

UNIVERSIDADE FEDERAL DO MARANHÃO - UFMA  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE ELETRICIDADE

ALEX MARTINS SANTOS

---

***DIAGNÓSTICO DE DEGENERAÇÃO MACULAR  
RELACIONADA À IDADE A PARTIR DE TOMOGRAFIA  
DE COERÊNCIA ÓPTICA USANDO GEOESTATÍSTICA E  
CAPSULE NETWORKS***

---

SÃO LUÍS - MA  
2019

ALEX MARTINS SANTOS

**DIAGNÓSTICO DE DEGENERAÇÃO MACULAR RELACIONADA À  
IDADE A PARTIR DE TOMOGRAFIA DE COERÊNCIA ÓPTICA  
USANDO GEOESTATÍSTICA E CAPSULE NETWORKS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão como requisito parcial para obtenção do título de Doutor em Engenharia de Eletricidade.

Orientador: Prof. Dr. Anselmo Cardoso de Paiva  
Coorientador: Prof. Dr. Aristófanês Corrêa Silva

São Luís - MA  
2019

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).  
Núcleo Integrado de Bibliotecas/UFMA

Santos, Alex Martins.

Diagnóstico de degeneração macular relacionada à idade a partir de tomografia de coerência óptica usando geoestatística e Capsule Networks / Alex Martins Santos. - 2019.

106 f.

Coorientador(a): Aristófanês Corrêa Silva.

Orientador(a): Anselmo Cardoso de Paiva.

Tese (Doutorado) - Programa de Pós-graduação em Engenharia de Eletricidade/ccet, Universidade Federal do Maranhão, São Luís, 2019.

1. CapsuleNet. 2. DMRI. 3. OCT. 4. Retina. I. Paiva, Anselmo Cardoso de. II. Silva, Aristófanês Corrêa. III. Título.

**DIAGNÓSTICO DE DEGENERAÇÃO MACULAR RELACIONADA À  
IDADE A PARTIR DE TOMOGRAFIA DE COERÊNCIA ÓPTICA  
USANDO GEOESTATÍSTICA E CAPSULE NETWORKS**

ALEX MARTINS SANTOS

Tese aprovada em 08 de maio de 2019.

**Prof. Dr. Anselmo Cardoso de Paiva**  
(Orientador)

**Prof. Dr. Aristófanês Corrêa Silva**  
(Coorientador)

**Prof<sup>ª</sup>. Dr<sup>a</sup>. Aura Conci**  
(Membro da Banca Examinadora)

**Prof. Dr. Rodrigo de Melo Souza Veras**  
(Membro da Banca Examinadora)

**Prof. Dr. Ewaldo Eder Carvalho Santana**  
(Membro da Banca Examinadora)

**Prof. Dr. Stelmo Magalhães Barros Netto**  
(Membro da Banca Examinadora)

**Prof. Dr. João Viana da Fonseca Neto**  
(Membro da Banca Examinadora)

SÃO LUÍS - MA  
2019

*Dedico esta tese em amor e honra aos meus pais Severiano e Marilene.*

## AGRADECIMENTOS

Agradeço em primeiro lugar a Deus por sua bondade que se manifesta em bênçãos inmerecidas e por ter providenciado todas as coisas que concorreram para o desenvolvimento desta tese.

Agradeço ao meu orientador prof. Anselmo Paiva pela orientação, amizade e paciência. Sem sua dedicação este projeto não teria acontecido.

Agradeço a minha esposa Luciana pelo amor, compreensão e por ser fortaleza nos momentos de dificuldade.

Agradeço a minha filha Alexia por seu carinho e por ser fonte de motivação para a conclusão desta tese.

Agradeço aos meus pais Severiano e Marilene, irmãs Andreia e Adriana, meus sogros Oliveira e Guilhermina pelo suporte durante este período.

Agradeço aos meus amigos Daniel Jr., Steve e Thamila pelo companheirismo e colaboração nesta jornada.

Agradeço a meu coorientador prof. Aristófanês pelas ideias e discussões sobre os métodos utilizados.

Agradeço aos membros da banca avaliadora. Suas correções e apontamentos nos permitiram corrigir caminhos e aprimorar a pesquisa.

Agradeço ao IFMA pela liberação e à FAPEMA pelo auxílio concedido para a finalização desta tese.

Agradeço ao corpo docente do PPGEE e seus colaboradores, em especial a Alcides por sua dedicação e apoio.

Agradeço aos meus irmãos em Cristo da Igreja Presbiteriana do Cruzeiro do Anil pelo incentivo e a todos meus colegas do Núcleo de Computação Aplicada.

Por fim, agradeço a todos que indiretamente colaboraram para a realização deste projeto.

*“Os olhos são a lâmpada do corpo. Portanto, se teus olhos forem bons, teu corpo será pleno de luz. (Mateus 6:23)”*

## RESUMO

Nesta tese, propomos um método completamente automático para diagnóstico de degeneração macular relacionada à idade (DMRI), por meio de técnicas de processamento de imagens e aprendizado computacional aplicadas a exames de tomografia de coerência óptica (OCT). A DMRI é uma doença degenerativa da retina considerada como uma das principais causas de cegueira da população idosa mundial. A OCT é um dos principais exames de imagem para a detecção e acompanhamento da DMRI, sendo uma tecnologia de captação de imagens de tecidos biológicos em alta resolução que permite a visualização tridimensional das estruturas internas do olho em escala micrométrica. A avaliação das imagens geradas se dá através da avaliação, por parte do especialista, dos sucessivos cortes seccionais na busca por alterações morfológicas na região macular. Assim, o objetivo principal desta tese é o desenvolvimento e validação de um método de detecção automática de DMRI, a partir de exames de OCT, que possua alta sensibilidade e especificidade para a detecção da doença, diminuindo a necessidade de visualização de grande quantidade de dados pelo especialista para rastreamento e detecção. Os principais problemas abordados neste trabalho são a segmentação da retina total e camada do epitélio pigmentado retiniano, assim como a metodologia para o diagnóstico de DMRI. O método de segmentação proposto é baseado na utilização de redes de aprendizado profundo Capsnet e corte de grafo. O método de diagnóstico automático é baseado em descritores geoestatísticos calculados sobre mapas topográficos da retina. A classificação utiliza a técnica de máquina de vetores de suporte. Testes realizados com 76 volumes de OCT alcançaram resultado de área sob a curva ROC de 0,996.

**Palavras-chave:** DMRI; OCT; Retina; CapsuleNet.

## ABSTRACT

In this thesis, we propose a fully automated method for age-related macular degeneration (AMD) diagnosis based on image processing and computer learning techniques applied to OCT. AMD is a degenerative retinal disease considered to be one of the leading causes of blindness in the world's elderly population. OCT is one of the major imaging tests for the detection and the monitoring of AMD, being a high-resolution biological imaging technology that allows the three-dimensional visualization of the internal structures of the eye on a micrometer scale. The evaluation of the generated images comes about through the specialist's evaluation of the successive image slices in the search for morphological alterations in the macular region. Thus, the main objective of this thesis is the development and validation of an automatic detection of AMD, based on OCT images, which has high sensitivity and specificity for disease detection, reducing the need for large quantity visualization by the specialist for screening and detection. The main issues addressed in this work are the segmentation of the total retina and retinal pigment epithelial layer, as well as the methodology for AMD diagnosis. The proposed segmentation method is based on the use of deep-learning network CapsNet and graph cutting technique. The diagnostic method is based on geostatistical descriptors calculated on topographic maps of the retina. The classification uses support vector machine. The performed tests with 76 volumes of OCT reached an area under the ROC curve value of 0.996.

**Keywords:** AMD; OCT; Retina; CapsuleNet.

## LISTA DE FIGURAS

Figura 1 – Esquema estrutural do olho humano (corte transversal). . . . .	30
Figura 2 – Retinografia de um olho humano saudável com destaque para o disco óptico e a mácula. . . . .	31
Figura 3 – Corte histológico da região macular da retina humana com identificação das camadas. . . . .	32
Figura 4 – Ilustração da percepção de imagem por paciente com DMRI, comparada à visão normal. . . . .	33
Figura 5 – Ilustração comparativa entre uma retina saudável (A), retina acometida por DMRI não-exsudativa com formação de drusas (B) e retina com neovascularização (C). . . . .	34
Figura 6 – Diagrama básico de um equipamento de OCT com destaque para os principais elementos presentes. . . . .	36
Figura 7 – Esquema de aquisição de um tomógrafo de coerência óptica no domínio espectral. . . . .	37
Figura 8 – Aquisição de SD-OCT. À esquerda, aquisição unidimensional ( <i>A-scan</i> ). Na imagem central, a imagem bidimensional ( <i>B-scan</i> ). Na imagem à direita, temos a aquisição 3D. . . . .	37
Figura 9 – <i>B-scan</i> da região central da mácula obtida pela tecnologia SD-OCT. . . . .	38
Figura 10 – Imagem de OCT com destaque para o ruído característico inserido pelo processo de aquisição (ampliação no canto inferior direito). . . . .	40
Figura 11 – Redução de ruídos em recorte de uma <i>B-scan</i> de OCT com o filtro wavelet para $\sigma = \{0.05, 0.10, 0.125\}$ . . . . .	42
Figura 12 – Esquema da arquitetura de uma rede CapsNet para identificação de algarismos. . . . .	45
Figura 13 – Roteamento dinâmico para envio de informações da cápsula de nível inferior para a cápsula de nível superior. . . . .	47
Figura 14 – Exemplo de imagem representada por um grafo. À esquerda uma imagem representada por seus <i>pixels</i> . À direita, a mesma imagem representada como um grafo. A marcação em verde corresponde a um caminho percorrido na imagem que a segmenta em duas regiões. . . . .	48
Figura 15 – Curva característica do semivariograma com destaque para o patamar, alcance e pepita. . . . .	50
Figura 16 – Parâmetros utilizados para o cálculo do semivariograma (2D). . . . .	50
Figura 17 – Transformação do espaço amostral (a) para o espaço multidimensional (b) de características através de uma função de mapeamento $\phi$ e separação das classes por um hiperplano. . . . .	52

Figura 18 – A curva ROC (vermelho) representando a relação entre a sensibilidade e a especificidade do classificador. . . . .	58
Figura 19 – Arquitetura da rede baseada em CapsNet para utilizada para a classificação de fragmentos de imagens de OCT. . . . .	60
Figura 20 – Fluxo do processo de treinamento da CapsNet para o reconhecimento de fragmentos de imagens de OCT. . . . .	61
Figura 21 – Fluxo do processo de segmentação das bordas da MLI, EPR e MB com a utilização de CapsNet. . . . .	62
Figura 22 – Sequência de processamento definição da região de interesse. A) imagem original; B) diminuição de ruídos (filtro wavelet); C) filtro gaussiano; D) binarização; E) dilatação e fechamento e F) região de interesse delimitada. . . . .	63
Figura 23 – Mapas de probabilidade para as bordas das camadas retinianas: 1) borda interna da MLI; 2) borda interna do EPR; 3) borda externa da MB e 4) região de "fundo". . . . .	64
Figura 24 – Aplicação de restrições aos mapas de probabilidade da borda do EPR (à esquerda) e da borda da MB (à direita). . . . .	66
Figura 25 – Bordas obtidas após o corte de grafos (curvas vermelhas). . . . .	67
Figura 26 – Delimitação de região de busca a resolução do problema de segmentação da borda da MB. . . . .	68
Figura 27 – Fluxograma básico da metodologia proposta para a classificação de imagens de OCT. . . . .	69
Figura 28 – Processo de construção do mapa topográfico para a retina total. As partes da imagem referentes ao fundo e às demais camadas são removidas e é calculada a espessura para cada ponto. . . . .	71
Figura 29 – Bordas utilizadas no processo de geração dos mapas topográficos. À esquerda, representação tridimensional. À direita, marcação das bordas em uma <i>B-scan</i> . . . . .	71
Figura 30 – Reconstrução <i>en face</i> da borda do epitélio pigmentado para um volume de SD-OCT sem DMRI (A) e um volume com DMRI (B). Visualmente é possível perceber pequenas variações na textura. . . . .	72
Figura 31 – Semivariograma experimental calculado sobre os mapas topográficos da retina humana com ângulo de direção 45°. . . . .	73
Figura 32 – <i>B-scan</i> 19 do exame “AMD_1081” da Base A com baixa qualidade de aquisição. . . . .	80
Figura 33 – A) <i>B-scan</i> 50 do exame “Controle_1055 da Base A”. B) Imagem com marcações das bordas da base (azul) e nosso método (vermelho). Mapa topográfico da RT: C) nosso método, D) com marcações da base. Mapa topográfico do EPR: E) nosso método, F) com marcações da base. . . . .	84

Figura 34 – A) <i>B-scan</i> 39 do exame “AMD_1131 da Base A”. B) Imagem com marcações das bordas da base (azul) e nosso método (vermelho). Mapa topográfico da RT: C) nosso método, D) com marcações da base. Mapa topográfico do EPR: E) nosso método, F) com marcações da base. . . . .	85
Figura 35 – A) <i>B-scan</i> 52 do exame “AMD_1034” da Base A com destaque para área com atrofia geográfica. B) Imagem com marcações das bordas da base (azul) e nosso método (vermelho). Mapa topográfico da RT: C) nosso método, D) com marcações da base. Mapa topográfico do EPR: E) nosso método, F) com marcações da base. . . . .	86
Figura 36 – Gráficos dos descritores semivariograma e semimadograma para mapas topográficos do EPR-CD. As linhas A e B apresentam exames de olhos com DMRI não-exsudativa, enquanto C e D apresentam exames considerados normais. . . . .	88
Figura 37 – Curva ROC obtida para cada experimento. O melhor valor obtido foi 0,989. . . . .	90
Figura 38 – Comparação da curva ROC do desempenho do semivariograma sobre o mapa EPR-CD em relação a outros métodos. O melhor valor obtido foi 0,989. . . . .	92
Figura 39 – Curva ROC do desempenho do semivariograma sobre o mapa EPR-CD de volumes segmentados com a metodologia proposta nesta tese. . . . .	92
Figura 40 – Máscara obtida a partir das marcações existentes na base. . . . .	105
Figura 41 – Valores encontrados de SNR, CNR e PSNR para a filtragem <i>wavelet</i> com parâmetro $\sigma \in [0, 8; 0, 15]$ . . . . .	106

## LISTA DE TABELAS

Tabela 1 – Comparação de trabalhos relacionados. . . . .	27
Tabela 1 – Comparação de trabalhos relacionados (continuação) . . . . .	28
Tabela 1 – Comparação de trabalhos relacionados (continuação) . . . . .	29
Tabela 2 – Funções núcleo utilizadas para separar o espaço de características. . . . .	53
Tabela 3 – Interpretação do valor de Kappa . . . . .	57
Tabela 4 – Qualidade do diagnóstico de acordo com a área sob a curva ROC. . . . .	58
Tabela 5 – Caracterização das imagens utilizadas na Base A . . . . .	77
Tabela 6 – Avaliação do erro de segmentação automática da base A em baixa resolução em comparação à segmentação fornecida por Farsiu et al. (2014). . . . .	79
Tabela 7 – Resultados da segmentação automática da base A em alta resolução comparada à segmentação da base. . . . .	80
Tabela 8 – Tempo de processamento por <i>B-scan</i> em segundos para a segmentação das bordas da retina. . . . .	81
Tabela 9 – Resultados da segmentação automática da base B vs. marcação manual do Especialista 1. . . . .	81
Tabela 10 – Resultados da segmentação automática da base B vs. marcação manual do Especialista 2. . . . .	82
Tabela 11 – Comparação entre marcações efetuadas por dois especialistas. . . . .	82
Tabela 12 – Comparação de erro de posicionamento de borda com outros métodos de segmentação baseados em aprendizado profundo. . . . .	83
Tabela 13 – Experimentos realizados para o diagnóstico automático de DMRI. . . . .	87
Tabela 14 – Resultados dos experimentos de diagnóstico de DMRI utilizando descritores geoestatísticos. . . . .	89
Tabela 15 – Comparação com outras técnicas utilizando o mesmo conjunto de exames. . . . .	91
Tabela 16 – Comparação de desempenho de diagnóstico de DMRI com os principais trabalhos relacionados. . . . .	93
Tabela 17 – Comparação das médias de SNR, CNR e PSNR antes e após a filtragem por <i>wavelets</i> . O desvio padrão é apresentado entre parênteses. . . . .	106

## LISTA DE ABREVIATURAS E SIGLAS

AUC	<i>Area under curve</i>
CNN	<i>Convolutional neural network</i>
CNR	<i>Contraste to noise ratio</i>
DMRI	Degeneração macular relacionada à idade
EM	Edema macular
EPR	Epitélio pigmentado retiniano
EPR-CD	Epitélio pigmentado da retina e complexo de drusas
FD-OCT	<i>Frequency domain optical coherence tomography</i>
FN	Falso negativo
FP	Falso positivo
LBP	<i>Local binary pattern</i>
MB	Membrana de Bruch
MLI	Membrana limitante interna
OCT	<i>Optical coherence tomography</i>
PCA	<i>Principal component analysis</i>
PDI	Processamento digital de imagens
PSNR	<i>Peak signal-to-noise ratio</i>
RN	Retina neurosenial
ROC	<i>Receiver operating characteristic</i>
SD-OCT	<i>Spectral domain optical coherence tomography</i>
SNR	<i>Signal to noise ratio</i>
SS-OCT	<i>Swept source optical coherence tomography</i>
SVM	<i>Support vector machine</i>
TD-OCT	<i>Time domain optical coherence tomography</i>
VN	Verdadeiro negativo
VP	Verdadeiro positivo

# SUMÁRIO

<b>1</b>	<b>Introdução</b>	<b>15</b>
1.1	Objetivos	17
1.2	Contribuições	17
1.3	Organização da Tese	18
<b>2</b>	<b>Trabalhos Relacionados</b>	<b>19</b>
2.1	Segmentação da Retina	19
2.2	Diagnóstico de DMRI	23
<b>3</b>	<b>Fundamentação Teórica</b>	<b>30</b>
3.1	O Olho Humano e a Retina	30
3.2	Degeneração Macular Relacionada à Idade	32
3.2.1	Fisiopatologia da DMRI	33
3.2.2	Diagnóstico e Acompanhamento	34
3.3	Tomografia de Coerência Óptica - OCT	35
3.4	Processamento digital de imagens	39
3.5	Eliminação de ruído em imagens	39
3.5.1	Filtro da média	40
3.5.2	Filtragem gaussiana	40
3.5.3	Supressão de ruído com <i>Wavelet</i>	41
3.5.4	Avaliação da redução de ruído	43
3.6	<i>Capsule Networks</i>	44
3.7	Segmentação por corte de grafo	47
3.8	Geoestatística	48
3.8.1	Semivariograma	49
3.8.2	Semimadograma	50
3.9	Máquina de Vetores de Suporte	51
3.10	Validação de Resultados	53
3.10.1	Avaliação de erro médio	54
3.10.2	Distância de Hausdorff	55
3.10.3	Validação do Classificador	55
3.10.3.1	Sensibilidade, Especificidade e Acurácia	56
3.10.3.2	Concordância Kappa de Cohen	57
3.10.4	Curva característica de operação do receptor (ROC)	57
<b>4</b>	<b>Método para a segmentação de camadas da retina</b>	<b>59</b>
4.1	Arquitetura da rede de cápsulas para a segmentação	59
4.2	Treinamento	61
4.3	Segmentação das bordas	62
4.3.1	Definição automática da região de interesse	62

4.3.2	Extração de fragmentos . . . . .	64
4.3.3	Classificação . . . . .	64
4.3.4	Corte de Grafo . . . . .	65
4.3.5	Estratégia para segmentação em duas resoluções . . . . .	66
<b>5</b>	<b>Diagnóstico de DMRI utilizando descritores geoestatísticos . . . . .</b>	<b>69</b>
5.1	Aquisição das imagens . . . . .	69
5.2	Representação . . . . .	70
5.2.1	Mapa topográfico . . . . .	70
5.2.2	Reconstrução “ <i>en face</i> ” da borda do EPR . . . . .	70
5.3	Extração de características . . . . .	72
5.4	Classificação utilizando Máquina de Vetores de Suporte (SVM) . . . . .	72
<b>6</b>	<b>Resultados e Discussão . . . . .</b>	<b>75</b>
6.1	Base de imagens de OCT . . . . .	75
6.1.1	Base A . . . . .	75
6.1.2	Base B . . . . .	77
6.2	Recursos de <i>Hardware</i> e <i>Software</i> . . . . .	77
6.3	Resultados da Segmentação da Retina . . . . .	78
6.4	Resultados do diagnóstico de DMRI não-exsudativa . . . . .	84
6.4.1	Determinação do modelo de classificação . . . . .	86
6.4.2	Comparação com descritores de textura clássicos . . . . .	91
6.4.3	Teste com o método de segmentação proposto . . . . .	91
6.4.4	Comparação com outros trabalhos . . . . .	93
<b>7</b>	<b>Conclusão . . . . .</b>	<b>95</b>
	<b>Referências . . . . .</b>	<b>98</b>
	<b>APÊNDICE A Parâmetro <math>\sigma</math> para pré-processamento . . . . .</b>	<b>105</b>

## 1 INTRODUÇÃO

A degeneração macular relacionada à idade (DMRI) é uma doença ocular degenerativa que danifica progressivamente a mácula, região central da retina onde a imagem é focalizada (LEEUEWEN et al., 2003). Os danos causados provocam a incapacidade de leitura, reconhecimento de faces e de objetos, além de provocar dificuldades na percepção de expressões faciais. Sua progressão pode levar à cegueira total e irreversível.

A DMRI é apontada como a principal causa de danos permanentes à visão da população idosa em países industrializados (EVANS; SYED, 2013). A idade avançada é o principal fator de risco, com incidência e gravidade aumentando conforme aumenta a idade do paciente (SOON-CHUN, 2015).

Wong et al. (2014) projetam que um total de 196 milhões de pessoas no mundo terão DMRI em 2020. No Brasil, estudos epidemiológicos são escassos. Porém, estima-se com base em dados locais e de outros países, uma prevalência de 2,2% na faixa entre 70-79 anos e de até 10,3% em indivíduos com 80 ou mais anos (BRASIL, 2018). Ainda segundo Wong et al. (2014), mundialmente, a expectativa para 2040 é que 288 milhões de pessoas serão acometidas pela doença.

A DMRI se manifesta em duas formas, a forma não-exsudativa (ou seca) ou exsudativa (também conhecida como úmida ou neovascular). Cerca de 85% de todos os casos de DMRI são manifestações do tipo não-exsudativa (DUKER; WAHEED; GOLDMAN, 2015). Nessa forma de degeneração, é comum o aparecimento de drusas (depósitos sub-retinianos amarelados de várias formas e tamanhos) e atrofia nas retinas acometidas. Já a DMRI do tipo úmida (ou exsudativa) é caracterizada principalmente pela neovascularização. O rompimento dos neovasos causa o extravasamento de sangue e líquido dentro da mácula, provocando a morte das células fotorreceptoras (AMBATI et al., 2003). A DMRI exsudativa é praticamente sempre acompanhada das manifestações clínicas do tipo seca (DUKER; WAHEED; GOLDMAN, 2015).

O diagnóstico clínico da DMRI é comumente feito pelo exame de fundo de olho com a técnica de biomicroscopia de mácula (BELLINI, 2008) e retinografia (fotografia do fundo do olho). Esses exames permitem a avaliação da formação de drusas e outras anormalidades na retina. As medições destes achados servem como indicadores da severidade da doença e da possibilidade de progressão para a forma úmida. O Ministério da Saúde recomenda além da biomicroscopia, outros exames como a angiofluoresceinografia (AF) e a tomografia

de coerência óptica (OCT, do inglês *optical coherence tomography*)(BRASIL, 2018).

A OCT é uma tecnologia não-invasiva de captação de imagens que permite a visualização de tecidos biológicos *in vivo* em cortes seccionais com alta resolução. A captação da imagem é feita através de sondagem por radiação luminosa inofensiva de forma indolor e confortável para o paciente (FREITAS; AMARAL; RAELE, 2014).

A tecnologia empregada na OCT é baseada na análise do retroespalhamento de um sinal luminoso emitido através de meios transparentes ou semitransparentes, como tecidos biológicos (DUKER; WAHEED; GOLDMAN, 2015). A interferência provocada pelo sinal refletido em um sinal de referência é utilizada para gerar uma imagem da profundidade (*A-scan*) da retina em um único ponto. A combinação de *A-scans* de pontos consecutivos da retina em uma direção gera uma imagem seccional chamada *B-scan* ou “linha de varredura”. O agrupamento de sucessivas *B-scans* fornece uma representação volumétrica da região amostrada.

Imagens de OCT da região macular revelam em detalhes as duas principais camadas da retina: o epitélio pigmentado retiniano (EPR) e a retina neurosensorial (RN). O EPR é uma camada de células cuboides localizado externamente à RN. Esta, por sua vez, é formada por neurônios e células estruturais que transformam o estímulo luminoso em impulsos elétricos para levar as informações ao córtex cerebral.

Especificamente para o diagnóstico da DMRI, a imagem de OCT permite a avaliação de alterações no EPR, drusas moles e deslocamentos da RN (HEE et al., 1996). Em relação aos outros exames recomendados, foi evidenciado que a OCT tem maior sensibilidade à detecção das alterações patológicas relacionadas à DMRI. Foi verificado que 57% dos olhos avaliados apresentavam alterações percebidas somente através da OCT e não visualizadas através de outros exames, como a angiofluoresceinografia (ETER; SPAIDE, 2005). Outra vantagem em relação à AF é a rápida aquisição, dispensando a utilização de contrastes e, assim, os riscos de reações alérgicas.

A avaliação das alterações causadas por DMRI, a partir de imagens de OCT, se dá através da avaliação dos sucessivos cortes seccionais (*B-scans*) na busca por alterações morfológicas provocadas no EPR. Por se tratar de um método de alta resolução, a OCT acaba por gerar um grande número de *B-scans*. Um escaneamento macular volumétrico tipicamente produz cerca de 100 imagens de  $512 \times 1.000$  *pixels* (protocolo de aquisição rápida).

A quantidade de imagens geradas submete o especialista à necessidade de visualizar

e avaliar uma grande quantidade de fatias do volume a procura de alterações, especialmente se elas são mínimas e esparsas nos estágios iniciais da manifestação. Se considerarmos ainda que um médico é responsável pelo atendimento e acompanhamento de dezenas de pacientes, verifica-se a necessidade de tecnologias que apoiem a prática clínica da análise de imagens de OCT. Assim, o desenvolvimento de uma metodologia para o diagnóstico automático de DMRI não-exsudativo se constitui no problema abordado nesta tese.

## 1.1 Objetivos

O objetivo desta tese é o desenvolvimento e validação de uma metodologia automática, que possua alta sensibilidade e especificidade para o diagnóstico de DMRI não-exsudativa, a partir de exames de OCT volumétricos da região macular.

A solução proposta para atingir o objetivo se dá através da utilização de descritores geoestatísticos de textura extraídos a partir da segmentação do EPR. Uma base de exames, contendo segmentação precisa da retina e diagnóstico quanto à presença de DMRI revisada por especialistas, é utilizada para a validação da metodologia.

De maneira específica, pretendemos:

- Propor um método para segmentação da imagem de OCT de forma a delimitar a retina total (RT) e camadas de interesse para a detecção da DMRI (RN e o EPR);
- Propor a utilização de descritores de textura geoestatísticos para a caracterização de mapas topográficos (de espessura), gerados a partir das camadas segmentadas e, com isso realizar o diagnóstico automático de DMRI;
- Comparar o desempenho obtido pela utilização de descritores geoestatísticos aplicados sobre a RT, RN e EPR; e
- Validar a metodologia proposta, utilizando base de exames segmentada e classificada por especialistas.

## 1.2 Contribuições

As principais contribuições desta tese estão relacionadas ao método de segmentação proposto para as camadas da retina, e à extração de descritores baseados em descritores

geoestatísticos para o diagnóstico de DMRI a partir dos mapas topográficos da retina. Mais especificamente, propomos:

- Um novo método para a segmentação do EPR em imagens de OCT baseado no algoritmo de aprendizado profundo *Capsule Nets* e busca em grafos;
- A utilização descritores de textura geoestatísticos para o diagnóstico de DMRI em imagens de OCT baseados na classificação de textura do mapa topográfica das camadas da retina e da reconstrução *en face* da borda do epitélio pigmentado retinal;

### 1.3 Organização da Tese

O restante desta tese está dividido em mais seis capítulos. O Capítulo 2 alista os principais trabalhos relacionados com tema desta tese. O Capítulo 3 apresenta conceitos teóricos que sustentam a metodologia proposta e que são empregados nesta pesquisa. O Capítulo 4 apresenta o método desenvolvido para a segmentação automática das camadas retinianas e o Capítulo 5 descreve a metodologia para utilização de descritores geoestatísticos aplicados ao diagnóstico de DMRI. O Capítulo 6 apresenta os testes realizados e discute seus resultados. Por fim, o Capítulo 7 apresenta as conclusões desta pesquisa, suas contribuições, limitações e propostas para trabalhos futuros.

## 2 TRABALHOS RELACIONADOS

Nas últimas duas décadas, um número significativo de métodos para análise automática da retina humana em exames de OCT tem sido desenvolvido. As técnicas apresentadas se aplicam principalmente aos problemas da segmentação da retina e diagnóstico de doenças oculares.

A segmentação da retina na imagem de OCT foi o primeiro problema enfrentado no processo de análise dessas imagens (HUANG et al., 1991; HEE et al., 1995b). Como a retina é uma estrutura multicamadas, a identificação e separação de suas camadas são importantes para análise anatômica detalhada. A segmentação também é importante para o diagnóstico automatizado de DMRI não-exsudativa. Para isso, pode ser feita simplesmente uma separação da retina do fundo da imagem (para métodos que analisam globalmente a *B-scan*) ou o enfrentamento mais complexo da segmentação do EPR. Essa é a camada que sofre as maiores alterações provocadas pela DMRI não-exsudativa, sendo importante para o diagnóstico automatizado.

O problema de diagnosticar a DMRI de forma automática, por sua vez, tem sido enfrentado com técnicas de reconhecimento de padrões tradicionais e métodos baseados em redes neurais de aprendizado profundo. A seguir, apresentaremos os principais trabalhos desenvolvidos relacionados à proposta desta tese de desenvolver e avaliar métodos para segmentação da retina e diagnóstico automático de DMRI não-exsudativa em imagens de OCT.

### 2.1 Segmentação da Retina

A segmentação das camadas da retina é tarefa desafiadora, dada a complexidade da imagem e a grande quantidade de ruídos presentes devido à natureza de sua aquisição. Comumente, essa tarefa envolve um pré-processamento com vistas à eliminação do ruído e evidenciação das bordas para sua posterior segmentação.

Os trabalhos pioneiros de segmentação desenvolvidos foram apresentados na década de 1990 (HEE et al., 1995a; HEE et al., 1995b). Esses trabalhos são baseados nas variações de intensidade da imagem para a medição da espessura da retina total e também a medição apenas da camada de fibras nervosas, camada de interesse para o diagnóstico de glaucoma.

Esses trabalhos utilizaram um operador convolucional unidimensional sobre cada coluna da imagem de forma a detectar arestas que separam as camadas. Após filtragem das arestas encontradas, 90% das bordas das subcamadas foram identificadas.

Abordagem semelhante foi utilizada por Huang et al. (1998). Os autores efetuaram medições da espessura da retina total e a segmentação conjunta das camadas de fotoreceptores, epitélio pigmentado e coróide para estudo sobre degeneração retiniana em humanos e animais. Para a detecção das bordas, foi utilizada uma suavização gaussiana, cálculo da primeira derivada e detecção de amplitude.

No trabalho de Fernández, Salinas e Puliafito (2005) é apresentado um método de busca das bordas das camadas através da informação da coerência local da imagem. As imagens foram analisadas utilizando uma combinação de tensor estrutural com filtro de difusão complexa para mitigar a presença de ruído na imagem. Um algoritmo de detecção de picos é utilizado em cada coluna (*A-scan*) para a detecção das bordas. O método foi testado com 72 volumes obtidos a partir de voluntários saudáveis e os autores apresentaram casos de sucesso e falha na detecção. Os resultados são discutidos, no entanto não são comparados a outras técnicas.

O método proposto por Farsiu et al. (2008) enfrenta o problema da segmentação de imagens de OCT de pacientes com DMRI com manifestação de drusas. Como primeiro passo, é aplicado à imagem um filtro baseado no gradiente de Xu-Prince (XU; PRINCE, 1997) com o objetivo de diminuição dos ruídos. Um recorte a partir da localização da camada de fibras nervosas reduz a área de processamento. O método desenvolvido é semi-automatizado e delimita as bordas do EPR com drusas através de modelos deformáveis. Este trabalho não apresentou resultados quantitativos nem comparação com outros métodos.

A abordagem feita por Garvin et al. (2009) foi baseada em corte de grafos tridimensionais cujos parâmetros foram estimados a partir de um conjunto de 13 volumes de OCT. Os resultados obtidos são semelhantes à variação de medidas encontradas entre segmentações manuais feitas por especialistas diferentes. Foi encontrado erro médio de determinação das bordas de  $5,69 \mu m$  com desvio padrão de  $2,41 \mu m$ .

O método desenvolvido por Chiu et al. (2010) também se baseia em corte de grafo e técnicas de programação dinâmica para contornar condições específicas da imagem de posicionamento e região de captura. A segmentação é feita a partir de imagens bidimensionais. Inicialmente, são calculadas as derivadas parciais das imagens e em seguida efetuada a diminuição de ruídos através de um filtro de média retangular. A partir da imagem

resultante, é construído um grafo com os valores dos *pixels* como pesos. A pesquisa por bordas é realizada sequencialmente e são detectadas oito camadas. O método foi validado com uma base contendo dez volumes de pessoas consideradas saudáveis segmentados por dois especialistas. Considerando todas as camadas, o algoritmo apresentou uma diferença média de 0,95 *pixels*, um valor menor até mesmo que a diferença entre a segmentação efetuada por dois especialistas. Posteriormente, o método é modificado para segmentação de imagens de OCT com DMRI<sup>1</sup> (CHIU et al., 2012). Esses trabalhos, em conjunto com o desenvolvido por Farsiu et al. (2008), são referentes ao desenvolvimento do software DOC-TRAP (*Duke OCT Retinal Analysis Program*) projetado para a utilização em pesquisas sobre segmentação em OCT.

Alonso-Caneiro, Read e Collins (2013) utilizaram uma técnica baseada em corte de grafos, de maneira semelhante a Chiu et al. (2010), com filtros específicos para evidenciação da coroide humana. Os autores destacam o interesse clínico dessa aplicação como a detecção de doenças e erros refrativos. No método proposto, cada *B-Scan* é pré-processado com filtro de média retangular de forma a evidenciar a borda interna da coroide (BIC) e a borda externa, ao mesmo tempo em que são minimizadas as presenças de artefatos e ruídos. Após essa etapa, um filtro simples de detecção de arestas na direção vertical é utilizado para a criação do mapa de pesos para a etapa de segmentação por grafo. Foram processadas 1083 *B-scan* obtidos de olhos de crianças e 90 *B-scan* de pacientes adultos. Os resultados alcançados foram de 0,79 e 0,82 *pixels* de erro médio absoluto para BIC de olhos infantis e BIC de olhos adultos, respectivamente. Já para a segmentação da borda externa, os resultados encontrados foram de 3,26 e 4,23 *pixels* de erro médio absoluto para crianças e adultos, respectivamente.

Antony et al. (2013) apresentaram uma metodologia que utiliza técnicas de aprendizado de máquina em conjunto com a segmentação por corte de grafos. A imagem é processada através de um conjunto de filtros. Os autores utilizaram filtros como descritores de Haar, Gabor, filtros gaussianos direcionais além dos descritores de intensidade clássicos, tais como intensidade média, variância, entropia, obliquidade e curtose. Esses descritores são utilizados para treinamento de um classificador do tipo Florestas Aleatórias que, por sua vez, determina um mapa de pesos a serem passados para a próxima etapa de segmentação por corte de grafos. O método foi avaliado a partir de imagens de pessoas, ratos

<sup>1</sup> Os autores disponibilizam a base de imagens utilizadas no endereço eletrônico: <[http://www.duke.edu/sf59/Chiu\\_IOVS\\_2011\\_dataset.htm](http://www.duke.edu/sf59/Chiu_IOVS_2011_dataset.htm)>

e cães com resultados favoráveis. A metodologia alcançou erros médios de segmentação de  $6,45 \pm 1,87\mu m$ ,  $3,35 \pm 0,62\mu m$  e  $9,75 \pm 3,18\mu m$ . Os autores demonstraram que a combinação de técnicas de aprendizado de máquina com o corte de grafos representa uma vantagem por essa abordagem possuir maior flexibilidade de adaptação a outros modelos e imagens de várias espécies.

No trabalho de Srinivasan et al. (2014a), temos a apresentação de uma abordagem para a segmentação da retina de roedores com a utilização de um filtro esparsos de ruídos, máquinas de vetores de suporte (SVM) e corte de grafos. Nesse trabalho, os autores partem do *framework* desenvolvido por Chiu et al. (2010) e conseguem segmentar até dez camadas das retinas desses animais. As amostras foram divididas em dois grupos. No primeiro grupo, as retinas apresentavam todas as camadas (Grupo A) e no segundo, as retinas possuíam apenas oito camadas devido a alterações na retina. A segmentação apresentou erro médio absoluto de  $1,90\mu m$  para um grupo de pacientes e  $2,17\mu m$  para outro grupo, em relação às marcações efetuadas por um especialista experiente através do software DOCTRAP.

Fang et al. (2017a) apresentaram um *framework* que combina redes neurais convolucionais (CNN, do inglês *convolutional neural networks*) com algoritmos de busca em grafos visando segmentar nove bordas da retina em imagens de OCT. Primeiramente, a CNN é treinada com as características das bordas das camadas de forma a estimar o posicionamento das bordas das oito camadas. Em seguida, esses valores são submetidos ao algoritmo de busca em grafo para a definição final das bordas. O método foi validado com 60 volumes (2915 *B-Scans*) de vinte olhos de pessoas com DMRI. Os resultados encontrados foram comparados aos resultados obtidos pelos softwares de segmentação DOCTRAP e OCTExplorer. Os resultados encontrados são considerados satisfatórios, sendo superiores aos obtidos pelo software OCTExplorer porém ainda inferiores aos obtidos pelo software DOCTRAP.

Proposto por Roy et al. (2017), o ReLayNet utilizou uma arquitetura de rede neural totalmente convolucional (FCN, do inglês *Fully-convolutional networks*) para realizar a segmentação semântica das camadas da retina e do fluido intrarretiniano em imagens de OCT macular, concluindo que a ReLayNet pode segmentar de forma confiável, mesmo na presença de um alto grau de patologia que afeta severamente a estrutura normal da retina. Redes totalmente convolucionais (XU et al., 2017) também foram utilizadas, em uma estrutura de aprendizado profundo em dois estágios para a segmentação do descolamento

do EPR.

Kugelman et al. (2018a) propõem o uso de uma rede neural recorrente treinada (RNN, do inglês *Recurrent Neural network*) como um classificador baseado em fragmentos para segmentar sete limites da camada retiniana em crianças saudáveis e três bordas da camada retiniana de pacientes com DMRI em imagens de OCT. Os resultados indicaram que a arquitetura RNN é uma alternativa viável à CNN para tarefas de classificação de imagens que exibem estrutura sequencial clara. Comparado à CNN, a RNN mostrou uma acurácia média geral ligeiramente superior.

Um problema comum enfrentado pelos métodos apresentados nesta seção é o tratamento e a diminuição dos ruídos nas imagens de OCT. Diversas técnicas são propostas como a utilização de filtro passa-baixa simples (HEE et al., 1995b; FARSIU et al., 2008), filtro não linear de difusão anisotrópica (FERNÁNDEZ; SALINAS; PULIAFITO, 2005). Já o método apresentado por Chiu et al. (2010) utiliza filtragem gaussiana com núcleo retangular.

Nem todos os trabalhos pesquisados apresentam dados quantitativos sobre seus resultados. As bases de imagens utilizadas possuem quantidades de arquivos diferentes e muitas apresentam condições específicas de aquisição. A maioria das pesquisas não torna pública as imagens utilizadas.

As condições específicas das imagens utilizadas, presença de doenças, protocolo de aquisição, quantidade de fatias por volume e objetivo do procedimento determinam abordagens diferentes. Essa diversidade de bases e condições dificultam a comparação dos resultados de forma direta. Quanto aos trabalhos que apresentam resultados quantitativos (ALONSO-CANEIRO; READ; COLLINS, 2013; CHIU et al., 2010; FANG et al., 2017a), os erros médios encontrados são próximos aos valores encontrados pela comparação de segmentações manuais feitas entre especialistas distintos.

## 2.2 Diagnóstico de DMRI

O desenvolvimento de métodos automáticos para o diagnóstico de DMRI, utilizando imagens de OCT, envolve a definição de descritores relevantes para a caracterização das imagens, o emprego de classificadores e a validação do método quanto à geração de resultados confiáveis a partir de uma base de imagens devidamente classificada.

O método desenvolvido por Liu et al. (2011) utilizou uma única fatia extraída

da região central da mácula. As *B-scans* foram processadas com filtro Canny, pirâmides espaciais multirresolução e padrões binários locais (LBP, do inglês *Local Binary Pattern*). Após o processamento de características, os descritores foram submetidos à redução de dimensionalidade por Análise de Componentes Principais (PCA, do inglês *Principal Component Analysis*) e classificação por máquina de vetores de suporte (SVM, do inglês, *support vector machine*). A base utilizada continha 326 imagens para treinamento e 131 para testes. As imagens foram devidamente diagnosticadas por especialistas. A avaliação foi feita a partir do cálculo da área sob a curva característica de operação do receptor (ROC, do inglês *Receiver Operating Characteristic*) que resultou no valor de 0,975 para a classificação. No entanto, esse método apresenta como limitações a necessidade de uma seleção manual de uma fatia centralizada na fóvea com uma boa qualidade de sinal. Apenas uma imagem é utilizada para a detecção de doenças no olho inteiro. Outro trabalho que utilizou LBP como descritor para o diagnóstico de DMRI foi desenvolvido por Albarrak et al. (2013). Nesse trabalho, foi realizado o processamento de 140 volumes OCT dos quais 72 eram previamente diagnosticados com DMRI e o restante utilizado como amostras de controle. Uma etapa de pré-processamento, utilizando o algoritmo de variação total de Bregman (VT), foi empregada para descartar as partes da imagem que não correspondem à retina (fundo). Após a extração de características, as instâncias são submetidas a um classificador de rede bayesiana. Como resultado, o método alcançou acurácia de 91,4%, sensibilidade de 92,4%, 90,5% de especificidade correspondendo a uma área sob a curva ROC igual a 0,944.

Já Serrano-Aguilar et al. (2012) utilizaram descritores Haar-like e de textura clássica para representar as características de 200 imagens de OCT. Uma etapa de normalização foi utilizada como pré-processamento para ajustar a intensidade dos *pixels* e contornar erros de registro e alinhamento. A classificação efetuada por árvores de decisão resultou em uma sensibilidade de 96% e especificidade de 92%.

Srinivasan et al. (2014b) apresentaram um método para diferenciação de edema macular (EM), DMRI e olhos normais através do processamento de 45 imagens de OCT volumétricas. Foram utilizados como descritores histogramas de gradiente orientado (HOG, do inglês *Histogram of oriented gradient*) em multirresolução. Antes, as imagens foram pré-processadas pelo algoritmo BM3D e achatadas para reduzir a variação entre pacientes. As imagens são classificadas com SVM, alcançando sensibilidade de 100% para as amostras com DMRI e EM. A sensibilidade para volumes normais foi de 86,7% .

No estudo conduzido por Farsiu et al. (2014), 384 volumes de OCT (269 com DMRI) foram segmentados de forma semiautomática. As bordas encontradas delimitaram a retina total e camada do epitélio pigmentado com complexo de drusas (EPR-CD) para análise de espessura. Eles avaliaram tanto a espessura como o volume do EPR e efetuaram a classificação com regressão logística. Esse experimento resultou em uma área sob a curva ROC de 0,99. A mesma base de imagens foi utilizada por Venhuizen et al. (2015), que realizou classificação de forma independente de segmentação, através da clusterização não supervisionada de fragmentos das imagens. Após a determinação de vetor de características pelo histograma da incidência desses fragmentos, a classificação por Florestas Aleatórias permitiu alcançar área sob a curva ROC igual a 0,984.

Wang et al. (2016) desenvolveram um método para diferenciar DMRI, EM e máculas saudáveis utilizando padrões de configuração linear (LCP, do inglês *linear configuration pattern*). Um classificador de otimização mínima sequencial foi utilizado alcançando a acurácia de 99,3% para as três classes. Especificamente para o diagnóstico de DMRI, a acurácia foi de 93,3% com AUCROC de 0,995. Os testes foram aplicados sobre as mesmas imagens utilizadas por Srinivasan et al. (2014b).

A abordagem descrita por Sun, Li e Sun (2017) fragmentou as *B-scans* em regiões pequenas e utilizou descritores SIFT (LOWE, 1999) com classificador SVM (SUN; LI; SUN, 2017). Os autores reportaram encontrar valores de acurácia de 95,4%, sensibilidade de 92,2% e especificidade de 96,9%.

Fang et al. (2017b) apresentaram um método automático usando método de aprendizado profundo PCANet-CK (do inglês *principal component analysis network with composite kernel*). Para o conjunto de imagens disponibilizado por Srinivasan et al. (2014b), foi alcançado 100% de acurácia para a detecção de DMRI. Para a outra base utilizada nesse trabalho (base HUCM), foi encontrada acurácia de 95,4%, sensibilidade de 92,2% e especificidade de 96,9%.

Lee, Baughman e Lee (2017) utilizaram cerca de 100.000 *B-scans* para treinar e validar uma rede CNN. O resultado alcançado foi de uma acurácia de 88,98%, sensibilidade de 85,41% e especificidade de 93,82%. A área sob a curva ROC encontrada foi de 0,938.

Outra estratégia utilizando CNN é o retreinamento de uma rede construída para um problema de classificação de outro domínio. Essa estratégia é conhecida como transferência de aprendizado (ou *transfer learning*, em inglês). Foi esta a estratégia empregada por Karri, Chakraborty e Chatterjee (2017) que utilizaram a rede GoogLeNet para a classificação,

obtendo acurácia de 89%. Já Kermany et al. (2018) processaram 200.000 *B-Scans* obtendo acurácia de 99% para a detecção de imagens com drusas.

Esses trabalhos relacionados à detecção de DMRI não-exsudativa estão resumidos de forma comparativa no Quadro 1. Podemos verificar que, à exceção de Farsiu et al. (2014), que utilizam a caracterização topográfica das camadas da retina para a estabelecer a classificação, a maioria dos métodos para detecção de DMRI em exames de OCT mencionados anteriormente realizou análise individualizada das *B-scans* dos volumes OCT (WANG et al., 2016; SUN; LI; SUN, 2017; RAVENSCROFT et al., ; RAVENSCROFT et al., 2017). Eles utilizaram técnicas de segmentação simples baseada na eliminação do fundo (ALBARRAK et al., 2013) ou recorte de dimensões pré-definidas na região de maior intensidade de *pixels* (SRINIVASAN et al., 2014b).

Dessa forma, a correlação entre *B-scans* sucessivos é desconsiderada. O motivo é que normalmente as fatias sucessivas sofrem de um desalinhamento provocado pelo movimento dos olhos durante a captura. Assim, o processamento individualizado de cada *B-scan* não leva em consideração esse desalinhamento. Além disso, é mais simples a definição de uma região de interesse a partir de uma imagem bidimensional do que de uma representação completa do volume.

A principal desvantagem da análise individualizada de cortes seccionais é desconsiderar a correlação entre *pixels* de *B-scans* sucessivas, o que fragmenta a informação e pode levar a diagnósticos errados. Além disso, após a etapa de detecção de características indicadoras da presença de DMRI por *B-scans*, é necessária uma nova classificação para a indicação da presença da doença por volume.

Tabela 1 – Comparação de trabalhos relacionados.

Trabalho	Dimen.	Pré-processamento	Descritores	Classificação	Amos- tras	Publ. disp.
Liu (2011)	2D	Deformação da imagem	Pirâmides espaciais multi-escala, LBP + PCA.	SVM não-linear	457	Não
Serrano-Aguilar (2012)	2D	Normalização de histograma	Descritores Haar-Like e Haralick	Árvore de decisão	200	Não
Albarrak (2013)	3D	Varição total isotrópica <i>Split Bregman</i> e realinhamento da imagem por método dos mínimos quadrados	Histograms LBP com gradiente orientado	Rede Bayesiana	140	No
Zhang (2014)	3D	Varição total isotrópica (Bregman) com mínimos quadrados	LBP de três planos ortogonais (LBP-TOP), quantização de fase local (LPQ) e pirâmide espacial multiescala (MSSP)	Conjunto de modelos PCA de núcleo de classe única(KPCA)	140	Não
Farsiou (2014)	3D	Segmentação de três bordas da retina	Gradação de espessura do EPR-CD	Regressão linear generalizada	384	Sim
Srinivasan (2014)	3D	BM3D	HOG	Três SVM de uma classe	45	Sim

(quadro continua)

Tabela 1 – Comparação de trabalhos relacionados (continuação)

Trabalho	Dimen.	Pré-processamento	Descritores	Classificação	Amos- tras	Publ. disp.
Venhuizen (2015)	2D	Filtro de gradiente gaussiano de primeira ordem	Auto-aprendizado de descritores baseado em fragmento das imagens	Florestas aleatórias	384	Sim
Wang (2016)	2D	-	(LCP, do inglês <i>linear configuration patterns</i> ) em multi-escala	Otimização sequencial mínima (SMO)	45	Sim
Sun (2017)	2D	Alinhamento da retina e definição de ROI	SIFT	Três classificadores SVM de duas classes	678 <i>B-scans</i>	Não
Ravenscroft (2017)	2D	Segmentação manual da coroide	Auto-aprendizado de descritores por CNN	Redes neurais	75	Não
Fang (2017)	3D	Remoção da média de fragmentos	PCA	Máquina de aprendizagem extrema	54	Não
Karri (2017)	2D	Aproximação do EPR baseada em itensidade e redução de ruídos com BM3D	Auto-aprendizado de descritores	CNN (retreinamento da GoogLeNet)	45	Sim
Lee (2017)	2D	-	Auto-aprendizado de descritores	CNN	100,000 <i>B-scans</i>	Não

(quadro continua)

Tabela 1 – Comparação de trabalhos relacionados (continuação)

Trabalho	Dimen.	Pré-processamento	Descritores	Classificação	Amos- tras	Publ. disp.
Kermary et al. (2018) Kermary (2018)	2D	-	Auto-aprendizado de descritores	CNN (retreinamento)	207,130 <i>B-</i> <i>scans</i>	Sim

### 3 FUNDAMENTAÇÃO TEÓRICA

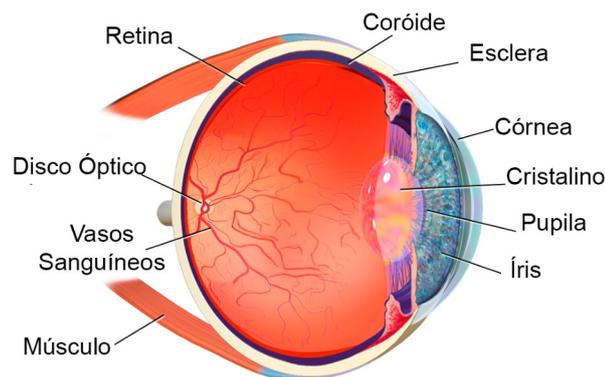
Neste capítulo apresentamos os fundamentos teóricos que sustentam a metodologia de detecção automatizada de DMRI proposta por esta tese.

Iniciamos por uma breve apresentação sobre o olho humano e a DMRI. Em seguida, abordamos a OCT, o tratamento de ruídos característicos dessas imagens e apresentamos técnicas de processamento digital de imagens, caracterização de textura por descritores geoestatísticos (semivariograma e semimadograma), *Capsule Networks* (redes de aprendizado profundo baseado em cápsulas), classificação de padrões com Máquina de Vetores de Suporte (SVM) e, por fim, técnicas de validação de resultados.

#### 3.1 O Olho Humano e a Retina

O olho humano é um órgão de formato aproximadamente esférico que possui cerca de 25 milímetros de diâmetro e ocupa a maior parte da porção frontal da órbita. Suas estruturas são responsáveis pela refração e focalização dos raios luminosos e, por fim, sua conversão em impulsos elétricos que são transmitidos ao cérebro através do nervo óptico.

Figura 1 – Esquema estrutural do olho humano (corte transversal).



Fonte: Staff (2014)

A estrutura do globo ocular (Figura 1) pode ser dividida em três camadas (ou túnicas) distintas: camada fibrosa, formada pela córnea e pela esclera, responsável pelo revestimento do ocular; camada vascular ou trato uveal, formada pela íris, corpo ciliar e coróide, e túnica interna, que é constituída pela retina.

A retina é uma estrutura multicamadas composta por células nervosas que recebem os estímulos luminosos e os convertem em estímulos elétricos. Essa estrutura contém duas partes: a parte óptica da retina, situada nas regiões posterior e lateral, e a parte cega, situada na região frontal da retina. A Figura 2 apresenta uma fotografia do fundo do olho centralizada na parte óptica da retina. Nela podemos perceber dois pontos de maior destaque, o disco óptico e a mácula.

O disco óptico é o ponto onde o nervo óptico sai da retina. Nesse local não possui células fotorreceptoras, sendo conhecido como o ponto cego da retina. Esta região tem a coloração mais clara e do seu centro partem ramos da artéria central da retina. A região da retina que apresenta maior espessura é denominada mácula.

Figura 2 – Retinografia de um olho humano saudável com destaque para o disco óptico e a mácula.

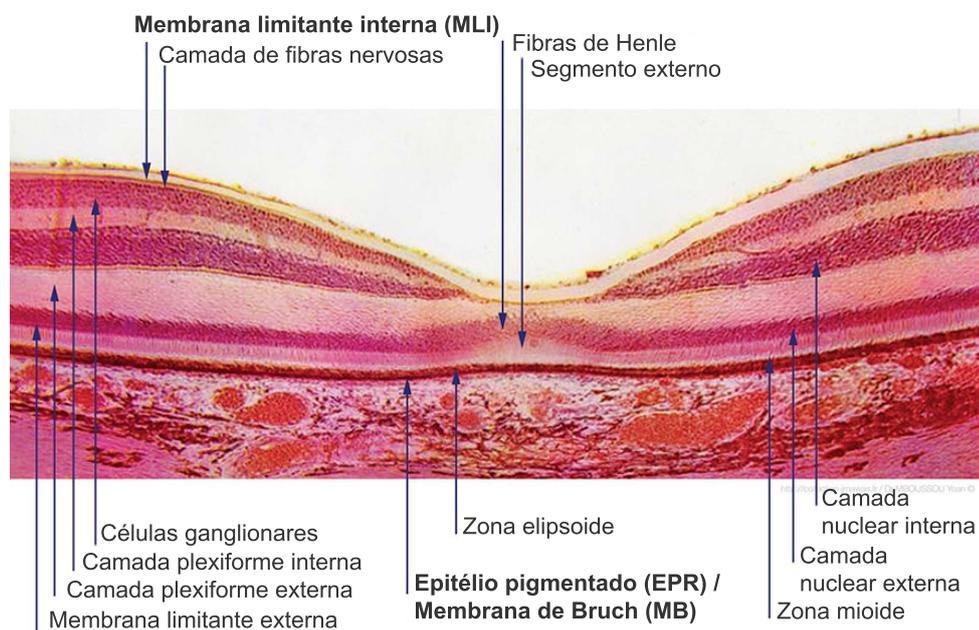


Fonte – Zeiss (2017)

A mácula situa-se temporalmente ao disco óptico e possui formato ovalado com aproximadamente 5,5 *mm* de diâmetro. No centro da mácula encontra-se a fóvea, uma depressão da retina que constitui o local de maior acuidade visual. Alterações morfológicas nas camadas retinianas desta área estão diretamente relacionados à diminuição ou perda da capacidade visual. Na Figura 3 é apresentado um corte histológico da região de uma mácula saudável com a identificação das camadas.

As camadas de interesse para esta tese são a membrana limitante interna (MLI), o epitélio pigmentado retiniano (EPR) e membrana de Bruch (MB). A MLI delimita a retina internamente e a MB externamente. O EPR é a camada que sofre alterações diretamente quando há o acometimento por DMRI.

Figura 3 – Corte histológico da região macular da retina humana com identificação das camadas.



Fonte – Adaptado de Mawas (19-)

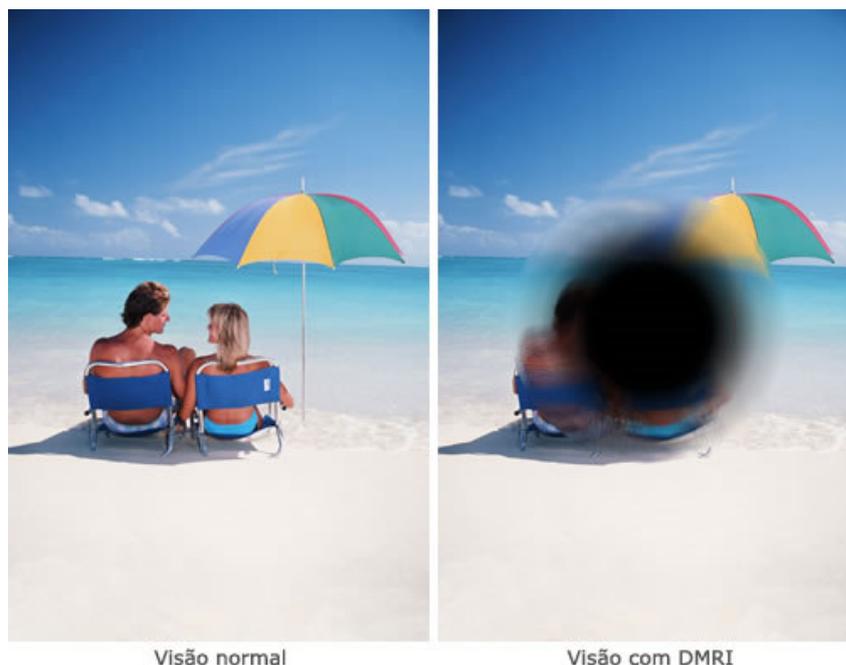
### 3.2 Degeneração Macular Relacionada à Idade

A Degeneração Macular Relacionada à Idade (DMRI) é uma das principais causas de cegueira. Nos Estados Unidos, por exemplo, é responsável por mais de 50% dos casos de cegueira (MEHTA, 2015).

A idade avançada é o principal fator de risco para a DMRI. A incidência e gravidade aumentam conforme aumenta a idade do paciente (SOON-CHUN, 2015). Outros fatores de risco são: ascendência caucasiana, tabagismo, história familiar de DMRI, sedentarismo, obesidade e doenças cardiovasculares, como hipertensão e elevação do nível de colesterol no sangue (SOON-CHUN, 2015).

A DMRI atinge a região macular, área central da retina, provocando a degeneração progressiva de fotorreceptores, EPR e coróide. Frequentemente, a DMRI provoca atrofia do EPR e/ou neovascularização na coróide, levando à perda de visão central (YANOFF; DUKER; AUGSBURGER, 2014). O comprometimento da visão é ilustrado pela Figura 4, onde se percebe perda de visão que começa na região central e se estende na direção das regiões periféricas.

Figura 4 – Ilustração da percepção de imagem por paciente com DMRI, comparada à visão normal.



Fonte – CBO (2017)

### 3.2.1 Fisiopatologia da DMRI

A mácula é responsável pela nitidez da visão, leitura e identificação de cores, possuindo o dobro de fotorreceptores que o resto da retina. O processo de desenvolvimento da DMRI se inicia com o depósito de lipídios na MB. Esses depósitos são conhecidos como depósitos basais laminares que, posteriormente, serão classificados como drusas (YANOFF; DUKER; AUGSBURGER, 2014).

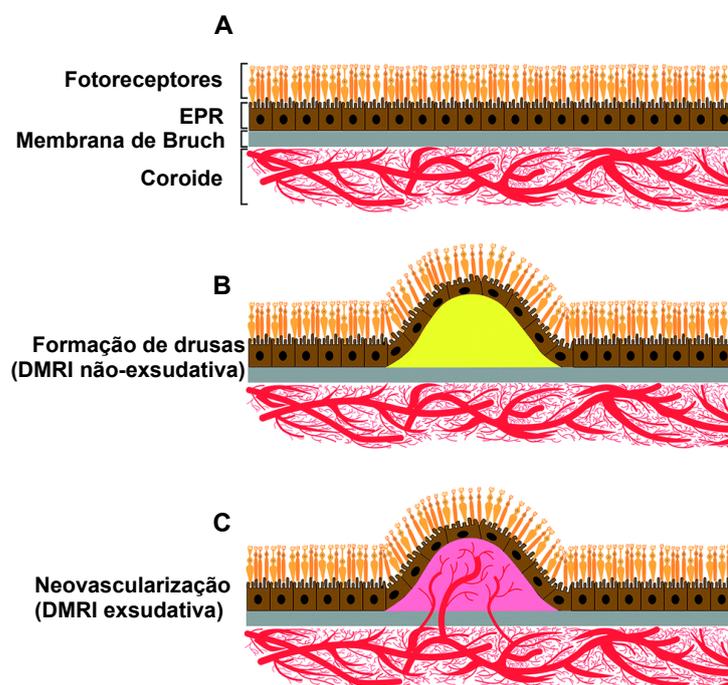
As drusas são formadas entre a lâmina basal do EPR e a camada colágena interna da MB. Podem ser divididas em duras ou moles. As primeiras são redondas e pequenas, de cor branco-amarelada e geralmente inócuas. Já as drusas moles são maiores e sem margens distintas, podendo aumentar de tamanho e agregar-se no chamado complexo de drusas (drusenoide). Além disso, podem passar por processo de calcificação (KANSKI, 2012).

A DMRI pode ser classificada em não-exsudativa (ou seca) ou exsudativa (ou úmida). A manifestação não-exsudativa é caracterizada pela presença de drusas geralmente duras, e atrofia do EPR, com perda de fotorreceptores e progressiva diminuição da visão central.

Já na DMRI exsudativa, ocorre o aparecimento de uma membrana neovascular sub-retiniana, com extravasamento de sangue e/ou soro, alterando a anatomia macular,

com perda devastadora de fotorreceptores locais. Nesse tipo da doença, a consequente baixa da visão é muito mais rápida e acentuada quando comparada à forma não-exsudativa. A presença de drusas moles é considerada fator predisponente para o desenvolvimento de DMRI exsudativa (NEHEMY, 2006; KANSKI, 2012). A Figura 5 ilustra a manifestação de drusas nas formas exsudativa e não exsudativa da DMRI. Podemos observar o atrofiamento dos fotorreceptores provocado pela presença das drusas (B e C) em comparação com uma retina saudável (A).

Figura 5 – Ilustração comparativa entre uma retina saudável (A), retina acometida por DMRI não-exsudativa com formação de drusas (B) e retina com neovascularização (C).



Fonte – Farjood e Vargis (2018)

### 3.2.2 Diagnóstico e Acompanhamento

O diagnóstico e acompanhamento da DMRI são feitos com base no exame de fundo do olho, seja por meio de oftalmoscopia direta ou indireta. Esta análise permite identificar a presença de drusas, mesmo antes da presença de sintomas (CBO, 2013).

Outro exame importante é a angiofluoresceinografia, muito utilizado para diagnóstico e acompanhamento de DMRI exsudativa, que utiliza administração de contraste endovenoso para a visualização, estudo e avaliação funcional da circulação sanguínea da retina e coróide. Esse exame permite ainda identificar edemas maculares em vários níveis,

bem como a caracterização precisa da formação de membrana neovascular (KANSKI, 2012).

A OCT é uma modalidade de exame que apresenta alta sensibilidade na detecção de doenças maculares e possui alta correlação com a angiofluoresceinografia. Para casos de DMRI não-exsudativa, a OCT torna possível caracterizar o volume e área das drusas com precisão. Também permite o acompanhamento da progressão da doença, desde seu estágio mais precoce até sua forma mais severa, com caracterização detalhada de áreas de atrofia, perda de fotorreceptores e deformações do EPR. (YANOFF; DUKER; AUGSBURGER, 2014).

A capacidade de realização de exame sem contato físico com o olho e sua alta resolução tornaram a OCT muito adequada à prática clínica oftalmológica, podendo ajudar na identificação, monitoramento e quantificação da DMRI (DEBUC, 2011). Isso tem expandido sua utilização em consultórios clínicos e, hoje, existe uma variedade de modelos de tomógrafos OCT disponíveis comercialmente.

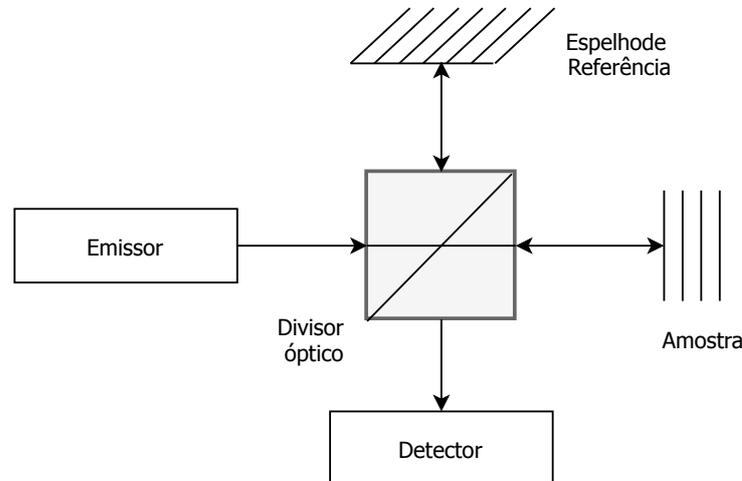
### 3.3 Tomografia de Coerência Óptica - OCT

A tomografia de coerência óptica (OCT) é uma modalidade não-invasiva de aquisição de imagens de estruturas semitransparentes (e.g., tecidos biológicos) (HUANG et al., 1991). As imagens são capturadas em resolução micrométrica em duas ou três dimensões a partir da sondagem por um sinal luminoso de baixa coerência. Essa tecnologia possui aplicações na indústria e na área médica e boa aceitação na área oftalmológica.

A Figura 6 apresenta um diagrama básico do funcionamento de um tomógrafo de OCT. Uma fonte de luz de baixa coerência emite um feixe que passa por um divisor óptico onde é dividido em dois. Uma parte é direcionada a um ponto da amostra e a outra a um espelho de referência. A luz refletida pela amostra é recombinada à que foi refletida pelo espelho de referência dando origem a um padrão de interferência que é captado pelo detector. O emissor aponta para um ponto imediatamente vizinho e o processo é repetido por toda a região de interesse.

Na tecnologia empregada na OCT, há duas formas em que a interferência do sinal é detectado: no domínio do tempo (TD-OCT, do inglês *time domain optical coherence tomography*) e no domínio da frequência (FD-OCT, do inglês *Frequency Domain Optical Coherence Tomography*).

Figura 6 – Diagrama básico de um equipamento de OCT com destaque para os principais elementos presentes.



Fonte: Produzida pelo autor.

Os tomógrafos ópticos atuais se baseiam na captação no domínio da frequência e se dividem em duas categorias: OCT de domínio espectral (SD-OCT, do inglês *spectral domain optical coherence tomography*) e OCT de fonte de varredura (SS-OCT, do inglês *swept source optical coherence tomography*). A SD-OCT é a tecnologia mais utilizada atualmente e as imagens utilizadas nesta tese foram adquiridas utilizando essa tecnologia.

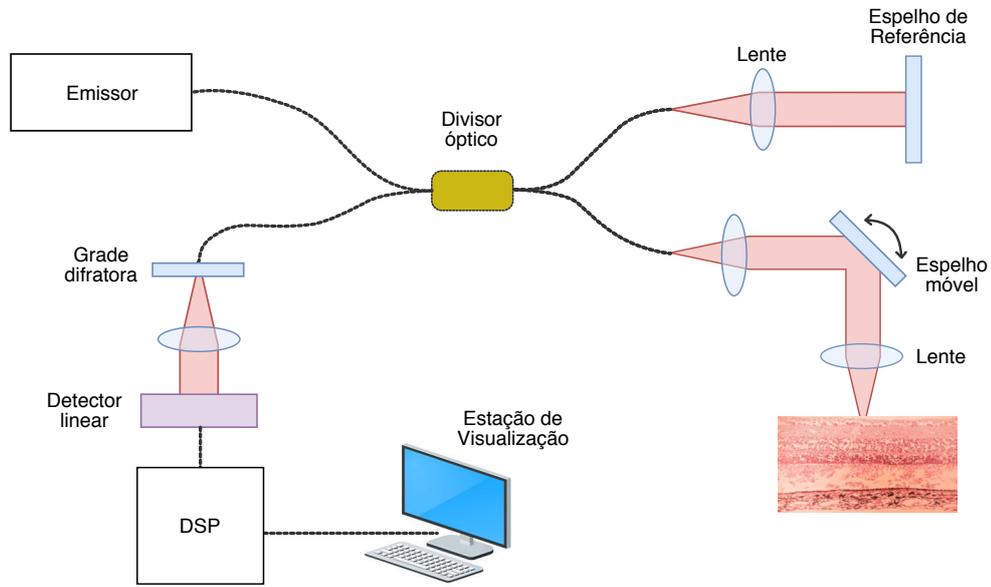
Na SD-OCT, a aquisição da imagem também se baseia na divisão de um feixe de luz em dois segmentos e na captura do padrão de interferência capturado. Entretanto, conforme apresentado na Figura 7, o sinal de interferência é disperso por uma grade difratora e analisado por um espectrômetro de detecção linear. Cada linha de profundidade da imagem (*A-scan*) é reconstruída pelo cálculo da transformada inversa de Fourier (DUKER; WAHEED; GOLDMAN, 2015) obtida pelo detector. A partir dos *A-scans* adquiridos sucessivamente, é realizada a reconstrução da imagem tomográfica (*B-scan*) ou até mesmo uma reconstrução volumétrica da região escaneada (Figura 8).

O sinal  $I(\omega)$  adquirido no domínio da frequência pode ser representado por (ALI; PARLAPALLI, 2010):

$$I(\omega) \approx K_1 S(\omega) + K_2 \text{Re}\{S(\omega)H(\omega)\} \quad (3.1)$$

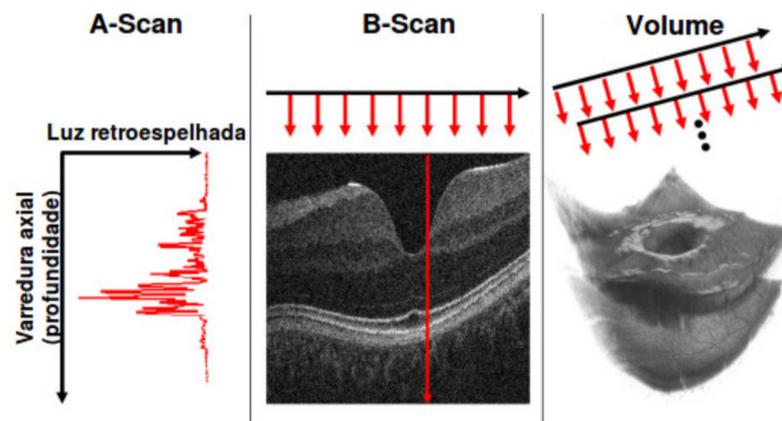
O termo  $K_1 S(\omega)$  representa um valor DC (valor médio) que é comumente subtraído

Figura 7 – Esquema de aquisição de um tomógrafo de coerência óptica no domínio espectral.



Fonte: Produzida pelo autor.

Figura 8 – Aquisição de SD-OCT. À esquerda, aquisição unidimensional (*A-scan*). Na imagem central, a imagem bidimensional (*B-scan*). Na imagem à direita, temos a aquisição 3D.



Fonte: Adaptado de (KRAUS et al., 2012).

do sinal capturado para eliminação de ruído e sinal de fundo, resultando em:

$$I'(\omega) = K_2 \text{Re}\{S(\omega)H(\omega)\} \quad (3.2)$$

onde  $K_2 \text{Re}\{S(\omega)H(\omega)\}$  representa a parte real da interferência capturada entre o sinal de referência  $S(\omega)$  e o sinal refletido pela amostra  $H(\omega)$ .

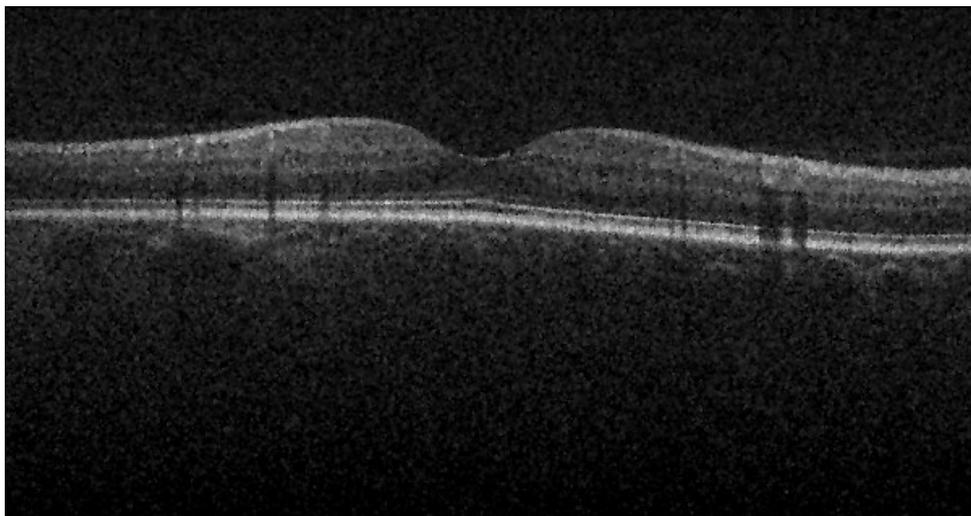
A reconstrução de uma (*A-scan*) é realizada pela aplicação da transformada inversa de Fourier:

$$i(t) = FT^{-1}(I'(\omega)) \quad (3.3)$$

A principal vantagem da SD-OCT é que a aquisição de uma *A-scan* individual é realizada inteiramente a partir da análise do sinal em um único espectro e sem a necessidade da movimentação mecânica do conjunto de detecção. O sinal medido a partir da amostra representa a combinação das propriedades de refletividade, absorção e espalhamento do feixe de luz pelas camadas do tecido estudado.

A imagem de SD-OCT resultante é um mapa da refletividade da amostra representado em escala de cinza. Áreas escuras representam material homogêneo com baixa refletividade (ar ou fluidos) e áreas mais claras representam as áreas de maior refletividade como fibras colágenas, paredes e núcleos celulares (DEBUC, 2011). O sinal luminoso é atenuado nos tecidos de forma exponencial, reduzindo a intensidade da imagem com a profundidade. Vasos sanguíneos apresentam maior atenuação que tecidos com colágeno e gordura e fluídos são os que possuem menor atenuação. Uma imagem típica de OCT da região macular adquirida pela tecnologia espectral é apresentada na Figura 9. Nesta imagem podemos observar a variação de intensidade entre as camadas da estrutura da retina e áreas de sombra devido à presença de estruturas de alta refletividade.

Figura 9 – *B-scan* da região central da mácula obtida pela tecnologia SD-OCT.



Fonte – Farsiu et al. (2014)

### 3.4 Processamento digital de imagens

O processamento digital de imagens (PDI) é a aplicação de técnicas computacionais para o tratamento de sinais discretizados de duas ou mais dimensões (imagens). Assim, aspectos da imagem serão modificados de modo que possam beneficiar a análise feita por um usuário humano ou até mesmo gerar uma saída que será utilizada como entrada para outro processo.

Existem vários níveis de processamento de imagens. Partindo de níveis mais básicos, temos operações de filtragem de ruídos e realce de bordas como formas de melhoramento de imagens, inclusive processamentos envolvendo aprendizado de máquina e inteligência artificial Gonzalez e Woods (2006).

O processamento digital aplicado a imagens médicas traz uma série de benefícios para a prática clínica e de pesquisa: promove auxílio ao diagnóstico, planejamento de tratamentos, simulação de cirurgias, compressão de imagens em bancos de exames etc. Além disso, ferramentas de detecção, baseadas em processamento de imagens, auxiliam os especialistas a encontrarem automaticamente estruturas e alterações morfológicas e funcionais de interesse, como nervos, vasos etc.

Nesta tese, uma série de técnicas de processamento de imagens são aplicadas com o objetivo de efetuar a segmentação da retina e diagnóstico automático de DMRI. Essas técnicas serão discutidas a seguir. Começaremos por técnicas de pré-processamento, ou seja, filtros e procedimentos que visam eliminar ruídos das imagens de forma a permitir um processamento posterior.

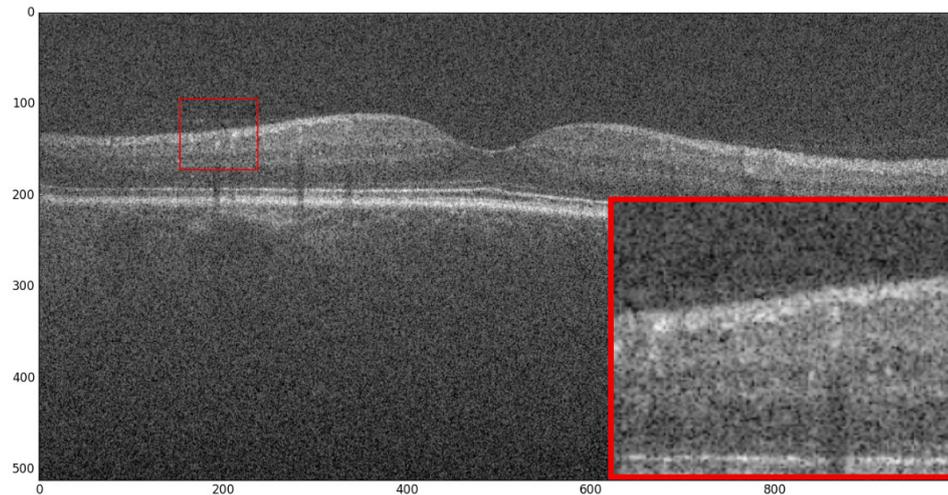
### 3.5 Eliminação de ruído em imagens

A presença de ruído em um sinal é uma característica indesejada. Ruídos podem ser introduzidos através do processo de aquisição, armazenamento, transmissão etc. (RUDIN; OSHER; FATEMI, 1992). São anomalias que dificultam o processamento das imagens, podendo diminuir a eficiência de um método ou, dependendo da intensidade, até mesmo impedir o processamento.

As imagens de OCT sofrem da presença de ruído introduzido durante o processo de aquisição devido à interferência destrutiva entre vários feixes de luz pulsante. As imagens de OCT são contaminadas por variações entre pontos escuros e de alto brilho (Figura 10)

que degradam a qualidade das imagens para a análise clínica e pesquisa (GONG; ZHANG; YAO, 2015).

Figura 10 – Imagem de OCT com destaque para o ruído característico inserido pelo processo de aquisição (ampliação no canto inferior direito).



Fonte – Farsiu et al. (2014). Montagem produzida pelo autor.

### 3.5.1 Filtro da média

O filtro da média é um dos filtros mais simples para a diminuição do ruído de uma imagem. Calcula para cada *pixel*  $x$  a média das intensidades dos seus  $N$  vizinhos adjacentes, conforme:

$$M(x) = \frac{1}{N} \sum_{i=1}^N z_i \quad (3.4)$$

onde  $z_i$  corresponde ao valor de cada *pixel* de uma vizinhança.

Podem ser atribuídos pesos  $w$  aos *pixels* vizinhos. Assim, o filtro passa a ser considerado com média ponderada, que é representado por:

$$M_p(x) = \frac{1}{N} \sum_{i=1}^N w_{z_i} z_i \quad (3.5)$$

### 3.5.2 Filtragem gaussiana

Outro filtro comum é filtro de suavização gaussiana. Ele se baseia na convolução da imagem por um operador de duas dimensões que diminui a nitidez da imagem, porém

remove ruído e detalhes ao mesmo tempo. O operador é uma superfície gaussiana definido pela equação:

$$g_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.6)$$

onde o par  $(x, y)$  corresponde à coordenada do *pixel* e  $\sigma$  ao desvio padrão da distribuição gaussiana.

O resultado da aplicação da suavização gaussiana em uma imagem corresponde ao cálculo da média ponderada da vizinhança de cada *pixel*, com um peso maior para os *pixels* centrais contrastando com filtro da média que considera a vizinhança de maneira uniforme. Essa característica permite uma suavização que preserva melhor os detalhes, sendo mais adequada para a detecção de bordas.

### 3.5.3 Supressão de ruído com Wavelet

A transformada *wavelet* (WT, do inglês *wavelet transform*) (DAUBECHIES, 1990) é uma ferramenta poderosa para o processamento de sinais e imagens, podendo inclusive ser empregada no tratamento de ruído.

A função de transformação WT contínua pode ser definida como:

$$(W_{\psi}f)(a, b) = \int_{-\infty}^{\infty} f(t) \cdot \psi_{a,b}^*(t) dt \quad (3.7)$$

Onde o parâmetro  $b$  representa o deslocamento no espaço, o parâmetro  $a$  representa o fator de escala ( $a > 0$  corresponde à dilatação e  $a < 0$  para a contração de  $\psi(t)$ ). O fator de multiplicação  $\frac{1}{\sqrt{a}}$  é para normalização da energia através das diferentes escalas.

A supressão de ruído por *wavelets* pode ser definida da seguinte forma. Seja  $x(t)$  o sinal de dados corrompido por ruído:

$$x(t) = s(t) + n(t) \quad (3.8)$$

onde  $s(t)$  é o sinal original, e  $n(t)$  o ruído, ambas funções do tempo a serem amostrados.

Sejam  $W(\ )$  o operador de transformação direto e  $W^{-1}(\ )$  o inverso que representam a transformada de *wavelet*, escolhidos apropriadamente. Seja ainda  $D(\ , \omega)$  o operador que representa a supressão de ruído para um limiar  $\sigma$ . Como o objetivo é suprimir o ruído

de  $x(t)$  para obter uma estimativa  $\hat{s}(t)$  de  $s(t)$ , então o procedimento completo pode ser resumido nas três equações a seguir (DIAS, 2003):

$$C = W(x) \quad (3.9)$$

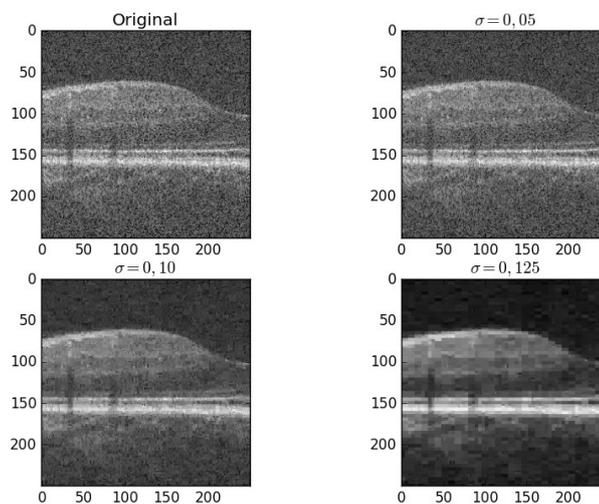
$$C_d = D(C, \sigma) \quad (3.10)$$

$$\hat{s}(t) = W^{-1}(C_d) \quad (3.11)$$

Inicialmente os coeficientes de decomposição *wavelet*  $C$  são calculados. Em seguida, o operador de supressão de ruído  $D()$  é aplicado a  $C$ , de forma a eliminar o que for menor que um limiar  $\sigma$  previamente estabelecido. Por fim, a estimativa  $\hat{s}(t)$  do sinal original é obtida a partir da transformada inversa dos coeficientes com ruído suprimido  $C_d$  (DIAS, 2003).

Na Figura 11 apresentamos um exemplo de supressão de ruído em uma imagem OCT. Vários valores de  $\sigma$  foram utilizados testados a fim de promover a diminuição do ruído na imagem. Uma análise visual permite perceber que, para esta imagem, valores de  $\sigma \leq 0,10$  diminuem a presença de ruído, porém, acima desse valor, tem-se perda visível de detalhes.

Figura 11 – Redução de ruídos em recorte de uma *B-scan* de OCT com o filtro wavelet para  $\sigma = \{0,05, 0,10, 0,125\}$ .



Fonte – Produzida pelo autor.

### 3.5.4 Avaliação da redução de ruído

Para a avaliação sobre a redução do ruído das imagens, neste trabalho utilizaremos métricas empregadas em processamento de sinais: os quocientes sinal/ruído (SNR, do inglês *signal to noise ratio*), contraste/ruído (CNR, do inglês *contrast to noise ratio*) e a relação sinal-ruído de pico (PSNR, do inglês *Peak Signal-to-Noise Ratio*) (JIN et al., 2013).

O SNR é dado por:

$$SNR = 20 \log_{10} \frac{\mu_f}{\sigma_b} \quad (3.12)$$

onde,  $\mu_f$  corresponde à média das intensidades da região de interesse na imagem, ou seja, o sinal. E  $\sigma_b$  corresponde ao desvio padrão do fundo da imagem.

A medida do CNR é dada por:

$$CNR = \frac{|\mu_f - \mu_b|}{\sqrt{\sigma_f^2 + \sigma_b^2}} \quad (3.13)$$

onde,  $\mu_f$  corresponde à média das intensidades da região de interesse na imagem.

O sinal,  $\mu_b$  é a média das intensidades do fundo e  $\sigma_b^2$  e  $\sigma_f^2$  correspondem às variâncias do fundo e da região de interesse, respectivamente.

Já a medida PSNR efetua a comparação entre dois sinais, avaliando relação entre a máxima energia de um sinal ( $MAX_I$ ) e o ruído detectado através do cálculo da raiz do erro quadrático médio (RMSE, do inglês *root mean square error*) que afeta a sua representação. Pode ser definida como:

$$PSNR = 20 \log_{10} \frac{MAX_I}{RMSE} \quad (3.14)$$

$MAX_I$  se refere à intensidade máxima que um *pixel* pode obter em uma imagem, enquanto  $RMSE$  é definida por:

$$RMSE = \left[ n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]^{1/2} \quad (3.15)$$

A comparação de imagens com grande semelhança implica em valores baixos de  $RMSE$  e valores elevados de  $PSNR$ . Quando a semelhança entre as imagens é total, o valor  $RMSE = 0$ . Nesse caso o valor de PSNR não existe. Valores típicos variam entre 20 dB e 40 dB (SALOMON, 2002).

### 3.6 Capsule Networks

Redes de Cápsulas (*CapsNet* ou *Capsule Networks* em inglês) são uma arquitetura de redes neurais de aprendizado profundo proposta por Sabour, Frosst e Hinton (2017) para superar limitações das redes neurais convolucionais (CNN, do inglês *convolutional neural networks*) (LECUN; BENGIO; HINTON, 2015).

Apesar de terem um desempenho excepcional em muitas aplicações, as CNNs têm vários inconvenientes importantes, principalmente devido ao procedimento de roteamento de dados. Elas realizam o roteamento entre camadas por meio de operações de *pooling* que descartam informações como localização e posicionamento dos objetos que podem ser valiosos para fins de classificação ou detecção.

A ideia principal da CapsNet é codificar as relações relativas a locais, escalas, orientações, por exemplo, entre as partes locais e o objeto inteiro. Codificar essas relações prepara o modelo com uma compreensão interna do espaço. Isso permite que a CapsNet reconheça objetos em diferentes visualizações que não foram vistas nos dados de treinamento. Devido a essas características, várias aplicações têm sido propostas para o processamento de imagens médicas com resultados que alcançam ou superam o desempenho de redes convolucionais tradicionais (AFSHAR; MOHAMMADI; PLATANIOTIS, 2018; IESMANTAS; ALZBUTAS, 2018; MOBINY; NGUYEN, 2018).

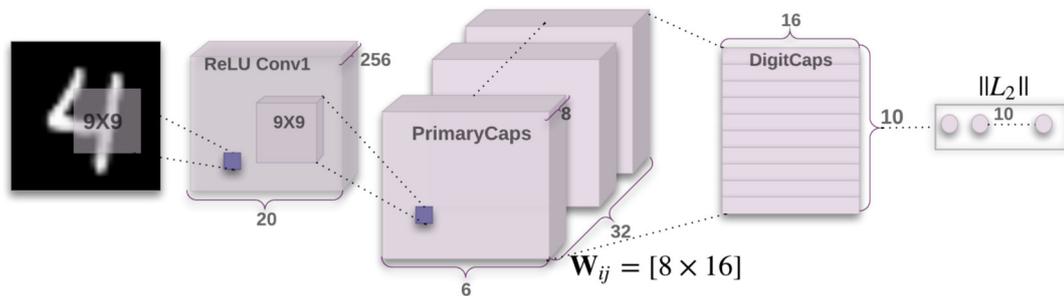
A CapsNet é uma rede composta por cápsulas que são um grupo de neurônios especializados em detectar um objeto particular em uma região da imagem. Cada cápsula responde com um vetor que representa a probabilidade desse objeto estar presente nessa região e também parâmetros de pose como localização, rotação etc.

Como em outras redes, a CapsNet é organizada em várias camadas. As cápsulas de nível mais baixo são chamadas cápsulas primárias. Elas são responsáveis por receber uma região pequena da imagem e tentar detectar a presença e pose de determinado objeto (GÉRON, 2018). Essa detecção é invariante quanto à mudança de posicionamento do objeto.

As camadas de nível mais elevado detectam objetos mais complexos levando em consideração a detecção das camadas inferiores. A Figura 12 apresenta o arranjo da rede proposta por Sabour, Frosst e Hinton (2017) para o reconhecimento de dígitos. Nessa configuração as cápsulas primárias responsáveis pelo processamento de regiões de  $9 \times 9$  *pixels* interagem com cápsulas de alto nível (“DigitCaps”) para a identificação de algarismos

manuscritos.

Figura 12 – Esquema da arquitetura de uma rede CapsNet para identificação de algarismos.



Fonte: (SABOUR; FROSST; HINTON, 2017)

O fato de a saída de uma cápsula ser um vetor torna possível usar um poderoso mecanismo de roteamento dinâmico para garantir que a saída da cápsula seja enviada para uma “mãe” apropriada na camada superior. Inicialmente, a saída é roteada para todas as possíveis mães, mas é reduzida pelos coeficientes de acoplamento que somam 1.

Para cada possível “mãe”, a cápsula calcula um “vetor de previsão”, multiplicando sua própria saída por uma matriz de ponderação. Se esse vetor de previsão tiver um produto escalar grande para a saída de uma possível “mãe”, haverá uma realimentação de cima para baixo que aumenta o coeficiente de acoplamento para essa “mãe” e o diminui para outras. Isso aumenta a contribuição que a cápsula faz para a cápsula superior, aumentando ainda mais o produto escalar da previsão da cápsula com a saída da mãe.

Esse tipo de “roteamento por acordo” é mais eficaz do que a forma muito simples de roteamento implementada pelo *pool* máximo (CNN), que permite que os neurônios em uma camada ignorem todos, exceto o detector mais ativo em um conjunto local na camada abaixo.

A relação entre a  $i$ -ésima cápsula em uma camada inferior e a  $j$ -ésima cápsula na próxima camada superior é codificada usando uma matriz de transformação linear  $W_{ij}$ . A informação é propagada como:  $\hat{u}_{j|i} = W_{ij}u_i$ . O vetor  $\hat{u}_{ij}$  representa a semelhança da  $i$ -ésima cápsula em uma camada inferior sobre a  $j$ -ésima cápsula na camada superior.

Tendo computado os vetores de previsão, as cápsulas de nível inferior encaminham suas informações para as cápsulas-mães que mais concordam com suas previsões. O mecanismo que garante que as saídas das cápsulas filhas sejam enviadas para as cápsulas “mães” apropriadas é denominado *Roteamento dinâmico*. A variável  $c_{ij}$  denota o coeficiente de roteamento da  $i$ -ésima cápsula na camada inferior à  $j$ -ésima cápsula na camada superior,

onde  $\sum_j c_{ij} = 1$  e  $c_{ij} \geq 0, \forall j$ . Quando  $c_{ij} = 1$ , todas as informações da  $i$ -cápsula serão enviadas para a  $j$ -cápsula, enquanto que para  $c_{ij} = 0$ , não há informações fluindo entre as duas cápsulas.

O método de roteamento dinâmico ajusta iterativamente os coeficientes  $c_{ij}$  e direciona as saídas das cápsulas filhas para a cápsula apropriada na próxima camada, de modo que elas obtenham uma entrada mais clara, determinando, assim, a pose dos objetos com mais precisão.

CapsNets calcula os coeficientes de roteamento da seguinte forma (MOBINY; NGUYEN, 2018):

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|}, s_j = \sum_i c_{ij} \hat{u}_{j|i}, \quad (3.16)$$

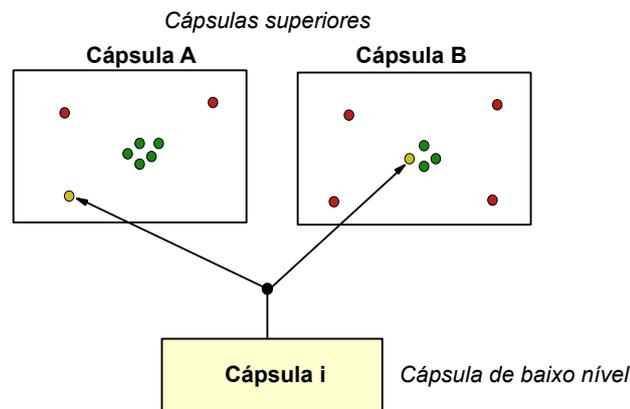
$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})}, b_{ij} \leftarrow b_{ij} + v_j \cdot \hat{u}_{j|i}. \quad (3.17)$$

A saída de cada cápsula mãe  $v_j$  é computada como a soma ponderada de todas as predições de cápsulas filhas, passando então por uma função *squash* (Equação 3.16) não-linear. A compactação garante que o vetor de saída tenha comprimento não maior que 1 (de modo que o comprimento pode ser interpretado como a probabilidade de que uma determinada característica seja detectada pela cápsula) sem alterar sua direção. As cápsulas mãe recebem previsões de todas cápsulas filhas. Esses vetores são representados por pontos na Figura 13. O mecanismo de roteamento dinâmico aumentará o coeficiente de roteamento para a cápsula  $j$  “mãe” por um fator de  $\hat{u}_{j|i} \cdot v_j$ . Assim, uma cápsula filha enviará mais informações para a cápsula “mãe” cuja saída  $v_j$  é mais semelhante à sua predição  $\hat{u}_{j|i}$  (MOBINY; NGUYEN, 2018).

Para todos, exceto a primeira camada de cápsulas, a entrada total para uma cápsula  $s_j$  é uma soma ponderada sobre todos os "vetores de previsão"  $\hat{u}_{j|i}$  das cápsulas na camada abaixo e é produzida pela multiplicação da saída  $u_i$  de uma cápsula na camada abaixo por uma matriz de peso  $W_{ij}$  onde o  $c_{ij}$  são coeficientes de acoplamento determinados pelo processo de roteamento dinâmico iterativo.

A Figura 13 mostra uma cápsula de nível inferior tomando a decisão de enviar a sua saída para as cápsulas “mãe”. Essa decisão é feita ajustando os coeficientes de roteamento,  $c_{ij} = 0$ , que serão multiplicados pelos vetores de previsão antes de enviá-los para cápsulas de alto nível. Esse roteamento dinâmico é realizado pelo algoritmo apresentado no artigo

Figura 13 – Roteamento dinâmico para envio de informações da cápsula de nível inferior para a cápsula de nível superior.



Fonte: Adaptado de Mobiny e Nguyen (2018)

apresentado por Sabour, Frosst e Hinton (2017).

### 3.7 Segmentação por corte de grafo

A segmentação de imagens a partir de técnicas de corte em grafos tem desempenhado um papel importante no processamento de imagens médicas. A técnica interpreta os *pixels* ou regiões da imagem como nós de um grafo e transforma um problema de segmentação em um problema de identificação de marcação de nós de acordo com suas propriedades (CHEN; PAN, 2018).

Um grafo  $G(V, A)$  pode ser definido pelo par de conjuntos  $V$ : nós (ou vértices) e  $A$ : pares ordenados  $a = (s, t)$  que representam as arestas, ou seja, a ligação entre os nós adjacentes  $s$  e  $t$ .

Uma imagem pode ser vista como um grafo  $G_{img}(V, A)$ , onde os vértices são os *pixels* (ou *voxels*) e as arestas os pares ordenados de *pixels*  $(s, t) \in A$  considerando uma relação de adjacência em uma região de vizinhança. A região de vizinhança pode ser a vizinhança-4 (onde são considerados os vizinhos imediatamente acima, abaixo, à direita e à esquerda), a vizinhança-8 (além da vizinhança anterior, são introduzidos os vizinhos da posição nordeste, sudeste, sudoeste e noroeste), ou outra vizinhança pré-determinada, incluindo a expansão das vizinhanças anteriores para três ou mais dimensões.

A cada arco  $(s, t) \in A$  do grafo está relacionado um peso  $\omega(s, t) \geq 0$  que corresponde a uma medida de conexão ou similaridade entre os *pixels*. Por exemplo, as arestas formadas

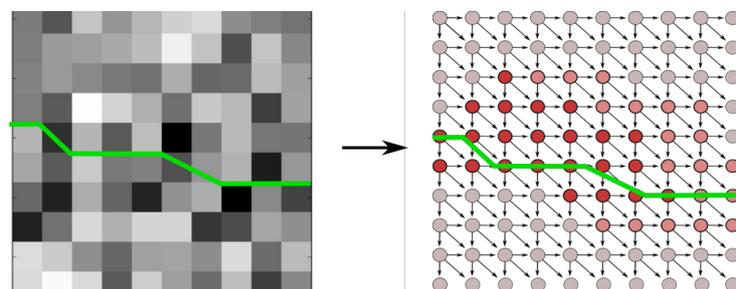
por *pixels* vizinho de intensidade semelhante, um peso  $\omega(s, t) = |I(s) - I(t)|$  pode ser atribuído, onde  $I(t)$  se refere à intensidade de um *pixel* aleatório  $t$  da imagem.

Sobre o grafo formado a partir da imagem podem ser definidos caminhos  $\pi = \{t_1, t_2, t_3, \dots, t_n\}$ , onde os valores  $t_i$  pertencem ao conjunto de *pixels* da imagem. O vetor  $\pi$  corresponde a uma sequência de vértices; nesse caso, os *pixels* que correspondem a um caminho sobre a imagem capaz de dividi-la em duas regiões. Uma restrição impõe que cada *pixel* só poderá ser incluído uma única vez em um determinado caminho  $\pi$ .

A segmentação de imagens por corte de grafo utilizada nesta tese efetua a busca por um caminho ótimo (pesquisa em grafo, ou *graph search*) que parte de um ponto inicial a um ponto final definidos na imagem. Nesse caminho, a soma dos pesos das arestas percorridas é a menor possível. Uma ilustração deste conceito é apresentada na Figura 14, na qual, dado um nó escolhido à borda esquerda da imagem e outro escolhido à borda direita, é calculado o menor caminho (considerando os custos das arestas do grafo) que segmenta a imagem considerando os pontos de chegada e partida. Na imagem da esquerda temos uma representação por *pixels*, enquanto na direita, a mesma imagem é representada como um grafo. As setas que ligam os nós representam a conectividade estabelecida entre os *pixels*. Em verde destacamos o menor caminho que segmenta a imagem.

Para a solução desse problema utilizamos uma técnica clássica e simples, o algoritmo de Dijkstra (CORMEN et al., 2002).

Figura 14 – Exemplo de imagem representada por um grafo. À esquerda uma imagem representada por seus *pixels*. À direita, a mesma imagem representada como um grafo. A marcação em verde corresponde a um caminho percorrido na imagem que a segmenta em duas regiões.



Fonte – Produzida pelo autor.

### 3.8 Geoestatística

As propriedades estatísticas de dados com localização espacial têm sido estudadas e modeladas por anos em aplicações como mineração e exploração de recursos não-renováveis

(petróleo e gás) (SCHOWENGERDT, 2006). Esse campo de investigação de estatística de dados espaciais é conhecido como geoestatística.

A geoestatística fornece ferramentas para descrever a continuidade espacial de fenômenos naturais (ISAAKS; SRIVASTAVA, 1989). O princípio básico é que eventos ocorridos em uma determinada localização tendem a ter comportamento similar em localizações vizinhas, estabelecendo uma relação de dependência espacial.

No campo do processamento de imagens, uma imagem pode ser visualizada como uma superfície e descritores geoestatísticos podem ser empregados para descrever a textura espacial da imagem. Textura é um conceito intuitivo que descreve suavidade, rugosidade e regularidade de uma imagem. É uma característica muito importante para classificação e reconhecimento de objetos e se caracteriza pela variação local dos valores dos *pixels* da região que se repete de maneira regular ou aleatória (HARALICK; SHANMUGAM et al., 1973). Nesta tese, utilizamos os descritores geoestatísticos semivariograma e semimadograma para a caracterização dos padrões existentes nas imagens.

### 3.8.1 Semivariograma

O semivariograma pode ser descrito como (ISAAKS; SRIVASTAVA, 1989):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (x_i - y_i)^2 \quad (3.18)$$

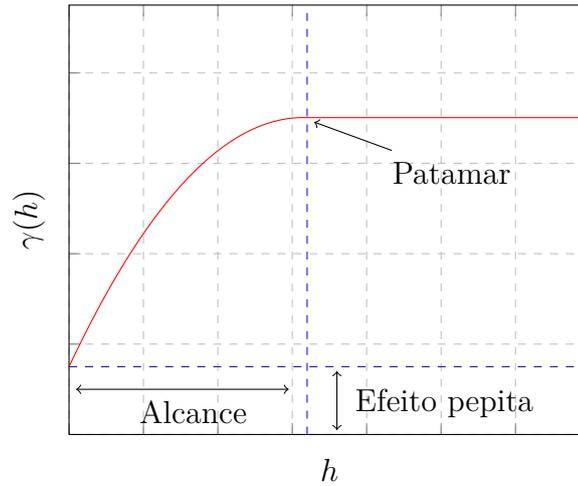
onde,  $h$  é o *lag*,  $x_i$  e  $y_i$  são os valores de *pixels* de uma imagem separados por  $h$ .

O semivariograma normalmente exhibe uma curva característica (Figura 15) na qual a semivariância  $\gamma$  cresce em função da distância. Em determinada distância a semivariância se torna aproximadamente constante. O platô onde  $\gamma$  se estabiliza é denominado patamar.

A distância na qual o patamar é alcançado é denominada alcance. Nesse ponto, a semivariância é igual à variância dos dados da amostra considerada, significando que não existe qualquer relação entre os pares de observações considerados a esta distância (SILVA; CARVALHO; GATTASS, 2005). Já se a distância  $h$  é muito pequena, os pontos comparados são muito semelhantes entre si, o que implica em valores muito baixos de semivariância para essas distâncias iniciais, logo o valor do semivariograma iniciará com valores muito baixos  $\gamma > 0$ . Esse valor, assim como o alcance e o patamar, também representa uma característica do semivariograma denominada efeito pepita.

Para o cálculo do semivariograma, alguns parâmetros são utilizados: a direção, o *lag*,

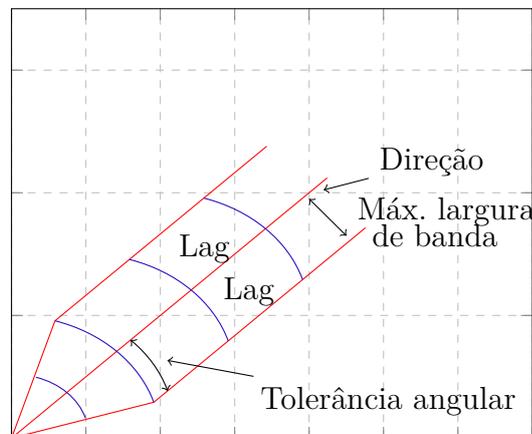
Figura 15 – Curva característica do semivariograma com destaque para o patamar, alcance e pepita.



Fonte – Produzida pelo autor.

a largura de banda máxima, a tolerância do *lag* e a tolerância angular. A definição desses valores limita quais *pixels* serão considerados para o cálculo. Uma ilustração dos limites aplicados por estes parâmetros aos pares de *pixels* pode ser visualizada na Figura 16.

Figura 16 – Parâmetros utilizados para o cálculo do semivariograma (2D).



Fonte – Produzida pelo autor.

### 3.8.2 Semimadograma

O semimadograma é a média da diferença entre os pares da amostra, como uma função de distância e direção (SILVA; CARVALHO; GATTASS, 2005). A função é definida por:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} |x_i - y_i| \tag{3.19}$$

onde  $h$  é o *lag*,  $x_i$  e  $y_i$  são os valores de *pixels* de uma imagem separados por  $h$ .  $N(h)$  é o número de pares na distância  $h$ .

### 3.9 Máquina de Vetores de Suporte

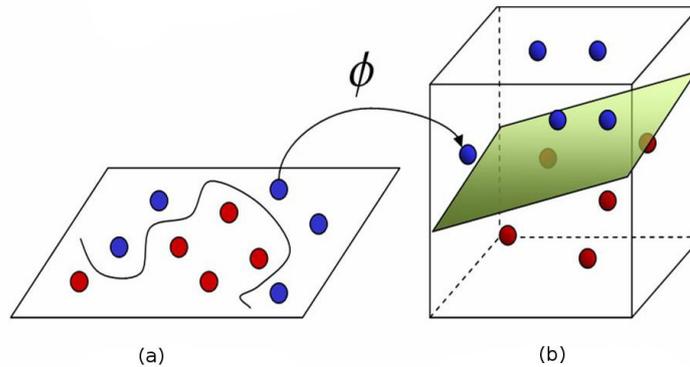
Máquina de Vetores de Suporte (SVM, do inglês *support vector machine*) (VAPNIK, 2013) é um método de aprendizagem supervisionada que pode ser aplicado à classificação e reconhecimento de padrões. Seu funcionamento se baseia em separar dados em um espaço amostral a partir de um conjunto de dados previamente rotuladas. A separação desses dados é feita através de hiperplanos que segmentam o espaço amostral em partições, separando assim os dados em classes distintas (WANG, 2005).

A operação básica da SVM é composta por duas etapas. Primeiramente é realizado o mapeamento dos vetores de entrada (amostras) para um espaço multidimensional de características. Nas aplicações de classificação, como neste trabalho, os vetores de entrada são na verdade, vetores de características dos objetos aos quais se quer classificar. Já o mapeamento para um espaço multidimensional é feito através de uma função de mapeamento  $\phi$  pré-determinada. Após o mapeamento, já no novo espaço multidimensional de características, são construídos hiperplanos de separação entre classes, de modo que a distância entre as classes seja máxima, ou seja, hiperplanos de separação ótimos. A Figura 17 ilustra o mapeamento e a divisão do espaço multidimensional por um hiperplano.

Em casos em que o conjunto de amostras é composto por duas classes separáveis, um classificador SVM é capaz de encontrar um hiperplano baseado em um conjunto de pontos denominados “vetores de suporte”, o qual maximiza a margem de separação entre as classes. Esses hiperplanos são chamados de hiperplanos de margem máxima. Porém, mesmo quando as duas classes não são separáveis, a SVM é capaz de encontrar um hiperplano aproximado através do uso de técnicas de otimização. A Figura 17 ilustra hiperplanos de separação entre duas classes linearmente separáveis. O hiperplano ótimo (linha mais escura), não somente separa as duas classes, mas mantém a maior distância possível com relação aos pontos da amostra.

Em uma descrição mais formal, podemos definir o espaço amostral de entrada como o conjunto  $(x_i, y_i)$ , sendo  $x_i \in \mathbb{R}^n$  o vetor de entrada,  $y_i$  o rótulo das amostras e  $i = 1, 2, \dots, n$  o índice de cada ponto amostral. O objetivo da classificação é estimar a função  $f : \mathbb{R}^n \rightarrow \pm 1$ , que separe corretamente os exemplos de teste em classes distintas. A etapa

Figura 17 – Transformação do espaço amostral (a) para o espaço multidimensional (b) de características através de uma função de mapeamento  $\phi$  e separação das classes por um hiperplano.



de treinamento estima a função  $f(x) = (w \cdot x) + b$ , procurando por valores de  $w$  e  $b$ , de tal forma que satisfaçam a seguinte relação:

$$y_i((w \cdot x_i) + b) \geq 1 \quad (3.20)$$

Os valores ótimos de  $w$  e  $b$  serão encontrados ao minimizar a Equação 3.21, de acordo com a restrição dada pela Equação 3.20.

$$\phi(w) = \frac{w^2}{2} \quad (3.21)$$

A máquina de vetores de suporte possibilita ainda encontrar um hiperplano que minimize a ocorrência de erros de classificação nos casos em que uma perfeita separação entre as duas classes não for possível. Isso graças à inclusão de variáveis de folga, que permitem que as restrições presentes na Equação 3.20 sejam quebradas.

O problema de otimização passa a ser então a minimização da Equação 3.22, de acordo com a restrição imposta pela Equação 3.23.  $C$  é um parâmetro de treinamento que estabelece um equilíbrio entre a complexidade do modelo e o erro de treinamento e deve ser selecionado pelo usuário.

$$\phi(w, \xi) = \frac{w^2}{2} + C \sum_{i=1}^N \xi_i \quad (3.22)$$

$$y_i((w \cdot x_i) + b) + \xi_i \geq 1 \quad (3.23)$$

Através da teoria dos multiplicadores de Lagrange, chega-se:

$$w(\alpha) = \sum_{i=1}^N \alpha_i + \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (3.24)$$

sujeita a:

$$\sum_{i=1}^N \alpha_i x_i = 0, 0 \leq \alpha_i \leq C \quad (3.25)$$

onde  $K$  é uma função núcleo utilizada para que a SVM possa classificar amostras que não são linearmente separáveis, fazendo uma transformação não-linear que transforme o espaço entrada (dados) para um novo espaço (espaço de características). Esse espaço deve apresentar dimensão suficientemente grande para que, através dele, a amostra possa ser separável linearmente.

Dessa maneira, o hiperplano de separação é definido como uma função linear de vetores retirados do espaço de características, em vez do espaço de entrada original. Essa construção depende do cálculo de  $K$  (HAYKIN, 2007). A função  $K$  realiza o mapeamento das amostras para um espaço de dimensão muito elevada sem aumentar a complexidade dos cálculos.

As funções núcleo mais utilizadas são apresentadas na Tabela 2, a seguir:

Tabela 2 – Funções núcleo utilizadas para separar o espaço de características.

Tipo de <i>kernel</i>	Função
Linear	$K(x_i, x_j) = x_i^T x_j$
Poligonal	$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$
Radial	$K(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2), \gamma > 0$
Sigmoide	$K(x_i, x_j) = \tan h(\gamma x_i^T x_j + r)$

### 3.10 Validação de Resultados

Para alcançar o objetivo de desenvolver uma metodologia para o diagnóstico automático de DMRI não-exsudativa, esta tese aborda os problemas da segmentação de camadas da retina e classificação para diagnóstico da doença. A validação dos resultados obtidos é feita a partir de técnicas de avaliação de erro médio e distância de Hausdorff para a segmentação e métricas de desempenho de classificação para o diagnóstico automático. Essas medidas são apresentadas a seguir.

### 3.10.1 Avaliação de erro médio

A comparação entre dois conjuntos de dados observados sobre o mesmo fenômeno pode ser feita a partir do cálculo do erro entre os pares correspondentes (avaliação de erro médio) ou a partir de uma métrica de distância entre os próprios conjuntos. A seguir falaremos de duas estratégias: a avaliação do erro médio (sumarização de erros entre pares correspondentes) e a medida de distância de Hausdorff (distância entre conjuntos).

Dadas duas variáveis contínuas  $A$  e  $B$ , o erro entre pares correspondentes é dado por  $e_i = a_i - b_i$ . A sumarização dos valores  $e_i (i = 1, 2, 3, \dots, n)$  permite estabelecer uma medida do erro médio entre as variáveis. Essa medida é representada de forma genérica por (WILLMOTT; MATSUURA, 2005):

$$\bar{e}_\gamma = \left[ \frac{\sum_{i=1}^n w_i |e_i|^\gamma}{\sum_{i=1}^n w_i} \right]^{1/\gamma} \quad (3.26)$$

onde  $\gamma \geq 1$  e  $w_i$  é um peso designado para cada  $|e_i|^\gamma$  referente a influência daquele elemento no conjunto.

A partir da representação genérica oferecida pela Equação 3.26, obtemos expressões comumente usadas, como a raiz do erro quadrático médio (RMSE, do inglês “*Root Mean Square Error*”), o erro médio (EM) e o erro médio absoluto (EMA).

A medida *RMSE* é obtida pela Equação 3.26 quando  $\gamma = 2$  e  $w_i = 1, \forall i \in \{1, 2, \dots, n\}$  correspondendo a:

$$RMSE = \left[ n^{-1} \sum_{i=1}^n |e_i|^2 \right]^{1/2} \quad (3.27)$$

O EM é uma medida mais simples obtida quando  $\gamma = 1$  e  $w_i = 1, \forall i \in \{1, 2, \dots, n\}$  e os sinais dos erros são preservados. O cálculo de EM é dado por:

$$EM = n^{-1} \sum_{i=1}^n e_i = \bar{A} - \bar{B} \quad (3.28)$$

Já o EMA mantém os  $\gamma = 2$  e  $w_i = 1, \forall i \in \{1, 2, \dots, n\}$  e despreza os sinais dos erros, sendo representado por:

$$EMA = n^{-1} \sum_{i=1}^n |e_i| \quad (3.29)$$

A medida de EM é útil para revelar o viés médio dos erros entre os pares corres-

pondentes. O valor de RMSE gera valores mais consistentes para erros com distribuição gaussiana (CHAI; DRAXLER, 2014). Já o cálculo de EMA permite uma interpretação intuitiva na mesma escala dos valores a serem comparados (WILLMOTT; MATSUURA, 2005).

### 3.10.2 Distância de Hausdorff

Diferentemente de outros métodos de comparação de similaridades entre conjuntos de dados, para o cálculo da medida de distância de Hausdorff não se estabelece qualquer pareamento de pontos de  $A$  com pontos de  $B$ .

A medida de distância de Hausdorff (CONCI et al., 2015; CONCI; KUBRUSLY, 2017) quantifica o quanto dois conjuntos de pontos estão distantes um do outro podendo ser utilizada para expressar a semelhança entre dois objetos sobrepostos, sendo definida por  $h: \wp(X) \setminus \emptyset \times \wp(X) \setminus \emptyset \rightarrow \overline{\mathbb{R}}$  é definida para  $A, B \in \wp(X)$ :

$$h(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A) \right\} \quad (3.30)$$

$$= \max \left\{ \sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b) \right\}. \quad (3.31)$$

onde,  $d$  é uma medida de distância entre dois pontos (*e.g.* a medida de distância euclidiana).

A função  $h(A, B)$  determina a distância máxima entre os dois conjuntos de pontos através da medida da distância de um ponto de  $A$  que está mais afastado de qualquer ponto de  $B$ . Se  $h(A, B) = d$ , então cada ponto  $a \in A$  deve estar no máximo a uma distância  $d$  de qualquer ponto  $b \in B$  e *vice-versa*.

### 3.10.3 Validação do Classificador

A validação do classificador é essencial para verificar sua eficiência em situações do mundo real. Em problemas relacionados à classificação de dados e diagnóstico de doenças, como neste caso em que os exames devem ser categorizados em duas classes, como portadores ou não de DMRI, a validação é baseada em determinar o quão bom é o método proposto nesta tese para essa tarefa.

Nesse tipo de problema, existe a presença de uma variável preditora (resultado do teste) e uma variável resultante (ALMEIDA, 2013). A variável preditora é analisada considerando-se duas hipóteses (resultado positivo e negativo) e definindo-se os critérios

de corte que confirmam ou não a ocorrência de um evento. A variável resultante do teste diagnóstico é a confirmação de uma das duas hipóteses mediante a comparação com o resultado de um exame considerado padrão-ouro ("*Gold Standard*").

Na avaliação de um teste de classificação, existem quatro possibilidades:

- i. O teste é positivo e esse resultado corresponde ao padrão-ouro - Verdadeiro Positivo (VP);
- ii. O teste é positivo e esse resultado não corresponde ao padrão-ouro - Falso Positivo (FP);
- iii. O teste é negativo e esse resultado não corresponde ao padrão-ouro - Falso Negativo (FN) e
- iv. O teste é negativo assim como apontado pelo padrão-ouro - Verdadeiro Negativo (VN).

### 3.10.3.1 Sensibilidade, Especificidade e Acurácia

Para avaliar o desempenho de modelos de classificação, normalmente são utilizadas estatísticas descritivas, tais como, Sensibilidade (S), Especificidade (E) e Acurácia (A).

A sensibilidade de um teste refere-se à razão entre a quantidade de diagnósticos corretos (VP) e o número total de casos positivos (VP + FN). Assim, esse indicador representa a eficácia do método em classificar diagnósticos positivos (a presença de determinada doença ou condição) corretamente. É expresso por:

$$S = \frac{VP}{VP + FN} \quad (3.32)$$

De maneira semelhante, a especificidade se refere à proporção entre a quantidade de instâncias corretamente classificadas pelo teste como negativo (VN) e o número total de instâncias negativas (VN + FP). Essa medida serve para indicar a eficácia do método em identificar corretamente casos que não apresentam a patologia estudada. A especificidade  $E$  é formulada por:

$$E = \frac{VN}{VN + FP} \quad (3.33)$$

Já a acurácia de um teste remete à taxa de classificação correta sendo calculada dividindo-se o número de casos corretamente classificados ( $VN + VP$ ) pelo número total de casos ( $VP + VN + FP + FN$ ):

$$A = \frac{VN + VP}{VP + VN + FP + FN} \quad (3.34)$$

### 3.10.3.2 Concordância Kappa de Cohen

O valor Kappa de Cohen (MCHUGH, 2012) mede a concordância entre dois árbitros (e.g. classificadores humanos *vs.* automáticos) que classificam  $N$  instâncias em categorias mutuamente exclusivas. É dado por:

$$\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (3.35)$$

onde,  $Pr(a)$  corresponde à concordância entre os classificadores (equivalente a acurácia) e  $Pr(e)$  é a probabilidade de uma classificação aleatória. Uma interpretação do valor  $\kappa$  é apresentada na Tabela 3. Quanto mais o valor  $\kappa$  se aproxima da unidade, maior é a concordância entre as classificações comparadas.

Tabela 3 – Interpretação do valor de Kappa

$\kappa$	Concordância
$< 0$	Fraca
0,01–0,20	Leve
0,21– 0,40	Razoável
0,41–0,60	Moderada
0,61–0,80	Substancial
0,81–0,99	Quase perfeita

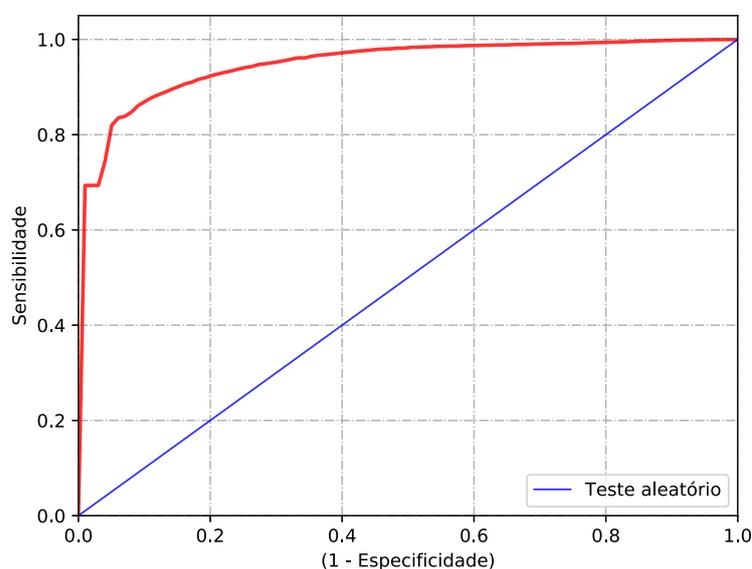
Fonte – Landis e Koch (1977)

### 3.10.4 Curva característica de operação do receptor (ROC)

A curva característica de operação do receptor (ROC, do inglês *Receiver Operating Characteristic*) é um método utilizado para avaliar o desempenho de um teste dicotômico. Se considerarmos que o resultado do teste será um intervalo contínuo, deve existir um valor que diferencie os indivíduos pertencentes aos dois grupos, chamado de “ponto de corte”.

Considerando-se todos os “pontos de corte” possíveis para aquele teste, pode-se delinear uma curva (ROC) que relaciona sensibilidade e especificidade em cada um desses pontos (PRATI; BATISTA; MONARD, 2008). Um teste de melhor qualidade é aquele cuja área sob a curva ROC aproxima-se de 1,0. Já uma área com valor próximo de 0,5 representa um teste sem significância, cujo resultado é aleatório (Tabela 4). A Figura 18 mostra a construção de uma curva ROC (em vermelho). A linha em azul é uma referência para um classificador aleatório. A Tabela 4 fornece uma interpretação para o valor da área sob a curva ROC (AUROC, do inglês “*area under ROC curve*”).

Figura 18 – A curva ROC (vermelho) representando a relação entre a sensibilidade e a especificidade do classificador.



Fonte – Produzida pelo autor.

Tabela 4 – Qualidade do diagnóstico de acordo com a área sob a curva ROC.

Área sob a Curva	Qualidade do teste
0,9 – 1,0	Excelente
0,8 – 0,9	Bom
0,7 – 0,8	Regular
0,6 – 0,7	Ruim
0,5 – 0,6	Insignificante

Fonte – Rocha (2014)

## 4 MÉTODO PARA A SEGMENTAÇÃO DE CAMADAS DA RETINA

Este capítulo apresenta um método para a segmentação automática de camadas da retina humana em *B-scans* de OCT. O objetivo é realizar a segmentação das bordas que delimitam a RT e o EPR. Para isso, são identificadas as bordas internas da MLI e do EPR e borda externa da MB. Com a obtenção das bordas do EPR e da MB é possível delimitar o EPR-CD (FARSIU et al., 2014). A delimitação dessa camada pode ser utilizada para avaliação de variações locais de espessura e alterações morfológicas causadas pela DMRI.

O método proposto realiza a busca pelas bordas das camadas utilizando uma rede de aprendizado profundo, baseada no roteamento entre cápsulas (CapsNet), gerando mapas de probabilidade de localização dessas bordas. Em seguida, um algoritmo de corte de grafos permite a identificação de cada borda.

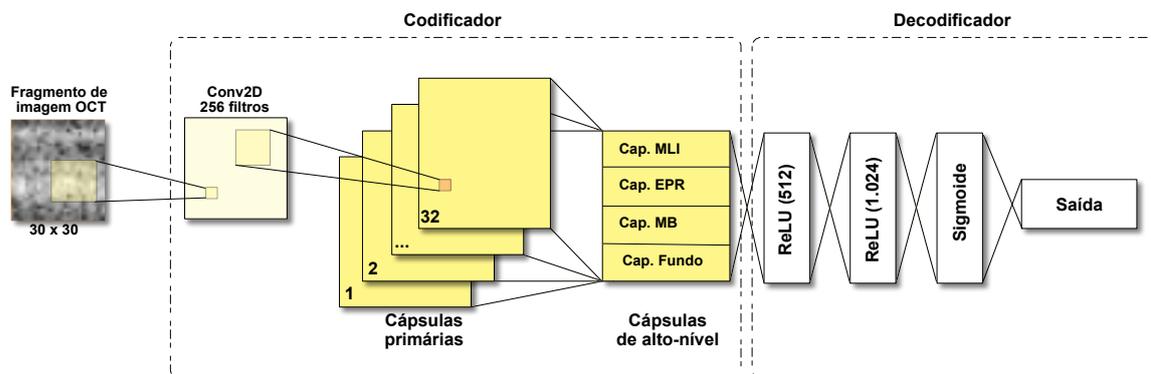
O processo é dividido em duas etapas: treinamento da rede e a segmentação de bordas. No centro dessas etapas está a rede de aprendizado profundo CapsNet. Ela é responsável por aprender as características das bordas das camadas procuradas, a partir de um conjunto de imagens de treinamento, e identificar corretamente essas bordas em um conjunto de imagens de testes. Nas seções seguintes é apresentada a arquitetura da rede utilizada, e são detalhadas as etapas de treinamento e segmentação de bordas propostas na metodologia.

### 4.1 Arquitetura da rede de cápsulas para a segmentação

A arquitetura da rede utilizada nesta etapa é baseada na arquitetura apresentada por Sabour, Frosst e Hinton (2017). A rede é composta por um codificador e um decodificador (Figura 19). O codificador recebe como entrada regiões da imagem para a extração de descritores vetoriais. Já o decodificador é responsável por identificar as quais bordas (ou região de fundo) cada parte da imagem corresponde e fornecer a probabilidade de pertencimento às classes como saída da rede.

O codificador é composto por três camadas e lida com uma pequena região da imagem por vez, a qual denominamos fragmento. Sua primeira camada é responsável por detectar características básicas de cada fragmento de  $30 \times 30$  *pixels*. Essa camada opera uma convolução com operador ReLU de 256 núcleos de tamanho  $16 \times 16$ , convertendo os valores de intensidades de *pixel* em descritores de características locais de baixo nível.

Figura 19 – Arquitetura da rede baseada em CapsNet para utilizada para a classificação de fragmentos de imagens de OCT.



Fonte – Produzida pelo autor.

A segunda camada contém 32 cápsulas primárias. Sua função é gerar combinações de descritores a partir das características básicas extraídas pela camada anterior. As cápsulas primárias compõem o nível mais baixo de entidades multidimensionais. Elas estão dispostas em 16 canais, onde cada cápsula principal aplica 8 filtros convolucionais de tamanho  $9 \times 9$  e passo 2 à saída da camada anterior.

A última camada do codificador é composta por quatro cápsulas de alto-nível. As três primeiras são responsáveis por codificar descritores vetoriais que representem as bordas das camadas que serão segmentadas (MLI, EPR e MB) e a última corresponde a representações das demais áreas da imagem que não são identificados como bordas (“fundo”).

O algoritmo de roteamento dinâmico (SABOUR; FROSST; HINTON, 2017) é empregado para codificar as entradas das cápsulas primárias em saídas das cápsulas de alto-nível (de 16 dimensões). Os comprimentos desses vetores de previsão são usados no decodificador a fim de determinar a classe prevista para cada fragmento.

O decodificador é composto por três camadas. Ele recebe a saída vetorial do codificador e o decodifica em um rótulo que corresponde à borda ou região de fundo do fragmento.

## 4.2 Treinamento

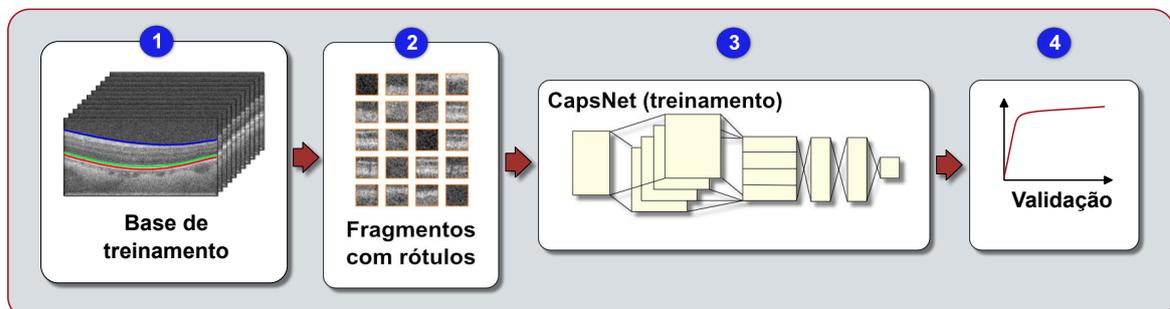
O treinamento é realizado a partir de um conjunto de *B-scans* de OCT reservados para este fim. Cada imagem desse conjunto possui marcações das bordas da MLI, do EPR e da MB. As imagens são divididas em  $N$  fragmentos  $x_n$  de  $30 \times 30$  *pixels* com rótulos  $y_n \in \{1, 2, 3, 4\}$  que correspondem a classes possíveis para cada fragmento: MLI(1), EPR(2), MB(3) e fundo(4).

Para cada três fragmentos centralizados nas três bordas, é selecionado um fragmento da região de fundo com o objetivo de proporcionar representações balanceadas para as três classes. Antes de serem submetidos ao treinamento, cada fragmento  $x_i$  tem seu valor médio subtraído ( $x_n - \bar{x}_n$ ).

O processo de treinamento é apresentado na Figura 20 em forma de fluxo com indicação da ordem das etapas. Após o processamento dos fragmentos da imagem, esses são submetidos à CapsNet para o treinamento, ou seja, o ajuste dos pesos da rede, de forma que a rede possa classificar corretamente os fragmentos correspondentes às bordas.

O treinamento é realizado em um número determinado de épocas. A cada época, todo o conjunto de fragmentos de treino é exposto à rede de forma a ajustar seus pesos através do algoritmo de otimização Adam (KINGMA; BA, 2015) e o desempenho da rede é medido em termos de acurácia. É esperado que a acurácia aumente com a sucessão de épocas. A validação da rede é dada por sua capacidade de classificar corretamente os fragmentos da imagem (valor de acurácia  $\geq 0.95\%$ ).

Figura 20 – Fluxo do processo de treinamento da CapsNet para o reconhecimento de fragmentos de imagens de OCT.

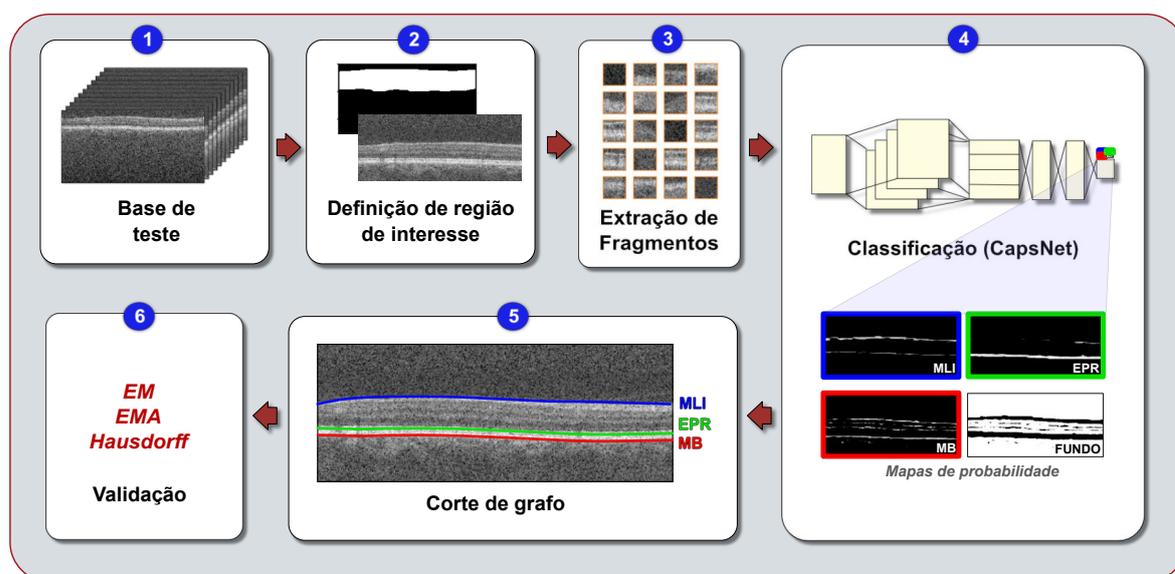


Fonte – Produzida pelo autor.

### 4.3 Segmentação das bordas

Após o treinamento, a rede CapsNet é utilizada como um classificador associado a um algoritmo de corte de grafos para a segmentação das bordas das camadas de interesse. O fluxo desse processo é apresentado na Figura 21. Cada *B-scan* tem o processamento limitado a uma região de interesse e depois é dividida em fragmentos para o processamento pela rede. A CapsNet classifica cada fragmento e gera um mapa de probabilidades para a localização de cada borda e das regiões consideradas como fundo. Por fim, um algoritmo de corte de grafos aplicado em cada mapa de probabilidade permite identificar cada borda procurada. Cada etapa é detalhada a seguir.

Figura 21 – Fluxo do processo de segmentação das bordas da MLI, EPR e MB com a utilização de CapsNet.



Fonte – Produzida pelo autor.

#### 4.3.1 Definição automática da região de interesse

A definição de uma região de interesse visa reduzir a área a ser processada. Isso evita processamento desnecessário de regiões muito distantes da retina que representam o “fundo” da imagem.

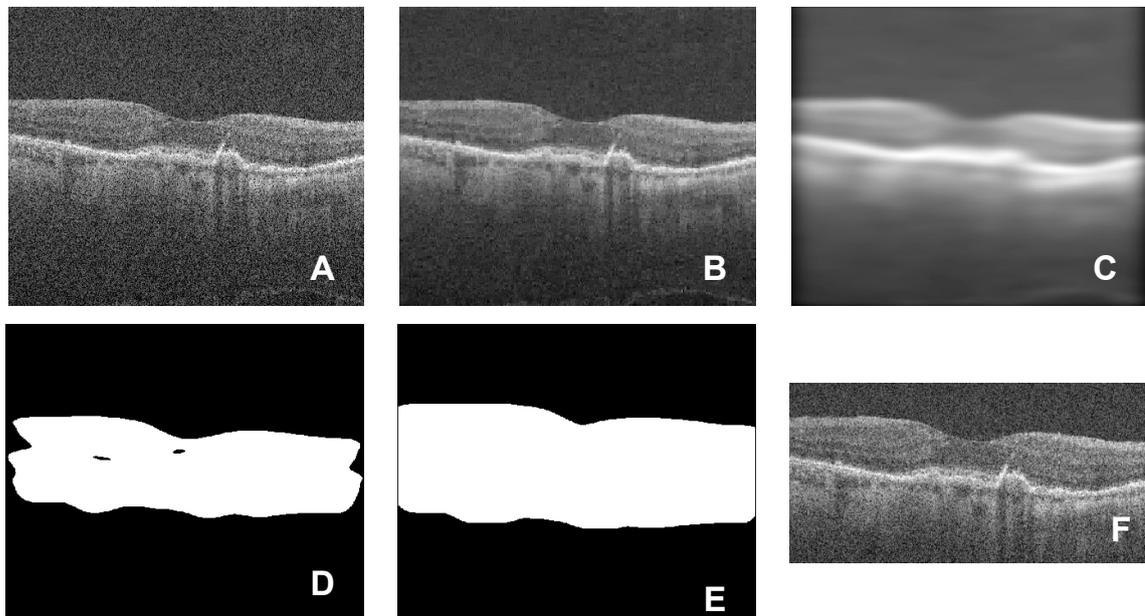
O processo para a delimitação da região de interesse é feito em cinco etapas. A primeira consiste em uma filtragem de ruídos utilizando o filtro *wavelet* com o parâmetro<sup>1</sup>

<sup>1</sup> A escolha deste parâmetro é apresentada no Apêndice A

$\sigma = 0.11$ . Em seguida, a imagem é processada por um filtro gaussiano com um núcleo retangular ( $\sigma_x = 20$  e  $\sigma_y = 5$ ). Essa filtragem permite tornar as regiões da imagem mais homogêneas para a etapa seguinte. O próximo passo é a criação de uma máscara binária para delimitar a região de interesse. Isso é realizado aplicando um limiar correspondente a 55% da intensidade máxima existente na imagem. Valores maiores que o limiar são considerados como parte da retina e valores menores como “fundo”. A máscara binária, gerada pela binarização da imagem, sofre operações morfológicas de dilatação (com elemento de  $40 \times 40$  *pixels*) e fechamento (com elemento de  $20 \times 20$  *pixels*). Por fim, são calculadas a altura máxima e mínima para a máscara e a imagem original é reduzida a uma região mínima para o processamento com margem de segurança de 50 *pixels* para cima e para baixo.

A Figura 22 apresenta os resultados intermediários do processo de delimitação da região de interesse para uma *B-scan* de OCT. A imagem *A* apresenta a *B-scan* original. De *B* até *E* são apresentadas as etapas intermediárias, e, em *F*, a imagem recortada pronta para a etapa seguinte de extração de fragmentos.

Figura 22 – Sequência de processamento definição da região de interesse. A) imagem original; B) diminuição de ruídos (filtro wavelet); C) filtro gaussiano; D) binarização; E) dilatação e fechamento e F) região de interesse delimitada.



Fonte – Produzida pelo autor.

### 4.3.2 Extração de fragmentos

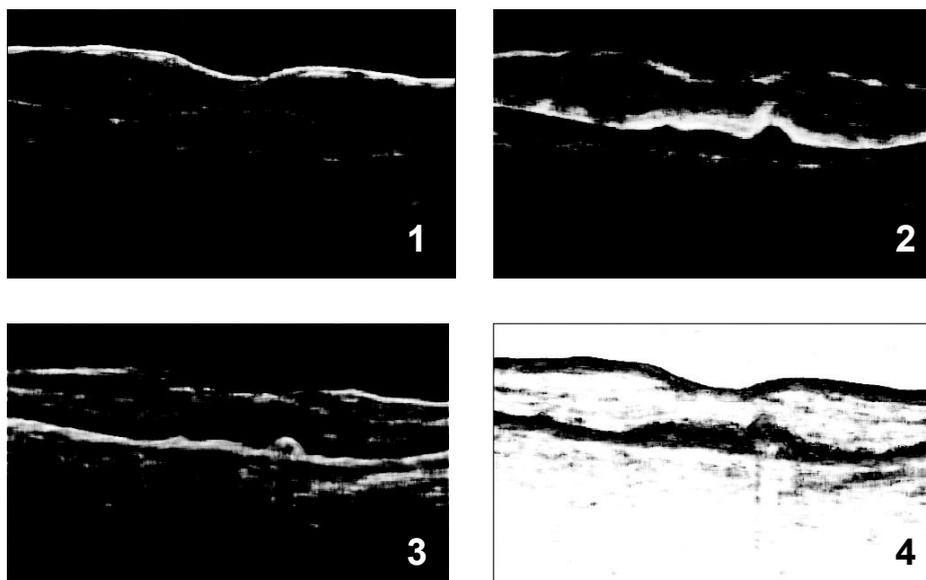
Assim como efetuado para o treinamento da rede, a região de interesse é dividida em fragmentos de  $30 \times 30$  *pixels* com distanciamento de 1 *pixel*. Todos os *pixels* de cada fragmento têm seu valor de intensidade subtraído do valor correspondente à media de intensidades do referido fragmento. O distanciamento pequeno entre os fragmentos gera a sobreposição de parte deles.

### 4.3.3 Classificação

Os fragmentos extraídos são submetidos à CapsNet treinada para classificação. A saída da rede fornece quatro mapas de probabilidade, um para cada borda e outro para o “fundo” (regiões da imagem que não fazem parte da retina ou não estão centralizadas nas bordas de interesse). Os mapas gerados podem ser representados como imagens bidimensionais onde cada *pixel*  $p$  corresponde à probabilidade  $P(x_n)$  de um fragmento  $x_n$  pertencer à determinada borda ou à região de fundo ( $0 \geq P(x_n) \leq 1$ ).

A Figura 23 apresenta um exemplo de mapas de probabilidade gerados pela etapa de classificação de fragmentos. As áreas mais claras correspondem a regiões com maior probabilidade de pertencerem às bordas pesquisadas (imagens 1 a 3). A imagem 4 representa o mapa de probabilidade de regiões de “fundo”.

Figura 23 – Mapas de probabilidade para as bordas das camadas retinianas: 1) borda interna da MLI; 2) borda interna do EPR; 3) borda externa da MB e 4) região de "fundo".



Fonte – Produzida pelo autor.

#### 4.3.4 Corte de Grafo

Para a identificação das bordas, os mapas de probabilidade são representados como um grafo<sup>2</sup>, onde a probabilidade  $Pr(x_n)$  de cada fragmento corresponde a um nó e o custo entre dois nós adjacentes é definido como  $w(x_n, x_{n+1}) = 1 - Pr(x_{n+1})$ . Essa operação atribui um peso menor para arestas que levam a pontos no mapa que apresentam probabilidade elevada de pertencerem a uma borda. Já as arestas que levam a pontos com baixa probabilidade recebem pesos maiores.

O grafo é implementado pela lógica de lista de adjacência. Para cada nó, são criadas adjacências apenas para os nós imediatamente vizinhos. Uma restrição é que, para cada nó, não possam existir adjacências para outros nós localizados na mesma coluna da imagem. Essa restrição garante que para cada coluna da imagem (*A-scan*), haverá apenas um *pixel* selecionado para compor uma borda específica.

Seguindo a lógica adotada em (CHIU, 2014), acrescentamos duas colunas temporárias na imagem a ser processada: uma coluna à direita e outra à esquerda. O objetivo do acréscimo dessas colunas é que sejam utilizadas para acomodar, respectivamente, o ponto inicial e o ponto final que servirão para o algoritmo de seleção do menor caminho entre esses pontos. Os pontos inicial e final selecionados são os primeiros pontos da coluna inicial e final, respectivamente, embora seja possível estabelecer esses pontos em outros locais. Essas colunas não estão sujeitas à restrição anterior e são estabelecidas adjacências para nós na mesma coluna com pesos próximos a zero. As colunas adicionadas são desconsideradas após o término do processamento.

Para detectar uma borda é necessário efetuar uma busca no grafo de forma a encontrar o melhor caminho que leva do ponto inicial ao ponto final. A busca é efetuada de forma simples através do algoritmo de Dijkstra (CORMEN et al., 2002). O melhor caminho corresponde ao de menor custo, ou seja, aquela cujas somas dos pesos das arestas percorridas seja o menor possível.

Os mapas de probabilidade são processados na seguinte ordem: MLI, EPR e MB. O mapa que se refere à probabilidade de fragmentos pertencerem à classe “fundo” é descartado por não representar nenhuma borda em particular.

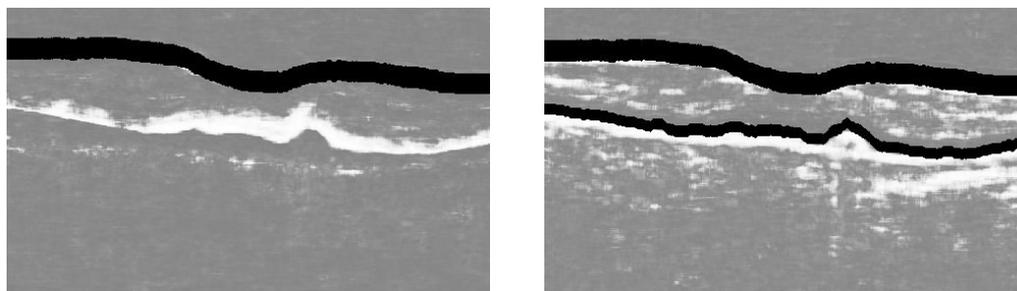
Uma vez encontrado um caminho sobre a imagem que corresponde à borda da MLI, o mapa da borda do EPR é alterado para o próximo processamento. Os pesos das arestas

<sup>2</sup> A segmentação de imagens de OCT foi apresentada inicialmente por Chiu et al. (2010). Nesta tese, utilizamos a técnica modificada para a utilização com os mapas de probabilidade providos pela CapsNet

que compõem o caminho que segmentou a MLI e as arestas muito próximas recebem um peso elevado ( $10^4$ ), para que na próxima busca pelo melhor caminho, essa região seja desconsiderada e não exerça influência na detecção da próxima borda. Isso é necessário porque as bordas representam variações de intensidade que podem encontrar semelhanças em outros pontos. Essa abordagem evita que uma borda seja detectada mais de uma vez.

Além disso, essa estratégia impede o cruzamento de caminhos na imagem, que poderia permitir a detecção de bordas que se cruzam, um comportamento que não existe na realidade. Esse processo de busca pelo melhor caminho e elevação dos pesos do caminho encontrado é repetido no mapa da MB, para a busca dessa borda. Na Figura 24 é apresentada a ilustração desse processo. A imagem à esquerda corresponde ao mapa de probabilidade utilizado para encontrar a borda interna do EPR. A área mais escura corresponde à região onde foi segmentada a borda da MLI que recebeu pesos elevados a fim de ser excluída da próxima busca. Na imagem à direita está o mapa de probabilidade para a segmentação da borda da MB. Antes de iniciar a busca por essa borda, é feita a exclusão das regiões onde foram segmentadas bordas anteriores (áreas escuras).

Figura 24 – Aplicação de restrições aos mapas de probabilidade da borda do EPR (à esquerda) e da borda da MB (à direita).



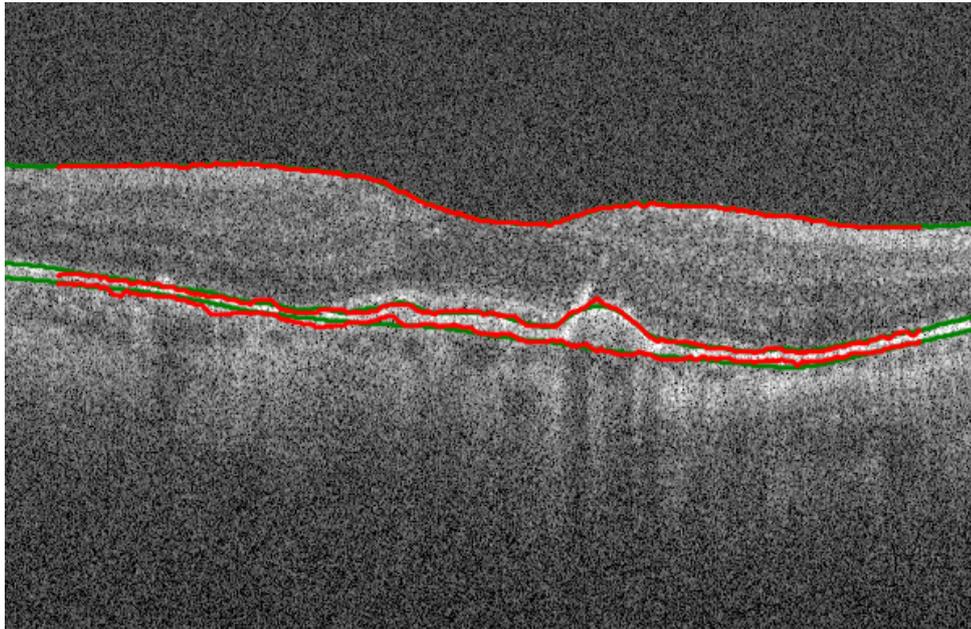
Fonte – Produzida pelo autor.

Na Figura 25 é apresentada o fim do processo. As bordas de cor vermelha indicam a segmentação esperada por esta metodologia.

#### 4.3.5 Estratégia para segmentação em duas resoluções

Um problema de segmentação da borda da MB pode ocorrer na presença de drusas. Essas estruturas podem atenuar o sinal luminoso que chega a MB, diminuindo o contraste e dificultando a segmentação da borda. Em alguns casos, pode ocorrer, erroneamente, a segmentação da borda externa do EPR sem incluir o complexo de drusas (Figura 26 - A).

Figura 25 – Bordas obtidas após o corte de grafos (curvas vermelhas).



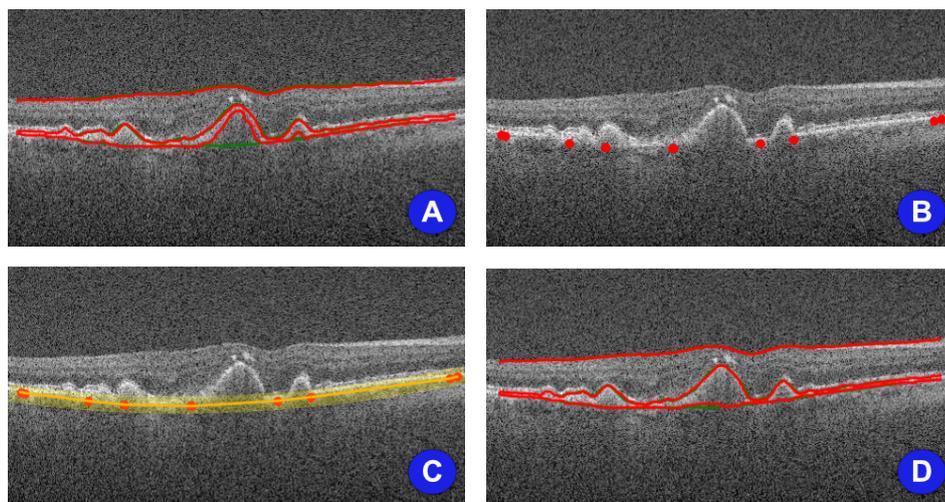
Fonte – Produzida pelo autor.

Uma estratégia para mitigar esse problema é realizar a segmentação em duas etapas. A primeira etapa é processada em baixa resolução ( $1/4$  da resolução original) e define uma segmentação inicial a ser refinada na etapa seguinte.

Na segunda etapa, para cada borda segmentada na primeira etapa, é definida uma região com 8 *pixels* acima e abaixo para efetuar a segmentação novamente. Entretanto, para a definição da região de busca da borda da MB, detectam-se os pontos mais externos da borda segmentada em baixa resolução utilizando-se a técnica de fecho convexo (GONZALEZ; WOODS, 2006), e determina-se uma nova borda aproximada, interpolando linearmente os valores entre os pontos mais externos. Assim, define-se uma nova região de busca para ser processada na resolução original.

A Figura 26 ilustra o processo de delimitação de uma nova região de busca para a borda da MB. A imagem *A* nos permite perceber que a atenuação causada por uma drusa grande na região central do *B-scan* impediu a segmentação correta da borda da MB (marcação em verde). A imagem *B* mostra a detecção dos pontos mais externos da borda. Na imagem *C*, é definida uma região (na cor amarela) para uma nova procura pela borda da MB. Por fim, na imagem *D*, apresentamos o resultado do processo, que corrigiu o problema de segmentação da imagem *A*. Essa estratégia, além de permitir a correção do problema de segmentação da MB causado pela presença de drusas grandes, ainda diminuiu em cerca de seis vezes o tempo de execução necessário.

Figura 26 – Delimitação de região de busca a resolução do problema de segmentação da borda da MB.



Fonte – Produzida pelo autor.

Para avaliação e validação do método proposto, são empregadas as medidas de Erro Médio (EM), Erro Médio Absoluto (EMA) e Distância de Hausdorff.

## 5 DIAGNÓSTICO DE DMRI UTILIZANDO DESCRITORES GEO-ESTATÍSTICOS

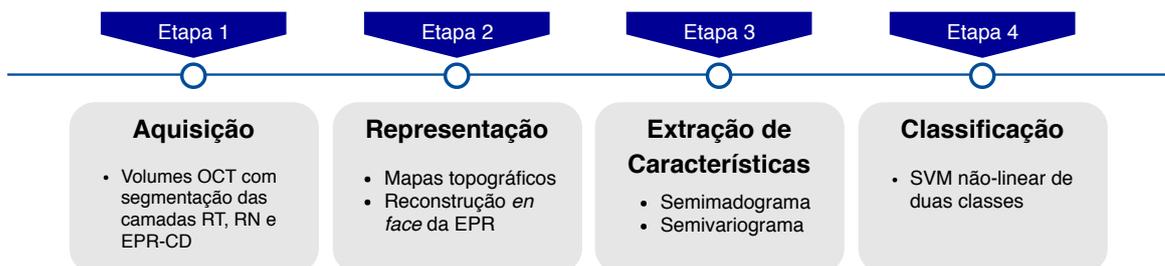
Este capítulo apresenta o método proposto para a classificação de exames de OCT da retina humana entre acometidas ou não por DMRI não-exsudativa, através de descritores geostatísticos semivariogram e semimadograma.

Para a treinamento e validação do método, são necessárias imagens com segmentação das bordas que delimitam a RT e o EPR-DC. A segmentação dessas estruturas permite a representação dos volumes OCT em mapas topográficos, dos quais serão extraídos os descritores geostatísticos.

Além da representação por mapas topográficos de espessura, também realizamos a reconstrução da *en face* do EPR, obtendo um mapa que corresponde à intensidade dos *pixels* na borda da camada. Com isso, foi possível realizar experimentos de classificação com SVM, de forma a determinar se os descritores geostatísticos empregados permitem a diferenciação das imagens acometidas pela doença, e qual combinação mapa/descritor geostatístico possui maior acurácia. A avaliação do desempenho do método permitiu uma comparação com métodos clássicos de descrição de textura e outros trabalhos relacionados.

Uma representação simplificada das etapas que compõem a metodologia proposta pode ser vista na Figura 27 e são descritas a seguir.

Figura 27 – Fluxograma básico da metodologia proposta para a classificação de imagens de OCT.



Fonte – Produzida pelo autor.

### 5.1 Aquisição das imagens

Para a metodologia proposta, as imagens devem ser do tipo SD-OCT com tamanho mínimo *B-scans* de  $1.000 \times 512$  *pixels*. A aquisição deve ser centrada na mácula com

diâmetro mínimo de 5 *mm* sem a presença de artefatos e com uma quantidade mínima de 100 *B-scans* por exame para a reconstrução volumétrica da retina.

## 5.2 Representação

Esta etapa visa construir representações em duas dimensões (2D) do volume de OCT (3D) a partir das bordas da MLI, do EPR e da MB. Isso permite representar o volume através de mapas topográficos (mapas de espessuras) da RT e do EPR-CD, assim como reconstruir *en face* a borda da camada do EPR com seus valores de reflectância.

### 5.2.1 Mapa topográfico

A representação do volume pelo mapa topográfico consiste em calcular a espessura para cada ponto da superfície da retina. Isso se dá pela identificação da quantidade de *voxels* existentes entre duas bordas.

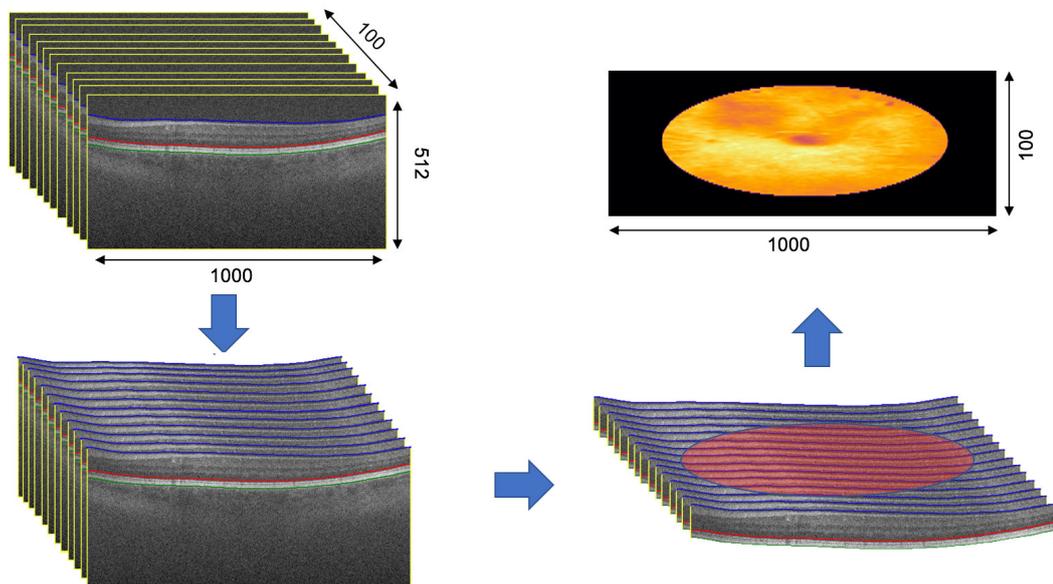
A partir das superfícies que definem as bordas das camadas, são geradas três novas superfícies  $L_i(x, y)$ , onde  $i \in \{1, 2, 3\}$  para representar cada volume OCT. Os valores  $z_i = L_i(x, y)$  correspondem à profundidade axial em um ponto com coordenadas  $(x, y)$  em uma subestrutura da retina. Isso é feito pela soma dos *voxels* existentes, ou simplesmente  $L_1 = B_3 - B_1$ ,  $L_2 = B_2 - B_1$  e  $L_3 = B_3 - B_2$ .

O primeiro mapa topográfico gerado (superfície  $L_1$ ) representa a espessura da retina total, que corresponde à quantidade de *voxels* existentes entre a borda da MLI e da MB. O segundo mapa (superfície  $L_2$ ) corresponde ao mapeamento da espessura da RN, região compreendida entre a MLI e a borda interna do EPR. O último mapa topográfico gerado (superfície  $L_3$ ) corresponde ao EPR. A Figura 28 sumariza o processo de geração dos mapas topográficos e a Figura 29 ilustra as bordas das camadas retinianas utilizadas na criação dos mapas topográficos.

### 5.2.2 Reconstrução “*en face*” da borda do EPR

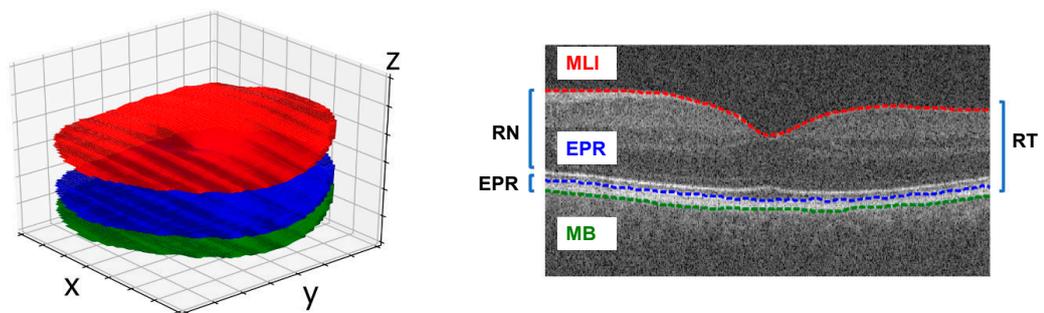
A reconstrução *en face* da borda do EPR gera uma representação bidimensional do volume OCT em que são levantados os valores de reflectância existentes nessa borda. Assim, é extraída uma imagem que guarda semelhança a uma fotografia de fundo do olho. No entanto, cada *pixel* apresentado corresponde à intensidade do *voxel* situado sobre a borda.

Figura 28 – Processo de construção do mapa topográfico para a retina total. As partes da imagem referentes ao fundo e às demais camadas são removidas e é calculada a espessura para cada ponto.



Fonte: Produzida pelo autor.

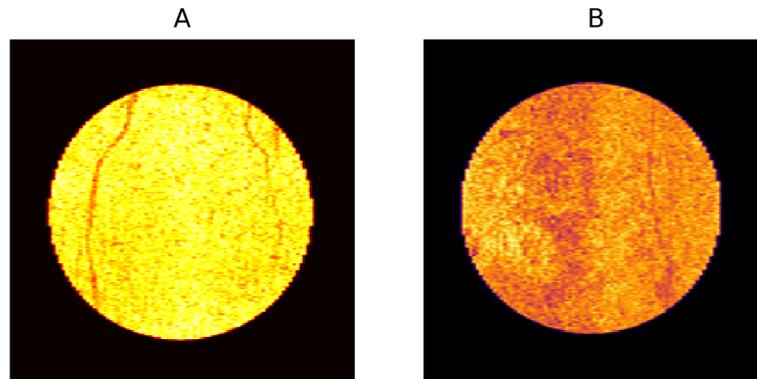
Figura 29 – Bordas utilizadas no processo de geração dos mapas topográficos. À esquerda, representação tridimensional. À direita, marcação das bordas em uma *B-scan*.



Fonte: Produzida pelo autor.

A Figura 30 apresenta a reconstrução *en face* da borda do EPR para um volume com DMRI e outro considerado normal. É possível perceber visualmente variações na textura entre as duas imagens, onde as regiões mais escuras correspondem a pontos de menor reflectância, como vasos sanguíneos ou formação de drusas.

Figura 30 – Reconstrução *en face* da borda do epitélio pigmentado para um volume de SD-OCT sem DMRI (A) e um volume com DMRI (B). Visualmente é possível perceber pequenas variações na textura.



Fonte: Produzida pelo autor.

### 5.3 Extração de características

Nesta etapa são produzidos descritores destinados à composição de vetores de características a serem utilizados pela etapa de classificação. Foram utilizados o semivariograma ( $\gamma$ ) e o semimadograma ( $m$ ) calculados em quatro direções ( $0^\circ, 45^\circ, 90^\circ$  e  $135^\circ$ ) com  $lag = h \in \{0, 13, 27, \dots, 200\}$ , a tolerância para o ângulo de direção foi igual a  $7^\circ$  e a largura de banda máxima de 3 *pixels*. Para cada volume é gerado um vetor de características composto pelos valores do semivariograma e semimadograma calculados para cada *lag*.

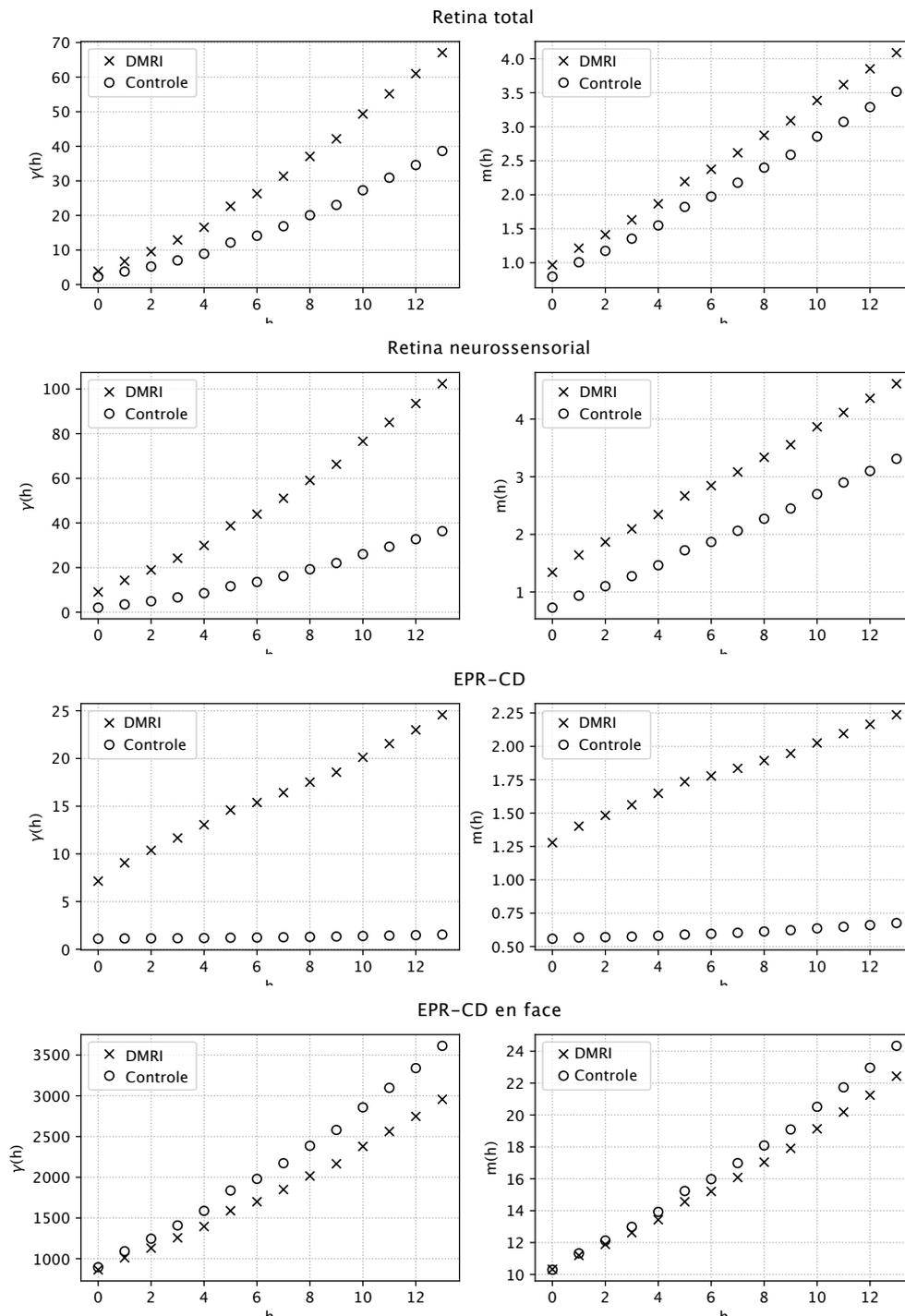
Para ilustração da variação desses valores, apresentamos na Figura 31 um gráfico com os valores obtidos para dois volumes: um diagnosticado com DMRI e outro do grupo de controle. Percebemos, para este exemplo, uma maior variação nos gráficos do semivariograma do volume com DMRI quando aplicado sobre o mapa topográfico do EPR-CD.

### 5.4 Classificação utilizando Máquina de Vetores de Suporte (SVM)

A classificação é feita através da técnica de aprendizado computacional Máquina de Vetores de Suporte. A função *kernel* empregada é a função linear  $K(x_i, x_j) = x_i^T x_j$ . O classificador recebe como entrada vetores de características extraídas na etapa anterior.

Para avaliação da capacidade de generalização da metodologia em desenvolvimento, utilizamos o método *k-fold*, um caso específico de validação cruzada (KOHAVI et al.,

Figura 31 – Semivariograma experimental calculado sobre os mapas topográficos da retina humana com ângulo de direção 45°.



Fonte: Produzida pelo autor.

1995). Nessa abordagem, o conjunto dos vetores de características é particionado em  $k$  subconjuntos mutuamente exclusivos, sendo uma parte destinada ao treinamento e outra destinada à validação do modelo.

Para medir o desempenho do classificador são utilizadas as medidas de VP, VN,

FP, FN, Sensibilidade, Especificidade, Acurácia, índice de concordância Kappa e área sob a curva ROC.

## 6 RESULTADOS E DISCUSSÃO

Este capítulo apresenta e discute os resultados obtidos em testes realizados para avaliação da metodologia para o diagnóstico DMRI.

Em primeiro lugar, o método proposto para a segmentação da retina é validado. Isso é feito avaliando o erro de segmentação em comparação a duas bases de volumes de OCT com marcações das bordas da MLI, do EPR e da MB. As bases utilizadas contêm marcações feitas de forma automática (revisada por especialistas) e marcações realizadas manualmente por especialistas.

Para validar o método de diagnóstico de DMRI, realizamos experimentos baseados em duas abordagens básicas. Uma abordagem avalia o método de diagnóstico de DMRI utilizando a segmentação de camadas realizada automaticamente pelo software DOCTRAP e revisada por especialistas. Uma outra abordagem utilizada nos experimentos visa avaliar o método automático de diagnóstico da DMRI que combina o algoritmo de segmentação e a metodologia de diagnóstico propostos nesta tese, configurando, assim, um método completamente automático de diagnóstico de DMRI a partir de imagens de OCT.

Primeiramente são apresentadas informações sobre as imagens utilizadas, assim como sobre *software* e *hardware* empregados no processamento.

### 6.1 Base de imagens de OCT

Duas bases de exames de SD-OCT foram utilizadas para a avaliação dos métodos propostos. A primeira base (A) é dividida em dois grupos (DMRI e Controle), já a segunda (B) contém somente *B-scans* de OCT de pacientes acometidos por DMRI. Ambas possuem marcações das bordas da RT e do EPR. As bases A e B foram utilizadas para avaliação do método de segmentação e a Base B foi empregada para a avaliação do diagnóstico de DMRI. A seguir, serão apresentados mais detalhes sobre os conjuntos de imagens utilizados.

#### 6.1.1 Base A

A base de imagens de OCT disponibilizada por Farsiu et al. (2014) é composta por 269 volumes de olhos com DMRI do tipo seca e 115 volumes de controle (olhos considerados normais). Trata-se de um subconjunto do estudo AREDS2 (CHEW et al., 2012) e foi

coletado com o propósito de promover um estudo sobre indicadores quantitativos da presença de DMRI em adultos.

As principais características da população do estudo AREDS2 são:

- idade média de 74 anos;
- 57% são mulheres;
- 97% brancos;
- 7% fumantes;
- 19% com patologia cardiovascular;
- 44% utilizam medicamentos para regulação dos níveis de colesterol;
- 50% utilizam aspirina;
- 59% são acometidos de drusas bilaterais grandes e
- 32% com o DMRI em estágio avançado em pelo menos um olho.

As imagens foram adquiridas em quatro clínicas por tomógrafos SD-OCT fabricados pela empresa Bioptigen, Inc (Research Triangle Park, NC). Os volumes foram adquiridos em uma região retangular de 6,7 mm x 6,7 mm centrada na fóvea com protocolo de *scan* rápido, resultando em volumes de tamanho 1.000 x 512 x 100.

O subconjunto disponibilizado por Farsiu et al. (2014), possui imagens de pacientes entre 50 e 85 anos exibindo DMRI intermediária, com drusas maiores que 125  $\mu m$  nos dois olhos ou em apenas um, quando o outro olho está acometido por DMRI em estágio avançado. Os pacientes arrolados não possuem histórico de cirurgia vitreoretiniana ou outra condição patológica que afete a acuidade visual dos olhos. O grupo de controle utiliza os mesmos critérios de inclusão do estudo AREDS2, contudo, não apresentam nenhuma evidência da presença de drusas.

Os exames SD-OCT fornecidos também disponibilizam a marcação das bordas que limitam a retina total (entre a membrana limitante interna e a membrana de Bruch) assim como delimitam o epitélio pigmentado retiniano associado com o complexo de drusas (quando existente). A segmentação foi inicialmente executada de forma automatizada pelo software DOCTRAP, desenvolvido pela Universidade de Duke, e submetida à revisão e

ajustes por especialistas certificados pelo *Duke Advanced Research in Spectral Domain OCT Imaging Laboratory*.

Após a segmentação, os especialistas marcaram manualmente o ponto central da fóvea e definiram como região de interesse em cada volume, um cilindro centrado nesse ponto com 5 mm de diâmetro. Essa revisão criteriosa, com a correção dos possíveis erros cometidos pelo método automatizado, permitem a criação de uma representação da retina em mapas topográficos com precisão adequada para a validação dessa metodologia.

### 6.1.2 Base B

A base disponibilizada por Chiu et al. (2012) possui 19 volumes oriundos do estudo A2A que pesquisou sobre a DMRI intermediária. As imagens foram adquiridas em quatro clínicas utilizando equipamentos do fabricante Bioptgen, Inc.

Os volumes foram avaliados por especialistas e classificados em quatro grupos com características diversas quanto à patologia e qualidade de aquisição (Tabela 5). Os critérios de classificação foram: (1) centralização da fóvea, (2) baixa resolução e saturação, (3) presença de artefatos produzidos por piscadas do paciente, (4) presença de artefatos produzidos por movimento ou perda de fixação, (5) artefatos produzidos pelo equipamento, (6) cortes ou *B-scans* em branco e (7) situações não-classificáveis nas categorias anteriores.

As bordas internas da MLI e do EPR, assim como a borda externa da MB, estão segmentadas por dois especialistas. Por conter duas marcações realizadas por especialistas, essa base foi utilizada para a validação e comparação com marcações efetuadas pelos especialistas.

Tabela 5 – Caracterização das imagens utilizadas na Base A

	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>
<b>Total de volumes</b>	4	5	5	5
<b>Qualidade</b>	Alta	Baixa	Alta	Baixa
<b>Patologia</b>	Drusas	Drusas	Drusas + AG	Drusas + AG

AG: atrofia geográfica do EPR.

## 6.2 Recursos de *Hardware e Software*

Os métodos desenvolvidos nessa tese foram implementados e executados utilizando a linguagem de programação Python 3.5 (distribuição Anaconda). Algumas bibliotecas es-

pecíficas também foram utilizadas durante a implementação, como a Numpy (OLIPHANT, 2006) para o processamento de matrizes e operações matemáticas, Scikit-image (WALT et al., 2014) e Scipy (JONES; OLIPHANT; PETERSON, 2001) para o processamento de imagens e filtragem de ruídos, Scikit-learn (PEDREGOSA et al., 2011) e libSVM (CHANG; LIN, 2011) para o aprendizado computacional e o *framework* CUDA 9.0 da Nvidia (NVIDIA, 2010) em conjunto com as bibliotecas Keras (GULLI; PAL, 2017) e Tensorflow (ABADI et al., 2016) para a implementação da rede de aprendizado profundo.

Os testes foram executados em uma máquina com processador Intel Core i7 com 16 GB de memória RAM e duas placas de processamento gráfico do tipo Nvidia Geforce GTX 1080Ti.

### 6.3 Resultados da Segmentação da Retina

Nesta seção são apresentados os resultados encontrados para experimentos efetuados com o método de segmentação das camadas da retina apresentado no Capítulo 4. Para os testes efetuados, a base A foi utilizada para etapa de treinamento e testes. A base B, por conter uma quantidade menor de exames e apenas volumes acometidos por DMRI, foi utilizada apenas para testes.

A base A foi dividida em dois grupos. Um grupo composto por 308 volumes de OCT foi destinado ao treinamento e outro com 76 volumes foi reservado para testes. A seleção para a composição dos grupos foi realizada de forma aleatória, porém, respeitando a proporção da distribuição de exames com DMRI e do grupo de controle existentes na base (269:115). Logo, o conjunto de imagens para o treinamento foi composto por 220 volumes com DMRI e 88 do grupo de controle. Já o conjunto de teste foi composto por 50 volumes com DMRI e 26 do grupo de controle.

Para cada volume do conjunto de treinamento foram selecionadas 10 *B-scans* das quais foram retirados fragmentos de  $30 \times 30$  *pixels* centralizados sobre as bordas pesquisadas e também da região de fundo. O total de fragmentos extraídos foi de 1.811.384, dos quais 362.277 foram reservados para a validação da rede e não foram expostos ao treinamento. A validação da rede CapsNet alcançou acurácia de 98,89% indicando uma boa capacidade de identificação das regiões pertencentes às bordas e ao fundo da imagem.

Uma vez que cada volume da base A contém 100 *B-scans* cada, os testes de segmentação para esta base avaliaram 7.600 *B-scans*. Já da base B, que possui 11 *B-scans*

por volume, a quantidade de imagens testadas foi de 209. As imagens processadas têm dimensão de  $1.000 \times 512$  *pixels*.

A Tabela 6 apresenta os resultados intermediários obtidos para a segmentação da borda da MLI, da borda interna do EPR e da borda externa da MB com 1/4 da resolução original. Apresenta-se o erro calculado entre nosso método de segmentação automática e as marcações fornecidas pela base.

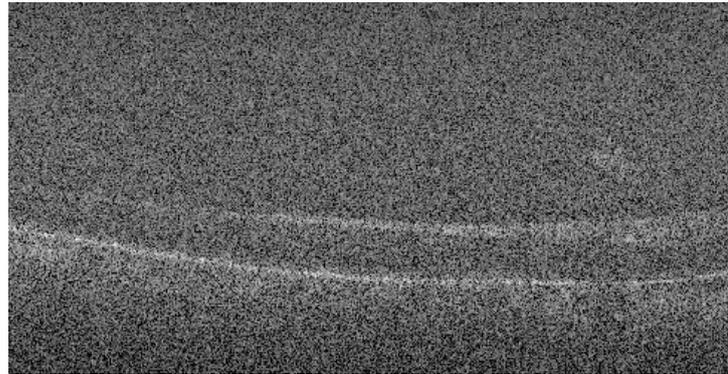
Tabela 6 – Avaliação do erro de segmentação automática da base A em baixa resolução em comparação à segmentação fornecida por Farsiu et al. (2014).

Grupo	Borda	EM			EMA			Hausdorff		
		$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$
DMRI	MLI	1,03	3,44	1,16	1,39	3,33	1,22	6,52	11,47	5,00
	EPR	0,35	0,81	0,29	0,96	0,68	0,86	6,85	5,50	5,00
	MB	-0,49	2,38	0,01	1,47	2,22	0,98	9,57	10,26	5,83
Controle	MLI	1,14	3,56	1,23	1,48	3,44	1,27	6,95	14,48	6,00
	EPR	0,69	0,88	0,61	1,04	0,70	0,92	4,90	2,95	4,47
	MB	0,39	0,85	0,33	0,81	0,73	0,76	4,21	4,01	4,00
Todos	MLI	1,07	3,49	1,19	1,42	3,37	1,24	6,67	12,59	5,00
	EPR	0,47	0,85	0,38	0,99	0,69	0,88	6,18	4,86	5,00
	MB	-0,19	2,03	0,14	1,24	1,87	0,88	7,73	9,00	5,00

Valores em *pixels*. ( $\mu$  - média;  $\sigma$  - desvio padrão;  $M$  - mediana).

Os resultados apresentam baixos valores de EM para todas as bordas. Os valores mais elevados encontrados se referem à borda da MLI. Isso se deve a algumas *B-scans* próximas às extremidades da região de aquisição que apresentam nível de ruído muito elevado. A Figura 32 apresenta uma dessas *B-scans* com presença elevada de ruído que gerou falha na segmentação da borda da MLI. Essa imagem apresentou valores de SNR de 8,37 dB e CNR de 0,295.

Os valores positivos das medianas do EM indicam uma tendência de as bordas segmentadas estarem posicionadas ligeiramente acima das bordas anotadas na base. Os valores de EMA na Tabela 6 apresentam valores mais elevados para as bordas da MLI e da MB. Destaca-se, também, que os valores da medida de distância de Hausdorff da MB são maiores nos volumes acometidos por DMRI do que os do grupo de controle. Isso se deve a falhas na segmentação da MB em imagens com drusas que prejudicam a detecção da MB, problema este enfrentado com a estratégia de processamento em duas resoluções apresentadas na subseção 4.3.5.

Figura 32 – *B-scan* 19 do exame “AMD\_1081” da Base A com baixa qualidade de aquisição.

Fonte: Produzida pelo autor.

A Tabela 7 apresenta o resultado final da segmentação das bordas utilizando a estratégia de segmentação em duas resoluções. O emprego dessa estratégia diminuiu os valores médios de EMA para todas as bordas. Além disso, as distâncias de Hausdorff da borda da MB diminuiram mais acentuadamente nos exames acometidos por DMRI, o que indica que há diminuição das falhas de segmentação dessa borda.

Tabela 7 – Resultados da segmentação automática da base A em alta resolução comparada à segmentação da base.

Grupo	Borda	EM			EMA			Hausdorff		
		$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$
DMRI	MLI	0,37	3,42	0,33	0,80	3,36	0,46	5,23	11,67	3,16
	EPR	0,35	0,73	0,30	0,90	0,67	0,81	6,69	5,60	5,00
	MB	-0,54	2,16	-0,33	1,25	2,08	0,90	7,57	8,72	5,00
Controle	MLI	0,23	3,56	0,29	0,65	3,52	0,40	4,40	14,75	2,24
	EPR	0,66	1,20	0,60	1,01	1,13	0,95	4,63	6,40	4,00
	MB	-0,36	3,14	-0,31	0,78	3,11	0,62	4,23	9,48	3,61
Todos	MLI	0,32	3,47	0,31	0,75	3,42	0,43	4,95	12,82	3,00
	EPR	0,46	0,93	0,37	0,93	0,86	0,85	5,98	5,97	4,12
	MB	-0,48	2,54	-0,32	1,09	2,49	0,77	6,42	9,13	4,12

Valores em *pixels*. ( $\mu$  - média;  $\sigma$  - desvio padrão;  $M$  - mediana).

O tempo de processamento médio para uma *B-scan* é apresentado na Tabela 8. A estratégia de segmentação em duas resoluções permitiu que cada *B-scan* fosse processada em cerca de 65 segundos, um valor cinco vezes menor que o tempo empregado no processamento com a resolução original.

As tabelas 9 e 10 apresentam a comparação da segmentação automática executada pelo nosso método com marcações realizadas manualmente por dois especialistas. Os

Tabela 8 – Tempo de processamento por *B-scan* em segundos para a segmentação das bordas da retina.

Resolução	Tempo médio de Processamento (s)
original	353
1/4	22
refinamento	43

valores médios de EMA são superiores aos valores encontrados quando comparados à segmentação automática revisada por especialistas da Base A.

Tabela 9 – Resultados da segmentação automática da base B vs. marcação manual do Especialista 1.

Grupo	Borda	EM			EMA			Hausdorff		
		$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$
Grupo 1	MLI	-0,14	0,49	-0,14	0,85	0,22	0,84	3,79	1,67	3,0
	EPR	0,42	0,68	0,45	1,54	0,42	1,62	7,05	3,21	7,0
	MB	-1,10	0,95	-0,78	1,51	0,78	1,26	8,64	7,75	4,5
Grupo 2	MLI	-2,70	19,73	0,05	3,61	19,60	0,93	7,35	25,26	4,00
	EPR	0,26	0,70	0,31	1,91	0,54	1,78	8,65	2,97	8,00
	MB	-1,27	1,22	-1,08	1,87	0,94	1,61	9,19	6,74	6,71
Grupo 3	MLI	-1,56	11,54	-0,03	2,47	11,40	0,93	5,83	14,11	3,61
	EPR	0,92	1,04	0,96	2,25	0,58	2,20	10,97	4,40	10,77
	MB	-1,56	7,42	-0,62	3,10	7,07	1,88	11,07	10,09	9,00
Grupo 4	MLI	-0,30	0,44	-0,27	0,99	0,63	0,92	3,84	2,89	3,00
	EPR	0,78	0,83	0,91	2,10	0,52	2,20	10,36	3,98	10,00
	MB	0,15	1,15	0,29	1,62	0,86	1,29	8,80	6,52	6,08
Todos	MLI	-1,23	11,69	-0,14	2,04	11,61	0,92	5,28	14,91	3,61
	EPR	0,60	0,87	0,53	1,97	0,58	1,92	9,37	3,97	9,00
	MB	-0,94	3,95	-0,56	2,05	3,74	1,52	9,47	7,91	7,00

Valores em *pixels*. ( $\mu$  - média;  $\sigma$  - desvio padrão;  $M$  - mediana).

Destacamos que dois grupos de imagens da base B (grupo 2 e 4) possuem baixa qualidade de aquisição. Isso dificulta a segmentação. Além disso, a marcação manual foi efetuada com curvas mais suavizadas que as marcações automáticas, o que sugere que a ferramenta utilizada não possuía uma boa resolução para marcação. Também destacamos que existe certa subjetividade na marcação efetuada por pessoas. Isso pode ser observado pela variação das medianas de EM que indica uma tendência dos especialistas de marcarem as bordas acima ou abaixo da segmentação automática.

Em geral, a comparação com marcações efetuadas por especialistas apresentou valores médios de EMA próximos a 2 *pixels*. Esses valores são próximos aos valores encontrados

Tabela 10 – Resultados da segmentação automática da base B vs. marcação manual do Especialista 2.

Grupo	Borda	EM			EMA			Hausdorff		
		$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$
Grupo 1	MLI	0,43	0,88	0,33	1,17	0,59	1,05	4,70	2,04	4,00
	EPR	1,69	1,22	1,83	2,27	0,76	2,11	8,05	3,28	7,07
	MB	-0,85	1,21	-0,86	1,72	0,76	1,50	9,25	7,60	5,00
Grupo 2	MLI	-2,73	19,67	-0,06	3,80	19,51	1,13	7,96	25,18	4,00
	EPR	0,67	1,06	0,69	2,11	0,53	2,05	8,60	3,13	8,00
	MB	-1,20	1,80	-0,96	2,30	1,20	2,15	9,75	7,03	8,00
Grupo 3	MLI	-1,50	11,50	0,02	2,57	11,35	1,05	6,26	14,18	4,00
	EPR	1,55	1,28	1,49	2,49	0,88	2,36	10,02	3,49	9,85
	MB	-1,47	7,42	-0,49	3,20	7,05	2,02	10,85	9,96	8,00
Grupo 4	MLI	0,03	0,57	0,15	1,19	0,81	1,03	4,57	3,03	4,00
	EPR	1,13	1,13	1,05	2,45	0,77	2,47	10,41	3,28	10,20
	MB	-0,13	1,51	0,11	1,94	0,98	1,57	9,09	6,66	6,32
Todos	MLI	-1,02	11,69	0,07	2,24	11,56	1,09	5,94	14,90	4,00
	EPR	1,24	1,23	1,20	2,33	0,76	2,20	9,33	3,41	9,00
	MB	-0,91	4,04	-0,48	2,32	3,74	1,94	9,76	7,91	7,00

Valores em *pixels*. ( $\mu$  - média;  $\sigma$  - desvio padrão;  $M$  - mediana).

pela comparação entre as marcações efetuadas pelos dois especialistas (Tabela 11).

Percebemos que nossa metodologia apresentou um desvio padrão elevado ( $>19$ ) para os valores de EMA referentes à borda da MLI em imagens do Grupo 2, o que atribuímos à baixa qualidade das imagens, sugerindo a necessidade de melhoria quanto ao pré-processamento dessas *B-scans*. Outro ponto a ser destacado é o resultado obtido para o processamento dos grupos 3 e 4 da base B. Esse grupos contêm exames com a presença de atrofia geográfica do EPR. Isso elevou os valores das distâncias de Hausdorff para as bordas do EPR e da MB, que merecem ser melhorados. Esses são dois problemas na nossa metodologia.

Tabela 11 – Comparação entre marcações efetuadas por dois especialistas.

Borda	EM			EMA			Hausdorff		
	$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$	$\mu$	$\sigma$	$M$
MLI	-0,23	0,76	-0,28	1,30	0,43	1,23	3,95	1,23	4,00
EPR	-0,75	1,34	-0,67	2,34	0,78	2,25	8,78	3,64	8,00
MB	-0,02	1,15	-0,02	1,61	0,59	1,54	4,44	1,57	4,00

Valores em *pixels*. ( $\mu$  - média;  $\sigma$  - desvio padrão;  $M$  - mediana).

A Tabela 12 apresenta uma comparação do nosso método com implementações de

outras redes de aprendizado profundo disponibilizados por Kugelman et al. (2018b) com a mesma base. Os resultados são promissores e demonstram que a CapsuleNet permite realizar a segmentação das bordas com excelente precisão. Nossa metodologia conseguiu apresentar menores valores de EMA para as bordas do EPR e da MB as principais bordas de interesse para o diagnóstico de DMRI. Por fim, apresentamos alguns casos para ilustrar o desempenho do nosso método.

Tabela 12 – Comparação de erro de posicionamento de borda com outros métodos de segmentação baseados em aprendizado profundo.

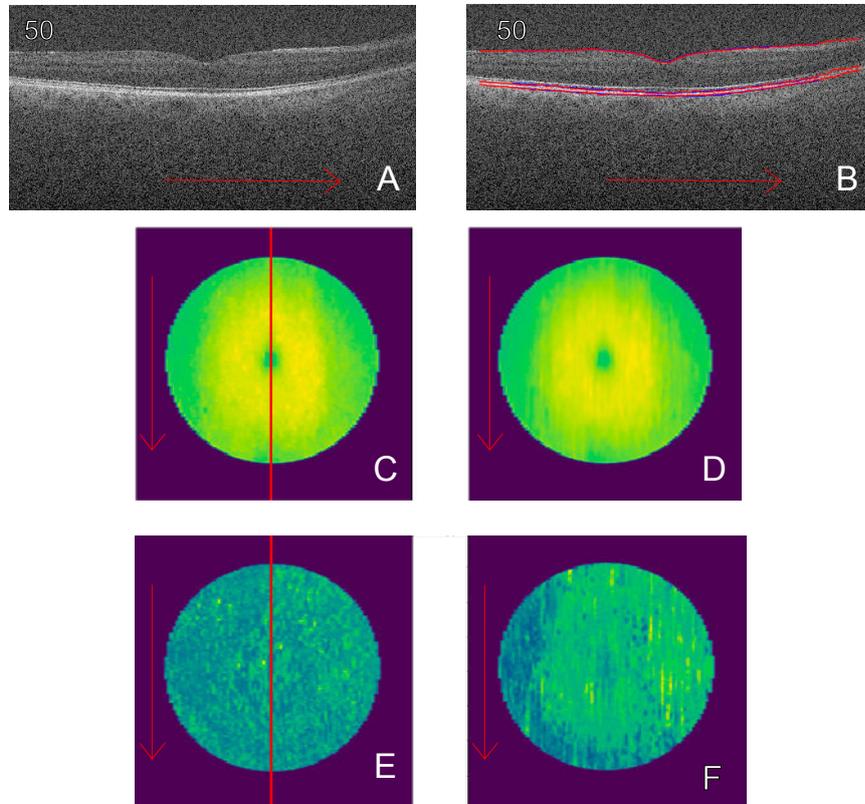
Método	Borda	EM		EMA	
		$\mu$	$\sigma$	$\mu$	$\sigma$
RNN-GS	MLI	-0,03	0,99	0,38	0,92
	EPR	-0,52	3,05	1,05	2,91
	MB	0,91	4,69	2,07	4,31
CNN-GS	MLI	-0,68	7,26	1,10	7,21
	EPR	-0,64	3,30	1,17	3,15
	MB	1,19	5,01	2,31	4,60
FCN-GS	MLI	0,14	4,29	0,65	4,24
	EPR	-0,43	3,12	1,03	2,97
	MB	0,46	3,79	1,53	3,50
Nosso método	MLI	0,32	3,47	0,75	3,42
	EPR	0,46	0,93	0,93	0,86
	MB	-0,48	2,54	1,09	2,49

Valores em *pixels*. ( $\mu$  - média;  $\sigma$  - desvio padrão).

A Figura 33 apresenta um caso saudável (exame Control\_1055) no qual a segmentação das bordas apresentou sucesso. Na figura apresentamos uma comparação visual entre as bordas geradas pelo nosso método e as disponíveis na base (imagem B). Também apresentamos a representação pelos mapas de espessura da RT (imagem C) e do EPR-CD (imagem E). As imagens D e F correspondem a mapas topográficos gerados a partir das marcações anotadas na base. A Figura 34 apresenta um caso de segmentação com sucesso para um exame acometido por DMRI. Podemos perceber que a metodologia consegue segmentar adequadamente uma drusa localizada na parte central da imagem. A representação em mapa topográfico do EPR-CD (imagem E) consegue destacar bem as drusas com menor quantidade de ruído que o mapa gerado a partir das marcações da base (imagem F).

A Figura 35 apresenta um caso (exame AMD\_1034) no qual a segmentação da borda da MB apresenta falhas devido à presença de atrofia geográfica (perdas no tecido

Figura 33 – A) *B-scan* 50 do exame “Controle\_1055 da Base A”. B) Imagem com marcações das bordas da base (azul) e nosso método (vermelho). Mapa topográfico da RT: C) nosso método, D) com marcações da base. Mapa topográfico do EPR: E) nosso método, F) com marcações da base.



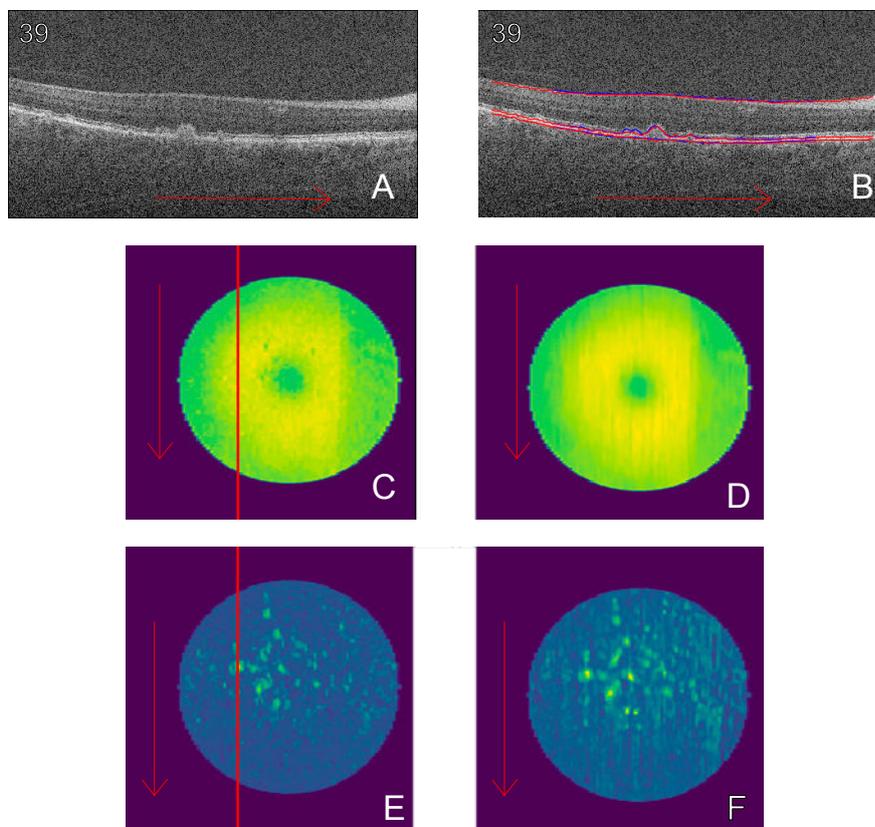
Fonte: Produzida pelo autor.

do EPR). A ausência do EPR em alguns pontos faz com que o sinal luminoso penetre mais profundamente no tecido causando um efeito de sombreamento reverso. Isso pode ser percebido na região de maior intensidade abaixo da MB na imagem A. O nosso método apresentou problemas ao segmentar a MB nessas áreas. As zonas de alta intensidade abaixo do EPR não permitiram a construção de um mapa de probabilidade consistente e, por consequência, a etapa de corte de grafos não pode segmentar corretamente essa região. O EMA para a borda da MB encontrado para este caso foi 1,86 *pixel*; já a distância de Hausdorff média foi 18,48.

#### 6.4 Resultados do diagnóstico de DMRI não-exsudativa

Para avaliar o desempenho do método proposto para o diagnóstico de DMRI não-exsudativa, realizamos experimentos com 383 volumes de OCT fornecidos da base A.

Figura 34 – A) *B-scan* 39 do exame “AMD\_1131 da Base A”. B) Imagem com marcações das bordas da base (azul) e nosso método (vermelho). Mapa topográfico da RT: C) nosso método, D) com marcações da base. Mapa topográfico do EPR: E) nosso método, F) com marcações da base.



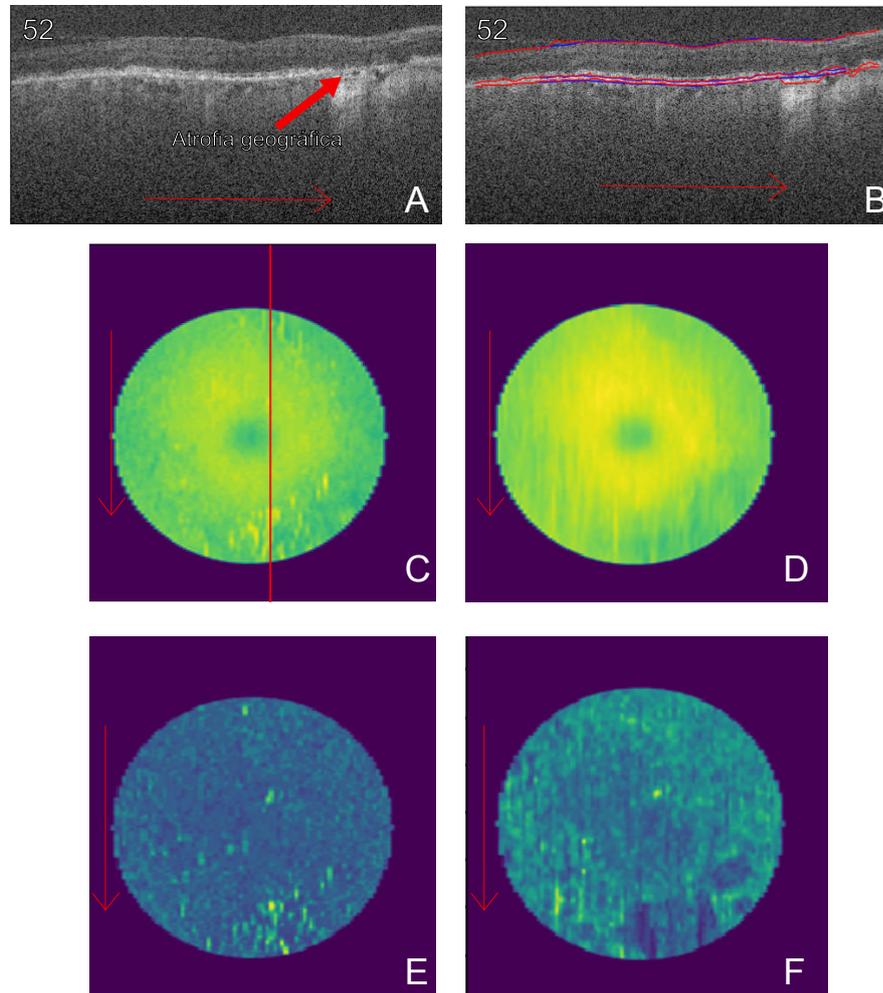
Fonte: Produzida pelo autor.

O volume faltante não foi usado devido possuir um número menor de *B-scans* quando comparado com os demais (80 *B-scans* versus 100 nos outros volumes).

Os experimentos utilizaram os mapas topográficos da RT, da RN e EPR extraídos dos volumes da base A com descritores geoestatísticos semivariograma e semimadograma conforme apresentados no Capítulo 5.

Foram realizados experimentos com diferentes modelos para determinar qual a melhor combinação de mapa topográfico/descritor geoestatístico. Também realizamos experimentos com descritores clássicos de textura para comparação com o melhor modelo gerado pela nossa metodologia e, por fim, realizamos um teste utilizando 76 volumes segmentados por nosso método de segmentação automática de forma a avaliar o método de diagnóstico completamente automático. A seguir detalharemos esses experimentos.

Figura 35 – A) *B-scan* 52 do exame “AMD\_1034” da Base A com destaque para área com atrofia geográfica. B) Imagem com marcações das bordas da base (azul) e nosso método (vermelho). Mapa topográfico da RT: C) nosso método, D) com marcações da base. Mapa topográfico do EPR: E) nosso método, F) com marcações da base.



Fonte: Produzida pelo autor.

#### 6.4.1 Determinação do modelo de classificação

Foram avaliados oito experimentos para construir um modelo robusto de diagnóstico de DMRI não-exsudativa. Cada experimento corresponde a uma combinação mapa topográfico/função geoestatística para avaliação de sua eficácia em distinguir exames afetados pela DMRI daqueles considerados saudáveis, conforme apresentados na Tabela 13.

Para cada experimento, as funções geoestatísticas fornecem valores que mostram dependência espacial das estruturas locais dos mapas topográficos. Nesses experimentos, os mapas são construídos com bases nas marcações já existentes na base, efetuadas pelo *software* DOCTRAP e revisada por especialistas. Na Figura 36, apresentamos a saída

Tabela 13 – Experimentos realizados para o diagnóstico automático de DMRI.

n	Nome	Descrição do cálculo
1	RT (SV)	semivariograma sobre o mapa da RT
2	RT (SM)	semimadograma sobre o mapa da RT
3	RN (SV)	semivariograma sobre o mapa da RN
4	RN (SM)	semimadograma sobre o mapa da RN
5	EPR-CD (SV)	semivariograma sobre o mapa do EPR-CD
6	EPR-CD (SM)	semimadograma sobre o mapa do EPR-CD
7	EPR-CD en face (SV)	semivariograma sobre a reconstrução <i>en face</i> do EPR
8	EPR-CD en face (SM)	semimadograma sobre a reconstrução <i>en face</i> do EPR

dada pelas funções geoestatísticas em dois casos diagnosticados com DMRI e dois casos controle. Esses dados foram então submetidos ao classificador SVM, permitindo, assim a realização de testes e validação.

Os testes foram realizados usando a técnica *k-fold* (KIM, 2009) com repetições. Para validação, dividimos os dados em 5 grupos com 100 repetições, estabelecendo uma proporção de 80% para treino e 20% para testes em cada interação. O uso de repetições diminui a apresentação de viés, dado o fato de que isto diminui a contribuição dos *folds* que apresentam *overfitting* no resultado geral. Calculamos a média e desvio padrão para cada teste. Esses valores são apresentados na Tabela 14 e na Figura 37.

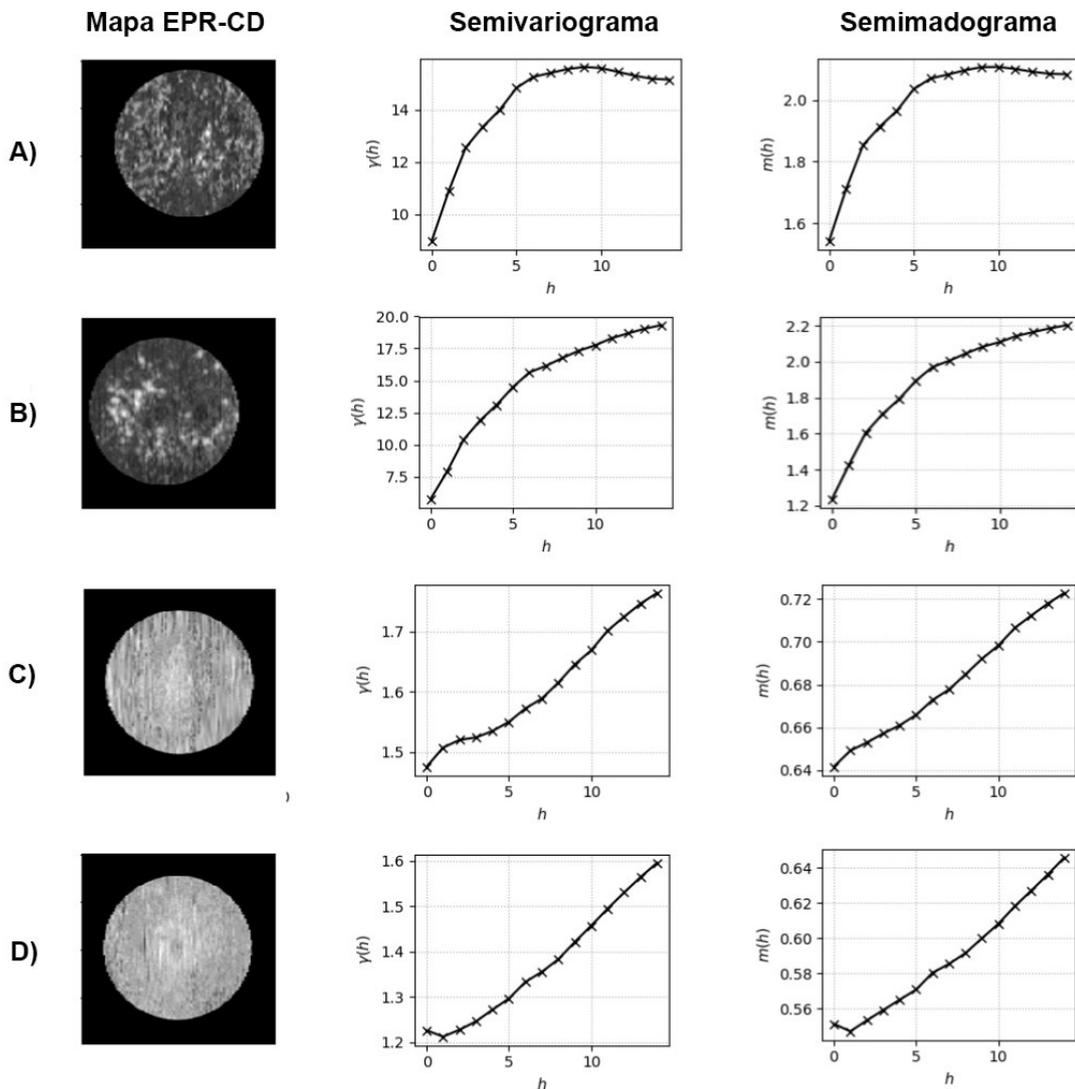
Como pode ser visto na Tabela 14, tanto semivariograma quanto semimadograma, quando calculados sobre o mapa topográfico da RT, apresentam baixa capacidade de diferenciação entre as duas classes. As AUROC médias encontradas para estes experimentos são, respectivamente 0,576 e 0,756.

Quando o mapa topográfico da RN é utilizado, obtivemos melhores resultados em termos de sensibilidade, especificidade e AUROC. Os melhores resultados para este mapa foram obtidos pela função semimadograma que apresentou sensibilidade média de 90,0%, especificidade média de 84,3%, acurácia média de 88,3% e AUROC média de 0,950. O índice Kappa de valor 0,724 indica uma concordância substancial.

Obtivemos o melhor resultado utilizando o mapa topográfico do EPR-CD, especialmente com a função semivariograma. A acurácia média encontrada de 95,2% mostra que este modelo tem excelente capacidade de diferenciação diagnóstica para DMRI, apresentando AUROC média de 0,989.

Já as funções semivariograma e semimadograma aplicadas à reconstrução da borda *en face* do EPR, que considerou a refletância dos pontos na borda superior dessa camada,

Figura 36 – Gráficos dos descritores semivariograma e semimadograma para mapas topográficos do EPR-CD. As linhas A e B apresentam exames de olhos com DMRI não-exsudativa, enquanto C e D apresentam exames considerados normais.



Fonte – Produzida pelo autor.

não demonstraram capacidade expressiva de diferenciar os controles dos exames afetados pela DMRI (AUROC média de 0,505 e 0,778, respectivamente).

Para uma comparação visual, na Figura 36 apresentamos as respostas de semivariograma e semimadograma calculadas para os mapas topográficos. Estas foram calculadas para a direção  $45^\circ$  e um volume com DMRI e um do grupo controle são visualmente comparados.

Através da análise da Tabela 14 percebemos que tanto o semivariograma quanto semimadograma, quando calculados no mapa topográfico da RT, apresentam baixa capaci-

Tabela 14 – Resultados dos experimentos de diagnóstico de DMRI utilizando descritores geoestatísticos.

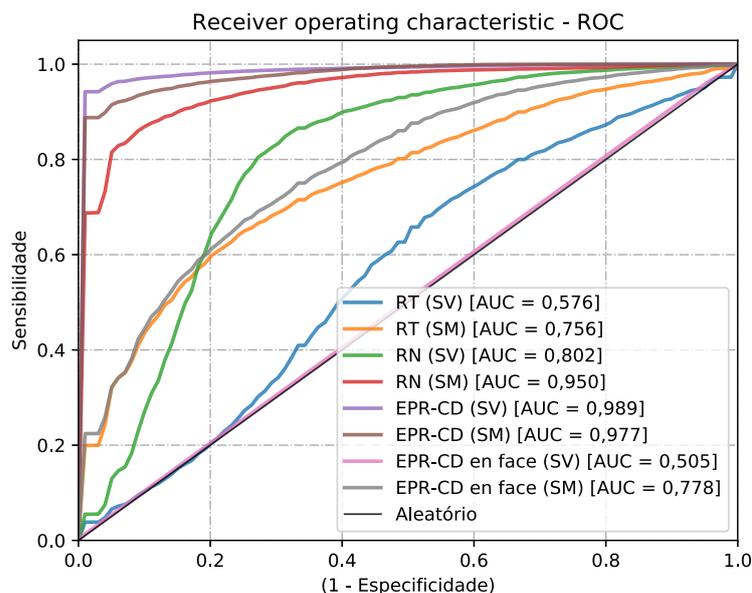
Experimento		Sens.	Esp.	Acu.	AUC	Kappa
RT (SV)	$\mu$	76,0%	42,3%	65,8%	0,576	0,180
	$\sigma$	5,7%	10,2%	4,7%	0,071	0,101
	máx. acu.	84,5%	61,1%	78,9%	0,306	0,439
	máx. kappa	84,5%	61,1%	78,9%	0,306	0,439
RT (SM)	$\mu$	91,9%	26,0%	71,7%	0,756	0,204
	$\sigma$	5,7%	13,9%	5,3%	0,052	0,123
	máx. acu.	96,7%	50,0%	86,8%	0,846	0,541
	máx. kappa	91,2%	63,2%	84,2%	0,821	0,564
RN (SV)	$\mu$	88,8%	65,7%	81,9%	0,802	0,554
	$\sigma$	4,1%	9,6%	3,8%	0,060	0,091
	máx. acu.	96,5%	80,0%	92,2%	0,861	0,791
	máx. kappa	96,5%	80,0%	92,2%	0,861	0,791
RN (SM)	$\mu$	90,0%	84,3%	88,3%	0,950	0,724
	$\sigma$	4,0%	7,0%	3,2%	0,023	0,076
	máx. acu.	98,2%	95,5%	97,4%	0,993	0,936
	máx. kappa	98,0%	96,3%	97,4%	0,996	0,943
<b>EPR-CD (SV)</b>	$\mu$	<b>94,2%</b>	<b>97,5%</b>	<b>95,2%</b>	<b>0,989</b>	<b>0,886</b>
	$\sigma$	3,1%	3,2%	2,3%	0,010	0,054
	máx. acu.	100,0%	100,0%	100,0%	1,000	1,000
	máx. kappa	100,0%	100,0%	100,0%	1,000	1,000
EPR-CD (SM)	$\mu$	90,2%	91,6%	90,5%	0,977	0,780
	$\sigma$	4,5%	7,5%	3,5%	0,014	0,080
	máx. acu.	100,0%	95,7%	98,7%	0,999	0,969
	máx. kappa	98,0%	100,0%	98,7%	0,992	0,971
EPR-CD <i>en face</i> (SV)	$\mu$	100,0%	0,0%	70,2%	0,505	0,000
	$\sigma$	0,0%	0,0%	4,7%	0,008	0,000
	máx. acu.	100,0%	0,0%	84,4%	0,542	0,000
	máx. kappa	100,0%	0,0%	71,4%	0,500	0,000
EPR-CD <i>en face</i> (SM)	$\mu$	90,2%	41,5%	75,5%	0,778	0,347
	$\sigma$	4,7%	10,1%	4,3%	0,051	0,098
	máx. acu.	95,1%	62,5%	88,3%	0,844	0,619
	máx. kappa	95,1%	62,5%	88,3%	0,844	0,619

Os dados em negrito correspondem aos melhores resultados obtidos. ( $\mu$  - média;  $\sigma$  - desvio padrão; máx. acu. - máxima acurácia; máx. kappa. - máximo kappa).

dade de diferenciação para detecção de DMRI. Isso também foi demonstrado em (FARSIU et al., 2014) usando mapas de espessura média. Especificamente em nosso trabalho, este resultado indica que não há diferença significativa em termos de dependência espacial entre classes quando considerada a espessura da RT.

Quando utilizado o mapa topográfico da RN, o índice  $\kappa$  obtido corresponde a 0,724, isso indica uma concordância substancial. Uma diferença importante entre os resultados

Figura 37 – Curva ROC obtida para cada experimento. O melhor valor obtido foi 0,989.



Fonte – Produzida pelo autor.

obtidos pelos mapas da RT e da RN é que o último mapa é calculado a partir da borda do EPR, uma camada que é deformada diretamente pelo acúmulo de drusas.

O melhor resultado foi apresentado pelo experimento “EPR-CD (SV)” que alcançou AUROC de 0,989, um resultado muito próximo daquele alcançado pelo método proposto por Farsiu et al. (2014). O valor de  $kappa = 0,886$  também indica que a performance do modelo possui uma alta concordância com o diagnóstico do especialista.

Por outro lado, o semivariograma aplicado sobre a reconstrução *en face* do EPR, que considerou apenas a refletância dos pontos da borda superior dessa camada, não demonstrou habilidade expressiva em diferenciar os volumes do grupo controle dos exames afetados pela DMRI. Especialmente o semivariograma, que resultou em uma AUROC média de 0,505, que é um resultado muito similar a um classificador aleatório. Uma AUROC levemente melhor (0,778) foi alcançada quando utilizado o semimadograma aplicado a essa reconstrução. Porém, um valor menor de  $kappa$  (0,347) indica que o modelo gerado é tendencioso e tende a apresentar um maior número de classificações divergentes do especialista.

O pior desempenho desse último teste, em contraste com resultados promissores obtidos com o mapa topográfico da EPR-CD, demonstra que a DMRI não causa mudanças significativas na refletância dos pontos na borda do EPR como nos mapas de espessura. Essas variações são melhor observadas quando avaliamos as deformações causadas nessa

camada pelo aparecimento das drusas e, possivelmente, neovascularização coroidal (para casos avançados).

Considerando os resultados obtidos, verificamos que o melhor modelo dentre os demais apresentados nesta tese para o diagnóstico de DMRI não-exsudativa usando mapas topográficos é o definido no experimento 5 (EPR-CD/Semivariograma) que será comparado com outras abordagens e empregado para teste completamente automático de diagnóstico de DMRI.

#### 6.4.2 Comparação com descritores de textura clássicos

Comparamos o melhor modelo gerado “EPR-CD (SV)” com a utilização de descritores clássicos de textura (Haralick)(GONZALEZ; WOODS, 2006; COELHO, 2013) e com os padrões binários locais (LBP)(COELHO, 2013) para avaliar o desempenho do nosso método frente aos descritores comumente utilizados para descrição de textura.

Os resultados desses experimentos (Tabela 15) mostraram um desempenho inferior dos descritores de Haralick e LBP em comparação com o semivariograma aplicado ao mapa topográfico do EPR-CD. Especificamente, os descritores LBP e Haralick apresentaram AUROCs de 0,968 e 0,862, respectivamente. Esses valores estão abaixo dos encontrados no resultado apresentado no experimento 5: AUROC=0,989. A Tabela 15 apresenta a comparação entre esses resultados, e a Figura 38 apresenta uma comparação gráfica dos valores AUROC encontrados.

Tabela 15 – Comparação com outras técnicas utilizando o mesmo conjunto de exames.

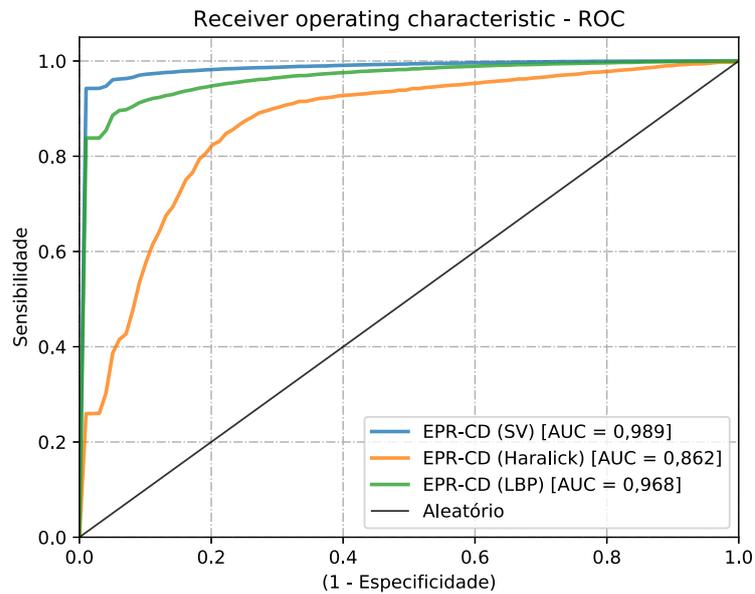
Teste	Sen.	Esp.	Acu.	AUROC	Kappa
Descritores Clássicos de Textura	88,7%	77,4%	85,3%	0,862	0,651
LBP	93,0%	86,1%	90,9%	0,968	0,784
<b>Semivariograma (Este Trabalho)</b>	<b>94,2%</b>	<b>97,5%</b>	<b>95,2%</b>	<b>0,989</b>	<b>0,886</b>

#### 6.4.3 Teste com o método de segmentação proposto

Uma vez definido o modelo mais adequado ao diagnóstico automático de DMRI (EPR-CD/SV), foi realizado um teste com esse modelo utilizando os volumes processados pelo método de segmentação proposto nesta tese.

Nesse teste, os 76 volumes segmentados foram utilizados como amostras de teste e os demais destinados ao treinamento do classificador SVM. Os parâmetros do classificador foram os mesmos dos demais experimentos. O teste apresentou um resultado de acurácia

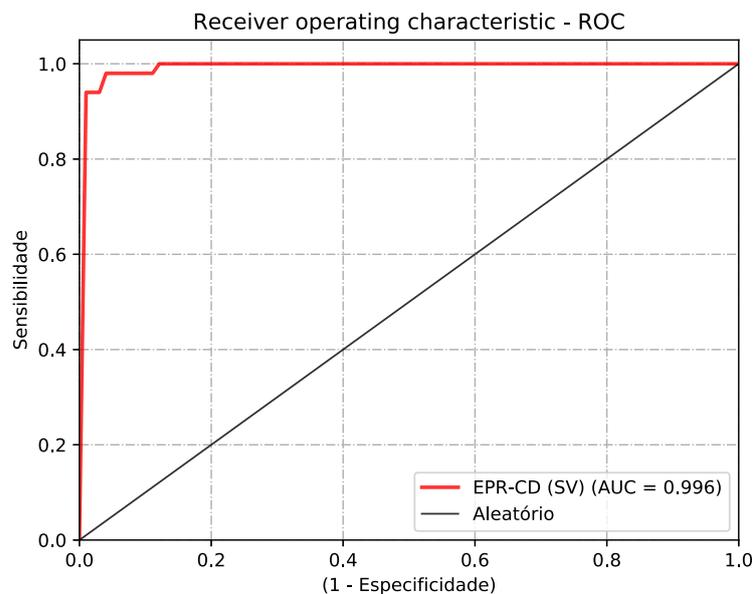
Figura 38 – Comparação da curva ROC do desempenho do semivariograma sobre o mapa EPR-CD em relação a outros métodos. O melhor valor obtido foi 0,989.



Fonte – Produzida pelo autor.

de 96%, sensibilidade de 98%, especificidade de 92,3%, índice de concordância de Kappa de 0,91 e uma AUCROC de 0,996 (Figura 39).

Figura 39 – Curva ROC do desempenho do semivariograma sobre o mapa EPR-CD de volumes segmentados com a metodologia proposta nesta tese.



Fonte – Produzida pelo autor.

#### 6.4.4 Comparação com outros trabalhos

Uma comparação de desempenho com outros trabalhos é apresentada na Tabela 16. Verificamos que o resultado do nosso melhor experimento (EPR-CD/SV), quando utiliza mapas topográficos gerados a partir da base disponibilizada por Farsiu et al. (2014), são superiores aos apresentados em (LIU et al., 2011), (SRINIVASAN et al., 2014b), (ALBARRAK et al., 2013), (KARRI; CHAKRABORTY; CHATTERJEE, 2017), (ZHANG et al., 2014) e (SERRANO-AGUILAR et al., 2012) e desempenho ligeiramente inferior aos apresentados em (FARSIU et al., 2014), (VENHUIZEN et al., 2015), (KERMANY et al., 2018) e (FANG et al., 2017a).

Já quando empregado o nosso método de segmentação em conjunto com o método de diagnóstico, nosso trabalho apresenta resultados superiores aos de (FARSIU et al., 2014), (VENHUIZEN et al., 2015) e (FANG et al., 2017a). A AUCROC de 0,996 é um resultado animador. Destacamos ainda que nosso método não empregou nenhum ajuste de segmentação manual por especialistas. O que indica que o método de segmentação proposto é preciso o suficiente para representar as alterações provocadas pela DMRI não-exsudativa.

Tabela 16 – Comparação de desempenho de diagnóstico de DMRI com os principais trabalhos relacionados.

<b>Trabalho</b>	<b>Amostras</b>	<b>Acur.</b>	<b>Sen.</b>	<b>Esp.</b>	<b>AUROC</b>
Liu et al. (2011)	457	89,3%	-	-	0,975
Serrano et al. (2012)	200	-	96,0%	92,0%	-
Albarrak et al. (2013)	140	91,4%	92,4%	90,5%	0,944
Zhang et al. (2014)	140	92,06%	91,82%	92,3%	-
Farsiu et al. (2014)	384	-	-	-	0,991
Srinivasan et al. (2014)	45	-	100%	-	-
Venhuizen et al. (2015)	384	-	-	-	0,984
Wang et al. (2016)	45	93,3%	-	-	0,995
Sun et al. (2017)	45	100%	-	-	-
	678 <i>B-scans</i>	99,6%	-	-	-
Ravenscroft et al. (2017)	75	83,3%	-	-	-
	45	100%	100%	100%	1,00
Fang et al. (2017)	54	92,2%	96,9%	95,4%	-
Karri et al. (2017)	45	89,0%	-	-	-
Lee et al. (2017)	100.000 <i>B-scans</i>	88,98%	85,41%	93,82%	0,938
Kermany et al. (2018)	207.130 <i>B-scans</i>	99%	98%	99,2%	0,999
<b>Este trabalho (C)</b>	<b>383</b>	<b>95,2%</b>	<b>94,2%</b>	<b>97,5%</b>	<b>0,989</b>
<b>Este trabalho (S + C)</b>	<b>76</b>	<b>96,0%</b>	<b>98,0%</b>	<b>92,3%</b>	<b>0,996</b>

(C) - Classificação/(S + C) - Segmentação e Classificação

Entretanto, não é possível estabelecer uma comparação direta com todos os trabalhos

tendo em vista que utilizam bases de dados e quantidades diferentes de amostras. Alguns reportam valores de acurácia um pouco maiores que nossa metodologia, contudo utilizam bases menores para a validação e se baseiam na análise de *B-scans* individualizadas.

Entre os trabalhos que utilizam o mesmo conjunto de exames OCT (FARSIU et al., 2014; VENHUIZEN et al., 2015), considerando também os testes realizados com os descritores clássicos de Haralick e a técnica LBP, conseguimos evidenciar que a utilização da função semivariograma sobre o mapa topográfico do EPR-CD é um método robusto para a detecção de DMRI não-exsudativa em volumes OCT.

Uma comparação direta entre os valores de AUROC nos permite perceber que os resultados da metodologia proposta é competitiva com os demais trabalhos. O trabalho de Farsiu et al. (2014) é baseado na espessura média das camadas retinianas. Essas medidas podem apresentar variação dentro da mesma classe e não levam em consideração a variação das espessuras locais. Acreditamos que nosso método possui a vantagem de atuar sobre a distribuição das alterações anatômicas causadas pela DMRI.

Os descritores clássicos de Haralick também não fazem uso da distribuição espacial em sua análise. Já o LBP, apesar de considerar a localização espacial dos dados, não foi capaz de promover uma discriminação adequada entre as retinas comprometidas por DMRI não-exsudativa, o que nos leva a considerar a necessidade do uso de algum tipo de pré-processamento com essa técnica.

Os resultados indicam que o método desenvolvido, baseado na caracterização da variação da espessura local de mapas topográficos com funções geoestatísticas, é promissor. Observamos que a função semivariograma consegue caracterizar bem essas variações e nosso método de segmentação é capaz de gerar mapas topográficos adequados para serem empregados no diagnóstico de DMRI. No próximo capítulo apresentaremos as conclusões e considerações finais desta tese.

## 7 CONCLUSÃO

Nesta tese, propusemos uma metodologia para o diagnóstico automático de DMRI não-exsudativa em imagens de OCT. Partimos da desafiadora tarefa de segmentar camadas da retina humana submetidas a deformações anatômicas causadas pela DMRI e efetuamos a classificação entre exames acometidos pela formação de drusas e exames saudáveis. Assim, desenvolvemos um método completamente automático para o diagnóstico de DMRI.

O método proposto para a segmentação das bordas das camadas da retina alcançou bons resultados, com erros médios aproximados aos encontrados no estado da arte. A segmentação das bordas das camadas permitiu que fossem gerados mapas topográficos, com precisão, que pudessem ser utilizados para o diagnóstico automático através da caracterização por descritores geoestatísticos.

Para o diagnóstico de DMRI, empregamos descritores geoestatísticos semivariograma e semimadograma para caracterizar os mapas topográficos da RT, da RN e do EPR-CD da região macular, assim como reconstruções *en face* da superfície do EPR. Os mapas topográficos descrevem como sua espessura varia ao longo da região macular. Desenvolvemos uma maneira de descrever essas variações com base no pressuposto de que a representação da textura desses mapas é uma forte indicação de pacientes saudáveis e não saudáveis.

A partir das representações geradas, determinamos um modelo utilizando SVM para efetuar o diagnóstico de DMRI não-exsudativa. O modelo utiliza o cálculo do semivariograma sobre o mapa topográfico do EPR, uma nova maneira de representar a retina globalmente e, traz como vantagem a análise das variações de espessura nas camadas.

Assim, a principal contribuição desta tese foi o desenvolvimento de um método para diagnóstico de DMRI completamente automático com taxa de acerto (AUROC= 0,996) superior a estado da arte atual. Evidenciou, ademais, que descritores geoestatísticos podem ser adotados com sucesso para a caracterização das alterações provocadas pela DMRI não-exsudativa. O emprego da rede de aprendizado profundo, baseado no roteamento entre cápsulas para a segmentação das bordas das camadas é, também, outra contribuição deste trabalho.

Acreditamos ter alcançado os objetivos propostos para esta tese, apontando métodos para a segmentação da retina e diagnóstico automatizado da DMRI que atingiram bons resultados. A validação do modelo de classificação demonstra que a metodologia proposta

pode realizar o diagnóstico com grande precisão. Entretanto, ressaltamos que ainda precisamos realizar mais testes, com bases de dados maiores de modo a confirmar esses resultados e aprimorar os métodos apresentados. É preciso ainda aprimorar o processo de segmentação para exames com atrofia geográfica do EPR, de forma a garantir que a MB seja segmentada corretamente para esses casos.

Entretanto, a metodologia proposta foi testada apenas em volumes saudáveis ou acometidos por DMRI não-exsudativa. É provável que esse método também consiga diagnosticar casos da manifestação exsudativa, uma vez que a presença de drusas manifestada na forma não-exsudativa é fator predisponente para a manifestação avançada. Entretanto, é necessária a realização de testes com exames com DMRI exsudativa disponíveis e novo treinamento do classificador para a inclusão desse tipo de doença.

Embora tenhamos conseguido atingir bons resultados nesta tese, existem alguns pontos a serem abordados em trabalhos futuros:

- Empregar um tratamento de ruído eficiente para imagens de baixa qualidade de aquisição, especialmente as *B-scans* próximas aos limites da região de aquisição;
- Melhorar o tratamento de falhas de segmentação da MB devido à atenuação do sinal provocada pela presença de drusas. Isso pode ser feito com o aprimoramento da rede de classificação;
- Efetuar a segmentação das drusas e efetuar medidas de tamanho e volume para facilitar o acompanhamento da evolução da doença;
- Realizar testes com outros bancos de imagens e com diversas condições, como a manifestação exsudativa da DMRI e descolamento da retina neurosensorial;
- Identificar adequadamente os casos de atrofia do EPR para evitar problemas de segmentação na sua borda e na MB;
- Expandir a rede de classificação baseada em Cápsulas para segmentar as demais camadas da retina.
- Testar a metodologia com outras modalidades de aquisição de OCT como a SS-OCT; esse método de aquisição apresenta melhor captação da imagem das camadas mais externas da retina, como a MB.

- Efetuar testes com novas abordagens com representações *en face* do EPR, como integração axial de valores de intensidade de *pixels* e até mesmo utilização de imagens capturadas por equipamentos que fazem a reconstrução *en face* automatizadas (i. e. OCT-A<sup>1</sup>).

Acreditamos que essa metodologia é útil para distinguir de forma robusta os casos de DMRI, podendo ser aprimorada para fornecer mais informações ao especialista, apoiando sua decisão clínica, facilitando o rastreamento, o diagnóstico remoto e até mesmo a detecção precoce da doença.

---

<sup>1</sup> do inglês *Optical Coherence Tomography Angiography*, é uma aplicação da SD-OCT que através de aquisições múltiplas, consegue efetuar uma reconstrução *en face* do EPR com aumento de contraste para os vasos sanguíneos.

## REFERÊNCIAS

- ABADI, M. et al. Tensorflow: A system for large-scale machine learning. In: SYMPOSIUM ON OPERATING SYSTEMS DESIGN AND IMPLEMENTATION (OSDI), 12., 2016, Savannah, GA, EUA. *Proceedings...* Berkeley: USENIX, 2016. p. 265–283. ISBN 978-1-931971-33-1.
- AFSHAR, P.; MOHAMMADI, A.; PLATANIOTIS, K. N. Brain tumor type classification via capsule networks. In: IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP), 25., 2018, Atenas, Grécia. *Proceedings...* Atenas: IEEE, 2018. p. 3129–3133.
- ALBARRAK, A. et al. Age-related macular degeneration identification in volumetric optical coherence tomography using decomposition and local feature extraction. In: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE, UNDERSTANDING AND ANALYSIS, 2013, Birmingham, Reino Unido. *Proceedings...* Birmingham: BMVA, 2013. p. 59–64.
- ALI, M.; PARLAPALLI, R. Signal processing overview of optical coherence tomography systems for medical imaging. In: SPRABB9. Dallas: Texas Instruments, 2010.
- ALMEIDA, J. D. S. de. *Metodologia Computacional para Detecção e Diagnóstico Automáticos e Planejamento Cirúrgico do Estrabismo*. Tese (Doutorado) — Programa de Pós-graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão, 2013.
- ALONSO-CANEIRO, D.; READ, S. A.; COLLINS, M. J. Automatic segmentation of choroidal thickness in optical coherence tomography. *Biomedical Optics Express*, OSA, v. 4, n. 12, p. 2795–2812, dez. 2013.
- AMBATI, J. et al. Age-related macular degeneration: Etiology, pathogenesis, and therapeutic strategies. *Survey of Ophthalmology*, v. 48, n. 3, p. 257 – 293, 2003. ISSN 0039-6257.
- ANTONY, B. J. et al. A combined machine-learning and graph-based framework for the segmentation of retinal surfaces in SD-OCT volumes. *Biomedical Optics Express*, OSA, v. 4, n. 12, p. 2712–2728, dez. 2013.
- BELLINI, L. P. Atualização no diagnóstico e tratamento da degeneração macular relacionada à idade highlights on age-related macular degeneration diagnosis and treatment. *Revista da AMRIGS*, v. 52, n. 3, p. 204–208, 2008.
- BRASIL. Ministério da Saúde. Portaria conjunta nº 18, de 02 de julho de 2018. Aprova o Protocolo Clínico e Diretrizes Terapêuticas da Degeneração Macular Relacionada com a Idade (forma neovascular). *Diário Oficial da União* — Seção 1, Poder Executivo, Brasília, DF, n. 221, p. 173, 19 nov. 2018.
- CBO. Conselho Brasileiro de Oftalmologia. Degeneração macular relacionada à idade. *Revista da Associação Médica Brasileira*, scielo, v. 59, p. 106 – 111, 04 2013. ISSN 0104-4230.
- \_\_\_\_\_. DMRI. 2017. Disponível em: <<http://www.cbo.net.br/novo/publico-geral/dmri.php>>. Acesso em: 03 jun. 2017.

- CHAI, T.; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?—arguments against avoiding RMSE in the literature. *Geoscientific model development*, Copernicus GmbH, v. 7, n. 3, p. 1247–1250, 2014.
- CHANG, C.-C.; LIN, C.-J. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, ACM, v. 2, n. 3, p. 27, 2011.
- CHEN, X.; PAN, L. A survey of graph cuts/graph search based medical image segmentation. *IEEE reviews in biomedical engineering*, IEEE, v. 11, p. 112–124, 2018.
- CHEW, E. Y. et al. The age-related eye disease study 2 (AREDS2): Study design and baseline characteristics (AREDS2 report number 1). *Ophthalmology*, v. 119, n. 11, p. 2282 – 2289, 2012. ISSN 0161-6420.
- CHIU, S. J. et al. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. *Investigative Ophthalmology & Visual Science*, v. 53, n. 1, p. 53, 2012.
- \_\_\_\_\_. Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. *Biomedical Optics Express*, OSA, v. 18, n. 18, p. 19413–19428, aug 2010.
- CHIU, S. J.-Y. *Graph Theory and Dynamic Programming Framework for Automated Segmentation of Ophthalmic Imaging Biomarkers*. Tese (Doutorado) — Duke University, 2014.
- COELHO, L. P. Mahotas: Open source software for scriptable computer vision. *Journal of Open Research Software*, Ubiquity, v. 1, 2013. Disponível em: <<http://doi.org/10.5334/jors.ac>>. Acesso em: 01 jan. 2016.
- CONCI, A. et al. A new measure for comparing biomedical regions of interest in segmentation of digital images. *Discrete Applied Mathematics*, v. 197, p. 103 – 113, 2015. ISSN 0166-218X.
- CONCI, A.; KUBRUSLY, C. Distance between sets—a survey. *Advances in Mathematical Sciences and Applications*, v. 20, n. 1, p. 1 – 18, 2017.
- CORMEN, T. H. et al. *Algoritmos: teoria e prática*. 2. ed. Rio de Janeiro: Elsevier, 2002.
- DAUBECHIES, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, v. 36, n. 5, p. 961–1005, set. 1990.
- DEBUC, D. C. *A review of algorithms for segmentation of retinal image data using optical coherence tomography*. INTECH Open Access, 2011. Disponível em: <<http://cdn.intechopen.com/pdfs-wm/15371.pdf>>. Acesso em: 01 jan. 2016.
- DIAS, M. C. *Estimação das respostas do canal real de propagação rádio-móvel nos domínios espacial e temporal. Análise da supressão de ruído por decomposição wavelet como técnica complementar de processamento*. Tese (Doutorado) — PUC-Rio, Rio de Janeiro, 2003.
- DUKER, J.; WAHEED, N.; GOLDMAN, D. *Manual de OCT em Retina*. Rio de Janeiro: Elsevier, 2015. ISBN 9788535283334.

- ETER, N.; SPAIDE, R. F. Comparison of fluorescein angiography and optical coherence tomography for patients with choroidal neovascularization after photodynamic therapy. *Retina*, LWW, v. 25, n. 6, p. 691–696, 2005.
- EVANS, J. B.; SYED, B. A. New hope for dry amd? *Nature Reviews Drug Discovery*, Nature Publishing Group, 2013.
- FANG, L. et al. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomedical Optics Express*, OSA, v. 8, n. 5, p. 2732–2744, 2017.
- \_\_\_\_\_. Automatic classification of retinal three-dimensional optical coherence tomography images using principal component analysis network with composite kernels. *Journal of biomedical optics*, International Society for Optics and Photonics, v. 22, n. 11, p. 116011, 2017.
- FARJOOD, F.; VARGIS, E. Novel devices for studying acute and chronic mechanical stress in retinal pigment epithelial cells. *Lab on a Chip*, Royal Society of Chemistry, v. 18, n. 22, p. 3413–3424, 2018.
- FARSIU, S. et al. Fast detection and segmentation of drusen in retinal optical coherence tomography images. In: SPIE BIOS – OPHTHALMIC TECHNOLOGIES, 18., 2008, San Jose, Califórnia. *Proceedings...* Bellingham, Washington: SPIE, 2008. v. 6844, p. 1–12.
- \_\_\_\_\_. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. *Ophthalmology*, v. 121, n. 1, p. 162 – 172, 2014. ISSN 0161-6420.
- FERNÁNDEZ, D. C.; SALINAS, H. M.; PULIAFITO, C. A. Automated detection of retinal layer structures on optical coherence tomography images. *Biomedical Optics Express*, OSA, v. 13, n. 25, p. 10200–10216, Dec 2005.
- FREITAS, A. Z. d.; AMARAL, M. M.; RAELE, M. P. Optical coherence tomography: development and applications. InTech, 2014.
- GARVIN, M. K. et al. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE transactions on medical imaging*, IEEE, v. 28, n. 9, p. 1436–1447, 2009.
- GÉRON, A. Introducing capsule networks. 2018. Disponível em: <<https://www.oreilly.com/ideas/introducing-capsule-networks>>. Acesso em: 02 jan. 2019.
- GONG, G.; ZHANG, H.; YAO, M. Speckle noise reduction algorithm with total variation regularization in optical coherence tomography. *Biomedical Optics Express*, OSA, v. 23, n. 19, p. 24699–24712, 2015.
- GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 3. ed. Upper Saddle River: Prentice-Hall, 2006.
- GULLI, A.; PAL, S. *Deep Learning with Keras*. Birmingham: Packt, 2017.
- HARALICK, R. M.; SHANMUGAM, K. et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, IEEE, n. 6, p. 610–621, 1973.

- HAYKIN, S. *Neural Networks: A comprehensive foundation*. 3. ed. Upper Saddle River: Prentice-Hall, 2007.
- HEE, M. R. et al. Optical coherence tomography of age-related macular degeneration and choroidal neovascularization. *Ophthalmology*, Elsevier, v. 103, n. 8, p. 1260–1270, 1996.
- \_\_\_\_\_. Optical coherence tomography for micron-resolution ophthalmic imaging. *IEEE Engineering in Medicine and Biology*, IEEE, v. 14, p. 67–76, 1995.
- \_\_\_\_\_. Optical coherence tomography of the human retina. *Archives of Ophthalmology*, v. 113, n. 3, p. 325–332, 1995.
- HUANG, D. et al. Optical coherence tomography. *Science*, NIH Public Access, v. 254, n. 5035, p. 1178, 1991.
- HUANG, Y. et al. Relation of optical coherence tomography to microanatomy in normal and rd chickens. *Investigative ophthalmology & visual science*, ARVO, v. 39, n. 12, p. 2405–2416, 1998.
- IESMANTAS, T.; ALZBUTAS, R. Convolutional capsule network for classification of breast cancer histology images. In: CAMPILHO, A.; KARRAY, F.; ROMENY, B. ter H. (Ed.). *Image Analysis and Recognition*. Cham: Springer, 2018. p. 853–860.
- ISAAKS, E. H.; SRIVASTAVA, R. M. *Applied Geostatistics*. New York: Oxford University, 1989. ISBN 978-0-19-505013-4.
- SOCIETY FOR EXPERIMENTAL MECHANICS. *Advancement of Optical Methods in Experimental Mechanics*, v. 3. [S.l.]: Springer, 2013. ISBN 9783319007687.
- JONES, E.; OLIPHANT, T.; PETERSON, P. *SciPy: Open source scientific tools for Python*. 2001. Disponível em: <<http://www.scipy.org>>. Acesso em: 01 jan. 2016.
- KANSKI, J. J. *Oftalmologia Clínica*. Rio de Janeiro: Elsevier, 2012.
- KARRI, S. P. K.; CHAKRABORTY, D.; CHATTERJEE, J. Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. *Biomedical Optics Express*, OSA, Washington, v. 8, n. 2, p. 579–592, 2017.
- KERMANY, D. S. et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, Elsevier, v. 172, n. 5, p. 1122–1131, 2018.
- KIM, J.-H. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational statistics & data analysis*, Elsevier, v. 53, n. 11, p. 3735–3745, 2009.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS (ICLR), 3., 2015, San Diego, CA. *Proceedings...* 2015. Disponível em: <<http://arxiv.org/abs/1412.6980>>. Acesso em: 01 fev. 2018.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th international joint conference on Artificial intelligence – IJCAI'95*. San Francisco: Morgan Kaufmann, 1995. v. 2, p. 1137–1145.

- KRAUS, M. F. et al. Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns. *Biomedical Optics Express*, OSA, v. 3, n. 6, p. 1182–1199, 2012.
- KUGELMAN, J. et al. Automatic segmentation of oct retinal boundaries using recurrent neural networks and graph search. *Biomedical Optics Express*, OSA, v. 9, n. 11, p. 5759–5777, 2018.
- \_\_\_\_\_. Automatic segmentation of oct retinal boundaries using recurrent neural networks and graph search. *Biomedical Optics Express*, OSA, v. 9, n. 11, p. 5759–5777, nov. 2018.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, JSTOR, p. 159–174, 1977.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 2015.
- LEE, C. S.; BAUGHMAN, D. M.; LEE, A. Y. Deep learning is effective for classifying normal versus age-related macular degeneration oct images. *Ophthalmology Retina*, Elsevier, v. 1, n. 4, p. 322–327, 2017.
- LEEUWEN, R. van et al. Epidemiology of age-related maculopathy: a review. *European journal of epidemiology*, Springer, v. 18, n. 9, p. 845–854, 2003.
- LIU, Y.-Y. et al. Computerized macular pathology diagnosis in spectral domain optical coherence tomography scans based on multiscale texture and shape features. *Investigative ophthalmology & visual science*, ARVO, v. 52, n. 11, p. 8316–8322, 2011.
- LOWE, D. G. Object recognition from local scale-invariant features. In: *The proceedings of the seventh IEEE international conference on Computer vision*. Kerkyra: IEEE, 1999. v. 2, p. 1150–1157.
- MAWAS, J. Région maculaire couleur. 19–. Disponível em: <<http://collection-jmawas.for.paris>>. Acesso em: 09-mar-2019.
- MCHUGH, M. L. Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, Medicinska naklada, v. 22, n. 3, p. 276–282, 2012.
- MEHTA, S. Age-related macular degeneration. *Primary Care: Clinics in Office Practice*, Elsevier, v. 42, n. 3, p. 377–391, 2015.
- MOBINY, A.; NGUYEN, H. V. Fast CapsNet for lung cancer screening. In: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER ASSISTED & INTERVENTION – MICCAI, 21., 2018, Granada, Espanha. *Proceedings...* Cham, Switzerland: Springer, 2018. p. 741–749.
- NEHEMY, M. B. Degeneração macular relacionada à idade: novas perspectivas. *Arquivos Brasileiros de Oftalmologia*, v. 69, n. 6, p. 955–958, 2006.
- NVIDIA. *NVIDIA CUDA C Programming guide*. Santa Clara: NVIDIA, 2010.
- OLIPHANT, T. E. *A guide to NumPy*. 2. ed. Scotts Valley: CreateSpace, 2006.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, v. 12, n. Oct, p. 2825–2830, 2011.

PRATI, R.; BATISTA, G.; MONARD, M. Curvas roc para avaliação de classificadores. *Revista IEEE América Latina*, v. 6, n. 2, p. 215–222, 2008.

RAVENSCROFT, D. et al. AMD classification in choroidal OCT using hierarchical texton mining. In: INTERNATIONAL CONFERENCE ON ADVANCED CONCEPTS FOR INTELLIGENT VISION SYSTEMS, 18., 2017, Antwerp, Belgium. *Proceedings...* Cham: Springer. p. 237–248.

\_\_\_\_\_. Learning feature extractors for AMD classification in OCT using convolutional neural networks. In: EUROPEAN SIGNAL PROCESSING CONFERENCE (EUSIPCO), 25., 2017, Cos, Grécia. *Proceedings...* Cos: IEEE, 2017. p. 51–55.

ROCHA, S. V. d. *Diferenciação do padrão de malignidade e benignidade de massas em imagens de mamografias usando padrões locais binários, geoestatística e índice de diversidade*. Tese (Doutorado) — Programa de Pós-graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão, 2014.

ROY, A. G. et al. Relaynet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomedical Optics Express*, OSA, v. 8, n. 8, p. 3627–3642, 2017.

RUDIN, L. I.; OSHER, S.; FATEMI, E. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, v. 60, n. 1, p. 259 – 268, 1992.

SABOUR, S.; FROSST, N.; HINTON, G. E. Dynamic routing between capsules. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017. *Proceedings...* New York: Curran, 2017. p. 3856–3866.

SALOMON, D. *A Guide to Data Compression Methods*. New York: Springer, 2002. v. 1. ISBN 9780387952604.

SCHOWENGERDT, R. *Remote Sensing: Models and Methods for Image Processing*. Burlington, MA, USA: Elsevier Science, 2006. ISBN 9780080480589.

SERRANO-AGUILAR, P. et al. Development and validation of a computer-aided diagnostic tool to screen for age-related macular degeneration by optical coherence tomography. *British Journal of Ophthalmology*, BMJ, v. 96, n. 4, p. 503–507, 2012.

SILVA, A. C.; CARVALHO, P. C. P.; GATTASS, M. Diagnosis of lung nodule using semivariogram and geometric measures in computerized tomography images. *Computer methods and programs in biomedicine*, Elsevier, v. 79, n. 1, p. 31–38, 2005.

SOON-CHUN, P. *Handbook Of Vitreo-retinal Disorder Management: A Practical Reference Guide*. Singapura: World Scientific, 2015. ISBN 9789814663311.

SRINIVASAN, P. P. et al. Automatic segmentation of up to ten layer boundaries in sd-oct images of the mouse retina with and without missing layers due to pathology. *Biomedical Optics Express*, OSA, v. 5, n. 2, p. 348–365, fev. 2014.

\_\_\_\_\_. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomedical Optics Express*, OSA, v. 5, n. 10, p. 3568–3577, 2014.

STAFF, B. Blausen gallery 2014. *Wikiversity Journal of Medicine*, v. 1, 2014.

- SUN, Y.; LI, S.; SUN, Z. Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning. *Journal of biomedical optics*, International Society for Optics and Photonics, v. 22, n. 1, p. 016012, 2017.
- VAPNIK, V. *The nature of statistical learning theory*. 2. ed. New York: Springer, 2013.
- VENHUIZEN, F. G. et al. Automated age-related macular degeneration classification in OCT using unsupervised feature learning. In: MEDICAL IMAGING 2015: COMPUTER-AIDED DIAGNOSIS, Orlando, FL, Estados Unidos. *Proceedings of SPIE Vol. 9414*. Washington, DC, Estados Unidos: International Society for Optics and Photonics, 2015. v. 9414, p. 94141I.
- WALT, S. Van der et al. Scikit-image: image processing in python. *PeerJ*, PeerJ, v. 2, p. e453, 2014.
- WANG, L. *Support Vector Machines: Theory and Applications (Studies in Fuzziness and Soft Computing)*. Berlin, Heidelberg: Springer-Verlag, 2005. ISBN 3540243887.
- WANG, Y. et al. Machine learning based detection of age-related macular degeneration (AMD) and diabetic macular edema (DME) from optical coherence tomography (OCT) images. *Biomedical Optics Express*, v. 7, n. 12, p. 4928–4940, dez. 2016.
- WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, Luhe, v. 30, n. 1, p. 79–82, 2005.
- WONG, W. L. et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *The Lancet Global Health*, Londres, Reino Unido, v. 2, n. 2, p. e106 – e116, 2014. ISSN 2214-109X.
- XU, C.; PRINCE, J. L. Gradient vector flow: A new external force for snakes. In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 97., 1997, San Juan, Puerto Rico. *Proceedings...* Los Alamitos, CA, Estados Unidos: IEEE, 1997. p. 66–71.
- XU, Y. et al. Dual-stage deep learning framework for pigment epithelium detachment segmentation in polypoidal choroidal vasculopathy. *Biomedical Optics Express*, v. 8, n. 9, p. 4061–4076, 2017.
- YANOFF, M.; DUKER, J.; AUGSBURGER, J. *Ophthalmology*. 4. ed. Philadelphia: Saunders Elsevier, 2014.
- ZEISS. Zeiss Clarus 500. 2017. Disponível em: <<https://www.zeiss.com/meditec/us/products/ophthalmology-optometry/retina/diagnostics/fundus-imaging/clarus-500.html>>. Acesso em: 03-mar-2019.
- ZHANG, Y. et al. One-class kernel subspace ensemble for medical image classification. *EURASIP Journal on Advances in Signal Processing*, Springer, v. 2014, n. 1, p. 17, 2014.

## APÊNDICE A – PARÂMETRO $\sigma$ PARA PRÉ-PROCESSAMENTO

Nesta tese utilizamos a filtragem baseada em *wavelet* (ver seção 3.5) para tentar suprimir ruídos existentes nas imagens de OCT.

Com o objetivo de definir qual o parâmetro  $\sigma$  a ser utilizado, efetuamos testes utilizando o valor de  $\sigma$  no intervalo  $[0, 8; 0, 15]$  sobre o banco de imagens para mensurar qual o impacto da suavização do ruído de forma que não comprometesse o conteúdo da imagem. As métricas definidas na subseção 3.5.4, SNR e CNR foram utilizadas para a verificação da eliminação do ruído, enquanto o PSNR para a comparação entre a imagem original e a imagem filtrada.

Para estimar os valores da média e variância do fundo e do primeiro plano das imagens, utilizamos marcações existentes na base que delimitam a retina total. Um exemplo destas marcações pode ser visualizado na Figura 40. A imagem à direita apresenta uma *B-scan* sem pré-processamento. À esquerda, a máscara, definindo quais regiões são consideradas como fundo (azul) e como primeiro plano da imagem (verde). A região marcada em vermelho está fora da marcação da base e, portanto, foi desconsiderada.

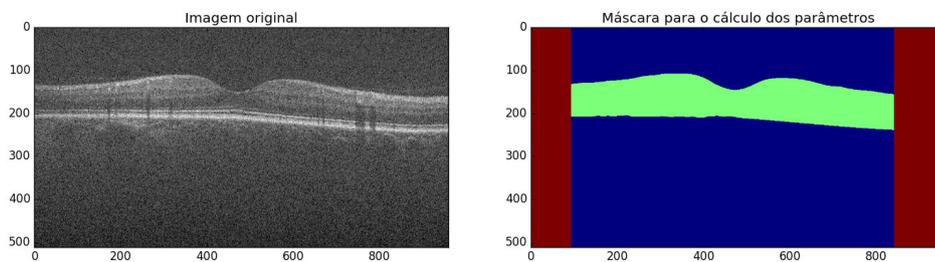


Figura 40 – Máscara obtida a partir das marcações existentes na base.

Os gráficos apresentados na Figura 41 mostram os valores encontrados para o SNR, CNR e PSNR para a filtragem *wavelet* com parâmetro  $\sigma \in [0, 8; 0, 15]$ . Para a seleção do valor a ser adotado, verificamos o que apresenta maior SNR e CNR e cujo PSNR esteja mais próximo do valor típico  $20dB$  (SALOMON, 2002). O valor encontrado foi  $\sigma = 0, 11$ .

Na Tabela 17, comparamos as médias dos valores de SNR, CNR e PSNR encontrados antes e depois da filtragem para as imagens do banco. Nota-se que há um ganho na relação sinal-ruído assim como na relação contraste-ruído. Com o PSNR permanecendo próximo do valor típico apresentado em (SALOMON, 2002).

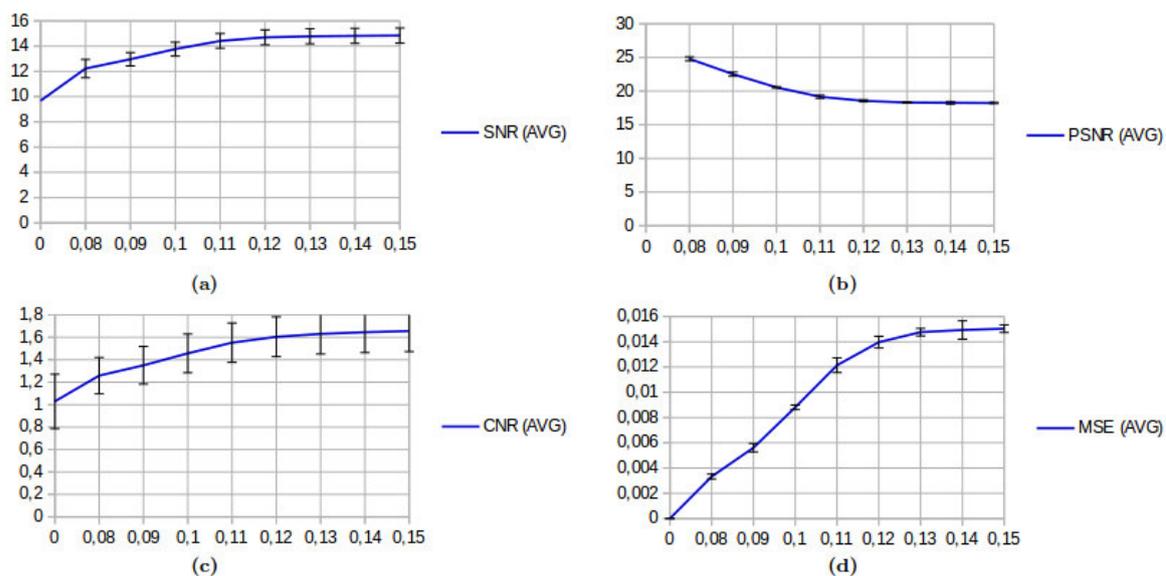


Figura 41 – Valores encontrados de SNR, CNR e PSNR para a filtragem *wavelet* com parâmetro  $\sigma \in [0, 8; 0, 15]$ .

Tabela 17 – Comparação das médias de SNR, CNR e PSNR antes e após a filtragem por *wavelets*. O desvio padrão é apresentado entre parênteses.

Pré-processamento	SNR(dB)	CNR	PSNR
Originais	9,678 (0,723)	1,029 (0,243)	$\infty$
Wavelet	14,413 (0,598)	1,552 (0,175)	19,164 (0,249)