**Classificação de imagens de fluorescência do citoesqueleto
através de técnicas em processamento de imagens**

Filomen Incahuanaco Quispe

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências
de Computação e Matemática Computacional (PPG-CCMC)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Filomen Incahuanaco Quispe

Classificação de imagens de fluorescência do citoesqueleto através de técnicas em processamento de imagens

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Afonso Paiva Neto

USP – São Carlos
Novembro de 2017

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

I35c Incahuanaco Quispe, Filomen
Classificação de imagens de fluorescência do
citoesqueleto através de técnicas em processamento
de imagens / Filomen Incahuanaco Quispe; orientador
Afonso Paiva Neto. -- São Carlos, 2017.
80 p.

Dissertação (Mestrado - Programa de Pós-Graduação
em Ciências de Computação e Matemática
Computacional) -- Instituto de Ciências Matemáticas
e de Computação, Universidade de São Paulo, 2017.

1. Imagens microscópicas fluorescentes. 2. BoVW.
3. LBP. 4. Textons. 5. Redes Complexas. I. Paiva
Neto, Afonso, orient. II. Título.

Filomen Incahuanaco Quispe

**Classification of cytoskeleton in fluorescence images with
image analysis techniques**

Master dissertation submitted to the Instituto de
Ciências Matemáticas e de Computação – ICMC-
USP, in partial fulfillment of the requirements for the
degree of the Master Program in Computer Science
and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and
Computational Mathematics

Advisor: Prof. Dr. Afonso Paiva Neto

USP – São Carlos
November 2017

Este trabalho é dedicado a minha família: Meu pai Guillermo(in memoriam), minha mãe Rosa Elena e meus irmãos Luz Mary, Jorge Wilbert e Lourdes Sarita.

AGRADECIMENTOS

Agradeço primeiramente a Deus, pelas provas postas em meu caminho para amadurecer, finalmente, conseguir encontrar o caminho certo na minha preparação acadêmica e pessoal.

A minha orientadora, Profa. Dra. Cynthia Oliveira, pelo seu permanente apoio e atenção dispensada no decorrer deste trabalho.

A meu co-orientador Prof. Dr. Afonso Paiva Neto, pelo suporte e sua disponibilidade irrestrita, sua forma amiga, exigente e crítica.

Ao Prof. Dr. João Batista, pelo apoio acadêmico e disposição praticamente imediata para consultas em relação a pesquisa.

Aos professores Dra. Maria Cristina, Dr. Gustavo Batista pelas orientações durante o programa de aperfeiçoamento de ensino (PAE), onde tive a oportunidade de crescer no âmbito do ensino.

Ao prof. Dr. César Beltran Castañon, pelo seu permanente apoio para eu reiniciar minha formação acadêmica, a sua disponibilidade irrestrita, sua forma amiga, exigente e crítica, fundamental contribuição no meu crescimento enquanto pesquisador.

Aos colegas da toda a vida Angel e Omar que deram o suporte emocional, durante o período deste trabalho que não fiquei em casa.

Aos colegas Ricardo, Mayra, Rommel, Maicol pela amizade fomentada desde muito antes do começo da pós no ICMC.

Aos colegas Jorge e Junior pela ajuda e orientação em relação a minha formação e pesquisa.

Aos colegas do ICMC no período da minha preparação acadêmica: Misael, Joelson, Evinton, Lucas, Francisco(Chico), pelos momentos vividos durante o período desta etapa.

Aos colegas do ICMC André, Jonathan, Paulo e Diego pelas correções da minha escrita e amizade.

Aos colegas do laboratório de Visualização, Imagens e Computação Gráfica (VICG - ICMC/USP 1007 Bloco 1) pelas extensas horas de trabalho compartilhadas aos 23 graus obrigatórios que foram me gelando dia a dia.

Para os senhores(as) da segurança do bloco 1 e 4 do ICMC, pelo trato amável e disposição.

Finalmente e, não menos importante A CAPES, pelo apoio financeiro.

“ Si piensas que estas vencido, lo estas. ”
(Christian Barnard)

RESUMO

INCAHUANACO, F. **Classificação de imagens de fluorescência do citoesqueleto através de técnicas em processamento de imagens**. 2017. 80 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

O citoesqueleto é a estrutura celular mais importante em células eucariotas e é responsável por manter a forma da célula e as junções celulares, auxiliando nos movimentos celulares. Esta é composta de filamentos de Actina, Microtúbulos e filamentos intermediários. Recentemente, a análise de duas dessas estruturas tornaram-se importantes, pois é possível obter micrografias usando microscópios de alta resolução, que contém microscopia de fluorescência, em combinação com métodos complexos de aplicação de substâncias de contraste para rotulagem e posterior análises visuais. A combinação dessas técnicas, entretanto, limita-se a ser descritiva e subjetiva. Neste trabalho, são avaliadas cinco técnicas de análise de imagens, as quais são: *Bag of Visual Words* (BoVW), *Local Binary Local* (LBP), Textons baseados em *Discrete Fourier Transform* (TDFT), Textons baseados em *Gabor Filter Banks* (TGFB) e Textons baseados em *Complex Networks* (TCN) sobre o conjunto de dados *2D Hela* e *FDIG Olympus*. Experimentos extensivos foram conduzidos em ambos os conjuntos de dados, e seus resultados podem servir de base para futuras pesquisas como análises do citoesqueleto em imagens de microscopia fluorescente. Neste trabalho, é apresentada uma comparação quantitativa e qualitativa dos métodos acima mencionados para entender o comportamento desses métodos e propriedades dos microfilamentos de actina (MA) e Microtúbulos (MT) em ambos os conjuntos de dados. Os resultados obtidos evidenciam que é possível classificar o conjunto de dados da *FDIG Olympus* com uma precisão de até 90.07% e 98.94% para *2D Hela*, além de obter 86.05% e 96.84%, respectivamente, de precisão, usando teoria de redes complexas.

Palavras-chave: Imagens microscópicas fluorescentes, BoVW, LBP, Textons, Redes Complexas.

ABSTRACT

INCAHUANACO, F. **Classification of cytoskeleton in fluorescence images with image analysis techniques**. 2017. 80 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

The cytoskeleton is the most important cellular structure in eukaryotic cells and is responsible for maintaining the shape of the cell and cellular junctions, aiding in cell movements. This is composed of filaments of Actin, Microtubules and intermediate filaments. Recently, the analysis of two of these structures has become important because it is possible to obtain micrographs using microscopes of high resolution and fluorescence technology, in combination with complex methods of application of substances of contrast for labeling and later visual analysis. The use of these techniques, however, is limited to being descriptive and subjective. In this work, we evaluate some of the most popular image analysis techniques such as Bag of Visual Words (BoVW), Local Binary Pattern (LBP), Textons based on Discrete Fourier Transform (TDFT), Gabor Filter banks (TGFB), and approaches based on Complex Networks theory (TCN) over the famous dataset 2D Hela and FDIG Olympus. Extensive experiments were conducted on both datasets in which their results can serve as a baseline for future research with cytoskeleton classification in microscopy fluorescence images. In this work, we present the quantitative and qualitative comparison of above mentioned methods for better understand the behavior of these methods and the properties of Actin microfilaments (MA) and Microtubules (MT) on both datasets. The results showed that it is possible to classify the FDIG Olympus data set with accuracy of up to 90.07% and 98.94% for 2D Hela, in addition to reaching 86.05% and 96.84% respectively, using complex network theory.

Keywords: Microscopy fluorescence image, BoVW, LBP, Textons, Complex Networks.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de imagens fluorescentes	24
Figura 2 – Diagrama da dispersão da luz através de um prisma	28
Figura 3 – Espectro Contínuo	29
Figura 4 – Espectro Contínuo visível	29
Figura 5 – Deslocamento de Stokes	30
Figura 6 – Esquema do Microscópio de fluorescência	31
Figura 7 – Uma célula animal generalizada.	32
Figura 8 – O citoesqueleto em uma célula eucariótica.	34
Figura 9 – De esquerda para direita microfilamentos de actina(vermelho), filamentos intermediários(verde) e microtúbulos(azul).	34
Figura 10 – Exemplo da aquisição de imagem digital	35
Figura 11 – Exemplo de tipos de textura em imagens digitais	36
Figura 12 – K-Means, com $K=3$ e $K=4$	38
Figura 13 – PCA 2D.	39
Figura 14 – Esquema de construção dos dicionários visuais ou <i>codebooks</i>	42
Figura 15 – Exemplo do problema do detector de Harris em um canto escalado.	42
Figura 16 – Conjunto de imagens de espaço em escala mostradas à esquerda e as imagens da diferença de gaussianas à direita.	43
Figura 17 – Máximo e mínimo das imagens da diferença gaussiana são detectados comparando um pixel (marcado com X) com seus 26 vizinhos em regiões 3×3 nas escalas atual e adjacente (marcadas com círculos).	44
Figura 18 – Esta figura mostra uma matriz do descritor 2×2 (direita) calculada a partir de um conjunto de amostras 8×8 (esquerda).	45
Figura 19 – Demonstração do algoritmo SIFT para detecção de <i>keypoints</i> no banco de dados 2D Hela(esquerda) e FDIG Olympus(direita). Os círculos brancos são os <i>keypoints</i>	46
Figura 20 – Vizinhança de $R = \{1, 2, 3\}$ e $P = \{8, 12, 20\}$ respectivamente	46
Figura 21 – Vizinhança circular	47
Figura 22 – Demonstração do operador LBP sobre uma imagem 2D Hela(22a) na esquerda e sobre outra imagem do Olympus(22b) na direita.	47
Figura 23 – Aprendizado dos dicionários de textons	49
Figura 24 – Geração do Modelo	49

Figura 25 – Textons de tamanho 20x20 para 2D Hela(a,c) e Olympus(b,d) usando média e variância em tons de cinza.	50
Figura 26 – Espectro do espaço frequência gerada pela transformada de Fourier 2D . . .	51
Figura 27 – Exemplo de uso dos filtros de Gabor.	53
Figura 28 – Evolução da rede dinâmica como um <i>threshold</i> T_l e área de <i>zoom</i> : (a) $T_l = 0.1$; (b) $T_l = 0.15$ and (c) $T_l = 0.2$	56
Figura 29 – (a) Nervo na folha; (b)-(d) Rede complexa obtida com diferentes valores para <i>thresholds</i> t , $t = \{0.050, 0.075, 0.100\}$	56
Figura 30 – Esquema da metodologia: Linha azul claro seguida por <i>LBP</i> e linha laranja seguida por <i>BoVW</i> em conjunto com os algoritmos baseados nos textons . .	62
Figura 31 – Espectro do espaço frequência gerada pela transformada de fourier em 2D .	64
Figura 32 – Etapas do pré-processamento para a geração das redes complexas das imagens	64
Figura 33 – Exemplo do banco de imagens 2D Hela, Microfilamentos de Actina(linha superior) e Tubulina(linha inferior)	68
Figura 34 – Filamentos de actina (fila superior) e Microtubulos(fila de baixo) do banco de dados <i>FDIG Olympus</i> (Olympus America, Inc.©)	69
Figura 35 – Resumo dos cinco métodos (BoVW, LBP, TDFT, TGFB and TCN) aplicados ao banco de dados 2D Hela e <i>FDIG Olympus</i>	72

LISTA DE TABELAS

Tabela 1 – Exemplo do banco de imagens <i>2D Hela</i>	68
Tabela 2 – Detalhes do banco de imagens <i>FDIG Olympus</i>	68
Tabela 3 – Comparação das matrizes de confusão obtidas, banco de dados (as columnas <i>2D Hela</i> e <i>FDIG Olympus</i>) por método(linhas BoVW, LBP, TDFT, TGFB e TCN)	71
Tabela 4 – Comparação de Precisão, <i>Recall</i> e <i>F-Measure</i> de ambos bancos de dados(colunas) por métodos(linhas)	72

LISTA DE ABREVIATURAS E SIGLAS

ATP	<i>Adenosine triphosphate</i> ou Trifosfato de adenosina
BoVW	<i>Bag of Visual Words</i>
CBIR	<i>Content Based Image Retrieval</i>
CMU	<i>Carnegie Mellon University</i>
DoG	<i>Difference of Gaussians</i>
LoG	<i>Laplacian of Gaussians</i>
PCA	<i>Principal Component Analysis</i>
PCA	<i>Principal Components Analysis</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
SVD	<i>Singular value decomposition</i>
SVM	<i>Support vector machines</i>
TCN	<i>Textons based on Complex Network</i>
TDFT	<i>Textons based on Discrete Fourier Transform</i>
TGFB	<i>Textons based on Gabor Filter Bank</i>

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Objetivos e Contribuições	25
1.2	Organização do trabalho	25
2	FUNDAMENTAÇÃO TEÓRICA	27
2.1	Considerações Iniciais	27
2.2	Luz	27
2.3	Cor	28
2.4	Fluorescência	29
2.4.1	<i>Aplicações</i>	30
2.4.2	<i>Microscópio de fluorescência</i>	31
2.5	Citoesqueleto	32
2.6	Imagem digital	35
2.7	Textura	36
2.8	K-Means	37
2.9	PCA	38
2.10	Considerações finais	40
3	DESCRITORES DE CARACTERÍSTICAS	41
3.1	Considerações Iniciais	41
3.2	<i>Bag of Visual Words</i>	41
3.2.1	<i>Scale Invariant Feature Transform (SIFT)</i>	42
3.2.1.1	<i>Identificação dos keypoints</i>	43
3.2.1.2	<i>Localização dos Keypoints</i>	44
3.2.1.3	<i>Atribuição de orientação</i>	44
3.2.1.4	<i>Descritor do Keypoint</i>	45
3.3	Local Binary Pattern	46
3.4	Dicionário de textons	48
3.5	2-D Discrete Fourier Transform	50
3.6	Filtros de Gabor	52
3.7	Redes Complexas	54
3.7.1	<i>Representação de rede complexa para a forma</i>	55
3.7.2	<i>Métricas</i>	57

3.7.2.1	<i>Distribuição do Grau</i>	57
3.7.2.2	<i>Segundo Momento da Distribuição do Grau</i>	57
3.7.2.3	<i>Entropia da distribuição das arestas (H_E)</i>	57
3.7.2.4	<i>Entropia da distribuição de Grau (H_D)</i>	57
3.7.2.5	<i>Grau Médio</i>	58
3.7.2.6	<i>Coeficientes de Aglomeração local</i>	58
3.7.2.7	<i>Coeficiente de Aglomeração global</i>	58
3.7.2.8	<i>Média dos menores caminhos</i>	58
3.7.2.9	<i>Eficiência</i>	58
3.7.2.10	<i>Diâmetro</i>	59
3.8	Considerações finais	59
4	METODOLOGIA	61
4.1	<i>Pre-processing</i>	62
4.2	Feature Extractor	62
4.2.1	Bag of Visual Words	62
4.2.2	<i>Local Binary Pattern</i>	63
4.2.3	<i>Dicionários de Textons</i>	63
4.2.4	<i>Classifier</i>	65
5	EXPERIMENTAÇÃO E RESULTADOS	67
5.1	Banco de imagens fluorescentes	67
5.1.1	<i>2D Hela</i>	67
5.1.2	<i>FDIG Olympus</i>	68
5.2	Configuração de experimentos	69
5.3	Resultados	71
6	CONCLUSÃO	73
	REFERÊNCIAS	75

INTRODUÇÃO

O olho humano exige contraste para perceber detalhes de objetos. O método de contraste mais simples e eficaz é o chamado "campo escuro" ou *dark field*. Esse método aproveita a dispersão de luz em pequenas partículas que diferem do seu ambiente com um índice de refração. Este fenômeno é conhecido na física como efeito Tyndall ([ATKINS, 2012](#)). Nesse sentido a técnica de contraste mais popular atualmente é a fluorescência. A fluorescência requer o uso dos chamados *fluorochromes* ou *fluorophores*, que são corantes, substâncias que adicionam contraste, que absorvem a luz em um intervalo de comprimento de onda específico, e reemitem-a essa mesma porção de luz com menor energia, deslocando ela para um comprimento de onda mais longo. Atualmente, estão disponíveis um número muito grande de corantes com absorção da região ultravioleta (UV) para o infravermelho, e ainda estão sendo desenvolvidos mais *fluorochromes* com novas propriedades. As principais vantagens desta abordagem são um contraste, sensibilidade, especificidade e seletividade muito elevados ([KUBITSCHKE, 2013](#)).

O microscópio é um instrumento utilizado para ampliar e observar estruturas pequenas dificilmente visíveis ou invisíveis a olho nú. O microscópio ótico utiliza luz visível e um sistema de lentes de vidro que ampliam a imagem das amostras. A microscopia de fluorescência tornou-se uma ferramenta essencial na biologia, bem como na ciência dos materiais, pois possui atributos que não estão prontamente disponíveis em outras técnicas de microscopia óptica. É possível marcar de forma fluorescente a(s) proteína(s), e identificar o organelo onde a proteína reside. Isto fornece uma pista importante para a sua possível função. O citoesqueleto desempenha um papel importante em numerosos processos fisiológicos e patológicos, e suas características morfológicas são, portanto, de primordial importância para compreender numerosos fenômenos celulares básicos, como a adaptação celular ao estresse físico ou químico. No entanto, a quantificação e análise do citoesqueleto estão longe de serem diretas e são necessários algoritmos sofisticados para cumprir estas tarefas. A Figura 1 ilustra diferentes células com estruturas citoesqueléticas.

A identificação de organelas sub-celulares é importante, por exemplo para caracterizar

Figura 1 – Exemplo de imagens fluorescentes

(a) 2D *Hela*(b) *FDIG Olympus* (Olympus America, Inc.©)

Fonte: Elaborada pelo autor.

genes ou genes recentemente descobertos com uma função desconhecida. É possível marcar de forma fluorescente a(s) proteína(s), e identificar a organela onde a proteína reside. Isto fornece uma pista importante para a sua possível função. É importante observar que os especialistas humanos têm problemas para distinguir algumas organelas como por exemplo *Endosomes* e *Lysosomes*. Robert F. Murphy da *Carnegie Mellon University* (CMU), é pioneiro no campo de detecção e classificação de proteínas (BOLAND; MURPHY, 2001; COELHO *et al.*, 2013).

Dada a quantidade de organelos e a complexidade delas, surgiram pesquisas mais especializadas, por exemplo Breuer *et al.* (2014), focando-se nas estruturas do citoesqueleto, onde eles encontraram propriedades relevantes de filamentos de actina e microtúbulos para entender a organização do citoesqueleto, caracterizando estruturas como redes. Por outro lado, Lan *et al.* (2015) usa análise de textura e quatro métodos distintos: histogramas de nível de cinza (GLHs), uma matriz de coocorrência de nível de cinza (GLCM), Número de Euler (EN) e pacote wavelet para análise de energia (WPEA), combinados com *Principal Components Analysis* (PCA) para descobrir se o processo celular é normal ou patológico baseado na análise de microtúbulos.

A textura é uma das formas mais comuns de reconhecer imagens digitais que podem ser consideradas como propriedades de imagem de baixo nível. Embora a textura possa ser facilmente interpretada pelos seres humanos, criar sistemas automáticos para executar o mesmo papel é difícil e desafiador (GONZALEZ; WOODS, 2006).

Um passo fundamental na análise de textura é criar descritores adequados para representá-las. Alguns exemplos de descritores de texturas no domínio espacial são *Gray Level Co-occurrence Matrix* (GLCM), *Local Binary Pattern* (LBP), autocorrelação, *MRF*, entre outros. Por outro lado, existe outra estratégia como a modelagem de distribuição estatística, que é desenvolvida transformando as imagens com alguma das transformações, como *Discrete Fourier Transform* (DFT), *Discrete Cosine Transform*, Transformada Wavelet, Bancos dos filtros do Gabor e outros.

Além dos métodos baseados em textura descritos no parágrafo anterior, também podemos usar o *Bag of Visual Words* (BoVW), com base em dicionários para caracterizar imagens originalmente usadas para textos (CSURKA *et al.*, 2004). O BoVW usa um conjunto de descritores locais (por exemplo, SIFT (LOWE, 2004)) para criar um dicionário ou *codebook* por meio de

algum método de agrupamento (por exemplo, *k-Means*) e eventualmente produzir um histograma para cada imagem e assim poder classificá-las.

Outro grupo importante baseia-se em modelos de grafos, onde as imagens são caracterizadas como um grafo e, mais recentemente, baseadas em teoria de rede complexa para análise de formas (BACKES; CASANOVA; BRUNO, 2009), reconhecimento de padrões (CASANOVA; BACKES; BRUNO, 2013) e análise de textura (GONÇALVES *et al.*, 2016). Por exemplo, recentemente, (SCABINI; GONÇALVES; CASTRO, 2015) propuseram um método inovador usando o *Bag of Visual Words* e a teoria das redes complexas para análise de textura. As abordagens discutidas nos trabalhos anteriormente citados, criam redes a partir de imagens e, em seguida, aplicam o *BoVW*, em vez de extrair medidas estatísticas globais de redes complexas.

Pelo exposto anteriormente visa-se avaliar cinco métodos para descrever imagens com o objetivo de poder classificá-las, os quais são: *Bag of Visual Words*, *Local Binary Patterns*, *Discrete Fourier Transform*, Filtro de Gabor e Redes Complexas. Os detalhes da implementação, utilização, avaliação e o porquê foram considerados serão desenvolvidos durante os seguintes capítulos.

1.1 Objetivos e Contribuições

Neste trabalho tem como objetivo principal a rotulagem de uma imagem dada como entrada como microfilamento de Actina (MA) ou Microtúbulos (MT).

A principal contribuição do presente trabalho de mestrado é a análise de dois tipos de conjuntos de dados em diferentes ambientes controlados (*2D Hela*) e não controlado (*FDIG Olympus*). Para tal fim são avaliados cinco métodos dentro da área de processamento de imagens, visão computacional e teoria de redes complexas para poder descrever essas imagens. Essa abordagem ajuda a encontrar o método mais apropriado para descrever ambos conjuntos de dados, a qual poderá também ser aplicada para demais tipos de bases de imagens fluorescentes.

1.2 Organização do trabalho

O restante deste texto está dividido da seguinte forma: No [Capítulo 2](#) mostramos um resumo dos conceitos essenciais para o melhor entendimento desta dissertação. No [Capítulo 3](#) apresentamos uma revisão do estado da arte em relação a cada um dos métodos considerados. A metodologia seguida no [Capítulo 4](#). No [Capítulo 5](#) é descrito em detalhe como foram executados os experimentos, incluindo a nova proposta baseada em redes complexas. Por fim no [Capítulo 6](#), apresentamos as conclusões e trabalhos futuros.

FUNDAMENTAÇÃO TEÓRICA

2.1 Considerações Iniciais

Nesse capítulo são apresentados conceitos fundamentais como luz e cor, de uma perspectiva ligada à fluorescência que faz possível observar ou obter o citoesqueleto, sendo essa estrutura celular a mais importante do presente trabalho de mestrado. Dessa forma, serão abordados esses conceitos com foco biológico e em relação a computação imagem digital, textura, *k-means* e *PCA*.

2.2 Luz

Do latim *lux*, a luz tal como o som é um fenômeno de natureza ondulatória. A luz é uma radiação eletromagnética, que se propaga através de diferentes meios materiais, como o ar ou a água e também se propaga através do vazio. Existem alguns tipos de sólidos (chamados de opacos) que não se deixam atravessar pela luz.

Ao longo dos anos, muitos cientistas procuraram respostas para esta questão. Surgiram duas teorias, suportadas por diferentes experiências e diferentes cientistas que são atualmente válidas, uma vez que a luz tanto se pode comportar como uma onda em determinadas ocasiões como também se pode comportar como uma partícula noutras. Essas duas teorias são:

- Teoria corpuscular da luz - Considera que a luz é constituída por pequenas partículas - chamadas fótons - de características muito especiais.
- Teoria ondulatória da luz - Considera que a luz é uma manifestação de energia, constituída por ondas semelhantes às do som, mas com comprimentos de onda muitíssimo menores do que as características das ondas sonoras.

No caso específico da luz, uma amplitude se identifica com o brilho e uma frequência com uma cor.

Figura 2 – Diagrama da dispersão da luz através de um prisma

Fonte: [Gonzalez e Woods \(2006\)](#)

Um raio de luz é uma trajetória da luz em determinado espaço e sua representação indica de onde uma luz é criada (fonte) e para onde ela se dirige. Propagando-se em meio homogêneo, a luz percorre trajetórias retilíneas; só em meios não-homogêneos a luz pode descrever trajetórias curvas. A Figura 2 mostra a representação da dispersão da luz ou também chamada dispersão da luz branca.

2.3 Cor

A cor é a característica da percepção visual humana descrita através de categorias de cores, com nomes como vermelho, amarelo, roxo ou azul. Esta percepção de cor deriva da estimulação de células cônicas no olho humano por radiação eletromagnética no espectro de luz. As categorias de cores e as especificações físicas da cor estão associadas a objetos através do comprimento de onda da luz que é refletida a partir deles. Esta reflexão é governada pelas propriedades físicas do objeto, tais como absorção de luz, espectros de emissão, entre outras ([WYSZECKI; STILES, 2000](#); [HUNT, 2005](#)).

Ao definir um espaço de cores, as cores podem ser identificadas numericamente por coordenadas. O espaço de cores RGB, por exemplo, é um espaço de cores correspondente à tricromia humana e aos três tipos de células cônicas que correspondem a três bandas de luz: longos comprimentos de onda, atingindo um pico próximo de 564-580 nm (vermelho); comprimento de onda médio, atingindo um pico próximo de 534-545 nm (verde); e luz de comprimento de onda curto, perto de 420-440 nm (azul), como mostra-se na Figura 3.

Figura 3 – Espectro Contínuo

Fonte: (GONZALEZ; WOODS, 2006)

A Figura 4 é uma representação diagramática da região da luz visível da radiação eletromagnética, que cobre uma gama de comprimento de onda de aproximadamente 400 a 800 nanômetros. Em torno da região visível se concentra a maior energia.

Figura 4 – Espectro Contínuo visível

Fonte: Adaptação do (GONZALEZ; WOODS, 2006)

2.4 Fluorescência

Quando as seres vivos ou não vivos, orgânicos ou inorgânicos, absorvem e refletem luz, o processo é descrito como fotoluminescência. Se a emissão de luz persiste por até alguns segundos após a energia de excitação (luz) ser interrompida, o fenômeno é conhecido como fosforescência. Se a emissão de luz que continua apenas durante a absorção da luz de excitação, o fenômeno é conhecido como fluorescência. O intervalo de tempo entre a absorção de luz de excitação e a emissão de luz re-irradiada em fluorescência é de duração extraordinariamente curta, geralmente inferior a um milionésimo de segundo. Adicionalmente a fluorescência é a capacidade de uma substância de emitir luz quando e exposta a radiações do tipo ultravioleta (UV), raios catódicos ou raios X. Essas radiações absorvidas (invisíveis ao olho humano) transformam-se em luz visível (OLYMPUS, 2017).

Figura 5 – Deslocamento de Stokes

Fonte: Adaptação do (OLYMPUS, 2017)

Na Figura 5, um fóton de radiação ultravioleta (roxo) colide com um elétron em um átomo simples, excitando e elevando o elétron para um nível de energia mais elevado. Posteriormente, o elétron excitado relaxa a um nível mais baixo e emite luz na forma de um fóton de baixa energia (vermelho) na região da luz visível. O fenômeno da fluorescência era conhecido em meados do século XIX. O cientista britânico Sir George G. Stokes fez a primeira observação de que o mineral espato fluor (fluorina) exibe fluorescência quando é iluminado com luz ultravioleta, e ele cunhou a palavra "fluorescência". Stokes observou que a luz fluorescente tem comprimentos de onda maiores do que a luz de excitação, um fenômeno que se tornou conhecido como o deslocamento de Stokes.

A microscopia de fluorescência é um excelente método de estudo de material que pode ser feito para fluorescer, seja na sua forma natural (denominada fluorescência primária ou autofluorescência) ou tratados com produtos químicos capazes de fluorescência (fluorescência secundária). A microscopia de fluorescência é agora uma ferramenta importante e até indispensável na biologia celular.

2.4.1 Aplicações

Um exemplo, é o fenômeno que faz com que certos materiais brilhem à exposição de UV emitida por uma lâmpada "luz negra". A aplicação mais habitual deste fenômeno são as lâmpadas fluorescentes, onde uma substância branca que recobre o seu interior de cristal emite luz quando se cria uma corrente elétrica no interior do tubo. Outro uso prático da fluorescência é de detectar bilhetes falsos, já que só os verdadeiros levam impressos uma tinta fluorescente que é visível apenas com auxílio de uma "luz negra".

2.4.2 Microscópio de fluorescência

As primeiras pesquisas que deram origem à invenção do microscópio de fluorescência atribuem-se ao cientista alemão August Köhler. Köhler realizou experiências com radiação ultravioleta que, devido ao seu comprimento de onda mais curto, poderiam contribuir para o incremento do poder de resolução, face ao que já era observado na microscopia óptica. Seria a partir destas pesquisas que surgiria o microscópio de ultravioleta (UV) (entre 1901 e 1904). Em 1904, descobriu-se que o tecido demonstra fluorescência quando irradiado por luz ultravioleta, sendo este fenômeno conhecido como autofluorescência. Esta propriedade foi pesquisada na investigação de bactérias, protozoários, plantas, tecidos animais e em várias substâncias bi-orgânicas como, por exemplo, a albumina, a elastina e a queratina.

Em 1914, Stanislav Von Prowazek utilizou o microscópio de fluorescência para estudar a ligação dos corantes às células vivas, chegando à conclusão que seções de tecido corados poderiam ser examinados recorrendo ao microscópio de fluorescência, caso os corantes utilizados fossem fluorescentes. Estes corantes foram denominados fluorocromos, de modo a distingui-los dos diacromos (tintas coloridas que se tornam visíveis por absorção de luz) (SPRING, 2003).

A Figura 6 mostra o esquema geral do funcionamento do microscópio de fluorescência, para a obtenção de imagens digitais fluorescentes.

Figura 6 – Esquema do Microscópio de fluorescência

Fonte: Adaptação do (OLYMPUS, 2017)

2.5 Citoesqueleto

A célula é a menor parte de um ser vivo capaz de desenvolver-se e reproduzir. De acordo com a organização estrutural, as células são divididas em: eucarióticas e procarióticas. As células eucarióticas são consideravelmente mais complexas. Os organismos eucarióticos são atualmente classificados em quatro reinos: animais, plantas, fungos e protistas. A célula animal na Figura 7 mostra características comuns às células de três reinos. A maior diferença entre a célula animal (ou qualquer célula eucariótica) das células procariotas, é que as células procarióticas tem a presença de compartimentos internamente ligados à membrana ou organelas. O mais proeminente é o núcleo, que abriga o DNA. Tradicionalmente, assumiu-se que a maioria dos genomas eucariotas pode variar de 10 a 100×10^6 nucleótidos (10-100 Mb) de comprimento total.

Figura 7 – Uma célula animal generalizada.

Fonte: Copyright © 2006 Pearson Education, Inc., publishing as Benjamin Cummings

O núcleo é delimitado por uma membrana de duas camadas (a maioria das outras organelas são delimitadas por uma única membrana) que é contínua com o retículo endoplasmático (ER). O retículo endoplasmático é subdividido em rugoso (RER) e suave (SER) com base na aparência em micrografias eletrônicas. Os "parafusos" no RER são ribossomos, que são a maquinaria molecular para fazer proteínas na célula.

O complexo de Golgi, ou os corpos de Golgi, embora fisicamente independentes, são uma extensão funcional ao processamento e triagem de proteínas que ocorre no ER. As proteínas deixam o Golgi em vesículas ligadas à membrana celular ou a outras organelas. As vesículas, embora ligadas à membrana, geralmente não são contadas como organelas: são simplesmente pequenos pacotes de transporte.

As mitocôndrias são organelas complexas que não são apenas delimitadas por uma membrana, mas também contêm uma segunda membrana altamente crenulada. As mitocôndrias tornam possível a respiração aeróbica, usando oxigênio como oxidante para produzir energia química (ou seja, ATP) de forma muito mais eficiente do que os processos anaeróbicos utilizados pela maioria dos procariotas. Essa capacidade de produzir mais energia a partir da mesma quantidade de alimento permite que células eucarióticas cresçam mais do que as procariotas.

Os lisossomos são ácidos e contêm enzimas digestivas que quebram grandes moléculas de alimentos particularmente proteínas e gorduras, para torná-las utilizáveis pelo resto da célula. Essas enzimas funcionam otimamente em condições ácidas, o que atua como uma espécie de segurança. Se um lisossoma quebra e libera suas enzimas no citoplasma, eles não quebrarão os componentes celulares, porque o *pH* citoplasmático é próximo ao neutro e as enzimas não funcionam bem. Uma vez que se pensava que era exclusivo das células animais, os lisossomos já foram descritos em todas as células de todos os reinos eucarióticos.

Os peroxissomos são organelas membranosas presentes no citoplasma das células vegetais e animais, formando vesículas arredondadas, cuja função está relacionada ao armazenamento de enzimas que catalisam o peróxido de hidrogênio (água oxigenada - H_2O_2), uma substância tóxica que necessita ser degradada.

As células vegetais possuem todas as organelas citadas acima, mas também podem ter outros dois tipos de organelas: cloroplastos e vacúolos. Além disso, as células da planta também possuem uma parede celular rígida externa à membrana celular. Os cloroplastos são semelhantes às mitocôndrias em forma e estrutura (membranas dentro da membrana externa). No entanto, sua função é muito diferente: os cloroplastos absorvem energia de luz do sol (ou outras fontes de luz) e convertem-na em energia química sob a forma de açúcares simples para que a célula armazene e use mais tarde, em essência, transformando a célula em um armazém de energia e centro de distribuição. Como as plantas e algumas bactérias fotossintéticas são os únicos organismos capazes de converter a energia solar em uma forma útil às células vivas, elas são cruciais para a sobrevivência.

Finalmente, tem-se o citoplasma. Uma vez considerado apenas o ambiente aquoso em que as moléculas ou organelas importantes flutuavam, ele é composto por importantes elementos estruturais e de transporte (Figura 8). O citoesqueleto fornece não apenas uma estrutura física interna, mas também um sistema de transporte para mover moléculas, vesículas e até organelas para onde elas são necessárias. O citoplasma é preenchido por uma matéria coloidal e semi-fluida denominada citosol, na qual estão suspensos os organelos celulares. Nos eucariontes, em oposição ao protoplasma, o citoplasma não inclui o núcleo celular, cujo interior é formado por nucleoplasma. No geral, o citoplasma é tudo o que compreende a célula menos o núcleo e a membrana plasmática. Filamentos de actina, presente no citoplasma das células eucariotas sob a forma de feixes de filamentos paralelos ou redes de filamentos anastomosados com 5-7 nm de diâmetro, constituídos pela polimerização da proteína globular actina G, originando os filamentos

Figura 8 – O citoesqueleto em uma célula eucariótica.

Fonte: Adaptação do Purves *et al.* (2004)

de actina F. O modelo mais aceitável é o de um filamento helicoidal formado por uma cadeia simples de monômeros. Estes monômeros são constituídos por uma cadeia polipeptídica de 374 (375 no músculo esquelético) unidades. A actina é o maior componente dos filamentos finos das células musculares e do sistema citoesquelético de células não musculares, e está presente em todos os eucariontes.

Por outro lado, os microtúbulos são estruturas proteicas que fazem parte do citoesqueleto nas células. São filamentos com diâmetro de, aproximadamente, 24 nm e comprimentos variados, de vários micrômetros até alguns milímetros nos axônios das células nervosas. Microtúbulos são formados pela polimerização das proteínas tubulina e albetralopina. As extremidades de um microtúbulo são designadas como (+) (a que se polimeriza mais rapidamente) e (–) (a que se polimeriza mais vagarosamente).

Figura 9 – De esquerda para direita microfilamentos de actina(vermelho), filamentos intermediários(verde) e microtúbulos(azul).

Fonte: Adaptação do Purves *et al.* (2004)

A Figura 9 apresenta imagens digitais das estruturas mais importantes do citoesqueleto,

obtidas por meio da câmera de um microscópio de fluorescência.

Os microtúbulos são pequenas estruturas cilíndricas e ocas formadas por proteínas chamadas Tubulinas. Existem 2 tipos de Tubulinas que se associam formando dímeros, a α (alfa) e a β (beta); estes se polimerizam formando protofilamentos. São necessários 13 protofilamentos para se formar o microtúbulo. Os microtúbulos são polimerizados a partir de um centríolo que geralmente fica no centro do citoplasma celular. Além da função estrutural, os microtúbulos têm outras funções. Eles formam um substrato onde proteínas motoras celulares (Dineínas e Cinesinas) podem interagir e, assim, são usados no transporte intracelular. Os microtúbulos também fazem parte dos flagelos e cílios das células eucarióticas (LODISH *et al.*, 2008).

2.6 Imagem digital

A visão é o mais avançado dos nossos sentidos, por isso as imagens desempenham o papel mais importante na percepção humana. No entanto, ao contrário dos seres humanos, que estão limitados à banda visual do espectro eletromagnético, máquinas de imagem cobrem quase todo o espectro eletromagnético, variando de gama a ondas de rádio. Elas podem operar em imagens geradas por fontes que os seres humanos não estão acostumados a associar com imagens. Estes incluem ultra-som, microscopia eletrônica e imagens geradas por computador. Assim, o processamento de imagem digital abrange um vasto e variado campo de aplicações. A Figura 10 mostra um exemplo do processo de aquisição da imagem digital.

Figura 10 – Exemplo da aquisição de imagem digital

Fonte: Adaptação do (GONZALEZ; WOODS, 2006)

Uma imagem pode ser definida como uma função bidimensional $f(x,y)$, onde x e y são coordenadas espaciais (plano), e a amplitude de f em qualquer par de coordenadas (x,y) é chamado de intensidade ou nível de cinza da imagem nesse ponto. Quando x,y , e os valores de intensidade de f são todos finitos, quantidades discretas, chamamos a imagem uma imagem digital. Note que uma imagem digital é composta por um número finito de elementos, cada um dos quais tem uma localização específica e valor. Esses elementos são chamados *picture elements*, *image elements*, *pels*, e *pixels*. *Pixel* é o termo mais amplamente usado para denotar os elementos de uma imagem digital. É possível ampliar essas definições em termos mais formais no livro de (GONZALEZ; WOODS, 2006).

2.7 Textura

Não existe uma definição clara de textura, mas os autores concordam em defini-la como as mudanças na intensidade da imagem que formam determinados padrões repetitivos (CHEN, 2010). Esses padrões podem ser o resultado de propriedades físicas da superfície do objeto (rugosidade), ou ser o resultado de diferenças de reflexão tal como a cor na superfície. A Figura 11 mostra alguns exemplos de imagens com textura. Embora seja fácil, para as pessoas, o reconhecimento de textura, isto não acontece com procedimentos automáticos onde esta tarefa, às vezes, precisa de técnicas computacionais complexas. Nos sistemas de recuperação de imagens baseada em conteúdo (do inglês *content-based image retrieval - CBIR*) as características de textura são muito úteis, pelo fato que a grande maioria de superfícies naturais exibem texturas.

Figura 11 – Exemplo de tipos de textura em imagens digitais

Fonte: Adaptação do (LIN *et al.*, 2006)

2.8 K-Means

O método *k-means* é uma técnica de clustering amplamente utilizada que busca minimizar a distância média quadrática entre pontos no mesmo cluster. Embora não ofereça garantias de precisão, sua simplicidade e velocidade são muito atraentes na prática. O objetivo é agrupar n observações em k grupos distintos onde cada observação pertence ao grupo mais próximo da média, de forma a minimizar o erro quadrático médio entre o posicionamento encontrado e o posicionamento ótimo (BOTTOU; BENGIO, 1995).

Este algoritmo é uma simplificação do algoritmo *Expectation Maximization(EM)* para estimativa de parâmetros para uma mistura de gaussianas. Neste caso, deseja-se estimar as k médias dos grupos que tem maior probabilidade de gerar os dados observados. A hipótese é que os dados são originários de distribuições gaussianas com matrizes de covariância proporcionais à identidade. O que equivale a dizer que os dados são isotrópicos, ou circularmente simétricos. Sob essa hipótese, o posicionamento ótimo do centro de um grupo corresponde ao seu centróide (KANUNGO *et al.*, 2002).

Inicialmente, o algoritmo seleciona k pontos, $\mu_1, \mu_2, \dots, \mu_k$, aleatoriamente. A partir disso, para cada $x_j \in \mathfrak{R}^n$ pertencente ao conjunto de dados, deseja-se escolher μ_{i*} tal que

$$\|x_j - \mu_{i*}\| < \|x_j - \mu_i\|, \quad i = 1, \dots, k \quad (2.1)$$

o que é equivalente a dizer que

$$\sum_{m=1}^n (x_{j_m} - \mu_{i_m^*})^2 < \sum_{m=1}^n (x_{j_m} - \mu_{i_m})^2, \quad i = 1, \dots, k, \quad (2.2)$$

Ou seja, deseja-se minimizar o erro quadrático entre o centro encontrado, e o centro verdadeiro da distribuição. Esta atribuição corresponde à etapa do algoritmo *Expectation-Maximization*. A partir desta atribuição, para cada um dos k grupos, calcula-se

$$\mu_i = \frac{1}{|X_i|} \sum_{x_j \in X_i} x_j \quad (2.3)$$

onde x_i é o conjunto formado por todos os pontos atribuídos ao grupo i . Dado um conjunto de inicialização, o algoritmo garantidamente converge para o erro mínimo. Porém, não há garantias de que este mínimo seja o mínimo global (BOTTOU; BENGIO, 1995; KANUNGO *et al.*, 2002). Dessa forma, a escolha dos k centros tem grande importância na qualidade do resultado final obtido.

Na Figura 12 é possível observar 4 exemplos de execução do *k-means* para n pontos bidimensionais com 3(12a, 12b) e 4(12c, 12d) grupos. Foram geradas com inicializações aleatórias(ponto pretos) e a principal observação sobre elas é que se o parâmetro k for maior aos grupos reais, os resultados não sempre serão iguais, daí a importância desse parâmetro.

Figura 12 – K-Means, com K=3 e K=4.

Fonte: Elaborada pelo autor.

A mais recente melhoria do algoritmo *k-means* é *kmeans++* onde se faz a seleção de forma probabilística dos próximos pontos durante o processo de reajuste dos centroides, conseguindo um custo de até $O(\log k)$ (ARTHUR; VASSILVITSKII, 2007).

2.9 PCA

A análise dos componentes principais (do inglês *Principal Component Analysis - PCA*) é uma técnica que tem como objetivo a análise dos dados usados visando sua redução, eliminação de sobreposições e a escolha das formas mais representativas de dados a partir de combinações lineares das variáveis originais. O *PCA* tem sido aplicado a uma ampla classe de problemas de visão computacional, incluindo seleção de características, reconhecimento de padrões, reconhecimento de objetos e reconhecimento de faces (JOLLIFFE, 1986).

O *PCA* funciona com base na premissa de que uma base dimensional baixa é suficiente para aproximar a matriz de covariância das amostras, proporcionando assim uma representação compacta. Dadas as M imagens observadas, o *PCA* diagonaliza a matriz de covariância resolvendo a seguinte equação de autovalor:

$$C = \frac{1}{M} \sum_{j=1}^M y_j y_j^T \quad (2.4)$$

onde y_j pode ser considerado um *patch* de imagem vetorizada, e sem perda de generalidade, assume-se que ele é pré-processado para ter como média zero (BELHUMEUR;

HESPANHA; KRIEGMAN, 1997). A transformada de *Karhunen-Loeve* (KL) é um método eficiente para calcular a base (componentes principais), que pode ser realizada usando decomposição de valor singular *SVD* (BISHOP, 1995). Os passos para calcular as componentes principais são:

- Obter os dados ou as M amostras de vetores de dimensão n ;
- Calcular a média ou o vetor médio destes dados;
- Subtrair a média de todos os itens de dados;
- Calcular a matriz de covariância usando todas as subtrações. Ela é o resultado da média do produto de cada subtração por ela mesma e terá dimensão $n \times n$.
- Calcular os autovalores e autovetores da matriz de covariância.
- Arranjar a matriz da transformada de Hotellings, cujas linhas são formadas a partir dos autovetores da matriz de covariância arranjados de modo que a primeira linha, o elemento $(0,0)$, seja o auto vetor correspondente ao maior autovalor, e assim sucessivamente até que a última linha corresponda ao menor autovalor.

Figura 13 – PCA 2D.

Fonte: Adaptação do (JOLLIFFE, 1986)

O autovetor com o maior auto valor associado corresponde à componente principal do conjunto de dados usado. Isso significa que esse é o relacionamento mais significativo entre as dimensões dos dados.

2.10 Considerações finais

Neste Capítulo foram apresentados conceitos importantes que ajudam ao leitor a conhecer termos relacionados ao processo de obtenção das imagens de fluorescência(luz , cor, fluorescência e microscópio de fluorescência), e todos os conceitos relacionados com as imagens como: textura, algoritmos *K-means*, e *PCA*.

DESCRITORES DE CARACTERÍSTICAS

3.1 Considerações Iniciais

A detecção e descrição das características de imagem desempenham um papel vital em vários domínios de aplicação, como processamento de imagem, visão computacional, reconhecimento de padrões e aprendizado de máquina. Existem dois tipos de características que podem ser extraídas de uma imagem; características globais e locais. As características globais descrevem a imagem como um todo e podem ser interpretadas como uma propriedade particular da imagem envolvendo todos os pixels; enquanto que os recursos locais visam detectar pontos-chave ou *keypoints* dentro da imagem e descrevem as regiões ao redor desses pontos-chave. Depois de extrair as características e seus descritores de imagens, a correspondência de estruturas comuns entre imagens (ou seja, combinações de características) é o próximo passo para essas aplicações.

Nesse capítulo são apresentados os cinco métodos selecionados para análises das imagens, *Bag of Visual Words*, *Local Binary Patterns*, *Discrete Fourier Transform*, filtros de Gabor e a teoria de redes complexas, considerados neste trabalho.

3.2 *Bag of Visual Words*

Os dicionários visuais ou *Bag of Visual Words*(BoVW) é uma das técnicas mais conhecidas para representar imagens com o objetivo de classificá-las(CSURKA *et al.*, 2004). O BoVW consiste em extrair um conjunto de descritores locais, como por exemplo o descritor *SIFT* (LOWE, 2004). Tendo uma imagem e atribuindo cada descritor à entrada mais próxima em um vocabulário visual, esse vocabulário ou *codebook* é criado *off-line* por agrupamento (por exemplo *k-means*), como é mostrado na Figura 14. Em seguida, o vocabulário é utilizado para construir um histograma para cada imagem.

Figura 14 – Esquema de construção dos dicionários visuais ou *codebooks*

Fonte: Adaptação do [Csurka et al. \(2004\)](#)

3.2.1 Scale Invariant Feature Transform (SIFT)

O *SIFT* é conhecido como descritor local, mas é composto por um detector de *keypoints* em combinação do descritor propriamente dito. Um detector muito conhecido é o detector de cantos de Harris ([HARRIS; STEPHENS, 1988](#)), ele é invariante à rotação, o que significa que, mesmo que a imagem seja girada, podemos encontrar os mesmos cantos, o que é óbvio porque os cantos permanecem os mesmos na imagem girada também. Mas isso não acontece na escala, pois um canto pode não ser o mesmo se a imagem for redimensionada. Por exemplo, verifique a [Figura 15](#). Um canto em uma imagem pequena dentro de uma pequena janela é plano quando este é ampliado na mesma janela. Então, detectores como Harris não são invariantes em escala.

Figura 15 – Exemplo do problema do detector de Harris em um canto escalado.

Fonte: [Lowe \(2004\)](#)

Existem principalmente quatro etapas envolvidas no algoritmo *SIFT*, as quais serão detalhadas a seguir.

3.2.1.1 Identificação dos keypoints

Da [Figura 15](#), infere-se que não se pode usar a mesma janela para detectar *keypoints* com diferentes escalas. Ele funciona bem com um pequeno canto, mas para detectar cantos maiores precisa-se de janelas maiores. Para isso, é utilizada a filtragem de espaço em escala. Nela, o *Laplacian of Gaussian (LoG)* é gerado com vários valores σ . O *LoG* atua como um detector de *blobs* que detecta bolhas em vários tamanhos devido à mudança do σ . Em suma, o kernel gaussiano σ atua como um parâmetro de escala. Por exemplo, na [Figura 15](#), o kernel gaussiano com σ baixo dá alto valor para um canto pequeno, enquanto o kernel gaussiano com alto σ se encaixa bem para um canto maior. Assim, podemos encontrar o máximo local em toda o espaço escala que nos dá uma lista de valores (x, y, σ) , o que significa que existe um *keypoint* potencial em (x, y) na escala σ . Mas o *LoG* é muito caro em termos de processamento, por isso o algoritmo *SIFT* usa diferença de gaussianas para fazer uma aproximação do *LoG*. A diferença de gaussianas ou *DoG* é obtido como a diferença de gaussianas suaves de uma imagem com dois σ diferentes, que seja σ e $k\sigma$. Este processo é feito para diferentes oitavas da imagem em uma pirâmide gaussiana, como visto na [Figura 16](#).

Figura 16 – Conjunto de imagens de espaço em escala mostradas à esquerda e as imagens da diferença de gaussianas à direita.

Fonte: [Lowe \(2004\)](#)

Uma vez que o *DoG* é calculado, as imagens são pesquisadas por extremo local sobre escala e espaço. Por exemplo, um pixel de uma imagem é comparado com seus oito vizinhos,

bem como 9 pixels na próxima escala e 9 pixels em escalas anteriores. Se é um lugar extremo, é um ponto-chave em potencial. Isso significa que, basicamente, o *keypoint*, é melhor representado na escala, como é mostrado na Figura 17. Em relação aos parâmetros, o artigo original baseia-se em dados empíricos, sugerindo o número de *oitavas* = 4, o número de níveis de escala = 5, σ inicial = 1.6, e $k = \sqrt{2}$ como valores ótimos.

Figura 17 – Máximo e mínimo das imagens da diferença gaussiana são detectados comparando um pixel (marcado com X) com seus 26 vizinhos em regiões 3x3 nas escalas atual e adjacente (marcadas com círculos).

Fonte: Lowe (2004)

3.2.1.2 Localização dos Keypoints

Depois que os pontos-chave locais são encontrados, eles tem que ser refinados para obter resultados mais precisos. No artigo original (LOWE, 2004), os autores usaram a expansão da série de Taylor do espaço escala para obter a localização mais precisa do extremo, e se a intensidade neste extremo é inferior a um valor limiar (0.03 por exemplo), ela é rejeitada. Este limiar é denominado *contrast threshold* neste trabalho. O *DoG* tem maior resposta para as bordas, de modo que as bordas precisam ser removidas também. Para isso, um conceito semelhante ao detector de cantos de Harris (HARRIS; STEPHENS, 1988) é usado. Utiliza-se uma matriz hessiana (H) de 2x2 para calcular a curvatura principal. Sabemos por detector de cantos de Harris que para bordas, um autovalor é maior que o outro. Então eles usaram uma função simples, se esta relação é maior do que um limiar, então isso é descartado. No presente trabalho de mestrado isso é chamado *edge threshold*. Assim, esse método elimina quaisquer *keypoints* de baixo contraste e bordas-chave e o que resta são os *keypoints* de interesse final.

3.2.1.3 Atribuição de orientação

Agora uma orientação para cada ponto-chave é atribuída para alcançar a invariância na rotação da imagem. A vizinhança ao redor do *keypoint* local é levada em consideração dependendo da escala, e a magnitude e a direção do gradiente é calculada nessa região. Um

histograma de orientação com 36 bins abrangendo 360 graus é criado. Ele é ponderado pela magnitude do gradiente e gaussiana ponderada de janela circular com $\sigma = 1.5$ vezes a dimensão do ponto chave. O pico mais alto no histograma é tomado e qualquer pico acima de 80% também é considerado para calcular a orientação. Ele cria pontos chave com a mesma localização e escala, mas diferentes direções, contribuindo para a estabilidade da correspondência.

3.2.1.4 Descritor do Keypoint

É tomado uma vizinhança de 16x16 ao redor do ponto chave. Ele é dividido em 16 sub-blocos de tamanho 4x4. Para cada sub-bloco, um histograma de orientação de 8-bins é criado. Assim, um total de 128 valores bin estão disponíveis. Ele é representado como um vetor para formar o descritor do *keypoint*. Em adição a isso, várias medidas são tomadas para conseguir robustez contra mudanças de iluminação, rotação, entre outras. A correspondência de *keypoints* entre duas imagens são acompanhadas por identificação de seus vizinhos mais próximos. Mas em alguns casos, o segundo mais próximo do grupo pode ser muito próximo ao primeiro. Isso pode acontecer devido a ruído ou algumas outras razões. Nesse caso, a relação de distância mais próxima é tomada. Se for maior do que 0.8, eles são rejeitados. Esse processo elimina cerca de 90% de falsos positivos, enquanto apenas 5% descarta combinações corretas.

Um descritor de *keypoint* é criado primeiro; computando a magnitude e a orientação do gradiente em cada ponto de amostra de imagem em uma região em torno da localização do *keypoint*, como se mostra à esquerda na [Figura 18](#). Estes são ponderados por uma janela gaussiana, indicada pelo círculo superposto. Essas amostras são então acumuladas em histogramas de orientação que resumem os conteúdos em sub-regiões 4x4, como mostrado à direita da [Figura 18](#), com o comprimento de cada seta correspondente à soma das magnitudes de gradiente perto dessa direção dentro da região.

Figura 18 – Esta figura mostra uma matriz do descritor 2x2(direita) calculada a partir de um conjunto de amostras 8x8(esquerda).

Fonte: [Lowe \(2004\)](#)

Figura 19 – Demonstração do algoritmo SIFT para detecção de *keypoints* no banco de dados 2D Hela(esquerda) e FDIG Olympus(direita). Os círculos brancos são os *keypoints*.

Fonte: Elaborada pelo autor usando OpenCV (ITSEEZ, 2015)

3.3 Local Binary Pattern

O operador *Local Binary Patterns (LBP)* é usado como um tipo de descritor visual para classificar textura em imagens, sendo uma das técnicas mais usadas devido a sua simplicidade e bom desempenho (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002). Foram criadas diversas variantes e melhorias do LBP (SILVA; BOUWMANS; FRÉLICOT, 2015). Esta Seção apresenta a versão *LBP* original, onde se considera os padrões uniformes e invariantes a mudanças na intensidade dos *pixels*(escala e rotação).

Figura 20 – Vizinhança de $R = \{1, 2, 3\}$ e $P = \{8, 12, 20\}$ respectivamente

Fonte: Elaborada pelo autor.

Na formulação do *LBP*, uma textura $LBP_{P,R}$ para um pixel g_c é a distribuição conjunta de g_c e uma vizinhança de tamanho P , dada pela equação abaixo:

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c) 2^i \quad (3.1)$$

Onde g_c é o valor cinza do pixel central e g_i é o valor cinza de cada pixel vizinho, e s é uma função limiar ou *thresholding* definida como:

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

De (3.1), é fácil mostrar que o número de termos binários a serem somados é $\sum_{i=0}^{P-1} 2^i = 2^P - 1$, de modo que o comprimento do histograma resultante (incluindo a posição do bin-0) é de 2^P .

Além disso, o arranjo espacial desta vizinhança é geralmente circular ou quadrada. Uma vizinhança circular é definida por uma amostragem de pontos P , igualmente espaçados em um círculo de raio R , que é centrada em g_c . Um exemplo dessa vizinhança é ilustrado na Figura 21. Os pontos não localizados no centro de um *pixel* precisam ter seus valores interpolados (por exemplo, por uma interpolação bi-linear). Uma vizinhança quadrada é obtida de forma semelhante, onde os vizinhos P são igualmente espaçados ao longo de um quadrado de lado R .

Figura 21 – Vizinhança circular

Fonte: Silva, Bouwmans e Frélicot (2015)

Figura 22 – Demonstração do operador LBP sobre uma imagem 2D Hela(22a) na esquerda e sobre outra imagem do Olympus(22b) na direita.

Fonte: Elaborada pelo autor.

3.4 Dicionário de textons

Texton foi definido como a unidade básica da percepção visual humana pré-ativo, aderido por Julesz (1981), e referem-se a micro-estruturas fundamentais em imagens naturais. No trabalho Zhu *et al.* (2005), foi feita uma extensão do modelo, onde é criado um modelo de imagem generativa de três níveis para aprender textons de imagens de textura. Eles estudaram as estruturas geométricas, dinâmicas e fotométricas da representação de textons:

- Para as estruturas geométricas, um texton consiste em várias bases de imagens com configurações espaciais deformáveis. As estruturas geométricas são aprendidas a partir de imagens de textura estática.
- Para as estruturas dinâmicas, o movimento de um texton é caracterizado por um modelo de cadeias de Markov no tempo, que às vezes pode alternar configurações geométricas durante o movimento. Chamamos os textons em movimento como “ motons ”. Os modelos dinâmicos são aprendidos usando as trajetórias dos textos inferidos a partir da sequência de vídeos.
- Para estruturas fotométricas, um texton representa o conjunto de imagens de um elemento de superfície 3D sob iluminação variável e é chamado de “ lighton ”.

No presente trabalho de mestrado consideramos apenas estruturas geométricas, o que significa o uso de um pequeno número de bases de imagem com configurações espaciais deformáveis. As estruturas geométricas são aprendidas a partir de uma imagem de textura estática com elementos repetidos.

A forma dos textons inicialmente não é conhecida, e eles são aprendidos através de respostas a um conjunto de filtros lineares, e as respostas resultantes são agrupadas. Os centros dos *clusters* são então selecionados como textons. A abordagem tem sido utilizada com sucesso em alguns campos de pesquisa como classificação de texturas (VARMA; ZISSERMAN, 2005; JAVED; KHAN, 2011; XIE *et al.*, 2015), segmentação de imagens, entre outros. No trabalho do Varma e Zisserman (2005) são definidos duas etapas de aprendizado, as quais são:

Geração do dicionário de textons Múltiplas imagens não registradas do conjunto de treinamento de uma classe de textura específica são convolucionadas com um banco de filtros. As respostas dos filtros aplicados são agregadas e agrupadas em textons usando o algoritmo *k-means*. Os textons de diferentes classes de textura são combinados para formar o dicionário de textons, como mostra a Figura 23.

Figura 23 – Aprendizado dos dicionários de textons

Fonte: (VARMA; ZISSERMAN, 2005)

Geração do modelo Dada uma imagem de treinamento, seu modelo correspondente é gerado pela primeira convolução com um banco de filtros e, em seguida, rotulando cada resposta de filtro com o texton que fica mais próximo dele no espaço de resposta do filtro. O histograma de texto, isto é, a frequência com que cada texton ocorre na rotulagem, forma o modelo correspondente à imagem de treinamento.

Figura 24 – Geração do Modelo

Fonte: (VARMA; ZISSERMAN, 2005)

Como foi descrito na Seção 1, um dos objetivos deste trabalho de mestrado é criar dicionários baseado em textons usando diferentes técnicas tais como: Transformada discreta de Fourier, filtros de Gabor, e teoria de redes complexas, aos quais denominaremos *TDFT*, *TGFB* e *TCN*.

Figura 25 – Textons de tamanho 20x20 para 2D Hela(a,c) e Olympus(b,d) usando média e variância em tons de cinza.

Fonte: Elaborada pelo autor.

A Figura 25 apresenta quatro dicionários de textons correspondentes a Microfilamentos de Actina e Microtúbulos do banco de dados *2D Hela* (Figuras 25a e 25b), além do seus correspondentes do banco de dados *FDIG Olympus* (Figuras 25c e 25d). Fazendo uma análise visual simples, evidencia-se que a Figura 25a e Figura 25b são fáceis de diferenciar mas não é a mesma situação das Figura 25c e Figura 25d. Nas próximas Seções apresentamos cada uma das três técnicas selecionadas para gerar os dicionários de textons.

3.5 2-D Discrete Fourier Transform

A transformada de Fourier irá decompor uma imagem em seus componentes senos e cossenos. Em outras palavras, ele irá transformar uma imagem de seu domínio espacial para o seu domínio de frequência. A ideia é que qualquer função pode ser aproximada exatamente com a soma de funções infinitas de seno e coseno. Formalmente a transformada de Fourier de imagens bidimensionais é dada por:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad (3.2)$$

onde :

$f(x, y)$ = é a imagem digital de dimensões $M \times N$

$$u = 0, 1, 2, \dots, M-1$$

$$v = 0, 1, 2, \dots, N-1$$

$$e^{jx} = \cos(x) + j \sin(x)$$

Nota-se que f é o valor da imagem no seu domínio espacial e F no seu domínio de frequência. O resultado da transformação são números complexos. Exibindo isso, é possível representar através de uma imagem real, uma imagem complexa ou através de uma magnitude e uma imagem de fase. No entanto, no contexto de algoritmos de processamento de imagem apenas a imagem magnitude é interessante, pois contém todas as informações que precisamos

sobre a estrutura geométrica de imagens. Para fazer algumas modificações da imagem desta forma, é necessário uma nova transformação, porém é necessário preservar ambas partes, real e imaginária. Na Figura 26 mostra-se a imagem de magnitude da transformada de Fourier 2D, obtida sobre duas imagens do banco de dados *FDIG Olympus*.

Figura 26 – Espectro do espaço frequência gerada pela transformada de Fourier 2D

(a) Microfilamentos de Actina (FDIG Olympus)

(b) Microtubulos (FDIG Olympus)

Fonte: Elaborada pelo autor usando (ITSEEZ, 2015)

No caso de imagens digitais, as cores são discretas; isso significa que eles podem assumir um valor de um dado valor de domínio. Por exemplo, numa escala de cinza, os valores de imagem normalmente estão entre 0 e 255. Portanto, a transformada de Fourier também precisa ser de um tipo discreto resultando em uma transformada de Fourier discreta (do inglês *Discrete Fourier Transform - DFT*). Isso sempre será utilizado para determinar a estrutura de uma imagem, de um ponto de vista geométrico. Dada a transformada $F(u, v)$, podemos obter $f(x, y)$ usando o inverso *inverse discrete Fourier transform* (IDFT)

$$f(x, y) = \frac{1}{MxN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} F(u, v) e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad (3.3)$$

onde :

$$x = 0, 1, 2, \dots, M - 1$$

$$y = 0, 1, 2, \dots, N - 1$$

$$e^{jx} = \cos(x) + j \sin(x)$$

Algumas das propriedades mais importantes são revisadas em (GONZALEZ; WOODS, 2006, 4).

A transformada de Fourier é amplamente utilizada para análise de forma (KADIR, 2015) aplicada em classificação do folhas, reconhecimento de caracteres (RAJPUT; HORAKERI, 2011; DONG *et al.*, 2013), recuperação de imagens baseada em conteúdo, compressão, e muitos outros.

Na análise de forma, os coeficientes transformados de Fourier formam os descritores de Fourier da forma. Estes descritores representam a forma em um domínio de frequência. Os descritores de frequência mais baixa contêm informações sobre as características gerais da forma e os descritores de frequência mais elevada contêm informações sobre detalhes mais finos da forma. Embora o número de coeficientes gerados a partir da transformação seja geralmente grande, um subconjunto dos coeficientes é suficiente para capturar as características gerais da forma. A transformada de Fourier é uma ferramenta de processamento de imagem significativa que é usada para representar uma imagem em seus componentes de seno e cosseno.

3.6 Filtros de Gabor

A Textura é muitas vezes caracterizada por suas respostas a um conjunto de orientação e filtro linear seletivo de frequência espacial que é inspirado por várias evidências de processamento semelhante no sistema de visão humana (JULESZ, 1981; ZHU *et al.*, 2005). Os *kernels* de Gabor são definidos da seguinte forma:

$$\Psi_{u,v}(x,y) = \frac{\|K_{u,v}\|^2}{\sigma^2} \exp\left(-\frac{\|K_{u,v}\|^2 \|z\|^2}{2\sigma^2}\right) [\exp(iK_{u,v}.z) - \exp(-\frac{\sigma^2}{2})] \quad (3.4)$$

onde :

$\|x\|$ = denota a norma $l2$ do vetor x

$.$ = denota produto ponto

$z = (x, y)$

$K_{u,v} = k_v \exp(i\phi_u)$

$k_v = \frac{k_{max}}{f^v}$

$\phi_u = \pi \frac{u}{U}$

$u = 0, 1, \dots, U - 1,$

$v = 0, 1, \dots, V - 1$

u e v definem a orientação e escala dos *kernels* de Gabor.

Os *kernels* de Gabor na Equação 3.4 são todos auto-similares, uma vez que podem ser gerados a partir de um filtro, a *wavelet* principal, escalando e girando através do vetor de onda $k_{u,v}$. Cada *kernel* é um produto de um envelope gaussiano e uma onda plana complexa, e pode ser separado em partes reais e imaginárias. Assim, uma faixa de filtros de Gabor é gerada por um conjunto de várias escalas e rotações. Os filtros de Gabor são usados em reconhecimento

de face (LEI *et al.*, 2007; YI; SU, 2013), expressões faciais, uma das mais recentes técnicas biométricas pesquisadas atualmente, o reconhecimento de padrões de veias (HAN; LEE, 2012), e ainda continua desenvolvendo-se.

Lei *et al.* (2007) fizeram um trabalho baseado em filtros de gabor e textons onde eles usam cinco escalas $v \in \{0, 1, 2, 3, 4\}$ e quatro orientações $u \in \{0, 2, 4, 6\}$ com o parâmetro $\sigma = 2\pi$ para gerar 40 filtros de Gabor. Logo, por convolução das imagens de face com os diferentes grãos de Gabor gerados correspondentes, para cada pixel de imagem obtém-se 40 coeficientes de Gabor que logo são agrupados mediante *k-means* para formar textons de Gabor.

Em Han e Lee (2012), faz-se o reconhecimento de padrões de veias, considerando a veia da palma da mão como uma instancia de textura e aplicando técnicas de extração de características baseadas em textura. Os filtros de Gabor fornecem uma definição ótima do domínio espacial e frequência, sendo uma base para extrair características locais no reconhecimento das veias da palma. No entanto, os filtros de Gabor tem muitas combinações de parâmetros potenciais, porém é uma prática comum usar vários filtros de Gabor ou determinar a combinação ótima de forma empírica. O objetivo do trabalho é discutir o algoritmo de otimização que determina os melhores parâmetros de um único filtro de Gabor para o reconhecimento da veia da palma da mão. Para obter um padrão eficaz de palma vascular, propôs-se um robusto método de filtro Gabor adaptativo para codificar as características da veia da palma da mão, como uma sequência de bits. A representação de sequência de bits, chamada *VeinCode*, oferece uma rápida correspondência de modelos e permite um armazenamento e recuperação de resultados mais eficazes. A similaridade de dois *VeinCodes* é medida pela distância normalizada de *Hamming*.

Na Figura 27 é possível observar o resultado depois de aplicar um banco de 16 filtros de Gabor (Figura 27a) sobre duas imagens uma do banco de dados 2D Hela (Figura 27b) e outra do FDIG Olympus (Figura 27c).

Figura 27 – Exemplo de uso dos filtros de Gabor.

Fonte: Elaborada pelo autor usando (ITSEEZ, 2015)

O presente trabalho de mestrado considera principalmente os trabalhos Lei *et al.* (2007),

Han e Lee (2012) e Yi e Su (2013), para geração dos textons baseados em filtros de Gabor. As configurações e detalhes dos experimentos realizados para avaliar estes filtros são apresentados na Seção 5.

3.7 Redes Complexas

O estudo das redes complexas se relaciona com diferentes áreas de conhecimento e vem sendo aplicada com êxito em cada uma delas. Em 1736, *Euler* propôs uma solução para o problema das pontes de *Königsberg*, que era atravessar as sete pontes da cidade sem repetir nenhuma delas. Euler provou que era impossível solucionar tal problema e este evento se tornou o início da teoria dos grafos (BIGGS; LLOYD; WILSON, 1986).

As redes complexas são grafos que apresentam uma estrutura não trivial e podem ser modeladas para a resolução de problemas específicos. A principal razão para a popularidade das redes complexas é sua flexibilidade e generalidade para representar qualquer estrutura, natural ou discreta, incluindo aquelas que sofrem mudanças dinâmicas de topologia. A pesquisa sobre redes complexas pode ser definida como a interseção entre a teoria de grafos e o mecanismo estatístico (COSTA *et al.*, 2007). É possível notar três grandes grupos que contribuíram para aumentar a pesquisa em redes complexas:

- Os primeiros modelos, dentre os quais pode ser citado o trabalho de Erdős e Rényi (1959), Erdős e Rényi (1960), Erdős e Rényi (1961), têm o propósito de gerar redes a partir de uma regra aleatória. Esse é o modelo mais simples que uma rede complexa pode assumir. Nesse modelo, arestas não direcionadas são adicionadas aleatoriamente entre um número fixo de N vértices. Cada aresta é independentemente representada com base em uma probabilidade p . O número de arestas que conectam cada vértice na rede, denominado grau do vértice, segue a distribuição de Poisson com um limite máximo N .
- Contudo, os modelos aleatórios não garantem uma estrutura topológica que representem bem as redes reais. Estudos sobre redes sociais motivaram alguns pesquisadores a desenvolver regras de formação que capturam o efeito do mundo pequeno (LATORA; MARCHIORI, 2001)(do inglês *small-world*) como o modelo *Watts-Strogatz*. O efeito *small-world* é caracterizado por baixo comprimento médio de caminhos e alto coeficiente de agrupamento.
- A própria internet não é bem representada por modelos como o de *Watts-Strogatz*. Na estrutura da Internet, existem muitos nós com poucas conexões e poucos nós com muitas conexões (*hubs*). Este tipo de comportamento é bem representado por modelos que apresentam uma distribuição de grau dos nós em lei de potência, como o modelo de Barabási e Albert (1999), Girvan e Newman (2002), Lewis (2009), Barabási e Bonabeau (2003).

Recentemente, a teoria de rede complexas foi utilizada para análise de formas (BACKES; CASANOVA; BRUNO, 2009), reconhecimento de padrões (CASANOVA; BACKES; BRUNO, 2013) e análise de textura (BACKES; CASANOVA; BRUNO, 2013; GONÇALVES *et al.*, 2016). Por exemplo, recentemente, Scabini *et al.* (SCABINI; GONÇALVES; CASTRO, 2015) propôs um método inovador usando o Bag of Visual Words e a teoria da rede complexa para análise de textura. Eles criam redes a partir de imagens e, em seguida, aplicam o *BoVW*, em vez de extrair medidas estatísticas globais de redes complexas.

3.7.1 Representação de rede complexa para a forma

Seja P um conjunto de N pontos no espaço R^2 , representando o contorno de uma imagem. Cada ponto é um vetor típico na forma de $p_i = [x_i, y_i]$. Para aplicar a Teoria das redes complexas, uma representação do contorno P deve ser construída como o grafo $G = \langle V, E \rangle$, cada ponto no conjunto P como um vértice no grafo (ou seja, $P = V$). Um conjunto de arestas não direcionadas E são vinculados a cada par de vértices compondo a rede. Este conjunto E é calculado usando a distância euclidiana (Equação 3.5) entre eles:

$$w(e_{i,j}) = |p_i - p_j| \quad (3.5)$$

Para garantir a invariância de escala da rede, normaliza-se o peso no intervalo $[0, 1]$:

$$w(e_{i,j}) = \frac{w(e_{i,j})}{\max_{w(e_{i,j}) \in E}} \quad (3.6)$$

Para extrair propriedades relevantes, é aplicado um processo de evolução dinâmico (COSTA *et al.*, 2007) a esta rede, visto que ao início, foi uma rede regular. Isso significa que todos os vértices da rede estão conectados entre si. Em seguida, aplica-se um limite t sobre as bordas E para alcançar um novo conjunto de bordas E^* , onde

$$E^* = \delta_t(E) = e_{i,j} \in E | w(e_{i,j}) \leq t \quad (3.7)$$

A partir deste novo conjunto de bordas E^* , $E^* \subseteq E$, surge uma nova rede $G^* = (V, E^*)$ que pode ser interpretada como um passo intermediário na evolução da rede G e possui propriedades relevantes da topologia da rede. Depois disso, pode-se usar medidas de grau. Por exemplo, no trabalho de Backes, Casanova e Bruno (2009) usa-se o grau máximo para caracterizar folhas de prantas, como é amostrado na Figura 28.

Figura 28 – Evolução da rede dinâmica como um *threshold* T_l e área de *zoom*: (a) $T_l = 0.1$; (b) $T_l = 0.15$ and (c) $T_l = 0.2$.

Fonte: Backes, Casanova e Bruno (2009)

Um dos mais recentes trabalhos relacionados a redes complexas é de Casanova, Backes e Bruno (2013). O trabalho propõe uma ferramenta matemática para caracterização de sinais, curvas e conjunto de pontos. Para avaliar o poder do descritor de padrões da proposta, é realizado um experimento de identificação de plantas com base na imagem das veias da folha. A veia da folha é uma característica usada para a identificação da planta (taxonomia), e uma de suas características é que essas estruturas são complexas e difíceis de serem representadas com sinais ou curvas, e de dessa maneira, serem analisadas com uma abordagem clássica de reconhecimento de padrões. Eles modelaram as veias como um conjunto de pontos e modelaram esses pontos como grafo. Como características, usaram as medidas de grau e juntaram as medidas de grau com o método de evolução dinâmica. A Figura 29 apresenta a ideia de como obter as características mediante o uso de *thresholds*.

Figura 29 – (a) Nervo na folha; (b)-(d) Rede complexa obtida com diferentes valores para *thresholds* t , $t = \{0.050, 0.075, 0.100\}$.

Fonte: Casanova, Backes e Bruno (2013)

3.7.2 Métricas

Neste trabalho de mestrado para o descritor baseado em redes complexas é usado duas métricas de grau e grau máximo, no entanto nesta Seção se estende a outras adicionais. As métricas mais comuns dentro da teoria das redes complexas são introduzidas a seguir.

3.7.2.1 Distribuição do Grau

A distribuição do grau $P(K)$, é definida como a probabilidade de que um vértice escolhido de maneira uniforme e aleatoriamente tenha grau K . A representação gráfica de $P(K)$ para qualquer rede estudada poderá ser descrita através de um histograma de graus dos vértices. Tal tipo de histograma é o que se denomina por distribuição do grau de uma rede.

3.7.2.2 Segundo Momento da Distribuição do Grau

Com a distribuição do grau é possível definir os diversos momentos para K . O primeiro momento da distribuição do grau, o qual representa o valor esperado, pode ser dado por:

$$\langle K \rangle = E(K) = \sum_{k=0}^{\infty} KP(K)$$

O segundo momento e dado pelo:

$$\langle K^2 \rangle = \sum_{k=0}^{\infty} K^2 P(K)$$

E de maneira geral

$$\langle K^m \rangle = \sum_{k=0}^{\infty} K^m P(K) \quad (3.8)$$

3.7.2.3 Entropia da distribuição das arestas (H_E)

Métrica proposta pelo [Preusse \(2012\)](#), onde foram apresentados várias medidas de equidade e desigualdade com base na distribuição de grau em redes.

$$H_E = - \sum_i \frac{K_i}{2|E|} \log \frac{K_i}{2|E|} \quad (3.9)$$

3.7.2.4 Entropia da distribuição de Grau (H_D)

Conhecida como a entropia de Shannon.

$$H_D = - \sum_k P(K) \ln(P(K)) \quad (3.10)$$

3.7.2.5 Grau Médio

Uma medida para caracterização da estrutura de redes é dada pela média do número de conexões entre os vértices, denominada grau médio $\langle k \rangle$. O grau de um vértice i para uma rede não-dirigida pode ser dado por:

$$k_i = \sum_j A_{ij} \quad (3.11)$$

Onde A_{ij} corresponde a matriz de adjacência, e o grau médio é:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N K_i \quad (3.12)$$

3.7.2.6 Coeficientes de Aglomeração local

Dado:

$$cc(i) = \frac{2\rho_i}{K_i(K_i - 1)}$$

Onde ρ_i é número de arestas entre os vizinhos de K_i , então define-se o coeficiente de aglomeração local pela equação 3.13:

$$\langle cc \rangle = \frac{1}{N} \sum_{i=1}^N cc(i) \quad (3.13)$$

3.7.2.7 Coeficiente de Aglomeração global

$$C = \frac{3N_{\Delta}}{N_3} \quad (3.14)$$

Onde N_3 é o número do triplas e N_{Δ} é o número de triângulos na rede.

3.7.2.8 Média dos menores caminhos

O comprimento de um caminho que conecta dois vértices é dado pelo número de arestas visitadas ao longo desse caminho. O comprimento do menor caminho entre dois vértices é d_{ij} , onde i e j representam o origem e destino. A distribuição dos menores caminhos entre todos os vértices em uma rede podem ser representados através de uma matriz de distâncias D , cujos elementos d_{ij} correspondem ao valor do menor caminho entre os vértices i e j .

$$l = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \quad (3.15)$$

3.7.2.9 Eficiência

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \quad (3.16)$$

3.7.2.10 Diâmetro

$$d = \max_{i,j} \{d_{ij}\} \quad (3.17)$$

3.8 Considerações finais

Neste Capítulo foram apresentadas os principais métodos utilizados neste trabalho com relação a descritores de características. Foram vistos *Bag of Visual Words*, *Local Binary Patterns*, *Discrete Fourier Transform*, filtros de Gabor e a teoria de redes complexas.

Atualmente não existe uma técnica capaz de classificar todos os tipos de imagens. Dependendo das características do problema de classificação de imagens a ser resolvido, um método específico para descrever as características das imagens é necessário. No próximo Capítulo é apresentada a metodologia do presente trabalho de mestrado.

METODOLOGIA

O presente trabalho de mestrado tem como objetivo a classificação de imagens de microscópio fluorescentes que contenham componentes do citoesqueleto, tais como Microfilamento de actina (MA) e Microtúbulos(MT). Porém, planeja-se implementar e avaliar algoritmos da área de processamento de imagens, visão computacional e teoria de redes complexas que permitam obter os descritores de características das imagens e posteriormente classificá-las. O processo de obtenção de esses descritores de características é chamado "caracterização de imagens". A principal vantagem de usar os descritores em vez das próprias imagens é o ganho em termos de tempo e complexidade de processamento. Então é necessário a escolha e avaliação dos algoritmos que permitam a obtenção dos descritores de características para as imagens consideradas neste trabalho(2D *Hela* e *FDIG Olympus*). Pelo dito acima, visa-se implementar:

1. Dicionários visuais (do inglês *Bag of Visual Words - BoVW*).
2. *Local Binary Patterns (LBP)*.
3. Transformada discreta de Fourier (do inglês *Discrete Fourier Transform - DFT*).
4. Filtros de Gabor.
5. Redes complexas (do inglês *Complex Networks*).

Para avaliá-los são considerados dois bancos de dados, o 2D *Hela* e *FDIG Olympus*. Planeja-se usar o *pipeline* clássico no processamento de imagem, sendo os passos principais: (1) Pré-processamento, (2) Extração de características para representar cada imagem mediante um descritor específico; e finalmente (3) Classificação. A Figura 30 mostra o esquema gráfico da metodologia. A seguir é detalhado cada um dos passos nesta metodologia.

Figura 30 – Esquema da metodologia: Linha azul claro seguida por *LBP* e linha laranja seguida por *BoVW* em conjunto com os algoritmos baseados nos textons

Fonte:Elaborada pelo autor

4.1 Pre-processing

Esta primeira etapa de pré-processamento (do inglês *pre-processing*) visa preparar as imagens de fluorescência, aplicando filtros lineares ou não lineares para melhorar, aumentar o contraste, normalizar o tamanho ou equalizar essas imagens. Dado que as imagens do banco de dados *2D Hela* foram obtidas em condições controladas e estão em tons de cinza, apenas aplica-se operações para acrescentar o contraste, não sendo o caso do *FDIG Olympus*, visto que estas imagens possuem vários níveis de escala, variedade de coloração(substância de contraste), e a principal dificuldade deste banco de imagens é a pouca quantidade de amostras (maiores detalhes no [Capítulo 5](#)).

4.2 Feature Extractor

Como é descrito no [Capítulo 3](#), o propósito e utilidade dos descritores de características é representar as imagens para seu processamento ao invés de processar a imagem original. Com isso busca-se reduzir o custo de processamento. A seguir há mais detalhes de como planeja-se implementar e usar cada um dos algoritmos.

4.2.1 Bag of Visual Words

Para o descritor baseado em *BoVW*, considera-se a proposta original de [Fei-Fei e Perona \(2005\)](#). Com uma variante essencial no processo de criação dos dicionários, criam-se dicionários por partes para depois uni-los, ou seja, para cada classe geram-se dicionários parciais, um para MA e outro para MT, e eles são unidos para gerar o dicionário geral(ou *codebook*). Este critério

permite acrescentar até 3% de precisão no processo final de classificação. A ideia foi obtida do trabalho de [Javed e Khan \(2011\)](#), onde são usados dicionários de textons para classificação de densidade populacional em imagens de satélite.

O processo inicia-se com a obtenção dos *keypoints* locais por cada imagem do grupo de treinamento usando SIFT ([LOWE, 2004](#)). Logo depois de localizar e descrever os *keypoints*, aplica-se o algoritmo *k-means* com *k* centroides para definir o dicionário. Considera-se tamanhos para *k* pequenos ($k = \{30, 40, 50, 60\}$), por cada classe no processo de treinamento. Finalmente, após da criação do dicionário, o objetivo é quantizar as incidências dos *keypoints* das imagens de treinamento e teste no dicionário, que serve para a geração de um histograma correspondente a cada imagem. Esse histograma é o descritor da imagem.

4.2.2 Local Binary Pattern

No caso do descritor baseado em *Local Binary Pattern (LBP)*, introduzido por [Ojala, Pietikäinen e Mäenpää \(2002\)](#), considera-se gerar imagens multi-escala adicionais para cada imagem usando uma interpolação bi-cúbica, em seguida aplicando o operador *LBP* sobre cada imagem original e as geradas(multi-escala), obten-se os histogramas correspondentes. Os histogramas são unidos e isso constitui o descritor *LBP* para cada imagem.

4.2.3 Dicionários de Textons

Por definição os textons são obtidos a partir de *patches* que tenham sido agrupados previamente, como é explicado no [Capítulo 3](#). O processo inicia-se com a divisão das imagens em pequenos *patches* de tamanhos 10x10, 12x12, 15x15 ou 20x20, e depois disso, executa-se o descritor correspondente.

Discrete Fourier Transform

Consideram-se os trabalhos de [Rajput e Horakeri \(2011\)](#) e [Dong et al. \(2013\)](#) para a transformada discreta de Fourier, dado que os *patches* e algumas letras apresentam uma morfologia muito próxima. Na Figura [31a](#) e [31b](#), é possível observar a similaridade dos espectros do Fourier para a classe MA, assim como nas Figuras [31c](#) e [31d](#) para a classe MT, além da diferença entre classes MA e MT.

Filtros de Gabor

Para o caso dos filtros de Gabor, planeja-se usar 16 filtros por cada imagem, gerando 16 imagens resposta adicionais por cada imagem original. O próximo passo é dividir essas imagens para gerar os *patches* de tamanhos 10x10, 12x12, 15x15, até 20x20 por exemplo, em seguida estimando a média e variância de cada *patch*, para então aplicar *clustering* com o objetivo de obter as *k* médias e variâncias mais representativas, sendo eles o dicionário(ou textons) para

Figura 31 – Espectro do espaço frequência gerada pela transformada de fourier em 2D

Fonte: Elaborada pelo autor usando (ITSEEZ, 2015)

este método. Finalmente, quantiza-se os textons para gerar o histograma correspondente a cada imagem do conjunto de treinamento e teste.

Redes Complexas

Para os textons baseados em redes complexas, é necessária previamente a transformação da imagem original em tons de cinza e em seguida em binária para uma posterior esqueletização(ou esqueletonização).

Figura 32 – Etapas do pré-processamento para a geração das redes complexas das imagens

Fonte: Elaborada pelo autor usando (ITSEEZ, 2015)

É necessário fazer um pré-processamento para aplicar a teoria de redes complexas, como é formulado no trabalho de Backes, Casanova e Bruno (2009). A Figura 32 mostra cada um deles. A binarização (Figura 32b) é seguida da esqueletização (Figura 32c), de acordo com a revisão de Saha, Borgefors e Baja (2015), considera-se o afinamento de Voronoi (do inglês *Voronoi thinning*) (RAMEY, 2013) para esqueletização. Esse algoritmo tem como base o trabalho de Zhang e Suen (1984) e diagramas de Voronoi. Consegue-se assim uma ótima saída para nosso propósito. Feita a esqueletização das imagens, é necessário a divisão de elas para a geração dos *patches*, sendo cada *path* transformado em uma rede complexa seguindo o algoritmo de (BACKES; CASANOVA; BRUNO, 2009), podendo agora usar as métricas de redes complexas para cada *patch*. Backes, Casanova e Bruno (2009) usaram isso para classificar formas, porém neste trabalho foram feitas alterações no algoritmo original para dar suporte a estruturas muito irregulares como linhas pseudo paralelas, estreladas, ou de malha, como mostra-se na Figura 32d.

4.2.4 Classifier

Cada um dos algoritmos selecionados nesta metodologia geram um descritor de características representados por um vetor de características ou *features vector*. Essas saídas agora serão a entrada de um classificador como Máquina de Vetores de Suporte (do inglês *Support Vector Machine - SVM*). Para cada um deles, é adotado o método de validação cruzada 5x2 (DIETTERICH, 1998) para avaliar os algoritmos. Todos os resultados são mostrados na Seção 5.3.

EXPERIMENTAÇÃO E RESULTADOS

Os experimentos foram feitos sobre dois bancos de imagens *Hela* e *FDIG Olympus*, os quais são detalhados na Seção 5.1. As implementações dos métodos propostos foram feitos na linguagem de programação C/C++ no sistema operacional Linux, com o compilador GNU gcc/g++, versão 4.9.2. em conjunto com a biblioteca *standard* de Visão Computacional OpenCV v2.4.9 (ITSEEZ, 2015). Os *scripts* e código fonte estão disponíveis no repositório *Github*¹.

O objetivo deste Capítulo é detalhar cada um dos métodos usados, como suas configurações, além dos resultados com as taxas de acerto por cada algoritmo, conforme definido no Capítulo 4.

5.1 Banco de imagens fluorescentes

Para comparar os cinco descritores de características obtidos, baseados em *Bag of Visual Words*, *Local Binary Local*, *Discrete Fourier Transform*, filtro de Gabor e Redes complexas, detalhados no Capítulo 4, usa-se os seguintes bancos de imagens fluorescentes: *2D Hela*, introduzido por (BOLAND; MURPHY, 2001), que é um dos conjuntos de dados mais comuns para imagens de fluorescência; e o segundo banco de dados *FDIG Olympus* (OLYMPUS, 2017), uma galeria de imagens da companhia *Olympus America Inc.* A seguir descrevemos cada uma delas.

5.1.1 2D Hela

2D Hela é um conjunto de dados com 904 de imagens de microscopia de fluorescência de células Hela coradas, com vários corantes fluorescentes específicos de organela, que são núcleo, retículo endoplasmático, giantin, (cis / medial golgi), cis golgi, lisossomos, mitocôndria, nucleolina, actina, endossomas e tubulina. Consideramos só duas dessas categorias para nossa análise: Microfilamentos de actina (MA) e Microtúbulos (MT). A Tabela 1 mostra mais detalhes

¹ <<https://github.com/fincahuanaco/fluorescence-image-analysis>>

em quantidade por classe. A Figura 33 apresenta alguns exemplos de cada uma dessas imagens. O banco de dados está disponível no site do laboratório de Robert R. Murphy denominado *Murphy Lab*² da *Carnegie Mellon University*.

Tabela 1 – Exemplo do banco de imagens *2D HeLa*

Category	Label	Quantity
Actin Microfilaments	MA	98
Microtubules	MT	91

Figura 33 – Exemplo do banco de imagens *2D HeLa*, Microfilamentos de Actina(linha superior) e Tubulina(linha inferior)

5.1.2 FDIG Olympus

No caso do *Fluorescence Digital Image Gallery* (OLYMPUS, 2017), há tipos diferentes de imagens micrográficas, que são imagens de fluorescência. Coletamos imagens da galeria de imagens oficial, especialmente das categorias de Microfilamentos de Actina(MA) e Microtúbulos(MT). A Tabela 2 mostra a quantidade de imagens coletadas válidas para nossos experimentos.

Tabela 2 – Detalhes do banco de imagens *FDIG Olympus*

Category	Label	Quantity
Actin Microfilaments	MA	48
Microtubules	MT	38

As imagens originais estão em formato de cor, com 670x500 *pixels* de tamanho. A principal vantagem é que as imagens são rotuladas, fornecendo um *ground truth* sobre o que elas contém, mas, por outro lado, não incluem informações sobre as condições de como elas foram obtidas ou informações de escala ou perspectiva das imagens. A Figura 34 mostra alguns exemplos desse banco de imagens. O banco de imagens está disponível no site do *Olympus*³.

² *2D HeLa* <<http://murphylab.web.cmu.edu/data/>>

³ *FDIG Olympus* <<http://www.olympusmicro.com/galleries/index.html>>

Figura 34 – Filamentos de actina (fila superior) e Microtubulos(fila de baixo) do banco de dados *FDIG Olympus* (Olympus America, Inc.©)

Nas Figuras 33 e 34 é possível identificar a diferença entre ambos os conjuntos de dados. Observe que os conjuntos de dados *2D Hela* e *FDIG Olympus* são diferentes. *2D Hela* está em escala de tons de cinza e *FDIG Olympus* está em formato de cor (RGB). O *2D Hela* é intuitivamente fácil de entender a partir da perspectiva do processamento de imagens, este conjunto de dados é bem conhecido e usado (por exemplo, [Lan et al. \(2015\)](#)), em contraste com o banco de dados da *FDIG Olympus*. O *FDIG Olympus* é um formato maior e colorido, mas mistura diferentes componentes do citoesqueleto como um núcleo celular, redes mitocôndrias e outros, o que significa um desafio para o nosso trabalho para compreendê-los.

5.2 Configuração de experimentos

No [Capítulo 3](#), detalha-se o funcionamento de cada um dos métodos, e no [Capítulo 4](#) como serão usados para obter os descritores de características para as imagens de os dois conjuntos de dados descritos na Seção 5.1. Esta Seção conduz os experimentos para comparar esses métodos em ambos os conjuntos de dados (*2D Hela* vs *FDIG Olympus*). A seguir o detalhe de como foram configurados cada um dos métodos para executar os experimentos:

- *Bag of Visual Words*. Dado que considera-se a proposta original de [Fei-Fei e Perona \(2005\)](#) onde o *BoVW* requer *SIFT* ([LOWE, 2004](#)) para a localização dos *keypoints*, é fundamental a configuração dos parâmetros do algoritmo *SIFT* para localizar os *keypoints* para obter a máxima resposta. Define-se os parâmetros *number keypoints* = 600, *Scale levels* = 3, *contrast threshold* = 0.04, *edge threshold* = 10 e *sigma* = 0.6. A quantidade de *keypoints* por imagem foi obtido mediante uma busca exaustiva. Da mesma forma foram testados a influência da quantidade de palavras visuais por dicionário por cada classe, concluindo o tamanho de dicionário com *size* = 120, 60 palavras por classe.
- *Local Binary Pattern*. Primeiro é necessário aplicar o filtro Gabor para aumentar o contraste de bordas, depois executar o operador LBP ([NANNI; LUMINI, 2008](#)) para construir histogramas. Com uma pequena alteração, consideramos multi-escala em duas escalas para cada imagem, obtendo o respectivo histograma e juntando-os. Para uma comparação

com os outros métodos, usamos o mesmo esquema dividindo em dois grupos 50% para treinamento e 50% para teste.

- Textons baseados em *Discrete Fourier Transform*. Para os métodos baseados em textons, dividimos em pequenos *patches* quadrados de $n \times n$ pixels (ou seja, 10x10, 12x12, 15x15 e 20x20 como tamanho máximo), depois aplica-se o descritor correspondente (*Discrete Fourier Transform*, filtros de Gabor ou redes complexas). Essa saída é agrupada usando *k-means*. Os centroides que chamamos textons (ZHU *et al.*, 2005) são usados para quantificar e obter histogramas por cada imagem.
- Textons baseados em filtros de Gabor. Na Figura 27a da Seção 3.6 foi apresentado um banco com 16 filtros, usados para o conjunto de dados 2D *Hela*. Da mesma forma usa-se um banco de 16 filtros para cada banco de dados, e eles são aplicados em cada imagem obtendo 16 respostas como mostra-se nas Figuras 27b e 27c. Então geram-se os *patches* também de tamanhos 10x10, 12x12, 15x15 e 20x20 e suas métricas correspondentes por cada *patch* (ou seja, média e variância) para criar o vetor de características. O próximo passo é aplicar o método de agrupamento *k-means* para obter os centroides, e assim criar o dicionário.
- Textons baseados em redes complexas. É necessária algumas etapas adicionais no pré-processamento, como transformação em uma imagem binária e a esqueletização da mesma (detalhado nas Seções 3.7 e 4 da metodologia). Originalmente, o algoritmo foi usado na forma de análise que adaptamos para nossas estruturas.

Para o uso do classificador *SVM*, usamos libSVM (CHANG; LIN, 2011). Para avaliar o aprendizado do classificador *SVM*, realiza-se uma validação cruzada 5x2 (DIETTERICH, 1998) com cada método por cada conjunto de dados, obtendo ao final 10 resultados, o que significa dividir aleatoriamente as amostras em duas camadas, usando rotativamente uma camada para treinar e a outra para teste cinco vezes. Consideramos o *kernel* linear para todos os casos. Para manter uma comparação mais justa possível, foi escolhido esse método de validação por que o banco de dados *FDIG Olympus* tem uma baixa quantidade de amostras por classe.

5.3 Resultados

Depois de executar os experimentos, obtive-se como melhor caso para classificação, para o conjunto de dados *FDIG Olympus* com *Bag of Visual Words*, conseguindo até 90.07% de precisão. No entanto, para o conjunto de dados *2D Hela* obteve uma resposta de até 95.79%. Como é possível apreciar nas primeiras linhas nas Tabelas 3 e 4. *BoVW* perde em relação ao *LBP* que atinge até 98.94%, o que evidencia que a classificação das duas classes (MA e MT) para o banco de dados *2D Hela* está praticamente resolvido só usando *LBP*.

Para os outros métodos baseados em textons como a transformada discreta de Fourier e os filtros Gabor, obtive-se resultados muito próximos para o banco de dados *2D Hela*, mas os mais baixos para *FDIG Olympus*. Em relação ao método baseado em redes complexas, este supera o *Bag of Visual Words* em *2D Hela* com 96.84%, mas para *FDIG Olympus* cai para 86.05%.

O conteúdo da Tabela 3 mostra as matrizes de confusão das classes MA e MT, onde é verificável que a precisão é maior em *2D Hela* do que *FDIG Olympus*.

Tabela 3 – Comparação das matrizes de confusão obtidas, banco de dados (as columnas *2D Hela* e *FDIG Olympus*) por método(linhas *BoVW*, *LBP*, *TDFT*, *TGFB* e *TCN*)

Method	2D Hela			<i>FDIG Olympus</i>		
	MA	MT		MA	MT	
BoVW	48	1	MA	22	2	MA
	3	43	MT	2	17	MT
LBP	49	0	MA	20	4	MA
	1	44	MT	1	18	MT
TDFT	48	1	MA	15	9	MA
	3	43	MT	2	17	MT
TGFB	47	2	MA	24	0	MA
	0	46	MT	6	13	MT
TCN	45	4	MA	23	1	MA
	5	41	MT	5	14	MT

Fonte: Dados da pesquisa.

A Figura 35 apresenta o resumo geral do presente trabalho de mestrado, onde é possível ver as barras em tons de cinza as precisões dos métodos para o banco de imagens *2D Hela* e em cor verde para o *FDIG Olympus*.

É importante realçar que as condições de como os dados foram obtidos são muito diferentes, uma em condições controladas (*2D Hela*) e outro não controlados (*FDIG Olympus*), o que caracteriza o desafio deste trabalho. Os resultados apresentados na Seção 5.3 corroboram nossa afirmação.

Tabela 4 – Comparação de Precisão, *Recall* e *F-Measure* de ambos bancos de dados(colunas) por métodos(linhas)

Method	2D Hela				FDIG Olympus			
	Precision	Recall	F-Measure	Class	Precision	Recall	F-Measure	Class
BoVW	0.94	0.98	0.96	MA	0.92	0.92	0.92	MA
	0.98	0.94	0.96	MT	0.90	0.90	0.90	MT
	Accuracy 95.79%				Accuracy 90.70%			
LBP	0.98	1	0.99	MA	0.95	0.83	0.89	MA
	1	0.98	0.99	MT	0.82	0.95	0.88	MT
	Accuracy 98.94%				Accuracy 88.37%			
TDFT	0.94	0.98	0.96	MA	0.88	0.63	0.73	MA
	0.98	0.94	0.956	MT	0.65	0.90	0.76	MT
	Accuracy 95.79%				Accuracy 74.42%			
TGFB	1	0.96	0.98	MA	0.82	0.96	0.89	MA
	0.96	1	0.98	MT	0.93	0.74	0.82	MT
	Accuracy 97.89%				Accuracy 86.05%			
TCN	0.90	0.918	0.91	MA	0.821	0.958	0.885	MA
	0.91	0.891	0.90	MT	0.933	0.737	0.824	MT
	Accuracy 96.84%				Accuracy 86.05%			

Fonte: Dados da pesquisa.

Figura 35 – Resumo dos cinco métodos (BoVW, LBP, TDFT, TGFB and TCN) aplicados ao banco de dados 2D Hela e FDIG Olympus

Fonte: Dados da pesquisa.

CONCLUSÃO

O presente trabalho de mestrado avalia cinco métodos de extração de características sobre dois bancos de dados (*2D Hela* e *FDIG Olympus*), com objetivo de poder classificá-los e gerar uma linha base ou referente para futuras tentativas com este tipo de imagens.

Todos os métodos selecionados funcionam muito bem com o conjunto de dados *2D Hela*, mas não é o caso do banco de dados *FDIG Olympus*, devido a: (a) A substância usada para acrescentar o contraste dentro do processo de obtenção das imagens de fluorescência varia de uma para outra imagem (rotulagem de cores), e (b) presença de níveis elevados de escala entre as imagens, além do número reduzido de amostras.

Os resultados obtidos evidenciam que é possível classificar o conjunto de dados da *FDIG Olympus* com uma precisão de até 90.07%, tendo como método principal o *BoVW*. No entanto, a natureza multi-escala do algoritmo *SIFT* para a filtragem dos *keypoints* também poderia ser usado nos outros descritores para acrescentar o resultado, que por enquanto ficam até 2% abaixo do *BoVW*.

A teoria de redes complexas que foi usada para criar textons com objetivo de descrever imagens, obteve bons resultados em relação aos bancos de dados, conseguindo terceiro *ranking* em *2D Hela* e *FDIG Olympus*, considerando que foram usados só grau médio e máximo de conexões em diferentes intervalos como métricas.

Em relação aos trabalhos futuros, planeja-se adicionar um método baseado em *Deep Learning* para a caracterização das imagens, e assim acrescentar nossa galeria de métodos de caracterização de imagens com o objetivo de criar um *benchmark* para este tipo de imagens, ou seja imagens fluorescentes do citoesqueleto.

Finalmente visa-se a criação de um descritor que combine a teoria de redes complexas com o descritor *LBP*, aproveitando a simplicidade do *LBP* e a robustez das redes complexas.

REFERÊNCIAS

ARTHUR, D.; VASSILVITSKII, S. K-means++: The advantages of careful seeding. In: **Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms**. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2007. (SODA '07), p. 1027–1035. ISBN 978-0-898716-24-5. Disponível em: <http://dl.acm.org/citation.cfm?id=1283383.1283494>. Citado na página 38.

ATKINS, L. J. P. **Princípios de Química: Questionando a Vida Moderna e o Meio Ambiente**. 5ª. ed. Bookman, 2012. ISBN 978-85-407-0054-3. Disponível em: <http://gen.lib.rus.ec/book/index.php?md5=8DF3539D4758E3B578BB14C9E0FED656>. Citado na página 23.

BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. A complex network-based approach for boundary shape analysis. **Pattern Recognition**, Elsevier Science Inc., New York, NY, USA, v. 42, n. 1, p. 54–67, jan. 2009. ISSN 0031-3203. Disponível em: <http://dx.doi.org/10.1016/j.patcog.2008.07.006>. Citado nas páginas 25, 55, 56 e 65.

_____. Texture analysis and classification: A complex network-based approach. **Information Sciences**, v. 219, p. 168 – 180, 2013. ISSN 0020-0255. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0020025512004677>. Citado na página 55.

BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, p. 509–512, 1999. Citado na página 54.

BARABÁSI, A.-L.; BONABEAU, E. Scale-free networks. **Scientific American**, v. 288, p. 50–59, 2003. Citado na página 54.

BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. **IEEE Trans. Pattern Anal. Mach. Intell.**, v. 19, n. 7, p. 711–720, 1997. Disponível em: <http://dblp.uni-trier.de/db/journals/pami/pami19.html#BelhumeurHK97>. Citado na página 39.

BIGGS, N.; LLOYD, E. K.; WILSON, R. J. **Graph Theory, 1736-1936**. New York, NY, USA: Clarendon Press, 1986. ISBN 0-198-53916-9. Citado na página 54.

BISHOP, C. M. **Neural Networks for Pattern Recognition**. [S.l.]: Oxford University Press, 1995. Citado na página 39.

BOLAND, M. V.; MURPHY, R. F. A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of hela cells. **Bioinformatics**, v. 17, n. 12, p. 1213, 2001. Disponível em: <http://dx.doi.org/10.1093/bioinformatics/17.12.1213>. Citado nas páginas 24 e 67.

BOTTOU, L.; BENGIO, Y. Convergence properties of the k-means algorithms. In: **Advances in Neural Information Processing Systems 7**. [S.l.]: MIT Press, 1995. p. 585–592. Citado na página 37.

BREUER, D.; IVAKOV, A.; SAMPATHKUMAR, A.; HOLLANDT, F.; PERSSON, S.; NIKOLOSKI, Z. Quantitative analyses of the plant cytoskeleton reveal underlying organizational principles. **Journal of The Royal Society Interface**, The Royal Society, v. 11, n. 97, p. 20140362, 2014. Citado na página 24.

CASANOVA, D.; BACKES, A. R.; BRUNO, O. M. Pattern recognition tool based on complex network-based approach. **Journal of Physics: Conference Series**, v. 410, n. 1, p. 012048, 2013. Disponível em: <<http://stacks.iop.org/1742-6596/410/i=1/a=012048>>. Citado nas páginas 25, 55 e 56.

CHANG, C.-C.; LIN, C.-J. LIBSVM: A library for support vector machines. **ACM Transactions on Intelligent Systems and Technology**, v. 2, p. 27:1–27:27, 2011. Software available at <<http://www.csie.ntu.edu.tw/~cjlin/libsvm>>. Citado na página 70.

CHEN, C. H. **Handbook of Pattern Recognition and Computer Vision**. 4th. ed. River Edge, NJ, USA: World Scientific Publishing Co., Inc., 2010. ISBN 9789814273381, 9814273384. Citado na página 36.

COELHO, L. P.; KANGAS, J. D.; NAIK, A. W.; OSUNA-HIGHLEY, E.; GLORY-AFSHAR, E.; FUHRMAN, M.; SIMHA, R.; BERGET, P. B.; JARVIK, J. W.; MURPHY, R. F. Determining the subcellular location of new proteins from microscope images using local features. **Bioinformatics**, Oxford University Press (OUP), v. 29, n. 18, p. 2343–2349, jul 2013. Disponível em: <<https://doi.org/10.1093/bioinformatics/btt392>>. Citado na página 24.

COSTA, L. D.; RODRIGUES, F. A.; TRAVIESO, G.; BOAS, P. R. V. Characterization of complex networks: a survey of measurements. **Advances in Physics**, v. 56, p. 167–242, 2007. Citado nas páginas 54 e 55.

CSURKA, G.; DANCE, C.; FAN, L.; WILLAMOWSKI, J.; BRAY, C. Visual categorization with bags of keypoints. In: PRAGUE. **Workshop on statistical learning in computer vision, ECCV**. [S.l.], 2004. v. 1, n. 1-22, p. 1–2. Citado nas páginas 24, 41 e 42.

DIETTERICH, T. G. Approximate statistical tests for comparing supervised classification learning algorithms. **Neural Comput.**, MIT Press, Cambridge, MA, USA, v. 10, n. 7, p. 1895–1923, out. 1998. ISSN 0899-7667. Disponível em: <<http://dx.doi.org/10.1162/089976698300017197>>. Citado nas páginas 65 e 70.

DONG, L.; WANG, J.; LI, Y.; TANG, Y. Y. Sector projection fourier descriptor for chinese character recognition. In: **2013 IEEE International Conference on Cybernetics (CYBCO)**. [S.l.: s.n.], 2013. p. 162–167. Citado nas páginas 52 e 63.

ERDŐS, P.; RÉNYI, A. On random graphs. **Publicationes Mathematicae (Debrecen)**, v. 6, p. 290–297, 1959. Citado na página 54.

_____. On the evolution of random graphs. **Publication of the Mathematical Institute of the Hungarian Academy of Sciences**, v. 5, p. 17–61, 1960. Citado na página 54.

_____. On the strength of connectedness of a random graph. **Acta Mathematica Hungarica**, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV, Formerly Kluwer Academic Publishers BV, v. 12, n. 1-2, p. 261–267, 1961. Citado na página 54.

FEI-FEI, L.; PERONA, P. A bayesian hierarchical model for learning natural scene categories. In: **2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)**. [S.l.: s.n.], 2005. v. 2, p. 524–531 vol. 2. ISSN 1063-6919. Citado nas páginas 62 e 69.

GIRVAN, M.; NEWMAN, M. E. J. Community structure in social and biological networks. **Proceedings of the National Academy of Sciences**, v. 99, n. 12, p. 7821–7826, 2002. Citado na página 54.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing (3rd Edition)**. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. ISBN 013168728X. Citado nas páginas 24, 28, 29, 35, 36 e 51.

GONÇALVES, W. N.; SILVA, N. R. da; COSTA, L. da F.; BRUNO, O. M. Texture recognition based on diffusion in networks. **Inf. Sci.**, Elsevier Science Inc., New York, NY, USA, v. 364, n. C, p. 51–71, out. 2016. ISSN 0020-0255. Disponível em: <<http://dx.doi.org/10.1016/j.ins.2016.04.052>>. Citado nas páginas 25 e 55.

HAN, W.-Y.; LEE, J.-C. Palm vein recognition using adaptive gabor filter. **Expert Syst. Appl.**, Pergamon Press, Inc., Tarrytown, NY, USA, v. 39, n. 18, p. 13225–13234, dez. 2012. ISSN 0957-4174. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2012.05.079>>. Citado nas páginas 53 e 54.

HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: **Proceedings of the Alvey Vision Conference 1988**. Alvey Vision Club, 1988. Disponível em: <<https://doi.org/10.5244/c.2.23>>. Citado nas páginas 42 e 44.

HUNT, R. **The Reproduction of Colour**. Wiley, 2005. (The Wiley-IS&T Series in Imaging Science and Technology). ISBN 9780470024263. Disponível em: <<https://books.google.com.br/books?id=nFtW4LG24fEC>>. Citado na página 28.

ITSEEZ. **Open Source Computer Vision Library**. 2015. <<https://github.com/itseez/opencv>>. Citado nas páginas 46, 51, 53, 64 e 67.

JAVED, Y.; KHAN, M. M. Image texture classification using textons. In: **2011 7th International Conference on Emerging Technologies**. [S.l.: s.n.], 2011. p. 1–5. Citado nas páginas 48 e 63.

JOLLIFFE, I. **Principal Component Analysis**. [S.l.]: Springer Verlag, 1986. Citado nas páginas 38 e 39.

JULESZ, B. Textons, the elements of texture perception, and their interactions. **Nature**, v. 290, n. 5802, p. 91–97, Mar 1981. Disponível em: <<http://dx.doi.org/10.1038/290091a0>>. Citado nas páginas 48 e 52.

KADIR, A. Leaf identification using fourier descriptors and other shape features. **Gate to Computer Vision and Pattern Recognition**, Science Gate Publishing PC, v. 1, n. 1, p. 3–7, apr 2015. Disponível em: <<https://doi.org/10.15579/gtcvpr.0101.003007>>. Citado na página 52.

KANUNGO, T.; MOUNT, D. M.; NETANYAHU, N. S.; PIATKO, C. D.; SILVERMAN, R.; WU, A. Y. An efficient k-means clustering algorithm: analysis and implementation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 7, p. 881–892, Jul 2002. ISSN 0162-8828. Citado na página 37.

- KUBITSCHKE, U. **Fluorescence Microscopy : from principles to biological applications**. Weinheim: Wiley-Blackwell, 2013. ISBN 978-3-527-32922-9. Citado na página 23.
- LAN, X.; LI, L.; HU, J.; ZHANG, Q.; DANG, Y.; HUANG, Y. A quantitative method for microtubule analysis in fluorescence images. **Microscopy and Microanalysis**, Cambridge University Press, v. 21, n. 6, p. 1582–1590, 2015. Citado nas páginas 24 e 69.
- LATORA, V.; MARCHIORI, M. Efficient behavior of small-world networks. **Physical Review Letters**, v. 87, n. 19, p. 198701–1, 2001. Citado na página 54.
- LEI, Z.; LI, S. Z.; CHU, R.; ZHU, X. Face recognition with local gabor textures. In: _____. **Advances in Biometrics: International Conference, ICB 2007, Seoul, Korea, August 27-29, 2007. Proceedings**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. p. 49–57. ISBN 978-3-540-74549-5. Disponível em: <http://dx.doi.org/10.1007/978-3-540-74549-5_6>. Citado na página 53.
- LEWIS, T. G. **Network Science: Theory and Applications**. [S.l.]: Wiley Publishing, 2009. ISBN 0470331887, 9780470331880. Citado na página 54.
- LIN, W.-C.; HAYS, J.; WU, C.; LIU, Y.; KWATRA, V. Quantitative evaluation of near regular texture synthesis algorithms. In: **2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)**. [S.l.: s.n.], 2006. v. 1, p. 427–434. ISSN 1063-6919. Citado na página 36.
- LODISH, H.; BERK, A.; MATSUDAIRA, P.; KAISER, C. A.; KRIEGER, M.; SCOTT, M. P.; ZIPURSKY, L.; DARNELL, J. **Molecular Cell Biology**. Fifth edition. W. H. Freeman, 2008. ISBN 0716743663, 9780716743668. Disponível em: <<http://gen.lib.rus.ec/book/index.php?md5=FCEA1ADE54D39A86C422E3C5AD8D6822>>. Citado na página 35.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. **Int. J. Comput. Vision**, Kluwer Academic Publishers, Hingham, MA, USA, v. 60, n. 2, p. 91–110, nov. 2004. ISSN 0920-5691. Disponível em: <<http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>>. Citado nas páginas 24, 41, 42, 43, 44, 45, 63 e 69.
- NANNI, L.; LUMINI, A. A reliable method for cell phenotype image classification. **Artif. Intell. Med.**, Elsevier Science Publishers Ltd., Essex, UK, v. 43, n. 2, p. 87–97, jun. 2008. ISSN 0933-3657. Disponível em: <<http://dx.doi.org/10.1016/j.artmed.2008.03.005>>. Citado na página 69.
- OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Trans. Pattern Anal. Mach. Intell.**, IEEE Computer Society, Washington, DC, USA, v. 24, n. 7, p. 971–987, jul. 2002. ISSN 0162-8828. Disponível em: <<http://dx.doi.org/10.1109/TPAMI.2002.1017623>>. Citado nas páginas 46 e 63.
- OLYMPUS. **Microscopy Resource Center**. 2017. Disponível em: <<http://www.olympusmicro.com/primer/lightandcolor/fluorointroduction.html>>. Citado nas páginas 29, 30, 31, 67 e 68.
- PREUSSE, J. Fairness on the Web: Alternatives to the Power Law. 2012. Citado na página 57.
- PURVES, W. K.; SADAVA, D.; ORIAN, G. H.; HELLER, H. C. **Life, the science of biology**. 7th ed. ed. Sinauer Associates; W.H. Freeman and Co, 2004. ISBN 9780716798569, 0716798565. Disponível em: <<http://gen.lib.rus.ec/book/index.php?md5=906B377C74C651DCFE4ACEFB3F3524E6>>. Citado na página 34.

RAJPUT, G. G.; HORAKERI, R. Shape descriptors based handwritten character recognition engine with application to kannada characters. In: **2011 2nd International Conference on Computer and Communication Technology (ICCCT-2011)**. [S.l.: s.n.], 2011. p. 135–141. Citado nas páginas 52 e 63.

RAMEY, A. **Super-fast thinning implementation (Zhang-Suen, Guo-Hall)**. 2013. <<https://github.com/arnaud-ramey/voronoi>>. Citado na página 65.

SAHA, P. K.; BORGEFORS, G.; BAJA, G. S. di. A survey on skeletonization algorithms and their applications. **Pattern Recognition Letters**, p. –, 2015. ISSN 0167-8655. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167865515001233>>. Citado na página 65.

SCABINI, L. F. S.; GONÇALVES, W. N.; CASTRO, A. A. Texture analysis by bag-of-visual-words of complex networks. In: _____. **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 20th Iberoamerican Congress, CIARP 2015, Montevideo, Uruguay, November 9-12, 2015, Proceedings**. Cham: Springer International Publishing, 2015. p. 485–492. ISBN 978-3-319-25751-8. Disponível em: <http://dx.doi.org/10.1007/978-3-319-25751-8_58>. Citado nas páginas 25 e 55.

SILVA, C.; BOUWMANS, T.; FRÉLICOT, C. An eXtended Center-Symmetric Local Binary Pattern for Background Modeling and Subtraction in Videos. In: **International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISAPP 2015**. Berlin, Germany: [s.n.], 2015. Disponível em: <<https://hal.archives-ouvertes.fr/hal-01227955>>. Citado nas páginas 46 e 47.

SPRING, K. R. **Fluorescence Microscopy**. National Institutes of Health, Bethesda, Maryland, U.S.A.: Marcel Dekker, Inc., 2003. Citado na página 31.

VARMA, M.; ZISSERMAN, A. A statistical approach to texture classification from single images. **International Journal of Computer Vision**, v. 62, n. 1, p. 61–81, Apr 2005. ISSN 1573-1405. Disponível em: <<https://doi.org/10.1023/B:VISI.0000046589.39864.ee>>. Citado nas páginas 48 e 49.

WYSZECKI, G.; STILES, W. **Color Science: Concepts and Methods, Quantitative Data and Formulae**. Wiley, 2000. (Wiley Series in Pure and Applied Optics). ISBN 9780471399186. Disponível em: <https://books.google.com.br/books?id=_51HDcjWZPwC>. Citado na página 28.

XIE, J.; ZHANG, L.; YOU, J.; SHIU, S. Effective texture classification by texton encoding induced statistical features. **Pattern Recognition**, v. 48, n. 2, p. 447 – 457, 2015. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0031320314003288>>. Citado na página 48.

YI, J.; SU, F. Gabor orientation histogram for face representation and recognition. **Proceedings of the 3rd International Conference on Multimedia Technology (ICMT 2013). Lecture Notes in Electrical Engineering**, Springer, Berlin, Heidelberg, v. 278, n. 3, p. 35–43, 2013. Citado nas páginas 53 e 54.

ZHANG, T. Y.; SUEN, C. Y. A fast parallel algorithm for thinning digital patterns. **Commun. ACM**, ACM, New York, NY, USA, v. 27, n. 3, p. 236–239, mar. 1984. ISSN 0001-0782. Disponível em: <<http://doi.acm.org/10.1145/357994.358023>>. Citado na página 65.

ZHU, S.-C.; GUO, C.-e.; WANG, Y.; XU, Z. What are textons? **International Journal of Computer Vision**, v. 62, n. 1, p. 121–143, 2005. ISSN 1573-1405. Disponível em: <<http://dx.doi.org/10.1023/B:VISI.0000046592.70770.61>>. Citado nas páginas 48, 52 e 70.