seria baseada em atributos de cor, pois o espaço cromático RGB (3-dimensional), seria suficiente para discriminar as duas imagens (SOLOMON; BRECKON, 2013).

Para a maioria das aplicações práticas, as variáveis de entradas são tipicamente pré-processadas, com a finalidade de transformá-las em um novo espaço de características, onde se espera que o problema de RP se torne mais fácil de resolver. Este estágio de pré-processamento muitas vezes é denominado de extração de características (BISHOP, 2006).

O objetivo principal da extração de características consiste em categorizar um padrão através de medidas que são similares para padrões em uma mesma categoria, e diferentes para padrões em categorias distintas. O problema consiste em procurar por características distintivas que sejam invariantes a transformações irrelevantes no padrão de entrada, uma tarefa mais dependente do problema e do domínio do que da classificação propriamente dita, o que torna essencial o conhecimento do domínio (DUDA; HART; STORK, 2000).

2.2 DESCRIPTORES DE TEXTURA

A representação de características em imagens é uma tarefa importante em sistemas de visão computacional e processamento de imagens digitais. Entre as diversas abordagens comumente encontradas na literatura estão características baseadas em cor, características baseadas em textura, características baseadas em descritores de forma e características baseadas em relações espaciais. Descritores de características baseadas em textura desempenham papel importante na análise de imagens naturais, entretanto a obtenção efetiva destes descritores continua sendo um tópico ativo de pesquisa (SONG *et al.*, 2013).

A textura é um fenômeno comum, fácil de reconhecer, porém difícil de definir, causando divergência entre os autores. Tipicamente para um efeito ser definido como textura, ele depende da escala na qual é observado. As texturas, no contexto de imagens digitais, geralmente são formadas por um padrão organizado de subelementos bastante regulares (*textons*). Uma abordagem para representação de texturas consiste em encontrar os *textons* e em seguida descrever a forma em que eles são definidos (FORSYTH; PONCE, 2002).

Os *textons*, também conhecidos como primitivas de textura ou elemento de textura, são altamente dependentes da quantidade de subelementos que compõem a primitiva, conceito conhecido como escala de textura. Características distintivas podem ser obtidas através dos tons e estruturas de uma textura, em que os tons são relacionados a intensidade de cinza de um determinado píxel e as estruturas definem as relações espaciais entre primitivas (PATRA; RAJU; MARUTHUPERUMAL, 2013).

Segundo Hawkins (1969), a noção a respeito de textura parece depender de três ingredientes: (i) uma ordem local repetida sobre uma região maior que esta ordem; (ii) a ordem consiste em um arranjo não aleatório de partes elementares e (iii) as partes são entidades uniformes com dimensões aproximadas, dentro da região de textura.

Segundo Kaur e Gupta (2012), no campo de análise de imagens, abordagens de textura são normalmente classificadas em duas categorias principais:

- **Estatística** – neste grupo as texturas são categorizadas por medidas estatísticas, relacionadas aos níveis de cinza dos pontos que compreendem uma superfície, computadas através de técnicas baseadas em histogramas; e
- **Estrutural** – nesta categoria, as texturas são textons determinísticos, que se repetem com alguma regra de localização. Tais regras determinam o relacionamento entre os textons, que podem ser adjacentes, com uma distância aproximada ou periódica.

Além destas abordagens, Gonzalez e Woods (2010) apontam uma terceira categoria:

- **Espectral** – nesta categoria, texturas são caracterizadas por propriedades do espectro de Frequência, sendo usadas, principalmente, para identificar a periodicidade global em uma imagem, através de picos de alta energia no espectro.

Sharma e Singh (2001) apresentam uma avaliação de cinco métodos (entre as categorias citadas), para extração de textura: Autocorrelação, Matriz de Coocorrência, Frequência de Borda, Comprimento Primitivo e Método de Law. Os testes realizados demonstraram um melhor desempenho para os métodos de Coocorrência e Law (categoria estatística).

Dentre os diversos métodos para extração e caracterização de textura, há um método emergente e popular denominado Filtros de Gabor. Na prática, um banco de filtros é definido através da variação de um grupo de parâmetros, incluindo frequência, orientação e suavização. Diferentes configurações destes parâmetros vêm sendo propostas na literatura em geral, com resultados aceitáveis. Comumente, a construção destes bancos combina cinco frequências e oito orientações (LI; RUAN; RUAN, 2010; ZHAO; PARK; LEE, 2011).

Outro método popular na descrição de texturas, encontrado na literatura, é denominado Padrão Binário Local (Local Binary Pattern - LBP). Nesse método, um descritor invariante a níveis de cinza é composto através da comparação de um píxel central com seus píxeis vizinhos. O método LBP demonstra ser um descritor de baixa complexidade computacional e menos sensitivo a variações de luminosidade que muitos descritores, e vem sendo utilizado em muitas tarefas de RP, com resultados promissores (GU; LIU, 2013; NANNI; BRAHNAM; LUMINI, 2012).

2.2.1 Matriz de coocorrênci a

Haralick, Shanmugam e Distein (1973) estabeleceram uma classe de características baseadas em textura, passíveis de aplicação geral em uma gama de tipos de imagens. A ideia principal consiste na ocorrência de probabilidade conjunta, em pares de níveis de cinza i e j para dois píxeis, com uma relação espacial em uma imagem. Tais relações são descritas em função de uma distância d e um ângulo θ (SHARMA; SINGH, 2001).

Supondo uma imagem I_g , em tons de cinza, com n níveis de cinza possíveis, desta forma se constrói uma matriz de coocorrênci a $M_{d,\theta}(i,j)$ que dependa de uma distância d e uma direção θ com dimensões $N \times N$, como ilustrado na Figura 2. Assim cada imagem é representada por quatro matrizes normalizadas de coocorrênci a ($M_{d,0^\circ}; M_{d,45^\circ}; M_{d,90^\circ}$ e $M_{d,135^\circ}$).

A notação $\#(i, j)$ na Figura 2-b representa o total da ocorrência dos valores de cinza i e j como vizinhos, de acordo com a distância e o ângulo selecionados, Figura 2-c (PRASETIYO *et al.*, 2010).

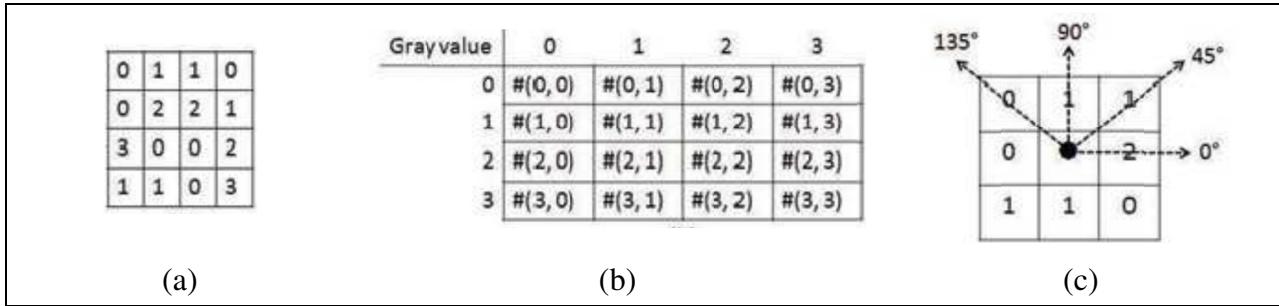


Figura 2. (a) imagem com níveis de cinza; (b) notação de matriz de coocorrência e (c) direções angulares

Fonte: Adaptado de Prasetiyo *et al.*, (2010).

Após a geração das matrizes de coocorrência, Haralick, Shanmugam e Distein (1973) propuseram 14 medidas F tomadas sobre estas matrizes. Três destas medidas são apresentadas em (3), (4), (5):

$$F_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \left(\frac{M(i,j)}{R} \right)^2 \quad (3)$$

$$F_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{|i-j|=n} \left(\frac{M(i,j)}{R} \right) \right\} \quad (4)$$

$$F_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [ijM(i,j)/R] - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (5)$$

em que μ_x , μ_y , σ_x e σ_y são a média e o desvio padrão da distribuição marginal $M(i, j) / R$, e R é uma constante de normalização. Estas 14 medidas caracterizam as dependências espaciais dos níveis de cinza em uma imagem: a medida apresentada em (3) calcula a homogeneidade e é definida como segundo momento angular; a medida em (4) é definida como contraste e aponta a quantidade de variações locais em uma imagem; em (5) a medida aponta uma correlação de dependência linear entre níveis de cinza em uma imagem.

Métodos para descrição de textura, baseados em coocorrência, vêm sendo utilizados com sucesso, em abordagens de segmentação de imagens, entretanto uma desvantagem consiste no custo

computacional. Algumas restrições causam a impossibilidade de implementação, para abordagens de segmentação ou classificação píxel a píxel (MOKJI; BAKAR, 2007).

2.2.2 Texturas de Laws

O método para descrição de características baseadas em textura, apresentado por Laws (1980), requer algumas convoluções com pequenas máscaras (normalmente de dimensões 5x5), de coeficientes inteiros para a mensuração dos valores de textura. As máscaras são separáveis e podem ser implementadas com convoluções 1-dimensional (horizontal ou vertical) ou 3-dimensional.

Neste método, vinte e cinco máscaras de convolução são aplicadas sobre a região de textura, com a finalidade de acentuar micro estruturas, fornecendo assim uma imagem na qual estas micro estruturas são mensuradas a partir de cinco métricas estatísticas, proporcionando um descritor de características com cento e vinte e cinco medidas (SHARMA; SINGH, 2001).

Com a finalidade de obter as respostas de textura do método Laws, são definidas máscaras de convolução, 1-dimensional (1D) de comprimento cinco (p. ex. L5 = [1, 4, 6, 4, 1]; E5 = [-1, -2, 0, 2, 1]; S5 = [-1, 0, 2, 0, -1]; W5 = [-1, 2, 0, -2, 1] e R5 = [1, -4, 6, -4, 1]). Estas máscaras de convolução permitem calcular aspectos como média local, detecção de borda, detecção de pontos e detecção de ondulações.

Vinte e cinco máscaras 2-dimensional (2D) de dimensões 5x5 são obtidas através do produto entre as máscaras 1D. A Figura 3-a ilustra estas 25 máscaras e a Figura 3-b mostra a máscara L5E5 propriamente dita (VIRMANI *et al.*, 2011)..

$L5L5$	$E5L5$	$S5L5$	$W5L5$	$R5L5$	$L5 * E5 =$	$\begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 8 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix}$
$L5E5$	$E5E5$	$S5E5$	$W5E5$	$R5E5$		
$L5S5$	$E5S5$	$S5S5$	$W5S5$	$R5S5$		
$L5W5$	$E5W5$	$S5W5$	$W5W5$	$R5W5$		
$L5R5$	$E5R5$	$S5R5$	$W5R5$	$R5R5$		
(a)					(b)	

Figura 3. (a) 25 combinações 1D que geram as máscaras de Laws e (b) exemplo da máscara L5E5

Fonte: Adaptado de Virmani *et al.*, (2011).

O resultado da convolução destas máscaras, sobre a imagem em análise, fornece a energia de textura de cada píxel. A medida desta energia é obtida através de (6):

$$E_k(i, j) = I(i, j) * H_k(i, j) \quad (6)$$

em que $I(i, j)$ é a intensidade do píxel de coordenada i e j , e $H(i, j)$ é o resultado da convolução de uma das máscaras em questão, com centro na coordenada i e j . O valor da energia geral E de uma imagem é obtido através de (7) e (8):

$$E_k(i, j) = \frac{1}{Pk \cdot N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} K_k^2(i, j) \quad (7)$$

$$P_k = \sum_{i,j} H_k(i, j)^2 \quad (8)$$

2.2.3 Padrão Binário Local

O método denominado LBP (inicialmente proposto por Ojala *et al.*, 1996 e 2002), consiste em um operador de imagem que a transforma em uma nova imagem rotulada píxel a píxel, descrevendo informações de textura em pequena escala. Cada píxel na imagem recebe um rótulo de acordo com uma avaliação da sua vizinhança, dado um píxel central $x_{0,0}$ e um grupo de vizinhos p ($x_{r,0}, \dots, x_{r,p-1}$), separados de forma equidistante ao redor de um círculo de raio r como ilustra a Figura 4-a. Todos os vizinho com valores (tons de cinza) acima ou igual ao do píxel $x_{0,0}$ recebem um valor 1, e 0 caso contrário. Estes valores binários são então agrupados sequencialmente, (no sentido horário) formando o padrão binário local (PIETIKÄINEN *et al.*, 2011; LIU *et al.*, 2012).

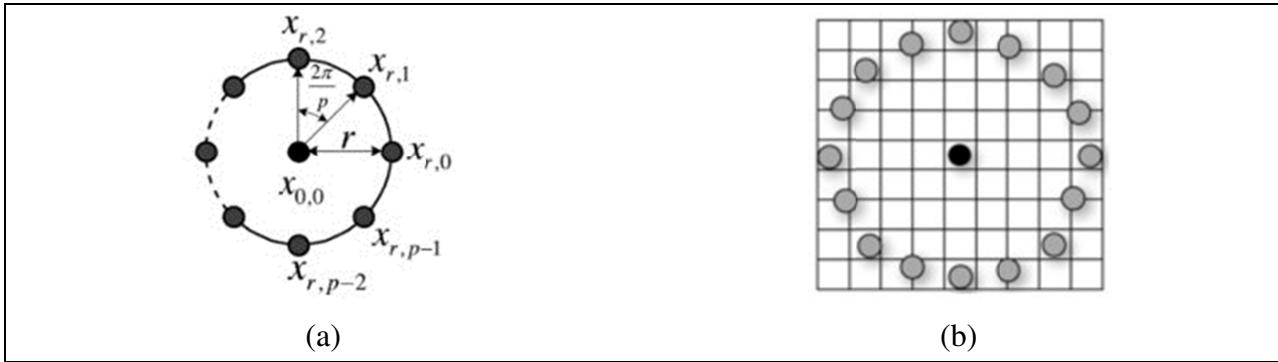


Figura 4. Um píxel central $x_{0,0}$ e sua vizinhança espaçada por um raio r

Fonte: Adaptado de Liu *et al.*, 2012 e Nanni, Brahnam, Lumini (2012).

O valor do operador LBP, para um determinado píxel, é alcançado através da conversão do valor binário para decimal, formalmente apresentado em (9). Usando uma vizinhança circular e interpolando linearmente o valor do píxel, é possível usar qualquer valor de r (raio) e p para formar uma vizinhança p, r (LIU; TANG; TANG, 2013). A Figura 4 ilustra uma vizinhança 16, 4 (16 vizinhos e raio 4). A Figura 5 ilustra o cálculo do valor LBP (8,1) para um píxel central de valor 166 ($x_{0,0}$).

$$LBP_{p,r} = \sum_{n=0}^{p-1} (x_{r,n} - x_{0,0}) 2^n, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (9)$$

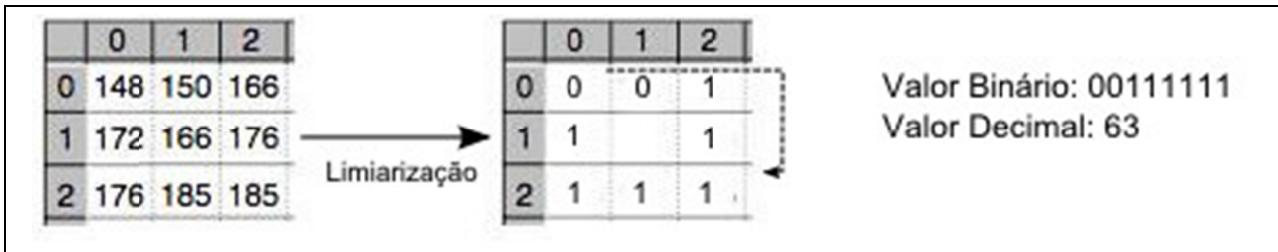


Figura 5. Ilustração do cálculo do valor $LBP_{(8,1)}$ para um píxel central de valor 5

Fonte: Liu; Tang; Tang, 2013.

O método LBP vem sendo amplamente utilizado, para descrição de textura, em diversas áreas de aplicação como medicina, reconhecimento de face, reconhecimento de impressão digital, armas inteligentes e outros. Este método tem demonstrando superioridade em estudos comparativos, provando ser um descritor poderoso, de baixo custo computacional e menos sensitivo as mudanças de iluminação que muitos métodos. Uma vez que é invariante a condições de iluminação, o método

LBP se torna uma boa escolha para a descrição de detalhes finos (NANNI; BRAHNAM; LUMINI, 2012).

2.3 CLASSIFICAÇÃO DE PADRÕES

De modo geral, qualquer método que incorpore informação de um grupo de amostras, para a construção do classificador, emprega métodos de aprendizagem. Desta forma, a construção de classificadores envolve a definição de um modelo e a utilização de um CT⁶, para estimar os parâmetros do referido modelo.

A aprendizagem consiste em um algoritmo que busca a redução de erro em um CT, podendo ser agrupada em duas categorias principais: supervisionada e não supervisionada. Na classificação supervisionada, padrões do CT são rotulados as suas respectivas classes, na esperança que o método de classificação possa ser generalizado, para reconhecer novos padrões. Em abordagens não supervisionadas, como métodos baseados em aglomeração (*clustering*), não existe a rotulação de classes (SOLOMON; BRECKON, 2013; DUDA; HART; STORK, 2001).

O Quadro 1 apresenta alguns termos típicos relacionados à classificação de padrões e sua descrição sumária. A Figura 6 apresenta o fluxo típico no processo de criação de classificadores, e as etapas são definidas da seguinte forma (SOLOMON; BRECKON, 2013):

- **CLASSE:** nesta etapa ocorre a definição das classes a serem classificadas, intimamente relacionadas ao problema abordado;
- **DADOS:** nesta etapa são definidos os atributos que permitam a separação dos dados em classes específicas, sendo os mais comuns: intensidade, cor, textura, ou uma combinação deles;
- **CARACTERÍSTICAS:** a quantidade de medidas que irão compor o DC define o espaço de características. Usualmente, nesta etapa é aplicada alguma técnica de redução de dimensionalidade, a fim de obter um DC compacto;

⁶ conjunto de treinamento

- **CLASSIFICADOR:** nesta etapa é construído o classificador propriamente dito. Em abordagens supervisionadas, um conjunto de treinamento rotulado é empregado, para treinamento do classificador; e

TESTES: nesta etapa o desempenho do classificador é avaliado, empregando um conjunto de dados não rotulado (não utilizado no processo de aprendizagem). Dependendo do resultado obtido, todo processo é revisto até atingir um nível estimado de acordo com a aplicação.

Quadro 1. Termos comumente utilizados no processo de classificação

Termo	Descrição
Descriptor de Características (DC)	Vetor n -dimensional, onde cada posição especifica uma medida relacionada ao padrão.
Espaço de Características (EC)	Espaço matemático abstrato (n -dimensional), que contém os descritores.
Conjunto de Treinamento (CT)	Descritores empregados no processo de aprendizagem.
Conjunto de Experimentos (DT)	Descritores utilizados no processo de teste.
Padrão de Classe (PC)	Grupos de descritores que compartilham uma origem estatística ou semântica.
Espaço de Classificação (EC)	Subespaço do EC que delimita uma classe.
Função Discriminante (FD)	Função cujo valor determina a atribuição de um padrão a uma classe ou outra.

Fonte: Adaptado Solomon; Breckon (2013).

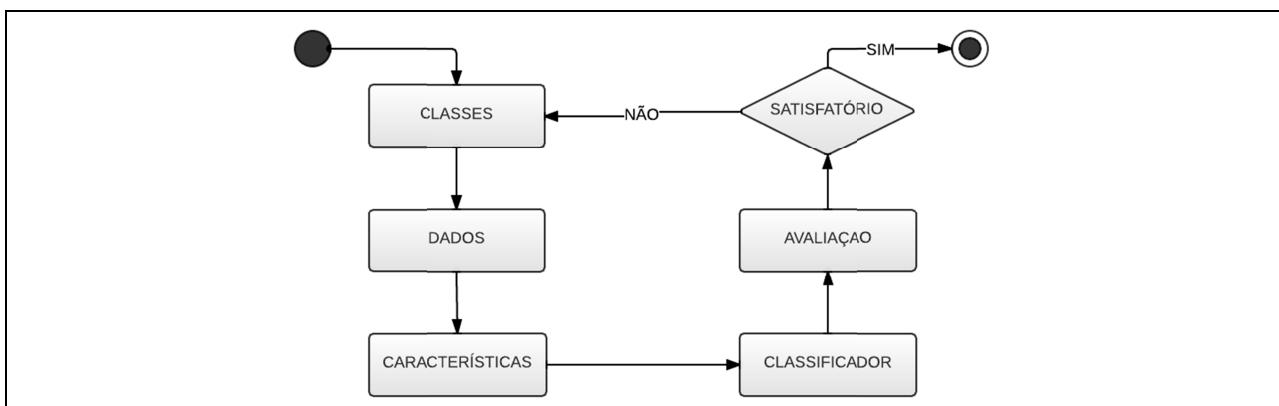


Figura 6. Fluxo no processo de construção de classificadores

Fonte: Adaptado Solomon e Breckon (2013).

Uma abordagem simples para a classificação é conhecida como distância mínima. Nesta abordagem são construídos protótipos de classe, a partir do vetor médio destas classes, denominados de \mathbf{m}_j , formalmente:

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{x \in \omega_j} x_j \quad j = 1, 2, \dots, W \quad (10)$$

onde N_j corresponde ao número de vetores atribuídos a classe ω_j , e W corresponde ao número de classes tratadas. Para calcular pertinência de um padrão a uma determinada classe, basta associar o padrão ao seu protótipo mais próximo, empregando métricas de similaridade, usando a Distância Euclidiana (DE), a fim de estimar tal similaridade se reduz o problema ao cálculo de distância:

$$D_j = \|\mathbf{x} - \mathbf{m}_j\| \quad j = 1, 2, \dots, W \quad (11)$$

onde $\|\mathbf{a}\| = (\mathbf{a}^T \mathbf{a})^{1/2}$ é a norma Euclidiana. Desta forma o padrão \mathbf{x} é atribuído a classe ω_i , se $D_i(\mathbf{x})$ for a menor distância (GONZALEZ; WOODS, 2010).

Entretanto, a consideração simplesmente da média dos DC, sem levar em conta a distribuição estatística dos DC em torno da média, geralmente não são ótimos. Especialmente em se tratando de distribuições não lineares, no espaço de características, uma alternativa é usar métrica de similaridade que considere uma medida probabilística como a Distância de Mahalanobis (DM):

$$D_{mh} = (\mathbf{x} - \boldsymbol{\mu})^T S^{-1} (\mathbf{x} - \boldsymbol{\mu}) \quad (12)$$

onde S representa a matriz de covariância e é definida por:

$$S = \langle (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T \rangle = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & \sigma_N^2 \end{bmatrix} \quad (13)$$

em que $\sigma_{ij} = \langle (\mathbf{x}_i - \boldsymbol{\mu}_i)(\mathbf{x}_j - \boldsymbol{\mu}_j) \rangle$, e $\langle . \rangle$ denota o valor esperado, ou a média das amostras. O emprego desta medida de distância fornece uma métrica mais intuitiva, uma vez que utiliza

informações da distribuição estatística dos DC no espaço, o princípio da classificação bayesiana (SOLOMON; BRECKON, 2013).

2.4 POLINOMIAL MAHALANOBIS

Na Figura 7 é possível observar a comparação entre a distância euclidiana, a distância de Mahalanobis e a distância polinomial de Mahalanobis (a partir da geração de uma distribuição sintética) em espaço de duas dimensões. Na Figura 7-a, a DE não segue a distribuição dos pontos e usa pesos iguais em todas as direções. Na Figura 7-b, a DM, por empregar a matriz de covariância, encolhe a esfera euclidiana para uma elipsoide.

Entretanto as métricas referentes a Distância Euclidiana e a Distância de Mahalanobis (sem a abordagem polinomial) apresentam limitações ao acompanhar distribuições com características não lineares, a abordagem polinomial ameniza esta restrição, mapeando os descritores de características em termos polinomiais de altas ordens, como pode ser observado nas Figura 7-c e Figura 7-d (quanto mais branco o fundo, mais próximo do ponto de referência representado pelo quadrado azul).

Uma das principais desvantagens da DPM é que sua precisão está diretamente relacionada ao tamanho e a representatividade do CT e a ordem polinomial (RICHTER; PETYOVSKY; MIKSIK, 2011).

Uma forma para calcular similaridade, empregando polinomiais de grande ordem, foi apresentada por Grudic e Mulligan (2006), considerando um grupo de N DC ($\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$) e um \mathbf{x}_0 usado para cálculos de covariância, a DPM entre dois descritores x_i e x_j é expressa por:

$$DPM(x_i, x_j) = DM_{\sigma^2}(x_i, x_j) + \sum_{l=1}^L DM_{\sigma^2}(g_l^i, g_l^j) \quad (14)$$

onde DM_{σ^2} , representa a distância de Mahalanobis computada com um pequeno escalar positivo σ^2 , usado para anular limitações de inversão de matrizes se algum valor for zero, $L > 0$ é o número máximo de projeções de q -ordem ($q = 2^L$), e os argumentos g_l^i e g_l^j são as projeções dos DC x_i e x_j em seus termos polinomiais.

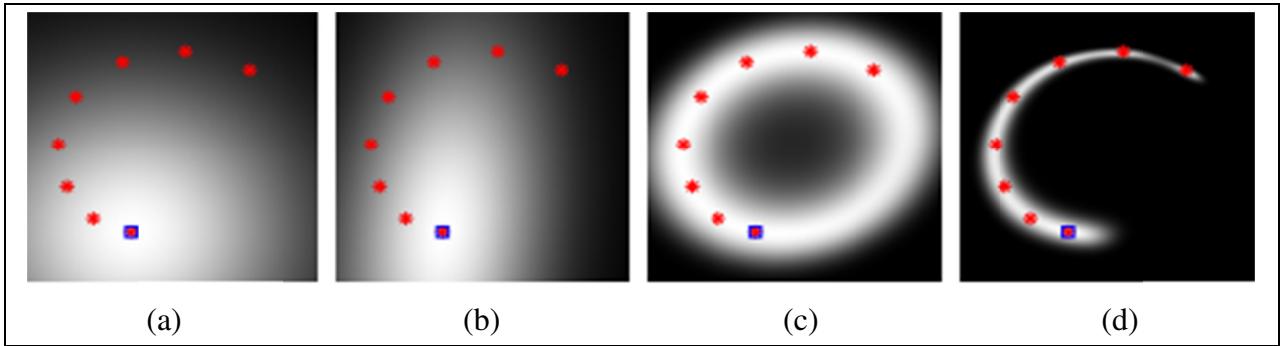


Figura 7. Comparação entre DE (a), DM (b), DPM ordem 2 (c) e DPM ordem 8 (d)

Fonte: Adaptado de Richter, Petyovsky e Miksik, (2011).

Suponha um conjunto S de DC m -dimensional ($s_1, s_2, s_3, \dots, s_N$), a primeira etapa consiste em calcular a DM, como apresentado em (12) entre dois DC de S , obtendo assim a primeira q -ordem (ordem 1 ou DM entre x_i e x_j). A segunda q -ordem é obtida com o mapeamento de todas as m dimensões, em todas as q -ordens pretendidas, por exemplo: um DC com duas dimensões d_1 e d_2 ($DC = d_1, d_2$) tem seu mapeamento na segunda ordem polinomial como $DC_{2\text{-ordem}} = d_1, d_2, d_1^2, d_2^2, d_1, d_2$.

Entretanto de um ponto de vista computacional, altas dimensões (m) no DC e altas ordens polinomiais (q -ordem) tornam o cálculo impraticável (SOBIERANSKI *et al.*, 2009). A fim de contornar o problema, Grudic e Mulligan (2006) propuseram uma abordagem que emprega o conceito de decomposição singular de valor, permitindo altas ordens de projeção polinomial.

2.5 AVALIAÇÃO DE CLASSIFICADORES

Diversas métricas para a avaliação de classificadores são propostas na literatura, entretanto a escolha da mais adequada continua sendo um problema de pesquisa em aberto. Entre algumas métricas estabelecidas estão a acurácia, acurácia não ponderada, erro quadrático médio e a área sob a curva Característica de Operação do Receptor (Receiver Operating Characteristic - ROC).

Muitos classificadores podem ser enxergados como uma função que tem como saída um escore e uma classe, este escore representa a probabilidade de uma amostra pertencer a uma determinada classe e a atribuição acontece de acordo com alguma regra de decisão estabelecida (HERNÁNDEZ-ORALLO *et al.*, 2012).

Considerando um exemplo de classificação binária (somente duas classes), formalmente cada instância (I) é mapeada para uma de duas possibilidades, um rótulo positivo (p) ou negativo (n). Um modelo de classificação pode ser visto como o mapeamento de uma instância para uma predição de classe. Alguns modelos têm uma saída contínua (normalmente um valor de probabilidade), por isso, valores de limiares devem ser definidos, para a previsão do classificador. Outros modelos têm uma saída com valor discreto, indicando a pertinência ou não a classe em questão (FAWCETT, 2006).

Normalmente, durante um teste do modelo de classificação, um grupo de amostras é submetido ao julgamento do classificador. Desta forma, o resultado das predições pode ser comparado com os rótulos verdadeiros (anotações de metadados sobre as amostras), empregando uma tabela de contingência (matriz de confusão), conforme ilustra a Tabela 1 (BRADLEY, 1997).

Tabela 1. Tabela de contingência ou matriz de confusão

	Positivo Anotado	Negativo Anotado
Positivo Preditó	Verdadeiro Positivo (VP)	Falso Positivo (FP)
Negativo Preditó	Falso Negativo (FN)	Verdadeiro Negativo (VN)
Total Anotado	Positivos (P) ($VP+FN$)	Negativos (N) ($FP+VN$)

Fonte: Adaptado de Arbel e Rokach (2006);

Dado um classificador e uma amostra para julgamento, existem quatro possibilidades de resultado, caso a predição seja positiva e o valor anotado para a amostra seja positivo, a classificação é computada como Verdadeiro Positivo (VP).

Quando a predição é negativa e a anotação da mesma forma, o julgamento é considerado Verdadeiro Negativo (VN). Com uma predição negativa e uma anotação positiva, o julgamento é computado como Falso Negativo (FN).

Por fim, quando a previsão é positiva e o valor anotado é negativo, o resultado de classificação é dito Falso Positivo (FP). A partir da avaliação de um conjunto de dados anotados (conforme a Tabela 1) é possível construir diversas métricas de desempenho F , relacionadas a predições corretas e incorretas. Formalmente apresentadas em (15), (16), (17), (18), (19), (20), (21), (22), (23), (24), (25) e (26).

$$F_{tvp} = \frac{VP}{P} \quad (15)$$

$$F_{tfp} = \frac{FP}{N} \quad (16)$$

$$F_p = \frac{VP}{VP + FP} \quad (17)$$

$$F_r = \frac{VP}{P} \quad (18)$$

$$F_a = \frac{VP + VN}{P + N} \quad (19)$$

$$F_f = \frac{2}{\frac{1}{precisão} + \frac{1}{recal}} \quad (20)$$

$$F_e = \frac{2 \times VP}{2 \times VP + FP + FN} \quad (21)$$

$$F_{mcc} = \frac{VP \times VN - FP \times FN}{\sqrt{(VP + FP) \times (VP + FN) \times (VN + FP) \times (FN + VN)}} \quad (22)$$

$$F_{tfp} = \frac{FP}{N} \quad (23)$$

$$F_{sen} = F_r \quad (24)$$

$$F_{esp} = \frac{VN}{FP + VN} \quad (25)$$

$$F_{tfd} = \frac{FP}{FP + VP} \quad (26)$$

Os valores de VP e VN correspondem ao número de decisões corretas do classificador. Por outro lado, os seus opostos são os erros cometidos (confusão entre classes), também conhecidos como erro do tipo I e erro do tipo II, FP e FN respectivamente.

As métricas F_{vp} , F_p , F_r , F_a , F_f , F_e , F_{mcc} , F_{sen} e F_{esp} estão relacionadas a previsões corretas do classificador, neste caso, quanto mais próximo de 1 melhor o desempenho de classificação (sendo 1, a classificação perfeita). As medidas de F_{fp} , F_{fd} e F_{fn} estão relacionadas a previsões incorretas (taxas de erro), portanto, um bom desempenho de classificação deve apresentar valores próximos a zero.

2.6 IMUNO-HISTOQUÍMICA

O processo de IHQ consiste em identificar抗ígenos alvos em amostras de tecido, empregando o uso de anticorpos específicos e avaliando o resultado desta interação (antígeno/anticorpo). Patologistas vêm dependendo mais e mais da análise de imagens microscópicas, para a avaliação da presença e atividade dos referidos抗ígenos, com aplicações importantes na análise de tumores, bem como diversos propósitos de pesquisa (SHU; QUI; MOHAMMAD, 2013; CATALDO; FICARRA; MACII, 2012).

Imuno-histoquímica consiste em uma atividade padrão para o diagnóstico e prognóstico de doenças como câncer e hepatite. A progressão e a resposta ao tratamento dependem do monitoramento de mudanças em抗ígenos específicos. Quando existe a suspeita de uma doença, uma biópsia do tecido correspondente é adquirida, corada através de um método de IHQ e analisada visualmente por um especialista, com microscópio. Os resultados desta análise constituem uma importante diretriz no planejamento do tratamento (HUI; ZHGUO; LONGMEI, 2011).

A visualização da relação抗ígeno/anticorpo pode ser alcançada através de diversas formas e métodos, dependendo do tipo de tecido que está sendo investigado. Comumente um anticorpo é conjugado a uma enzima capaz de catalisar uma reação que produz cor. Tipicamente o procedimento de IHQ consiste em duas etapas: (I) preparação das amostras e (II) imunocoloração (Figura 90). Ambas as etapas são compostas por diversos procedimentos mecânicos e químicos, os quais tem impacto direto na qualidade e reprodutibilidade das análises de IHQ. (CATALDO; FICARRA; MACII, 2012).

Na etapa de preparação das amostras, o primeiro procedimento consiste em imergir a biopsia em um fixador, com a finalidade de garantir a preservação da estrutura do tecido e da morfologia das células. Em seguida, as amostras são revestidas com parafina ou resina e cortadas em finas camadas, com um micrótomo. Após o corte, o processo de revestimento é desfeito e opcionalmente tratado com um recuperador, para fortalecer a ligação antígeno/anticorpo. No fim da etapa de preparação das amostras, o processo de incubação previne a ligação de proteínas não específicas (CATALDO; FICARRA; MACII, 2012).

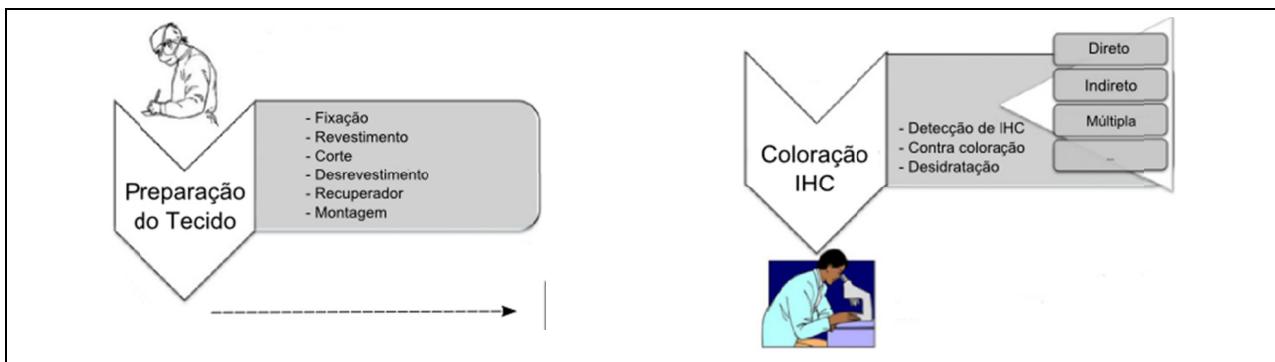


Figura 8. Etapas básicas no processo de IHQ

Fonte: Cataldo; Ficarra; Macii, (2012).

Na segunda etapa (imunocoloração), os métodos para coloração podem ser diretos ou indiretos. No método direto apenas um reagente (anticorpo primário) é aplicado à amostra, reagindo diretamente com o antígeno. No método indireto, um segundo reagente (anticorpo secundário) é aplicado, reagindo com o anticorpo primário. Um processo de contra coloração (opcional) pode ser realizado, para aumentar o contraste das reações a serem observadas. Finalmente as amostras são montadas para observação no microscópio, dependendo da mídia de montagem pode ocorrer a necessidade de desidratação dessas amostras, utilizando álcool (CATALDO; FICARRA; MACII, 2012).

Na análise de IHQ, mais importante do que identificar a presença ou ausência da coloração específica, é quantificar a intensidade desta expressão, bem como a sua localização (membrana celular, citoplasma ou núcleo). O resultado desta análise é obtido diretamente pelo especialista, através da observação com o microscópio, um processo laborioso e sujeito a vieses por diversos fatores subjetivos. Esta subjetividade pode ser evitada, empregando análises quantitativas guiadas por computador (MATOS *et al.*, 2010; DONG *et al.*, 2010 a).

A Figura 9 mostra dois exemplos de imagens de IHQ. A Figura 9-a apresenta uma amostra de tecido com carcinoma de pulmão, usando o cromógeno diaminobenzidina (DAB), coloração marrom. A Figura 9-b apresenta uma amostra de pele, empregando o cromógeno amino-etil-carbazol (AEC), coloração vermelha.

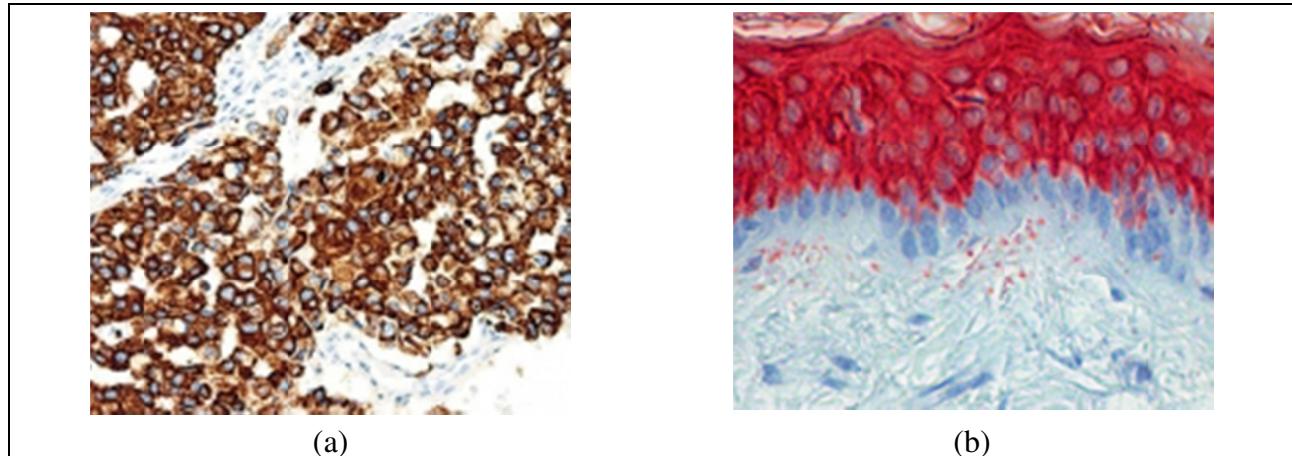


Figura 9- Imagens de IHQ: a) Carcinoma de pulmão com DAB; b) Amostra de pele com AEC

Fonte: Biogenex (2014).

Um sistema de graduação adotado pela World Health Organization divide os estágios de Linfoma Folicular (LF) em três níveis histológicos, de acordo com a média de Centroblastos (CB), em dez Campos de Observação (COb) distintos de uma mesma amostra, empregando microscópio. Até 5 CB/COb é considerado nível I, entre 6 e 15 CB/COb, nível II e acima de 15 CB/COb, nível III. Os dois primeiros níveis têm grande taxa de sobrevivência e não necessitam de quimioterapia, já o nível III é considerado agressivo e mortal, se não for tratado imediatamente (SAMSI *et al.*, 2010).

Na análise de IHQ, para avaliação de câncer de mama, os núcleos celulares são corados com reagentes marrom e azul, sendo identificadas como células positivas e células negativas respectivamente. Especialistas formulam diagnósticos e prognósticos, através da relação de porcentagem, entre o total de células e o total de células positivas. A American Society of Clinical Oncology (ASCO) recomenda a utilização de um protocolo de graduação de cinco níveis, iniciando no nível 0, para uma faixa de 0-5%, até os níveis 1, 2, 3, 4 para faixas 6-10%, 11-33%, 34-66%, 67-100% respectivamente (JITAREE *et al.*, 2013; KOSTOPOULOS *et al.*, 2007).

3 REVISÃO BIBLIOGRÁFICA

Conforme apresentado na Subseção 1.3.2 um dos procedimentos metodológicos definidos para a execução do trabalho foi uma revisão bibliográfica, conduzida através de uma RSL. O Capítulo 3 apresenta o resultado desta revisão, o resumo de sete trabalhos elencados como relacionados ao problema de pesquisa, e ainda algumas considerações a respeito da conclusão desta etapa.

3.1 REVISÃO SISTEMÁTICA DA LITERATURA

A motivação para a realização da RSL foi a identificação de trabalhos relacionados, seus métodos e métricas, relativos especificamente à classificação e quantificação de coloração imuno-histoquímica guiada por computador. Outros fatores apontados foram a identificação dos resultados alcançados por estes trabalhos, e ainda a prospecção de futuros problemas de pesquisa, através da identificação de possíveis tópicos em aberto.

Os critérios definidos no protocolo da RSL incluíram pergunta de pesquisa, população, resultados, estratégia para pesquisa, critérios e procedimentos para seleção, lista de verificação e procedimentos para avaliação da qualidade de estudos e reunião dos dados. O protocolo na íntegra pode ser conferido no Apêndice A. Somados os resultados das três fontes de pesquisa totalizam 62 trabalhos retornados, dos quais foram lidos o resumo e o título, para avaliação de inclusão (conforme protocolo). Destes 62 trabalhos, 20 foram selecionados para a análise (leitura completa), conforme ilustrado na Tabela 2.

Tabela 2. Resultado de pesquisa e seleção de trabalhos para análise

Fonte	Retorno	Excluídos	Selecionados
IEEE	25	10	15
ScienceDirect	18	15	3
ACM Digital Library	19	17	2
Total	62	42	20

Na Tabela 3 é apresentado um resumo da coleta de dados, dos vinte trabalhos analisados, seguindo a RSL, indicando métodos relacionados à segmentação e à quantificação de padrões de coloração, em imagens microscópicas de IHQ. Foram agrupados métodos, área de aplicação, ano de publicação e autor.

Tabela 3 Reunião de técnicas e campos de aplicação encontrados na revisão sistemática

T	Técnicas	Área de Aplicação	Referência
01	Textura GLCM e FD Rede Neural Multilayer Perceptron com Back Propagation	Identificação de regiões cancerígenas relativas ao câncer de mama	Jitaree <i>et al.</i> , (2013)
02	Espaço de cor híbrido (RGB, XYZ, HSI e Luv) Distância de Mahalanobis Algoritmo de crescimento de regiões baseado em watershed	Classificação de estruturas histológicas relativas a proteína P53	Zhao <i>et al.</i> , (2003)
03	Espaço de cor YCbCr Modelo estatístico/ deconvolução de cor	Detecção e segmentação de coloração IHQ ⁷	Shu; Qiu; Mohammad, (2013)
04	Operações morfológicas (filtros de cor e ruído, componentes conectados e detector de bordas) Espaço de cor HSI e limiarização Ridler	Identificação de regiões celulares (núcleo, citoplasma e membrana), classificação e quantificação da reação proteica de acordo com a localização.	Ficarra <i>et al.</i> , (2006)
05	Limiarização baseada no método Otsu. Espaço de cor RGB e filtro de Textura Laws. Variação do algoritmo k-Nearest Neighbours (com sistema de pesos).	Avaliação dos receptores estrogênicos para diagnóstico e prognóstico de câncer de mama	Kostopoulos <i>et al.</i> , (2007)
06	Espaço de cor RGB, convolução de kernels derivativos 2D-Gaussiana	Classificação e quantificação de câncer de mama em microarranjos de amostras de tecidos	Amaral <i>et al.</i> , (2013)
07	Espaço de cor HSV, medidas de textura baseadas em filtros de suavização e equalização de histograma. K-means , Watershed interativo e descritores de Fourier.	Detecção de regiões foliculares em imagens de IHQ	Samsi <i>et al.</i> , (2010)
08	Espaço de cor YUV e limiarização baseada no método Otsu.	Segmentação de imagens de IHQ relativas a câncer de ovário.	Dong <i>et al.</i> , (2010 a)
09	Espaço de cor RGB, algoritmo de clusterização ISODATA	Segmentação de imagens de IHQ relativas a câncer de ovário.	Dong <i>et al.</i> , (2010 b)
10	Convolução para identificação de formas elípticas, algoritmo de corte em grafo.	Quantificação de coloração de IHQ célula a célula.	Chang <i>et al.</i> , (2008)
11	Espaço de cor YIQ e RGB. Limiarização baseada no método Otsu. Watershed com marcação inicial.	Segmentação celular em imagens de IHQ, entre fundo e núcleo e entre núcleos sobrepostos	Mao; Zhao; Tan, (2006)
12	Espaços de cor RGB e HSV. Rede Neural Multilayer Perceptron.	Quantificação de coloração específica em imagens de IHQ multi reagentes.	Wemmert <i>et al.</i> , (2013)
13	Espaço de cor RGB. Métricas de textura (entropia, desvio padrão e intervalo de valores). Algoritmo de classificação SVM ⁸ com Linear Kernel.	Segmentação de <i>lumina</i> em imagens de IHQ adquiridas de amostras de fígado de rato.	Hui; Zhiguo; Longmei, (2011)
14	Espaço de cor YCbCr e limiarização baseada	Extração automática de células positivas	Liu; Wang;

⁷ Imuno-histoquímica⁸ Support Vector Machine

	no método de entropia máxima	(tumor) em imagens de IHQ.	Fang, (2006)
15	Espaço de cor RGB. Classificador baseado na distância polinomial de Mahalanobis	Identificação de padrões de coloração em imagens de IHQ	Sobieranski <i>et al.</i> , (2009)
16	Revisão	Revisão	Cataldo; Ficarra; Macii, (2010)
17	Rede neural multicamada, discriminante de Fisher, Geodesic Active Contour Model,	Classificação de núcleos cancerosos e segmentação de núcleos sobrepostos.	Mouelhi <i>et al.</i> , (2013)
18	Transformada Wavelet (com curva Gaussiana), Hidden Markov Tree Model e Majority Vote	Segmentação de imagens extensas (entre 40 e 100 Gb)	Signolle <i>et al.</i> , (2010)
19	Espaços de cor: CIELAB, CIELUV, CIEXYZ, CMY, CMYK, HSL, HSV, Hunter-LAB, NRGB, RGB, and SCT. Classificadores: SVMs, C4.5, e Naive Bayes.	Avaliação de espaços de cor para segmentação de imagens histológicas.	Mete; Topaloglu, (2010)
20	Revisão	Revisão de métodos e softwares para análise de imagens de IHQ.	Prasad; Prabhu, (2011)

As siglas que aparecem na Tabela 13 estão relacionadas ao nome de um método textura ou espaço de cor, e são apresentadas na lista de siglas. Os termos em inglês iniciados em maiúsculo estão relacionados a nomes de métodos de classificação.

3.2 TRABALHOS RELACIONADOS

Os trabalhos relacionados à pesquisa foram escolhidos posteriormente a etapa de análise e conclusão da RSL, através de dois critérios: o primeiro critério estava diretamente relacionado ao DC, e dependia da abordagem empregar atributos de cor e textura combinados. Por outro, lado o segundo critério estava diretamente relacionado ao método de classificação, e dependia da abordagem empregar alguma técnica de classificação supervisionada.

Seguindo os critérios estabelecidos, sete trabalhos foram selecionados para serem apresentados como relacionados a presente pesquisa. O Quadro 2 permite a identificação dos autores, o ano de publicação e o conceito do veículo de publicação (segundo a avaliação da CAPES para periódicos ⁹ e conferências ¹⁰), do resultado desta seleção.

⁹ Periódicos <http://goo.gl/W7HvP4>

¹⁰ Conferências <http://goo.gl/ahxIS6>

Quadro 2. Identificação dos sete trabalhos relacionados

Autores	Ano	Conceito
Jie Shu, Guoping Qiu, Ilyas Mohammad;	2013	B3
Kostopoulos, S. Cavouras, D. ; Daskalakis, A. ; Bougioukos, P. ; Georgiadis, P. ; Kagadis, G.C. ; Kalatzis, I. ; Ravazoula, P. ; Nikiforidis, G.;	2007	B1
Amaral, T. McKenna, S.J. ; Robertson, K. ; Thompson, A.;	2013	A1
Samsi, S. Lozanski, G. ; Shanarah, A. ; Krishnamurthy, A.K. ; Gurcan, M.N.;	2010	A1
Mao, K.Z. ; Peng Zhao ; Puay-Hoon Tan;	2006	A1
Wemmert, Cedric ; Kruger, Julianne M. ; Forestier, Germain ; Sternberger, Ludovic ; Feuerhake, Friedrich ; Gancarski, Pierre;	2013	A1
Sobieranski, A.C. ; Neto, S.L.M. ; Coser, L. ; Comunello, E. ; von Wangenheim, A. ; Cargnin-Ferreira, E. ; Di Giunta, G.;	2009	B1

As subseções de 3.2.1 até 3.2.7 apresentam um resumo de cada trabalho identificado no Quadro 2, bem como uma relação com a pesquisa desenvolvida.

3.2.1 A Semi-automatic Image Analysis Tool For Biomarker Detection In Immunohistochemistry Analysis

O problema abordado nesse trabalho consiste na detecção e segmentação de coloração imuno-histoquímica, em imagens microscópicas de amostras de tecido, empregando uma abordagem semiautomática. A solução proposta é constituída de duas etapas distintas: treinamento e reconhecimento.

Na etapa de treinamento é construído um modelo estatístico baseado em histogramas, quantificados de acordo com a identificação de píxeis representativos do padrão alvo pelo usuário. Após a construção do modelo estatístico, o reconhecimento de coloração pode ser computado automaticamente, a partir do cálculo de probabilidade. A Figura 10 ilustra o fluxo de processamento proposto e o resultado de reconhecimento em uma imagem.

Foram executados testes com modelos criados por quatro usuários, comparados com dois métodos específicos: *color deconvolution* e uma abordagem baseada no espaço de cor CMYK (Cyan Magenta Yellow Black). Para a avaliação dos resultados foram calculadas, com relação a um padrão ouro estabelecido, taxas de verdadeiro positivo e falso positivo plotadas como curvas ROC (Receiver Operator Characteristic). Bem como o cálculo da área abaixo da curva AUC (Area Under a Curve).

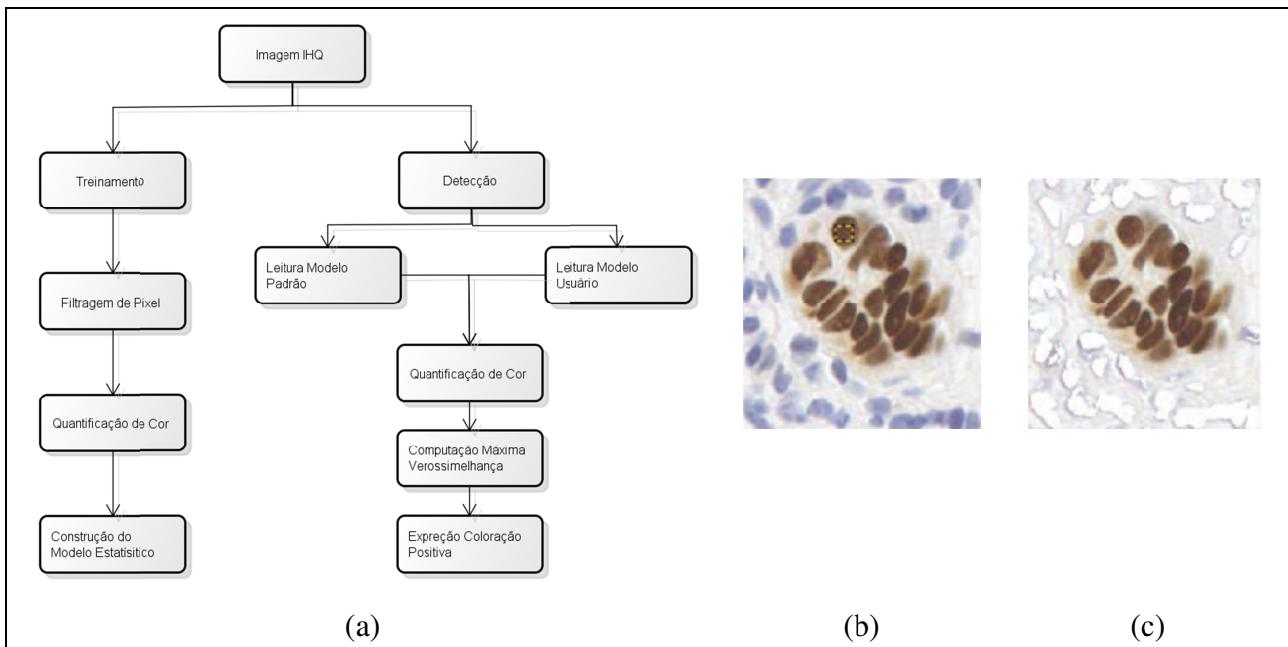


Figura 10. (a) fluxo de processamento; (b) imagem original e (c) imagem segmentada

Fonte: Shu, Qiu e Mohammad (2013).

Os resultados demonstrados para as métricas de AUC tiveram média máxima de 98,5% e mínima de 97,6%, onde 100% representaria uma classificação ótima. Mantendo resultados superiores aos métodos com os quais foram comparados.

Um problema apontado pelo trabalho diz respeito à acurácia do método, a qual depende diretamente da captação dos píxeis de treinamento, o que justifica a diferença encontrada entre modelos construídos por diferentes usuários. A abordagem foi desenvolvida em Java e disponibilizada como uma extensão do ImageJ¹¹, sem menção a disponibilidade do conjunto de imagens empregadas neste trabalho.

Abordagem empregada nesse trabalho relacionado é compatível com o trabalho desenvolvido nos seguintes aspectos:

¹¹ Software open source para processamento e análise de imagens, desenvolvido em Java.
<http://rsbweb.nih.gov/ij/index.html>

- Primeiro no emprego de um método de classificação estatístico com uma abordagem supervisionada, dividida entre treinamento e segmentação; e
- Análise de Resultados: testes feitos através de observações de verdadeiros positivos e falsos positivos, quando comparados a um padrão ouro fornecido.

3.2.2 Colour-Texture Based Image Analysis Method For Assessing The Hormone Receptors Status In Breast Tissue Sections

O problema abordado nesse trabalho consiste na avaliação da condição de receptores (proteínas) estrogênicos no diagnóstico e prognóstico de câncer de mama, empregando análise de imagens. A solução proposta emprega um método de análise de imagem baseada em atributos de cor e textura, composta por uma etapa de segmentação de núcleos celulares e uma etapa de classificação.

Inicialmente as imagens são convertidas em tons de cinza e limiarizadas com o algoritmo Otsu, a fim de segmentar núcleos celulares. Neste ponto é possível aplicar ajustes no parâmetro de limiarização, através da interface gráfica construída em MatLab¹². Para finalizar a segmentação, uma sequência de algoritmos morfológicos foi aplicada para correção de ruídos e eliminação de núcleos com menos de 300 píxeis, o parâmetro destes algoritmos morfológicos foram mantidos fixos para as imagens analisadas.

Para a extração de características foram empregadas medidas baseadas no método Laws de energia de textura, avaliando as cinco máscaras L5, E5, S5,W5 e R5. Adicionalmente são incorporadas medidas de cor aos descritores de características, ao final cada descritor contém 60 medidas (cor/textura). A fim de reduzir essa dimensionalidade, (60 dimensional) bem como o esforço computacional, os descritores são reduzidos empregando o teste estatístico não paramétrico Wilcoxon, escolhendo-se as 15 medidas mais bem ranqueadas pelo teste. Na Figura 11-a, é apresentado o resultado da avaliação dos descritores, indicando as combinações de métricas de textura empregado no DC.

¹² Linguagem de programação, visualização e computação numérica - <http://www.mathworks.com/products/matlab/>

Finalmente estes descritores de característica 15-dimensional são utilizados no treinamento e classificação dos núcleos celulares, utilizando uma variação do algoritmo k-Nearest Neighbours que emprega um sistema de votos. Foram avaliadas 22 imagens, totalizando 921 núcleos (1/3 destes foi usado para treinamento).

Nesse trabalho a avaliação foi conduzida, através de medidas de acurácia da ferramenta, contra uma avaliação feita manualmente, por um especialista médico. Como resultado foram classificadas corretamente 19/22 casos, 86,8% de acurácia. A Figura 11-b apresenta dois resultados de reconhecimento de núcleos celulares, para as colorações marrom e azul.

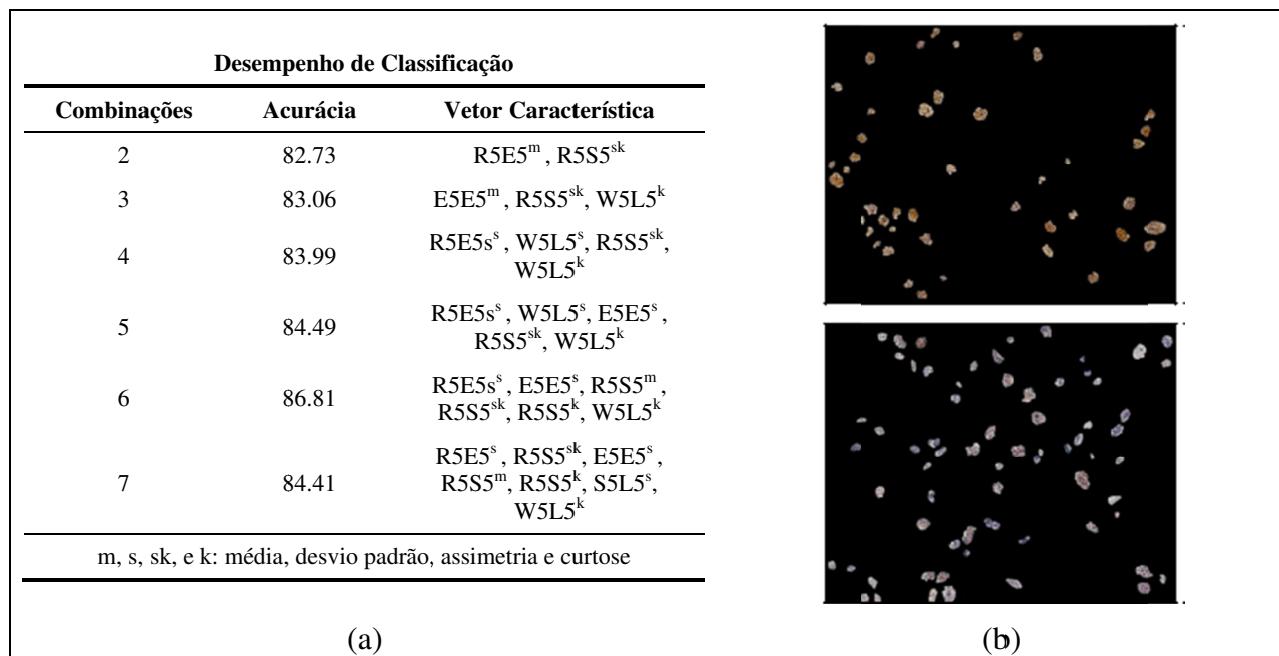


Figura 11. (a) máscaras, medidas estatísticas e desempenho e (b) imagem de núcleos reconhecidos
Fonte: Kostopoulos *et al.* (2007).

Abordagem empregada nesse trabalho relacionado é compatível com o trabalho desenvolvido nos seguintes aspectos:

- Prevendo um método supervisionado (etapa de treinamento), empregando atributos de cor e textura como descritores de características. Entretanto a abordagem vai além da quantificação de coloração imuno-histoquímica, provendo uma etapa para diagnóstico de câncer.

3.2.3 Classification And Immunohistochemical Scoring Of Breast Tissue Microarray Spots

O problema abordado nesse trabalho relacionado consiste na classificação e quantificação de câncer de mama, em imagens de micro arranjos de tecido. A quantificação imuno-histoquímica é realizada através da computação da proporção de epitélio nuclear corado, bem como a intensidade dessa coloração.

A solução proposta por esta abordagem emprega o conceito de *bag of visual words*, para a computação de descritores de características. Esses descritores são compostos de 15 dimensões, as três primeiras representam a cor no espaço RGB, as 12 restantes são compostas por medidas de convolução de um grupo de máscaras derivadas de primeira e segunda ordem (2D-Gaussiana). Estas medidas foram escolhidas devido à sua característica de invariância a rotação e orientação. A Figura 12 ilustra o fluxo de processamento da abordagem proposta.

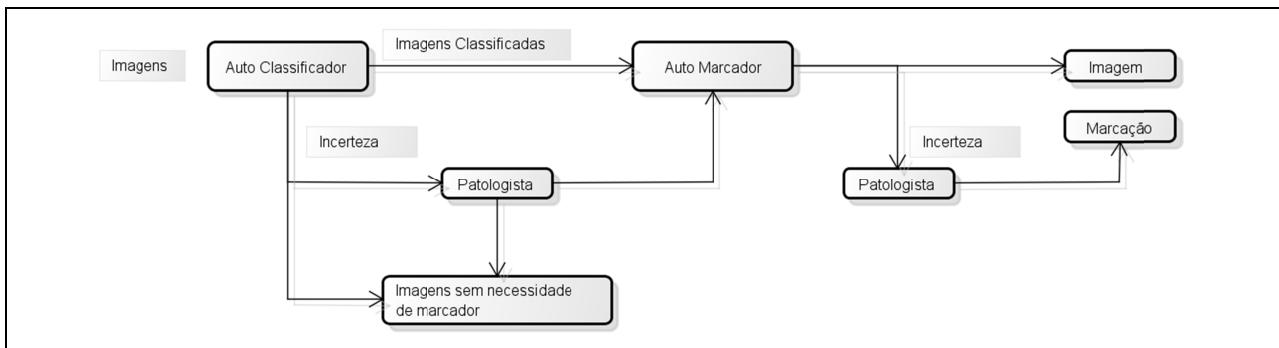


Figura 12. Visão geral do fluxo de processamento da abordagem proposta

Fonte: Amaral *et al.* (2013).

Para a avaliação da solução proposta foram executados testes entre três classificadores distintos, Multilayer Perceptron, Latent Topic Models e Support Vector Machine (SVM), na classificação dos núcleos do micro arranjo. Para a avaliação da quantificação, os classificadores foram comparados através dos métodos de regressão gaussiana e regressão linear, onde foram computadas medidas de acurácia de classificação, entre a abordagem proposta e os diferentes classificadores.

Como critério de comparação, para a avaliação dos resultados, foi empregando um conjunto de anotações feitas por um patologista. Os melhores resultados observados foram os computados através do algoritmo SVM.

Abordagem empregada nesse trabalho relacionado é compatível com o trabalho desenvolvido nos seguintes aspectos:

- Combinando atributos de cor e métricas semelhantes à textura em um único descritor de características (15-dimensional). Entretanto esta abordagem vai além da quantificação de padrões de coloração, proporcionando essa quantificação em localizações celulares específicas.

3.2.4 Detection of Follicles From IHQ-Stained Slides Of Follicular Lymphoma Using Iterative Watershed

O problema tratado por esse trabalho relacionado consiste na detecção de regiões foliculares em imagens microscópicas de imuno-histoquímica. A solução proposta é baseada na segmentação de Rough, empregando um descritor de características de três elementos (uma medida de cor e duas de textura).

Para a construção do DC inicialmente ocorre a conversão da imagem para o espaço de cor HSV (Hue Saturation Value). A informação do canal *hue* (H) é extraída para compor a primeira parte do descritor. A segunda parte do descritor é composta por uma medida de textura, a partir da aplicação de um filtro de suavização, seguido de uma equalização. Valores decorrentes da equalização do histograma são utilizados, como segunda parte do descritor.

A última parte do DC é computada calculando a matriz de coocorrência em níveis de cinza. Este cálculo provê um método estatístico para a caracterização da dependência espacial de níveis de cinza na imagem. A segmentação propriamente dita (regiões de folículo e não folículo) ocorre com o emprego do algoritmo *k*-Means, um algoritmo não supervisionado para o agrupamento de dados em *k* classes. Um algoritmo baseado em watershed iterativo é aplicado para executar a separação de folículos sobrepostos. A Figura 13-a ilustra o fluxo de processamento da abordagem e a Figura 13-b ilustra a plotagem das classes identificadas, após a aplicação do algoritmo *k*-Means e a imagem resultante.

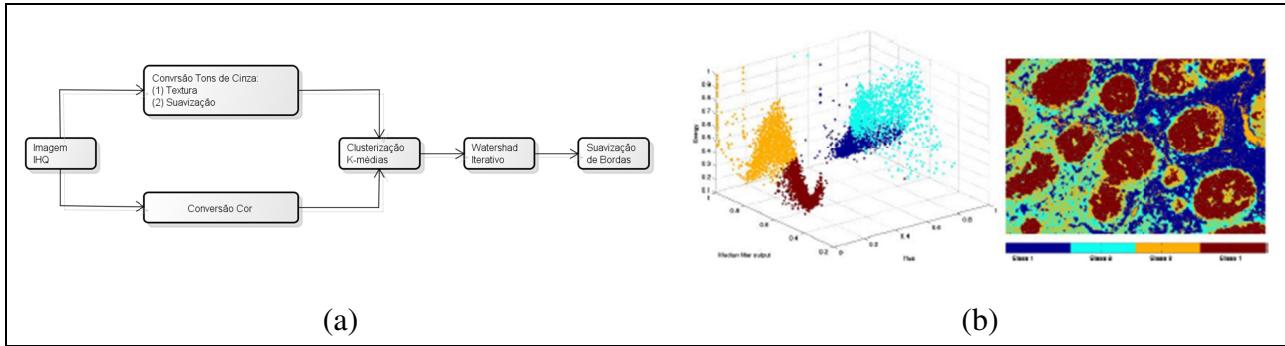


Figura 13. (a) esboço da abordagem proposta e (b) plotagem após clusterização

Fonte: Sansi *et al.* (2013).

Finalmente uma etapa de suavização, empregando descritores de Fourier, é executada para suavizar as delimitações dos folículos. Para a avaliação dessa proposta foi empregado um conjunto de imagens, comparando a segmentação do método proposto, contra a segmentação executada por um patologista. O resultado mostrou uma média de similaridade de 87,11%;

Abordagem empregada nesse trabalho relacionado é compatível com o trabalho desenvolvido nos seguintes aspectos:

- Utilizando um descritor de características, combinando atributos de cor e textura.

3.2.5 Supervised Learning-based Cell Image Segmentation For P53 Immunohistochemistry

O problema abordado neste trabalho relacionado trata da segmentação celular em imagens microscópicas de imuno-histoquímica, entre fundo, núcleo e entre núcleos sobrepostos, sendo apresentados dois novos métodos para segmentação de imagens celulares.

Inicialmente é apresentada a abordagem C-G-T, em que a métrica de transformação da imagem colorida (C), em tons de cinza (G), é aprendida sobre a distribuição global do espaço de cor, já a métrica de limiarização (T) é aprendida localmente, na distribuição de píxeis, em tons de cinza (empregando o algoritmo Otsu). Esta abordagem hibridizando técnicas de aprendizagem locais e globais mostra eficiência em classificar objetos de interesse (células) do fundo.

Entretanto esta primeira abordagem não é eficiente para a segmentação de núcleos sobrepostos ou que se tocam. Para esse problema, separação de núcleos, é apresentada uma abordagem baseada no algoritmo Watershed, no qual um conjunto de marcações é extraído nas

linhas de vales, tanto na imagem em tons de cinza, como na imagem da distância complementar, a fim de guiar a segmentação, a Figura 14-a e Figura 14-b demostram estas marcações e a plotagem topográfica da imagem.

Diversos testes e métricas foram utilizados para avaliar a solução proposta nesse trabalho. Inicialmente foi avaliada a abordagem de conversão (cor – tons de cinza), empregando três medidas: Geometric Margin (GM), Fisher Ratio (FR) e separabilidade de classe, baseada em dispersão (SM), comparadas contra cinco outros métodos.

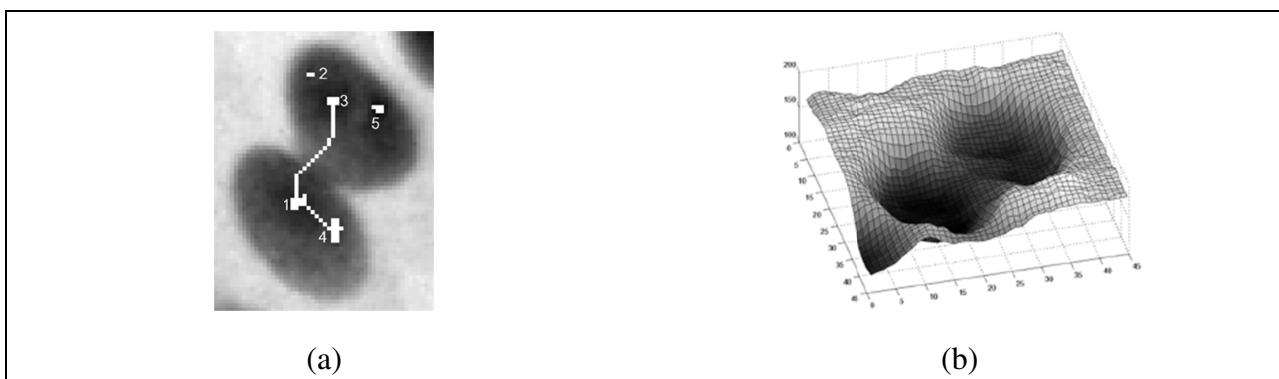


Figura 14. (a) extração da linha de vale e (b) superfície topográfica da imagem

Fonte: Mao; Zhao e Tan (2013).

O método proposto apresentou melhores resultados quando comparado a outros métodos, 10,977 - 28,683 - 14,370 para GM, FR e SM respectivamente. Contra 5,264 - 21,550 - 8,941 de GM, FR e SM respectivamente, no pior método avaliado.

Para a avaliação dos resultados de segmentação (limites) foram empregadas três métricas, distância de Hausdorff (DH), média da distância absoluta (DA), bem como uma avaliação estatística que leva em conta a quantidade de regiões segmentadas, a soma da distância Euclidiana entre os vetores de cor dos pixels das regiões e a área destas regiões, essa métrica foi denominada de Q(I). Para a avaliação dos resultados de segmentação (área) foram computadas as seguintes medidas de fração de área: falso negativo (FN), falso positivo (FP) e verdadeiro positivo (VP).

Novamente o método proposto teve melhor desempenho, quando comparado aos demais métodos, valores alcançados foram 2,234 - 0,837 - 0,022 - 0,044 - 0,965 para DH, DA, FP, FN e VP respectivamente. Contra 13,416 - 3,594 - 0,400 - 0,143 - 0,900 para DH, DA, FP, FN e VP

respectivamente (do pior método avaliado). As medidas de Q(I), para o método proposto tiveram a média (avaliadas 9 imagens) de 6,24, contra uma média de 13,47 do pior método avaliado.

Abordagem empregada nesse trabalho relacionado é compatível com o trabalho desenvolvido nos seguintes aspectos:

- Adotando um método supervisionado de classificação e uma etapa inicial de descrição de características empregando atributos de cor. Entretanto o foco fica na segmentação de núcleo celular.

3.2.6 Stain Unmixing In Bright Field Multiplexed Immunohistochemistry

O problema tratado nesse trabalho relacionado diz respeito à quantificação de colorações específicas em imagens de imuno-histocímica com múltiplos reagentes. A solução apresentada é baseada em três etapas, ocorrendo inicialmente a homogeneização de cores na pseudo imagem RGB, gerada a partir da aquisição feita pelo sistema de imagem multiespectral e a câmera CCD¹³. Esta homogeneização é alcançada usando uma técnica de especificação de histogramas, tanto no espaço de cores RGB, quanto no espaço de cores HSV.

Na segunda etapa ocorre o treinamento de uma rede neural (Multilayer Perceptron -MPL) para a aprendizagem da transformação de decomposição do espaço de cor, da imagem de entrada, nos três canais de interesse marrom, rosa e azul.

A rede neural é composta por três neurônios de entrada (um para cada valor de RGB), duas camadas intermediárias de seis neurônios, e na saída, três neurônios, que fornecem a separação dos padrões das classes possíveis (HLN-coloração azul; DAB-coloração marrom e PRD-coloração rosa), a Figura 15-a ilustra esta estrutura.

¹³ Dispositivo de Carga Acoplada – Charge Coupled Device, transforma luz em sinal digital.

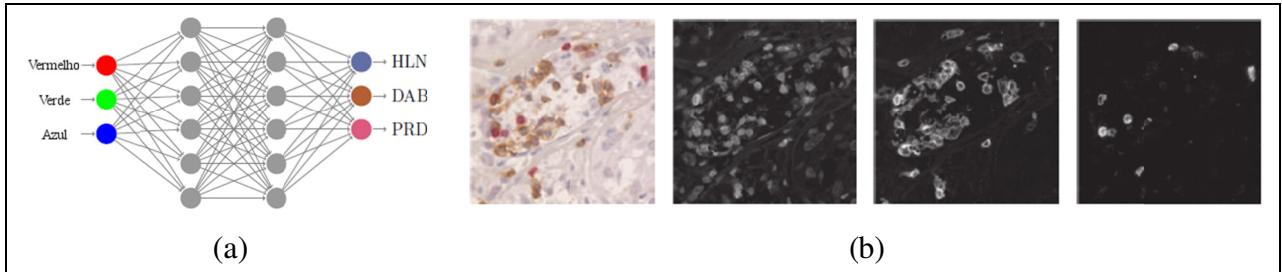


Figura 15. (a) estrutura da rede neural empregada e (b) resultados de classificação, imagem original, classe HLN, classe DAB e classe PRD (respectivamente)

Fonte: Wemmert *et al.* (2013).

Na terceira etapa ocorre a transformação de todas as imagens, a partir da métrica aprendida durante o treinamento. Os testes foram aplicados a um conjunto de 33 imagens comparadas, contra anotações; foram computadas a média e o desvio padrão da diferença entre cada imagem e seu respectivo GT, a Figura 15-b demonstra o resultado de identificação para uma imagem. Comparada a outros dois métodos, a abordagem proposta teve média de diferença de 0,171, contra 0,403 e 0,381 dos outros dois métodos (deconvolução de Laak e deconvolução de Rui frok);

Abordagem empregada nesse trabalho relacionado é compatível com o trabalho desenvolvido nos seguintes aspectos:

- Utilizando informações de cor (RGB) como descritor de características para a segmentação de imagens de imuno-histoquímica (IHQ), entretanto o DC utiliza somente atributos de cor; e
- Utilizando um método para classificação baseado em redes neurais (supervisionado).

3.2.7 Learning a Nonlinear Color Distance Metric For The Identification Of Skin Immunohistochemical Staining

O problema tratado nesse trabalho relacionado diz respeito à identificação de padrões de coloração em imagens microscópicas de IHQ, no espaço cromático RGB. A solução proposta é baseada em duas etapas (treinamento e classificação).

A primeira etapa é baseada em um processo de aprendizagem supervisionado, dividida em duas subetapas. Na primeira subetapa, o especialista indica um conjunto de pontos (na imagem) que definem o padrão alvo, na segunda subetapa o conjunto de pontos é utilizado para treinamento do

classificador (baseado na distância polinomial de Mahalanobis), a Figura 16-a apresenta uma visão geral da abordagem proposta, entre as etapas de treinamento e classificação e a Figura 16-b presenta um resultado de identificação do padrão alvo.

Na segunda etapa, a partir da métrica de distância aprendida na primeira etapa, o usuário fornece um grupo de imagens para classificação píxel a píxel, entre padrão de interesse ou não, de toda a imagem ou de uma região menor, que seja do mesmo contexto da imagem onde foram extraídos os pontos para treinamento.

Foram avaliadas 15 imagens microscópicas de amostras de pele, submetidas a um processo de IHQ, comparadas com um *Ground Truth* (GT) feito manualmente por um especialista. As métricas para avaliação foram correlação de Pearson (0,966), Sensibilidade (0,99), Especificidade (0,85) e *fitness* (0,92).

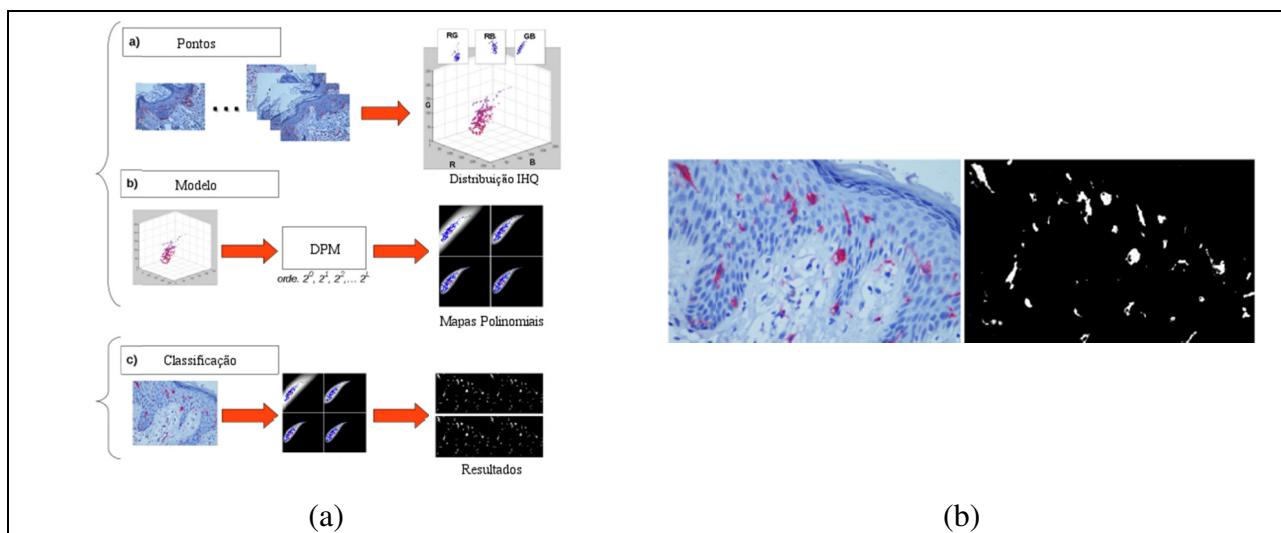


Figura 16. (a) visão geral da abordagem proposta e (b) resultado de reconhecimento, imagem original e segmentada

Fonte: Sobieranski *et al.* (2013).

Abordagem empregada nesse trabalho relacionado é compatível com o trabalho desenvolvido nos seguintes aspectos:

- A pesquisa desenvolvida estendeu esse trabalho, incorporando atributos de textura ao descriptor de características, utilizando o mesmo classificador.

3.3 ANÁLISE COMPARATIVA

Segundo Cataldo, Ficarra e Macii (2012) fluxos de processamento na análise de imagens de IHQ podem ser divididos em: (i) separação da contribuição de rótulo (segmentação de cor, deconvolução de cor e técnicas multiespectrais); (ii) localização de antígeno alvo (compartimentalização de tecidos e segmentação celular); e (iii) quantificação de IHQ (classificação entre corado/não corado, porcentagem de células coradas, continuidade de coloração e técnicas para quantificação).

Seguindo essa divisão o trabalho proposto pretendeu contribuir especificamente em dois pontos, na rotulação (segmentação de cor) e na quantificação de IHQ (classificação entre corado e não corado).

Dos 20 trabalhos avaliados, na RSL, 11 empregaram um método de classificação supervisionado (que depende de uma etapa de treinamento para construção do classificador), como por exemplo, Jitaree *et al.* (2013), Mouelhi, A. *et al.* (2013) e Wemmert *et al.* (2013), baseados em redes neurais. Bem como, Mete; Topaloglu (2010) e Hui; Zhiguo; Longmei (2011). baseados em SVM.

Dos 20 trabalhos avaliados, na RSL, nove não usaram um método de classificação supervisionado, adotando abordagens baseadas principalmente em três técnicas: (i) limiarização, com o método Otsu, como apresentado em Dong *et al.*, (2010) a; (ii) crescimento de regiões, principalmente baseado no algoritmo Watershed como apresentado em Samsi *et al.* (2010) e Zhao *et al.* (2003); (iii) ou ainda ,uma combinação das técnicas (i) e (ii), como apresentado em Mao; Zhao e Tan (2006).

Dos 20 trabalhos avaliados, 13 utilizam algum espaço de cor específico, durante o processo de segmentação/classificação e seis empregam alguma métrica de textura no processo de segmentação/classificação, tais como os apresentados por Jitaree *et al.* (2013), Kostopoulos *et al.* (2007), Amaral *et al.* (2013), Samsi *et al.* (2010), Hui, Zhiguo e Longmei (2011) e Signolle *et al.* (2010).

A maioria (11 de 20), dos trabalhos avaliados durante a RSL, empregaram algum método de classificação supervisionada, onde foi possível identificar dois métodos em particular, MP e SVM. Entre os trabalhos analisados, treze empregaram cor como descritor de características, e seis destes

combinaram métricas de textura no DC. Estes números contribuem para as justificativas das decisões tomadas para a pesquisa desenvolvida, combinação de cor e textura no DC, e classificação supervisionada.

Duas oportunidades de possíveis trabalhos futuros foram identificadas, que podem dar continuidade ao trabalho desenvolvido, a segmentação de estruturas celulares específicas, tais como a segmentação de núcleos celulares apresentada por Liu; Wang; Fang (2006) e Mouelhi *et al.* (2013).

A segunda possibilidade de pesquisa futura identificada foi a de trabalhar com processamento de imagens extensas, onde lâminas virtuais são criadas a partir da digitalização com equipamentos específicos. Estas lâminas virtuais são imagens que podem atingir tamanhos de 40Gb até 100 Gb, com dimensões de até 65000x45000 píxeis, conforme apresentado por Signolle *et al.* (2010), onde o processo de segmentação/classificação de um único slide virtual pode levar até 50 horas.

3.4 CONSIDERAÇÕES

A Subseção 1.3.2 aponta dois procedimentos metodológicos. Neste capítulo foi apresentado o resultado do primeiro deles (pesquisa bibliográfica). Conduzido através de uma revisão sistemática da literatura, bem como a identificação e apresentação de sete trabalhos relacionados. Esta etapa de trabalho contribuiu para contextualizar o problema de pesquisa, bem como as contribuições esperadas na área de aplicação definida (identificação e quantificação de padrões de coloração em IHQ).

O Quadro 3 compara a proposta de trabalho (última linha), com os sete trabalhos relacionados. A coluna T indica o trabalho, enquanto a coluna Descritor indica o método de segmentação/classificação e a abordagem de aprendizado, e a coluna Avaliação indica as métricas de avaliação empregadas nos trabalhos.

Os trabalhos R2, R3 ajudam a comprovar a viabilidade da proposta, uma vez que têm a mesma abordagem, empregando cor e textura com um método de classificação supervisionada. O trabalho R4 tem um princípio ligeiramente diferente, por empregar um método não supervisionado. Esses três trabalhos demonstram resultados, em pelo menos um dos testes apresentados, com alto

grau de acerto, onde um valor de 100% representaria uma classificação ótima. O trabalho R2 contribuiu ainda para a definição das medidas estatísticas utilizadas.

A revisão na literatura contribuiu para uma adequada contextualização do problema de pesquisa, na descrição e classificação de padrões em imagens de IHQ. A abordagem proposta segue os princípios de descrição e classificação encontrados na maioria dos trabalhos relacionados. Entretanto, o diferencial da proposta consiste em combinar o classificador baseado no método DPM, com atributos de cor e medidas de textura baseadas na abordagem LBP.

Quadro 3. Comparação entre os trabalhos relacionados e a proposta

T	Descriptor	Método	Avaliação
R1	Cor YCbCr	Modelo estatístico de deconvolução, supervisionado	Área sob a curva ROC
R2	Textura de Laws, medidas estatísticas e Cor RGB.	k-NN com votos ponderados, supervisionado	Erro médio, Acurácia.
R3	Cor RGB, Textura 2D-Gaussian.	BVS, k-Means clusterização, supervisionado	Erro Absoluto, Acurácia e Entropia
R4	Cor HSV, Textura Matriz de Coocorrência medidas estatísticas.	k-Means, Watershed, não supervisionado	Escore de similaridade
R5	Cor RGB, Tons de cinza	Limiarização, Watershed supervisionado	Média de distância absoluta, acurácia, super e sub segmentação
R6	Cor RGB	MLP, supervisionado	Diferença de segmentação entre o trabalho e metadados.
R7	Cor RGB	DPM, supervisionado	SEM, ESP e FIT
P	Cor RGB, Textura LBP, medidas estatísticas	DPM, supervisionado	SEM, ESP, PRE, ACC, TFP, TFN e TFD

Outra contribuição advinda da revisão da literatura diz respeito à identificação de duas linhas de pesquisas futuras, que podem ser investigadas como extensão do mestrado. A primeira e mais imediata consiste em classificar, não só padrões de IHQ, mas também sua localização espacial. Já a segunda, trata do processamento de alto desempenho (considerando velocidade) de grandes volumes de dados.

4 DESENVOLVIMENTO

Os objetivos específicos 1, 2 e 3, descritos na Subseção 1.2.2 , trataram da definição dos atributos de textura (método e medidas estatísticas) utilizados na proposta deste trabalho. A Seção 4.1 apresenta o descritor de características e na Seção 4.2 é apresentado o método de classificação empregado.

Na Seção 4.3 é apresentado o resultado da etapa de implementação do protótipo, necessário para atender os objetivos específicos 4 e 5. Na Seção 4.4 é apresentada a definição dos experimentos realizados e na Seção 4.5 uma descrição do conjunto de dados.

4.1 PROPOSTA DO DESCRIPTOR

O método escolhido para gerar as respostas de textura foi o LBP, uma vez que emprega uma abordagem simples de implementação, gerando uma de resultado, quando comparado ao método de Laws que gera 25 imagens de resultado. O método LBP apresenta ainda as vantagens de ser invariante a iluminação, rotação e ruídos, conforme referenciado na Subseção 2.2.3

As medidas estatísticas escolhidas para compor o DC foram baseadas naquelas utilizadas no trabalho relacionado, Kostopoulos *et. al* (2007), média, desvio padrão, assimetria, curtose, e energia. Tais medidas serão computadas em cinco vizinhanças distintas, a fim de descrever um único píxel. Desta forma, o DC foi definido com três componentes de cor e vinte componentes de textura.

4.1.1 Descritor Baseado no Método LPB

A partir da imagem filtrada (I_f), resultante da aplicação do método LBP, foram definidas cinco medidas estatísticas (F) que serão calculadas para compor o descritor do píxel, estas medidas serão computadas em quatro Janelas de Amostragem (JA) distintas de dimensões 3x3, 5x5, 7x7, 9x9, tendo o píxel, a ser descrito, como centro da janela. As medidas estatísticas definidas foram média (F_m), desvio padrão (F_d), assimetria (F_a), curtose (F_c) e energia (F_e), definidas em (27), (28), (29), (30) e (31) respectivamente:

$$F_m = \frac{\sum_{i=0}^M \sum_{j=0}^N [I_{f(i,j)}]}{M * N} \quad (27)$$

$$F_d = \sqrt{\frac{\sum_{i=0}^M \sum_{j=0}^N (I_{f(i,j)} - S_m)^2}{M * N}} \quad (28)$$

$$F_a = \frac{\sum_{i=0}^M \sum_{j=0}^N (I_{f(i,j)} - S_m)^3}{M * N * S_d^3} \quad (29)$$

$$F_c = \frac{\sum_{i=0}^M \sum_{j=0}^N (I_{f(i,j)} - S_m)^4}{M * N * S_d^4} \quad (30)$$

$$F_e = \frac{\sum_{i=0}^M \sum_{j=0}^N (I_{f(i,j)})^2}{M * N} \quad (31)$$

onde M e N são as dimensões de JA e $I_{f(i,j)}$ representa o valor do píxel na imagem filtrada com o método LBP, na posição i, j . A Figura 17 ilustra duas janelas de amostragens de dimensões distintas sobre o resultado de filtragem.

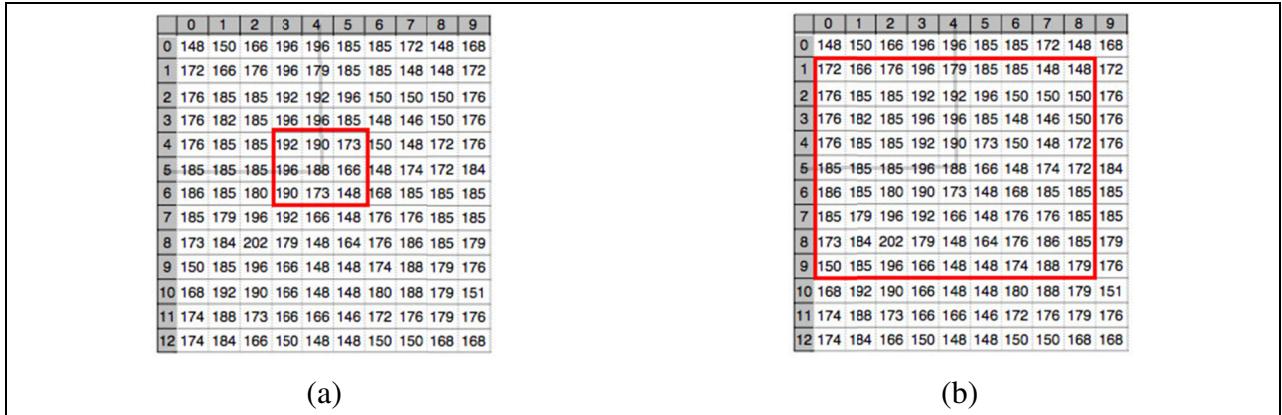


Figura 17. Exemplos de janelas de amostragem: a) dimensão 3x3; e b) dimensão 9x9

Na Figura 17 são representadas a menor e a maior janela de amostragem, em que os valores da matriz indicam a intensidade do píxel filtrado pelo método LBP, nas posições linha e coluna (indicadas em cinza). Na Figura 17-a o píxel sendo descrito é o de posição 5x4 (linha x coluna) de valor 188, as linhas vermelhas delimitam a vizinhança (JA) do píxel em questão, com dimensão de

3x3, onde serão computadas as medidas estatísticas apresentadas. A Figura 17-b segue o mesmo princípio, porém com uma JA com dimensão de 9x9 (maior JA utilizada neste trabalho).

O descritor de características baseado no método LBP foi definido em duas partes distintas, uma relacionada por cor no espaço cromático RGB e outra definida por medidas de textura, adotando a forma em:

$$\mathbf{d}_{23} = [r, g, b, \ t_1, \dots t_N] \quad (32)$$

onde r , g e b formam a componente definida por cor e representam os valores de vermelho verde e azul do píxel em questão. Os valores de t formam a componente definida por textura, e N representa a quantidade destas medidas usadas no descritor.

Inicialmente foram consideradas as quatro janelas de amostragem e as cinco medidas estatísticas, fornecendo um descritor inicial 23-dimensional, denominado de \mathbf{d}_{23} .

4.2 MÉTODO DE CLASSIFICAÇÃO

O método de classificação empregado é baseado no modelo Bayesiano, uma vez que ocorre em duas etapas: (i) Treinamento e (ii) Classificação. Segue a mesma abordagem apresentada em Sobieranski *et al.* (2009), diferindo no descritor de característica utilizado.

A etapa (i) consiste em um método de aprendizagem supervisionada, subdividida em duas fases: (i-a) geração dos descritores e (i-b) calibração do modelo. A Figura 18 e a Figura 19 mostram uma visão geral da etapa (i). A Figura 20 ilustra a etapa de classificação.

4.2.1 Etapa i: Treinamento

Na fase (i-a) são gerados os descritores de características para formar o conjunto de treinamento, cujos dados serão usados como entrada (S), na aprendizagem da métrica de distância. A Figura 18 ilustra uma imagem original, a seleção de pontos para treinamento (marcados em preto) e um histograma 3-dimensional, representando a distribuição de coloração alvo (avermelhada), para qual a função de discriminação foi modelada.

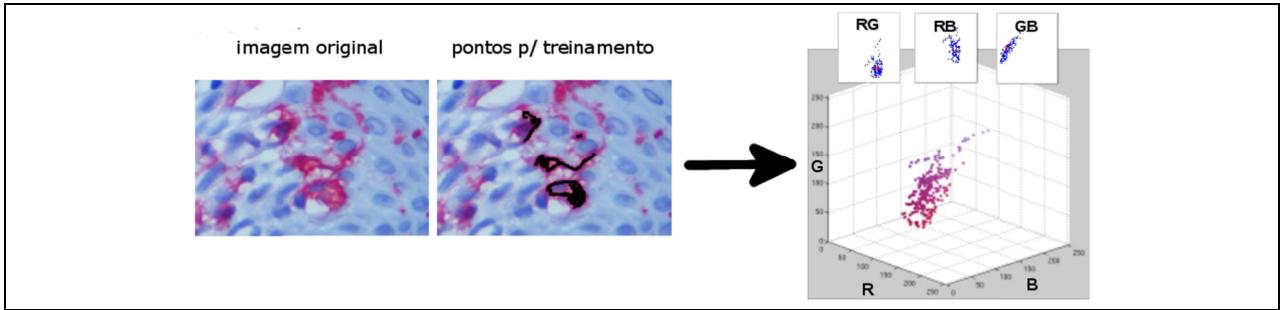


Figura 18. Visão geral da etapa de treinamento, extração dos descritores (fase i-a)

Fonte: Adaptado de Sobieranski *et al.* (2009).

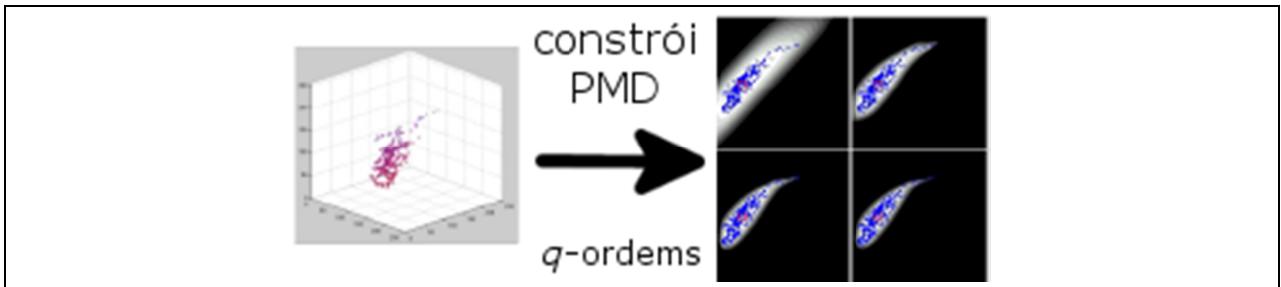


Figura 19. Visão geral da fase de treinamento, calibração do modelo, etapa (i-b)

Fonte: Adaptado de Sobieranski *et al.* (2009).

Na fase (i-b) da etapa de treinamento ocorre a calibração do modelo, onde a métrica de distância é estabelecida. A Figura 19 ilustra a construção dos mapas polinomiais, para cada ordem, onde os pontos em azul correspondem à distribuição do padrão alvo.

Pontos dentro da área branca são considerados pertencentes ao padrão alvo, as áreas em preto são consideradas não pertencentes (as quatro imagens à direita, com ordens polinomiais distintas). Esta fase pode ser explicada da seguinte maneira:

- Os descritores gerados são usados como entrada (S), conforme apresentado na Seção 2.4, na etapa de aprendizagem do método DPM;
- O parâmetro L , como explicado na Seção 2.4, controla o número de projeções polinomiais; e
- Os mapas polinomiais são então obtidos, um para cada q -ordem até L . A métrica é um método projetivo, todas as ordens construídas ($q \leq L$) ficam disponíveis para uso sem a necessidade de um novo treinamento. A Figura 19 demonstra a obtenção dos mapas construídos pela DPM. Os pontos em azul representam os dados, a coloração branca

representa o espaço de classificação e a coloração preta representa fundo (o que deve ser excluído do espaço de classificação).

4.2.2 Etapa ii: Classificação

Na etapa (ii) classificação, após a aprendizagem da métrica e a construção dos mapas polinomiais, durante a etapa de calibração do modelo (fase i-b), é possível classificar novas imagens desde que pertençam ao mesmo contexto do conjunto de treinamento. A Figura 20 apresenta esta etapa, definida da seguinte forma:

- Um grupo de imagens é selecionado para ser classificado, pelos mapas polinomiais, diferentes das imagens usadas no treinamento. A única restrição é que sejam imagens do mesmo contexto (p. ex. geradas com o mesmo agente de coloração);
- Com a definição das imagens cada DC é gerado e classificado (píxel a píxel), com a computação da similaridade, através da métrica aprendida; e
- Finalmente, para cada imagem um mapa de similaridade que corresponde as intensidades de coloração de IHQ, é obtido, relacionando as áreas das imagens reconhecidas como o padrão de interesse.

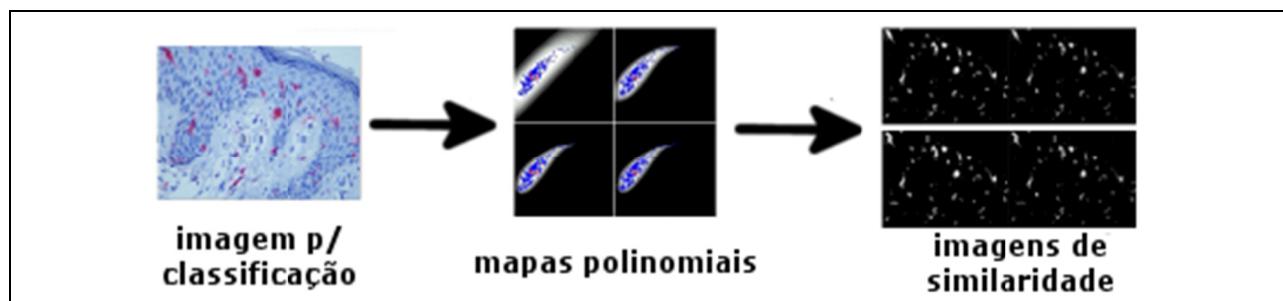


Figura 20. Visão geral da etapa de classificação

Fonte: Adaptado de Sobieranski *et al.* (2009).

As altas ordens polinomiais proporcionam alto grau de adaptação a distribuições não lineares, em um dado espaço de característica. Esta adaptação aumenta dependendo da q -ordem empregada, bem como da representatividade da amostra de treinamento. Uma etapa de normalização é efetuada, para aumentar a visualização, no mapa de similaridade, formalmente:

$$I_{(x,y)} = e^{(-\lambda d_{PM(x,y)})} \quad (33)$$

onde $d_{PM(x,y)}$ é a DPM, no píxel (x,y) , descrita na Seção 2.4 , e $\lambda > 0$ é uma constante usada para definir o grau do mapa de polinomial. Esta etapa de normalização resulta em valores variando entre 0 e 1, quanto mais próximo de 1 (píxeis em branco, nas imagens de similaridade da Figura 20) este valor, mais similar ao padrão alvo é o descritor.

Por outro lado, quanto mais próximo de zero, maior é o grau de não similaridade (píxeis em preto, nas imagens de similaridade da Figura 20). No Anexo A, é apresentado uma exemplo das etapas de treinamento e classificação usando a DPM e uma comparação com a distância euclidiana, que mantém um espaço de classificação circular e de pesos iguais em todas as direções, contra a capacidade de adequação ao padrão alvo da DPM à medida que aumenta o grau polinomial.

4.3 DESCRIÇÃO DA IMPLEMENTAÇÃO DO PROTÓTIPO

O código do algoritmo de classificação DPM utilizado foi o mesmo empregado por Carvalho *et al.* (2014). Esta versão trata a questão da variabilidade na quantidade de valores do descritor (dimensão). O restante das implementações efetuadas, no presente trabalho utilizaram o código DPM basicamente como uma biblioteca, fazendo chamadas a três funções:

- **carga:** função que lia um arquivo com os descritores para treinamento;
- **treinamento:** função que realizava a aprendizagem, para a métrica de distância de acordo com o grau polinomial definido; e
- **classificação:** função que recebia um descritor e retornava um valor entre 0 e 1. Quanto mais próximo a 1, menor a distância entre o descritor passado por parâmetro e o conjunto de treinamento.

O algoritmo para gerar a imagem resultante, do método de textura LBP tradicional, foi implementado do zero. Devido à resposta da abordagem tradicional, optou-se por utilizar uma

variação conhecida como LBP *Variance*. Uma implementação deste método está disponível no GitHub¹⁴ de Philipp Wagner¹⁵, e foi utilizada no presente trabalho.

Na Figura 21-a foi apresentada uma imagem original utilizada nos testes de desempenho. Na Figura 21-b foi possível observar o resultado do processo de textura para o método LBP tradicional e na Figura 21-c, mostra o resultado do método LBP baseado em variância.

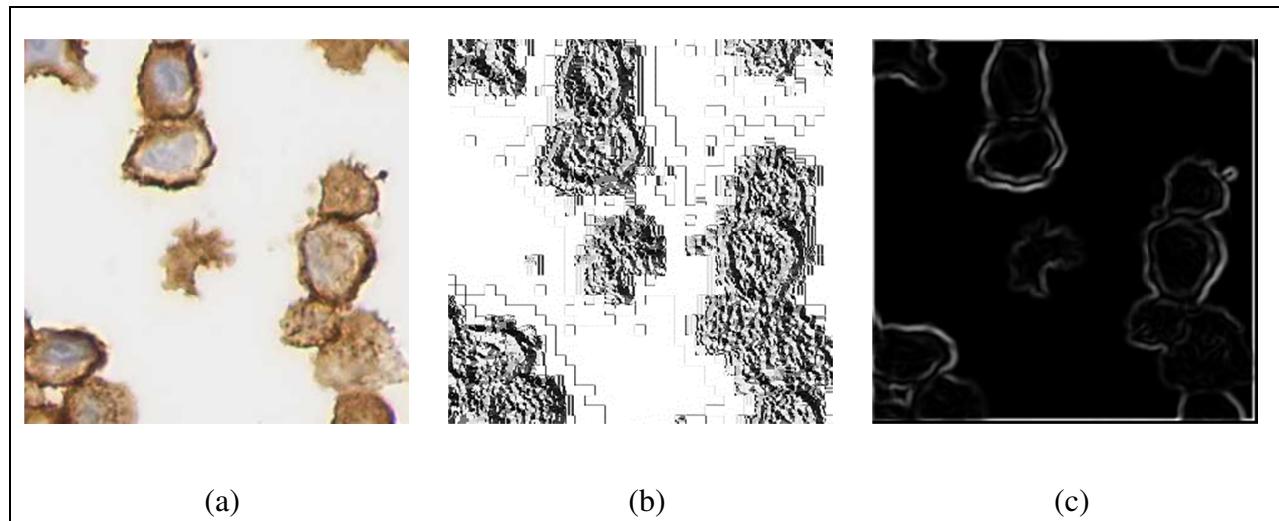


Figura 21. Imagens resultantes da filtragem por textura: a) Imagem original; b) Textura LBP tradicional e c) Textura LBP baseado em variância

O restante das implementações executadas compreenderam a geração do descritor, a criação dos arquivos para treinamento, a comparação das imagens resultantes do processo de classificação com as imagens de anotação, a geração das imagens temáticas e a criação dos arquivos com o resultado da matriz de contingência.

A Figura 22 apresenta um diagrama de classe simples, apontando em azul o que foi implementado para este trabalho, e em verde (DPM e LBP Variância) o que foi utilizado de implementação prévia (de terceiros). No Apêndice B é apresentado um diagrama de sequência entre as etapas de extração, persistência, e classificação.

¹⁴ <https://github.com/>

¹⁵ <https://github.com/bytewfish>

A Figura 23 mostra a ferramenta desenvolvida para a geração do conjunto de treinamento. Basicamente uma janela que responde a eventos de *mouse*, onde é possível selecionar os píxeis, (pontos marcados em verde) para a geração dos descritores. Dois critérios foram considerados para incluir um pixel no conjunto de treinamento: não ter sido escolhido previamente e ser um pixel marcado na imagem de anotação¹⁶. Um arquivo é salvo em disco, contendo os descritores.

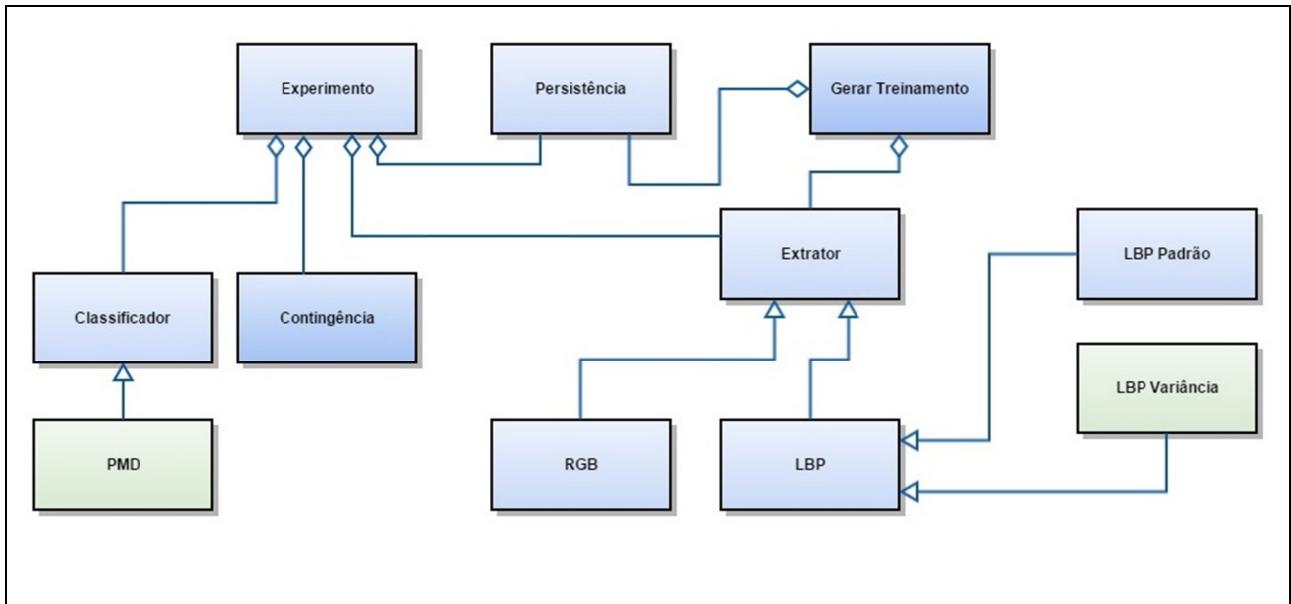


Figura 22. Diagrama de Classe simples das implementações efetuadas e utilizadas

Cada experimento recebe um conjunto de treinamento, um grupo de imagens e uma Combinação de Parâmetros (CP). Como resultado da etapa de classificação, são obtidas duas imagens, a resposta de classificação e a respectiva imagem de limiarização.

Na etapa de contingência, a imagem de limiarização é confrontada com a imagem de anotação, gerando como resultado uma imagem temática, com um código de cores para VP, VN, FP e FN. Na Figura 24 é apresentado um exemplo das referidas imagens, (Figura 24-a) classificação, (Figura 24-b) limiarização e (Figura 24-c) contingência.

¹⁶ O que garantiu que um pixel que não representasse o padrão de interesse fosse escolhido para treinamento.

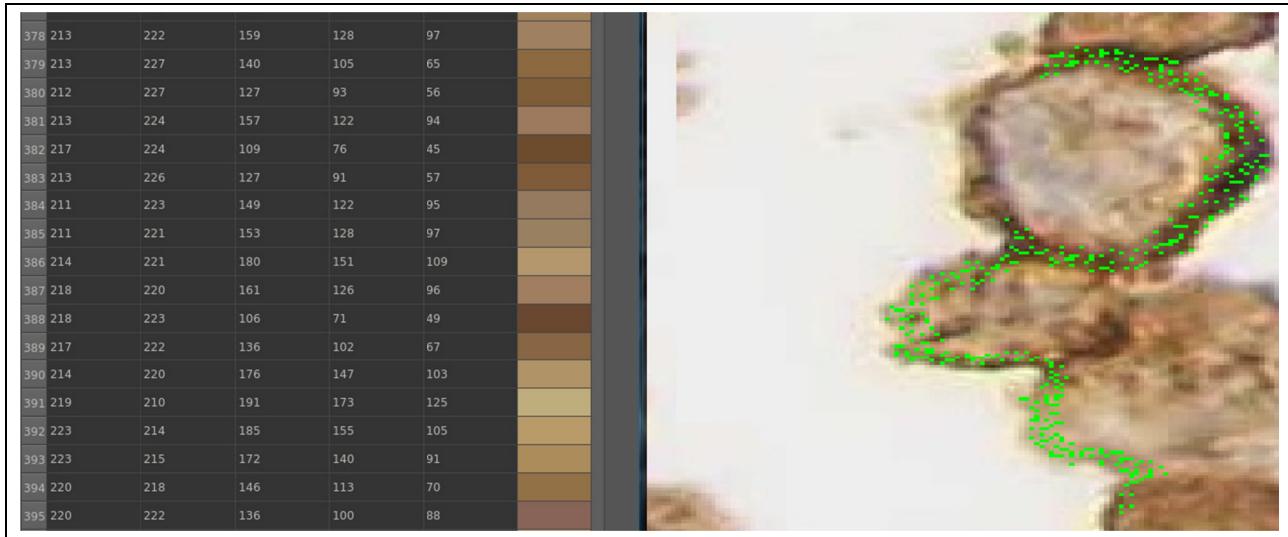


Figura 23. Ferramenta para seleção do conjunto de treinamento

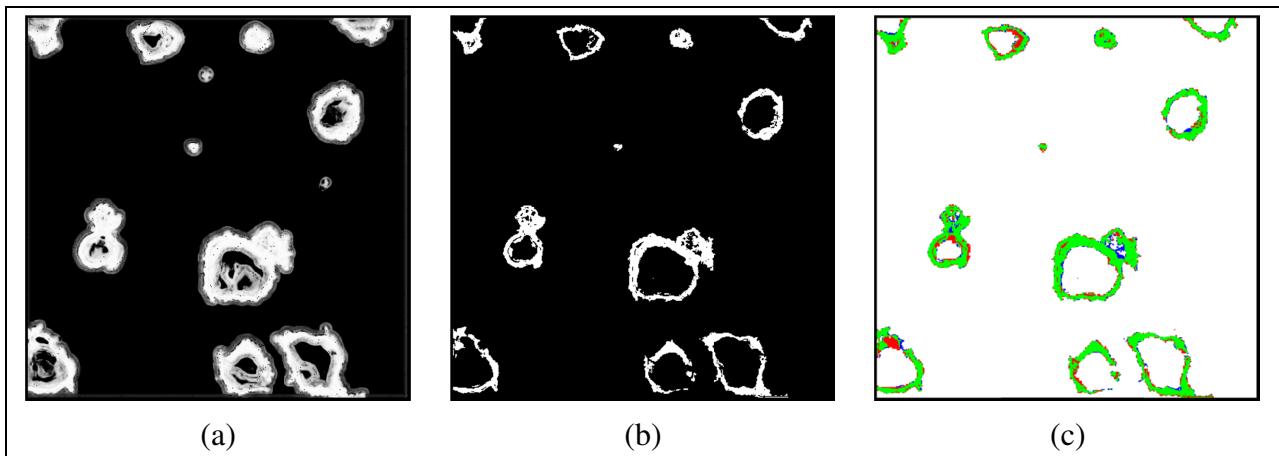


Figura 24. Resultados do processo de classificação e contingência

Como o método de classificação retorna um valor entre 0 e 1, representando a distância do descriptor em avaliação para a amostra de treinamento¹⁷, este valor é multiplicado por 255, a fim de gerar a imagem de resposta de classificação (Figura 24-a), que espera um valor entre 0 e 255.

Para gerar a imagem de limiarização, Figura 24-b, um valor de corte é definido, para aceitar o pixel como pertencente ao padrão de interesse (parâmetro avaliado nos experimentos). Por exemplo, um limiar de 240 considera, como pertencente ao padrão de interesse, qualquer píxel avaliado pelo classificador com uma distância superior a 0.9411.

¹⁷ Quanto mais próximo a 1, mais pertinente ao padrão de interesse.

A Figura 24-c, mostrou um exemplo do resultado da avaliação de contingência. Na Figura 25 é apresentado uma visão aproximada desta imagem, o mapa temático resultante tem a seguinte definição:

- **Verde:** píxeis verdadeiro positivo (VP);
- **Branco:** píxeis verdadeiro negativo (VN);
- **Vermelho:** Falso Positivo (FP); e
- **Azul:** Falso Negativo (FN).

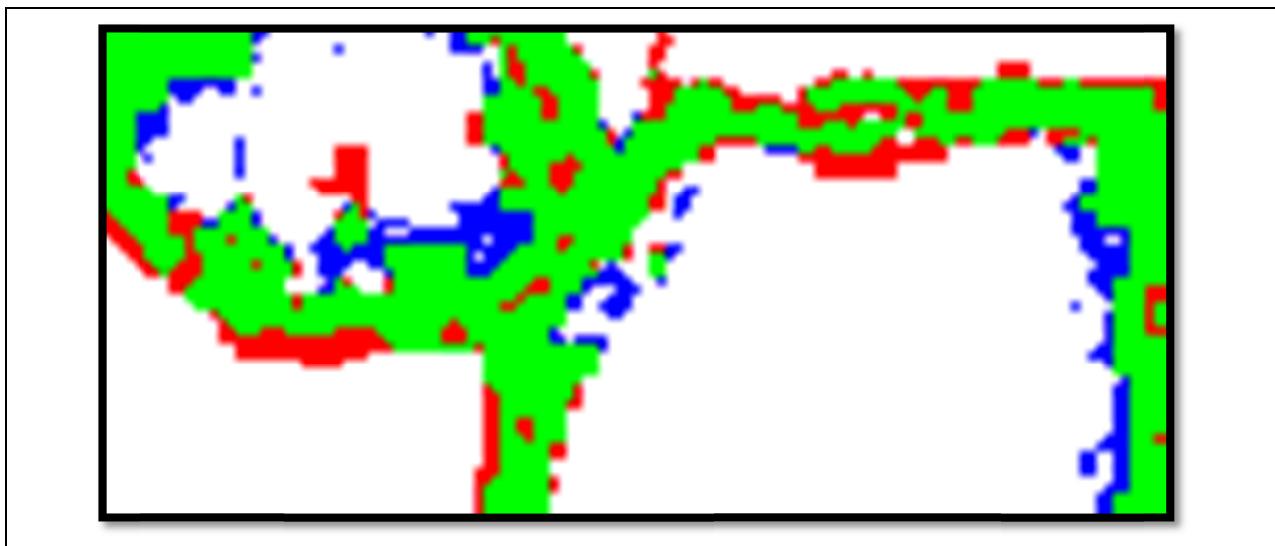


Figura 25. Vista aproximada do mapa temático

Ao final da execução do experimento, tem-se as imagens resultantes (mostradas na Figura 24) e um arquivo no formato CSV(*Comma Separated Values*), contendo uma linha para cada imagem avaliada. Cada linha deste arquivo é composta pelo nome da imagem e pelos seus respectivos valores de VP, FP, VN e FN, encontrados na etapa de contingência. Estes arquivos resultantes foram utilizados posteriormente na avaliação de classificação e testes de hipóteses.

4.4 DESCRIÇÃO DOS EXPERIMENTOS

Os objetivos específicos (4 e 5), apresentados na Subseção 1.2.2 tratam das avaliações que foram investigadas, a fim de comprovar a hipótese de pesquisa. Três grupos de experimentos foram elaborados para este fim: (i) dimensionalidade, que consiste em avaliar o grau de importância que cada dimensão do DC tem em relação à representatividade do padrão de interesse, inicialmente o

descriptor tem uma composição 23-dimensional (Seção 4.1); (ii) desempenho que consiste em executar uma série de testes, avaliando o resultado de classificação, empregando as métricas apresentadas na Seção 2.5; e (iii) variância que consiste em aceitar ou refutar H_P , a partir de uma análise de variância entre os testes realizados no segundo experimento (iii).

4.4.1 Experimento de Dimensionalidade

O objetivo deste experimento foi investigar a representatividade de cada medida definida para o descriptor inicial, com a finalidade de reduzir sua dimensão com a menor perda de informação possível. Este experimento definiu um segundo descriptor com 12 dimensões, os resultados obtidos foram descritos no Capítulo 5 .

4.4.2 Experimento de Desempenho

Usualmente a medida de acurácia é a mais comum para avaliação, entretanto ela não é suficiente, quando existe um desbalanceamento entre as classes. Nestes casos medidas mais recomendadas são Precisão (17) e *Recall* (18) (ARBEL; ROKACH, 2006).

As medidas empregadas para a avaliação de desempenho foram descritas no Capítulo 2 e no Capítulo 5 . Este experimento foi executado em duas etapas: um teste piloto (para estimação de parâmetros) e um teste principal, cujos resultados foram discutidos na Seção 5.2.

4.4.3 Experimento de Variância

Com as tabelas e gráficos gerados nos testes de desempenho foi possível fazer uma análise de variância, a fim de aceitar ou refutar a hipótese de pesquisa. O teste de hipótese pode assumir dois valores: H_0 , como hipótese nula ou verdadeira; ou H_a , como hipótese alternativa. Para aceitar a hipótese alternativa, o desempenho de classificação do método DPM, combinado com um DC híbrido (cor mais textura), deve ser maior que o desempenho de classificação que considera o descriptor simples (utilizando somente cor).

Inicialmente, a medida escolhida para realizar este experimento foi a Precisão, apresentada em (17). Esta decisão se embasou na adequação da medida em condições de desbalanceamento entre classes, pois descreve o grau de predições corretas do classificador, considerando também os falsos positivos. Entretanto o experimento de variância avaliou as medidas de Sensibilidade,

Especificidade, Precisão e Acurácia (medidas de acerto) e as Taxas de Falso Positivo, Falso Negativo e Falsa Descoberta.

O teste escolhido para investigar a hipótese de pesquisa foi o Mann-Whitney (MW) ou Wilcoxon-Mann-Whitney, por ser um teste não paramétrico, com a singularidade de avaliar variáveis quantitativas, afirmam Barbetta, Reis e Bornia (2008). A ideia principal é avaliar a posição central de duas populações, para isso, as hipóteses são colocadas como:

$$H_0: n_1 = n_2 \quad H_a: n_1 \neq n_2 \quad (34)$$

onde n_1 e n_2 são as medianas das populações 1 e 2 respectivamente, e H_a pode ser avaliada unilateralmente com $n_1 > n_2$ ou $n_1 < n_2$ de acordo com o que se deseja provar.

A hipótese nula, neste experimento, parte do princípio que as variâncias entre as populações são iguais, ou seja, o desempenho de classificação dos descritores avaliados é o mesmo. Já a hipótese alternativa investigada indica que o desempenho de classificação com o descritor baseado no método LBP é melhor do que o desempenho com o descritor RGB (que reflete a hipótese de pesquisa), formalmente $n_2 > n_1$. Com estas definições foi possível calcular a estatística U do teste, considerando um grau de confiança alfa de 0.05.

4.5 DESCRIÇÃO DO CONJUNTO DE DADOS

O projeto Human Protein Atlas¹⁸ (HPA) é financiado pela Fundação Knut e Alice Wallenberg e foi criado com a proposta de permitir a pesquisa sistemática no proteoma¹⁹ humano, através da análise de perfil de proteínas, em Microarranjos de Células ou Tecido (Tissue Microarrays - TMAs).

Diversas entidades hospedam o portal, tais como, o laboratório Rudbeck²⁰ (Universidade de Uppsala, Suécia), responsável por criar perfis da expressão de proteína humana, em larga escala

¹⁸ <http://www.proteinatlas.org/about/project>

¹⁹ Priberam – “Conjunto das proteínas produzidas por uma célula ou por um organismo.”

²⁰ <http://www.rudbeck.uu.se/>

empregando IHQ, e o Lab Surgpath²¹ (Mumbai, Índia) responsável por gerar as anotações das imagens de IHQ (PROTEIN ATLAS, 2014; NAVANI, 2011). No Anexo B são apresentadas ilustrações sobre TMAs usados no HPA.

Na versão sete do HPA estão incluídas mais de dez milhões de imagens coradas imuno-histoquimicamente de tecidos e células, baseadas em 13154 anticorpos, representando 10118 genes que compreendem aproximadamente 50% do genoma humano. O portal fornece uma base com imagens microscópicas de alta resolução que mostram a distribuição espacial de proteínas em 46 diferentes tipos de tecidos humanos normais e em 20 tipos com câncer, bem como 47 tipos diferentes de Linhas de Células Humanas (LCH). Os dados são liberados com validação de aplicação específica, incluindo imuno-histoquímica, análise de Western Blot e imunofluorescência confocal (UHLEN *et al.*, 2010; NAVANI, 2011).

Atualmente o atlas HPA, baseado em 185 amostras de tecidos e linhas de células, com mais de 21900 anticorpos analisados, relacionando proteínas em mais de 16600 genes humanos, cobrindo aproximadamente 82% do código proteico dos genes humanos²², está na 12^a versão, lançada em 05/12/2013.

As imagens empregadas neste trabalho foram adquiridas na categoria de LCH que fornecem amostras de indivíduos saudáveis e com leucemia/linfoma. Tais imagens são derivadas do DSMZ²³, ATCC²⁴ ou outros grupos de pesquisa. Nesta categoria são fornecidas imagens de sobreposição com anotações adicionais em duas modalidades distintas, a primeira baseada em objeto e a segunda baseada em área (PONTÉN; JIRSTRÖM e UHLEN, 2008).

A Figura 26 ilustra uma amostra de imagem original e as respectivas imagens de sobreposição. Os metadados anotados segmentam a reação de coloração alvo (amarronzada), a partir da imagem original Figura 26-a. Píxeis considerados da classe positiva (pertinente ao padrão de interesse) são marcados em preto, na Figura 26-b e píxeis da classe negativa são marcados com branco.

²¹ <http://www.labsurgpath.com/labsurgpath.html> (em construção)

²² <http://www.proteinatlas.org/about/releases>

²³ Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH - <http://www.dsmz.de/home.html>

²⁴ The essential of life science research - <http://www.atcc.org/>

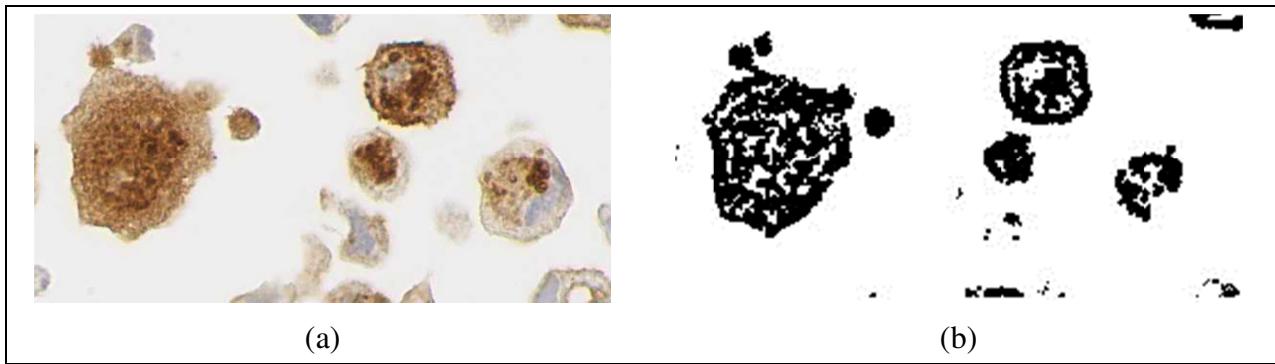


Figura 26. a) Imagem LCH-IHQ; e b) Imagem anotada respectiva

Fonte: Adaptado de Human Protein Atlas (2014).

Outras informações de anotações (metadados), fornecidas pelo HPA, na categoria LCH, além das imagens de sobreposição, são: (i) coloração do anticorpo; (ii) intensidade de coloração; (iii) número de células; (iv) porcentagem de células positivas; (v) gênero; e (vi) idade do paciente; para cada imagem (ver Anexo C). A categoria (LCH) foi escolhida por ser a única a fornecer anotações de metadados em forma de sobreposições de imagem. Mais informações sobre HPA e toda a gama de dados e metadados disponibilizados, podem ser encontradas nas citações anteriores e ainda em Uhlén *et al.* (2005), Lundberg *et al.* (2010) e Pontén *et al.* (2009), bem como no próprio site.

4.5.1 Conjunto de Dados

Todo o Conjunto de Dados (CD), selecionado para avaliação neste trabalho, foi extraído do HPA. As imagens utilizadas são relativas ao gene CD44, cuja proteína codificada por ele é classificada como um biomarcador para candidato a câncer. Estas imagens provém de 56 linhas de células distintas, de origens variadas (p. ex.²⁵ cérebro, fígado, mama, pele), e são relacionadas ao anticorpo CAB000316.

Para facilitar o processamento das etapas de experimento foi efetuada uma redução nestas imagens originais que foram cortadas, a partir do centro, para dimensões de 500x500 pixels. Após este corte foi realizada a separação das imagens resultantes em três grupos, respeitando os seguintes critérios:

²⁵ por exemplo – p. ex.

- **Grupo 01:** imagens com mais de 10% dos píxeis marcados como positivos;
- **Grupo 02:** imagens entre 10% e 5% dos píxeis marcados como positivos; e
- **Grupo 03:** imagens abaixo de 5% e acima de 1%.

O resultado dessa divisão gerou um grupo de sessenta imagens. As distribuições entre as classes são mostradas no Gráfico 1: onze imagens foram observadas no grupo 01; vinte e três imagens, no grupo 02 e vinte e seis imagens, no grupo 03.

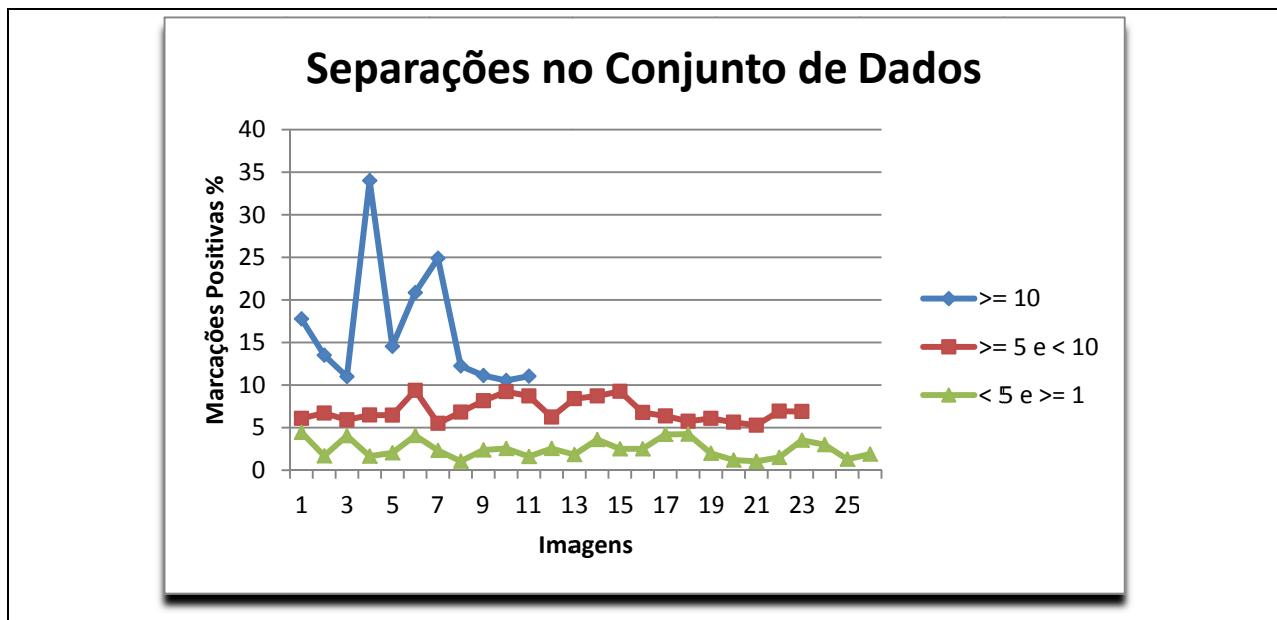


Gráfico 1. Distribuição da porcentagem de marcações positivas nas imagens

De acordo com a abordagem supervisionada adotada neste trabalho, a imagem com maior porcentagem de píxeis marcados como positivo foi utilizada para a geração da amostra de treinamento. Os experimentos que avaliaram o desempenho de classificação utilizaram o grupo 01 e o grupo 02 para avaliação (séries azul e vermelha no Gráfico 1).

4.6 CONSIDERAÇÕES

Tendo em vista que os objetivos específicos 1 e 2 tratam da definição do método de textura e das medidas estatísticas que foram empregadas no descritor da proposta, este capítulo apresentou tais definições demonstrando o DC 23-dimensional, baseado no método LBP. Além disso, apresenta o resultado da etapa de implementação, a origem e o conjunto de dados empregados, e a forma como foram conduzidos os experimentos.

5 ANÁLISE EXPERIMENTAL

O Capítulo 5 apresenta o desenvolvimento do projeto, mais especificamente com relação aos experimentos realizados entre análise de dimensionalidade, desempenho de classificação (taxas de acerto e erro) e os testes de hipóteses efetuados.

5.1 ANÁLISE DE DIMENSIONALIDADE

Em análises estatísticas, conjuntos de dados com altos espaços dimensionais podem ser simplificados, empregando um método definido como Análise de Componentes Principais (Principal Component Analysis - PCA), apresentado por Pearson (1901). A ideia principal consiste em prover um subespaço em que a maior parte da informação seja retida, uma vez que as medidas originais sejam projetadas neste subespaço.

Desta forma, os subespaços são calculados através de componentes. A linha (subespaço 1-dimensional), que carrega a maior representatividade dos dados, é definida como Primeira Componente Principal (PC1). A segunda linha mais representativa, quando combinada com a primeira, é aquela que retém a maior representatividade dos dados originais, em um subespaço 2-dimensional (ALFARO *et al.*, 2014).

O descritor definido inicialmente (\mathbf{d}_{23}) foi composto por 23 medidas (Seção 4.1). Um dos experimentos planejados consistiu em avaliar a representatividade dos 20 componentes compostos por medidas estatísticas, relativas ao filtro de textura. O método utilizado no experimento foi a Análise de Componentes Principais e teve como objetivo investigar a relevância de cada uma destas 20 medidas em relação ao padrão descrito.

O Gráfico 2 apresenta o resultado do experimento de PCA, o eixo y relaciona os Autovalores para cada componente resultante (eixo x). A linha em vermelho mostra a variabilidade acumulada do padrão descrito, ao longo das 20 componentes.

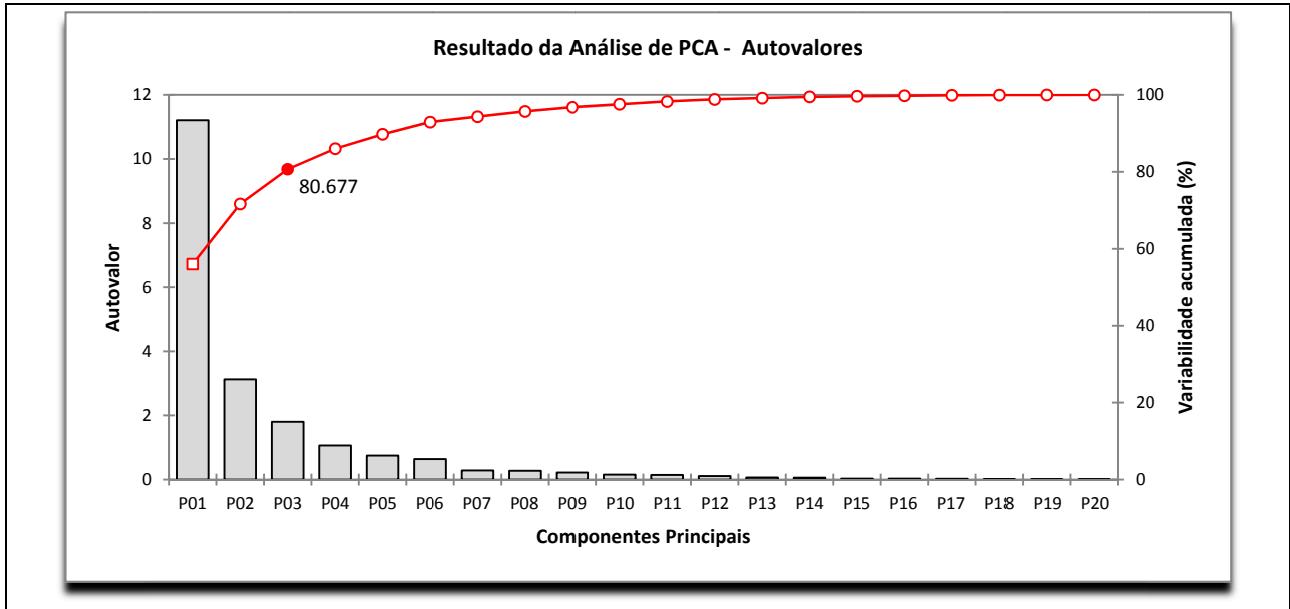


Gráfico 2. Resultado do experimento e PCA

Como pode ser observado no Gráfico 2, 80.677% (ponto preenchido com vermelho) da variabilidade (representatividade) do padrão descrito estavam contidas nas três primeiras componentes principais (P01, P02 e P03). A Tabela 4 mostra a contribuição de cada medida do descriptor (primeira coluna) relacionada às 10 primeiras componentes principais.

A partir da avaliação do resultado do experimento de PCA, foram identificadas as cinco maiores medidas com contribuição no descriptor: (i) três medidas, relativas à primeira componente principal; e (ii) duas medidas referentes à segunda componente principal. Os referidos valores foram destacados, em verde na Tabela 4, e são definidos a seguir:

- **J2_T2:** valor do desvio padrão referente a janela de amostragem com dimensão de 5x5;
- **J2_T5:** valor da energia referente a janela de amostragem com dimensão de 5x5;
- **J3_T2:** valor do desvio padrão referente a janela de amostragem com dimensão de 7x7;
- **J3_T5:** valor da energia referente a janela de amostragem com dimensão de 7x7; e
- **J4_T2:** valor do desvio padrão referente a janela de amostragem com dimensão de 9x9.

Tabela 4. Contribuição de cada variável para as 10 primeiras componentes (%)

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
J1-T1	5.049	7.520	0.052	1.773	1.851	16.570	0.103	3.747	0.005	20.693
J1-T2	6.888	2.340	0.206	4.028	0.006	0.000	20.286	5.943	0.291	1.310
J1-T3	3.710	0.137	0.301	20.599	2.030	41.852	18.683	3.165	1.752	2.611
J1-T4	6.606	0.451	0.201	7.430	0.100	3.045	28.913	0.000	4.665	0.148
J1-T5	0.166	16.778	21.430	0.000	0.052	0.677	4.007	5.284	3.652	2.226
J2-T1	6.671	5.643	0.610	0.527	0.011	4.857	0.504	2.128	0.344	7.816
J2-T2	7.961	1.077	0.234	0.033	2.127	0.000	3.285	0.954	2.086	7.218
J2-T3	3.540	2.441	0.813	26.626	24.530	0.357	0.397	1.799	7.061	2.824
J2-T4	7.018	0.704	0.208	3.820	9.096	2.016	0.643	0.185	2.362	10.010
J2-T5	0.025	25.265	4.979	5.202	2.004	4.753	0.378	0.315	0.949	0.294
J3-T1	7.347	2.364	2.002	0.030	0.466	3.251	2.970	7.110	0.695	0.190
J3-T2	8.216	0.110	0.386	0.600	0.995	0.177	0.130	1.209	4.331	18.102
J3-T3	3.801	6.203	5.385	7.085	13.636	0.013	0.075	21.210	10.457	0.943
J3-T4	6.415	2.289	2.381	1.469	3.189	0.840	1.531	5.318	25.309	1.937
J3-T5	0.015	18.032	5.023	17.144	5.187	14.684	1.444	1.875	1.021	0.001
J4-T1	7.583	0.232	2.658	0.084	0.173	3.698	8.203	5.945	0.031	1.184
J4-T2	7.891	0.587	0.483	0.027	0.256	0.831	7.085	2.671	1.612	20.828
J4-T3	4.688	4.428	3.716	0.244	24.260	0.585	0.481	0.018	22.174	0.147
J4-T4	6.388	3.345	2.202	0.038	2.909	1.726	0.003	23.836	5.879	0.979
J4-T5	0.023	0.053	46.732	3.243	7.122	0.070	0.881	7.287	5.324	0.539

O segundo descritor foi denominado de \mathbf{d}_{12} (RGB-LBP reduzido), baseado nas três medidas com maiores valores de contribuição, observadas na avaliação de PCA. Este novo descritor considerou as medidas de média, desvio padrão e energia nas janelas de 5x5 até 9x9, e foi utilizado no experimento de desempenho principal, e definido por:

$$\mathbf{d}_{12} = [r, g, b, t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9] \quad (35)$$

onde, r , g e b são as componentes de cor referente ao píxel descrito. As demais componentes são referentes as medidas efetuadas na imagem filtrada por textura.

5.2 DESEMPENHO DE CLASSIFICAÇÃO

Dois experimentos relacionados ao desempenho de classificação foram executados: (i) Classificação Piloto (C-PI); e (ii) Classificação Principal (C-PR). No experimento C-PI, os três descritores relacionados, \mathbf{d}_3 (descritor somente com os valores de RGB), \mathbf{d}_{23} (descritor RGB-LBP

completo) e \mathbf{d}_{12} (descriptor RGB-LBP reduzido), foram avaliados. O experimento C-PI teve a finalidade de definir o tamanho do conjunto de treinamento e a combinação de parâmetros, para a execução do experimento C-PR, onde apenas os descritores \mathbf{d}_3 e \mathbf{d}_{12} foram avaliados.

Para a análise do desempenho de classificação foram computadas as medidas apresentadas no Quadro 4, em que a segunda coluna mostra a sigla referente a cada medida. Essas siglas foram utilizadas nas tabelas, gráficos e no texto, no restante do documento.

Quadro 4. Medidas computadas para os experimentos de desempenho de classificação

Medida	Sigla	Descrição
Sensibilidade	SEM	Relaciona as predições positivas verdadeiras contra o valor observado. Equivalente à medida <i>Recall</i> .
Especificidade	ESP	Relaciona as predições negativas do classificador contra o valor observado.
Precisão	PRE	Relaciona as predições positivas pela soma dos valores de verdadeiro positivo e falso positivo.
Taxa de Predição Negativa	T-PN	Relaciona as predições negativas verdadeiras, pela soma dos valores de verdadeiro negativo e falso negativo.
Taxa de Falso Positivo	T-FP	Relaciona as falsas predições negativas contra os valores negativos anotados.
Taxa de Falsa Descoberta	T-FD	Relaciona os valores de falso positivo contra a soma dos valores de falso positivo e verdadeiro positivo.
Taxa de Falso Negativo	T-FN	Relaciona os valores de falso negativo contra os valores negativos observados.
Acurácia	ACC	Relaciona as predições corretas (positivas e negativas), contra anotações positivas e negativas.
F1 Escore	F1	Média harmônica entre PRE e SEM.
Coeficiente de Correlação de Matthews	MCC	Relaciona valores de verdadeiros, positivos e negativos, bem como falsos, positivos e negativos.

A Sensibilidade (SEN) pode ser entendida como a capacidade do modelo em prever amostras positivas, a partir do conjunto de teste. Do contrário, a Especificidade (ESP) pode ser entendida como a capacidade do classificador em predizer amostras negativas. Já a Acurácia (ACU) relaciona a taxa de acerto geral do modelo, entre predições positivas e negativas e o total de amostras classificadas (positivas e negativas anotadas).

Todas as medidas apresentadas no Quadro 4 assumem um valor entre 0 e 1. As medidas de SEM, ESP, PRE, T-PN, F1 e MCC estão relacionadas a predições corretas do classificador, neste caso, quanto mais próximo de 1 melhor o desempenho de classificação (sendo 1, a classificação

perfeita). As medidas de T-FP, T-FD e T-FN estão relacionadas a predições incorretas (taxas de erro), portanto, um bom desempenho de classificação deve apresentar valores próximo a 0.

5.2.1 Resultado do Experimento Piloto

Para o experimento preliminar foi definido um conjunto de treinamento com 5100 amostras e um conjunto de testes com 10 imagens (com dimensões de 508x508), referentes ao anticorpo CAB000316. O conjunto de testes do C-PI foi avaliado em diversas iterações, onde cada iteração tinha uma Combinação de Parâmetros (CP) específica.

No Grau Polinomial (GP) as combinações empregadas compreenderam as variações no grau de 3 a 5. Outra combinação foi relacionada aos parâmetros de Contraste (CN) e Limiar (LI), referentes ao algoritmo da DPM e valor de corte (CO) para aceitação do pixel como positivo (entre 0 e 1), respectivamente:

- Contraste -1.0 e Limiar 230;
- Contraste -1.0 e Limiar 240;
- Contraste -1.0 e Limiar 250;
- Contraste -0.5 e Limiar 230;
- Contraste -0.5 e Limiar 240; e
- Contraste -0.5 e Limiar 250.

Desta forma o conjunto de teste foi avaliado (segundo as medidas do Quadro 4), em um total de 54 iterações, executadas para os três descritores (\mathbf{d}_3 , \mathbf{d}_{23} , \mathbf{d}_{12}). As tabelas 5, 6 e 7(medidas relacionadas a acerto²⁶) resumem os resultados para o descritor \mathbf{d}_3 (RGB), nas combinações de parâmetros especificadas. Cada valor nas tabelas corresponde a média da medida para as 10 imagens do experimento piloto.

Os maiores valores encontrados para SEM e T-PN foram 0.9986 e 0.997, respectivamente, na combinação: GP 03 / CO -0.5 / LI 230. Nesta mesma combinação ocorreram os piores valores de

²⁶ Nas medidas referentes a acerto, quanto mais próximo a 1, o valor, melhor o desempenho do classificador.

ESP, PRE, ACC, F1 e MCC (valores em vermelho, na Tabela 5). Os maiores valores para F1 e MCC ocorreram com a CP: GP 05 / CO -1.0 / LI 240 (valores em verde, na Tabela 6). Os melhores valores para ESP, PRE e ACC foram 0.9719, 0.8594 e 0.9436, respectivamente, na CP: GP 05 / CO -1.0 e CO -0.5 / LI 250 (valores em verde, na Tabela 7).

Tabela 5. Resultados com o descriptor \mathbf{d}_3 (RGB), limiar 230

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.9934	0.8828	0.6243	0.9987	0.9049	0.7661	0.7380
0.5 / 03	0.9986	0.8491	0.5588	0.9997	0.8776	0.7153	0.6866
1.0 / 04	0.9764	0.9156	0.7018	0.9956	0.9295	0.8163	0.7885
0.5 / 04	0.9894	0.8980	0.6567	0.9980	0.9168	0.7891	0.7617
1.0 / 05	0.9619	0.9259	0.7295	0.9926	0.9358	0.8294	0.8005
0.5 / 05	0.9818	0.9105	0.6877	0.9965	0.9260	0.8085	0.7809

Tabela 6. Resultados com o descriptor \mathbf{d}_3 (RGB), limiar 240

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.9816	0.9058	0.6777	0.9966	0.9223	0.8013	0.7731
0.5 / 03	0.9949	0.8766	0.6112	0.9990	0.9000	0.7565	0.7284
1.0 / 04	0.9540	0.9294	0.7393	0.9914	0.9376	0.8326	0.8033
0.5 / 04	0.9799	0.9123	0.6930	0.9963	0.9273	0.8115	0.7839
1.0 / 05	0.9289	0.9391	0.7667	0.9862	0.9417	0.8395	0.8084
0.5 / 05	0.9671	0.9230	0.7216	0.9936	0.9342	0.8262	0.7977

Tabela 7. Resultados com o descriptor \mathbf{d}_3 (RGB), limiar 250

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.9816	0.9058	0.6777	0.9966	0.9223	0.8013	0.7731
0.5 / 03	0.9949	0.8766	0.6112	0.9990	0.9000	0.7565	0.7284
1.0 / 04	0.9540	0.9294	0.7393	0.9914	0.9376	0.8326	0.8033
0.5 / 04	0.9799	0.9123	0.6930	0.9963	0.9273	0.8115	0.7839
1.0 / 05	0.9289	0.9391	0.7667	0.9862	0.9417	0.8395	0.8084
0.5 / 05	0.9671	0.9230	0.7216	0.9936	0.9342	0.8262	0.7977

Da mesma forma a Tabela 8, a Tabela 9 e a Tabela 10 mostram a média dos resultados para o descriptor \mathbf{d}_{23} (RGB-LBP completo). A Tabela 11, a Tabela 12 e a Tabela 13 mostram os resultados para o descriptor \mathbf{d}_{12} (RGB-LBP reduzido), composto pelos valores de RGB adicionado aos valores de média, desvio padrão e energia.

Tabela 8. Resultados com o descriptor \mathbf{d}_{23} (RGB-LBP completo), limiar 230

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.9628	0.8623	0.5742	0.9918	0.8836	0.7193	0.6826
0.5 / 03	0.9890	0.8037	0.4853	0.9972	0.8385	0.6501	0.6164
1.0 / 04	0.9390	0.8960	0.6427	0.9872	0.9079	0.7629	0.7249
0.5 / 04	0.9747	0.8538	0.5609	0.9942	0.8783	0.7119	0.6772
1.0 / 05	0.9206	0.9109	0.6772	0.9837	0.9173	0.7799	0.7411
0.5 / 05	0.9614	0.8721	0.5934	0.9915	0.8915	0.7337	0.6976

Tabela 9. Resultados com o descriptor \mathbf{d}_{23} (RGB-LBP completo), limiar 240

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.9299	0.8980	0.6451	0.9854	0.9081	0.7614	0.7222
0.5 / 03	0.9692	0.8526	0.5573	0.9931	0.8765	0.7075	0.6716
1.0 / 04	0.8969	0.9250	0.7132	0.9795	0.9253	0.7936	0.7542
0.5 / 04	0.9469	0.8886	0.6266	0.9887	0.9029	0.7540	0.7166
1.0 / 05	0.8631	0.9393	0.7527	0.9735	0.9318	0.8026	0.7630
0.5 / 05	0.9297	0.9040	0.6609	0.9854	0.9131	0.7722	0.7338

Tabela 10. Resultados com o descriptor \mathbf{d}_{23} (RGB-LBP completo), limiar 250

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.7140	0.9677	0.8358	0.9482	0.9321	0.7669	0.7298
0.5 / 03	0.8783	0.9307	0.7267	0.9762	0.9272	0.7941	0.7538
1.0 / 04	0.5277	0.9844	0.8922	0.9183	0.9162	0.6593	0.6420
0.5 / 04	0.8065	0.9542	0.7965	0.9638	0.9354	0.7991	0.7598
1.0 / 05	0.3650	0.9923	0.9228	0.8937	0.8965	0.5196	0.5372
0.5 / 05	0.7243	0.9675	0.8380	0.9500	0.9336	0.7738	0.7371

Tabela 11. Resultados com o descriptor \mathbf{d}_{12} (RGB-LBP reduzido), limiar 230

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.7116	0.8904	0.5547	0.9395	0.8662	0.6229	0.5452
0.5 / 03	0.7173	0.8593	0.4905	0.9382	0.8410	0.5811	0.4965
1.0 / 04	0.9665	0.8892	0.6332	0.9918	0.9059	0.7649	0.7308
0.5 / 04	0.9737	0.8537	0.5605	0.9937	0.8776	0.7107	0.6760
1.0 / 05	0.0476	0.9972	0.7792	0.8487	0.8480	0.0867	0.1501
0.5 / 05	0.0491	0.9958	0.6753	0.8488	0.8471	0.0884	0.1396

Tabela 12. Resultados com o descritor \mathbf{d}_{12} (RGB-LBP reduzido), limiar 240

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.7018	0.9129	0.6137	0.9393	0.8835	0.6546	0.5829
0.5 / 03	0.7173	0.8842	0.5421	0.9400	0.8618	0.6166	0.5381
1.0 / 04	0.9440	0.9184	0.7049	0.9864	0.9260	0.8070	0.7720
0.5 / 04	0.9689	0.8823	0.6176	0.9924	0.9005	0.7541	0.7200
1.0 / 05	0.0450	0.9981	0.8534	0.8485	0.8484	0.0828	0.1550
0.5 / 05	0.0479	0.9969	0.7611	0.8488	0.8479	0.0871	0.1483

Tabela 13. Resultados com o descritor \mathbf{d}_{12} (RGB-LBP reduzido), limiar 250

CP	SEN	ESP	PRE	NPV	ACC	F1	MCC
1.0 / 03	0.6035	0.9637	0.7869	0.9254	0.9089	0.6824	0.6353
0.5 / 03	0.6895	0.9349	0.6846	0.9379	0.8991	0.6868	0.6234
1.0 / 04	0.6890	0.9742	0.8659	0.9408	0.9309	0.7659	0.7307
0.5 / 04	0.8779	0.9463	0.7823	0.9731	0.9380	0.8269	0.7890
1.0 / 05	0.0251	0.9996	0.9522	0.8463	0.8467	0.0480	0.1221
0.5 / 05	0.0388	0.9988	0.9007	0.8479	0.8481	0.0724	0.1494

A partir dos resultados do experimento preliminar foi possível observar que, utilizando o descritor baseado no método LBP, o processo de classificação se torna mais específico e preciso, de acordo com as medidas avaliadas. O que equivale dizer que: a adição de informação das componentes de textura proporcionou um melhor desempenho para classificar corretamente píxeis da classe negativa, bem como melhor desempenho considerando os valores de verdadeiro positivo quando relacionados aos valores de falso positivo.

A Tabela 14, e a Tabela 15, mostram os resultados da média das medidas de erro para os limiares de 230 e 250 (melhores e piores desempenhos na média), dos descritores propostos. Ao contrário das marcações de cores das tabelas de medidas de acerto, valores marcados com vermelho são mais altos, os valores em verde são mais baixos (o que se espera de uma taxa de erro é que ela seja mais próxima de zero).

Considerando os resultados de erro, foi possível observar melhor desempenho de classificação relativo à Taxa de Falso Positivo, o que tem impacto direto na medida de precisão e na medida de Especificidade (Taxa de Verdadeiro Negativo).

Tabela 14. Resultados de erro com o descritor d_{23} (RGB-LBP completo), limiares 230/250

CP	Limiar 230			Limiar 250		
	T-FP	FDR	T-FN	T-FP	FDR	T-FN
1.0 / 03	0.1377	0.4258	0.0068	0.0323	0.1642	0.0527
0.5 / 03	0.1963	0.5147	0.0020	0.0693	0.2733	0.0223
1.0 / 04	0.1040	0.3573	0.0112	0.0156	0.1078	0.0887
0.5 / 04	0.1462	0.4391	0.0047	0.0458	0.2035	0.0354
1.0 / 05	0.0891	0.3228	0.0146	0.0077	0.0772	0.1210
0.5 / 05	0.1279	0.4066	0.0071	0.0325	0.1620	0.0508

Tabela 15. Resultados de erro com o descritor d_{12} (RGB-LBP reduzido), limiares 230/250

CP	Limiar 230			Limiar 250		
	T-FP	FDR	T-FN	T-FP	FDR	T-FN
1.0 / 03	0.1096	0.4453	0.0563	0.0363	0.2131	0.0785
0.5 / 03	0.1407	0.5095	0.0555	0.0651	0.3154	0.0616
1.0 / 04	0.1108	0.3668	0.0069	0.0258	0.1341	0.0620
0.5 / 04	0.1463	0.4395	0.0051	0.0537	0.2177	0.0255
1.0 / 05	0.0028	0.2208	0.1866	0.0004	0.0478	0.1907
0.5 / 05	0.0042	0.3247	0.1863	0.0012	0.0993	0.1882

A Tabela 16 mostra os melhores desempenhos, nas medidas de acerto, para os três descritores avaliados no experimento preliminar, considerando as medidas de Sensibilidade (T-VP / Recall), Especificidade (T-VN), Precisão e Acurácia. A Tabela 17 mostra os melhores desempenhos para as medidas de erro, Taxa de Falso Positivo e Taxa de Falso Negativo.

Tabela 16. Maiores valores de ESP, SEN, PRE e ACC entre os três descritores

DESC	SEN	ESP	PRE	ACC
D ₃	0.9986	0.9719	0.8594	0.9436
D ₂₃	0.9890	0.9923	0.9228	0.9354
D ₁₂	0.9737	0.9996	0.9522	0.9380

Tabela 17. Menores valores de T-FP, T-FD e T-FN entre os três descritores

DESC	T-FP	T-FN
D ₃	0.0281	0.0003
D ₂₃	0.0077	0.0020
D ₁₂	0.0004	0.0051

As combinações de parâmetros que mais se destacaram entre os valores de melhor desempenho foram:

- GP 03 / CO -0.5 / LI 230;
- GP 04 / CO -0.5 / LI 230
- GP 04 / CO -0.5 / LI 250; e
- GP 05 / CO -1.0 / LI 250.

A Figura 27-a apresenta uma imagem das utilizadas no teste preliminar, a Figura 27-b, a imagem de anotação respectiva e a Figura 27-c, a imagem resultante da aplicação do filtro de textura LBP. A Figura 28 mostra as imagens de contingência resultantes, para os três descritores.

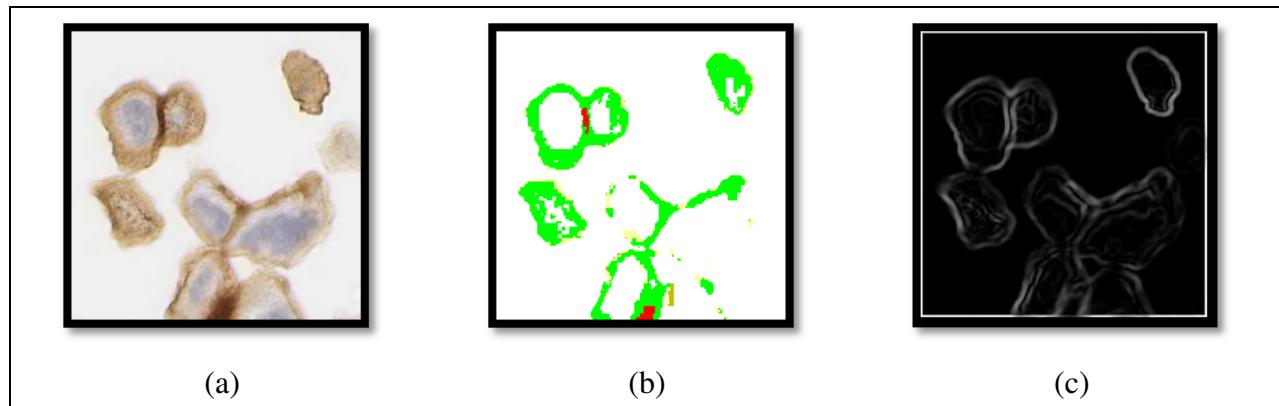


Figura 27.a) Imagem original; b) Imagem anotada e c) Resultado do filtro de textura LBP

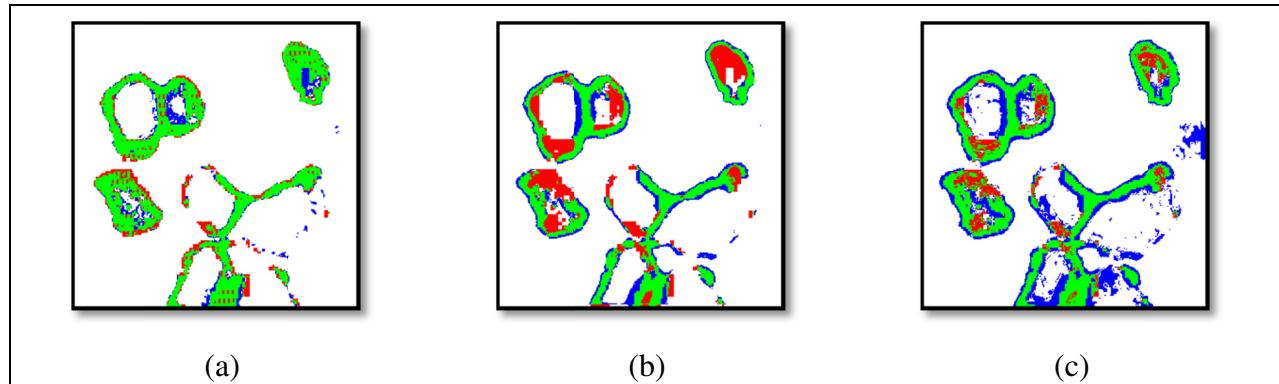


Figura 28. Mapa temático resultante para cada descritor: a) \mathbf{d}_3 , b) \mathbf{d}_{23} e c) \mathbf{d}_{12}

O Gráfico 3 apresenta a correlação entre Positivos anotados e Positivos preditos pelo classificador, na combinação de parâmetros GP 05 / CO -1.0 / LI 250 (para os descritores \mathbf{d}_3 e \mathbf{d}_{12}). O maior coeficiente de correlação (R^2) foi de 0.9933, utilizando o descritor \mathbf{d}_3 . No Gráfico 4 foi

apresentada a correlação entre Negativos anotados e Negativos preditos pelo classificador, na mesma combinação de parâmetros, maior coeficiente R² (correlação de Pearson) foi de 0.9998, utilizando o descritor d₂₃.

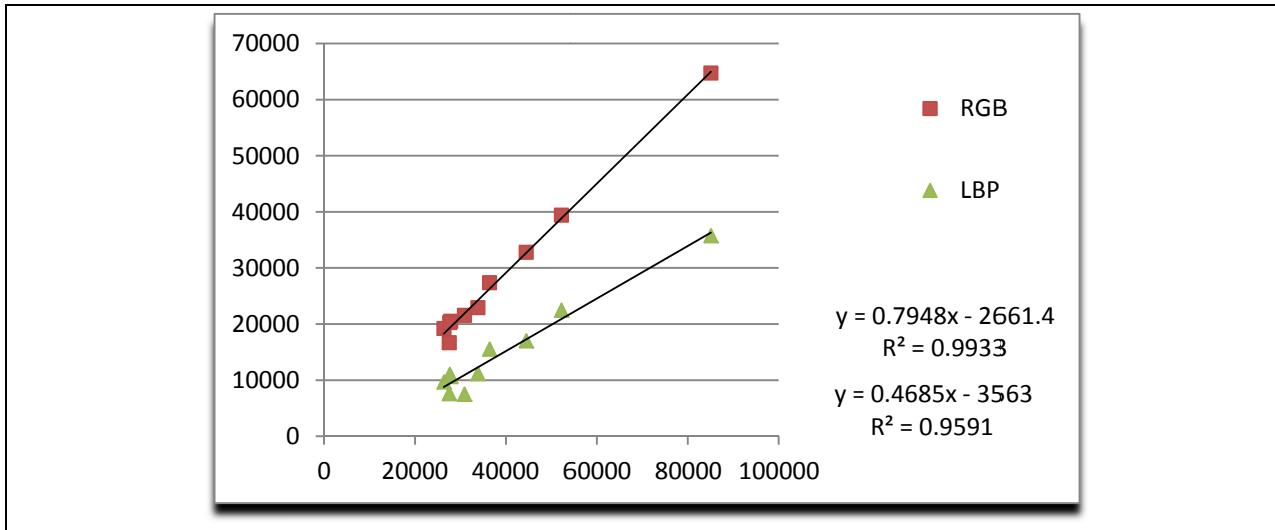


Gráfico 3. Correlação entre P e VP, descritores d₃ e d₂₃, parâmetros GP 05/ CO -1.0 / LI 250

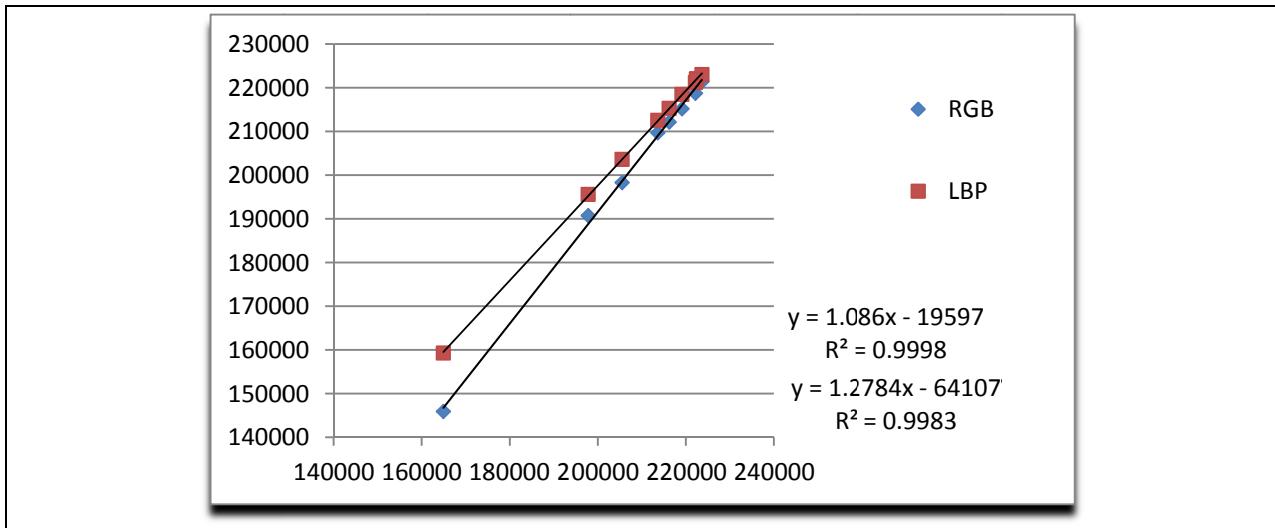


Gráfico 4. Correlação entre N e VN, descritores d₃ e d₂₃, parâmetros GP 05 / CO -1.0 / LI 250

5.2.2 Resultados do Experimento Principal

De acordo com os resultados encontrados, durante o experimento piloto, foi definido o experimento principal, com um conjunto de 23 imagens distintas das 10 imagens avaliadas preliminarmente. O conjunto de treinamento foi mantido o mesmo, 5100 amostras extraídas de uma

imagem não considerada nos testes (preliminar ou principal). A Tabela 18 apresenta a média, o desvio padrão, e a variância das medidas avaliadas.

Tabela 18. Resultados do Experimento principal (média para 23 imagens)

	SEN	ESP	PRE	T-PN	T-FP	T-FD	T-FN	ACC	F1	MCC
DESC-MDE										
Média	0.7014	0.9913	0.8660	0.9781	0.0087	0.1340	0.0223	0.9712	0.7673	0.7611
Desvio	0.1206	0.0040	0.0332	0.0079	0.0040	0.0332	0.0082	0.0073	0.0747	0.0653
Variância	0.0145	0.0000	0.0011	0.0001	0.0000	0.0011	0.0001	0.0001	0.0056	0.0043
DESC-RGB										
Média	0.8552	0.9830	0.8024	0.9893	0.0170	0.1976	0.0107	0.9744	0.8224	0.8123
Desvio	0.1004	0.0077	0.0450	0.0066	0.0077	0.0450	0.0068	0.0071	0.0461	0.0451
Variância	0.0101	0.0001	0.0020	0.0000	0.0001	0.0020	0.0000	0.0001	0.0021	0.0020

Nos gráficos 5 a 14 são apresentadas as distribuições dos valores observados, entre as medidas apresentadas na Tabela 18, para as 23 imagens. Em cada gráfico, a mesma medida é observada para ambos os descritores avaliados: em preto (a esquerda) considerando o descriptor baseado no método LBP (versão reduzida \mathbf{d}_{12}) e em vermelho (a direita) considerando o descriptor somente com os valores de RGB (\mathbf{d}_3).

Os gráficos 5 a 14 foram plotados com os seguintes parâmetros: a caixa maior comprehende 50% das observações (entre as 23 imagens), com 25% no limite inferior e com 75% no limite superior. A linha horizontal, dentro da caixa maior, indica a mediana e o quadro menor indica a média. As extremidades das linhas verticais que saem da caixa maior comprehendem 90% das observações, com 5% das observações no limite inferior e 95% no limite superior. Valores fora deste limite de 90% das amostras (marcações com “X”) são considerados valores atípicos²⁷.

Considerando os valores na Tabela 18 e a disposição dos resultados nos gráficos 5 a 14, foi possível observar maior média nas medidas de ESP e PRE; quando foi utilizado o descriptor baseado no método LBP, nas demais medidas de acerto, a maior média encontrada foi considerando somente os valores de cor RGB.

²⁷ Os cálculos e a plotagem propriamente dita foram efetuados com a ferramenta Qtiplot - <http://www.qtiplot.com/>

No Gráfico 6 é possível observar menor dispersão nos valores de Especificidade, utilizando o descritor baseado no método LBP, quando comparado com o descritor que utiliza somente informações RGB. Da mesma forma considerando o Gráfico 7, relativo à medida de precisão, corroborando o comportamento apresentado no experimento preliminar.

Nas medidas de erro, as menores médias encontradas para T-FP e T-FD foram observadas empregando o descritor baseado no método LBP. No Gráfico 13 é possível observar menor variabilidade nos valores da T-FP entre as 23 imagens. Estes resultados são condizentes com as avaliações executadas no experimento preliminar.

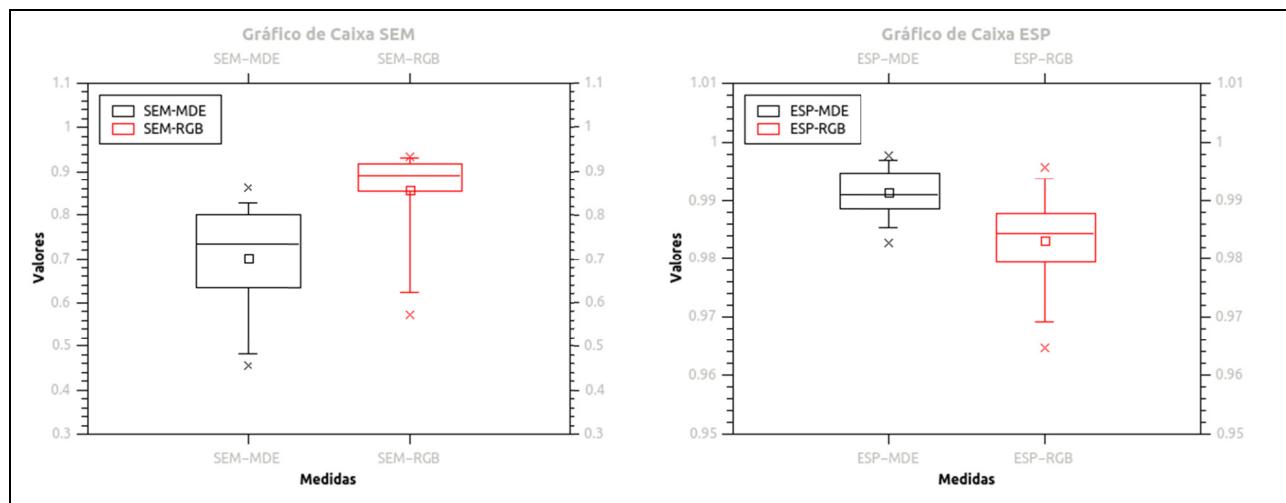


Gráfico 5. Gráfico de Caixa Sensibilidade

Gráfico 6. Gráfico de Caixa Especificidade

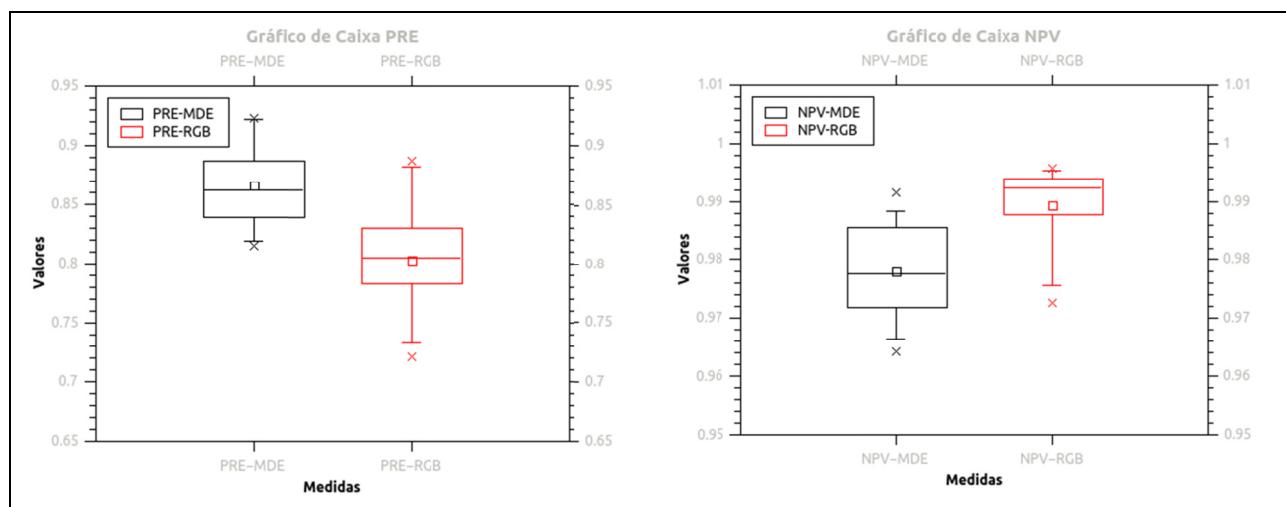


Gráfico 7. Gráfico de Caixa Precisão

Gráfico 8. Gráfico de Caixa Predição Negativa

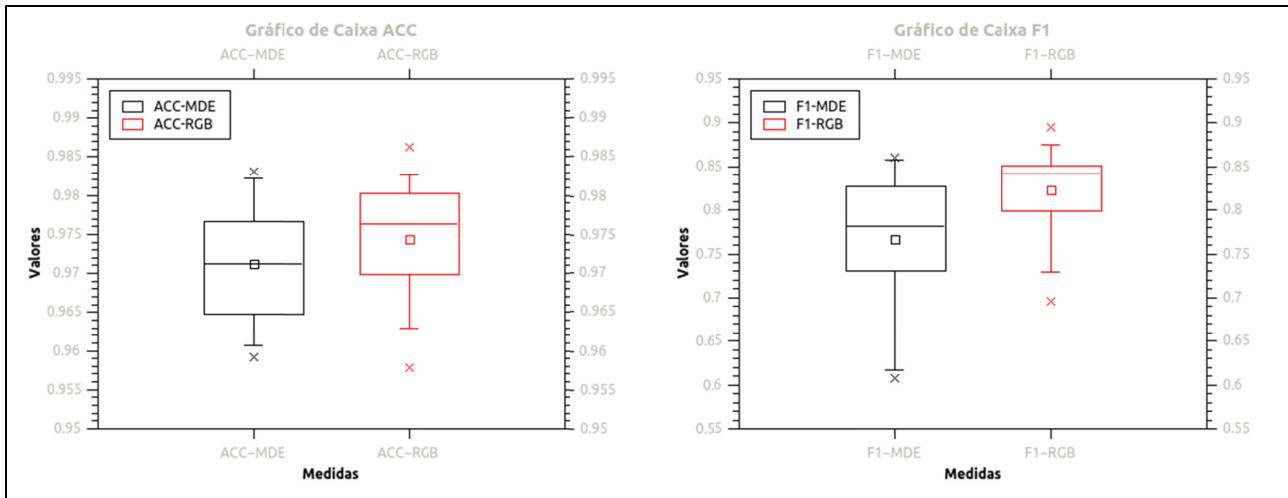


Gráfico 9. Gráfico de Caixa F1

Gráfico 10. Gráfico de Caixa Acurácia

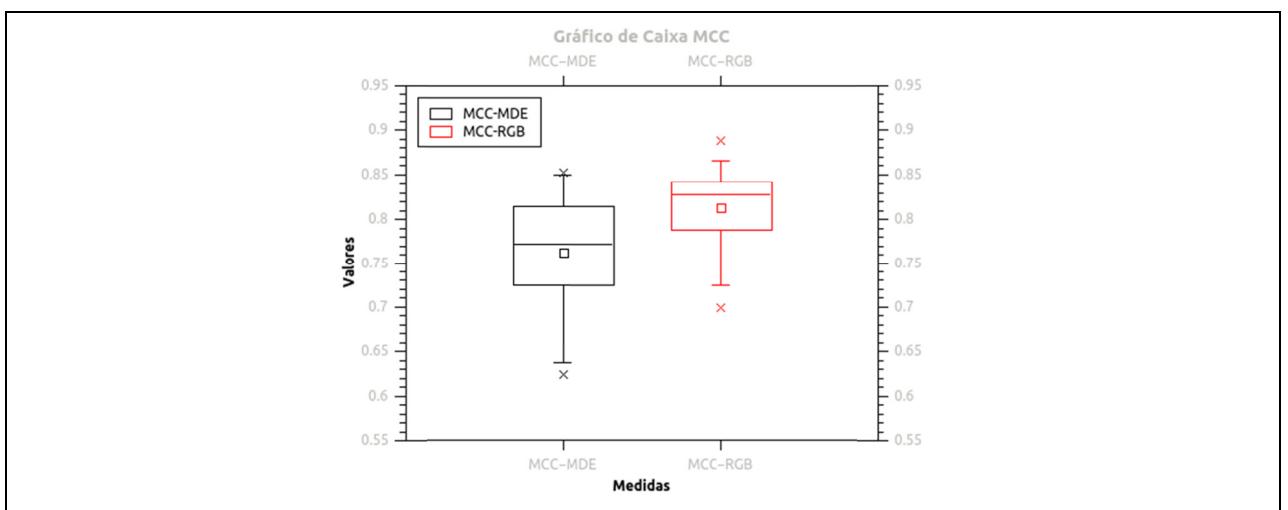


Gráfico 11. Gráfico de Caixa Correlação de Mattews

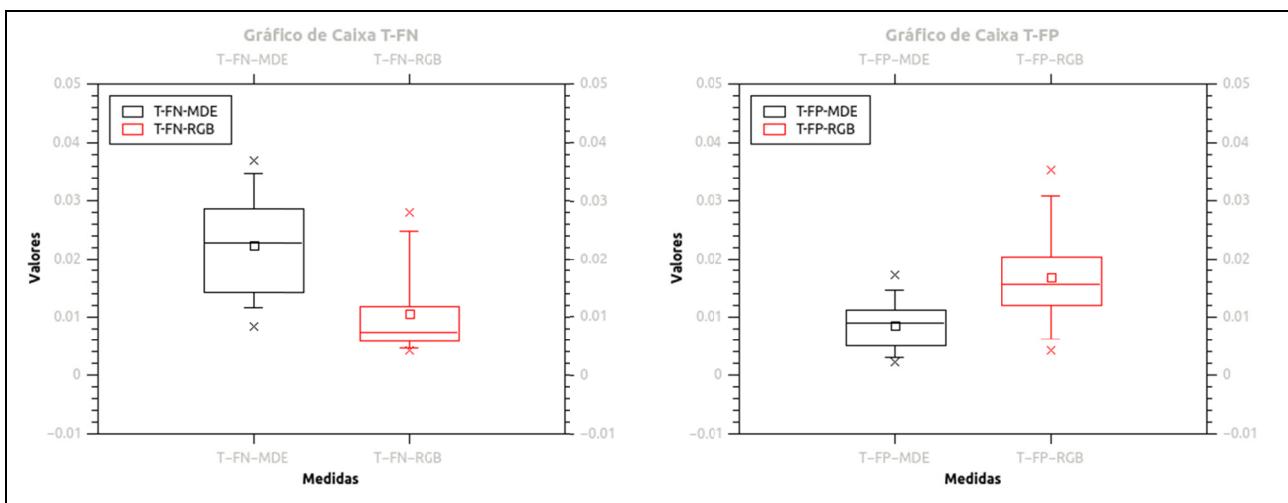


Gráfico 12. Gráfico de Caixa T-FN

Gráfico 13. Gráfico de Caixa T-FP

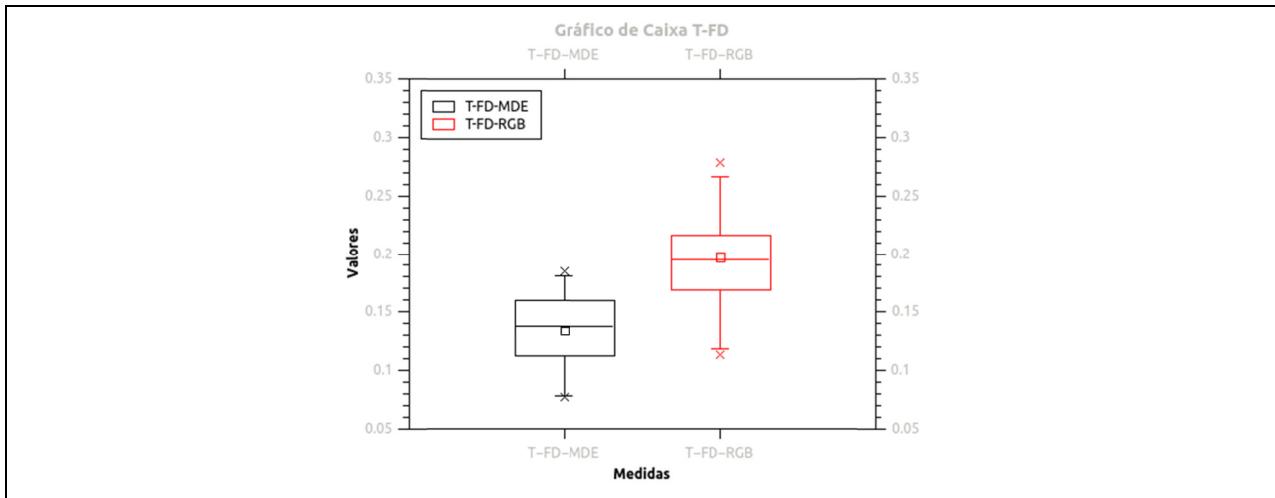


Gráfico 14. Gráfico de Caixa T-FD

No Gráfico 15 e no Gráfico 16 foi apresentada a correlação entre as previsões positivas do classificador contra as anotações positivas, e a correlação entre previsões negativas contra as anotações negativas, respectivamente. O maior coeficiente de determinação (R^2) na correlação positiva foi com o d_3 (RGB), 0.7996, contra 0.6916 do d_{12} (RGB-LBP reduzido). Na correlação entre negativos, o melhor coeficiente de determinação foi o de 0.973 para o d_{12} (RGB-LBP reduzido), contra 0.9373 do d_3 (RGB).

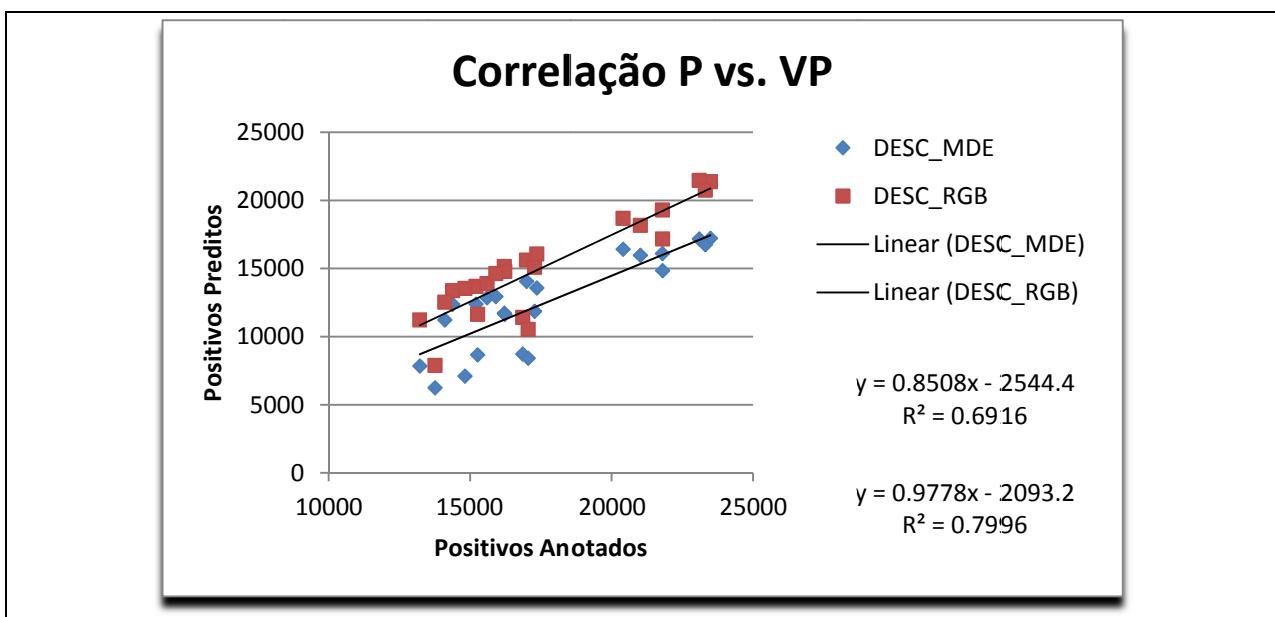


Gráfico 15. Correlação entre anotações e previsões positivas

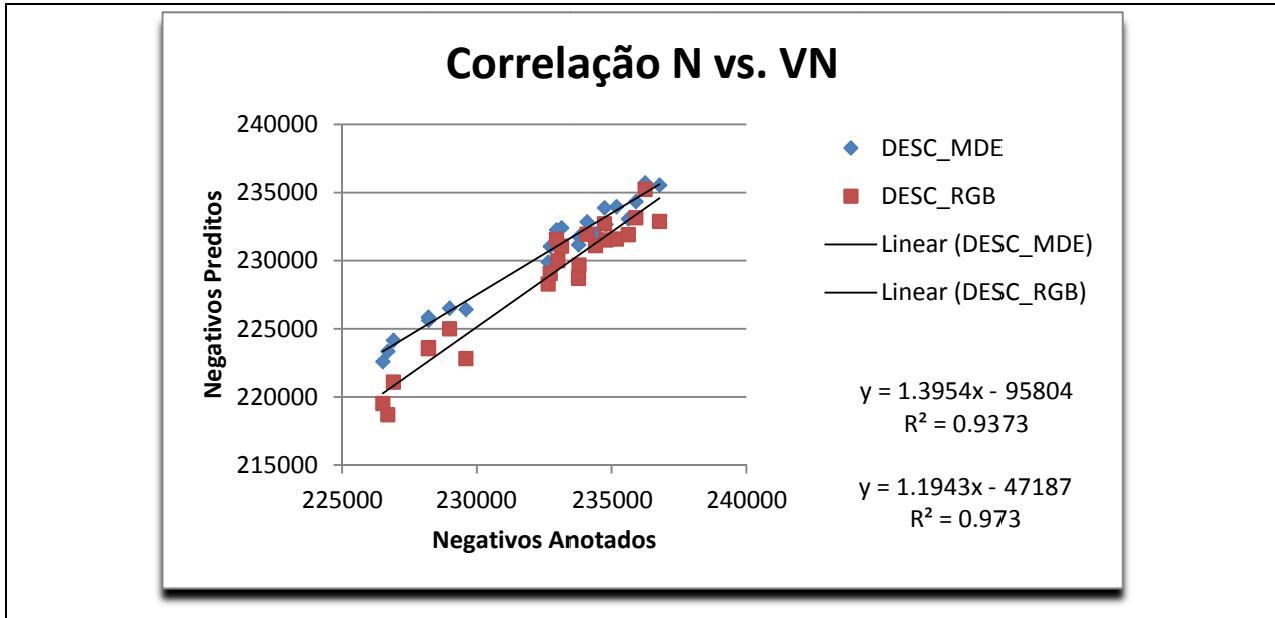


Gráfico 16. Correlação entre anotações e previsões negativas

A Figura 29 apresenta um exemplo de imagem utilizada no experimento principal:

- Na Figura 29-a é apresentada a imagem original; e
- na Figura 29-b, a respectiva imagem de anotação.

Na Figura 30 são apresentadas as respectivas imagens resultantes do processo de classificação:

- Figura 30-a utilizando somente RGB (**d₃** RGB-LBP reduzido); e
- Figura 30-b utilizando o descritor baseado no método LBP (**d₁₂** RGB-LBP reduzido).

Na Figura 31 o respectivo resultado do processo de contingência:

- Figura 31-a para o descritor RGB (**d₃** RGB-LBP reduzido); e
- Figura 31-b para o descritor baseado no método LBP (**d₁₂** RGB-LBP reduzido).

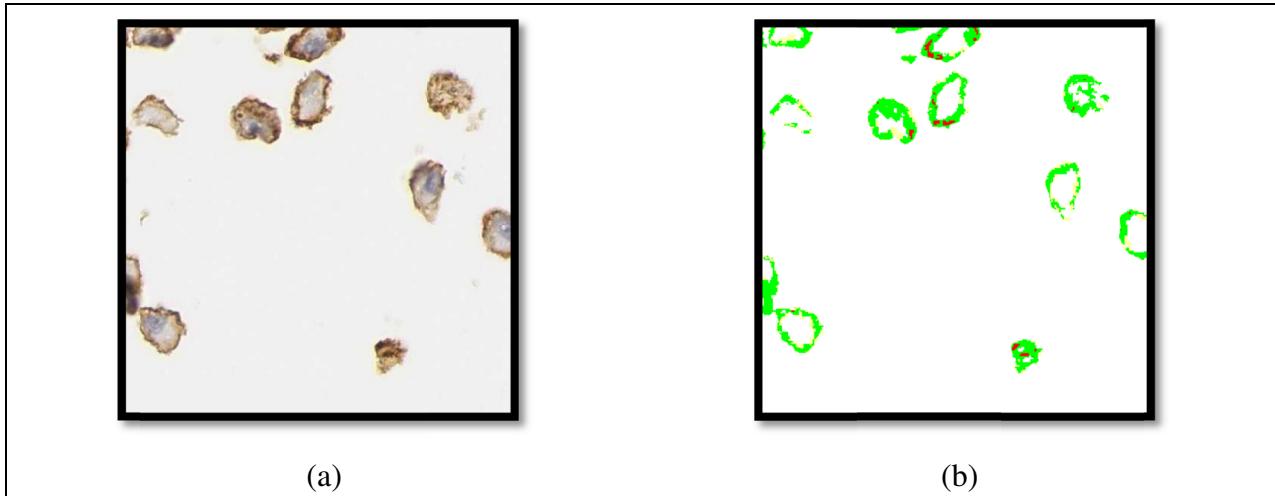


Figura 29. Imagem original utilizada no experimento e sua respectiva imagem de anotação

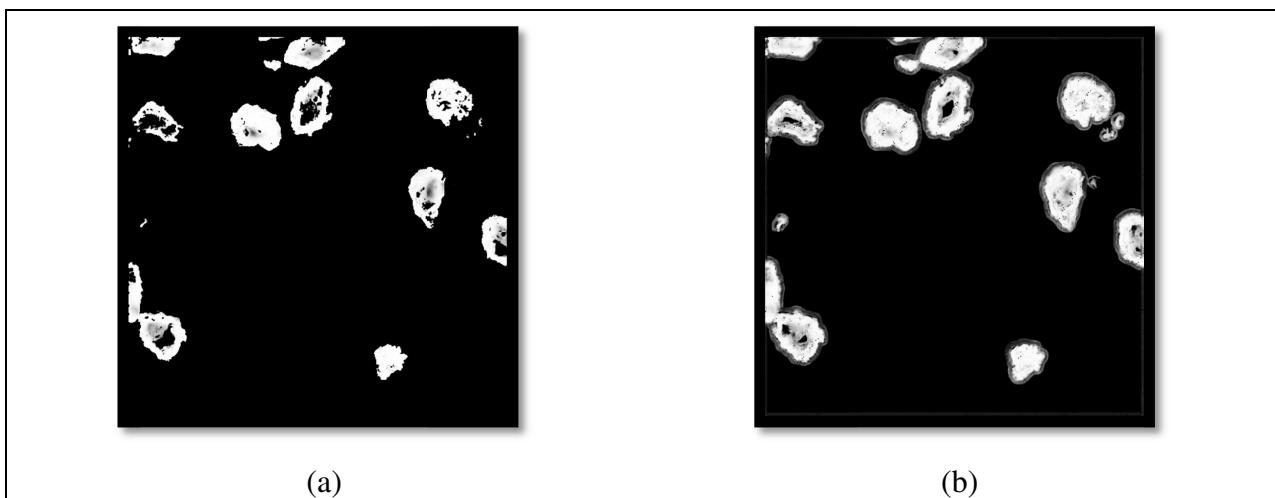


Figura 30. Exemplo de imagem resultante do processo de classificação, para o descriptor RGB e para o descriptor LBP reduzido, respectivamente.

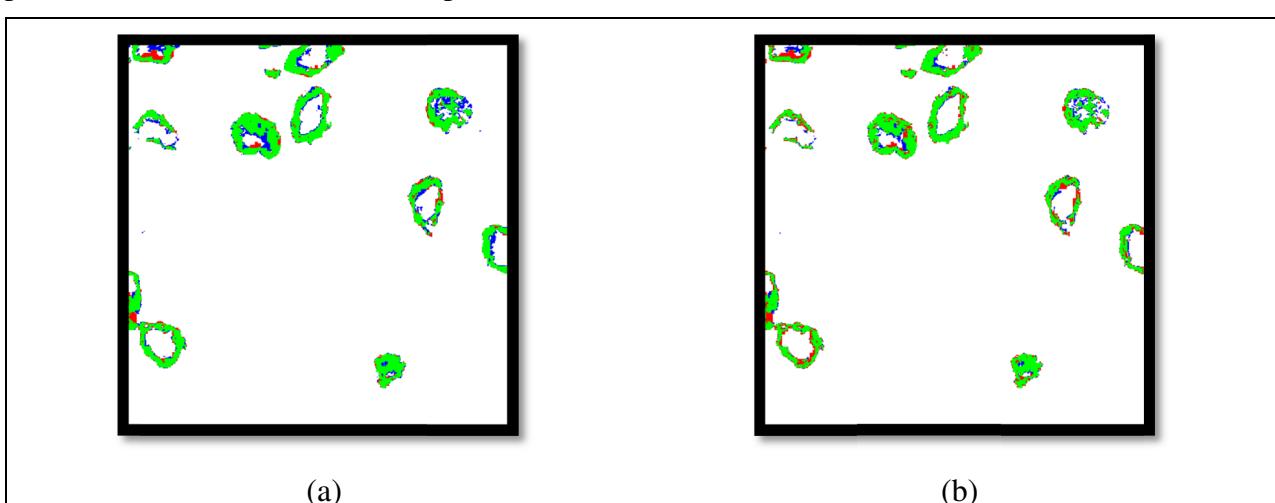


Figura 31. Exemplo de imagem resultante da avaliação de contingência, para o descriptor RGB e para o descriptor LBP Reduzido, respectivamente

5.3 TESTES DE HIPÓTESES

Usualmente, testes de hipóteses partem do princípio que a população submetida a amostragem (estimada aleatoriamente) segue uma distribuição conhecida (usualmente a distribuição normal), mesmo quando seus parâmetros (média e variância) são estimados na amostra. Tais testes são conhecidos como métodos *paramétricos* ou dependentes da distribuição, e tratam essencialmente variáveis contínuas.

Por outro lado quando não existe a necessidade de formular hipóteses sobre a forma de distribuição de uma população ou amostra, ou quando esta distribuição não segue uma forma conhecida (p. ex. normal), existe uma categoria de métodos conhecidos como não paramétricos ou não dependentes de distribuição. Na categoria não paramétrica é possível tratar, tanto variáveis quantitativas (contínuas), quanto variáveis categóricas (discretas como sim/não). Como exemplos podem ser citados: o teste dos sinais, testes de Wilcoxon e o teste de Kruskal-Wallis. (JEFFREY; CLARK, 2002; MONTGOMERY; RUNGER, 1999).

Os testes de hipóteses foram efetuados sobre o resultado do experimento principal de desempenho, considerando dois grupos de medidas: assertivas e de erro. No grupo das assertivas foram avaliadas as medidas de Sensibilidade, Especificidade, Precisão e Acurácia. No grupo das medidas de erro foram avaliadas a Taxa de Falso Positivo, a Taxa de Falsa Descoberta (T-FD) e a Taxa de Falso Negativo.

5.3.1 Teste de Medidas Assertivas

A hipótese nula H_0 para este teste considerou que a diferença entre as duas amostras é igual a zero, o que equivale a dizer que não há diferença em utilizar o descritor RGB ou o descritor baseado no método LBP. A hipótese alternativa H_a foi avaliada unilateralmente a esquerda, considerando que a diferença da amostra 1 (utilizando RGB), pela amostra 2 (utilizando LBP) é menor do que zero.

Aceitar a hipótese alternativa indica que foi observada uma diferença significativa, para mais, dos valores da amostra 2 de acordo com o grau de significância de 0.05. Portanto, um melhor desempenho de classificação, segunda a medida sendo avaliada, da amostra que utiliza o descritor baseado no método LBP. A Tabela 19 mostrou informações estatísticas das variáveis (medidas de acerto) e o respectivo descritor (primeira coluna) do experimento principal.

Tabela 19. Estatísticas descritivas do teste de hipótese para as medidas de acerto

Variável	Observação	com /falt	sem / falt	Mínimo	Máximo	Média	Desvio
SEN RGB	23	0	23	0.573	0.934	0.855	0.103
SEN MDE	23	0	23	0.454	0.861	0.701	0.123
ESP RGB	23	0	23	0.965	0.996	0.983	0.008
ESP MDE	23	0	23	0.983	0.998	0.991	0.004
PRE RGB	23	0	23	0.722	0.886	0.802	0.046
PRE MDE	23	0	23	0.814	0.923	0.866	0.034
ACC RGB	23	0	23	0.958	0.986	0.974	0.007
ACC MDE	23	0	23	0.959	0.983	0.971	0.007

A Tabela 20 mostrou os valores do teste de hipótese, a medida U, os valores de Esperança, Variância, p-valor e nível de significância alfa, para cada uma das quatro medidas consideradas, SEM, ESP, PRE e ACC.

Tabela 20. Resultados do teste de hipótese para as medidas de acerto

VTH	SEM	ESP	PRE	ACC
U	461.000	87.000	72.000	328.500
Esperança	264.500	264.500	264.500	264.500
Variância	2071.917	2071.917	2071.917	2071.789
p-valor	1.000	< 0.0001	< 0.0001	0.922
alfa	0.05	0.05	0.05	0.05

Sensibilidade

H_0 : A diferença da posição entre as amostras é igual a 0.

H_a : A diferença da posição entre as amostras é inferior a 0.

Como o p-valor calculado é maior que o nível de significância alfa=0.05, não se rejeita a hipótese nula H_0 . O risco de rejeitar a hipótese nula H_0 quando ela é verdadeira é de 100.00%.

Especificidade

H_0 : A diferença da posição entre as amostras é igual a 0.

H_a : A diferença da posição entre as amostras é inferior a 0.

Como o p-valor calculado é menor que o nível de significância alfa=0.05, deve-se rejeitar a hipótese nula H_0 em favor da hipótese alternativa H_a . O risco de rejeitar a hipótese nula H_0 quando ela é verdadeira é menor do que 0.01%.

Precisão

H_0 : A diferença da posição entre as amostras é igual a 0.

H_a : A diferença da posição entre as amostras é inferior a 0.

Como o p-valor calculado é menor que o nível de significância alfa=0.05, deve-se rejeitar a hipótese nula H_0 em favor da hipótese alternativa H_a . O risco de rejeitar a hipótese nula H_0 quando ela é verdadeira é menor do que 0.01%.

Acurácia

H_0 : A diferença da posição entre as amostras é igual a 0.

H_a : A diferença da posição entre as amostras é inferior a 0.

Como o p-valor calculado é maior que o nível de significância alfa=0.05, não se rejeita a hipótese nula H_0 . O risco de rejeitar a hipótese nula H_0 quando ela é verdadeira é de 92.18%.

5.3.2 Teste para Medidas de Erro

A hipótese nula H_0 para este teste considerou que a diferença entre as duas amostras é igual a zero, isso significa que não há diferença em utilizar o descritor RGB ou o descritor baseado no método LBP. A hipótese alternativa H_a foi avaliada unilateralmente a esquerda, considerando que a diferença da amostra 1 (utilizando RGB) pela amostra 2 (utilizando LBP) é maior do que zero.

Aceitar a hipótese alternativa indica que foi observado uma diferença significativa, para mais, dos valores da amostra 1 de acordo com o grau de significância de 0.05. Portanto, maior taxa de erro, segunda a media sendo avaliada, da amostra que utiliza o descritor baseado no método RGB. A Tabela 21 mostrou informações estatísticas das variáveis (medidas de erro) e o respectivo descritor (primeira coluna) do experimento principal.

Tabela 21. Estatísticas descritivas do teste de hipótese para as medidas de erro

Variável	Observação	com /falt	sem / falt	Mínimo	Máximo	Média	Desvio
T-FP RGB	23	0	23	0.004	0.035	0.017	0.008
T-FP MDE	23	0	23	0.002	0.017	0.009	0.004
T-FD RGB	23	0	23	0.114	0.278	0.198	0.046
T-FD MDE	23	0	23	0.077	0.186	0.134	0.034
T-FN RGB	23	0	23	0.004	0.028	0.011	0.007
T-FN MDE	23	0	23	0.008	0.037	0.022	0.008

A Tabela 22 mostrou os valores do teste de hipótese, a medida U, os valores de Esperança, Variância, p-valor e nível de significância alfa, para cada uma das três medidas consideradas, T-FP, T-FD e T-FN.

Tabela 22. Resultados do teste de hipótese para as medidas de erro

VTH	T-FP	T-FD	T-FN
U	442.000	457.000	66.000
Esperança	264.500	264.500	264.500
Variância	2071.917	2071.917	2071.917
p-valor	< 0.0001	< 0.0001	1.000
alfa	0.05	0.05	0.05

Taxa de Falso Positivo

H_0 : A diferença da posição entre as amostras é igual a 0.

H_a : A diferença da posição entre as amostras é superior a 0.

Como o p-valor calculado é menor que o nível de significância alfa=0.05, deve-se rejeitar a hipótese nula H_0 em favor da hipótese alternativa H_a . O risco de rejeitar a hipótese nula H_0 quando ela é verdadeira é menor do que 0.01%.

T-FD

H_0 : A diferença da posição entre as amostras é igual a 0.

H_a : A diferença da posição entre as amostras é superior a 0.

Como o p-valor calculado é menor que o nível de significância alfa=0.05, deve-se rejeitar a hipótese nula H_0 em favor da hipótese alternativa H_a . O risco de rejeitar a hipótese nula H_0 quando ela é verdadeira é menor do que 0.01%.

Taxa de Falso Negativo

H_0 : A diferença da posição entre as amostras é igual a 0.

H_a : A diferença da posição entre as amostras é superior à 0.

Como o p-valor calculado é maior que o nível de significância alfa=0.05, não se rejeita a hipótese nula H_0 . O risco de rejeitar a hipótese nula H_0 quando ela é verdadeira é de 100.00%.

5.4 CONSIDERAÇÕES

A proposta inicial previa a utilização do método LBP tradicional, entretanto o método LBP baseado em Variância se mostrou melhor no contexto das imagens utilizadas. Foi realizado um experimento piloto para avaliação de comportamento do classificador com a variação e definição de parâmetros, para o experimento principal.

No experimento principal foi possível observar, em média, melhor desempenho de classificação, considerando todas as medidas assertivas, para o descritor \mathbf{d}_3 (RGB). Nas medidas de erro o descritor \mathbf{d}_{12} (baseado no método LBP considerando a Média, o Desvio Padrão e a Energia) apresentou melhor desempenho (menores valores) em média, levando em conta as três medidas de erro.

Durante a qualificação foi definido que o teste de hipóteses seria realizado somente sobre a medida de Precisão, entretanto estes testes foram estendidos para as medidas assertivas de Especificidade, Sensibilidade e Acurácia, bem como para as três medidas de erros avaliadas.

Com o teste de hipóteses realizado foi possível comprovar a melhora de desempenho de classificação, segundo a medida de precisão, quando utilizado o descritor \mathbf{d}_{12} . Para as outras medidas assertivas, não se observou diferença significativa.

Para as medidas de erro, comprovou-se melhora de desempenho significativa, com o descriptor **d₁₂**, considerando as medidas de T-FN e T-FD. Para a medida de T-FP não se observou uma diferença significativa.

6 CONCLUSÃO

O problema de pesquisa abordado neste trabalho foi a análise de imagens microscópicas de imuno-histoquímica, cujo escopo focou na descrição e classificação de coloração em imagens desta categoria. As justificativas apresentadas apontaram a viabilidade da pesquisa, do ponto de vista computacional, ao empregar uma abordagem supervisionada de aprendizagem e classificação baseadas na distância polinomial de Mahalanobis, como apontados em Sobieranski *et al.* (2011) e Carvalho *et al.* (2014) que empregam a mesma abordagem. O presente trabalho relacionou diversos temas relevantes de pesquisa, tais como: o reconhecimento de padrões, processamento de imagens, descrição de características e análise de textura.

Outro trabalho que ajudou a demonstrar a viabilidade da pesquisa foi encontrado em Aguiar (2012), onde igual abordagem supervisionada, empregando cor e textura como descritor de características, foi utilizada. A diferença do presente trabalho consiste no operador de textura escolhido (LBP), bem como na combinação de janelas de amostragens em quatro dimensões distintas (Seção 4.1). Outro ponto de diferenciação entre os trabalhos está na área de aplicação, classificação de coloração em imagens de IHQ e classificação de vegetação em imagens de sensoriamento remoto (no trabalho citado).

Do ponto de vista da área de aplicação, a viabilidade de utilização do método de classificação definido, pode ser encontrada, não só em Sobieranski *et al.* (2009), como também na utilização do método por pesquisadores da área médica como referenciado em Di Giunta (2008) (coautora em Sobieranski *et al.* (2009)), e ainda em Martins *et al.* (2011).

No que se refere à revisão bibliográfica, uma das atividades desenvolvidas neste trabalho, na busca de resposta à pergunta de pesquisa, foi efetuada uma análise de trabalhos relacionados. Após a conclusão desta análise, as contribuições mais expressivas foram a adequada contextualização e delimitação de escopo do trabalho, do ponto de vista da área de aplicação, uma vez que a classificação de coloração específica é uma das etapas, no processo de análise de imagens de IHQ.

Dos vinte trabalhos avaliados, no Capítulo 3 foi possível observar um destaque para métodos de classificação supervisionada, mais especificamente empregando uma abordagem baseada em redes neurais, como Jitaree *et al.* (2013), Mouelhi, A.; *et al.* (2013) e Wemmert *et al.* (2013), ou baseado em SVM como Topaloglu (2010) e Hui; Zhiguo; Longmei (2011). As

abordagens mais semelhantes ao presente trabalho foram encontradas em Kostopoulos *et al.* (2007), Amaral *et al.* (2013) e Samsi *et al.* (2010), por apresentarem a mesma abordagem, na formação do descritor de características.

O objetivo do trabalho foi demonstrar a melhora no desempenho de classificação, considerando taxas de acertos e erros, do método DPM, de acordo com o escopo descrito na Subseção 1.1.2 . Para tanto, foi definido um descritor de características que considera informações de textura além das informações de cores, baseado no operador de textura Padrão Binário Local. Consequentemente, cinco objetivos específicos foram traçados e alcançados com ressalvas.

O primeiro objetivo específico tratou da definição do operador de textura, o método LBP, que foi escolhido por apresentar um baixo custo computacional (considerando o número de operações), com apenas uma operação de convolução por imagem. Este processo resultou em uma imagem de resposta de textura, menor custo consequentemente, quando comparado com o método de Matriz de coocorrência que gera quatro imagens de resposta de textura, ou quando comparado ao método de Laws, com 25 operações de convolução por imagem, o que resulta em 25 imagens de resposta.

Aliada a questão do custo de operações do filtro de textura LBP, outros fatores contribuíram para a sua escolha, principalmente a boa aceitação do método na literatura, identificada pelas inúmeras variações propostas em diversas áreas de aplicação, bem como o algoritmo simples do operador original apresentado em (9). Entretanto, após a etapa de implementação, percebeu-se que o método original não se mostrou adequado, no contexto das imagens utilizadas (Figura 21-b), neste ponto ocorreu a troca do operador original por uma extensão do mesmo, baseada em variância (Figura 21-c).

O segundo objetivo específico tratou da definição das medidas estatísticas que compuseram o descritor proposto inicialmente. Tais medidas foram baseadas nas mesmas medidas apresentadas no descritor proposto por Haralick; Shanmugam; Distein, (1973). Um subgrupo destas medidas foi utilizado por Kostopoulos *et al.* (2007), este mesmo subgrupo (média, desvio padrão, assimetria, curtose e energia) foi definido para compor o descritor utilizado no presente trabalho.

O terceiro objetivo específico tratou da definição propriamente dita do descritor, ou seja, de que forma as medidas tomadas na imagem de resposta de textura foram arranjadas. Como resultado

foi apresentado o descritor 23-dimensional, baseado no método LBP, composto pelos valores de RGB do píxel mais as 5 medidas estatísticas tomadas em sua vizinhança, em quatro dimensões distintas (3x3, 5x5, 7x7 e 9x9).

Uma análise de componentes principais foi realizada a fim de investigar a representatividade de cada uma das vinte componentes relacionadas à textura no descritor principal. Com o resultado desta análise foi possível perceber que mais de 80% da variabilidade estava concentrada nas três primeiras componentes, com destaque de contribuição para as medidas de Média, Desvio Padrão e Energia, consequentemente dois descritores reduzidos foram definidos.

O primeiro descritor reduzido foi composto pelos valores de RGB mais o valor da Média e da Energia, tomadas somente na vizinhança de 3x3 do píxel descrito, resultando em um descritor de características 5-dimensional. O segundo descritor reduzido definido foi composto pelos valores de RGB do píxel mais as medidas de Média, Desvio Padrão e Energia, tomadas nas vizinhanças de 5x5, 7x7 e 9x9. Resultando em um descritor 12-dimensional. Ambos os descritores reduzidos foram avaliados durante os experimentos de classificação, o primeiro no experimento piloto e o segundo no experimento principal.

Os dois últimos objetivos (quarto e quinto) trataram das avaliações de desempenho de classificação, em que foram avaliadas sete medidas relacionadas a taxas de acerto (positivos e negativos), Sensibilidade, Especificidade, Precisão, Taxa de Predição Negativa, Acurácia, F1 Escore e Coeficiente de correlação de Matteews. Para as avaliações de erros (positivos e negativos) foram avaliadas três medidas, Taxa de Falso Positivo, Taxa de Falso Negativo e Taxa de Falsa Descoberta (T-FD).

No teste piloto de classificação foi utilizado um grupo de 10 imagens, para avaliar o comportamento dos três descritores (RGB, RGB-LBP completo, e o primeiro descritor reduzido), em combinações distintas de parâmetros. O maior valor de Sensibilidade (Taxa de Verdadeiro Positivo ou Recall) foi encontrado com o descritor RGB simples (0.9986), a maior medida de Especificidade (Taxa de Verdadeiro Negativo) foi encontrada com o descritor LBP Completo (0.9995). A maior diferença entre descritores ocorreu na medida de Precisão, 0.9522 do descritor LBP reduzido (Média e Energia), contra 0.8594 do descritor RGB.

Nas taxas de erro avaliadas no teste piloto, somente na Taxa de Falso Negativo o descritor RGB apresentou o menor valor (0.0003). A maior diferença encontrada ocorreu na Taxa de Falsa Descoberta, 0.1406 com o descritor RGB, contra 0.0478, com o descritor LBP reduzido (\mathbf{d}_{12})

Durante a execução do teste piloto, pode-se observar uma melhora no desempenho de classificação, utilizando o descritor LBP, entretanto esta melhora não foi incondicional. Foi observada uma melhora nas medidas de Especificidade (Taxa de Verdadeiro Positivo) e Precisão. Já nas medidas de erros, o maior desempenho (menor valor, o que se espera de uma taxa de erro) encontrado, utilizando o descritor RGB, foi observado apenas na medida Taxa Falso Negativo.

O experimento principal, de desempenho de classificação, teve como objetivo gerar os dados para a execução do teste de hipóteses planejado. Foi possível observar maior média, considerando todas as medidas assertivas, para o descritor RGB. Nas medidas de erro, o descritor \mathbf{d}_{12} (RGB-LBP reduzido) apresentou melhor desempenho (menores valores) em média, considerando as três medidas de erro, um comportamento alinhando com o teste piloto.

Durante a qualificação foi definido que o teste de hipóteses seria realizado somente sobre a medida de Precisão, entretanto estes testes foram estendidos para as medidas assertivas de Especificidade, Sensibilidade e Acurácia, bem como para as três medidas de erros avaliadas. Com o teste de hipóteses realizado foi possível comprovar a melhora de desempenho de classificação, segundo a medida de precisão, quando utilizando o descritor \mathbf{d}_{12} (RGB-LBP reduzido). Para as outras medidas assertivas não se observou diferença significativa.

Para as medidas de erro, comprovou-se melhora de desempenho significativa com o descritor \mathbf{d}_{12} (RGB-LBP reduzido), considerando as medidas de T-FN e T-FD. Para a medida de T-FP não se observou diferença significativa. Todos os testes de hipóteses foram efetuados utilizando o método não paramétrico, Mann-Whitney, considerando um nível de significância alfa de 0.05. A escolha deste método se justifica na a característica de não balanceamento entre as classes. Nas 23 imagens do teste principal, a distribuição entre as classes foi, em média, 15.6888% para a classe positiva e 84.3112% para a classe negativa.

A hipótese de pesquisa sustentava que ao se adicionadas informações de textura, ao descritor de características, o desempenho de classificação (acertos e erros) aumentaria. Diante das argumentações apresentadas, até este ponto, foi possível verificar que:

- não houve aumento de desempenho incondicional, como foi apontado nos testes de classificação, o descritor RGB apresentou melhor valores em algumas das 10 medidas avaliadas;
- a hipótese de pesquisa é válida especificamente para o contexto de imagens apresentadas, como o descritor definido no documento; e
- a hipótese de pesquisa é válida quando empregado o método de classificação baseado na projeção polinomial da Distância de Mahalanobis (de acordo com a parametrização utilizada).

Diante dos resultados observados foi possível concluir ainda, que o emprego das informações de textura, no descritor, tornou o processo de classificação mais específico e preciso. Em contrapartida houve um aumento nos tempo de classificação, entretanto este impacto, considerando número de instruções e tempo de processamento não foram avaliados durante o trabalho.

6.1 CONTRIBUIÇÕES DA DISSERTAÇÃO

A principal contribuição alcançada com este trabalho de pesquisa foi aprimorar o método classificação de coloração em imagens de IHQ, considerando as medidas de Precisão, Taxa de Falso Negativo e Taxa de Falso Descobrimento, apresentado por Sobieranski *et al.* (2009).

Outra contribuição alcançada foi a definição de um descritor de características baseado no operador de textura LBP, que pode ser empregado em outros contextos e utilizado com outros classificadores, como os baseados em Redes Neurais e em SVM.

Com relação às publicações referentes à dissertação:

- DORNBUSCH L., M.; COMUNELLO, E. Classificação, Segmentação e Mensuração de Imagens Médicas de Imuno-histoquímica: Abordagem Supervisionada Empregando Atributos de Cor e Textura. In: XII Seminário de Iniciação Científica, 2013, Itajaí. Livro de Resumos 2013. Itajaí: Univali, 2013. p. 59-59.

DORNBUSCH L., M.; SOBIERANSKI, A.C.; COMUNELLO, M. Classificação de Coloração Imuno-histoquímica em Imagens Combinando Cor e Textura Como Descritor de Características. In: COMPUTER ON THE BEACH - COTB, 2015, *Anais* ..., 2015. p. 467-468. ISSN: 2358-0852

6.2 TRABALHOS FUTUROS

A localização espacial (núcleo, membrana, citoplasma), bem como a quantificação de intensidade dos padrões de coloração são essências para o diagnóstico, na análise de imagens de IHQ. A título de trabalho futuro, a pretensão é definir um método para a segmentação de estruturas celulares específicas, tais como a segmentação de núcleos celulares apresentada por Liu; Wang; Fang (2006) e Mouelhi *et al.* (2013) e, posteriormente, quantificar tais estruturas, com a mesma abordagem utilizada no presente trabalho.

Em aplicações clínicas, normalmente as análises de IHQ trabalham com o processamento de imagens extensas, nas quais lâminas virtuais são criadas a partir de digitalização, com equipamentos específicos. Estas lâminas virtuais são imagens que podem atingir tamanhos de 40Gb até 100 Gb, com dimensões de até 65000x45000 píxeis, conforme apresentado por Signolle *et al.* (2010), onde o processo de segmentação/classificação de um único slide virtual consumiu até 50 horas de processamento. Em trabalho futuro poderá ser conduzida a otimização do código, com alguma abordagem de processamento paralelo, e realizado experimento semelhante ao apresentado neste trabalho, utilizando imagens extensas.

Ainda, a título de trabalho futuro, pretende-se a extensão dos experimentos apresentados, em três frentes distintas: primeiro, aumentando o número de imagens e de conjuntos de treinamento; segundo, aumentando o número de medidas de acerto e erro e realizando teste de hipóteses para todas as medidas empregadas e combinações de parâmetros; terceiro, aumentando a avaliação de desempenho, considerando também consumo de memória e número de instruções, com a ferramenta para análise dinâmica de processamento Valgrind²⁸.

²⁸ <http://valgrind.org/>

REFERÊNCIAS BIBLIOGRÁFICAS

- AGUIAR, D. B. **Classificação de Superfícies Utilizando Polinomial Mahalanobis e Métricas de Textura.** Dissertação (Mestrado). Programa de Pós-Graduação em Ciência da Computação do Centro de Informática e Estatística da Universidade Federal de Santa Catarina, UFSC, Universidade Federal de Santa Catarina, Florianópolis, 2012.
- AMARAL, T.; MCKENNA, S. J.; ROBERTSON, K.; THOMPSON, A. Classification and immunohistochemical scoring of breast tissue microarray spots. **IEEE Transactions on Biomedical Engineering.** v. 60, n. 10, p. 2806–2814, Oct 2013. ISSN 0018-9294.
- ARBEL, R.; ROKACH, L. Classifier evaluation under limited resources. **Pattern Recognition Letters.** V. 27 p.1619-1631, 2006.
- BARR, M.; MCCLELLAN, S.; WINOKUR, T.; VAUGHN, G. An automated tissue pre-classification approach for telepathology: implementation and performance analysis. **IEEE: Transactions on Information Technology in Biomedicine.** vol.8, no.2, pp.97-102, 2004.
- BIOGENEX (U.S.A.). **Primary Antibodies.** 2014. Disponível em: <<http://www.biogenex.com/>>. Acesso em: 22 abr. 2014.
- BISHOP, C. B. **Pattern Recognition and Machine Learning.** New York: Springer+Business Media, 2006.
- BORGES, M. R.; VALLE, H. A.; FIORELLI, R. K. A.; OLIVEIRA, C. A. B. Telepatologia - Nova ferramenta diagnóstica para lesões do aparelho respiratório e importante elo tecnológico entre as áreas da Patologia, Clínica Cirúrgica e Pneumologia. **Revista Pulmão** Rio de Janeiro, vol. 14, no.3, 2005.
- BOVIK, A. **The essential guide to image processing.** Oxford: Elsevier, 2009.
- BRADLEY, A. The use of the area under the ROC curve in the evaluation of machine learning algorithms. **Pattern Recognition.** v.30, n. 7, 1997, p. 1145-1159.
- BRADSKI, G. **The OpenCV Library.** Dr. Dobb's Journal of Software Tools, citeulike:2236121, 2000.
- BRADSKI, G.; KAEHLER, A. **Learning opencv:** Computer Vision with the OpenCV Library. O'Reilly Media, 2008.
- CARVALHO, L.E.; NETO, S. L. M. WANGENHEIM, A. V.; SOBIERANSKI, A. C.; COSER, L.; COMUNELLO, E. Hybrid color segmentation method using a customized nonlinear similarity function. **International Journal of Image and Graphics.** v. 14, e. 01n02, 06/2014.
- CATALDO, S. D.; FICARRA, E.; MACII, E. Computer-aided techniques for chromogenic immunohistochemistry: Status and directions. **Computers in Biology and Medicine.** v.42, i. 10, p. 1210-1225, oct. 2012.

CHANG, H.; DEFILIPPIS, R. A.; TLSTY, T. D.; PARVIN, B. Scoring histological sections through immunohistochemistry. In: IEEE INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING: FROM NANO TO MACRO - ISBI, 5th, 2008. **Proceedings ... IEEE**, 2008. p. 344–347.

DI GIUNTA, G. **Células dendríticas, expressão da forma induzida da óxido nítrico sintase e padrão de citocinas nas lesões de pitiríase liquenóide**. Tese (Doutorado). Faculdade de Medicina da Universidade de São Paulo, FMUSP. Departamento de Patologia, Faculdade de Medicina da Universidade de São Paulo, 2008.

DONG, J.; Li, J.; Fu, A.; Lv, H. Automatic segmentation for ovarian cancer immunohistochemical image based on yuv color space. In: INTERNATIONAL CONFERENCE ON BIOMEDICAL ENGINEERING AND COMPUTER SCIENCE - ICBECS, 2010. **Proceeding ... 2010**. p. 1–4. a.

DONG, J.; Li, J.; Lu, J.; Fu, A. Automatic segmentation for ovarian cancer immunohistochemical image based on chroma criterion. In: INTERNATIONAL CONFERENCE ON ADVANCED COMPUTER CONTROL - ICACC, 2, 2010. **Proceedings ... 2010**. v. 2, p. 147–150. b.

DUDA, R. O.; HART, P.E.; STORK, D. G. **Pattern Classification**. 2 ed. New York: Jhon Wiley, 2001.

FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**. Elsevier, v.27, p.861-874, 2006.

FICARRA, E.; MACII, E. Computer-aided evaluation of protein expression in pathological tissue images. In: COMPUTER-BASED MEDICAL SYSTEMS – CBMS, 19., 2006. 19th IEEE International Symposium ... IEEE, 2006. p. 413–418. ISSN 1063-7125.

FORSYTH, D. A.; PONCE, J. **Computer Vision A Modern Approach**. New Jersey: Prentice Hall, 2002.

GONZALEZ, R. C.; WOODS, Richard E.. **Processamento Digital de Imagens**. 3. ed. São Paulo: Pearson Prentice Hall, 2010.

GONZALEZ, R. F.; WOODS R. E. **Processamento de imagens digitais**. Tradutor: CESAR JR., R.M.; COSTA, L. D. F. São Paulo: Edgard Blücher, 2000.

GRUDIC, G. Z.; MULLIGAN, J. Outdoor Path Labeling Using Polynomial Mahalanobis Distance. **Robotics: Science and Systems II**, pp. 16-19, 2006.

GU, J.; LIU, C. Feature local binary patterns with application to eye detection. **Neurocomputing**. v. 113, p. 138-152, 02/2013.

GURCAN, M., BOUCHERON, L., CAN, A., MADABHUSHI, A., RAJPOOT, N., YENER, B. Histopathological image analysis: A review. **IEEE Reviews in Biomedical Engineering**, v.2, p.147-171, out. 2009.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEIN, I. Textural features for image classification. **IEEE Transactions on Systems, Man and Cybernetics**. v. SMC-3, i. 6, p. 610-621, 11/1973.

Hawkins, J.K. **Textural properties for pattern recognition.** B. Lipkin, A. Rosenfeld (Eds.), Picture Processing and Psychopictorics, Academic Press, New York (1969), pp. 347–370

HERNÁNDEZ-ORALLO, J.; FLACH , P.; FERRI , C.; ELKAN C. A Unified View of Performance Metrics: Translating Threshold Choice into Expected Classification Loss. **Journal of Machine Learning Research.** 2012.

HUI, W.; ZHIGUO, Z.; LONGMEI, J. Automatic detection of lumina in mouse liver immunohistochemical color image using support vector machine and cellular neural network. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND SECURITY - CIS, 7. **Proceedings ...** 2011, p. 1086–1090.

JITAREE, S.; PHINYOMARK, A.; THONGNOO, K.; BOONYAPIPHAT, P.; PHUKPATTARANONT, P. Classifying breast cancer regions in microscopic image using texture analysis and neural network. In: BIOMEDICAL ENGINEERING INTERNATIONAL CONFERENCE - BMEICON, 6th, 2013. **Proceedings ...** 2013. p. 1–4.

KAUR, A.; GUPTA, S. Texture Classification Based on Gabor Wavelets. **International Journal Of Research In Computer Science.** Ny, p. 39-44. jan. 2012.

KAYSER,K., GÖRTLER, J., BOGOVAC, M., BOGOVAC, A., GOLDMANN, T., VOLLMER, E., KAYSER, G. AI (artificial intelligence) in histopathology – from image analysis to automated diagnosis . **Folia Histochemica et Cytobiologica.** v. 47, n. 3, p 355-361, 2009.

KOSTOPOULOS, S.; CAVOURAS, D. ; DASKALAKIS, A. ; BOUGIOUKOS, P. ; GEORGIADIS, P. ; KAGADIS, G.C. ; KALATZIS, I. ; RAVAZOULA, P. ; NIKIFORIDIS, G. Colour-texture based image analysis method for assessing the hormone receptors status in breast tissue sections. In: ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY – EMBS, 29th Annual International Conference of the IEEE. **Proceddingis ...**, 2007. p. 4985–4988.

KOTHARI, S., PHAN, J., MO_TT, R., STOKES, T., HASSBERGER, S., CHAUDRY, Q., YOUNG, A., WANG, M. Automatic batchinvariant color segmentation of histological cancer images. In: INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING: FROM NANO TO MACRO. 2011, Chicago, **Proceedings ...**, IEEE, 2011. p. 657 - 660.

LAWS, Kenneth I. Rapid Texture Identification. In: IMAGE PROCESSING FOR MISSILE GUIDANCE, 238., 1980. **Proceedings ...** Spie, 1980. v. 238, p. 376 – 380

LI, X. L.; RUAN, Q.; RUAN, C. Facial Expression Recognition with Local Gabor Filters. In: INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING, 10, 2010, Beijing. **Proceedings... IEEE**, 2010, p. 1013-1016.

LIU, B.; WANG, W.; FANG, X. Automatic extraction of positive cells in tumor immunohistochemical pathology image based on ycber. In: SIXTH WORLD CONGRESS ON INTELLIGENT CONTROL AND AUTOMATION - WCICA, 6, 2006, Dalian. **Congress ... IEEE**, 2006. v. 2, p. 9708–9712.

LIU, F.; TANG, Z.; TANG, J. WLBP: Weber local binary pattern for local image description. **Neurocomputing.** v.120, p. 325-335, 11/2013.

LIU, L.; ZHAO, L.; LONG, Y.; KUANG, G.; FIEGUTH, P. Extended local binary patterns for texture classification. **Image and Vision Computing.** v. 30, i. 2, p. 86-99, 02/2012.

MAO, K.; ZHAO, P.; TAN, P.-H. Supervised learning-based cell image segmentation for p53 immunohistochemistry. **IEEE Transactions on Biomedical Engineering.** v. 53, n. 6, p. 1153–1163, 2006. ISSN 0018-9294.

MARQUES, F. O. ; VIEIRA, N. H. **Processamento digital de imagens**, Rio de Janeiro: Brasport, 1999.

MARTINS, D.F., MAZZARDO-MARTINS, L., GADOTTI, V.M., NASCIMENTO, F.P., LIMA, D.A.N., SPECKHANN, B., FAVRETTO, G.A., BOBINSKI, F., CARGNIN-FERREIRA, E., BRESSAN, E., DUTRA, R.C., CALIXTO, J.A.B., SANTOS, A.R.S. Ankle joint mobilization reduces axonotmesis-induced neuropathic pain and glial activation in the spinal cord and enhances nerve regeneration in rats. **PAIN**, 2011.

MATOS, L. L.; TRUFELLI, D. C.; MATOS, M. G. L.; PINHAL, M. A. S. Immunohistochemistry as an Important Tool in Biomarkers Detection and Clinical Practice. **Biomarker Insights.** v.5, p. 9-20, fev.2010.

METE, M. ; TOPALOGLU, U. Statistical comparison of color model-classifier pairs in hematoxylin and eosin stained histological images. In: IEEE SYMPOSIUM ON COMPUTATIONAL INTELLIGENCE IN BIOINFORMATICS AND COMPUTATIONAL BIOLOGY - CIBCB, 9th, Neshville, 2009. **Proceedings ... IEEE Computational Intelligence Society**, 2009, p. 284 – 291.

MOKJI, M. M.; BAKAR, A. Gray Level Co-Occurrence Matrix Computation Based On Haar Wavelet. In: COMPUTER GRAPHICS, IMAGING AND VISUALISATION – CGIV, 07, 2007, Bangkok. **Proceedings ... IEEE Computer Society**, 2007, p. 273-279.

MOUELHI, A.; SAYADI, M.; MRAD, K.; ROMDHAME, B. K. Automatic image segmentation of nuclear stained breast tissue sections using color active contour model and an improved watershed method. **Biomedical Signal Processing and Control.** v.8, i.5, p.421-436, set. 2013.

MOUELHI, A.; SAYADIA, M.; FNAIECHA, F.; MRADB, K.; ROMDHANE, K. B. Automatic image segmentation of nuclear stained breast tissue sections using color active contour model and an improved watershed method. **Biomedical Signal Processing and Control.** v.8, i.5, p.421-436, set. 2013.

NANNI, L. BRAHNAM, S.; LUMINI, A. A simple method for improving local binary patterns by considering non-uniform patterns. **Pattern Recognition.** v. 45, 04/2012.

NAVANI, S. The human protein atlas. **The Journal of Obstetrics and Gynecology of India.** v. 61, i. 1, 02/2011, p. 27-31.

OJALA, T.; PIETIKAINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition.** v. 29, n. 1, p. 51-59. 01/1996.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence.** v.24 , i.7, p. 971-987, 07/2002

PATRA, S. K.; RAJU, J.; MARUTHUPERUMAL, S. Texture analysis using statistical and structural approaches. **International Journal of Latest Trends in Engineering and Technology (IJLTET).** v. 3, i. 1, 09/2013.

PEARSON, K. On lines and planes of closest fit to systems of points in space. **Philosophical Magazine.** v. 2, n. 6, 1901, p. 559-572.

PETROU, M. **Image processing the Fundamentals.** Chichester: Jhon Wiley and Sons LTDA,1999.

PHAM, N.A., MORRISON, A., SCHWOCK, J., AVIEL-RONEN, S., IAKOVLEV, V., TSAO, M.S.,HO, J., HEDLEY, D. Quantitative image analysis of immunohistochemical stains using a cmyk color model. **Diagnostic Pathology** p. 1-10, 2007.

PIETIKÄINEN, M.; HADID, A.; ZHAO, G.; AHONEN, T. Local Binary Patterns for Still Images. In: PIETIKÄINEN, M.; A.; ZHAO, G.; AHONEN, T. **Computer Vision Using Local Binary Patterns.** Springer Link, v. 40. 2011, p. 13-47.

PONTÉN, F.; JIRSTRÖM, K.; UHLEN, M. The Human Protein Atlas a tool for pathology. **The Journal of pathology.** v. 216, n. 4, 2008, p. 387-393.

PRASAD, K., PRABHU, G. Image analysis tools for evaluation of microscopic views of immunohistochemically stained specimen in medical research, a review. **Journal of Medical systems** p. 1-11, 2011.

PRASETYO; KHALID, M.; YUSOF, R.; MARIAUDEAU, F. A Comparative Study of Feature Extraction Methods for Wood Texture Classification. In: INTERNATIONAL CONFERENCE ON SIGNAL-IMAGE TECHNOLOGY AND INTERNET BASED SYSTEMS - SITIS, 6, 2010, Kuala Lumpur. **Proceedings ... IEEE Computer Society**, 2010, p. 23-29.

PROTEIN ATLAS (Suécia). **The human protein atlas Project.** 2014. Disponível em: <<http://www.proteinatlas.org/about/project>>. Acesso em: 22 jun. 2014.

RICHTER, M. PETYOVSKY, P.; MIKSIK, O. Adapting Polynomial Mahalanobis Distance for Self-supervised Learning in an Outdoor Environment. In: 10TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND APPLICATIONS AND WORKSHOPS - ICMLA, 12, 2011. **Proceedings ... IEEE Computer Society** v.1, p.448-453.

RUSSELL , S.; NORVING, P. **Artificial Intelligence: A Modern Approach.** Prentice Hall Press, Upper Saddle River, NJ, USA.1994.

SAMPAIO, R.F.; MANCINI, M.C. Estudos de Revisão Sistemática: Um Guia para Síntese Criteriosa da Evidência Científica. **Revista Brasileira de Fisioterapia.** São Carlos, v. 11, n.1, p. 83-89, jan./fev., 2007.

- SAMSI, S. LOZANSKI, G.; SHANARAH, A.; KRISHANMURTHY, A. K.; GURCAN, M. N. Detection of follicles from IHQ-stained slides of follicular lymphoma using iterative watershed. **IEEE Transactions on Biomedical Engineering.** v. 57, n. 10, p. 2609–2612, Oct 2010. ISSN 0018-9294.
- SHARMA, M.; SINGH, S. Evaluation of texture methods for image analysis. In: THE SEVENTH AUSTRALIAN AND NEW ZEALAND INTELLIGENT INFORMATION SYSTEMS CONFERENCE, 17., 2001, Nz. **Proceedings... .** Nz: Ieee, 2001. p. 117 - 121.
- SHU, J.; QIU, G.; MOHAMMAD, I. A semi-automatic image analysis tool for biomarker detection in immunohistochemistry analysis. In: INTERNATIONAL CONFERENCE ON IMAGE AND GRAPHICS - ICIG, 7. **Proceedings ...** Elsevier Science, 2013, p. 937–942.
- SIGNOLLE, N.; REVENUA, M.; PLANCOULAINEB, B.; HERLINB, P. Wavelet-based multiscale texture segmentation: Application to stromal compartment characterization on virtual slides. **Signal Processing.** v.90, i.8, p.2412-2422, 08/2010.
- SOBIERANSKI A. C.; COMUNELLO E.; VON WANGENHEIM, A. Learning a nonlinear distance metric for supervised region-merging image segmentation. **Computer Vision and Image Understanding.** v. 115, i.2, p.127-139, 2011.
- SOBIERANSKI, A.; NETO, S. L. M.; COSER, L.; COMUNELLO, E. WANGENHEIM A.V.; FERREIRA, E. C.; GIUNTA, G. D. Learning a nonlinear color distance metric for the identification of skin immunohistochemical staining. In: COMPUTER-BASED MEDICAL SYSTEMS – CBMS, 22, 2009. **IEEE International Symposium.** 2009. p. 1–7. ISSN 1063-7125.
- SOLOMON, C. ; BRICKON, T. **Fundamentos de Processamento Digital de Imagens:** Uma abordagem Prática com Exemplos em Matlab. São Paulo: LTC- Grupo GEM, 2013.
- SONG, K.; YAN, Y.; CHEN, W.; ZHANG, X. Research and perspective on local binary pattern. **Acta Automatica Sinica.** v. 39, i. 6, 06/2013.
- UHLEN, M.; OKSVOLD, P.; FAGERBERG, L.; LUNDBERG, E.; JONASSON, K.; FORSBERG, M.; ZWAHLEN, M.; KAMPF, C.; WESTER, K.; HOBER, S.; WERNERUS, H. BJORLING, L. PONTEN, F. Towards a knowledge-based Human Protein Atlas. **Nature Biotechnology,** v. 28, 02/2010, p. 1248-1250.
- VIRMANI, J.; KUMAR, V.; KARLA, N.; KHANDELWAL, N. Prediction of cirrhosis from liver ultrasound B-mode images based on Laws' masks analysis. In: INTERNATIONAL CONFERENCE ON IMAGE INFORMATION PROCESSING - ICIIP, 2011, Himachal Pradesh. **Proceedings ... IEEE,** 2011, p. 1-5.
- WAZLAWICK, R. S. **Metodologia de Pesquisa para Ciência de Computação.** Rio de Janeiro: Elsevier, 2009.
- WEMMERT, C.; *et al.* Stain unmixing in brightfield multiplexed immunohistochemistry. In: 20th IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING - ICIP, 20., 2013, Melbourne. **Proceedings... IEEE,** 2013. p. 1125 - 1129.

XU, Y.; ZHU, J. CHANG, E. I.; LAI, M.; TU, Z. Weakly Supervised Histopathology Cancer Image Segmentation and Classification. **Medical Image Analysis.** v. 18, i. 3, 04/2014, p. 591-604.

YAGI, Y.: **Critical comparison of digital pathology systems.** 2007. Disponível em: <http://www.conganat.org/eurotelepath/WG4/Yukako_Yagi_Critical_Comparison.pdf>. Acesso em: 06/07/2014.

YANG, H., ZHANG, X., WANG, X. LS-SVM-based Image Segmentation Using Píxel Color-Texture Descriptors. **Pattern Analysis and Applications.** September, 2012.

ZHAO, P.; MAO, K.Z.; KOH, T. S.; TAN, P. H. Automatic cell analysis for p53 immunohistochemistry in bladder inverted papilloma. In: ASIAN-PACIFIC CONFERENCE ON BIOMEDICAL ENGINEERING, 2003. **Proceedings ... 2003**, p. 168-169.

ZHAO, W.; PARK, J.; LEE, S. Fully automatic face detection and facial feature points extraction using local Gabor filter bank and PCA. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND CYBERNETICS – ICMLC. v. 4, 2011, Guilin. **Proceedings ... IEEE**, 2011, p. 1789 – 1792.

ZHENG, N.; XUE, J. **Statistical Learning and Pattern Analysis for Image and Video Processing:** Advances in Pattern Recognition. London: Limited, Springer-Verlag, 2009.

ZUVA,T., OLUGBARA, O., O., OJO, S. O., NGWIRA, S. M. Image Segmentation, Available Techniques, Developments and Open Issues. **Canadian Journal on Image Processing and Computer Vision.** v.2, n3, 2011.

GLOSSÁRIO

Confocal	Microscopia Confocal é uma técnica imagiológica patenteada por Marvin Minsky em 1955. ConfocalMemoir²⁹
Micrótomo	Instrumento que serve para cortar nos tecidos animais ou vegetais tiras delgadas para serem estudadas ao microscópio.
Proteoma	Conjunto das proteínas produzidas por uma célula ou por um organismo

²⁹ <https://web.media.mit.edu/~minsky/papers/ConfocalMemoir.html>

APÊNDICE A – PROTOCOLO DE BUSCA

O protocolo de busca elaborado para o presente trabalho definiu os seguintes critérios: (I) pergunta de pesquisa; (II) população; (III) resultados; (IV) estratégia para pesquisa; (V) critérios e procedimentos para seleção; (VI) lista de verificação e procedimentos para avaliação da qualidade de estudos e (VII) reunião dos dados. Os referidos critérios são:

1. Pergunta de pesquisa:

Primária: Quais métodos computacionais são empregados na análise de imagens de imuno-histoquímica, especificamente relacionados à segmentação e quantificação de coloração imuno-histoquímica?

2. População:

- Abordagens computacionais para análise de imagens microscópicas de imuno-histoquímica.

3. Resultados:

- Métodos computacionais para segmentação e quantificação de coloração imuno-histoquímica;
- Resultados alcançados por estes métodos; e
- Problemas em aberto apontados nos trabalhos selecionados.

4. Estratégia utilizada para pesquisa dos estudos primários:

- Palavras chaves: *immunohistochemistry staining, immunohistochemistry quantification, immunohistochemical staining, immunohistochemical quantification, pattern recognition, image analysis, semi-supervised learning, nonlinear color distance metric, texture analysis.*

- *String* de busca: Devido as diferentes sintaxes permitidas, nas três fontes de pesquisas, a *string* de busca foi adequada de acordo com cada mecanismo de busca empregado na revisão sistemática. Na Subseção C1 foram apresentadas as *strings* utilizadas em cada fonte de pesquisa, *string* de busca padrão estabelecida foi:

(

("immunohistochemistry staining" OR "immunohistochemistry quantification" OR
"immunohistochemical staining" OR "immunohistochemical quantification")

AND

("pattern recognition" OR "image analysis" OR "semi-supervised learning" OR
"nonlinear color distance metric" OR "texture analysis")

)

- Fontes:

ACM Digital Library <http://portal.acm.org>

IEEEExplore <http://ieeexplore.ieee.org>

ScienceDirect <http://www.sciencedirect.com/>

5. Critérios e procedimentos para seleção de estudos:

- Análise do título do trabalho;
- Análise do resumo do trabalho; e
- Pela data de publicação do trabalho (entre 2003 e 2014).
- Critérios para inclusão: Trabalhos cujos títulos e resumos contenham informações sobre métodos e técnicas computacionais, empregados na quantificação de coloração imuno-histoquímica em imagens microscópicas.
- Critérios para exclusão: Trabalhos cujos títulos e resumos eram conflitantes, ou seja, o título remetia a um assunto enquanto o resumo remetia a outro. Trabalhos que não tinham o foco voltado à segmentação e quantificação imuno-histoquímica, e ainda trabalhos especificamente voltados à área médica.

6. Lista de verificação e procedimentos para avaliação da qualidade dos estudos:

- Objetivos: O objetivo dos trabalhos estava focado em métodos computacionais para a quantificação de coloração imuno-histoquímica.
- Condução: Os trabalhos deveriam estar bem referenciados e possuir, preferencialmente, uma etapa de avaliação experimental com abordagem quantitativa.

7. Reunião de dados:

- Título;
- Ano de Publicação / Conceito (Qualis CAPES) ;
- Autores
- Principal problema abordado;
- Solução Proposta;
- Existência de implementação disponível;
- Métricas de avaliação;
- Resultados obtidos;
- Problemas em aberto; e
- Disponibilidade do conjunto de imagens utilizado.

C1 - STRINGS DE BUSCA

A seguir são apresentadas as *strings* específicas, utilizadas em cada mecanismo de busca. Bem como as restrições permitidas em cada mecanismo.

IEEE

Neste mecanismo é possível *restringir* a busca por título, resumo, termos indexados e ano de publicação, a *string* empregada foi:

```
7  (
8    ("Document Title": "pattern recognition" OR "semi-supervised learning" OR "image analysis" OR "nonlinear color distance metric" OR "texture analysis")
9    OR
10   ("Index Terms": "pattern recognition" OR "semi-supervised learning" OR "image analysis" OR "nonlinear color distance metric" OR "texture analysis")
11   OR
12   ("Abstract": "pattern recognition" OR "semi-supervised learning" OR "image analysis" OR "nonlinear color distance metric" OR "texture analysis")
13
14 )
15 AND
16
17 (
18   ("Document Title": "immunohistochemistry staining" OR "immunohistochemistry quantification" OR "immunohistochemical staining" OR "immunohistochemical quantification"
19   OR immunohistochemical OR immunohistochemistry)
20   OR
21   ("Index Terms": "immunohistochemistry staining" OR "immunohistochemistry quantification" OR "immunohistochemical staining" OR "immunohistochemical quantification"
22   OR immunohistochemical OR immunohistochemistry)
23   OR
24   ("Abstract": "immunohistochemistry staining" OR "immunohistochemistry quantification" OR "immunohistochemical staining" OR "immunohistochemical quantification"
25   OR immunohistochemical OR immunohistochemistry)
26   OR
27
28 )
```

Figura 32. *String* de busca IEEE

Science Direct

Neste mecanismo de busca foi possível aplicar a restrição de data de publicação e fonte (ciência da computação). A *string* de busca empregada foi:

```
91  (
92  ("immunohistochemistry staining" OR "immunohistochemistry quantification" OR "immunohistochemical staining" OR "immunohistochemical quantification" )
93  AND
94  ("pattern recognition" OR "image analysis" OR "semi-supervised learning" OR "nonlinear color distance metric" OR "texture analysis")
95  )
96
```

Figura 33. *String* Science Direct

ACM Digital Library

Neste mecanismo de busca foi possível aplicar restrição para título, resumo, palavras chaves e data de publicação. A *string* de busca empregada foi:

```
44
45 (
46 (
47 (Title: "pattern recognition" OR "semi-supervised learning" OR "image analysis" OR "nonlinear color distance metric" OR "texture analysis")
48 OR
49 (Abstract: "pattern recognition" OR "semi-supervised learning" OR "image analysis" OR "nonlinear color distance metric" OR "texture analysis")
50 OR
51 (Keywords: "pattern recognition" OR "semi-supervised learning" OR "image analysis" OR "nonlinear color distance metric" OR "texture analysis")
52 )
53
54 AND
55
56 (
57 (Title: "immunohistochemistry staining" OR "immunohistochemistry quantification" OR "immunohistochemical staining" OR "immunohistochemical quantification" )
58 OR
59 (Abstract: "immunohistochemistry staining" OR "immunohistochemistry quantification" OR "immunohistochemical staining" OR "immunohistochemical quantification" )
60 OR
61 (Keywords: "immunohistochemistry staining" OR "immunohistochemistry quantification" OR "immunohistochemical staining" OR "immunohistochemical quantification" )
62 )
63 )
64 )
```

Figura 34. *String ACM*

APÊNDICE B – DIAGRAMA DE SEQUÊNCIA

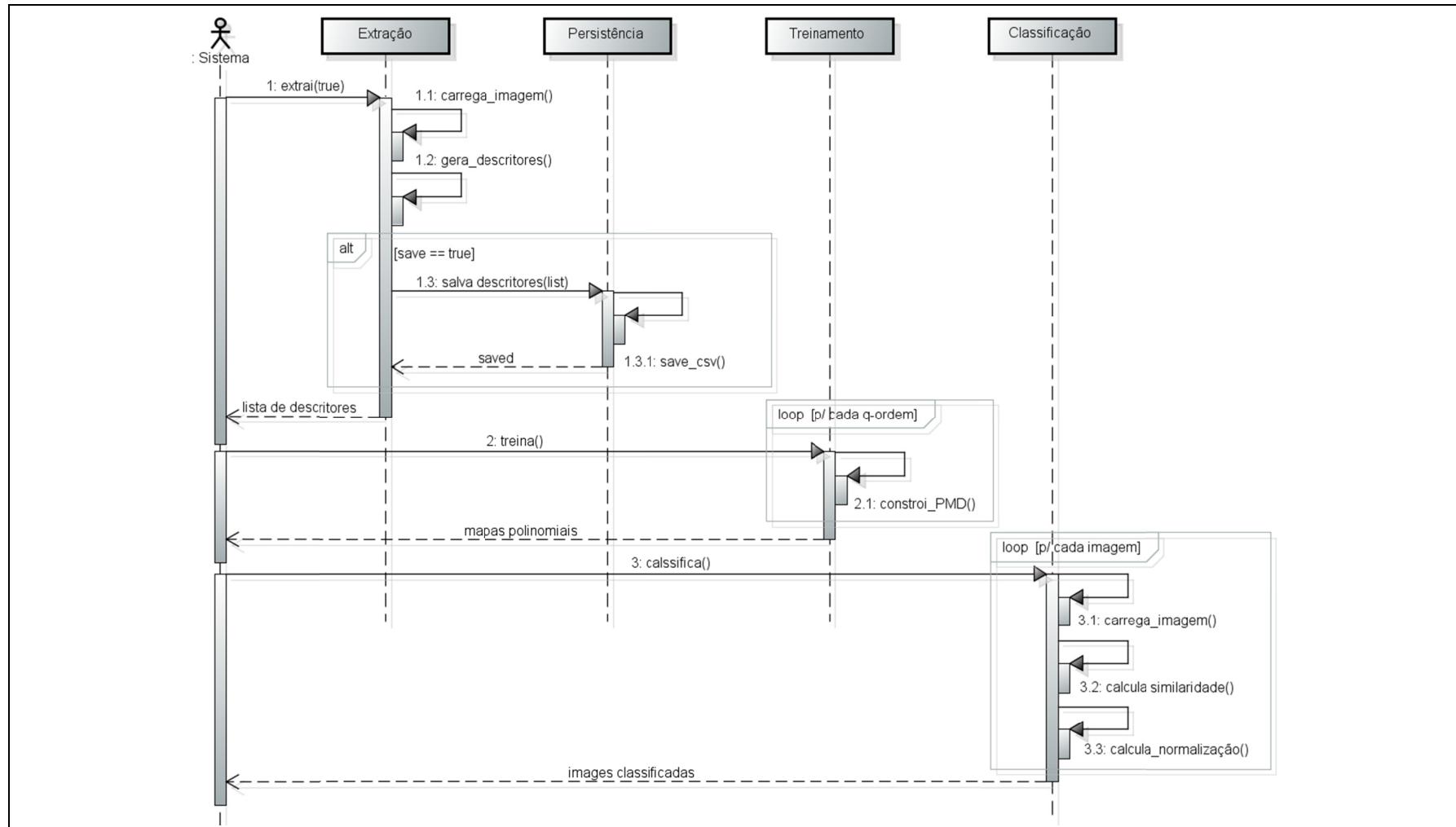


Figura 35. Sequência das etapas de extração de características, treinamento e classificação

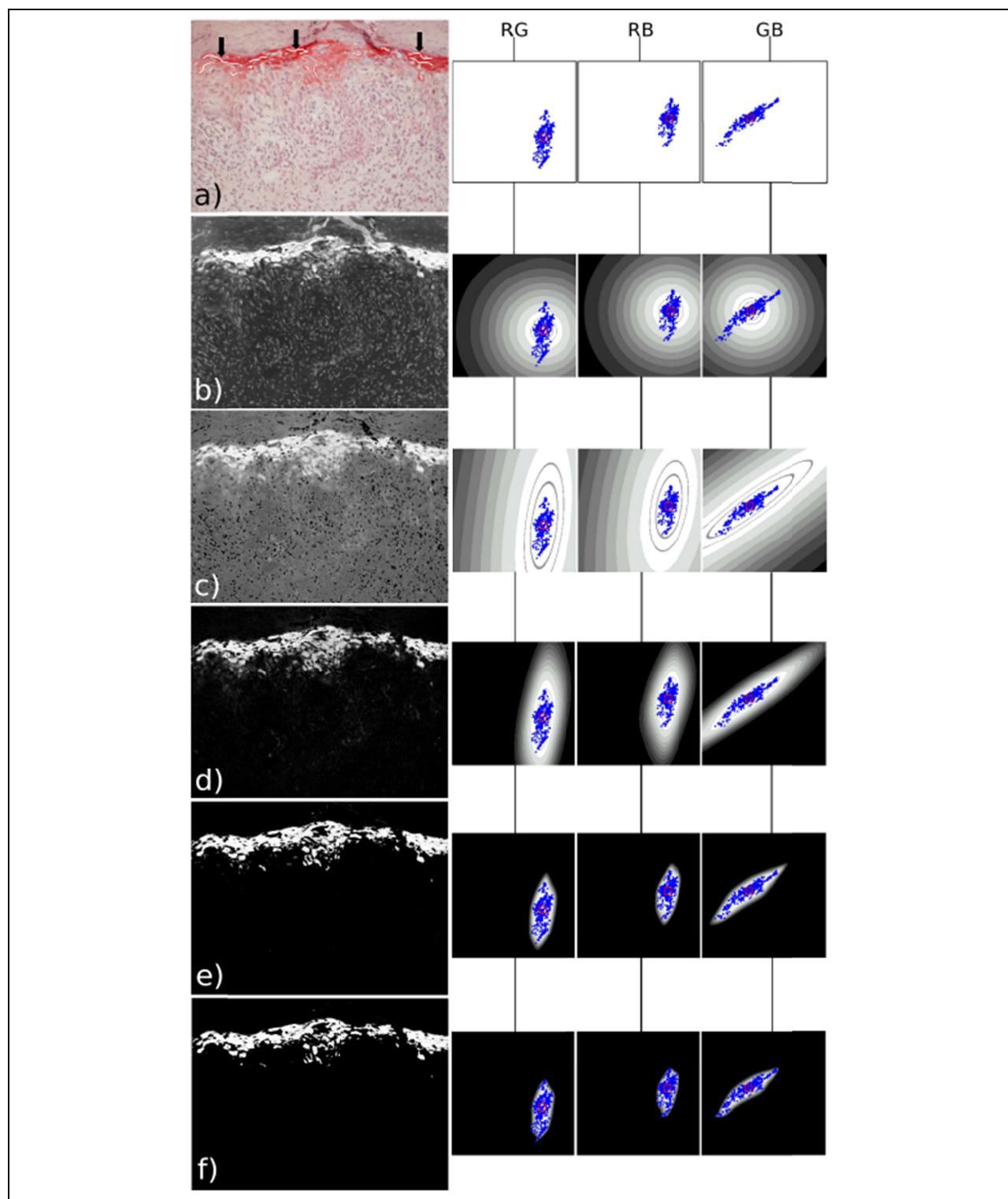
ANEXO A – DPM VS. DISTÂNCIA EUCLIDIANA

Figura 36. Comparação das etapas de treinamento e classificação da DPM vs. DE

Fonte: Sobieranski *et al.*, 2009.

A Figura 36 compara o resultado de classificação da métrica DPM, durante as criações das projeções polinomiais, contra o resultado da uma classificação, empregando a distância euclidiana. A imagem a) é a original (canto superior esquerdo), ilustra a amostra a ser classificada e a definição de pontos que representam o padrão de interesse (coloração avermelhada) e serão usados no treinamento, as linhas brancas indicam a marcação dos pontos. Nesta abordagem somente os atributos de cor são considerados no descritor de características.

No canto superior esquerdo é apresentada a decomposição da plotagem, em suas componentes bidimensionais (RG, RB e GB) do padrão selecionado, os pontos em azul representam o padrão alvo.

ANEXO B - ILUSTRAÇÃO DE TMAS USADOS NO HPA

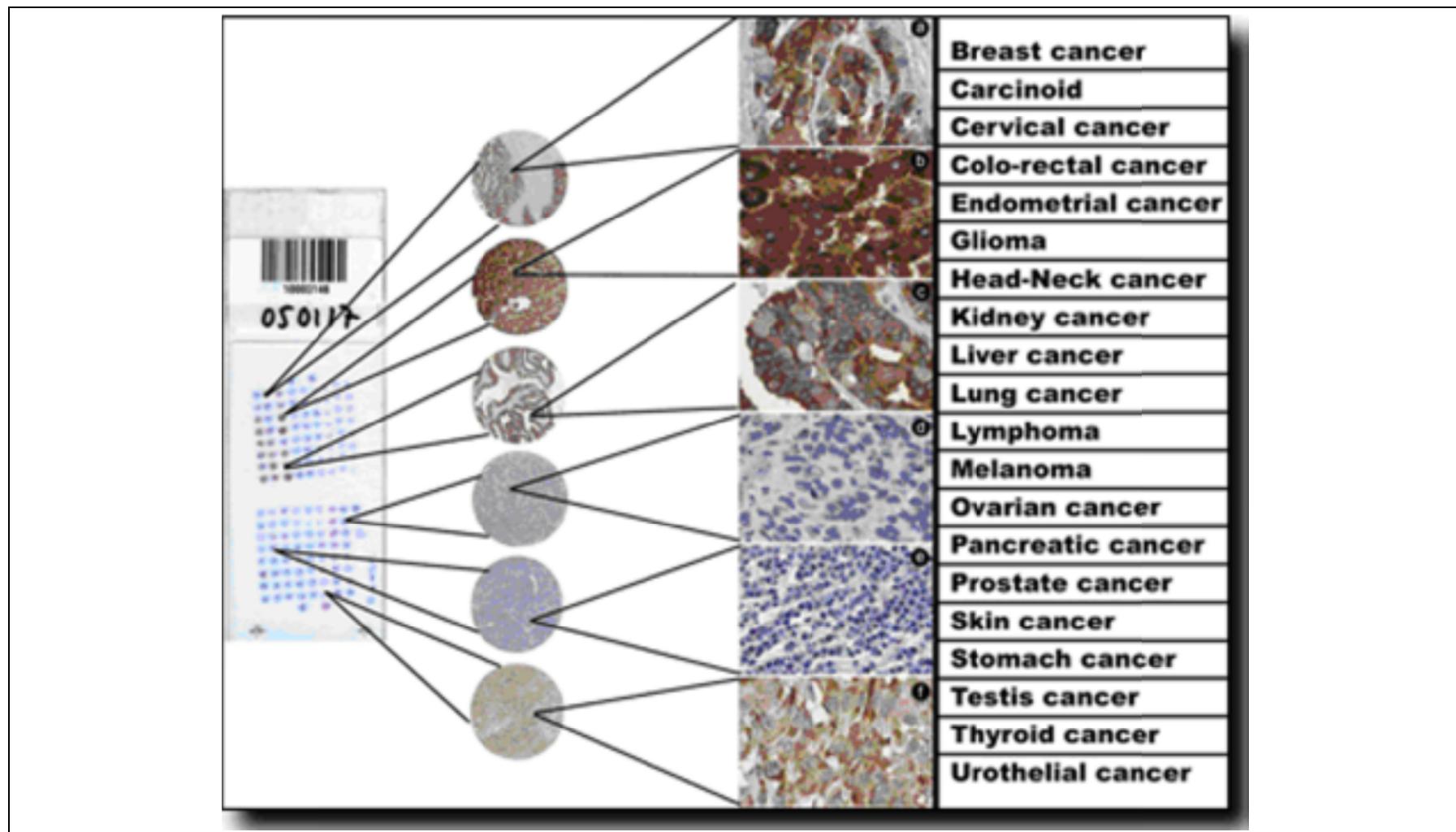


Figura 37. Representação dos tecidos cancerosos no HPA

Fonte: NAVANI (2011).

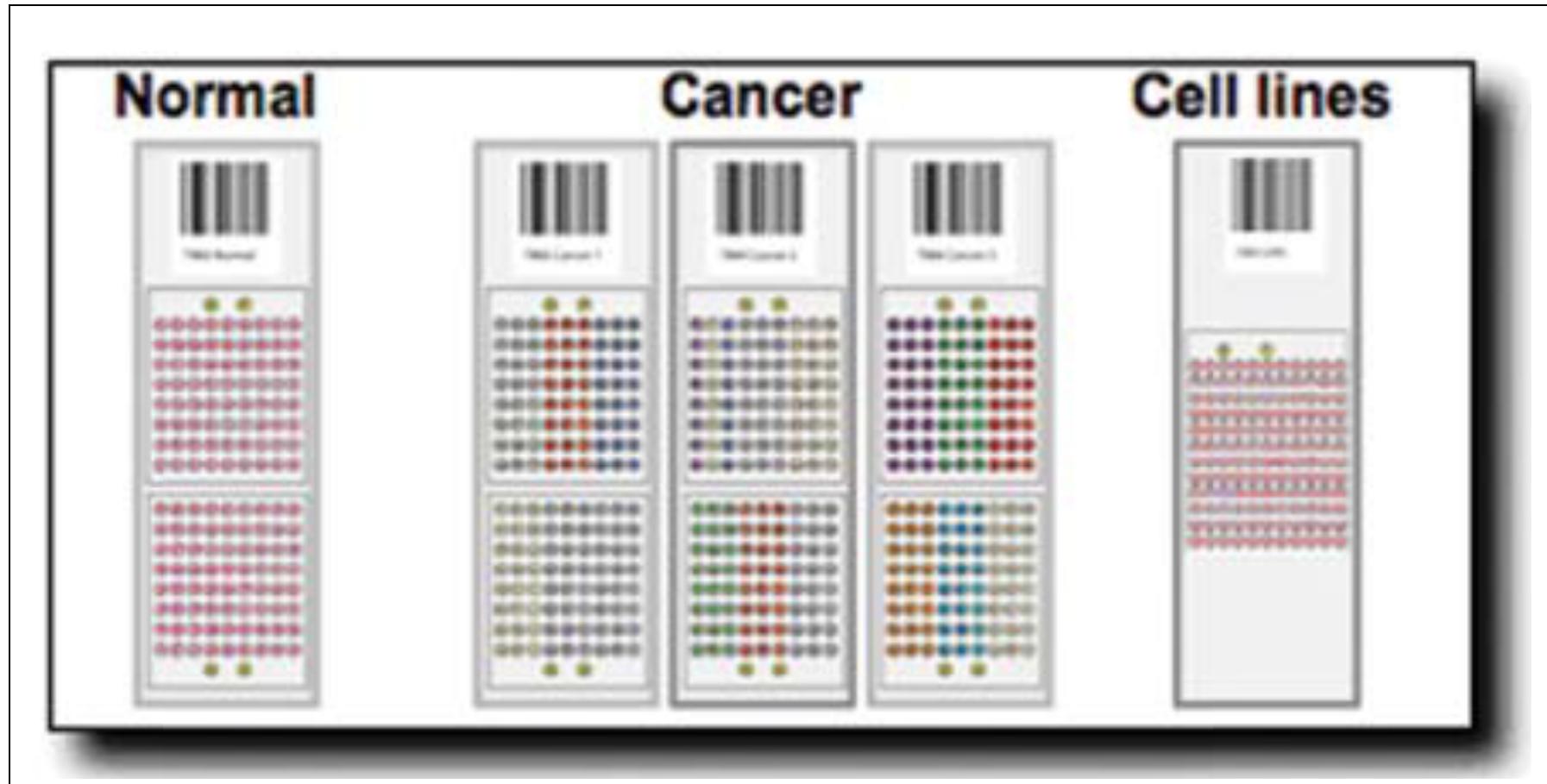


Figura 38. Lâminas de Microarranjo usadas no HPA

Fonte: NAVANI (2011).

Os microarranjos de tecidos consistem em blocos de parafina, nos quais núcleos de tecido são agrupados em disposição de arranjo em uma única lâmina. Este procedimento facilita a análise de IHQ para diferentes anticorpos em uma mesma seção de tecido. Usualmente um bloco contém 2 ou 3 núcleos, com medidas entre 0,6mm e 2mm de diâmetro (NAVANI, 2011).

ANEXO C - METADADOS HPA LINHAS DE CÉLULAS

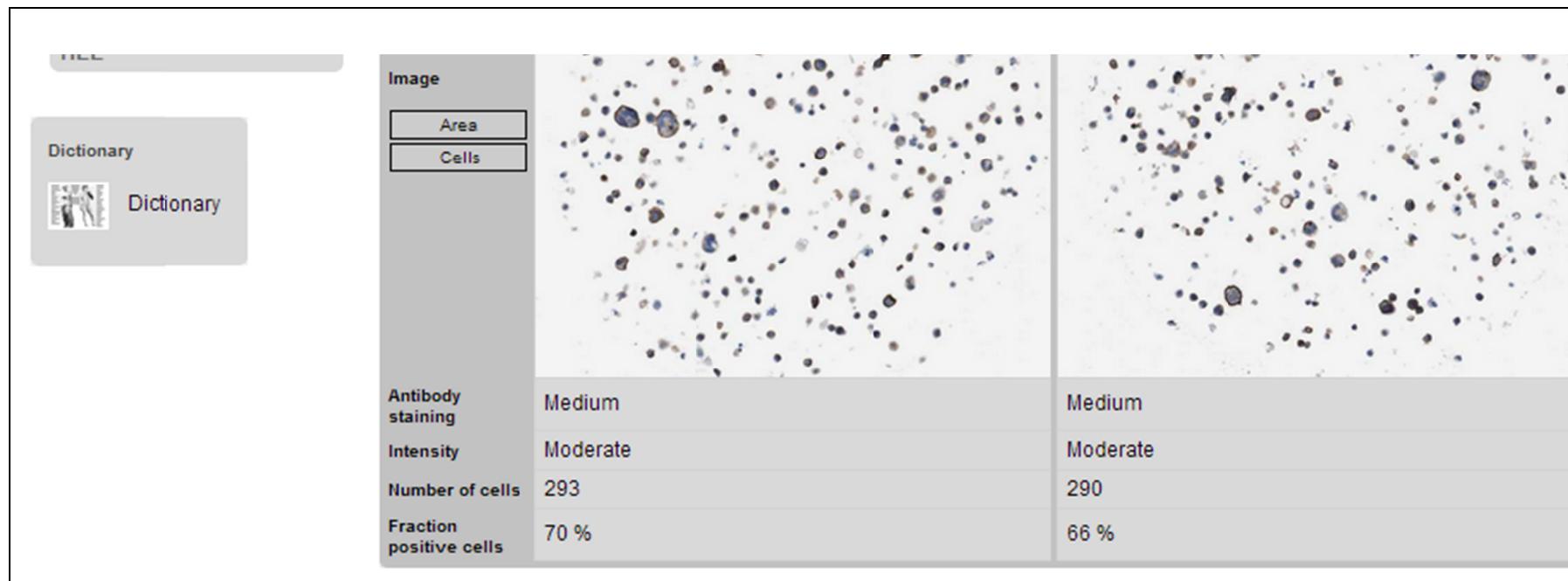


Figura 39. Metadados para as imagens na categoria linha da LCH, onde cada imagem diz respeito a um núcleo de TMAs

Fonte: Adaptado de Human Protein Atlas (2014)³⁰.

³⁰ <http://www.proteinatlas.org/ENSG00000026508/cell/HPA005785/HEL>

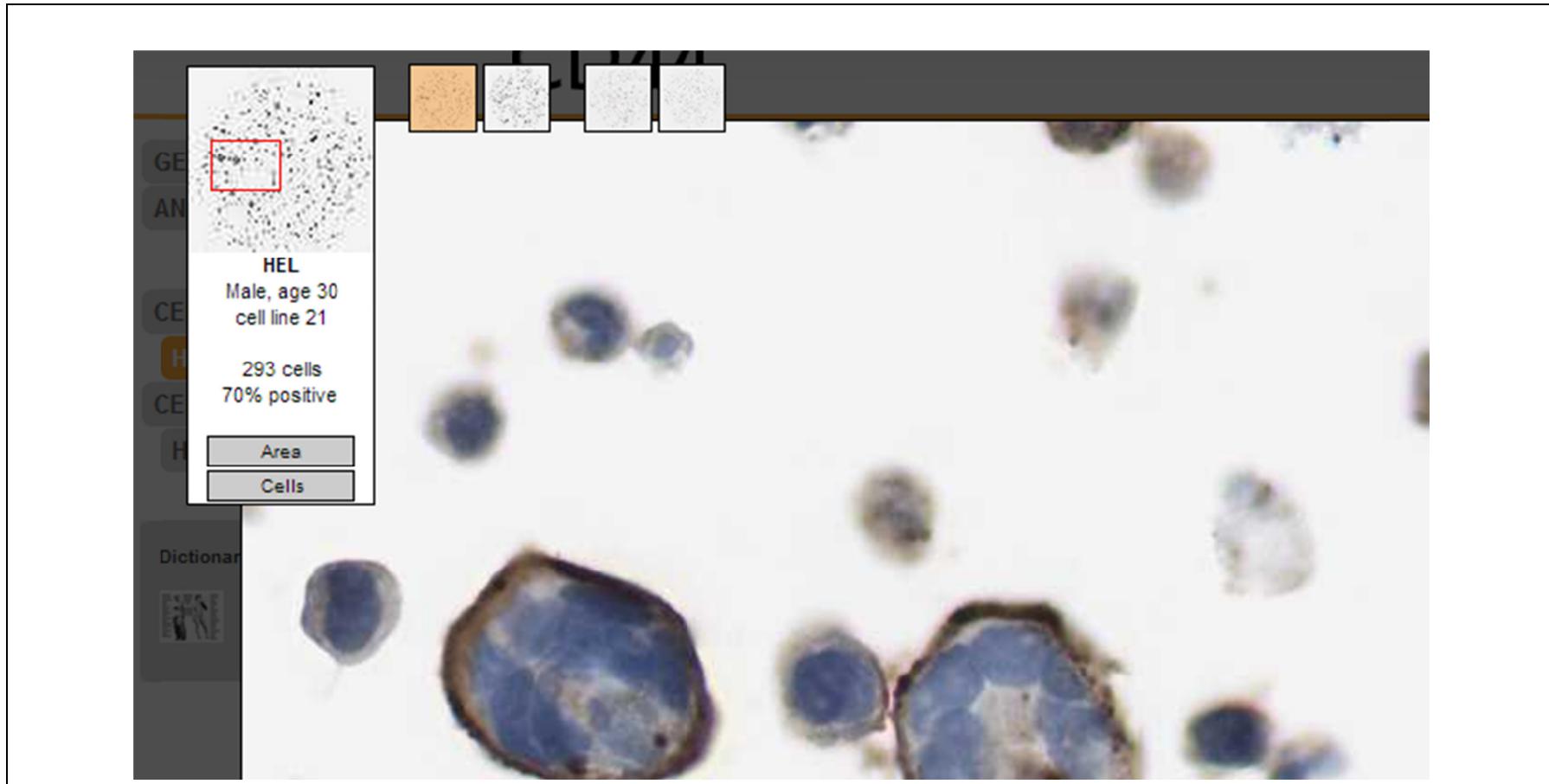


Figura 40. Detalhe de metadados para uma imagem

Fonte: Adaptado de Human Protein Atlas (2014)³¹.

³¹ <http://www.proteinatlas.org/ENSG00000026508/cell/HPA005785/HEL#img>