



Pós-Graduação em Ciência da Computação

Wendeson da Silva Oliveira

Consenso de segmentações de imagens usando classificação de padrões



Universidade federal de pernambuco

posgraduacao@cin.ufpe.br

[<www.cin.ufpe.br/~posgraduacao>](http://www.cin.ufpe.br/~posgraduacao)

Recife

2017

Wendeson da Silva Oliveira

**Consenso de segmentações de imagens usando
classificação de padrões**

Trabalho apresentado ao Programa de Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Ciência da Computação.

Orientador: Prof. Dr. Tsang Ing Ren
Coorientador: Prof. Dr. George Darmiton da Cunha Cavalcanti

**Recife
2017**

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

O48c Oliveira, Wendeson da Silva
 Consenso de segmentações de imagens usando classificação de padrões /
 Wendeson da Silva Oliveira. – 2016.
 131 f.: il., fig., tab.

 Orientador: Tsang Ing Ren.
 Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da
 Computação, Recife, 2015.
 Inclui referências e apêndices.

 1. Inteligência artificial. 2. Segmentação de imagens. 3. Avaliação de
 desempenho. I. Ren, Tsang Ing (orientador). II. Título.

 006.3 CDD (23. ed.) UFPE- MEI 2017-152

Wendeson da Silva Oliveira

Consenso de Segmentações de Imagens Usando Classificação de Padrões

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

Aprovado em: 22/02/2016.

Orientador: Prof. Dr. Tsang Ing Ren

BANCA EXAMINADORA

Prof. Dr. Cleber Zanchettin
Centro de Informática / UFPE

Prof. Dr. Hansenclever de França Bassani
Centro de Informática / UFPE

Prof. Dr. Abraão David Costa do Nascimento
Departamento de Estatística / UFPE

Prof. Dr. Juliano Bandeira Lima
Departamento de Eletrônica e Sistemas /UFPE

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

Este trabalho é dedicado a minha filha, Maria Alice.

AGRADECIMENTOS

Agradeço, como faço todo dia, à minha família, pela simples existência e presença em minha vida. Nenhuma palavra é capaz de expressar toda minha gratidão aos meus pais e aos meus irmãos, que sempre estiveram dispostos a me incentivar em qualquer atividade. Em especial a minha esposa, Joyce, pelo companheirismo e compreensão nas horas difíceis e a minha filha, Maria Alice, que me trouxe um novo significado do que é a vida.

Ao professor Tsang Ing Ren pela orientação, amizade e principalmente, pela paciência, sem a qual este trabalho não se realizaria.

Agradeço também ao professor George Darmiton e ao pesquisador Alexandre Cunha do *California Institute of Technology* pela imensa colaboração neste trabalho de tese.

Aos demais professores que ministraram disciplinas longo desses anos de doutorado pelos seus ensinamentos e aos funcionários do CIn que contribuíram de algum modo para o meu enriquecimento pessoal e profissional.

Por fim, mas não menos importante, a meus amigos e colegas, em especial, ao Josivan Reis, Bruno Almeida, Francisco Carlos, Roberto Nascimento, Pedro Lopes e Rafael Barreto.

*“O estudo, a busca da verdade e da beleza são domínios
em que nos é consentido sermos crianças por toda a vida.”*

Albert Einstein

RESUMO

O processamento e a análise de imagens caracterizam-se em 5 fases: aquisição, pré-processamento, segmentação, representação/descrição, e reconhecimento de imagens. Este trabalho situa-se na área de segmentação de imagens médicas 2D. O principal objetivo da segmentação consiste em dividir uma imagem em regiões ou objetos que a compõem. A segmentação de imagens é alvo de vários estudos e pesquisas para o desenvolvimento de métodos mais abrangentes e computacionalmente mais eficientes, principalmente em imagens médicas. Uma etapa crucial para a extração de informações automatizadas em imagens médicas, geralmente envolve a segmentação das regiões da imagem a fim de quantificar volumes e áreas de interesse de tecidos biológicos para posterior diagnóstico e localização de patologias. Um desafio importante que deve ser levado em consideração no desenvolvimento de algoritmos de segmentação de imagens médicas é como avaliar o desempenho do método. Este tipo de avaliação de desempenho geralmente é feito através da comparação dos resultados de algoritmos automáticos versus uma segmentação de referência, denominada de *ground-truth*. Em particular, a estimativa de *ground-truth* em imagens médicas pode ser obtida através do contorno manual ou ainda através de métodos interativos de segmentação. No entanto, essas abordagens podem ser extremamente trabalhosas e sujeitas a variabilidades inter e intra usuários. Uma alternativa para diminuir essa variabilidade é a segmentação utilizando o consenso da resposta de vários usuários para uma mesma imagem. Um método de consenso de segmentações tem a finalidade de combinar várias segmentações de uma mesma imagem. Este trabalho tem como principal objetivo propor uma abordagem para estimar o *ground-truth* em imagens médicas. A abordagem consiste nas duas principais contribuições deste trabalho. Na primeira, um novo sistema colaborativo COSE (*Collaborative Segmentation*) foi desenvolvido para obter várias segmentações de uma mesma imagem com o intuito de estimar o *ground-truth*. Já a segunda, propõe um novo método para o consenso de segmentações de imagens médicas, denominado ISECO (*Intelligent SEgmentation COnsensus*). Os principais métodos de consenso de segmentações utilizam apenas informações binárias. No entanto, o ISECO apresenta como principal diferencial o fato de utilizar informações extraídas da imagem em nível de cinza para auxiliar na classificação dos pixels em *foreground* ou *background* utilizando um modelo de aprendizagem. A extração de características é baseada em três extratores de textura: filtros de Gabor, *Local Binary Pattern* (LBP) e *Local Directional Pattern* (LDP). Outra abordagem para o consenso de segmentações, denominada *Distance Transform Merge* (DTM) também é proposta neste trabalho tendo como objetivo ser um método simples e rápido. Os métodos de consenso propostos são avaliados e comparados com outras abordagens disponíveis na literatura utilizando 4 bases de imagens. A avaliação das abordagens mostra que o ISECO apresenta resultado tão bom ou melhor que os principais métodos da literatura.

Palavras-chave: Segmentação de imagens médicas. Consenso de segmentações. Segmentação colaborativa. *Ground-truth*. Avaliação de desempenho.

ABSTRACT

Image processing is characterized by five steps: acquisition, pre-processing, segmentation, representation/description and image recognition. In this thesis, we work with the problem the 2D medical image segmentation. The main purpose of image segmentation is to divide an image into regions. Image segmentation has been a subject of several studies and research for the development of more comprehensive and computationally more efficient methods, especially in medical imaging. A challenge that must be taken into account in the development of medical image segmentation algorithms is how to evaluate the performance of the method. This type of performance evaluation is usually done by comparing the results of automatic algorithms versus a reference segmentation, called the ground-truth. In particular, the estimation of ground-truth in medical images can be obtained through manual segmentation or interactive segmentation methods. However, these approaches can be extremely subjective and subject to both intra and inter user variability. An alternative to reduce this variability is image segmentation using the consensus of multiple segmentations of a single image. A segmentation consensus method has the purpose of combining multiple segmentations of the same image. This work proposes an approach to estimate the ground-truth in medical images. The approach consists of two parts. At first, we develop a collaborative system called COSE (COllaborative SEgmentation) to obtain various segmentations of the same image in order to estimate the ground-truth. The second proposal is a new method for the consensus of medical image segmentations, called ISECO (Intelligent SEgmentation CCnsensus). The main consensus segmentation methods in the state-of-the-art uses only binary information. However, the ISECO proposes to use gray-scale image information to classify pixels in foreground or background. However, ISECO has as main differential the use of information extracted from the image in gray-scale to assist in the classification of pixels in foreground or background using a learning model. The features extraction is based on three texture extractors: Gabor filters, Local Binary Pattern (LBP) and Local Directional Pattern (LDP). Another approach to segmentation consensus, called Distance Transform Merge (DTM), is also proposed in this work, aiming to be a simple and fast method. The proposed consensus methods are evaluated and compared with other approaches available in the literature using 4 image database. Evaluation of the approaches shows that the ISECO presents better results than the main methods of the literature.

Keywords: Image segmentation. Consensus segmentation. Collaborative segmentation. Ground-truth. Performance evaluation.

LISTA DE ILUSTRAÇÕES

Figura 1 – Relacionamento entre o COSE (<i>COllaborative SEgmentation</i>) e o ISECO (<i>Intelligent SEgmentation COnsensus</i>).	22
Figura 2 – Exemplo de uma imagem obtida por ressonância magnética utilizada na segmentação do ventrículo esquerdo. Na imagem os destaques são o ventrículo esquerdo e o direito.	26
Figura 3 – Imagem do cérebro e seus tecidos obtida por ressonância magnética. . .	27
Figura 4 – Ilustração de uma imagem dos pulmões.	27
Figura 5 – Exemplo de uma imagem de ressonância magnética da próstata usada no problema de segmentação.	28
Figura 6 – Exemplo de uma imagem de raio-X dos ossos do joelho.	28
Figura 7 – Exemplo de uma imagem da retina com vasos sanguíneos.	29
Figura 8 – Exemplos de imagens médicas obtidas por diferentes modalidades de captura: (a) raio-X das mãos de uma pessoa; (b) ultra-sonografia de um feto de 12 semanas; (c) imagem de ressonância magnética do cérebro; (d) tomografia computadorizada do abdômen de uma pessoa adulta; (e) tomografia por emissão de pósitrons de um paciente com câncer no pulmão.	31
Figura 9 – Exemplos de máscaras utilizadas pelo operador laplaciano.	35
Figura 10 – Gráfico para a curva ROC.	41
Figura 11 – Composição da medida Variação da Informação (VINFO), representada pela soma das áreas em cinza. Na Figura, H e I correspondem a entropia e a informação mútua relacionadas a segmentação $\hat{S}$ e o <i>ground-truth</i> S	47
Figura 12 – Sequência de imagens usadas no experimento com as medidas baseadas em distância variando a proporção dos pixels do <i>foreground</i> . A primeira e a última imagens contêm aproximadamente 25% e 0,88% de pixels do <i>foreground</i> , respectivamente.	48
Figura 13 – Valores das medidas baseadas na similaridade de acordo com a proporção dos pixels do <i>foreground</i> . As linhas das medidas Dice e CCI aparecem sobrepostas.	49
Figura 14 – Três imagens de exemplo para o experimento com as medidas baseadas em distância: (a) <i>ground-truth</i> contendo um círculo; (b) imagem contendo um círculo com ruídos; (c) imagem com um quadrado.	50
Figura 15 – Ilustração das áreas administrativa e trabalho do COSE. O administrador cadastra imagens no sistema utilizando a área administrativa para que os usuários possam segmentá-las na área de trabalho.	53

Figura 16 – Exemplo da borda e da sobreposição em um retalho sendo segmentado. A parte translúcida mais externa corresponde a borda do retalho, enquanto que a área entre os quadrados vermelhos refere-se a área de sobreposição de retalhos. Os pontos e linhas verdes na imagem correspondem as interações do usuário.	54
Figura 17 – Exemplo de aplicação do algoritmo <i>Non-Local Means</i> : (a) imagem de um retalho da sépala de uma flor; (b) resultado da aplicação do algoritmo <i>Non-Local Means</i>	56
Figura 18 – Ilustração da interação ente um usuário e o computador. O usuário alimenta o sistema com informações que serão utilizadas pelo algoritmo de segmentação. O computador realiza o processamento e mostra para o usuário o resultado da segmentação. Fonte: (HE; KIM; KUO, 2013) .	58
Figura 19 – Ilustração do processo do Live-Wire para segmentar uma imagem. . . .	60
Figura 20 – Construção do mapa de custo. A partir de uma imagem contendo 4×4 pixels, é criado um grafo com pesos também 4×4 , onde as arestas de menor custo são correspondentes aos pixels mais claros na imagem. O custo das arestas do pixel p é dado em função dos pixels vizinhos q . A partir da criação do mapa de custo, o usuário pode especificar o contorno desejado.	61
Figura 21 – Imagem do ventrículo esquerdo sendo segmentada com o Live-Wire. os pontos s_1 e s_2 são definidos pelo usuário através do clique do botão esquerdo do mouse. O ponto livre é o ponto onde se encontra o cursor do mouse. A linha vermelha liga os ponto através do menor caminho no mapa de custo.	63
Figura 22 – Ilustração do processo de segmentação utilizando o <i>watershed</i>	65
Figura 23 – Aplicação do <i>watershed</i> em uma imagem das células da sépala. Os pontos verdes são os marcadores e as linha de watershed estão indicadas em vermelho.	65
Figura 24 – Exemplo de uma imagem da retina sendo segmentada utilizando o Live-Vessel. As cores representam o tamanho do raio. Quanto maior o raio mais azulada é a cor do pixel e quanto menor o raio, mais avermelhada é a cor do pixel. Fonte: Dickie, Hamarneh e Abugharbieh (2009). . . .	68
Figura 25 – Estrutura de um combinador de L classificadores. Cada classificador produz uma saída O_i , $i = 1, 2, \dots, L$ que é combinada para gerar uma saída $\hat{O}$	70
Figura 26 – Fluxograma do funcionamento do STAPLE.	73
Figura 27 – Fluxograma do funcionamento do SIMPLE.	75
Figura 28 – Distância $Dmap(P, F)$ entre um ponto P do <i>background</i> e o <i>foreground</i> F	76

Figura 29 – Abstração do funcionamento do DTM. Dados dois segmentos A e B que simulam a borda de um objeto, o DTM calcula a média do mapa de confiança dos dois segmentos. Após a aplicação do algoritmo de Otsu, o resultado do consenso é mostrado pelo contorno tracejado na figura. .	77
Figura 30 – Fluxograma da abordagem proposta.	80
Figura 31 – Ilustração dos pixels confusos em relação a fronteira de decisão para um problema com 5 votos. Os quadrados verdes e os círculos vermelhos representam os pixels classificados como <i>background</i> e <i>foreground</i> usando o voto majoritário, respectivamente.	81
Figura 32 – Exemplos de imagens médicas e biológicas: (a) exemplo de imagem da retina; (b) <i>ground-truth</i> da Figura 32a; (c) exemplo de imagem do ventrículo esquerdo; (d) <i>ground-truth</i> da Figura 32c; (e) imagem de uma sépala; (f) <i>ground-truth</i> da Figura 32e.	82
Figura 33 – Codificação usando o operador LBP.	84
Figura 34 – Exemplos da extensão do operador LBP com diferentes raios. Fonte: Ojala, Pietikäinen e Mäenpää (2002).	85
Figura 35 – Máscaras de Kirsch em oito direções. Fonte: Jabid, Kabir e Chae (2010). .	85
Figura 36 – Exemplo da geração do código LDP. Fonte: Jabid, Kabir e Chae (2010). .	86
Figura 37 – Exemplo da divisão de um espaço bidimensional. Na parte superior da figura é mostrado o processo de divisão do espaço baseado no <i>KD-tree</i> , enquanto que na parte inferior encontra-se a árvore formada após o processo.	88
Figura 38 – Ilustração do resultado final da abordagem proposta.	89
Figura 39 – Divisão dos dados no espaço de características do SVM utilizando um <i>kernel</i> ϕ . Os círculos vermelhos e azuis representam dados de duas classes distintas. Fonte: (JONES, 2014).	90
Figura 40 – Exemplo da criação de uma imagem sintética: (a) polígono com o número de lados e a localização dos vértices criados de modo aleatório; (b) campo randômico multi-fractal 2D; (c) resultado da adição das duas imagens.	93
Figura 41 – Exemplos de imagens das bases de dados com seus respectivos <i>ground-truths</i> : (a) imagem da retina em nível de cinza; (b) imagem do ventrículo esquerdo; (c) retalho da sépala em nível de cinza; (d) imagem da base sintética; (e) <i>ground-truth</i> da Figura 41a; (f) <i>ground-truth</i> da Figura 41b; (g) <i>ground-truth</i> da Figura 41c; (h) <i>ground-truth</i> da Figura 41d. .	94
Figura 42 – Comportamento do coeficiente de Dice à medida que adiciona-se novos raios.	95

Figura 43 – Coeficiente de Dice para todas as combinações das configurações descritas. As tabelas presentes na imagem são organizadas de acordo com o valor da variância ao longo dos eixos σ_x e σ_y . As colunas significam o número de orientações $\theta_n, n = 4, 6, 8$ e as linhas representam o número de frequências $f_m, m = 4, 5, 6$	96
Figura 44 – comportamento da taxa Dice para algoritmo ISECO de acordo com a variação do valor de λ	97
Figura 45 – Desempenho do ISECO para a base de dados da retina variando o valor de k	97
Figura 46 – Tempo de processamento médio para cada imagem da base de dados da retina variando o valor de k	98
Figura 47 – Exemplos de segmentações da imagem 41d obtidos pelos métodos de consenso para a base de imagens sintéticas: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.	99
Figura 48 – Exemplo de erros que podem acontecer na segmentação de bordas em imagens. A borda real do objeto é representada por uma linha sólida preta, enquanto que o <i>ground-truth</i> e a segmentação obtida são representadas por uma linha tracejada verde e uma linha pontilhada vermelha, respectivamente.	101
Figura 49 – Exemplo de um retalho da sépala rotulado erroneamente: (a) retalho da sépala; (b) <i>ground-truth</i> rotulado manualmente por um especialista; (c) resultado utilizando o ISECO.	102
Figura 50 – Exemplos de segmentações da Figura 41a obtidos pelos métodos de consenso para a base de imagens da retina: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.	103
Figura 51 – Exemplos de segmentações da Figura 41b obtidos pelos métodos de consenso para a base de imagens do ventrículo esquerdo: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.	104
Figura 52 – Exemplos de segmentações da imagem 41c obtidos pelos métodos de consenso para a base de imagens de retalhos da sépala: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.	105
Figura 53 – Coeficiente de Dice dos métodos de consenso à medida que o número de segmentações aumenta.	105
Figura 54 – Tempo de processamento à medida que o número de segmentações aumenta.	107

Figura 55 – Coeficiente de Dice para 3 usuários distintos na segmentação dos 10 primeiros retalhos da sépala apresentados pelo COSE.	108
Figura 56 – Visualização da área de trabalho com suas ferramentas: <i>undo</i> , <i>redo</i> , <i>smart tool</i> , <i>manual tracing</i> , <i>erase</i> , <i>clear</i> e <i>zoom</i> . A imagem a ser segmentada corresponde a maior parte da área de trabalho.	124
Figura 57 – Diagrama de classes do COSE.	127
Figura 56 – Fluxo de execução do COSE em telas: (a) Tela administrativa inicial; (b) Cadastro de tarefas; (c) Inserção de um problema de segmentação a uma tarefa; (d) Problema de segmentação após a inserção; (e) Criação dos retalhos; (f) Retalhos criados; (g) Retalho com pouca informações; (h) Início da segmentação utilizando o live-vessel; (h) Segmentação final para o retalho.	131

LISTA DE TABELAS

Tabela 1 – Desafios de segmentação de imagens médicas presentes na ferramenta <i>Grand Challenges in Biomedical Image Analysis</i>	37
Tabela 2 – Matriz de confusão entre a segmentação obtida e o <i>ground-truth</i>	40
Tabela 3 – Coeficiente de <i>Kappa</i> associados à qualidade da classificação.	42
Tabela 4 – Valores das medidas na comparação entre o <i>ground-truth</i> e as imagens das Figuras 14b e 14c.	50
Tabela 5 – Sensibilidade e especificidade de imagens segmentadas pelo voto majoritário.	82
Tabela 6 – Coeficiente de Dice variando o parâmetro Z entre 1 e 8.	96
Tabela 7 – Desempenho médio dos métodos de consenso na base de imagens sintéticas. O valor entre parênteses corresponde ao desvio padrão para cada medida.	99
Tabela 8 – Desempenho dos métodos de consenso na base de imagens da retina. O valor entre parênteses corresponde ao desvio padrão para cada medida.	101
Tabela 9 – Desempenho dos métodos de consenso na base de imagens do ventrículo esquerdo. O valor entre parênteses corresponde ao desvio padrão para cada medida.	101
Tabela 10 – Desempenho dos métodos de consenso na base de retalhos da sépala de uma planta. O valor entre parênteses corresponde ao desvio padrão para cada medida.	101
Tabela 11 – <i>P-values</i> obtidos utilizando o teste pareado de Wilcoxon entre a abordagem proposta e os demais métodos para as bases da retina e do ventrículo esquerdo.	102
Tabela 12 – Resultado da abordagem usando ISECO com SVM em comparação com os métodos ISECO usando KNN e STAPLE.	104
Tabela 13 – Tempo médio de processamento dos métodos de consenso para processar a base de dados da retina.	106
Tabela 14 – Desempenho dos segmentadores comparado com o resultado obtido pelo COSE.	108
Tabela 15 – Teste de hipótese de Wilcoxon entre os segmentadores e o COSE	108

LISTA DE ABREVIATURAS E SIGLAS

COLLATE	<i>C</i> onsensus <i>L</i> evel <i>L</i> abeler <i>A</i> ccuracy and <i>T</i> ruth <i>E</i> stimation
COSE	<i>C</i> ollaborative <i>S</i> Egmentation
ESP	Especificidade
ISECO	<i>I</i> ntelligent <i>S</i> Egmentation <i>C</i> onsensus
KNN	<i>K</i> -Nearest Neighbors
LBP	<i>L</i> ocal <i>B</i> inary <i>P</i> attern
LDP	<i>L</i> ocal <i>D</i> irectional <i>P</i> attern
MAH	Distância de Mahalanobis
SEN	Sensibilidade
SIMPLE	<i>S</i> elective and <i>I</i> terative <i>M</i> ethod for <i>P</i> erformance <i>L</i> evel <i>E</i> stimation
STAPLE	<i>S</i> imultaneous <i>T</i> ruth and <i>P</i> erformance <i>L</i> evel <i>E</i> stimation
SVM	<i>S</i> upport <i>V</i> ector <i>M</i> achine
VM	Voto Majoritário

SUMÁRIO

1	INTRODUÇÃO	20
1.1	Motivação	20
1.2	Objetivo	21
1.3	Contribuições	22
1.4	Organização da tese	23
2	SEGMENTAÇÃO DE IMAGENS MÉDICAS	25
2.1	Estruturas do corpo humano	25
2.1.1	Coração	25
2.1.2	Cérebro	26
2.1.3	Pulmões	27
2.1.4	Próstata	27
2.1.5	Ossos	28
2.1.6	Vasos sanguíneos	29
2.2	Modalidades de imagens médicas	29
2.2.1	Radiografia	30
2.2.2	Ultra-sonografia	30
2.2.3	Ressonância Magnética Nuclear	30
2.2.4	Tomografia Computadorizada	30
2.2.5	Tomografia por emissão de pósitrons	31
2.3	Métodos de segmentação de imagens	32
2.3.1	Segmentação baseada em regiões	32
2.3.1.1	Limiarização	32
2.3.1.2	Crescimento de regiões	33
2.3.1.3	Método da divisão e fusão	33
2.3.2	Segmentação baseada em bordas	34
2.3.2.1	Gradiente da imagem	34
2.3.2.2	Laplaciano	34
2.3.3	Outras técnicas	35
2.3.3.1	Modelos deformáveis	35
2.3.3.2	Segmentação utilizando redes neurais	35
2.3.3.3	Segmentação por textura	36
2.4	Desafios em segmentação de imagens médicas	37
2.5	Considerações finais	38
3	MEDIDAS DE AVALIAÇÃO BASEADAS EM <i>GROUND-TRUTHS</i>	39

3.1	Medidas de similaridade	39
3.1.1	Acurácia	40
3.1.2	Sensibilidade e especificidade	40
3.1.3	Área sob a curva ROC	40
3.1.4	Coeficiente de correlação de Matthews	41
3.1.5	Kappa	42
3.1.6	Dice	42
3.1.7	Índice de Jaccard	43
3.1.8	Índice de Rand	43
3.1.9	Índice ajustado de Rand	44
3.1.10	Coeficiente de Correlação Intraclasse	44
3.1.11	Informação mútua	45
3.2	Medidas baseadas em distâncias	45
3.2.1	Distância de Hausdorff	45
3.2.2	Distância média	45
3.2.3	Distância de Mahalanobis	46
3.2.4	Erro de consistência global	46
3.2.5	Variação da informação	47
3.3	Seleção das medidas	47
3.4	Considerações finais	50
4	SEGMENTAÇÃO COLABORATIVA	51
4.1	Área administrativa	53
4.2	Área de trabalho	56
4.3	Segmentação interativa	57
4.3.1	<i>Live-Wire</i>	59
4.3.2	<i>Watershed</i>	63
4.3.3	<i>Live-Vessel</i>	66
4.4	Considerações finais	67
5	CONSENSO DE SEGMENTAÇÕES	70
5.1	Voto Majoritário	71
5.2	STAPLE	71
5.3	COLLATE	73
5.4	SIMPLE	74
5.5	<i>Distance Transform Merge</i>	75
5.6	Considerações finais	77
6	CONSENSO INTELIGENTE DE SEGMENTAÇÕES	79
6.1	Pré-segmentação	79

6.2	Extração de características	83
6.2.1	Filtros de Gabor	83
6.2.2	<i>Local Binary Pattern</i> (LBP)	84
6.2.3	<i>Local Directional Pattern</i> (LDP)	85
6.3	Classificação	86
6.4	Outras abordagens do <i>Intelligent SEgmentation COnsensus</i>	89
6.5	Considerações finais	90
7	EXPERIMENTOS E RESULTADOS	92
7.1	Bases de dados e medidas para a avaliação	92
7.2	Análise das características	93
7.2.1	Seleção dos parâmetros dos extratores	94
7.3	Seleção dos parâmetros do ISECO	96
7.4	Avaliação dos métodos para o consenso de segmentações	98
7.4.1	Desempenho com imagens sintéticas	98
7.4.2	Desempenho utilizando base de dados reais	100
7.5	Análise da influência na quantidade de imagens segmentadas	102
7.5.1	Avaliação do ISECO usando SVM	103
7.6	Tempo de processamento	106
7.7	Avaliação do COSE como ferramenta para estimativa de <i>ground-truths</i>	106
8	CONCLUSÕES E DISCUSSÕES	109
	REFERÊNCIAS	112
	APÊNDICES	123
	APÊNDICE A – ÁREA DE TRABALHO DO COSE	124
	APÊNDICE B – ARQUITETURA DO COSE	126
	APÊNDICE C – FLUXO DE EXECUÇÃO DO COSE	128

1 INTRODUÇÃO

Este capítulo apresenta as principais motivações para a realização deste trabalho, descreve os objetivos e as contribuições, justificando a importância e a contribuição deste estudo para a área de processamento de imagens médicas, e, finalmente, demonstra a estruturação do restante da tese.

1.1 Motivação

O processamento e a análise de imagens caracterizam-se, de maneira geral, pelas fases de aquisição, pré-processamento, segmentação, representação/descrição e reconhecimento de imagens. Este trabalho situa-se na área de segmentação binária dos pixels em imagens médicas 2D.

Segundo [Gonzalez e Woods \(2006\)](#), o principal objetivo da segmentação consiste em dividir uma imagem em regiões ou objetos que a compõem. A segmentação pode ser de fundamental importância para etapas posteriores em processamento de imagens digitais. Esta etapa é alvo de vários estudos e pesquisas para o desenvolvimento de métodos mais abrangentes e computacionalmente mais eficientes, principalmente em imagens médicas.

A segmentação de imagens médicas pode ser considerada uma tarefa de classificação dos componentes da imagem (pixels ou voxels) em componentes anatômicos relevantes. Um papel crucial para a extração de informações automatizadas em imagens médicas geralmente envolve a segmentação das regiões da imagem a fim de quantificar volumes e áreas de interesse em tecidos biológicos para posterior diagnóstico e localização de patologias. A segmentação de imagens médicas extrai informações significativas e facilita a exibição dessas informações de uma forma clinicamente relevante. O conhecimento da localização, tamanho e forma das diferentes estruturas anatômicas é um passo fundamental na compreensão e análise de imagens médicas. O conhecimento explícito das estruturas segmentadas em imagens médicas nos permite fazer mais do que a avaliação visual qualitativa. A aplicação da segmentação de imagens médicas é tão variada, que é difícil listar os órgãos ou partes do corpo mais segmentadas, mas uma lista geral iria incluir, pelo menos, os seguintes: coração, pulmões, fígado, cérebro, ossos, próstata, vasos sanguíneos e as mamas ([GINNEKEN; KERKSTRA; SHNEIDER, 2016](#)).

Um fator importante que deve ser levado em consideração no desenvolvimento de algoritmos de segmentação de imagens médicas é como avaliar o desempenho dos métodos. Essa avaliação é necessária para comparar o desempenho de diferentes métodos de segmentação e estabelecer sua viabilidade na prática clínica ([WARFIELD; ZOU; WELLS, 2004](#)). A avaliação de desempenho geralmente baseia-se na comparação dos resultados

de um algoritmo de segmentação versus uma segmentação de referência, denominada de *ground-truth*, normalmente criada para ter alta precisão e que reflita as características da segmentação (CHALANA; KIM, 1997).

O *ground-truth* pode ser considerado uma segmentação que supõe-se ser verdadeira. Portanto, para obter o *ground-truth* de uma determinada imagem, deve-se utilizar alguma abordagem de segmentação. Em particular, a estimativa do *ground-truth* em imagens médicas pode ser obtida através do contorno manual (através do desenho de contornos e/ou preenchimento das regiões desejadas). Porém, tal abordagem, em geral, pode ser bastante custosa e estressante para o especialista que participa do processo.

Uma abordagem menos custosa é a segmentação das imagens utilizando métodos interativos (HE; KIM; KUO, 2013). Os métodos interativos de segmentação visam juntar o melhor das duas formas: a capacidade do ser humano em reconhecer objetos e a eficiência do computador em realizar cálculos e processamentos complexos. Enquanto o operador humano auxilia no processo, como por exemplo, inserindo pontos sementes ou ajustando parâmetros, o computador realiza a segmentação com base nessas informações. Durante a segmentação, o usuário tem total controle sobre o processo e é ele quem decide quando o algoritmo deve parar.

Tanto a segmentação manual quanto a segmentação interativa são consideradas abordagens consolidadas de segmentação. No entanto, esse processo pode ser subjetivo e sujeito a variabilidades inter e intra usuários. Uma alternativa para diminuir essa variabilidade é a segmentação utilizando o consenso das respostas de vários usuários para uma mesma imagem. Um método de consenso de segmentações tem a finalidade de combinar várias segmentações de uma mesma imagem. Em geral, de acordo com Mitchell (2012), uma imagem segmentada por um método de consenso apresenta melhor qualidade global do que qualquer uma das segmentações individuais.

Mais recentemente, algumas abordagens têm sido desenvolvidas na tentativa de segmentar imagens usando a colaboração baseada na Internet. Nesse contexto, propomos o COSE (*COllaborative Segmentation*). O COSE é um sistema *web* colaborativo para segmentação de imagens que alia algoritmos interativos e inteligência coletiva para realizar a tarefa de segmentar da melhor maneira possível uma imagem. O COSE pode ser utilizado de modo colaborativo para obter um conjunto de segmentações de uma mesma imagem que são combinadas utilizando um método de consenso para criar um *ground-truth*.

1.2 Objetivo

Este trabalho tem como principal objetivo propor uma abordagem para estimar segmentação verdadeira (*ground-truth*) em imagens médicas. A abordagem consiste em duas partes. Na primeira, utiliza-se o sistema colaborativo COSE para obter várias

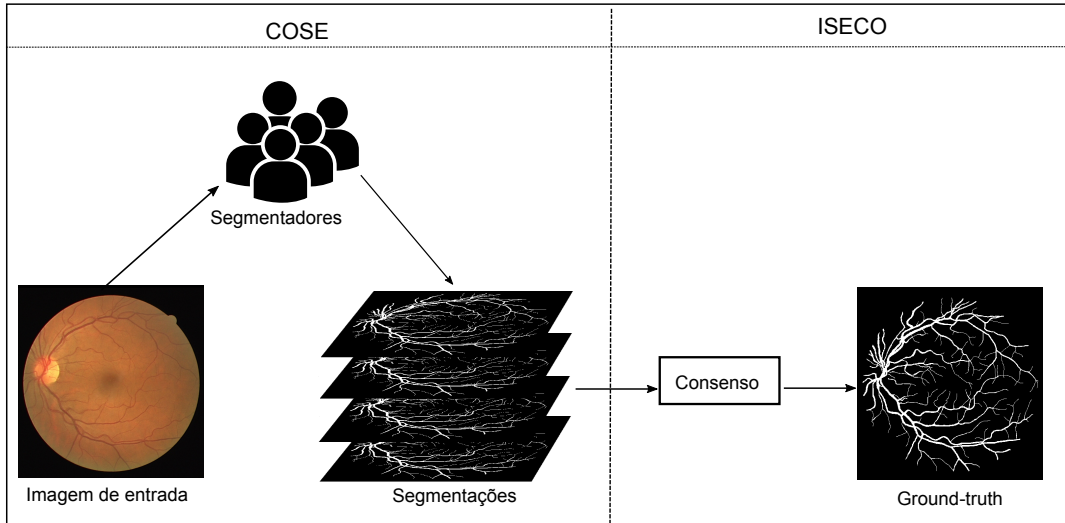


Figura 1: Relacionamento entre o COSE (*Collaborative SEgmentation*) e o ISECO (*Intelligent SEgmentation COnsensus*).

segmentações de uma mesma imagem com o intuito de estimar o *ground-truth*. Já a segunda, propõe um novo método para o consenso de segmentações de imagens médicas. O método proposto considera o problema do consenso de segmentações como um problema de classificação de padrões e é denominado *Intelligent SEgmentation COnsensus* (ISECO). Os principais métodos de consenso de segmentações utilizam apenas informações das imagens binárias. No entanto, a imagem em nível de cinza que originou essas imagens binárias não é utilizada. A proposta do novo método é diminuir o erro na fronteira de decisão e melhorar a assertividade utilizando informações extraídas da imagem em nível de cinza. Para demonstrar a eficácia do método proposto, o ISECO é comparado com outras abordagens de consenso de segmentações disponíveis na literatura. A Figura 1 ilustra o relacionamento entre o COSE e o ISECO. As segmentações obtidas pelo COSE através dos segmentadores¹ são combinadas pelo método proposto ISECO para estimar o *ground-truth*. Neste trabalho, os termos segmentação e *ground-truth* em alguns momentos se confundem, isso acontece porque todo *ground-truth* nesse contexto é uma segmentação, seja ela manual ou automática. No entanto, é importante deixar bem claro que o objetivo da segmentação aqui é a criação de *ground-truths* binários em imagens médicas.

1.3 Contribuições

Neste trabalho sobre estimativa de *ground-truths*, as principais contribuições são:

- Uma nova abordagem para o consenso de segmentações em imagens médicas utilizando reconhecimento de padrões: A abordagem proposta usa não somente informações das imagens segmentadas, mas também informações extraídas da imagem em

¹ Segmentadores podem ser humanos ou algoritmos automáticos.

nível de cinza para realizar o consenso de segmentações. Com base nessas informações e com o auxílio de um método de aprendizagem obtém-se o consenso;

- Uma nova abordagem para o consenso de segmentações em imagens médicas usando transformada de distâncias: O consenso baseia-se na transformada de distância de cada segmentação para criar um mapa de confiança. A segmentação final é obtida através de um método de limiarização;
- Aplicação dos novos métodos em quatro problemas de segmentação: As abordagens são avaliadas utilizando 4 problemas de segmentação com estruturas anatômicas diferentes. A primeira base de dados contém imagens criadas artificialmente. Na segunda base, utiliza-se imagens dos vasos da retina, enquanto a terceira é composta por imagens do ventrículo esquerdo. A última base de dados é formada por retalhos de uma imagem da sépala de uma flor;
- Desenvolvimento do sistema do colaborativo COSE (Collaborative SEgmentation) e a utilização dele como uma ferramenta para a estimativa de *ground-truths*: Um sistema colaborativo como o COSE tem a vantagem de possibilitar com que vários usuários possam participar e compartilhar informações para criar um *ground-truth*. Além disso, a subjetividade comum de cada usuário é amenizada através do consenso dessas informações;
- Revisão das principais medidas utilizadas na avaliação de métodos de segmentação baseados em *ground-truths*.

1.4 Organização da tese

O conteúdo deste trabalho de tese está organizado como segue:

O Capítulo 2 contextualiza o problema da segmentação de imagens médicas focando nas principais modalidades de aquisição de imagens médicas, métodos de segmentação e desafios encontrados na área.

O Capítulo 3 traz uma revisão das principais medidas para a avaliação de segmentações binárias baseadas em *ground-truth*. O objetivo do capítulo é situar o leitor sobre as principais medidas utilizadas no momento da avaliação de um método de segmentação de imagens médicas, tais como: coeficiente de Dice, acurácia, sensibilidade, especificidade, área sobre a curva ROC, coeficiente de correlação de Matthews, Kappa, índice de Jaccard, índice de Rand, índice ajustado de Rand, coeficiente de correlação interclasse, informação mútua, Distância de Hausdorff, distância média, Distância de Mahalanobis, erro de consistência global e variação da informação. No final, será feita uma análise das medidas com o intuito de selecionar um conjunto destas para avaliar a abordagem proposta.

No Capítulo 4 apresenta-se o sistema colaborativo COSE como uma ferramenta para estimação de *ground-truths* em imagens médicas. São descritas as principais funcionalidades e características do COSE, bem como as partes que compõem o sistema.

O Capítulo 5 aborda o estado da arte dos métodos para o consenso de segmentações. Neste capítulo, faz-se um resumo dos seguintes métodos: voto majoritário, STAPLE (*Simultaneous Truth and Performance Level Estimation*), SIMPLE (*Selective and Iterative Method for Performance Level Estimation*) e COLLATE (*COnsensus Level, Labeler Accuracy and Truth Estimation*). Neste capítulo também apresentamos uma abordagem baseada em transformada de distâncias para o problema do consenso de segmentações, denominada DTM (*Distance Transform Merge*).

O Capítulo 6 descreve em detalhes a abordagem proposta para o consenso de segmentações em imagens médicas, denominada ISECO (*Intelligent SEgmentation COnsensus*). Os métodos descritos no Capítulo 5 baseiam-se apenas em informações binárias extraídas das segmentações. No entanto, uma informação muito importante para a segmentação não é utilizada, a imagem que originou as segmentações. Essa informação pode ser bastante útil no momento de decidir se um pixel pertence ao *foreground*, principalmente na fronteira de decisão. O objetivo da abordagem proposta é realizar o consenso inteligente das segmentações adicionando informações extraídas da imagem em tons de cinza.

O Capítulo 7 apresenta os experimentos e resultados. Os experimentos têm como objetivo avaliar os algoritmos de consenso propostos de acordo com 4 bases de imagens, sendo que 3 dessas bases são compostas por imagens reais da retina, ventrículo esquerdo e sépala de uma planta, enquanto que a outra base contém imagens geradas artificialmente.

Finalmente, o Capítulo 8 apresenta as considerações finais e conclui este trabalho.

2 SEGMENTAÇÃO DE IMAGENS MÉDICAS

Segmentar significa extrair determinado objeto do resto da imagem, obedecendo algum critério para extração, como por exemplo, intensidade ou similaridade dos pixels do objeto (GONZALEZ; WOODS, 2006). Entende-se como objeto a parte da imagem que interessa para o problema a ser resolvido.

Segmentação de imagens é uma importante ferramenta em processamento de imagens médicas, sendo útil em diagnósticos (LI; NISHIKAWA, 2015), identificação de patologias (SKOUNAKIS et al., 2014), estudos anatômicos estruturais (PEREIRA et al., 2016), planejamento do tratamento (RAYLMAN et al., 2015), classificação automática de células sanguíneas (GAO et al., 2016), cirurgia auxiliada por computador (GAVAGHAN et al., 2011), etc. Por essa razão, grandes esforços têm sido prestados nos últimos anos nessa área.

Para desenvolver um novo método de segmentação de imagens médicas pode-se levar em consideração 3 fatores: estrutura anatômica do órgão a ser segmentado; modalidade da imagem e o estado da arte sobre o problema. As próximas seções descrevem esses fatores, bem como alguns desafios existentes na área da segmentação de imagens médicas.

2.1 Estruturas do corpo humano

A segmentação de imagens médicas é tão variada, que é difícil listar os órgãos ou partes do corpo mais segmentadas. No entanto, de acordo com trabalhos recentes (SUINESIAPUTRA et al., 2014; ORTIZ et al., 2011; BALAFAR, 2012; PETER et al., 2008; LEI et al., 2008; SOARES et al., 2006b; GHOSE et al., 2012), uma lista geral iria incluir, pelo menos, as estruturas descritas nas próximas seções.

2.1.1 Coração

De acordo com a OMS Organization (2015), as doenças cardiovasculares foram responsáveis por aproximadamente 17.5 milhões de mortes no mundo em 2012, representando 31% do total de mortes. O coração é o órgão central do sistema circulatório humano e é composto de quatro câmaras: dois átrios e dois ventrículos. O sangue entra pelos átrios e sai pelos ventrículos. O lado direito do coração é parte da circulação pulmonar, que recebe o sangue pobre de oxigênio e distribui para os pulmões. O lado esquerdo, por outro lado, recebe o sangue oxigenado do pulmão e o bombeia para diversos órgãos. O miocárdio é a parede muscular do coração. A superfície exterior lisa do miocárdio é chamada de epicárdio, enquanto que o revestimento interno é denominado endocárdio. Cada ciclo cardíaco é dividido em duas fases principais: diástole e sístole. Durante o período da diástole, os

átrios se contraem para empurrar o sangue para os ventrículos, que estão relaxados. Na sístole os ventrículos se contraem para bombear o sangue para os pulmões e outros órgãos do corpo, enquanto os átrios estão relaxados. O ventrículo esquerdo (VE) realiza um grande esforço durante a diástole. Justamente por isso possui um tamanho maior que as demais concavidades, o que facilita a segmentação. Uma medida quantitativa da saúde do VE é a fração de ejeção. Esta estatística mede a porcentagem de volume de sangue transmitido para fora da VE em um determinado ciclo cardíaco. Para calcular fração de ejeção, precisamos segmentar o VE em múltiplos pontos de um ciclo cardíaco. Além disso, observando como o VE evolui ao longo de um ciclo cardíaco, permite aos médicos determinar a saúde do miocárdio. A Figura 2 ilustra uma imagem típica do coração usada em problemas de segmentação (PENG et al., 2016).

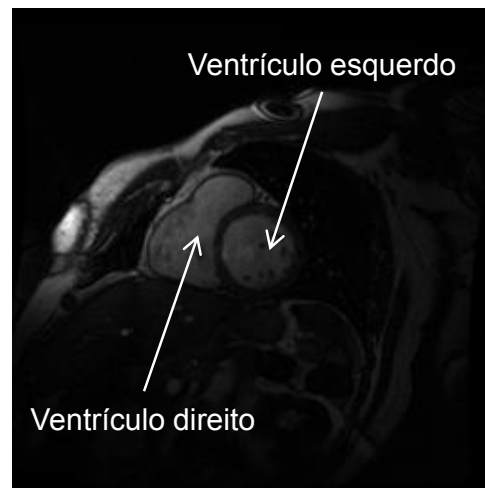


Figura 2: Exemplo de uma imagem obtida por ressonância magnética utilizada na segmentação do ventrículo esquerdo. Na imagem os destaques são o ventrículo esquerdo e o direito.

2.1.2 Cérebro

De acordo com Balafar et al. (2010), a identificação de estruturas cerebrais em imagens é muito importante na neurociência e tem muitas aplicações, tais como: mapeamento da ativação funcional sobre a anatomia do cérebro, o estudo do desenvolvimento do cérebro e a análise da variabilidade neuroanatômica em cérebros normais. A segmentação de imagens do cérebro também é útil no diagnóstico clínico de doenças neurodegenerativas, psiquiátricas, avaliação de tratamento e planejamento cirúrgico (HAN; FISCHL, 2007). Segundo Kapur et al. (1998), a segmentação de imagens do cérebro possui três objetivos. O primeiro é delimitar o cérebro em relação à caixa craniana. O segundo objetivo é identificar três tipos de tecidos cerebrais: tecido cinza, tecido branco e o fluido espinhal. O tecido cinza é composto principalmente por células nervosas. O tecido branco é o meio pelo qual os sinais do cérebro circulam. O fluido espinhal tem a função de sustentar e amortecer choques contra o crânio. O último objetivo da segmentação do cérebro é delinear estruturas

anatômicas específicas do cérebro como o córtex, o hipotálamo e os ventrículos. A Figura 3 exemplifica uma imagem do cérebro.

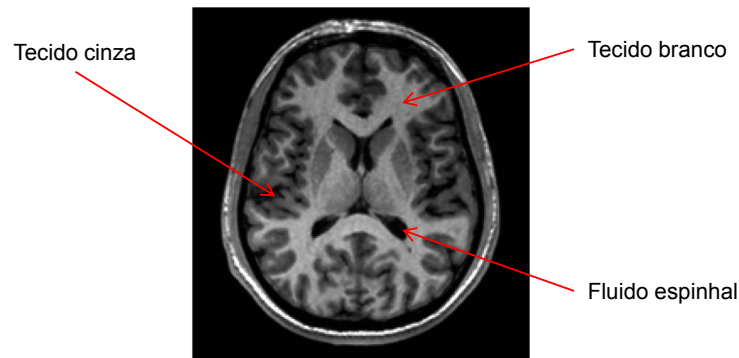


Figura 3: Imagem do cérebro e seus tecidos obtida por ressonância magnética.

2.1.3 Pulmões

Os pulmões são órgãos esponjosos e elásticos fundamentais na respiração. São duas vísceras situadas uma de cada lado, no interior do tórax e onde se dá o encontro do ar atmosférico com o sangue do corpo. Eles estendem-se do diafragma até um pouco acima das clavículas e estão protegidos pelas costelas. O pulmão direito apresenta-se dividido em três lobos (superior, médio e inferior), já o pulmão esquerdo está dividido em dois lobos (superior e inferior). Diversas patologias podem comprometer a correta função pulmonar. Entre elas estão os tumores, embolia pulmonar, colapso dos pulmões (atelectasia), pneumonia, enfisema, asma e muito mais. Em alguns casos, para um diagnóstico e tratamento adequado, as respectivas patologias devem ser identificadas e quantificadas (TOMASHEFSKI; FARVER, 2008). A Figura 4 ilustra uma imagem dos pulmões.



Figura 4: Ilustração de uma imagem dos pulmões.

2.1.4 Próstata

A próstata é uma glândula pertencente ao sistema reprodutor do homem, responsável por produzir um líquido que se junta à secreção da vesícula seminal para formar o sêmen e auxiliar no transporte dos espermatozoides durante a ejaculação. De acordo com Ghose et al. (2012), a segmentação da próstata é utilizada em diversos tratamentos de doenças da próstata, como braquiterapia, ultrassom focalizado de alta intensidade (HIFU- *high*

intensity focused ultrasound), crioterapia e na terapia transuretral por micro-ondas. Além disso, tanto o volume da próstata quanto o seu contorno são úteis no acompanhamento da braquiterapia de próstata. Outra vantagem é que a segmentação da glândula da próstata facilita a fusão da imagem multimodal para a localização do tumor na biópsia (AMIS, 1994). A Figura 5 apresenta uma imagem da próstata que pode ser utilizada em problemas de segmentação.

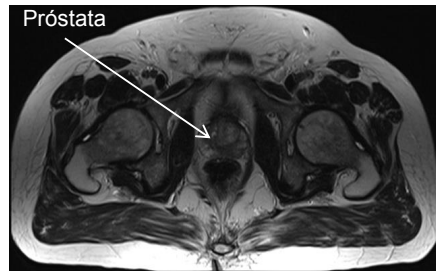


Figura 5: Exemplo de uma imagem de ressonância magnética da próstata usada no problema de segmentação.

2.1.5 Ossos

De acordo com Gilles, Moccozet e Magnenat-Thalmann (2006), deficiências músculo-esqueléticas afetam seriamente a maioria dos indivíduos com idade superior a 50 anos. A artrose é muitas vezes a origem destas deficiências, e seus sintomas típicos são a inflamação, dor, rigidez e perda de mobilidade (FELSON, 2006). A análise morfológica do osso é importante para compreender que fatores podem levar a artrose a se agravar. Além disso, a segmentação do osso pode servir como base para a análise de outras estruturas, tais como as cartilagens ou músculos (GILLES; MOCCOZET; MAGNENAT-THALMANN, 2006). A Figura 6 mostra uma imagem dos ossos do joelho que pode ser utilizada em problemas de segmentação.



Figura 6: Exemplo de uma imagem de raio-X dos ossos do joelho.

2.1.6 Vasos sanguíneos

A segmentação dos vasos sanguíneos é essencial para várias aplicações médicas, tais como o diagnóstico dos vasos e o registro de imagens de pacientes obtidas em diferentes momentos de tempo (CHAUDHURI et al., 1989). Os vasos da retina e as artérias coronárias são os dois principais tipos de vasos mais estudados na literatura em segmentação de imagens médicas (SOARES et al., 2006b). As doenças coronarianas são o tipo mais comum de doenças cardíacas. Elas geralmente são a consequência da aterosclerose que resulta no endurecimento e estreitamento das artérias que fornecem sangue ao músculo do coração (BOURAOUI et al., 2008). A análise da retina pode ser muito importante para os médicos, pois através dela é possível visualizar o olho humano por dentro e ter uma ideia de como está a circulação dos vasos sanguíneos do corpo (SOARES et al., 2006a). Ela permite examinar o tecido nervoso e vascular vivo de forma bastante ampliada. A análise dos vasos da retina auxilia no diagnóstico da hipertensão, diabetes, arteriosclerose e doenças cardiovasculares. A Figura 7 exemplifica uma imagem bastante utilizada na segmentação dos vasos sanguíneos da retina.

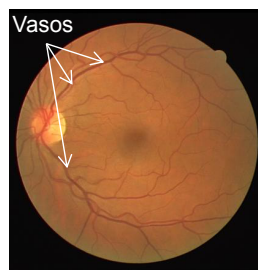


Figura 7: Exemplo de uma imagem da retina com vasos sanguíneos.

2.2 Modalidades de imagens médicas

Em 1895 Röntgen descobriu a possibilidade de usar a radiação eletromagnética, denominada raio-X, capaz de atravessar materiais, ser parcial ou totalmente absorvida por eles e fixar-se em anteparos (SMITH; WEBB, 2010). Graças a essa descoberta, a dinâmica dos exames médicos sofreu uma grande transformação, tornando possível a visualização e o diagnóstico de estruturas internas do corpo.

Desde então, o avanço na aquisição, processamento e armazenamento de imagens médicas vem permitindo o aperfeiçoamento de diagnósticos e tratamentos de diversas doenças. Cada modalidade de imagem médica atua de forma distinta, visando ressaltar as diferenças entre os diversos tipos de tecido do corpo humano, fazendo com que estruturas de interesse se sobressaíam e, assim, possam ser identificadas como eventuais anomalias. A seguir são descritas as principais modalidades de aquisição de imagens médicas (SMITH; WEBB, 2010).

2.2.1 Radiografia

Segundo [Smith e Webb \(2010\)](#), a radiografia utiliza radiação eletromagnética para produzir imagens da estrutura interna de uma pessoa enviando feixes de raios-X através do corpo, que são absorvidos em diferentes quantidades dependendo da densidade do material. Dentro do corpo humano, por exemplo, ossos, músculos e gordura absorvem os fótons com diferentes níveis de absorção, levando a diferentes atenuações e diferentes intensidades de radiação que chegam ao anteparo no qual a imagem será formada. Isso resulta em diferentes tons de cinza. É essa diferença que possibilita a formação do contraste, permitindo a identificação de estruturas anatômicas. Os raios-X são tipicamente utilizados para registrar: ossos quebrados; cavidades; objetos que foram engolidos; pulmões; vasos sanguíneos e mamas ([SMITH; WEBB, 2010](#)). A Figura 8a ilustra um exemplo de imagem das mãos de uma pessoa obtida por radiografia.

2.2.2 Ultra-sonografia

[Bega et al. \(2001\)](#) define a ultra-sonografia como uma modalidade de exame baseada na reflexão de som. A imagem é formada quando uma onda sonora é refletida ao chocar-se com a região anatômica em análise. A máquina de ultra-som envia ondas sonoras para o corpo e é capaz de converter os ecos de som retornados em uma imagem. A tecnologia de ultra-som também pode produzir sons audíveis do fluxo sanguíneo, permitindo que os profissionais médicos usem sons e imagens para avaliar a saúde do paciente. A ultra-sonografia é normalmente usada para avaliar: gravidez; anormalidade no coração e vasos sanguíneos; órgãos na pelve e abdômen; sintomas de dor, inchaço e infecção ([BEGA et al., 2001](#)). A Figura 8b apresenta uma ultra-sonografia de um feto de 12 semanas.

2.2.3 Ressonância Magnética Nuclear

[Dainty e Shaw \(1974\)](#) definem a ressonância magnética nuclear como uma tecnologia de imagem médica que usa ondas de rádio e um campo magnético para criar imagens detalhadas dos órgãos, oferecendo detalhes anatômicos e informações estruturais/fisiológicas de forma superior à fornecida por outras modalidades. A ressonância magnética provou ser altamente eficaz no diagnóstico de anormalidades, sendo capaz de mostrar a diferença entre tecidos moles normais e doentes. Dentre as aplicações da ressonância magnética pode-se destacar a avaliação de: vasos sanguíneos; tecidos anormais; ossos e articulações; lesões na coluna; problemas no tendão e no ligamento ([WANG, 2012](#)). A Figura 8c mostra uma ressonância magnética do cérebro humano.

2.2.4 Tomografia Computadorizada

Conforme define [Smith e Webb \(2010\)](#), a Tomografia Computadorizada é um método de imagem médica que combina várias projeções de raios-X tomadas de ângulos

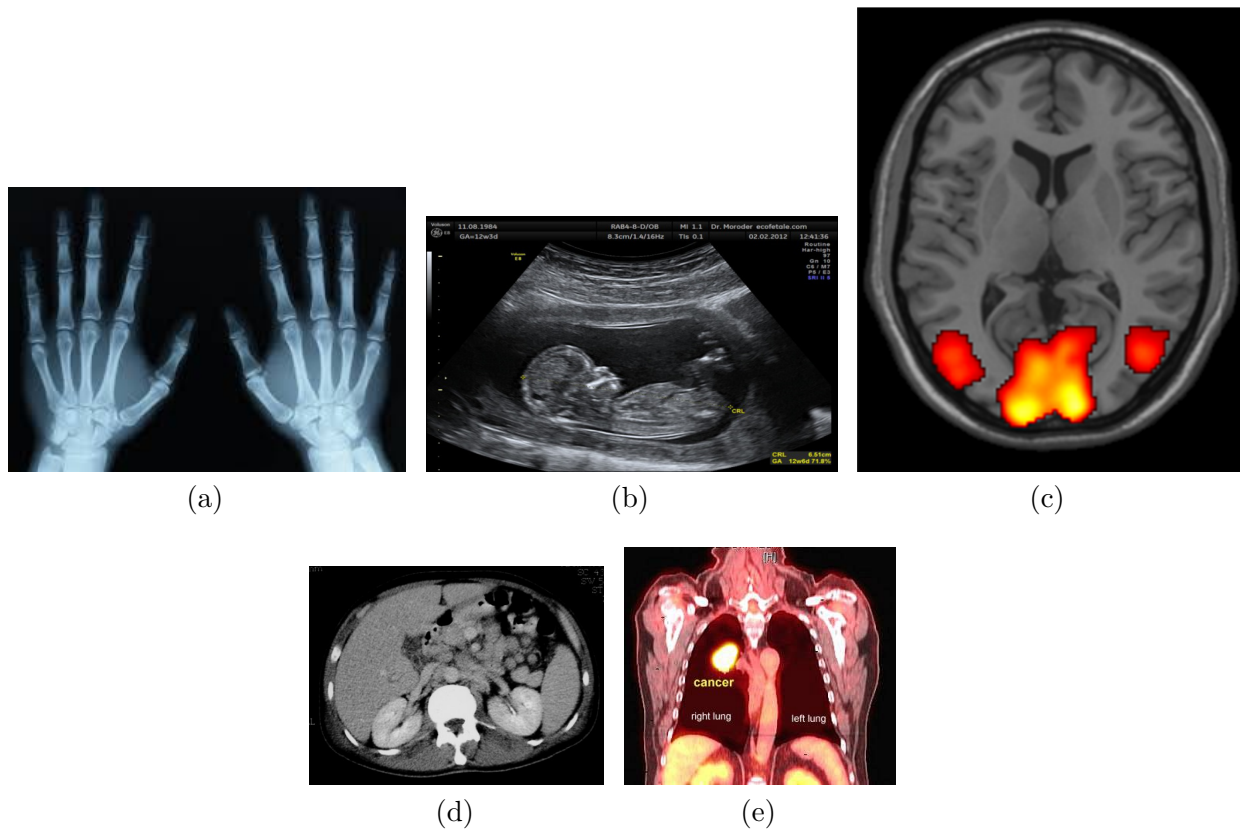


Figura 8: Exemplos de imagens médicas obtidas por diferentes modalidades de captura: (a) raio-X das mãos de uma pessoa; (b) ultra-sonografia de um feto de 12 semanas; (c) imagem de ressonância magnética do cérebro; (d) tomografia computadorizada do abdômen de uma pessoa adulta; (e) tomografia por emissão de pósitrons de um paciente com câncer no pulmão.

diferentes para produzir imagens detalhadas de áreas transversais dentro do corpo. Essas imagens permitem que os médicos obtenham visões tridimensionais muito precisas de certas partes do corpo, como tecidos moles, pelve, vasos sanguíneos, pulmões, cérebro, coração, abdômen e ossos. A Figura 8d exemplifica a tomografia computadorizada do abdômen de uma pessoa.

2.2.5 Tomografia por emissão de pósitrons

Tomografia por emissão de pósitrons é uma técnica de imagem nuclear que fornece aos médicos informações sobre como os tecidos e órgãos estão funcionando (SMITH; WEBB, 2010). Essa modalidade usa um *scanner* e uma pequena quantidade de radiofármacos que são injetados nos vasos sanguíneos do paciente para ajudar na captura de imagens detalhadas de áreas internas do corpo. Ela pode ser usada para avaliar: doenças neurológicas tais como Alzheimer e esclerose múltipla; câncer; eficácia de tratamentos e condições cardíacas (BAILEY et al., 2004). A Figura 8e mostra a tomografia por emissão de pósitrons de um paciente com câncer no pulmão.

2.3 Métodos de segmentação de imagens

Uma variedade de técnicas de segmentação de imagens médicas tem sido proposta, no entanto, nenhuma técnica está hábil a prover segmentação para todos os tipos de imagens, ou seja, para cada aplicação, tipo de aquisição e tipo da imagem, existe uma técnica apropriada.

De acordo com Rogowska (2009) as técnicas de segmentação podem ser divididas em muitas categorias dependendo do esquema de classificação:

- Manual, semi-automático e automático;
- Métodos baseados em pixels e baseados em região;
- Delineação Manual, segmentação em baixo nível (limiarização, crescimento de regiões, etc) e segmentação baseada em modelos (programação dinâmica, técnicas multipecstrais, etc.);
- Clássicos (limiarização, técnicas baseadas em borda ou regiões), técnicas baseadas em textura, modelos e de redes neurais.

Este trabalho irá se concentrar na última categoria de classificação, os métodos clássicos. Os métodos descritos a seguir foram escolhidos por serem alguns dos mais tradicionais (GONZALEZ; WOODS, 2006; PETROU; PETROU, 2010).

2.3.1 Segmentação baseada em regiões

A detecção de regiões em uma imagem pode ser feita com dois objetivos, extrair uma determinada região ou dividir a imagem num conjunto de regiões distintas (GONZALEZ; WOODS, 2006). Uma região de uma imagem é definida como um conjunto de pontos ligados onde, de qualquer ponto da região pode-se chegar a qualquer outro ponto por um caminho completamente contido nessa região. Essas regiões normalmente apresentam alguma característica homogênea, que comumente é uma certa continuidade do nível de cinza de seus pixels. A seguir são descritos alguns dos métodos mais conhecidos de segmentação baseados em região (GONZALEZ; WOODS, 2006).

2.3.1.1 Limiarização

Também conhecida como binarização, a limiarização é a técnica que divide a imagem em duas ou mais partes considerando valores de limiares (GONZALEZ; WOODS, 2006). Consiste, basicamente, em alterar os valores dos pixels de uma imagem deixando-a com uma quantidade menor de níveis com o objetivo de separar estruturas de interesse do fundo da imagem. No algoritmo básico dessa técnica o *foreground* e o *background* são

separados de acordo com um limiar, onde os resultados 1 e 0 correspondem ao *foreground* e ao *background*, respectivamente. É uma técnica vantajosa pois utiliza pouca memória e é um processo matematicamente simples, mas peca por necessitar de imagens com características específicas com bom contraste e boa nitidez. [Dawoud \(2011\)](#) utilizou uma abordagem de limiarização iterativa para segmentar os pulmões. A principal contribuição é a combinação de informações de intensidade e forma. Os resultados mostram que a técnica é mais assertiva que outros métodos, como o algoritmo de Otsu. [Roychowdhury, Koozekanani e Parhi \(2015\)](#) utilizam um método de binarização para segmentar vasos sanguíneos da retina. Primeiramente, os vasos são segmentados por um método de limiarização global, em seguida, utiliza-se um método de binarização iterativo. Os resultados mostram que o método apresenta acurácia superior aos demais métodos descritos no artigo.

2.3.1.2 Crescimento de regiões

Segundo [Gonzalez e Woods \(2006\)](#) crescimento de regiões é definido como um agrupamento de pixels ou sub-regiões em regiões maiores baseado em algum critério de similaridade. A aproximação de processamento mais simples é a agregação de pixels, isto é: escolhe-se um pixel ou um conjunto de pixels denominados sementes e faz-se o crescimento da região através da agregação de pixels vizinhos às sementes que possuem propriedades similares (intensidade, cor, textura, etc). O processo continua até que uma condição pré-estabelecida de parada seja alcançada, como por exemplo, um determinado nível de cinza ou uma distância específica.

A vantagem é que a imagem não precisa ser homogênea, pois as suas características são previamente analisadas e incluídas nos descritores de semelhança. As principais desvantagens são ([GONZALEZ; WOODS, 2006](#)): dificuldade na seleção dos pixels sementes (a aplicação deve ser conhecida); dificuldade no estabelecimento das propriedades de semelhança (a aplicação e os tipos de dados da imagem devem ser conhecidos) e dificuldade na determinação de condições de parada (depende da análise da imagem). Por isso, Crescimento de regiões, é normalmente usado para a detecção de estruturas simples. [Gibbs et al. \(1996\)](#) utilizou essa técnica para detectar tumores e lesões em pacientes. No trabalho de [Wu, Barba e Gil \(1996\)](#) imagens contendo células foram segmentadas usando uma abordagem de crescimento de regiões.

2.3.1.3 Método da divisão e fusão

Enquanto o método do crescimento de regiões é executado a partir de pontos sementes, o método da divisão e fusão agrupa e desagrupa os pixels até que se tenha regiões que obedeçam algum critério de homogeneidade. Esse método é normalmente associado a uma estrutura *quadtree*¹ ([GONZALEZ; WOODS, 2006](#)). Construir essa estrutura (*quadtree*)

¹ Uma árvore cujos nós têm exatamente quatro quadrantes.

significa dividir inicialmente uma região em quatro quadrantes e, para cada quadrante realizar sucessivas divisões e fusões de acordo com um determinado critério. [Manousakas et al. \(1998\)](#) utilizou o método da divisão e fusão para segmentar imagens de ressonância magnética do cérebro. Segundo [Hosseinzadeh e Khoshvaght \(2015\)](#), a principal limitação do método é o aspecto de blocos que a imagem segmentada pode apresentar.

2.3.2 Segmentação baseada em bordas

Contorno ou borda é definido como uma mudança brusca de níveis de cinza entre duas regiões relativamente homogêneas ([GONZALEZ; WOODS, 2006](#)). Ele pode aparecer como uma sequência de pontos, uma linha, um segmento, uma curva ou uma forte variação do nível de cinza médio. A seguir são descritos métodos de segmentação por detecção de bordas.

2.3.2.1 Gradiente da imagem

O operador gradiente é um dos procedimentos mais utilizados para detectar as discontinuidades entre regiões. O gradiente é uma derivada de primeira ordem. A ideia básica de segmentação através desse método é aplicar um processamento computacional baseado em um operador local de derivadas. Computacionalmente, consiste em obter derivadas parciais para todos os pixels da imagem. Considerando-se uma vizinhança de 3×3 pixels em torno de um ponto (x, y) , as equações 2.1 e 2.2 definem o operador gradiente na direção x e y , respectivamente ([GONZALEZ; WOODS, 2006](#)).

$$Gx = (x_7 + 2x_8 + x_9) - (x_1 + 2x_2 + x_3) \quad (2.1)$$

$$Gy = (x_3 + 2x_6 + x_9) - (x_1 + 2x_4 + x_7), \quad (2.2)$$

onde x_5 é o pixel sobre o qual está sendo aplicado o operador, e os demais x_i são os pixels pertencentes à vizinhança (oito pixels adjacentes centrados em x_5).

Uma vez detectadas as bordas da imagem, podem ser realizados processamentos para completá-las, se necessário, e interpretá-las, separando-se as partes desejadas. Esse método tem a vantagem de apresentar simplicidade na aplicação, apresentando baixa complexidade computacional. Como desvantagem está o fato de ser sensível a ruídos. [Liu e Li \(2006\)](#) usam informações do operador de gradiente para auxiliar na segmentação automática de imagens médicas.

2.3.2.2 Laplaciano

O laplaciano é uma derivada de segunda ordem. Segundo [Rogowska \(2009\)](#), quando aplica-se operador laplaciano na imagem, o resultado possui altas magnitudes nos picos e cruzamento por zero nos pontos de inflexão do contorno. O cruzamento por zero (*zero-crossing*) indica a presença de aresta. Quatro máscaras comumente usadas para calcular o laplaciano são ilustradas na Figura 9.

0	1	0	1	1	1	0	-1	0	-1	-1	-1
1	-4	1	1	-8	1	-1	4	-1	-1	8	-1
0	1	0	1	1	1	0	-1	0	-1	-1	-1

Figura 9: Exemplos de máscaras utilizadas pelo operador laplaciano.

O requisito básico para definir uma máscara laplaciana é que o coeficiente do pixel central tenha seu sinal diferente dos demais. Isso acontece devido ao fato do Laplaciano ser uma derivada, e como tal, sua soma tem que ser zero.

A mais importante contribuição do laplaciano em segmentação de imagens é encontrar o *zero-crossing*. No entanto, de acordo com Rogowska (2009) o cálculo do *zero-crossing* é em geral complicado e embora as posições dos *zero-crossing* estejam corretas para um contorno ideal, muitos erros podem acontecer. Lam e Yan (2008) utilizou o laplaciano para segmentar imagens da retina.

2.3.3 Outras técnicas

2.3.3.1 Modelos deformáveis

É uma técnica baseada em modelo, para a delineação do contorno do objeto. Também conhecido como contornos ativos, modelos deformáveis foram primeiramente introduzidos por Kass, Witkin e Terzopoulos (1988). Segundo Pham, Xu e Prince (2000), modelos deformáveis são curvas ou superfícies estabelecidas sobre uma imagem que se modificam, ou se deformam, diante da influência de forças internas e/ou externas, a fim de alinhar-se ao objeto de interesse. A correspondência do modelo com o contorno é obtida pela busca de um vetor de parâmetros que minimize a soma das energias internas e externas.

Modelos deformáveis têm sido amplamente utilizados em aplicações médicas, por exemplo, Gong et al. (2004) usaram modelos deformáveis para segmentar a próstata, enquanto que Cheng et al. (2015) os utilizaram para segmentar os vasos sanguíneos da retina.

2.3.3.2 Segmentação utilizando redes neurais

As Redes Neurais Artificiais (RNAs) (HAYKIN, 2007), são modelos matemáticos inspirados em redes neurais biológicas, em tecidos nervosos em geral, também conhecidas como connexionismo ou sistemas de processamento paralelo e distribuídos. Segundo Haykin (2007), uma RNA é um processador massiva paralelamente distribuído e composto por unidades de processamento simples, que têm a capacidade natural em armazenar conhecimento experimental. Trata-se de um modelo inspirado no cérebro, pois o conhecimento é adquirido a partir de um processo de aprendizagem e as forças das conexões inter-neurônios,

conhecidas como pesos sinápticos, são utilizados para armazenar o conhecimento adquirido. O trabalho de [Moeskops et al. \(2016\)](#) descreve uma abordagem para a segmentação de imagens do cérebro usando redes neurais.

Redes neurais podem ser usadas em diferentes contextos, tal como, na segmentação de imagens médicas. A abordagem mais amplamente usada é a supervisionada ([RODRÍGUEZ; CONDE; FRAILE, 2016](#)), onde os pesos são obtidos utilizando uma etapa de treinamento dos dados. De acordo com [Pham, Xu e Prince \(2000\)](#) uma rede neural pode também ser utilizada de modo não supervisionado, segmentando a imagem através de agrupamentos. Esse é o caso das redes de mapas auto-organizáveis (*Self-Organizing Maps* - SOM), que utilizam um algoritmo de aprendizado competitivo, em que os neurônios da camada de saída competem entre si para se tornarem ativos ([HAYKIN, 2007](#)). [Torbat, Ayatollahi e Kermani \(2014\)](#) desenvolveram uma rede neural para segmentação de imagens médicas baseada em mapas auto-organizáveis, enquanto que [Veredas, Mesa e Morente \(2010\)](#) segmentam imagens da pele para identificar feridas.

Nos últimos anos vem crescendo o número de trabalhos em segmentação de imagens médicas que utilizam redes *deep learning* ([XING; XIE; YANG, 2016](#); [MANSOOR et al., 2016](#)). Segundo [LeCun, Bengio e Hinton \(2015\)](#), *deep learning* permite que modelos computacionais compostos de várias camadas de processamento aprendam representações de dados com vários níveis de abstração. A rede *deep learning* mais utilizada em processamento de imagens, conhecida como *Convolutional Neural Network* (CNN), utiliza um conjunto de camadas de convolução e é capaz de aprender filtros para extração de características no momento do treinamento da rede. [Veredas, Mesa e Morente \(2010\)](#) utilizam CNN para segmentar os vasos sanguíneos da retina, enquanto que [Moeskops et al. \(2016\)](#) aplicaram CNN para segmentar os tecidos cerebrais.

2.3.3.3 Segmentação por textura

De acordo com [Gonzalez e Woods \(2006\)](#), textura refere-se a um padrão visual que pode ser caracterizado por variações locais em valores de pixels que se repetem de maneira regular ou aleatória ao longo do objeto ou imagem. É uma expressão relacionada com as propriedades que representam a superfície de um objeto. Apesar de não existir uma definição única e precisa, nem uma aproximação matemática formal para a quantificação da textura, as definições existentes associam a textura à impressão de rugosidade e contraste ou a repetição de padrões visuais sobre uma região. A análise da textura permite segmentar uma imagem considerando as propriedades de regularidade e de repetição presentes na imagem. Segundo [Gonzalez e Woods \(2006\)](#) os principais métodos para segmentar uma imagem pela textura são: estatístico, estrutural e espectral. O método estatístico baseia-se na distribuição de padrões de níveis de cinza na imagem. No método estrutural utiliza-se a ideia de que texturas são compostas de primitivas dispostas de forma aproximadamente

regular e repetitiva, de acordo com regras bem definidas. Já no espectral, uma ferramenta muito utilizada é o espectro de Fourier, que permite detectar a direção das texturas em uma imagem. Como exemplo de aplicação, Xie, Jiang e Tsui (2005) segmentam imagens dos rins utilizando um método baseado em textura.

2.4 Desafios em segmentação de imagens médicas

Vários aspectos das características da segmentação e algoritmos têm sido amplamente explorados durante o últimos anos em uma série de publicações. No entanto, de acordo com Pham, Xu e Prince (2000) o problema permanece desafiador, sem solução geral e única, devido a um grande e crescente número de diferentes objetos de interesse, variações nas propriedades das imagens, diferentes modalidades de imagens médicas e ruídos presentes nos objetos. O problema é tão desafiador que dentre os desafios propostos no *Grand Challenges in Biomedical Image Analysis*², os que envolvem segmentação de imagens correspondem a aproximadamente 45% do total. O *Grand Challenges in Biomedical Image Analysis* é uma ferramenta criada em 2007 que tem como objetivo auxiliar na validação e divulgação de métodos em processamento de imagens médicas. Portanto, a ideia é que com essa ferramenta seja possível publicar mais facilmente os resultados de uma pesquisa, bem como validar os métodos utilizando a mesma base de dado e a mesma medida de avaliação. Existem vários desafios envolvendo imagens médicas e a cada ano novos desafios são propostos. Dentre os problemas de segmentação de imagens médicas encontrados nessa ferramenta, pode-se destacar os apresentados na Tabela 1.

Tabela 1: Desafios de segmentação de imagens médicas presentes na ferramenta *Grand Challenges in Biomedical Image Analysis*.

Problema	Órgão	Modalidade da imagem	Ano
Segmentação de tumores no fígado	Fígado	Tomografia computadorizada	2017
Segmentação de esclerose múltipla	Cérebro	Ressonância magnética	2016
Segmentação da espessura da parede do átrio esquerdo	Coração	Ressonância magnética e tomografia computadorizada	2016
Segmentação de vértebras fraturada	Ossos	Tomografia computadorizada	2016
Localização e segmentação do disco intervertebral a partir de Imagens 3D	Ossos	Ressonância magnética	2016
Segmentação de nervos (Plexo braquial)	Nervos	Ultra-sonografia	2016
Segmentação automática da cabeça e pescoço	Cabeça	Tomografia computadorizada	2015
Segmentação de tumores cerebrais em imagens multimodais	Cérebro	Ressonância magnética	2015
Segmentação de lesões provocadas por um acidente vascular cerebral isquêmico	Cérebro	Ressonância magnética	2015
Segmentação de cistos da retina	Retina	Tomografia de coerência óptica	2015
Segmentação da matéria cinza, matéria branca e do líquido cefalorraquidiano	Cérebro	Ressonância magnética	2015
Segmentação de lesões de esclerose múltipla longitudinal	Cérebro	Ressonância magnética	2015
Segmentação da espinha e da vértebra	Ossos	Tomografia computadorizada	2015
Segmentação de tumores cerebrais	Cérebro	Ressonância magnética	2014
Segmentação do átrio esquerdo	Coração	Tomografia computadorizada	2013
Segmentação da estrutura anatômica de radiografias do tórax	Tórax	Radiografia	2013
Segmentação de estruturas da próstata	Próstata	Ressonância magnética	2013
Segmentação de neurites em imagens 3D	Cérebro	Microscópio eletrônico	2013
Segmentação dos vasos sanguíneos do pulmão	Vasos sanguíneos	Tomografia computadorizada	2012
Segmentação de imagens da próstata	Próstata	Ressonância magnética	2012
Segmentação do ventrículo direito	Coração	Ressonância magnética	2012
Segmentação de tecidos cerebrais neonatais	Cérebro	Ressonância magnética	2012
Segmentação de estruturas neuronais	Cérebro	Microscópios eletrônicos	2012
Segmentação dos pulmões e lóbulos em imagens do tórax	Pulmões	Tomografia computadorizada	2011
Segmentação das bordas do Lúmen e das laminadas elásticas externas	Vasos sanguíneos	Ultra-sonografia	2011
Segmentação da cartilagem do joelho	Ossos	Ressonância magnética	2010
Segmentação do ventrículo esquerdo	Coração	Ressonância magnética	2009

² <https://grand-challenge.org/All_Challenges/>

2.5 Considerações finais

Esse capítulo apresentou uma introdução ao problema da segmentação no contexto de imagens médicas, focando nos principais órgãos do corpo humano utilizados em problemas de segmentação, modalidades de imagens e métodos de segmentação. Apesar da segmentação na medicina ser objeto de estudo há muitos anos, o problema ainda continua desafiador, como pode-se perceber pelo número de desafios de segmentação encontrados na ferramenta *Grand Challenges in Biomedical Image Analysis*. Esse trabalho explora um problema específico em segmentação de imagens médicas que é a criação de *ground-truths* usando métodos de consenso. Os próximos capítulos se aprofundam mais no tema, com um enfoque nas medidas de avaliação de desempenho baseados em *ground-truth* e nos métodos de consenso de segmentações usados para a criação de *ground-truths*.

3 MEDIDAS DE AVALIAÇÃO BASEADAS EM *GROUND-TRUTHS*

Segmentação de imagens é um dos problemas mais estudados em processamento de imagens. No entanto, apesar de terem sido desenvolvidos vários métodos de segmentação, pouca atenção tem sido dada à forma de avaliar segmentação, o que dificulta afirmar que um determinado algoritmo é melhor que outro (ZHANG; FRITTS; GOLDMAN, 2008). Um dos métodos mais simples de avaliar uma segmentação é a avaliação subjetiva, no qual um especialista compara visualmente as segmentações. No entanto, de acordo com Zhang, Fritts e Goldman (2008), quando se compara vários métodos esse processo é lento e tedioso. Uma alternativa é utilizar um método de avaliação objetivo, em que a imagem segmentada é comparada com uma segmentação manual, denominada de *ground-truth*. Nessa abordagem, uma vez construído o *ground-truth*, pode-se avaliar vários métodos de segmentação de forma automática.

A avaliação objetiva da segmentação de imagens não é um procedimento simples. De acordo com Soares et al. (2006c), uma das dificuldades encontradas na avaliação dos métodos de segmentação de imagens é o estabelecimento de uma medida confiável a partir de imagens rotuladas manualmente. Existem diferentes critérios de qualidade descritos na literatura e cada um deles melhor avalia uma determinada característica. Segundo Khooshabi (2013), as medidas mais comuns incluem similaridade e distância. As medidas de similaridade comparam a consistência entre o *ground-truth* e a imagem segmentada por um determinado algoritmo. As medidas de distância medem a diferença entre o *ground-truth* e a imagem segmentada.

A estimação de *ground-truths* está ligada às medidas de avaliação. Portanto, o objetivo deste capítulo é fazer uma revisão das principais medidas para avaliação baseada em *ground-truths*. As medidas baseadas em similaridade e distância serão expostas nas próximas seções. No final, propõe-se o uso de um conjunto de medidas para avaliar a abordagem proposta para o consenso de segmentações em imagens médicas.

3.1 Medidas de similaridade

Nesta seção serão descritas as seguintes medidas de similaridade: acurácia (accuracy - *ACC*), sensibilidade (SEN), especificidade (ESP), área sob a curva ROC (*Area Under ROC* - *AUC*), coeficiente de correlação de Matthews (*Matthews Correlation Coefficient* - *MCC*), Kappa (*Kappa*), coeficiente de Dice (*Dice*), índice de Jaccard (*INJ*), índice de Rand (*INR*), índice ajustado de Rand (*IRA*), coeficiente de correlação interclasse (*CCI*) e informação mútua (*IFM*).

3.1.1 Acurácia

A acurácia mede o grau de conformidade entre a segmentação obtida e o *ground-truth* (ZHANG; FRITTS; GOLDMAN, 2008). O valor máximo para a acurácia é 1 e ela é definida por:

$$ACC = \frac{\text{Número de pixels corretamente classificados}}{\text{Número total de pixels}} = \frac{TP + TN}{TN + TP + FN + FP}, \quad (3.1)$$

onde TP (*True Positive*) é o número de pixels pertencentes ao *foreground* corretamente segmentados, FP (*False Positive*) é o número de pixels do *background* que são classificados de forma errada como *foreground*, FN (*False Negative*) é o número de pixels do *foreground* que são classificados como *background* e TN (*True Negative*) é o número de pixels do *background* que são corretamente segmentados. A Tabela 2 mostra as medidas TP , FP , FN e TN na matriz de confusão entre a segmentação obtida e o *ground-truth*.

Tabela 2: Matriz de confusão entre a segmentação obtida e o *ground-truth*.

		<i>Ground-truth</i>	
		<i>Foreground</i>	<i>Background</i>
Segmentação obtida	<i>Foreground</i>	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
	<i>Background</i>	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

3.1.2 Sensibilidade e especificidade

A sensibilidade (SEN) diz respeito à taxa de verdadeiros positivos, enquanto que a especificidade (ESP) mede a proporção de pixels do *background* corretamente classificados. As medidas SEN e ESP são definidas como segue:

$$SEN = \frac{\text{Número de pixels do } foreground \text{ corretamente classificados}}{\text{Número total de pixels do } foreground} = \frac{TP}{TP + FN}, \quad (3.2)$$

$$ESP = \frac{\text{Número de pixels do } background \text{ corretamente segmentados}}{\text{Número total de pixels do } background} = \frac{TN}{TN + FP}. \quad (3.3)$$

O valor máximo para as duas medidas é 1.

3.1.3 Área sob a curva ROC

De acordo com Fawcett (2006), a curva ROC (*Receiver Operating Characteristic*) relaciona a quantidade de falsos positivos e verdadeiros positivos, onde o percentual varia de acordo com o ponto de corte aplicado. Para se construir o gráfico ROC, plota-se no

eixo das ordenadas (eixo y) os valores verdadeiros positivos e no eixo das abscissas os valores falsos positivos considerando vários níveis de um limiar de classificação. A Figura 10 ilustra um gráfico para a curva ROC. Os pontos extremos (0,0) e (1,1) representam que um determinado método classificará todos os pixels ou como negativos ou como positivos, respectivamente. Já o ponto (0,1) do gráfico representa o algoritmo de segmentação ideal, enquanto que o ponto (1,0) informa que o algoritmo irá classificar todos os pixels de forma errada.

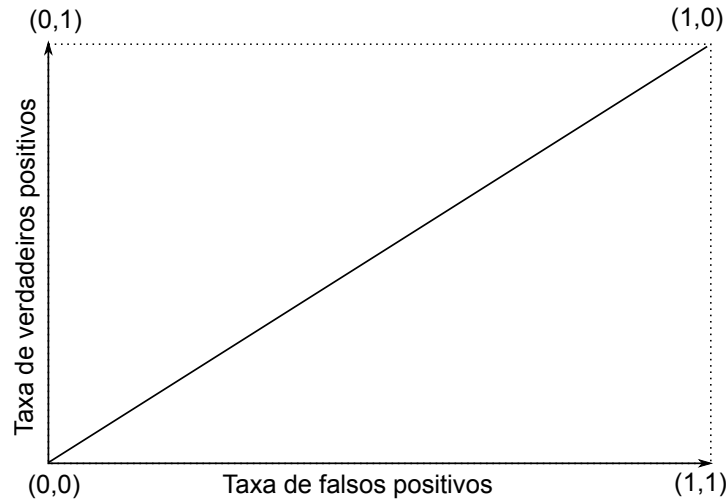


Figura 10: Gráfico para a curva ROC.

A linha diagonal do gráfico ROC descreve um algoritmo com comportamento aleatório. Portanto, pontos pertencentes à parte superior representam algoritmos melhores que o aleatório, já os pontos pertencentes à parte inferior significam que os algoritmos são piores que o aleatório.

A partir da curva ROC pode-se extrair uma medida de desempenho denominada área sob a curva ROC (*Area Under ROC* - *AUC*). A área sob a curva ROC é uma importante e largamente difundida medida para avaliação de problemas de classificação binária. A AUC de um determinado método é caracterizada pela área total da parte inferior à sua curva ROC. O valor ideal da área sob a curva ROC é 1. Um valor de AUC inferior a 0.5 informa que o método é inferior à segmentação aleatória.

3.1.4 Coeficiente de correlação de Matthews

O coeficiente de correlação de Matthews (*Matthews Correlation Coefficient* - *MCC*) mede a qualidade de uma classificação binária (BALDI et al., 2000) e é recomendada quando as amostras das duas classes são desbalanceadas, como é o caso de muitos problemas de segmentação de imagens médicas, onde o número de pixels pertencentes ao *background* é maior do que o número de pixels do *foreground*. O valor *MCC* varia entre -1 e +1, onde +1 representa uma previsão perfeita, 0 não é melhor que a previsão aleatória e -1

indica uma discordância total entre a classificação obtida e o *ground-truth*. Esta medida foi introduzida em 1975 por Matthews (MATTHEWS, 1975) e sua fórmula é definida como:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \quad (3.4)$$

3.1.5 Kappa

O coeficiente *Kappa* (COHEN, 1960) é uma medida estatística de concordância inter-segmentação e pode ser usada para estimar o grau de concordância entre a segmentação obtida por um algoritmo qualquer e o *ground-truth*. O valor de *Kappa* é baseado no número de respostas concordantes, ou seja, no número de pixels cujo resultado é o mesmo tanto na segmentação estimada quanto na segmentação do *ground-truth*. O coeficiente de *Kappa* é obtido por:

$$Kappa = \frac{P(A) - P(E)}{1 - P(E)}, \quad (3.5)$$

onde $P(A) = (TP + TN)/N$ é a proporção de concordância observada, enquanto $P(E) = (TP + FP)/N \times (TP + FN)/N + (1 - (TP + FP)/(P + N))(1 - (TP + FN)/N)$ é a proporção de concordância esperada no caso aleatório. O número total de pixels na imagem é representado por N .

O valor máximo dessa medida é 1 e isso significa total concordância. Valores próximos e até abaixo de 0 indicam nenhuma concordância, ou a concordância se assemelha à aleatória. Um valor de *Kappa* menor que zero sugere que a concordância foi menor do que a esperada por um método aleatório. Landis e Koch (1977) associam coeficientes de *Kappa* à qualidade da classificação de acordo com a Tabela 3 abaixo:

Tabela 3: Coeficiente de *Kappa* associados à qualidade da classificação.

Coeficientes de Kappa	Qualidade
< 0.00	Péssima
0.01 a 0.20	Ruim
0.21 a 0.40	Razoável
0.41 a 0.60	Boa
0.61 a 0.80	Muito boa
0.81 a 1.00	Excelente

3.1.6 Dice

O coeficiente de *Dice* (DICE, 1945) avalia a similaridade através da sobreposição espacial entre duas imagens binárias. Essa medida é definida como a média harmônica de

precisão e *recall* e é dada por:

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN}. \quad (3.6)$$

Precisão e *recall* medem a qualidade da imagem segmentada, sendo sensíveis a sobre-segmentação ou sub-segmentação. A sobre-segmentação esta relacionada ao baixo valor da precisão, enquanto a sub-segmentação diz respeito ao baixo valor do *recall*. O valor do coeficiente de *Dice* é uma quantificação normalizada, obtida pelo coeficiente entre a quantidade de pixels comuns às duas regiões e a média dos pixels em cada região. Quando nenhum pixel é comum para as duas regiões o coeficiente de *Dice* é 0. Quando todos os pixels de ambas as regiões são correspondentes o valor *Dice* é 1.

3.1.7 Índice de Jaccard

O índice de Jaccard (JACCARD, 1912) é outra medida estatística que avalia a sobreposição entre duas imagens binárias. Considere a imagem segmentada $\hat{S}$ e o *ground-truth* S , pode-se calcular o índice de Jaccard por:

$$INJ = \frac{\hat{S} \cap S}{\hat{S} \cup S}. \quad (3.7)$$

O valor de *INJ* é 0 quando as duas imagens não têm elementos em comum e 1 quando as duas imagens são idênticas. Existe uma relação entre o índice de Jaccard e o coeficiente de *Dice*, dada por:

$$Dice = \frac{2 \times INJ}{1 + INJ}. \quad (3.8)$$

3.1.8 Índice de Rand

O índice de Rand (RAND, 1971) mede a fração de pares de pixels cujos rótulos são consistentes entre a segmentação $\hat{S}$ e o *ground-truth* S . Matematicamente pode ser definido como:

$$INR = 1 - \frac{n_{11} + n_{00}}{N(N-1)/2} \quad (3.9)$$

onde n_{11} corresponde ao número de pixels que possuem o mesmo rótulo em $\hat{S}$ e S , n_{00} é o número de pixels que possuem rótulos distintos em $\hat{S}$ e S .

Esta medida tem valor 0 quando as duas segmentações não têm qualquer semelhança e valor 1 quando $\hat{S}$ e S são exatamente idênticas.

3.1.9 Índice ajustado de Rand

O valor esperado do índice de Rand para dois conjuntos aleatórios não assume um valor pré-definido (ou constante), como, por exemplo, zero. Assim, [Hubert e Arabie \(1985\)](#) propuseram um ajuste ao índice de Rand, no qual assume-se uma distribuição hipergeométrica generalizada como um modelo de aleatoriedade, isto é, as segmentações $\hat{S}$ e S são escolhidas ao acaso de tal modo que o número de pontos com qualquer rótulo específico é fixado para as duas segmentações. O índice ajustado de Rand pode ser obtido por:

$$IRA = \frac{\binom{n}{2}(n_{11} + n_{00}) - [(n_{11} + n_{10})(n_{11} + n_{01}) + (n_{01} + n_{10})(n_{10} + n_{00})]}{\binom{n}{2}^2(n_{11} + n_{00}) - [(n_{11} + n_{10})(n_{11} + n_{01}) + (n_{01} + n_{10})(n_{10} + n_{00})]} \quad (3.10)$$

onde n_{10} é o número de pixels marcados como *foreground* em S e como *background* em $\hat{S}$, n_{01} é o número de pixels classificados como *background* em S e marcados como *foreground* em $\hat{S}$.

3.1.10 Coeficiente de Correlação Intraclasse

O Coeficiente de Correlação Intraclasse (CCI) é uma medida estatística que mensura a confiabilidade entre duas segmentações. Sendo N o número de pixels na imagem, o CCI pode ser calculado como ([BLAND; ALTMAN, 1990](#)):

$$CCI = \frac{EMINTER - EMINTRA}{EMINTER + (N - 1) \cdot EMINTRA}, \quad (3.11)$$

com

$$EMINTER = \frac{2}{N - 1} \sum_i (m(i) - \mu)^2 \quad (3.12)$$

$$EMINTRA = \frac{1}{N} \sum_i (s_i - m(i))^2 + (\hat{s}_i - m(i))^2 \quad (3.13)$$

onde μ é a média das segmentações S e $\hat{S}$, e m é dado por:

$$m(i) = \frac{(s_i + \hat{s}_i)}{2}. \quad (3.14)$$

As equações $EMINTER$ e $EMINTRA$ correspondem ao erro médio quadrado inter e intra segmentação, respectivamente. O valor do CCI varia entre 0 e 1, indicando nenhuma e perfeita confiabilidade entre as segmentações, respectivamente.

3.1.11 Informação mútua

[Abdallah et al. \(2005\)](#) sugeriu usar a informação mútua como forma de medir as informações nas atribuições de classe. A informação mútua (*IFM*) é máxima quando cada estado da segmentação estimada é mapeado para um e somente um estado do *ground-truth* e se aproxima do valor de zero quando a distribuição conjunta sobre os rótulos é uniformemente aleatória. Sendo $\hat{S}$ a segmentação obtida por um determinado algoritmo, $\hat{s}_1, \hat{s}_2$ os rótulos presentes em $\hat{S}$, e S sendo o *ground-truth* com os seus possíveis rótulos s_1, s_2 , a informação mútua pode ser obtida por:

$$IFM = \sum_i \sum_j p(s_i, \hat{s}_i) \log \frac{p(s_i, \hat{s}_i)}{p(s_i)p(\hat{s}_i)}, \quad (3.15)$$

em que $s_i \in S$, $\hat{s}_i \in \hat{S}$, $p(s_i)$ e $p(\hat{s}_i)$ são probabilidades de um determinado pixel ser rotulado como s_i e $\hat{s}_i$, respectivamente.

3.2 Medidas baseadas em distâncias

Neste grupo encontram-se as seguintes medidas: Distância de Hausdorff (*HAU*), distância média (*DMED*), Distância de Mahalanobis (*MAH*), erro de consistência global (*ECG*) e variação da informação (*VINFO*).

3.2.1 Distância de Hausdorff

A distância de Hausdorff (*HAU*) foi proposta por [Beauchemin, Thomson e Edwards \(1998\)](#) para avaliar algoritmos de segmentação. Essa medida é definida como a máxima das distâncias mínimas em pixels obtidas a partir de qualquer ponto da imagem segmentada $\hat{S}$ para o *ground-truth* S e é expressa por:

$$HAU = \max(h(S, \hat{S}), h(\hat{S}, S)), \quad (3.16)$$

onde

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (3.17)$$

3.2.2 Distância média

A distância média, ou média da distância de Hausdorff, calcula a média de *HAU* para todos os pontos da imagem e é definida por:

$$DMED = \max(d(S, \hat{S}), d(\hat{S}, S)), \quad (3.18)$$

onde

$$d(A, B) = \frac{1}{N} \sum_{a \in A} \min_{b \in B} \|a - b\|. \quad (3.19)$$

3.2.3 Distância de Mahalanobis

A forma e orientação podem ser interessantes quando se tenta comparar distribuições, por exemplo, sempre que duas distribuições são centradas no mesmo ponto. A distância de Mahalanobis ([MAHALANOBIS, 1936](#)) introduz essas sutilezas utilizando como medida o inverso da matriz de covariância de uma determinada distribuição. Portanto, o quadrado da distância entre um ponto da amostra y e uma distribuição X é calculado como:

$$D^2(y, X) = (y - \bar{x})C_x^{-1}(y - \bar{x})^t, \quad (3.20)$$

onde $\bar{x}$ e C_x são a média e a matriz de covariância de X , respectivamente.

Uma vez que estas distâncias são bilineares, a distância média de uma distribuição de um conjunto de pontos $\{y_i\}_{i=1, \dots, n_y}$ é igual à distância média entre a média das amostras $\bar{y}$ e a distribuição, isto é, $\frac{1}{n_y} \sum_{i=1}^{n_y} D^2(y_i, X) = D^2(\bar{y}, X)$ ([VERGÉS-LLAHÍ; SANFELIU, 2005](#)).

Portanto, considerando a distância reversa $D^2(\bar{x}_i, Y)$ entre x_i pontos e a distribuição Y , a medida de similaridade utilizando distância de Mahalanobis pode ser calculada como:

$$MAH = \frac{1}{2}(D^2(\bar{x}_i, Y) + D^2(\bar{y}_i, X)). \quad (3.21)$$

3.2.4 Erro de consistência global

De acordo com [Martin et al. \(2001\)](#), a medida erro de consistência global calcula o grau com que uma segmentação pode ser vista como um refinamento de outra segmentação. Segmentações são consideradas consistentes, uma vez que elas podem representar a mesma imagem segmentada em diferentes escalas. O erro de consistência global (ECG) é definido como ([MARTIN et al., 2001](#)):

$$ECG = \frac{1}{N} \min \left\{ \sum_i E(S, \hat{S}, p), \sum_i E(\hat{S}, S, p) \right\}, \quad (3.22)$$

onde E é expresso por:

$$E(S_1, S_2, p_i) = \frac{|R(S_1, p_i) \setminus R(S_2, p_i)|}{R(S, p_i)}, \quad (3.23)$$

onde $R(S_j, p_i)$ é a região na segmentação j que contém o pixel i , o símbolo $\setminus$ denota a diferença de conjuntos e $|\cdot|$ representa a cardinalidade.

Nessa medida o valor 0 significa completa similaridade e o valor 1 representa a diferença máxima entre as duas segmentações. A medida *ECG* tem como vantagem o fato de ser tolerante a refinamentos na segmentação.

3.2.5 Variação da informação

A medida de variação da informação mede a soma da perda e do ganho de informação entre S e $\hat{S}$, isto é, ela mede aproximadamente o grau em que uma segmentação pode ser explicada pela outra. Ela é definida como (MEILA, 2007):

$$VINFO = H(\hat{S}) + H(S) - 2I(\hat{S}, S), \quad (3.24)$$

onde H e I correspondem a entropia e a informação mútua, respectivamente, associada entre a segmentação $\hat{S}$ e o *ground-truth* S . Usando essa medida, se as duas segmentações forem idênticas o valor de $VINFO$ será 0. A Figura 11 ilustra como é obtida a medida $VINFO$.

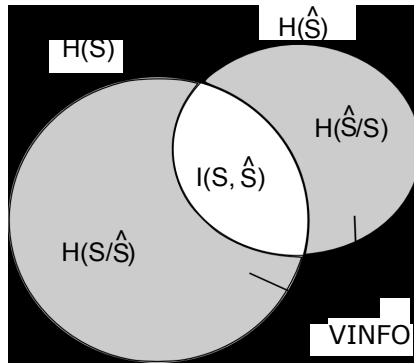


Figura 11: Composição da medida Variação da Informação ($VINFO$), representada pela soma das áreas em cinza. Na Figura, H e I correspondem a entropia e a informação mútua relacionadas a segmentação $\hat{S}$ e o *ground-truth* S .

3.3 Seleção das medidas

Nas seções anteriores foi apresentado um conjunto de medidas que podem ser utilizadas para a avaliação de métodos de segmentação baseados em *ground-truth*. No entanto, avaliar vários métodos de segmentação utilizando todas essas medidas pode ser bastante confuso. Portanto, nesta seção propomos selecionar 4 medidas que serão utilizadas para avaliar os métodos de consenso no Capítulo 7. As medidas serão selecionadas de acordo com a frequência de utilização delas na literatura e através de 2 experimentos.

As duas primeiras medidas selecionadas são a sensibilidade e a especificidade. Essas são importantes por serem capazes de informar o erro cometido por um determinado método de segmentação em cada classe. A sensibilidade diz respeito ao erro associado à classificação dos pixels do *foreground*, enquanto que a especificidade está relacionada ao erro na classificação dos pixels do *background*.

Para as outras duas medidas, selecionou-se um representante de cada grupo de medidas: similaridade e distância. Dentre as medidas baseadas em similaridade, optou-se por uma que fosse menos sensível ao problema do desbalanceamento de classes, comum em imagens médicas. Esse problema acontece quando uma classe possui muito mais elementos que a outra. Neste caso, realizou-se um experimento que consiste em avaliar as medidas de acordo com as duas classes em uma imagem binária: *foreground* e *background*. Nesse experimento, foi utilizada uma imagem de *ground-truth* contendo um círculo preenchido com raio igual a 5 pixels, como mostrado na primeira imagem da Figura 12. A partir dessa imagem, criamos outros 9 *ground-truths*, contendo o mesmo objeto, onde a única diferença entre eles é a proporção do *foreground* em relação ao *background*. No primeiro *ground-truth* da Figura 12, a proporção dos pixels do *foreground* corresponde a 25% do total, enquanto que no último *ground-truth* da Figura 12 essa proporção chega a aproximadamente 0.88% da imagem. O experimento consiste em mostrar a evolução das medidas considerando a proporção dos pixels do *foreground*. As imagens avaliadas foram criadas a partir dos *ground-truths*. Adicionou-se artificialmente um erro através da sobre-segmentação dessas imagens. Para simular a sobre-segmentação foi aplicado o operador morfológico de dilatação com o elemento estruturante de raio igual a 7 pixels em todas as imagens da Figura 12.

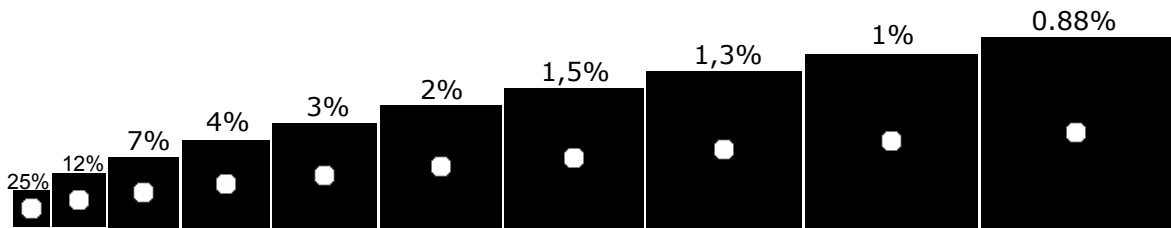


Figura 12: Sequência de imagens usadas no experimento com as medidas baseadas em distância variando a proporção dos pixels do *foreground*. A primeira e a última imagens contêm aproximadamente 25% e 0,88% de pixels do *foreground*, respectivamente.

Observando o problema do desbalanceamento de classes do experimento descrito utilizando as imagens da Figura 12, uma medida pode ser considerada boa se o perfil de curva dela em um gráfico permanece constante para diferentes proporções de pixels do *foreground* para um mesmo objeto. Como o objeto é o mesmo, espera-se que as medidas não apresentem grandes variações independentemente da quantidade de pixels do *background*. Uma medida que satisfaça esse requisito, evita que um método que segmentou todos os pixels da imagem como *background* tenha uma taxa próxima a um outro método que

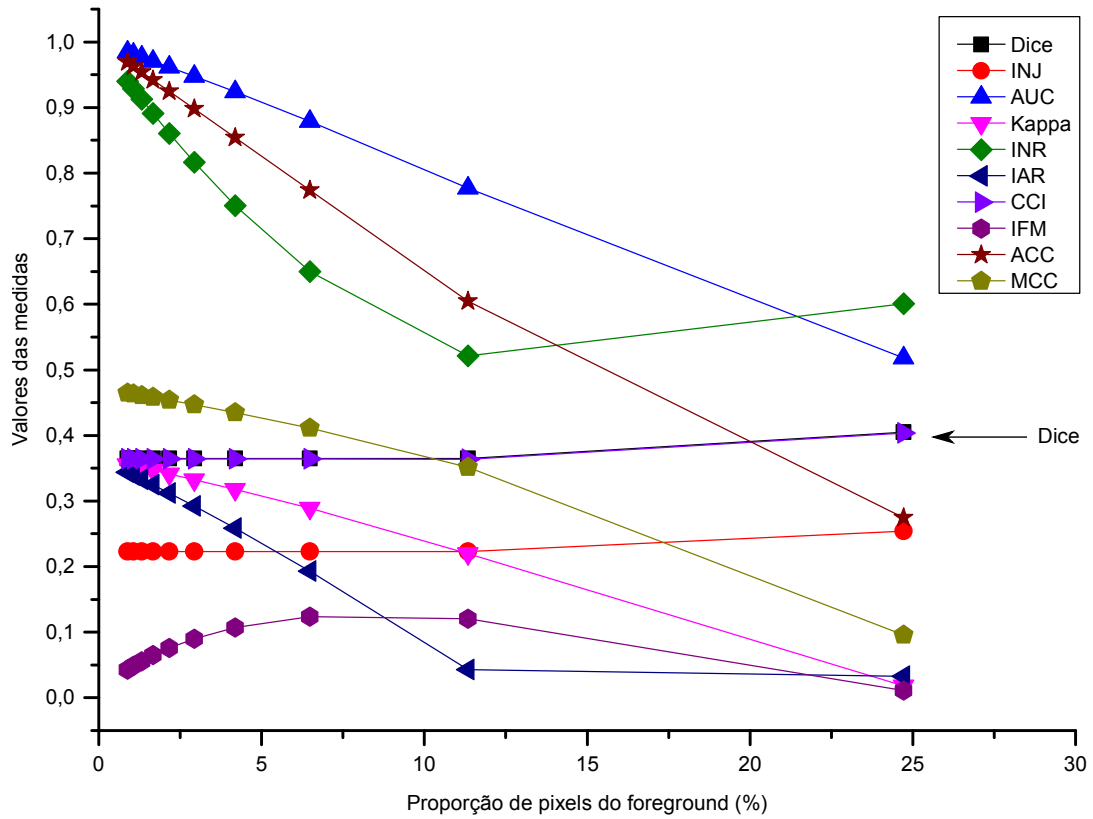


Figura 13: Valores das medidas baseadas na similaridade de acordo com a proporção dos pixels do *foreground*. As linhas das medidas Dice e CCI aparecem sobrepostas.

segmentou corretamente os pixels do *foreground*. De acordo com o gráfico da Figura 13, as medidas *Dice*, *CCI* e *INJ* são as que mais se aproximam desse critério. Dentre essas, optou-se pelo coeficiente de *Dice* por ser uma das medidas mais conhecidas e amplamente utilizadas em problemas de segmentação em imagens médicas (ZIJDENBOS et al., 1994; SABUNCU et al., 2010; YAZDANPANA et al., 2011; SUN; BAUER; BEICHEL, 2012; DODIN et al., 2010). Ainda considerando o gráfico, é importante observar que a acurácia (*ACC*) é a medida mais sensível ao problema de desbalanceamento de classes.

O experimento para selecionar a medida baseada em distância visa analisar as taxas em relação a ruídos do tipo *salt-and-pepper* presentes na imagem binária. Por exemplo, de acordo com Paumard (1997), a distância de Hausdorff é bastante sensível a esses ruídos. Nesse experimento, as imagens mostradas da Figura 14 serão utilizadas na avaliação das medidas baseadas em distância. O *ground-truth* do experimento é o círculo apresentado na Figura 14a. Espera-se que as medidas de distância mostrem que o círculo com ruído da Figura 14b é mais próximo do *ground-truth* do que o quadrado da Figura 14c. Apesar do erro provocado pelo ruído, supõe-se que é mais coerente afirmar que o círculo com ruído é propriamente um círculo de que um quadrado. A Tabela 4 apresenta os valores das medidas na comparação entre o *ground-truth* e as imagens das Figuras 14b e 14c. De acordo com essa tabela, as medidas *HAU* e *DMED* indicam que o *ground-truth* está mais

próximo do quadrado do que a imagem contendo o círculo com ruídos. Portanto, *HAU* e *DMED* não podem ser consideradas boas medidas para a avaliação de imagens binárias com ruídos. Dentre as restantes, escolheu-se a *MAH* por apresentar a maior diferença entre os valores das duas imagens.

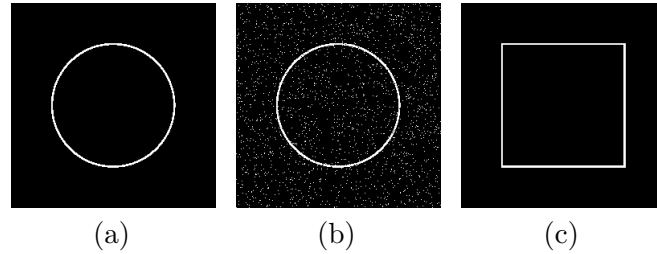


Figura 14: Três imagens de exemplo para o experimento com as medidas baseadas em distância: (a) *ground-truth* contendo um círculo; (b) imagem contendo um círculo com ruídos; (c) imagem com um quadrado.

Tabela 4: Valores das medidas na comparação entre o *ground-truth* e as imagens das Figuras 14b e 14c.

	Imagem	HAU	DMED	MAH	VINFO	ECG
Círculo com ruídos (Figura 14b)		102,53	10,82	0,03	0,25	0,04
Quadrado (Figura 14c)		31,83	8,81	0,18	0,28	0,05

3.4 Considerações finais

O objetivo desse capítulo foi apresentar um resumo das principais medidas utilizadas para avaliar métodos de segmentação. As medidas foram divididas em: medidas de similaridade e distância. De cada grupo foi selecionada uma medida que será utilizada na avaliação dos métodos de consenso de segmentações no Capítulo 7. De acordo com os experimentos e a literatura, as medidas escolhidas foram o coeficiente de Dice e a distância de Mahalanobis. As medidas de sensibilidade e especificidade também foram escolhidas por serem capazes de indicar o erro específico nas classes de *foreground* e *background*, respectivamente. Vale ressaltar que este conjunto não é uma proposta de medidas para uso geral. Sabe-se que diferentes problemas de segmentações têm características distintas e que as medidas usadas na avaliação devem levá-las em consideração no momento da avaliação.

4 SEGMENTAÇÃO COLABORATIVA

Com o avanço e proliferação da internet, sistemas colaborativos vêm ganhando bastante atenção nos últimos anos (DOAN; RAMAKRISHNAN; HALEVY, 2011). Um sistema colaborativo possibilita que vários usuários realizem atividades de forma compartilhada, geralmente a partir de locais remotos. O objetivo deste capítulo é apresentar um novo sistema colaborativo para segmentação de imagens. Segundo Bar e Maheswaran (2013), um sistema colaborativo pode ser classificado de acordo com: natureza da colaboração; tipo de problema; grau de esforço manual; e o papel do usuário. A natureza da colaboração pode ser implícita ou explícita. Na colaboração implícita os usuários resolvem um problema como um efeito secundário de alguma outra atividade que eles estão fazendo. Já na colaboração explícita, os usuários trabalham em conjunto para atingirem um objetivo em comum. O tipo do problema é qualquer um definido pelo administrador do sistema colaborativo. O grau de esforço do usuário pode ser desde muito baixo (automático) até o mais alto (manual). Um sistema colaborativo deve saber balancear entre a tarefa executada pelos usuários e a máquina. De acordo com Bar e Maheswaran (2013), o papel do usuário pode ser de 4 tipos:

- 1 Escravos: os usuários resolvem o problema através do paradigma de dividir-para-conquistar;
- 2 Fornecedores de perspectiva: os usuários contribuem com perspectivas diferentes, que quando combinadas podem produzir uma solução melhor;
- 3 Provedores de conteúdo: os usuários contribuem com conteúdos gerados por eles próprios;
- 4 Fornecedores de componentes: os usuários funcionam como componentes do problema, como uma rede social, ou simplesmente uma comunidade de usuários.

O COSE (*CO*llaborative *SE*gmentation) é um sistema web colaborativo para segmentação de imagens. Esse projeto iniciou em 2012 com o objetivo de construir uma plataforma que permitisse segmentar qualquer imagem 2D através da colaboração de vários usuários¹. O sistema alia algoritmos iterativos, interação humana e inteligência coletiva para realizar a tarefa de segmentar uma imagem. Considerando a classificação proposta por Bar e Maheswaran (2013), o COSE pode ser classificado como um sistema de colaboração explícita que utiliza fornecedores de perspectiva para segmentar iterativamente um

¹ O projeto foi desenvolvido durante o doutorado em parceria com o pesquisador Alexandre Cunha (*California Institute of Technology*) e do professor Tsang Ing Ren. Disponível em: <<http://cose.cacr.caltech.edu>>

conjunto de imagens. Em muitos problemas de segmentação de imagens não há algoritmos automáticos eficientes para a segmentação. Portanto, o objetivo do COSE é ser uma ferramenta eficaz para a segmentação de imagens na qual usuários participem do processo. A segmentação realizada por uma pessoa pode ser muito eficaz, no entanto, ela está sujeita à subjetividade inerente desse indivíduo. Com o COSE é possível segmentar a mesma imagem por vários usuários, fazendo com que o viés de subjetividade seja amenizado. Os usuários no COSE são chamados de segmentadores. No futuro, planeja-se que o COSE se apresente como um jogo e torne o processo de segmentação mais atrativo.

Analisando a literatura sobre segmentação colaborativa de imagens, pode-se destacar dois trabalhos. No primeiro, [Seung \(2016\)](#) propõe uma plataforma web, denominada *Eye Wire*, para ajudar na reconstrução 3D da rede neural humana utilizando algoritmos de processamento de imagens. O processo de segmentação é visto como um jogo onde os usuários pontuam enquanto mapeiam conexões entre neurônios. O objetivo é segmentar tridimensionalmente estas conexões. No entanto, esse sistema é voltado unicamente para o problema do mapeamento das conexões entre neurônios. Já [Han et al. \(2010\)](#) propõem um sistema colaborativo de segmentação em imagens médicas. Nesse sistema, a interação dos usuários se dá através de pontos sementes na imagem que são utilizados por um algoritmo de segmentação interativo baseado em modelos deformáveis, descrito em [Schmid et al. \(2009\)](#). Os pontos sementes dos usuários mais experientes têm peso maior no algoritmo. A desvantagem dessa abordagem é que os usuários compartilham os pontos entre si, possibilitando que pontos marcados por um usuário sejam vistos e modificados por outros usuários. O COSE diferencia desses métodos por ser uma ferramenta de propósito mais geral, escalável em termos de algoritmos de segmentação e que possibilita a segmentação de uma imagem por vários segmentadores distintos de forma independente.

Dentre as principais características do COSE, pode-se citar:

- *Framework* de algoritmos interativos: A arquitetura do COSE possui uma interface que permite adicionar métodos de segmentação interativa através do uso de *plugins*;
- Diferentes tipos de interação: Segundo [Olabarriaga e Smeulders \(2001\)](#) pode-se classificar os métodos interativos pelo tipo de interação: ferramenta de desenho (pontos, linhas, ou regiões); ajuste de parâmetros; e seleção do menu de opções através do mouse. O COSE dá suporte a esses 3 tipos de interação;
- Aplicação web: O usuário pode acessar o sistema de qualquer lugar do mundo através de um dispositivo conectado à internet. O dispositivo pode ser um computador, smartphone, tablet, entre outros, desde que possua um navegador ligado à internet;
- Resposta rápida às interações: Os algoritmos de segmentação são implementados na linguagem de programação javascript e portanto ficam no lado do cliente web. Com

isso não há necessidade de processamento no lado servidor, o que possibilita uma rápida resposta desde a interação do usuário até a visualização do resultado na tela do dispositivo;

- **Consenso de segmentações:** Utilizando algoritmos de combinação de segmentações, o COSE é capaz de obter segmentações melhores que as segmentações realizadas por cada um dos segmentadores;
- **Usabilidade:** O COSE apresenta uma interface simples e intuitiva. Um experimento realizado em uma escola secundarista dos Estados Unidos mostrou que crianças conseguem interagir com o sistema sem grandes dificuldades. A tarefa dos alunos era segmentar a imagem da sépala de uma flor utilizando o algoritmo iterativo Watershed. Nesse experimento também foi possível notar que mesmo sem serem solicitados, alguns alunos continuaram utilizando o sistema por alguns dias.

O COSE segue uma arquitetura cliente/servidor e é composto por duas grandes áreas: área administrativa e área de trabalho. A Figura 15 ilustra as principais funções dessas duas áreas.

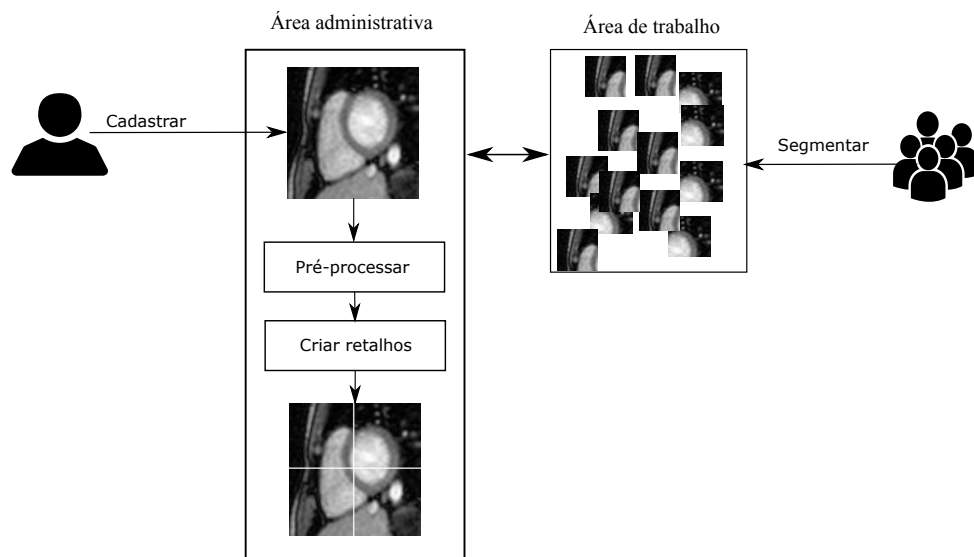


Figura 15: Ilustração das áreas administrativa e trabalho do COSE. O administrador cadastra imagens no sistema utilizando a área administrativa para que os usuários possam segmentá-las na área de trabalho.

4.1 Área administrativa

Na área administrativa, o COSE apresenta os seguintes conceitos: tarefas, problema de segmentação e retalhos. Uma tarefa é composta por um conjunto de problemas de segmentação. Exemplos de tarefas são: retina, ventrículo esquerdo, pulmão, próstata, entre outras. Já um problema de segmentação representa uma imagem a ser segmentada, por

exemplo, cada uma das N imagens de uma base de dados da retina. O menor elemento existente no COSE é o retalho. Ao segmentar imagens grandes, o usuário pode não perceber pequenos detalhes e segmentar apenas aquilo que mais chama a atenção na imagem. Com o objetivo de realçar os pequenos detalhes na imagem e tornar o processo de segmentação do usuário mais preciso, o COSE foi projetado para dividir um problema de segmentação R em um conjunto de N_R retalhos menores $R_i, i = 1, 2, \dots, N_R$ de modo que $\bigcup_{i=1}^{N_R} R_i = R$. Além disso, trabalhar com retalhos menores, melhora o desempenho tanto na transferência de imagens entre o servidor e os clientes quanto no tempo de resposta dos algoritmos interativos de segmentação. Em ambos os casos, considera-se que é mais rápido processar imagens menores do que grandes imagens. Atualmente o COSE permite apenas recortes quadrados para os problemas de segmentação. O tamanho do retalho é definido pelo administrador do sistema, assim como outros dois parâmetros: borda do retalho e sobreposição. O primeiro parâmetro serve para que o usuário tenha uma noção do contexto e consiga visualizar a imagem além dos limites de interação na imagem. Como consequência, a segmentação nas bordas se torna mais precisa. O parâmetro de sobreposição dos retalhos define uma porcentagem de sobreposição entre retalhos vizinhos. Isso evita que no momento da união dos retalhos para obter a imagem segmentada não haja descontinuidades entre retalhos adjacentes. A Figura 16 ilustra o efeito desses dois parâmetros no retalho.

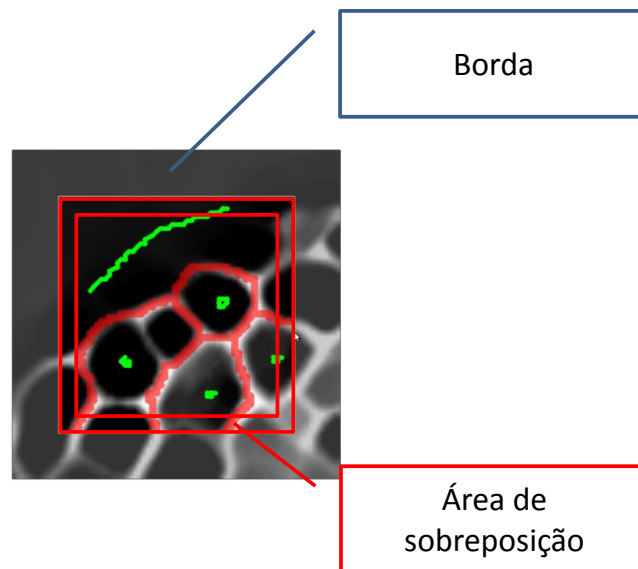


Figura 16: Exemplo da borda e da sobreposição em um retalho sendo segmentado. A parte translúcida mais externa corresponde a borda do retalho, enquanto que a área entre os quadrados vermelhos refere-se a área de sobreposição de retalhos. Os pontos e linhas verdes na imagem correspondem as interações do usuário.

Ainda na parte administrativa, pode-se determinar o número máximo de segmentações por retalho. Através desse parâmetro, o sistema controla até quando deve enviar o mesmo retalho para os segmentadores. Um retalho não deve ser segmentado pelo mesmo

segmentador mais de uma vez. Se o sistema contiver pelo menos três segmentações de um mesmo retalho, o administrador pode usar um algoritmo de consenso de segmentações para unir os resultados desse retalho. Supõe-se que quanto mais segmentações diferentes para um mesmo retalho houver, mais informações o algoritmo de consenso terá, possibilitando uma melhora no resultado. Cada sessão de um segmentador possui um determinado tempo limite para expirar, definido pelo administrador. Caso esse tempo seja excedido, a segmentação do usuário será descartada e o retalho será apresentado a outro usuário. Isso evita que um determinado retalho não alcance o número mínimo de resultados, como por exemplo, suponha que o número mínimo de resultados por retalho seja cinco e que quatro segmentadores já submeteram seus resultados. Se um quinto usuário estiver segmentando, nenhum outro usuário poderá realizar a segmentação do retalho. Se por algum motivo esse último segmentador não realizar a segmentação e nem sair do sistema, a ferramenta aguardaria a resposta desse usuário "eternamente", impossibilitando que se alcance o número mínimo de resultados para o retalho.

O processo de segmentação de imagens muitas vezes requer uma etapa de pré-processamento. O objetivo do pré-processamento é aumentar a qualidade da imagem, reduzindo a quantidade de ruídos que aparecem na imagem, enquanto mantém características importantes para a segmentação da imagem, como por exemplo, as bordas. Essa etapa é opcional no COSE e depende do nível de ruídos na imagem de entrada e da escolha de sua utilização pelo administrador do sistema. Neste trabalho foi utilizado o algoritmo de remoção de ruídos *Non-Local Means* (NLM), proposto por Buades, Coll e Morel (2005). Esse algoritmo considera que imagens naturais têm características que se repetem e que podem ser encontradas não só localmente, mas globalmente na imagem. Para remover o ruído de um pixel x , o algoritmo procura por características similares na imagem e atribui um peso para cada pixel y da imagem de acordo com a similaridade entre as vizinhanças de y e x . Portanto, a suavização do pixel x é baseada na média ponderada de todos os pixels na imagem. Dado um pixel x , o NLM pode ser expresso como:

$$NLM_v(x) = \sum_y w(x, y) v(y), \quad (4.1)$$

onde v é a imagem de referência. A função de pesos $w(x, y)$ depende da similaridade entre os pixels x e y , e satisfaz as condições $0 \leq w(x, y) \leq 1$ e $\sum_y w(x, y) = 1$. A função de pesos $w(x, y)$ é definida como:

$$w(x, y) = \frac{1}{Z(x)} \exp \left(- \frac{\|v(N_x) - v(N_y)\|_2^2}{h^2} \right), \quad (4.2)$$

onde N_i corresponde a uma vizinhança com a forma de um quadrado centrada no pixel i , h é um parâmetro que controla o decaimento dos pesos e $Z(x)$ é uma constante de

normalização, expressa por:

$$Z(x) = \sum_y \exp \left(-\frac{\|v(N_x) - v(N_y)\|_2^2}{h^2} \right). \quad (4.3)$$

O valor de h depende do nível de ruído presente na imagem. De acordo com os autores, para um ruído do tipo gaussiano com desvio padrão σ , o valor de h deve ser $\sigma * 10$. A Figura 17 mostra um exemplo da aplicação do algoritmo com $h = 7$.

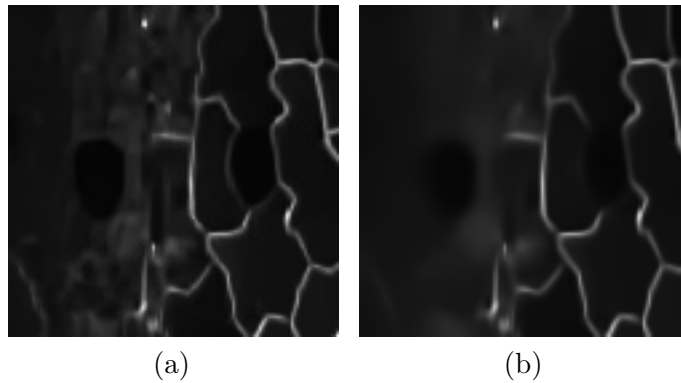


Figura 17: Exemplo de aplicação do algoritmo *Non-Local Means*: (a) imagem de um retalho da sépala de uma flor; (b) resultado da aplicação do algoritmo *Non-Local Means*.

4.2 Área de trabalho

A área de trabalho é responsável por apresentar as imagens a serem segmentadas e fornecer um conjunto de ferramentas para auxiliar os segmentadores. No COSE, o segmentador pode segmentar as imagens manualmente ou utilizando algoritmos interativos. A parte interativa é responsável por mediar informações entre o segmentador e o computador. Especificamente, ela traduz o resultado produzido pela parte computacional em resposta visual para o usuário, e da mesma forma, transforma a entrada de dados do segmentador em parâmetros para o sistema. A comunicação entre o computador e o usuário é feita através dos dispositivos de entrada e saída controlados pela interface do usuário, tais como: teclado, mouse, telas *touchscreen* e monitores. O Apêndice A apresenta a interface gráfica do COSE com os botões e suas funcionalidades.

Uma das ideias para tornar o COSE uma ferramenta mais interessante é “gamificar” o processo de segmentação. A gamificação consiste em dar características de jogo a uma aplicação com o objetivo de estimular o uso do sistema (SEABORN; FELS, 2015). Apesar de ser uma etapa inicial no processo de *gamificação*, o COSE apresenta um esquema de *score* para cada segmentador. O intuito é criar um *ranking* dos melhores segmentadores e

estimular a utilização com frequência do COSE para aumentar o *score*. O *score* de um usuário para uma dada imagem i é obtido de acordo com:

$$\text{score}_i = \alpha_i \times ACC_i, \quad (4.4)$$

onde α_i é o número de pixels segmentados como *foreground* pelo usuário e ACC_i é a acurácia em relação ao consenso das segmentações desta mesma imagem.

Atualmente, os retalhos são apresentados ao usuário de forma aleatória. Isso significa que as imagens consecutivas apresentadas ao segmentador podem ser de tarefas distintas. Por exemplo, o usuário pode começar segmentando um retalho de uma imagem da retina, o próximo retalho pode ser de uma imagem do pulmão e em seguida pode ser apresentado novamente um retalho da retina. No futuro, pretende-se criar grupos de segmentadores e associá-los às tarefas de segmentação. Portanto, para um determinado grupo só seriam apresentados os retalhos de problemas de segmentação aos quais o grupo estiver vinculado.

O Apêndice B apresenta a arquitetura do sistema COSE, enquanto que o Apêndice C mostra o fluxo de execução em telas do COSE desde o cadastro de uma tarefa até a segmentação por parte do usuário.

4.3 Segmentação interativa

O objetivo desta seção é descrever os métodos interativos já implementados e integrados ao COSE.

No COSE, uma forma simples do operador humano segmentar uma imagem é utilizando a segmentação manual. No entanto, esse procedimento pode ser extremamente laborioso e sujeito a erros devido ao cansaço. Uma alternativa menos cansativa é a utilização de métodos de segmentação interativos. Os métodos interativos visam juntar o melhor das duas abordagens: a capacidade do ser humano em reconhecer objetos e a eficiência do computador em realizar cálculos e processamentos complexos. Enquanto o operador humano auxilia no processo inserindo pontos sementes, por exemplo, o computador realiza a segmentação com base nessas informações. Durante a segmentação, o usuário tem total controle sobre o processo e é ele quem decide quando o algoritmo deve parar. A Figura 18 ilustra a interação entre o computador e o usuário.

De acordo com Mageswari, Sridevi e Mala (2013), os métodos de segmentação interativa podem ser divididos em:

- Baseados em *scribbles*;
- Baseados em borda;
- Baseados no método *Random-Walk*;

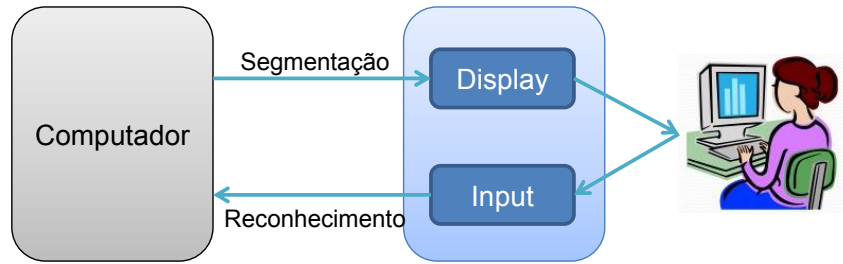


Figura 18: Ilustração da interação ente um usuário e o computador. O usuário alimenta o sistema com informações que serão utilizadas pelo algoritmo de segmentação. O computador realiza o processamento e mostra para o usuário o resultado da segmentação. Fonte: (HE; KIM; KUO, 2013)

- Baseados em regiões.

No caso particular da segmentação interativa de imagens médicas, as abordagens mais comuns podem ser agrupadas em (ZHAO; XIE, 2013):

- Baseadas em *level set*;
- Crescimento por região;
- Abordagens estatísticas;
- Baseadas em grafos.

Os métodos baseados em *level-set* trabalham com o conceito de uma superfície que se deforma de acordo com forças derivadas da imagem. De acordo com Zhao e Xie (2013), os métodos baseados em *level-set* dão a possibilidade de estimar diretamente as propriedades geométricas de estruturas anatômicas. Os métodos baseados em crescimento de região segmentam a imagem através da união de pixels vizinhos a partir de pontos sementes dado um conjunto de critérios de similaridade. Já as abordagens estatísticas envolvem um conhecimento a priori sobre a estrutura de interesse, como por exemplo a forma, textura, intensidade.

As próximas seções descrevem os métodos de segmentação interativa implementados no COSE: *Live-Wire*, *Live-Vessel* e *Watershed*. Os métodos *Live-Wire*, *Live-Vessel* e *Watershed* foram implementados por serem métodos simples de segmentação de borda, de estruturas vasculares e de regiões com transições abruptas, respectivamente. Estes métodos podem ser classificados como baseados em grafos. Nos métodos baseados em grafo, a imagem é vista como um grafo direcionado e ponderado $g = (V, B)$, onde V denota o conjunto de vértices e B denota o conjunto de arestas. Cada aresta b , onde $b \in B$, do grafo corresponde a um par ordenado em que um nó $v \in V$ no grafo está conectado a um de seus n vértices vizinhos adjacentes e distintos $u_k \in V, k = 1, 2, \dots, n$. Elementos de B

constituem vértices ponderados no intervalo $[0,1]$, ou seja, cada aresta b possui um valor que corresponde ao comprimento de distância l entre v e u_k . Um caminho P em um grafo é definido como um conjunto de ligações de arestas que conectam um nó inicial a um nó final.

4.3.1 *Live-Wire*

O *Live-Wire* é um algoritmo de segmentação interativo de imagens que utiliza a técnica de pesquisa em grafo. Mais precisamente corresponde a um algoritmo que utiliza programação dinâmica para encontrar o caminho mais curto no grafo (FALCÃO; UDUPA; MIYAZAWA, 2000). Esse caminho corresponde ao contorno do objeto de interesse.

A ideia de transformar o problema de detecção de contornos em um problema de encontrar o caminho mais curto em um grafo ponderado e orientado foi inicialmente proposta por Martelli (1972). Martelli propôs uma solução usando o algoritmo heurístico A^* e mostrou que essa estratégia global de encontrar segmentos de borda é extremamente robusta. Entretanto, na tentativa de reduzir o tempo de processamento do algoritmo A^* , variações do algoritmo de Martelli usando diferentes heurísticas foram subsequentemente propostas na literatura (ROSENFELD; KAK, 1982). Esses métodos comprometiam a otimalidade dos resultados e normalmente não garantiam nem que dados dois pontos sobre a borda, um segmento qualquer entre esses dois pontos seria encontrado. Outras abordagens globais para detecção ótima de segmentos de borda baseadas em programação dinâmica também foram propostas. Porém, impondo algumas restrições ao formato da borda e limitando a solução do problema para alguns tipos de aplicação (FALCÃO; STOLFI; LOTUFO, 2004).

Considerando o aumento da capacidade computacional e o desenvolvimento de algoritmos rápidos para resolver o problema do caminho mais curto, a ideia de Martelli ressurge em um método interativo de segmentação denominado *Live-Wire*, proposto inicialmente por (MORTENSEN et al., 1992; UDUPA; SAMARASEKERA; BARRETT, 1992). A solução contida nesse método pode ser vista tanto como uma abordagem de programação dinâmica 2D que não impõe restrições à borda e não depende da aplicação, como uma abordagem de grafos que sempre garante uma solução ótima ao problema. O algoritmo *Live-Wire* apresentado aqui é baseado no trabalho realizado por Mortensen e Barrett (1995).

O algoritmo requer inicialização manual. Dependendo da especificação, qualquer ponto caracterizado pelo usuário é refinado automaticamente. O resultado é mostrado em tempo real ao usuário. O *Live-Wire* pode ser dividido nas seguintes etapas: interação do usuário, cálculo do mapa de custo e pesquisa em grafo. A Figura 19 ilustra o funcionamento do *Live-Wire*.

No método *Live-Wire* a imagem é vista como um grafo g direcionado e ponderado.

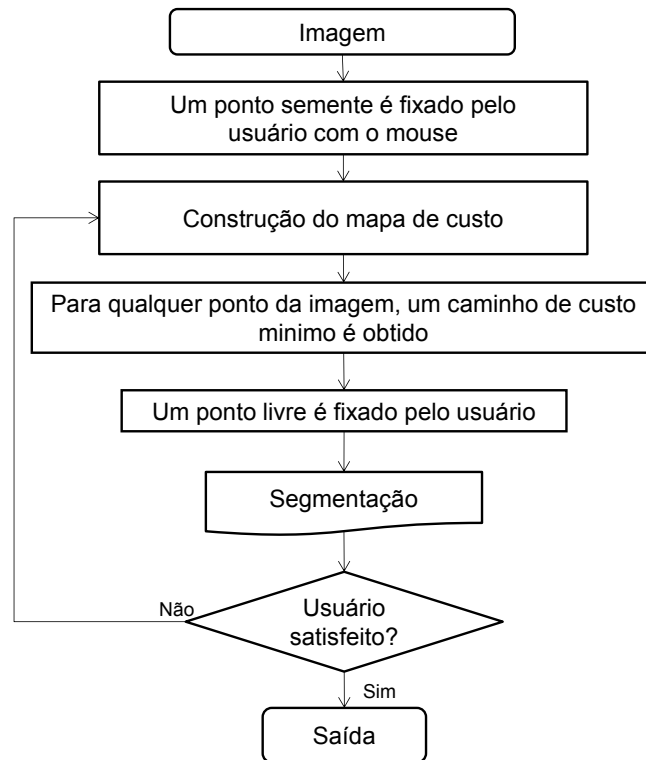


Figura 19: Ilustração do processo do Live-Wire para segmentar uma imagem.

Este método utiliza um algoritmo de menor caminho no grafo e é baseado em informações de borda, ou seja, quanto menor o custo de uma aresta b , maior a probabilidade desta pertencer à borda do objeto. Para o grafo ser criado, o usuário especifica um simples ponto na imagem. No próximo passo, caminhos ótimos conectando esse ponto a qualquer outro ponto definido pelo usuário, através do movimento do mouse, são encontrados. Um caminho P em um grafo é definido como um conjunto de ligações de arestas que conectam um nó inicial a um nó final. Cada caminho possível do grafo também possui um custo c . No problema de detecção de contorno, o contorno ótimo é associado ao caminho com menor custo, também designado de caminho ótimo. A Figura 20 mostra uma abstração do mapa de custo.

Um objeto a ser segmentado possui certos atributos que o discriminam de outros, sendo que um atributo é qualquer informação útil derivada a partir da imagem. Dessa forma, a segmentação pode ser vista como um problema de reconhecimento de padrões, onde é fundamental determinar quais atributos devem ser empregados para obter melhores resultados. Diversas funções podem ser usadas como atributos para ponderar as arestas de um grafo. A ponderação das arestas tem por objetivo associar valores no intervalo $[0,1]$ a cada aresta b . Isso representa o custo de ter b como elemento da borda desejada, sendo que uma boa ponderação associa valores de custo baixo para as arestas pertencentes à borda de interesse e valores de custo alto para as demais áreas da imagem. Esse processo, associa inicialmente um conjunto de atributos da imagem para todas as arestas. O *Live-Wire*

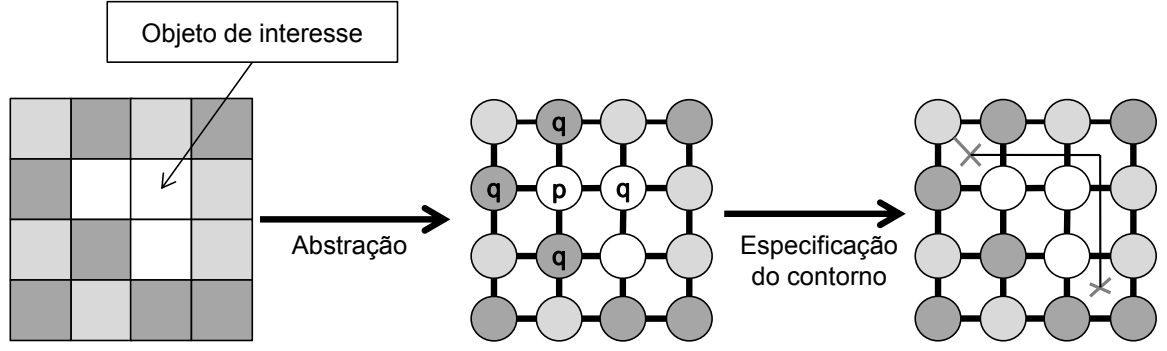


Figura 20: Construção do mapa de custo. A partir de uma imagem contendo 4×4 pixels, é criado um grafo com pesos também 4×4 , onde as arestas de menor custo são correspondentes aos pixels mais claros na imagem. O custo das arestas do pixel p é dado em função dos pixels vizinhos q . A partir da criação do mapa de custo, o usuário pode especificar o contorno desejado.

proposto por [Mortensen e Barrett \(1995\)](#) utiliza as seguintes características:

- *Zero-Crossing* laplaciano do pixel q .

$$f_z(q) = \begin{cases} 0 & \text{se } I_L(q) = 0 \\ 1 & \text{em outros casos} \end{cases} \quad (4.5)$$

- Magnitude do gradiente

$$f_G(q) = 1 - \frac{G(q)}{\max(G)} \quad (4.6)$$

- Direção do gradiente

$$f_D(p, q) = \frac{1}{\pi} \left\{ \cos[d_p(p, q)]^{-1} + \cos[d_q(p, q)]^{-1} \right\} \quad (4.7)$$

onde I_L é o laplaciano de uma imagem I e G é a magnitude do gradiente. As funções d_p e d_q são definidas por:

$$d_p(p, q) = D'(p) \cdot L(p, q), \quad (4.8)$$

$$d_q(p, q) = D'(q) \cdot L(p, q), \quad (4.9)$$

onde

$$L(p, q) = \begin{cases} q - p & \text{se } D'(p) \cdot (q - p) \geq 0 \\ p - q & \text{caso contrário} \end{cases} \quad (4.10)$$

onde $D'(p) = [I_y(p) - I_x(p)]$. As funções I_x e I_y são o gradiente da imagem nas direções x e y , respectivamente.

No final, um valor de custo conjunto é obtido para cada aresta $c(p, q)$ através da soma ponderada das características. A função de custo conjunto é definida como:

$$c(p, q) = w_z \cdot f_z(q) + w_D \cdot f_D(p, q) + w_G \cdot f_G(q), \quad (4.11)$$

onde w_i é um peso positivo indicando a ênfase dada à característica i , $i = Z, D, G$. Empiricamente, [Mortensen e Barrett \(1995\)](#), definiu os pesos como $w_Z = 0.43$, $w_D = 0.43$ e $w_G = 0.14$.

O método de pesquisa é similar ao algoritmo de Dijkstra ([CORMEN et al., 2001](#)). O algoritmo de Dijkstra é um algoritmo de programação dinâmica e como tal resolve o problema de forma otimizada. A implementação de [Falcão, Udupa e Miyazawa \(2000\)](#) do algoritmo de Dijkstra é aplicada neste trabalho. Ele utiliza uma fila de prioridade circular Q e tem complexidade computacional, no pior caso, de $O(m + nC)$, onde m é o número de arestas na imagem, n é o número de vértices na imagem e C é o custo máximo dentre todas as arestas. Este algoritmo funciona da seguinte forma:

Algoritmo 1: Algoritmo do caminho mais curto utilizado neste trabalho.

Input: Um vértice inicial v_s , e um vértice final v_e .

Data: Um array cc representando o custo atual do caminho mínimo encontrado até o momento de v_s para cada vértice, um array $dir(v)$ indicando qual é o vértice vizinho que está no caminho de custo mínimo partindo de v_s , uma fila circular Q , e uma lista L de vértices já processados.

Output: Um caminho P de custo mínimo de v_s a v_e .

```

1   $cc(v) \leftarrow \infty, cc(v_s) \leftarrow 0, Q \leftarrow v_s, N(v_s) \leftarrow \text{vizinhos adjacentes de } v$  while  $Q \neq \phi$  do
2     $v \leftarrow \min(Q), L(v) \leftarrow \text{true}$  for  $v', v' \in N(v)$  e  $!L(v')$  do
3       $tmpcc \leftarrow cc(v) + c(v')$  if  $tmpcc < cc(v')$  then
4         $cc(v') \leftarrow tmpcc$   $dir(v') \leftarrow \text{direcao de } v' \text{ para } v$  if  $v' \notin Q$  then
5           $Q \leftarrow v'$ 
6        end
7      end
8    end
9  end
10  $P \leftarrow v_e, v' \leftarrow v_e$  while  $v' \neq v_s$  do
11    $P \leftarrow v', v' \leftarrow dir(v')$ 
12 end
```

No algoritmo acima, a fila Q está sempre ordenada de forma crescente. O algoritmo é utilizado de forma iterativa pelo *Live-Wire* para segmentar bordas 2D. O resultado obtido com o Dijkstra constitui uma árvore de caminhos ótimos que conecta cada pixel da imagem ao ponto semente. Resumidamente, seleciona-se exclusivamente um caminho dentre o conjunto de todos os possíveis caminhos. Esse conjunto de caminhos é uma árvore

tendo como raiz o ponto semente.

O usuário seleciona inicialmente um vértice da borda do objeto desejado na imagem com o cursor do mouse. Para qualquer movimento subsequentemente do cursor, um caminho ótimo é obtido conectando o vértice inicial, ponto semente s_1 , ao vértice em que o cursor se encontra, ponto livre. Esse caminho é mostrado independente do usuário clicar ou não novamente na imagem, bastando apenas movimentar o cursor do mouse. O contorno do *Live-Wire* procura sempre aderir o máximo possível a borda do objeto. Quando o cursor é posicionado próximo à borda, o segmento de *Live-Wire* a adere. O resultado é mostrado em tempo real ao usuário que pode aceitá-lo e selecionar o vértice compreendido pelo cursor ou simplesmente movimentar o cursor até obter um caminho melhor. Se o usuário optar por aceitar o caminho, o vértice atual passa a ser o ponto semente s_2 e um novo segmento de borda deve ser obtido com os vértices que ainda não fazem parte do caminho segmentado. Esse processo continua até que se obtenha o contorno do objeto de acordo com os critérios do usuário. A Figura 21 ilustra uma imagem sendo segmentada pelo *Live-Wire*.

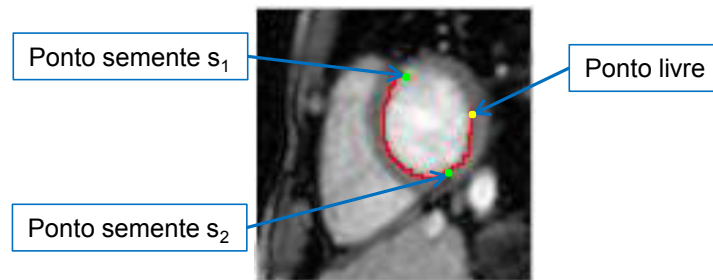


Figura 21: Imagem do ventrículo esquerdo sendo segmentada com o Live-Wire. os pontos s_1 e s_2 são definidos pelo usuário através do clique do botão esquerdo do mouse. O ponto livre é o ponto onde se encontra o cursor do mouse. A linha vermelha liga os ponto através do menor caminho no mapa de custo.

4.3.2 Watershed

O algoritmo *watershed*, originalmente proposto por Digabel e Lantuéjoul (1978), utiliza conceitos de detecção de bordas e morfologia matemática para particionar imagens em regiões homogêneas. De acordo com Bankman (2000), uma imagem no *watershed* pode ser vista como uma superfície, onde os pixels mais claros e escuros representam topos de montanhas e os os vales, respectivamente. Esta superfície é inundada através de buracos feitos nos vales. À medida que o nível da água sobe, as águas de duas bacias distintas tendem a se unir e quando isto ocorre uma barreira é construída para demarcar a linha do encontro. O processo termina quando apenas barreiras estiverem acima do nível da água.

Existem várias versões do *watershed* e uma boa revisão pode ser encontrada em Roerdink e Meijster (2000). Klava e Hirata (2009) propõem uma abordagem do *watershed* que utiliza *Image Transform Forest* (IFT) com um conjunto de marcadores. O IFT é uma

generalização do algoritmo de pesquisa Dijkstra, onde uma floresta é criada e, para cada nó no grafo, existe um caminho de custo mínimo até os marcadores. No IFT uma imagem pode ser vista como um grafo ponderado $g = (V, B, w)$, onde V é um conjunto de nós que representam os pixels da imagem e B é um conjunto de arestas ponderadas por w . A vizinhança de um nó v é representada por $N(v)$, que corresponde ao conjunto de nós adjacentes a v . Dois nós u e v são considerados adjacentes se existe uma aresta que os conecta. Sendo I o gradiente da imagem de interesse, a função de dissimilaridade w é dada por:

$$w(u, v) = \max\{I(u), I(v)\} \quad (4.12)$$

Considere $S \in V$ um conjunto de marcadores. Para um dado grafo ponderado g e um conjunto S de marcadores, o IFT retorna uma floresta F , na qual existe um caminho P de custo mínimo para cada nó $v \in V$ para um determinado marcador $s \in S$. Em outras palavras, o IFT retorna uma floresta de custo mínimo, onde cada árvore de custo mínimo está enraizada a um marcador. A função de custo é dada por:

$$f_{\max}(\pi_u \cdot \langle u, v \rangle) = \max\{f_{\max}(\pi_u), w(u, v)\}, \quad (4.13)$$

onde π_u é um caminho pré-existente até u .

Como o algoritmo do *watershed* descrito pode ter caminhos com custos iguais para diferentes marcadores em regiões de planalto, utilizou-se a função de custo $f_{lex} = (f_{\max}, f_d)$, onde f_d representa a distância em relação à borda do planalto e é responsável por fazer com que a água proveniente de diferentes marcadores flua com a mesma velocidade sobre o planalto (KLAVA; HIRATA, 2009):

$$f_d(\langle v_1, \dots, v_n \rangle) = \max_{k \in [0, n-1]} \{k, f_{\max}(\langle v_1, \dots, v_n \rangle) = f_{\max}(\langle v_1, \dots, v_{n-k} \rangle)\}, \quad (4.14)$$

onde $\langle v_1, \dots, v_n \rangle$ é um caminho ligando o vértice v_1 até o vértice v_n .

$$f_{lex}(\pi \cdot (v_{n-1}, v_n)) = \begin{cases} (w(v_{n-1}, v_n), 0) & \text{se } w(v_{n-1}, v_n) > f_{\max}(\pi) \\ (f_{\max}(\pi), f_d(\pi) + 1) & \text{caso contrário} \end{cases} \quad (4.15)$$

Para segmentar os pixels com *watershed* utilizando IFT, cada marcador recebe um rótulo. Após a execução do IFT, cada pixel é associado ao rótulo do marcador cujo caminho é de menor custo. A Figura 22 apresenta o fluxograma para segmentar uma imagem utilizando o *watershed* com IFT, enquanto que a Figura 23 ilustra uma imagem da sépala de uma planta sendo segmentada utilizando o *watershed* conforme descrito acima.

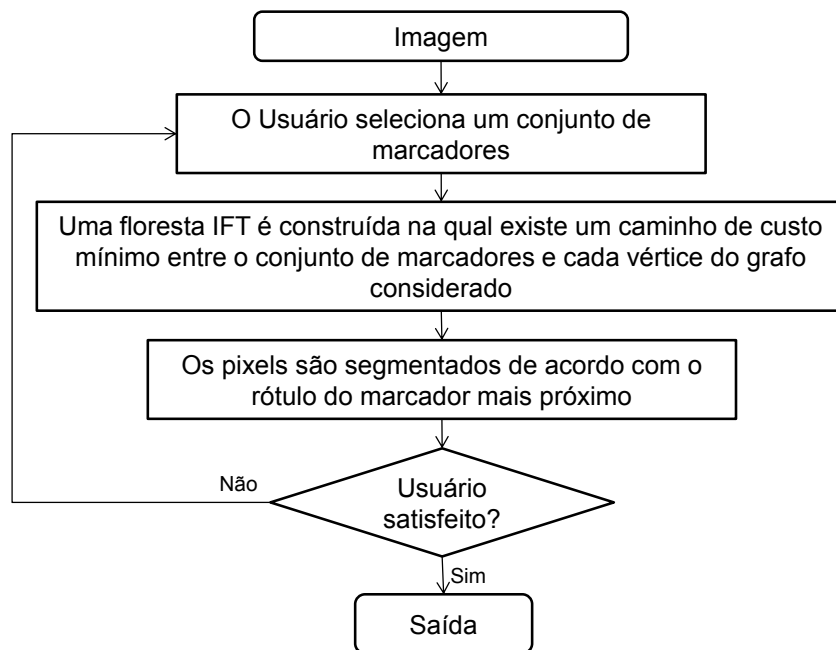


Figura 22: Ilustração do processo de segmentação utilizando o *watershed*.

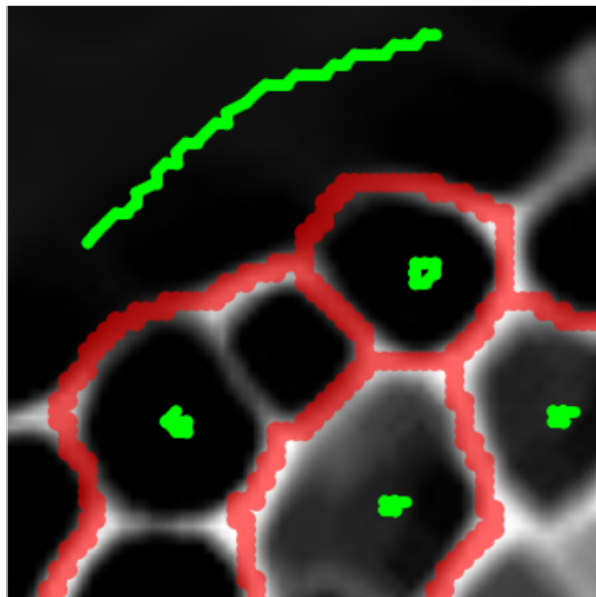


Figura 23: Aplicação do *watershed* em uma imagem das células da sépala. Os pontos verdes são os marcadores e as linha de watershed estão indicadas em vermelho.

4.3.3 Live-Vessel

O *Live-Wire* usa o algoritmo de Dijkstra para calcular globalmente o melhor caminho entre os dois pontos sementes. O *Live-Vessel* proposto por [Dickie, Hamarneh e Abugharbieh \(2009\)](#) estende o *Live-Wire* para um espaço de pesquisa 3D, onde simultaneamente encontra a linha central (x e y) e o raio do vaso (r). O algoritmo é específico para imagens que apresentam estruturas tubulares, tais como vasos sanguíneos. Considere x e y como sendo as coordenadas espaciais de um pixel p e r sendo o raio do vaso centrado em p , o custo de ir do nó p para um nó vizinho q é definido como:

$$Cost(q, p) = (w_1 C_v(q) + w_2 C_c(q) + w_3 C_e(q))(w_4 C_s(p, q) + w_5 C_r(p, q)), \quad (4.16)$$

onde w_i é o peso dado ao termo i , $i = \{C_v, C_c, C_e, C_s, C_r\}$.

A função C_v é baseada no filtro multi-escala de Frangi ([FRANGI et al., 1998](#)) e é definida como:

$$C_v(q) = \begin{cases} 1, & \text{if } \lambda_2 > 0 \\ 1 - e^{-\frac{\lambda_1/\lambda_2}{2\beta^2}} \left(1 - e^{-\frac{\sqrt{\lambda_1^2 + \lambda_2^2}}{2\alpha^2}} \right), & \text{caso contrário} \end{cases}, \quad (4.17)$$

onde o autovalor λ_k possui a k -ésima menor magnitude ($|\lambda_1| \leq |\lambda_2|$) da matriz Hessiana 2×2 em uma especificada escala σ . Os parâmetros β e α controlam a sensibilidade dos termos. A escala σ do filtro de Frangi é equivalente ao raio r dos vasos utilizado no *Live-Vessel*. Para construir a matriz Hessiana, Frangi utiliza a seguinte função gaussiana:

$$G(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{|x|^2}{2\sigma^2}}. \quad (4.18)$$

A segunda derivada dessa função, com escala σ , gera um kernel que mede o contraste das regiões dentro e fora do intervalo $(-\sigma, \sigma)$ na direção da derivada. A ideia por trás da análise de autovalores hessianos é extrair as principais orientações em que a estrutura local de segunda ordem da imagem pode ser decomposta. A matriz Hessiana para uma imagem 2D contendo informações das derivadas é definida como:

$$H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix}. \quad (4.19)$$

A função C_c é responsável por realçar estruturas alongadas através da estimação da direção dos vasos e é definida por:

$$C_c(q) = \min(Pos(R_l); Pos(R_r)), \quad (4.20)$$

onde $Pos(x) = \max(x, 0)$ e as funções R_l e R_r são respostas das bordas calculadas sobre a imagem suavizada na escala σ e são definidas por:

$$R_l(q) = \Delta f_\sigma(q + \sigma \vec{d}) \cdot \vec{d} \quad (4.21)$$

$$R_r(q) = -\Delta f_\sigma(q + \sigma \vec{d}) \cdot \vec{d}, \quad (4.22)$$

onde $\vec{d}$ é a direção obtida a partir dos autovetores da matriz hessiana.

A função C_e é uma medida utilizada para encontrar o centro de um determinado vaso com raio r e é expressa por:

$$C_e(q) = 1 - Edge(q \pm r \cdot e_1 \pm \{-\epsilon, 0, \epsilon\} \cdot e_2) \quad (4.23)$$

onde $Edge$ é o resultado da aplicação do detector de borda proposto por [Chodorowski et al. \(2005\)](#), os termos e_1 e e_2 são os autovetores da matriz Hessiana e ϵ é um parâmetro livre. Os termos de regularização espacial C_s e C_r suavizam o caminho e penalizam mudanças abruptas, respectivamente. O termo C_s é baseado na distância euclidiana entre os nós vizinhos p e q e é responsável pela suavização do caminho traçado. Já a função C_r suaviza mudanças abruptas dos raios estimados para p (r_p) e q (r_q) e é dada por:

$$C_r(p, q) = \frac{|r_q - r_p|}{r_{max} - r_{min}}, \quad (4.24)$$

onde r_{min} e r_{max} são os valores mínimo e máximo possíveis que o raio pode assumir, respectivamente.

Assim como o *Live-Wire*, na interação do *Live-Vessel*, o usuário inicialmente escolhe um ponto na imagem e para todos os pixels da imagem 3D um caminho de custo mínimo é criado. O contorno é mostrado em tempo real ao usuário, o qual pode mover o cursor do mouse e fixar o segundo ponto na imagem. O algoritmo refaz os caminhos de menor custo a partir desse novo ponto selecionado. Este processo é repetido até que os vasos estejam segmentados de acordo com o critério do usuário. A Figura 24 ilustra uma imagem sendo segmentada usando o *Live-Vessel*. No *Live-Vessel* o caminho de menor custo tenta percorrer os pixels centrais dos vasos e a espessura do contorno *live* é determinado pelo raio r , enquanto que no *Live-Wire*, a espessura do contorno é fixa e tenta sempre aderir à borda do objeto.

4.4 Considerações finais

Dentre as funções do COSE pode-se destacar: segmentação de imagens complexas e a criação de *ground-truths*. Uma imagem pode ser considerada complexa se não existirem

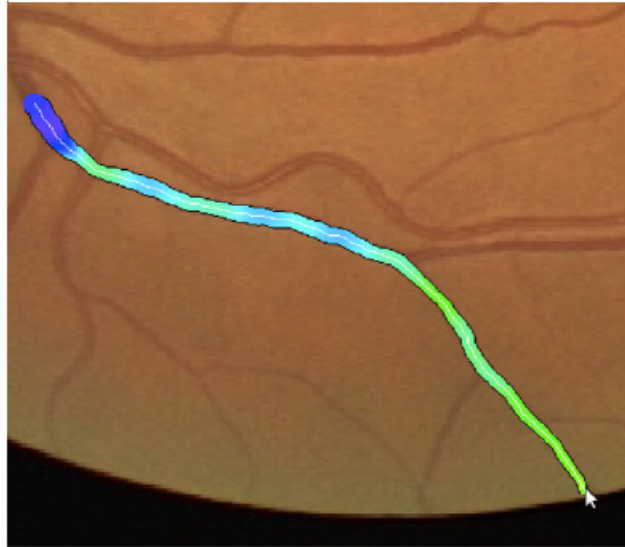


Figura 24: Exemplo de uma imagem da retina sendo segmentada utilizando o Live-Vessel. As cores representam o tamanho do raio. Quanto maior o raio mais azulada é a cor do pixel e quanto menor o raio, mais avermelhada é a cor do pixel. Fonte: [Dickie, Hamarneh e Abugharbieh \(2009\)](#).

métodos automáticos eficientes para segmentá-la. Neste caso, pode-se utilizar o COSE como uma ferramenta para segmentar de forma eficaz a imagem.

No segundo caso, o COSE serve como uma ferramenta de suporte para avaliação e aprimoramento de algoritmos de segmentação de imagens. Segundo [Zhang, Fritts e Goldman \(2008\)](#), os métodos de avaliação de segmentações podem ser classificados em dois grupos: supervisionados e não-supervisionados. A avaliação não-supervisionada permite a comparação de segmentações sem a necessidade de comparações visuais ou imagens segmentadas manualmente. Este tipo de abordagem utiliza medidas empíricas para satisfazer determinados critérios heurísticos. Alguns exemplos de medidas empíricas incluem: cor, textura e forma dos objetos. Portanto, os critérios são baseados principalmente em características comuns ou informações semânticas dos objetos, que não são suficientemente precisas para descrever os objetos em imagens complexas. Por outro lado, a avaliação supervisionada é amplamente utilizada na avaliação de imagens médicas e envolve um conhecimento a priori sobre o problema. Neste tipo de avaliação, a abordagem mais comum é calcular a diferença entre a imagem segmentada por um método de segmentação e um *ground truth*. Um dos métodos que pode ser utilizado para a criação de *ground truth* é a segmentação manual. No entanto, a segmentação manual em muitos casos é uma tarefa bastante laboriosa, repetitiva e está sujeita a subjetividade dos especialistas. O problema se agrava quando o *ground truth* é criado baseado na segmentação manual de um único especialista, como é o caso da base de imagens da retina Drive ([STAAL et al., 2004](#)).

Neste trabalho, propomos utilizar o COSE como uma ferramenta para a estimativa de *ground truth*. A vantagem de utilizar um sistema colaborativo de segmentações de

imagens como o COSE é possibilitar que vários usuários trabalhem em uma mesma imagem. Isso reduz os problemas causados pela subjetividade de um único especialista. Além disso, a segmentação de imagens utilizando métodos interativos é menos laboriosa e cansativa que a segmentação manual. No Capítulo 7 avaliamos o desempenho do COSE como uma ferramenta de estimativa de *ground-truths* em comparação às segmentações obtidas por cada segmentador.

5 CONSENSO DE SEGMENTAÇÕES

A combinação de múltiplos classificadores já demonstrou a capacidade de melhorar a precisão da classificação em muitos domínios de aplicação (KITTLER et al., 1998). Um combinador de classificadores tem a função de unir as decisões individuais de acordo com algum critério para fornecer como resultado uma decisão de consenso. A Figura 25 apresenta um esquema geral para o problema de combinação de classificadores. Dada uma entrada, os classificadores geram saídas individuais que são combinadas com base em uma determinada estratégia do combinador de classificadores. Existem várias abordagens para a combinação de classificadores, conforme descrito em Kittler et al. (1998).

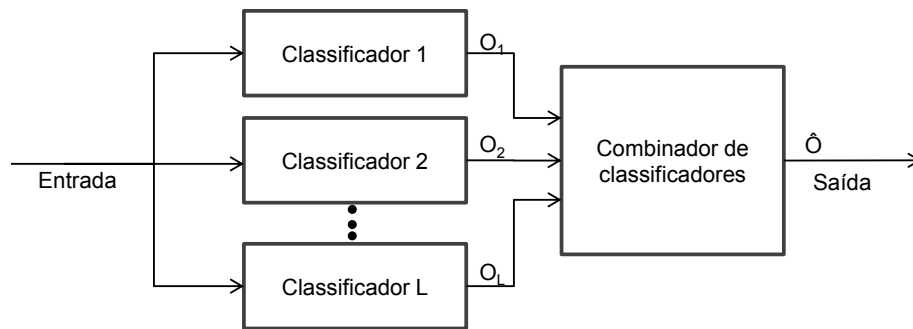


Figura 25: Estrutura de um combinador de L classificadores. Cada classificador produz uma saída O_i , $i = 1, 2, \dots, L$ que é combinada para gerar uma saída $\hat{O}$.

O consenso de segmentações pode ser visto como um problema de combinação de classificadores binários. Nesse sentido e aproveitando os conceitos de Warfield, Zou e Wells (2004) sobre combinação de classificadores, suponha um conjunto de segmentações binárias S , obtidas através de R métodos de segmentação. O propósito de um método de consenso de segmentações é obter uma imagem segmentada a partir do consenso das segmentações em S . A ideia é que essa imagem segmentada pelo algoritmo de consenso de segmentações seja melhor, de acordo com um determinado critério de avaliação, que as segmentações individuais.

Na literatura, existem alguns métodos de consenso de segmentações, dentre esses pode-se destacar: Voto majoritário, STAPLE (*Simultaneous Truth and Performance Level Estimation*), COLLATE (*C*Onsensus *L*evel, *L*abeler *A*ccuracy and *T*ruth *E*stimation) e SIMPLE (*S*elective and *I*terative *M*ethod for *P*erformance *L*evel *E*stimation). As próximas seções descrevem cada um desses métodos. Além disso, uma abordagem proposta simples para o problema do consenso de segmentações, denominada DTM (*D*istance *T*ransform *M*erge), também é descrita neste capítulo.

5.1 Voto Majoritário

Voto majoritário é uma estratégia simples de consenso, que atribui a um determinado pixel i um rótulo igual ao da maioria dos votos. Considere um conjunto de segmentações $S = S_1(i), \dots, S_R(i)$, onde R é o número de segmentadores¹, o consenso das R segmentações para o pixel i é calculado como:

$$T(i) = \begin{cases} 1, & \text{se } \hat{S} > 0.5 \\ 0, & \text{Caso contrário,} \end{cases} \quad (5.1)$$

onde

$$\hat{S}(i) = \frac{\sum_{j=1}^R S_j(i)}{R}. \quad (5.2)$$

5.2 STAPLE

O STAPLE originalmente proposto por [Warfield, Zou e Wells \(2004\)](#), utiliza um conjunto de imagens de um mesmo objeto segmentadas por segmentadores distintos para estimar a probabilidade da segmentação verdadeira e calcular o desempenho de cada segmentador usando o algoritmo *Expectation Maximization* (EM).

Seja $\mathbf{D}$ uma matriz binária $R \times N$ contendo a decisão de R segmentadores para cada um dos N pixels da imagem. Considere $\mathbf{T}$ um vetor com N elementos com a segmentação verdadeira da imagem a ser estimada. Seja $p = (p_1, p_2, \dots, p_R)^T$ e $q = (q_1, q_2, \dots, q_R)^T$ a sensibilidade e especificidade de cada segmentação candidata em R , respectivamente. A sensibilidade está relacionada a taxa dos pixels verdadeiros positivos e a especificidade se relaciona com a taxa de pixels verdadeiros negativos. Para estimar $\mathbf{T}$ é preciso maximizar a função de log-verossimilhança dos dados completos do problema $(\mathbf{D}, \mathbf{T})$, dado o conjunto de parâmetros (p, q) e a função de densidade $f(\mathbf{D}, \mathbf{T}|p, q)$. Essa maximização pode ser descrita como:

$$(\hat{p}, \hat{q}) = \arg \max_{p, q} \log f(\mathbf{D}, \mathbf{T}|p, q). \quad (5.3)$$

Se a segmentação verdadeira $\mathbf{T}$ é conhecida, a sensibilidade e a especificidade podem ser facilmente estimadas. Sendo D_{ij} a resposta binária do segmentador $j, j = 1, \dots, R$ para um dado pixel i , p e q podem ser representadas por:

$$\begin{aligned} p_j &= Pr(D_{ij} = 1 | T_i = 1) \\ q_j &= Pr(D_{ij} = 0 | T_i = 0) \end{aligned} \quad (5.4)$$

¹ Um segmentador pode ser uma pessoa ou um algoritmo de segmentação.

Considere agora $\theta_j = (p_j, q_j), \forall j \in 1, \dots, R$ sendo os parâmetros desconhecidos do segmentador j e $\theta = [\theta_1, \theta_2, \dots, \theta_R]$ os parâmetros desconhecidos para todos os R segmentadores, a função de log-verossimilhança pode ser reescrita como:

$$\hat{\theta} = \arg \max \log f(\mathbf{D}, \mathbf{T}|\theta). \quad (5.5)$$

A Equação 5.5 pode ser resolvida iterativamente utilizando um algoritmo *Expectation-Maximization* (EM). Sendo $\theta^{(k)}$ os parâmetros de desempenho estimados na iteração t , durante o *E-step* do EM, a distribuição condicional de um determinado pixel i em $\mathbf{T}$ é calculada e referenciada por uma variável ponderada, dada por:

$$W_i^{k-1} = \frac{a_i^{k-1}}{a_i^{k-1} + b_i^{k-1}}, \quad (5.6)$$

onde

$$a_i^k = f(T_i = 1) \prod_{j:D_{ij}=1} p_j^k \prod_{j:D_{i,j}=0} (1 - p_j^k) \quad (5.7)$$

e

$$b_i^k = f(T_i = 0) \prod_{j:D_{ij}=0} q_j^k \prod_{j:D_{i,j}=1} (1 - q_j^k) \quad (5.8)$$

Durante o *M-step* do EM, os parâmetros de performance p e q para um determinado pixel i no momento k são estimados de acordo com:

$$p_j^k = \frac{\sum_{i:D_{ij}=1} W_i^{k-1}}{\sum_i W_i^{k-1}} \quad (5.9)$$

e

$$q_j^k = \frac{\sum_{i:D_{ij}=0} (1 - W_i^{k-1})}{\sum_i (1 - W_i^{k-1})} \quad (5.10)$$

Essas duas etapas são executadas até o algoritmo convergir. Segundo Warfield, Zou e Wells (2004), uma boa estratégia para inicialização dos parâmetros de sensibilidade e especificidade é inicializá-los com os mesmos valores e próximos a 1, por exemplo, selecionando $p_j^{(0)} = q_j^{(0)} = 0.9999 \forall j$. O algoritmo converge quando não há modificação no mapa de probabilidade W entre a iteração atual e a anterior. A segmentação final pode ser obtida por:

$$STAPLE_i = W_i > t_W, \quad (5.11)$$

onde t_W é um limiar. O limiar sugerido por Warfield, Zou e Wells (2004) é $t_W = 0.5$. A Figura 26 apresenta um fluxograma do funcionamento do STAPLE.

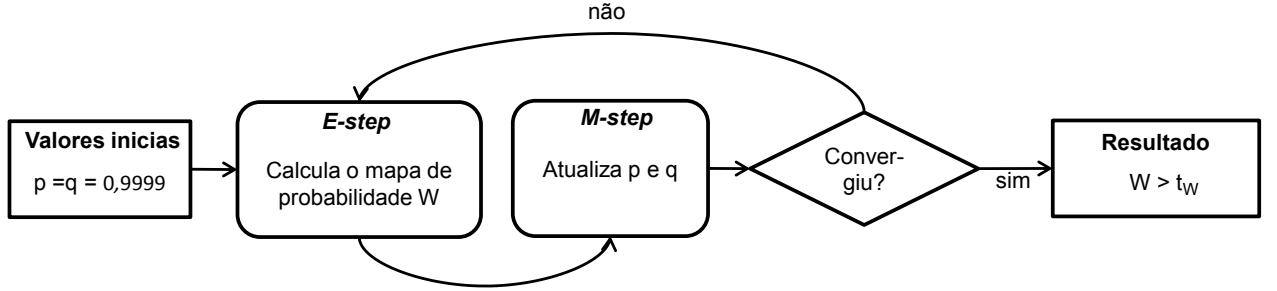


Figura 26: Fluxograma do funcionamento do STAPLE.

5.3 COLLATE

Consensus Level, Labeler Accuracy and Truth Estimation (COLLATE) proposto por Asman e Landman (2011) é uma extensão do STAPLE que considera que determinadas regiões da imagem são mais difíceis de segmentar, como por exemplo, áreas com baixo contraste, enquanto que outras regiões são intrinsecamente mais simples, como regiões homogêneas. Assim como o STAPLE, o COLLATE usa o algoritmo EM e o objetivo é maximizar a seguinte função:

$$\hat{\theta} = \arg \max_{\theta} \ln f(D, T, C | \theta), \quad (5.12)$$

onde C é um vetor que caracteriza a confusão de cada pixel em um dos R níveis de consenso, $C_i \in 0, 1, \dots, R-1$ e θ corresponde aos parâmetros desconhecidos (sensibilidade e especificidade) associada ao segmentador j , $\theta = [\theta_1, \theta_2, \dots, \theta_R]$.

No *E-step* do EM, calcula-se a seguinte função:

$$W_{si\zeta}^{k-1} = \frac{f(C_i = \zeta | T_i = s) f(T_i = s) \prod_j f(D_{ij} | T_i = s, C_i = \zeta, \theta_j^{(k-1)})}{\sum_{\zeta'} \sum_{s'} f(C_i = \zeta' | T_i = s') f(T_i = s') \prod_j f(D_{ij} | T_i = s', C_i = \zeta', \theta_j^{(k-1)})}, \quad (5.13)$$

onde $W_{si\zeta}^{k-1}$ é uma matriz de pesos que indica a probabilidade da segmentação verdadeira no pixel i ser igual ao rótulo s , com nível de confusão ζ .

No *M-step* o objetivo é maximizar os parâmetros de acordo com:

$$\theta_j^k = \arg \max_{\theta_j} \sum_{s'} \sum_{i; D_{ij}=s'} \sum_s \sum_{\zeta} W_{si\zeta}^{(k-1)} \ln((1 - p(\zeta)) I(s = s') + p(\zeta) \theta_{js's}). \quad (5.14)$$

Em seguida, maximiza-se a Equação 5.14 igualando-a a zero e usando o multiplicador de Lagrange λ com a restrição $\sum_{s'} \theta_{js's} = 1$:

$$0 = \sum_{i, D_{ij}=n'} \sum_{\zeta} W_{ni\zeta}^{(k-1)} \frac{p(\zeta)}{(1 - p(\zeta))I(n = n') + p(\zeta)\theta_{jn'n}}. \quad (5.15)$$

Essa solução pode ser dividida em dois componentes: (i) o componente fora da diagonal e (ii) o componente na diagonal. O componente fora da diagonal ($\theta_{jn,n \neq n'}$) pode ser simplificado para:

$$\theta_{jn'n, n \neq n'} = \frac{\sum_{i: D_{ij}=n'} \sum_{\zeta} W_{ni\zeta}^{k-1}}{-\lambda}. \quad (5.16)$$

Enquanto que o componente na diagonal pode ser solucionado por:

$$\theta_{jn'n, n=n'} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad (5.17)$$

onde:

$$a = \lambda \prod_{\zeta} p(\zeta) \quad (5.18)$$

$$b = \lambda \left(\sum_{\zeta} p(\zeta)^2 \right) + \left(\prod_{\zeta} p(\zeta) \right) \sum_{i: D_{ij}=n', n=n'} \sum_{\zeta} W_{ni\zeta}^{k-1} \quad (5.19)$$

$$c = \left(\sum_{i: D_{ij}=n', n=n'} \sum_{\zeta} p_{\zeta}^2 W_{ni\zeta}^{k-1} \right) + \lambda \prod_{\zeta} p(\zeta) \quad (5.20)$$

A solução para resolver λ é descrita no trabalho de [Asman e Landman \(2011\)](#).

5.4 SIMPLE

O algoritmo SIMPLE (*Selective and Iterative Method for Performance Level Estimation*), proposto por [Langerak et al. \(2010\)](#), assim como o STAPLE, usa uma estratégia iterativa baseada no desempenho das segmentações dadas como entrada em relação a segmentação verdadeira estimada. A principal diferença é que, a cada iteração, segmentações com baixo desempenho são descartadas. Sendo S^0 o conjunto inicial de segmentações obtidas por R^0 segmentadores, o algoritmo iterativo SIMPLE primeiramente computa uma estimativa de segmentação inicial V^0 utilizando o voto majoritário. Posteriormente, é calculado o valor do coeficiente de similaridade de Dice para cada segmentação S_j^0 , $j = 1, 2, \dots, R^0$ em relação a V^0 . As segmentações S_j com um valor Dice inferior ao limiar τ^0 são descartadas. Na próxima iteração, os valores de Dice são usados como pesos no voto majoritário ponderado para obter a segmentação V^1 . Esse processo continua até que

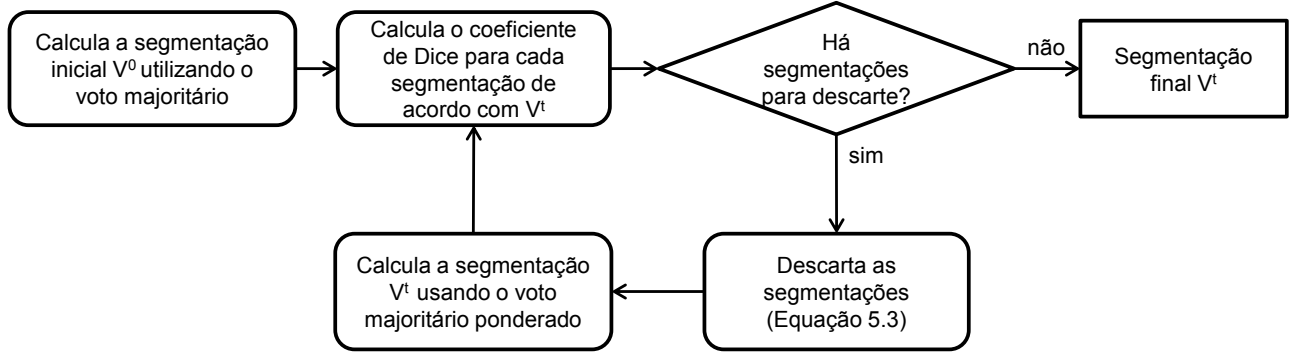


Figura 27: Fluxograma do funcionamento do SIMPLE.

não haja mais descarte de segmentações. O valor do coeficiente de similaridade Dice é calculado conforme a Equação 3.6, descrita no Capítulo 3.

Sendo S^t o conjunto das R^t segmentações que não foram descartadas na iteração $t - 1$, a decisão do voto majoritário ponderado para um pixel i é dada por:

$$V^t(i) = \begin{cases} 1, & \text{se } \hat{S}^t(i) > 0.5 \cdot W_{Dice}^t \\ 0, & \text{Caso contrário,} \end{cases} \quad (5.21)$$

onde

$$\hat{S}^t(i) = \frac{\sum_{j=1}^{R^t} S_j^t(i)}{R^t} \quad (5.22)$$

e

$$W_{Dice}^t = \sum_{j=1}^{N^t} Dice_j \quad (5.23)$$

Em cada iteração t , o limiar τ^t para descarte de segmentações é dado por:

$$\tau^t = \sigma_{Dice}^t - \alpha \cdot \sigma_{Dice}^t \quad (5.24)$$

onde σ_{Dice}^t e σ_{Dice}^t são a média e o desvio padrão dos coeficientes de Dice das segmentações S^t . Esse limiar também depende de um parâmetro α que influencia na velocidade de convergência do algoritmo. A Figura 27 apresenta um fluxograma do funcionamento do método SIMPLE.

5.5 Distance Transform Merge

O *Distance Transform Merge* (DTM) utiliza a transformada de distância das imagens segmentadas para obter o consenso. O algoritmo tem como proposta ser rápido e foi originalmente desenvolvido e integrado ao sistema COSE. Considere um conjunto de

segmentações $S = S_1, \dots, S_R$, onde R é o número de segmentadores. O nível de consenso de um pixel i pertencer à segmentação j em S é dado por:

$$C_j(i) = \frac{1}{1 + Dmap_j(i)^\alpha}, \quad (5.25)$$

onde $\alpha > 0$ controla a sensibilidade do método e $C_j(i) \in [0, 1]$. Quanto maior o valor de α , menor a importância do mapa de distância. Nos experimentos realizados, escolheu-se $\alpha = 1$. A função $Dmap_j$ é o mapa de distância da segmentação S_i obtida através de uma transformada de distância (*Distance Transform* - DT) (MAURER et al., 2003). A transformada de distância pode ser descrita como: dado o *foreground* F de uma imagem binária e algum ponto P do *background*, a distância $Dmap(P, S_j)$ entre o ponto P e o *foreground* S_j é definida como sendo a menor distância entre P para qualquer ponto em S_j . O conceito de distância entre um ponto qualquer e uma imagem binária é ilustrado na Figura 28.

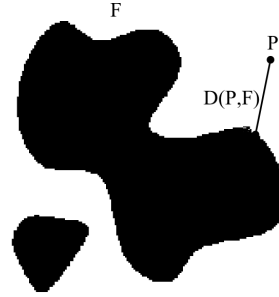


Figura 28: Distância $Dmap(P, F)$ entre um ponto P do *background* e o *foreground* F .

O mapa de distância apresenta valor 0 para pixels dentro dos objetos. À medida que se distancia do objeto, os valores de distância tornam-se maiores. Já o mapa de consenso apresenta um comportamento inverso, quanto mais próximo um pixel está do objeto, maior é o seu consenso, ou seja, se aproxima do valor máximo 1. Dado o mapa de confiança, o DTM realiza o consenso de acordo com a seguinte equação:

$$T_i = \begin{cases} 1, & \text{se } \hat{S}(i) > t \\ 0, & \text{Caso contrário,} \end{cases} \quad (5.26)$$

onde t é um limiar que pode ser obtido utilizando o método de limiarização de Otsu (OTSU, 1979) e $\hat{S}$ é dada por:

$$\hat{S}(i) = \frac{\sum_{j=1}^R C_j(i)}{R}. \quad (5.27)$$

A principal motivação para usar Otsu está na excelente relação entre velocidade e eficácia que o este algoritmo apresenta. É um método rápido que não requer a configuração de

parâmetros e é capaz de gerar resultados ótimos se as classes seguem uma distribuição gaussiana (OTSU, 1979).

A Figura 29 mostra uma abstração de como o DTM segmenta uma imagem baseada em dois segmentos: A e B. Uma média é calculada a partir dos mapas de confiança e, com base no limiar de Otsu, o consenso é realizado. Como a espessura do novo segmento pode ser maior que a espessura dos segmentos A e B, pode-se utilizar um algoritmo de esqueletização² para diminuir a espessura do novo segmento. O algoritmo de esqueletização utilizado neste trabalho é o descrito em (HARALICK; SHAPIRO, 1992).

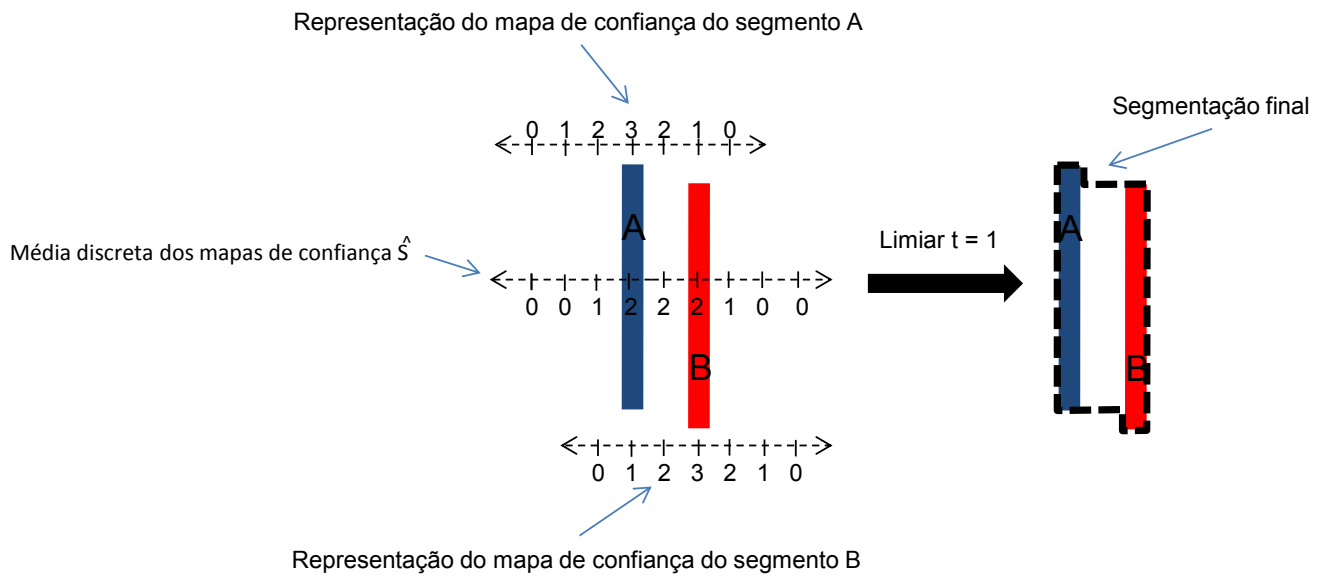


Figura 29: Abstração do funcionamento do DTM. Dados dois segmentos A e B que simulam a borda de um objeto, o DTM calcula a média do mapa de confiança dos dois segmentos. Após a aplicação do algoritmo de Otsu, o resultado do consenso é mostrado pelo contorno tracejado na figura.

5.6 Considerações finais

Este capítulo apresentou o estado da arte dos métodos de consenso para segmentações. Dentre as abordagens, as mais utilizadas são o STAPLE e o voto majoritário. O STAPLE tem a vantagem de calcular a sensibilidade e a especificidade de cada segmentação ao mesmo tempo que obtém o consenso. O algoritmo SIMPLE utiliza uma abordagem que descarta segmentações com baixo desempenho a cada iteração. Esse tipo de abordagem pode ser interessante quando as segmentações variam bastante de uma para outra e apresentam altas taxas de erros do tipo falso positivo e falso negativo. A abordagem proposta DTM foi desenvolvida juntamente com o projeto do COSE e tem a vantagem de obter o consenso com baixo tempo de processamento.

² De acordo com (GONZALEZ; WOODS, 2006), esqueletização é um processo de afinamento de numa imagem que consiste em remover todos os pixels redundantes produzindo uma nova imagem simplificada com o mínimo de pixels possível.

Todos os métodos descritos nesse capítulo ignoram informações da imagem em nível de cinza. A imagem em nível de cinza pode ser importante para extrair informações e ajudar no processo de decisão do algoritmo de consenso. Além disso, todas as abordagens utilizam apenas a informação do pixel para tomar a decisão no momento da classificação. No próximo capítulo, apresentamos uma nova abordagem de consenso que utiliza informações da imagem em nível de cinza para classificar um pixel que está próximo da fronteira de decisão entre *foreground* e *background*.

6 CONSENSO INTELIGENTE DE SEGMENTAÇÕES

As técnicas descritas no Capítulo 5 são baseadas unicamente em informações binárias extraídas das R segmentações. No entanto, uma informação bastante importante para a segmentação não é utilizada: a imagem que originou essas segmentações. Essa informação pode ser bastante útil no momento de decidir se um pixel pertence ao *foreground*, principalmente na fronteira de decisão. A utilização da imagem em nível de cinza abre outras possibilidades, como a extração de características baseadas em textura. A principal contribuição da abordagem proposta é um novo método para obter o consenso das R segmentações adicionando informações extraídas da imagem em nível de cinza. Os dados para o treinamento são aprendidos em tempo de execução e utilizados para classificar os demais dados que estão na fronteira de decisão.

A abordagem proposta considera o problema do consenso de segmentações como um problema de classificação de padrões e é denominada *Intelligent SEgmentation COnsensus* (ISECO). A proposta do método é diminuir o erro na fronteira de decisão através da análise das características dos pixels na imagem. A abordagem pode ser dividida em três etapas: pré-segmentação; extração de características; e classificação. A Figura 30 ilustra a abordagem proposta. A imagem é primeiramente segmentada por R segmentadores distintos. Em seguida, utiliza-se o voto majoritário para pré-segmentar a imagem e selecionar 3 conjuntos: F, B e C. O conjunto F contém os pixels classificados como *foreground* pelo voto majoritário, enquanto que o conjunto B contém os pixels que obtiveram nenhum voto. O conjunto C é formado pelos pixels com pelo menos um voto, mas que não obtiveram votos suficientes para serem classificados como *foreground*. Posteriormente, o método extrai características dos pixels nos 3 conjuntos para serem utilizadas por um classificador, que é treinado com base nos pixels em F e B para classificar os pixels presentes no conjunto C em *foreground* (F) ou *background* (B). A saída da abordagem proposta é a união do resultado dessa classificação com a pré-segmentação. As próximas seções descrevem com mais detalhes o ISECO.

6.1 Pré-segmentação

Como descrito no Capítulo 5, o voto majoritário é uma técnica simples, cuja combinação das segmentações é obtida através da decisão da maioria. Nesse método, muitos pixels rotulados como *foreground* pelos segmentadores são descartados. Sendo m o número mínimo de votos para um pixel ser classificado como *foreground*, um pixel que possui $m - 1$ votos é descartado. Isso levanta alguns questionamentos, tais como, será que esse pixel realmente não pertence ao *foreground*? Qual o motivo que levou pelo menos 1 segmentador a marcá-lo como *foreground*? Além disso, por que pixels com menos de $m - 1$

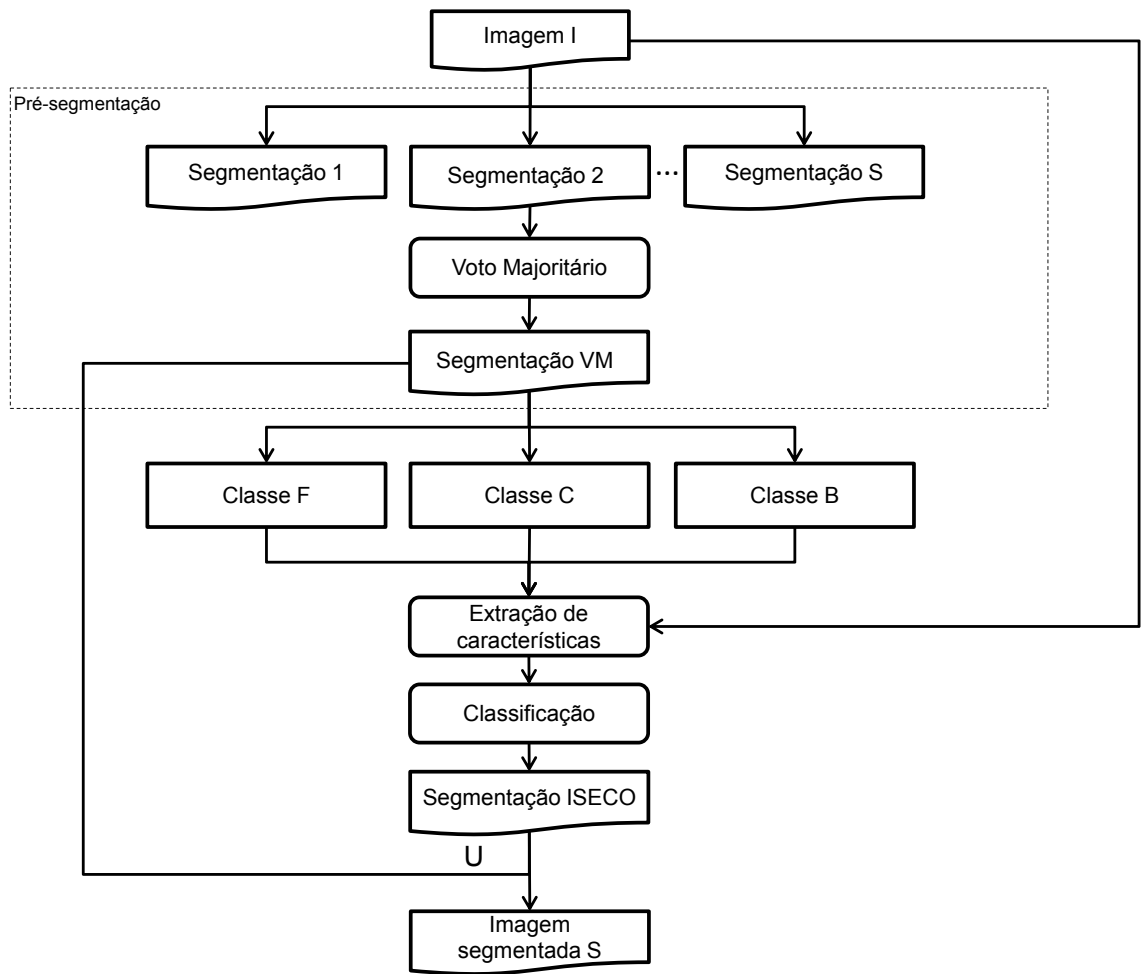


Figura 30: Fluxograma da abordagem proposta.

votos e que possuem atributos semelhantes aos pixels classificados como *foreground* pelo voto majoritário são rotulados como *background*? Neste trabalho, propomos um método inteligente de consenso que analisa os pixels de fronteira e decide se deve descartá-los ou não.

Sendo *VM* a segmentação obtida pelo voto majoritário e supondo que os pixels classificados como *foreground* estão corretos, podemos utilizar *VM* como referência para classificar os pixels confusos. Um pixel é considerado confuso se ele não foi classificado como *foreground* pelo voto majoritário e pelo menos um segmentador o classificou como *foreground*. O método visa decidir se um pixel confuso faz parte do *foreground* ou não. A Figura 31 ilustra os pixels confusos em relação a fronteira de decisão do voto majoritário. Dado um pixel p , a escolha é feita baseada nos pixels similares ao pixel p . Por exemplo, se existem vários pixels similares ao pixel p marcados como *foreground* em *VM*, tem-se um forte indício de que o pixel p pertence realmente ao *foreground* da imagem e portanto deve ser assim classificado.

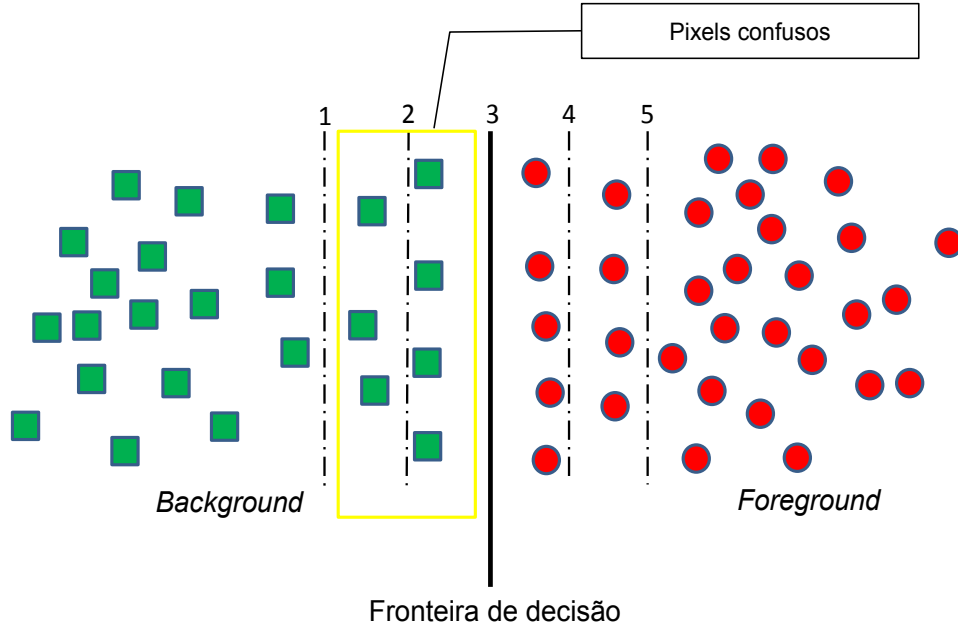


Figura 31: Ilustração dos pixels confusos em relação a fronteira de decisão para um problema com 5 votos. Os quadrados verdes e os círculos vermelhos representam os pixels classificados como *background* e *foreground* usando o voto majoritário, respectivamente.

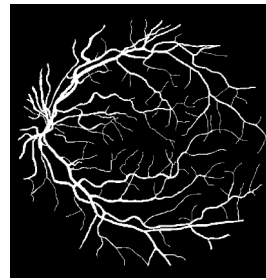
Considere I a imagem em nível de cinza e R o conjunto de segmentações obtidas a partir de I , três conjuntos são formados após a aplicação do voto majoritário: os pixels classificados como *foreground* F pelo voto majoritário, a classe formada pelos pixels confusos C e a classe dos pixels com nenhum voto B . O objetivo do ISECO é classificar os pixels em C de acordo com a similaridade dos pixels em F e B . A classe F representa o consenso da maioria dos segmentadores e por isso assume-se que os pixels dessa classe realmente pertence ao *foreground*. Analogamente, assume-se que os pixels presentes em B que não foram marcados como *foreground* por nenhum segmentador fazem realmente parte do *background*. Isso pode ser medido de acordo com as taxas de sensibilidade e especificidade. A sensibilidade diz respeito aos pixels marcados como *foreground* que são realmente *foreground*, enquanto que a especificidade afere os pixels marcados como *background* que realmente são *background*. A Tabela 5 apresenta as taxas de sensibilidade e especificidade para imagens da retina, ventrículo esquerdo e sépala segmentadas pelo consenso do voto majoritário. As imagens da retina, ventrículo esquerdo e sépala possuem 25, 8 e 5 segmentadores, respectivamente. Ambas as imagens possuem *ground-truth*. Quanto maior o valor da sensibilidade e especificidade, maior é a confiança nos conjuntos F e B , respectivamente. As imagens e os *ground-truths* utilizados para calcular a sensibilidade e especificidade para o retina, ventrículo esquerdo e a sépala são mostrados nas Figuras 32a, 32c e 32e, respectivamente.

Tabela 5: Sensibilidade e especificidade de imagens segmentadas pelo voto majoritário.

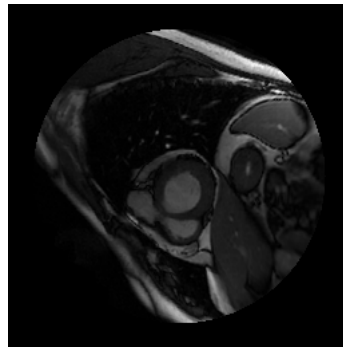
	Sensibilidade	Especificidade
Retina	0,91	0,94
Ventrículo esquerdo	0,95	0,99
Sépala	0,89	0,83



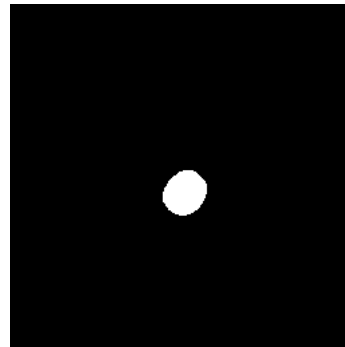
(a)



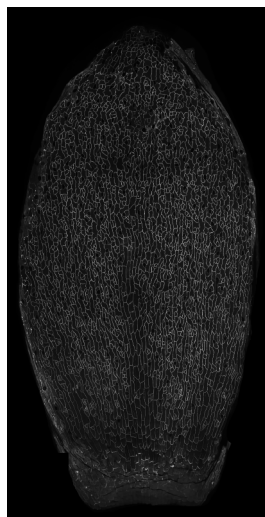
(b)



(c)



(d)



(e)



(f)

Figura 32: Exemplos de imagens médicas e biológicas: (a) exemplo de imagem da retina; (b) *ground-truth* da Figura 32a; (c) exemplo de imagem do ventrículo esquerdo; (d) *ground-truth* da Figura 32c; (e) imagem de uma sépala; (f) *ground-truth* da Figura 32e.

6.2 Extração de características

As características utilizadas pela abordagem proposta incluem informações binárias e não-binárias. As características binárias são obtidas a partir da marcação dos segmentadores. Portanto, se existem 5 segmentadores, serão 5 características binárias utilizadas pelo método. Já as características não-binárias são oriundas de métodos de extração de informações da imagem em nível de cinza. A primeira característica não-binária utilizada é o nível de cinza do pixel, enquanto que as demais são obtidas utilizando extratores de características. A extração de características pode ser entendida como um processo de captura de informações relevantes para realizar a classificação de um determinado dado. Um extrator pode gerar uma ou várias características. Existem várias abordagens para extrair características de imagens, e segundo [Choras \(2007\)](#), elas podem ser divididas em:

- características em nível de pixel - são calculadas para cada pixel da imagem;
- características locais - obtidas subdividindo a imagem;
- características globais - extração baseada na imagem inteira.

Por ser um processo de classificação pixel-a-pixel, neste trabalho estamos interessados nos extratores de características em nível de pixel. Desse grupo selecionamos 3 extratores de características bastante utilizados na literatura ([JAIN; FARROKHIA, 1991](#); [DUNN; HIGGINS; WAKELEY, 1994](#); [DUNN; HIGGINS, 1995](#); [ZHANG et al., 2010](#); [SURULIANDI; MEENA; ROSE, 2012](#); [YI; ERAMIAN, 2016](#); [COTE; ALBU, 2015](#)): filtros de Gabor, *Local Binary Pattern* (LBP) e *Local Directional Pattern* (LDP).

6.2.1 Filtros de Gabor

Os filtros de Gabor, introduzidos por Gabor em 1946 ([GABOR, 1946](#)), são grupos de *wavelets*, onde cada *wavelet* é responsável por capturar informações em uma dada frequência e direção. As funções utilizadas nos filtros de Gabor são senoides complexas e bidimensionais modeladas por uma função Gaussiana também bidimensional. Existem várias implementações dos filtros de Gabor na literatura e neste trabalho utilizamos a abordagem descrita por [Jain e Farrokhnia \(1991\)](#). Sendo x e y coordenadas espaciais, a função Gabor $2D$ pode ser descrita por:

$$g(x, y, \theta, f) = \exp\left(-\frac{1}{2}\left(\frac{\hat{x}^2}{\sigma_x^2} + \frac{\hat{y}^2}{\sigma_y^2}\right)\right) \exp(j2\pi f \hat{x}), \quad (6.1)$$

onde f representa a frequência da função sinusoidal, θ a orientação do filtro, j um número complexo ($\sqrt{-1}$), e σ_x e σ_y definem a extensão espacial da função de Gabor. Os termos $\hat{x}$ e $\hat{y}$ são definidos por:

$$\hat{x} = x \cos(\theta) + y \sin(\theta) \quad (6.2)$$

$$\hat{y} = -x \sin(\theta) + y \cos(\theta) \quad (6.3)$$

Uma das principais motivações para o uso dos filtros de Gabor é a possibilidade de extrair características de textura invariantes a iluminação, orientação, escala e translação (KAMARAINEN; KYRKI; KÄLVIÄINEN, 2006). Informações de textura podem ser muito importantes para a segmentação de imagens em níveis de cinza, como é o caso da maioria das imagens médicas. Através dela pode-se extrapolar o caráter discriminativo do nível de cinza de um pixel para um nível estrutural do objeto a ser segmentado.

6.2.2 Local Binary Pattern (LBP)

Introduzido por Ojala, Pietikäinen e Harwood (1996), o LBP tem se mostrado um poderoso descritor de texturas. Utilizando uma janela deslizante de tamanho 3×3 centralizada em um determinado pixel da imagem, os valores dos pixels da vizinhança são limiarizados pelo valor do pixel central da janela. No pixel correspondente da imagem LBP, o pixel central é codificado com o código LBP, que pode ser binário ou convertido para decimal. Sendo I uma determinada imagem em nível de cinza, i a localização espacial de um pixel dentro da janela LBP e c o pixel correspondente ao ponto central da janela, a Figura 33 ilustra um exemplo de codificação LBP.

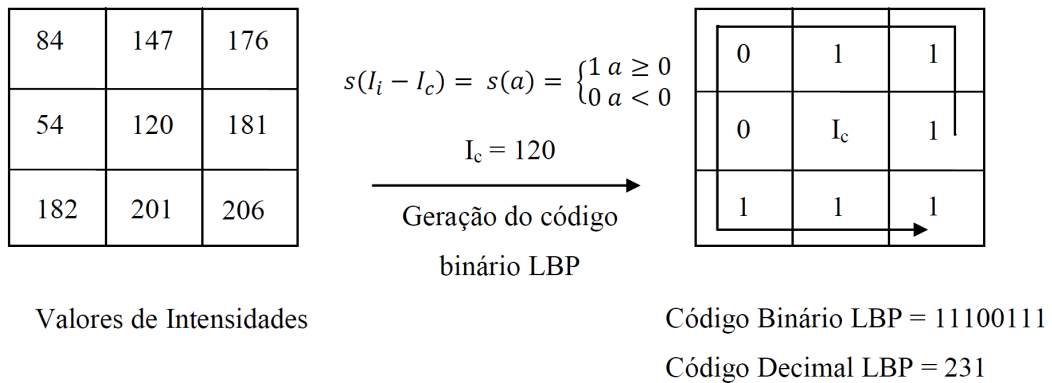


Figura 33: Codificação usando o operador LBP.

A janela de tamanho 3×3 é pequena, sendo assim, não é possível a captura de características dominantes em escalas maiores. Para solucionar este problema, Ojala, Pietikäinen e Mäenpää (2002) propôs uma extensão do operador LBP para trabalhar com

diferentes tamanhos de janelas. Desta forma ele é capaz de lidar com texturas em diferentes escalas. A vizinhança local é definida como um conjunto de pontos de amostragem que são uniformemente espaçados em um círculo centrado no pixel a ser analisado. Qualquer raio e número de pontos de amostragem são permitidos. Quando um ponto de amostragem não cair no centro do pixel, a interpolação bilinear é utilizada. A Figura 34 apresenta exemplos da extensão do LBP, onde a notação (P, R) se refere às vizinhanças do pixel e P significa pontos de amostragem em um círculo de raio R . Neste trabalho utilizamos a abordagem de Ojala, Pietikäinen e Mäenpää (2002) com o raio R das janelas variando de 1 e 5.

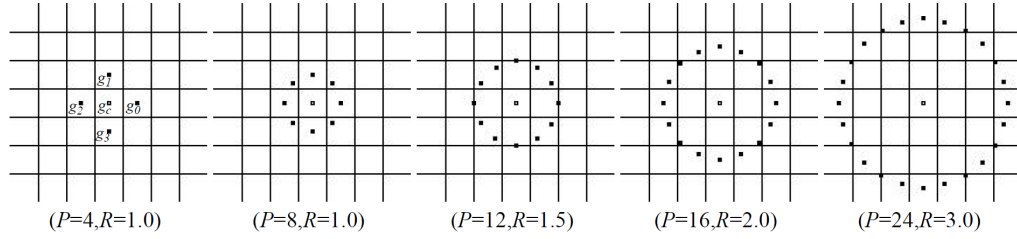


Figura 34: Exemplos da extensão do operador LBP com diferentes raios. Fonte: Ojala, Pietikäinen e Mäenpää (2002).

6.2.3 Local Directional Pattern (LDP)

O *Local Directional Pattern* (LDP) proposto por Jabid, Kabir e Chae (2010) é um código binário de oito bits atribuído a cada pixel de uma imagem de entrada. Este padrão é calculado comparando o valor de resposta da borda relativa de um pixel em diferentes direções. Para calcular os oito valores de borda são usadas as máscaras de Kirsch ($M0, \dots, M7$) descritas na Figura 35.

$$\begin{array}{cccc}
 \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} & \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \\
 M0 & M1 & M2 & M3 \\
 \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \\
 M4 & M5 & M6 & M7
 \end{array}$$

Figura 35: Máscaras de Kirsch em oitos direções. Fonte: Jabid, Kabir e Chae (2010).

Com base nos oito valores de borda $[m_0, m_1, \dots, m_7]$, verifica-se quais destes são mais significantes selecionando os Z maiores valores. Considerando um vetor de bits V de 8 posições, onde a primeira e a última posição representam as respostas obtidas pelos filtros $M7$ e $M0$, respectivamente, se o valor de borda para um dado filtro for um dos Z

maiores valores, o vetor V na posição do filtro recebe o valor 1. No final, o vetor de bits é convertido para um número decimal. A Figura 36 ilustra a geração do código LDP.

85	32	26	→	Índice	m7	m6	m5	m4	m3	m2	m1	m0
53	50	10		Valor	161	97	191	537	313	97	503	393
60	38	45		Rank	6	7	5	1	4	8	2	3
				Vetor V	0	0	0	1	0	0	1	1
				Código LDP	19							

Figura 36: Exemplo da geração do código LDP. Fonte: [Jabid, Kabir e Chae \(2010\)](#).

6.3 Classificação

Nessa etapa o objetivo é utilizar um algoritmo de reconhecimento de padrões para classificar um determinado pixel em *foreground* ou *background* de acordo com seu vetor de característica. Segundo [Jain e Farrokhnia \(1991\)](#), pode-se classificar o reconhecimento de padrões de acordo com seu modo de aprendizagem: aprendizagem supervisionada ou não supervisionada. Na aprendizagem supervisionada o padrão de entrada é rotulado como um membro de uma classe pré-definida por um especialista, enquanto que na aprendizagem não supervisionada, o padrão é associado a uma classe que é aprendida com base na similaridade entre os padrões de treinamento. O ISECO baseia-se em algoritmos de reconhecimento de padrões que utilizam aprendizagem supervisionada. No entanto, ao contrário do que é feito tradicionalmente, o conjunto de treinamento é criado de modo *on-line*, ou seja, durante a execução do algoritmo, os pixels da classe do *foreground* e *background* são selecionados logo após a pré-segmentação de acordo com o algoritmo do voto majoritário.

A abordagem dos K-vizinhos mais próximos ou KNN (*K-Nearest Neighbors*) introduzida por [Fix e Hodges \(1951\)](#) é um dos métodos de classificação de padrões mais simples existentes na literatura. Uma das vantagens do KNN é a robustez em relação a classes desbalanceadas, como é o caso da maioria dos problemas de segmentação de imagens médicas, onde o número de pixels do *foreground* é inferior ao número de pixels do *background*. A Figura 32 apresenta 3 problemas de segmentação: retina, ventrículo esquerdo e sépala. Considerando o *ground-truth*, a proporção de pixels do *foreground* é de 8.9%, 2.75% e 3.5% para as imagens da retina, ventrículo esquerdo e sépala, respectivamente.

Sendo o KNN o algoritmo utilizado neste trabalho para auxiliar no processo de classificação dos pixels confusos, suponha um conjunto X composto por N amostras de treinamento. Cada elemento X_i , $i = 1, 2, \dots, N$, é formado por um vetor de características $(x_{i1}, x_{i2}, \dots, x_{in})$ e uma classe c a qual pertence. O vetor de características X_i pode ser interpretado como um ponto num espaço n -dimensional. O tamanho do vetor de características n é definido no próximo capítulo através de experimentos. No KNN, um novo individuo Y é classificado, primeiramente, calculando-se as distâncias entre os

vetores de características de Y e todos os indivíduos de X . Seleciona-se as k amostras de treinamento mais similares a Y e dentre estes, verifica-se qual a classe que aparece com mais frequência. A amostra Y será classificada como pertencente a esta classe mais frequente. A similaridade entre as amostras pode ser calculada de acordo com a distância euclidiana:

$$\omega(X_i, Y) = \sqrt{\sum_{j=1}^n (X_{ij} - Y_j)^2}, \quad (6.4)$$

onde n é o número de características.

Uma das grandes desvantagens do KNN é a complexidade computacional para realizar uma busca, que é igual a $O(n)$. No entanto, os K-vizinhos mais próximos podem ser localizados de forma mais eficiente quando o conjunto de treinamento é representado através de uma árvore binária, denominada *KD-tree*. O *KD-tree* é uma estrutura de dados para armazenar um conjunto finito de pontos em um espaço n -dimensional, originalmente proposta por Bentley (1975). O nó raiz representa a totalidade do espaço, e os nós folhas representam subespaços contendo pequenos subconjuntos mutuamente exclusivos de X . Em qualquer nó, apenas uma das dimensões é usada como um discriminador para particionar o espaço. Normalmente, utiliza-se um limiar t_r , que pode ser a média ou a variância dos valores nessa dimensão, para particionar um determinado conjunto em dois sub-conjuntos: um superior que contém os indivíduos cujo valor nessa dimensão é superior a t_r e um inferior que contém o restante dos indivíduos. Este particionamento é feito de forma recursiva, agrupando os indivíduos em espaços cada vez menores, de acordo com a quantidade de níveis especificada na criação da árvore. A parte superior da Figura 37 exemplifica a divisão de um espaço 2-dimensional pelo *KD-tree*. No exemplo, de acordo com um critério qualquer, o espaço é primeiramente dividido na dimensão x . Em seguida, o espaço é dividido na dimensão y e x , respectivamente. Na parte inferior da Figura 37 é mostrada a representação da árvore *KD-tree* do exemplo.

Através do *KD-tree* pode-se localizar o vizinho mais próximo de um novo indivíduo Y iniciando a partir do nó raiz e selecionando recursivamente a região mais próxima que contém o novo ponto. O processo inicia pelo nó raiz e prossegue recursivamente pela árvore tentando encontrar o melhor candidato X' . Todo nó folha encontrado no caminho pode se tornar X' , desde que seja mais similar a Y do que o melhor candidato encontrado até o momento. No final do processo, tem-se a folha com o melhor candidato X' e pode-se encontrar os $k - 1$ vizinhos restantes de forma recursiva considerando os nós vizinhos e os nós ascendentes de X' .

Uma das vantagens dessa abordagem é que a construção da árvore é rápida, uma vez que o particionamento é executado apenas com base em uma dimensão dos dados, não sendo necessário calcular distâncias n -dimensionais. A complexidade da busca depende da

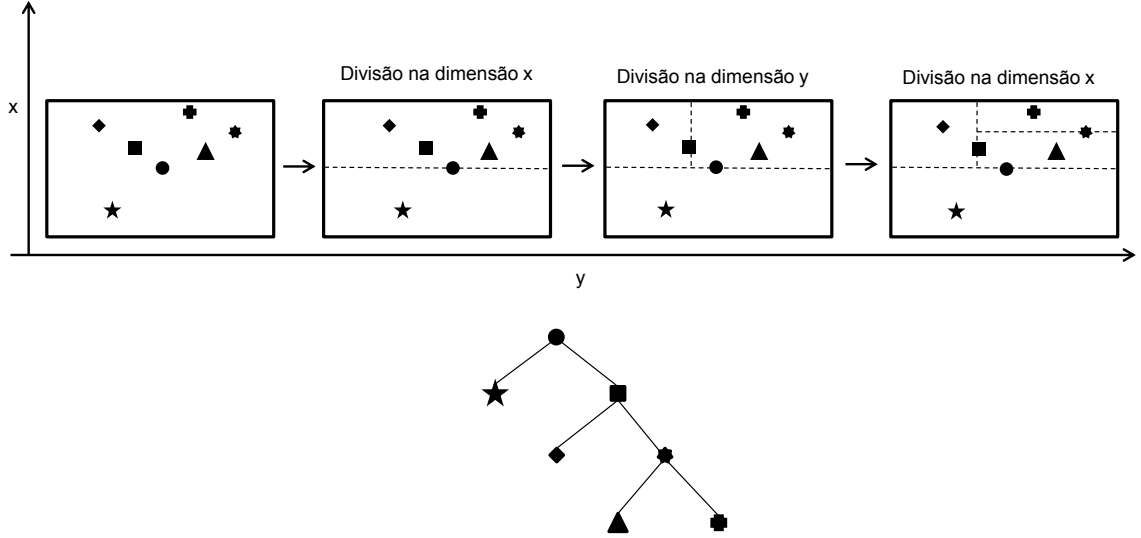


Figura 37: Exemplo da divisão de um espaço bidimensional. Na parte superior da figura é mostrado o processo de divisão do espaço baseado no *KD-tree*, enquanto que na parte inferior encontra-se a árvore formada após o processo.

profundidade da árvore. No entanto, se a árvore for balanceada, a complexidade é próxima a $\log_2 n$. No pior caso, a complexidade é n .

Neste trabalho, os pixels confusos em C são classificados em *foreground* F e *background* B com o auxílio de duas árvores *KD-tree*. A primeira árvore *KD-tree* (T_F) contém os pixels da classe F e a segunda árvore (T_B) armazena os pixels da classe B . Dado um pixel confuso p , em cada árvore é realizada uma busca e seleciona-se os K pixels mais similares a p . A similaridade é dada pela distância euclidiana do vetor de características do pixel p e o vetor de um dado pixel q presente em cada uma das árvores. Em seguida, o algoritmo computa a média das similaridades M_F e M_B para os pixels selecionados em T_F e T_B , respectivamente, ponderada pela restrição espacial:

$$M_i(p) = \frac{\sum_{q \in t_i} D_i(p, q) W_i(p, q)}{K} \quad (6.5)$$

onde t_i contém os K pixels mais similares em T_i e D_i é a similaridade entre os pixels p e q , sendo $i = F, B$. A função W_i é uma restrição espacial que indica que os pixels espacialmente vizinhos devem ter pouca ou nenhuma influência na etapa de classificação. Os pixels espacialmente vizinhos em uma mesma região apresentam características semelhantes, o que aumenta a chance de serem rotulados como *foreground* ou *background* pelo voto majoritário. Se um pixel foi rotulado erroneamente como *background* pelo voto majoritário, são grandes as chances de que os vizinhos desse pixel também estejam rotulados incorretamente como *background*. A ideia é que a classificação considere fortemente os pixels similares de outras regiões da imagem. Sendo c_x e c_y a localização espacial do pixel confuso p na imagem, a

restrição espacial em relação a um pixel q da classe F é dada por:

$$W_F(p, q) = 1 + \exp\left(-\frac{D_F(p, q)}{\lambda}\right) \quad (6.6)$$

onde $D_F(p, q) = \sqrt{(x_F - c_x)^2 + (y_F - c_y)^2}$ e (x_F, y_F) são coordenadas espaciais do pixel $q \in t_F$. O parâmetro λ controla o tamanho da restrição espacial. De forma semelhante calcula-se W_B .

Dado M_F e M_B , um pixel confuso p pode ser classificado de acordo com a seguinte equação:

$$ISECO(p) = \begin{cases} 1 & \text{se } M_F(p) \leq M_B(p); \\ 0 & \text{em outros casos.} \end{cases} \quad (6.7)$$

A segmentação final da abordagem proposta é dada por:

$$S = ISECO \cup VM. \quad (6.8)$$

A Figura 38 ilustra o resultado final da abordagem proposta.

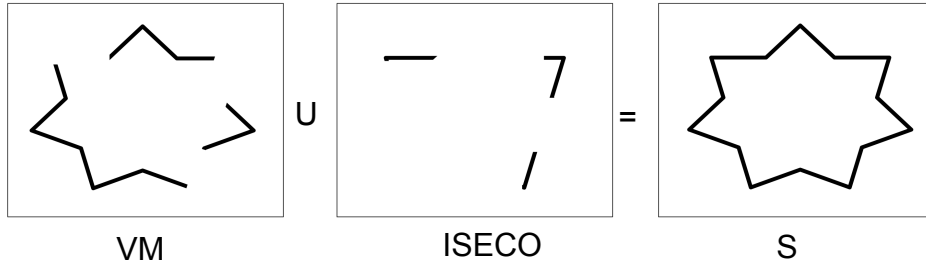


Figura 38: Ilustração do resultado final da abordagem proposta.

6.4 Outras abordagens do *Intelligent SEgmentation CO*nsensus

A abordagem desenvolvida é baseada em reconhecimento de padrões e por isso outros métodos de aprendizagem clássicas foram implementadas e testadas, tais como MLP (*Multilayer Perceptron*) e SVM (*Support Vector Machine*). Em ambos os casos, a etapa de classificação do ISECO mostrada no fluxo da Figura 30 é a única alteração do algoritmo.

De acordo com Haykin (2007), o Perceptron de Múltiplas Camadas (*Multilayer Perceptron* - MLP) consiste de um conjunto de unidades sensoriais (nós de entrada) que constituem a camada de entrada, uma ou mais camadas ocultas de nós de processamento e uma camada de saída contendo um ou mais nós. Nesse tipo de rede, o sinal de entrada se propaga através da rede para frente, da esquerda para a direita, camada por camada.

Um dos algoritmos mais difundidos para treinar uma rede MLP é o *Backpropagation*¹ (HAYKIN, 2007). No entanto, esse algoritmo não apresenta resultados satisfatórios quando comparado ao ISECO utilizando o *KNN*, como será mostrado no próximo capítulo.

Ao contrário da MLP, o ISECO usando SVM apresentou resultados próximos à abordagem descrita na seção anterior. Máquinas de Vetor de Suporte (*Support Vector Machine* – SVM) foram introduzidas por Boser, Guyon e Vapnik (1992) e tem como objetivo encontrar um hiperplano que separe perfeitamente os dados n -dimensional em duas classes. No entanto, os dados muitas vezes não são linearmente separáveis. Portanto, a SVM introduz a noção de um espaço de características criado a partir de um *kernel* que projeta os dados em um espaço de dimensão mais alta, onde é possível separá-los. A Figura 39 ilustra a divisão no espaço de características do SVM.

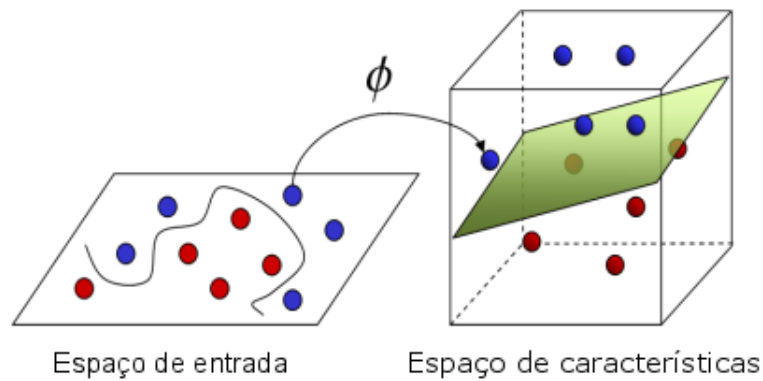


Figura 39: Divisão dos dados no espaço de características do SVM utilizando um *kernel* ϕ . Os círculos vermelhos e azuis representam dados de duas classes distintas. Fonte: (JONES, 2014).

6.5 Considerações finais

Este capítulo descreveu a abordagem proposta para o consenso de segmentações, denominada *Intelligent SEgmentation COnsensus* (ISECO). A abordagem consiste em utilizar características extraídas da imagem em nível de cinza para ajudar no processo de decisão dos pixels marcados como *foreground* por pelo menos um segmentador. O processo de decisão baseia-se nos K pixels mais similares extraídos de duas árvores *KNN*: uma para os pixels rotulados como *foreground* e outra para os pixels marcados como *background* pelo voto majoritário. Caso a distância euclidiana do pixel confuso p para o conjunto de pixels selecionados do *foreground* seja menor do que a distância em relação aos k pixels mais similares do *background*, o algoritmo classifica p como *foreground*. Como a ideia é criar

¹ O algoritmo *Backpropagation* pertence à classe dos algoritmos de aprendizagem supervisionada e baseia-se na minimização do erro no treinamento de acordo com gradiente descendente.

uma abordagem baseada em reconhecimento de padrões, outras métodos de aprendizagem clássicos foram testadas, tal como a SVM (*Support Vector Machine*), que apresentou desempenho próximo ao ISECO usando KNN. Outros métodos de aprendizagem podem ser utilizadas, como métodos de *ensemble* que combinam vários classificadores fracos, ou até mesmo redes *deep learning*. No entanto, o mais importante na abordagem é o fluxo de execução desenvolvido, que torna possível a utilização do ISECO com outros métodos de aprendizagem.

O método é baseado na reclassificação dos pixels do *background*. Intuitivamente pode-se imaginar que o mesmo processo deve ser feito para a reclassificação dos pixels do *foreground*. No entanto, partindo do pressuposto que para um especialista é mais fácil não segmentar uma determinada região do que segmentá-la, seja por cansaço ou dúvida, pixels com mais votos tendem a ser corretos, pois os especialistas fizeram um determinado esforço para segmentá-los. Além disso, nos testes feitos levando em consideração essa abordagem, o resultado foi inferior à abordagem descrita neste capítulo.

O próximo capítulo apresenta os experimentos e resultados envolvendo os parâmetros do ISECO e o desempenho desse método em comparação com os métodos descritos no Capítulo 5.

7 EXPERIMENTOS E RESULTADOS

Neste Capítulo, é apresentado um conjunto de experimentos para avaliar o desempenho da abordagem proposta *Intelligent SEgmentation COnsensus* (ISECO) na tarefa de combinar várias segmentações. Além disso, o COSE será avaliado como uma ferramenta para a estimativa de *ground-truths* capaz de diminuir o viés de subjetividade dos especialistas.

7.1 Bases de dados e medidas para a avaliação

Para avaliar o desempenho dos métodos para consenso de segmentações serão utilizadas três bases contendo imagens reais e uma base com imagens sintéticas. A primeira base real é composta por 20 imagens da retina¹. As imagens foram capturadas usando uma câmera Canon CR5 e possuem tamanho de 768 x 584 pixels. O campo de visão de cada imagem é circular com um diâmetro de aproximadamente 540 pixels. As imagens foram segmentadas por 5 algoritmos de segmentação de vasos na retina: (CHAUDHURI et al., 1989; MARTÍNEZ-PÉREZ et al., 1999; STAAL et al., 2004; ZANA; KLEIN, 2001; SOARES et al., 2006a). Para selecionar os parâmetros dos extratores e do ISECO são utilizadas 5 dessas imagens escolhidas aleatoriamente. As 15 imagens restantes são utilizadas para aferir o desempenho dos métodos de consenso.

A segunda base contém imagens do ventrículo esquerdo e é disponibilizada pelo *Sunnybrook Health Sciences Center* (RADAU et al., 2009). A base inclui imagens de 45 pacientes contendo 12 casos de insuficiência cardíaca com isquemia, 12 casos de insuficiência cardíaca sem isquemia, 12 pacientes com hipertrofia e 9 casos normais. Seis a doze imagens com tamanho 256 x 256 pixels foram obtidas a partir do anel atrioventricular até o ápice. Foram utilizados 8 segmentadores: 3 manuais e 5 algoritmos automáticos. Os métodos automáticos são descritos em Lebenberg et al. (2015).

A última base de dados reais é formada por retalhos de uma imagem da sépala. Essa imagem possui tamanho de 1608 x 3114 pixels e foi segmentada por um especialista para criar o *ground-truth*. A imagem foi recortada em 494 retalhos válidos (retalhos com pelo menos 1 pixel do *foreground*) de 124 x 124 pixels. Cada retalho foi pré-processado e segmentado por 5 a 8 segmentadores utilizando o COSE.

As bases de dados da retina e do ventrículo esquerdo foram escolhidas por possuírem um conjunto de imagens segmentadas publicamente disponíveis, enquanto que a base de dados da sépala foi selecionada por ser objeto de estudo de um dos membros do grupo de desenvolvimento do COSE.

¹ <<http://www.isi.uu.nl/Research/Databases/DRIVE/>>

A base de dados sintética é composta por 10 imagens. O objetivo nessa base é segmentar os polígonos criados artificialmente de maneira aleatória. Para tornar o problema de segmentação mais interessante, cada imagem foi corrompida com a adição de um campo randômico multi-fractal² 2D (SCHERTZER; LOVEJOY, 1989) de modo semelhante ao descrito no trabalho de Estrada e Jepson (2009). A Figura 40 ilustra a criação de uma imagem dessa base. Ao contrário do ruído *salt-and-pepper* empregado no Capítulo 3, esse campo fractal foi empregado com o objetivo de degradar as transições entre as regiões na imagem, fazendo com que as bordas das regiões fiquem mais suaves. As imagens possuem tamanho 100×100 pixels e associado a cada imagem existe um *ground-truth*. O *ground-truth* de uma imagem é o polígono criado artificialmente, como mostrado na Figura 40(a). As imagens foram retalhadas em um único retalho e não foram pré-processadas. Cada retalho foi segmentado por 5 segmentadores diferentes utilizando o COSE.

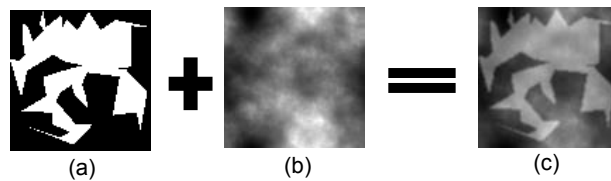


Figura 40: Exemplo da criação de uma imagem sintética: (a) polígono com o número de lados e a localização dos vértices criados de modo aleatório; (b) campo randômico multi-fractal 2D; (c) resultado da adição das duas imagens.

A Figura 41 exemplifica uma imagem de cada base de dados com seu respectivo *ground-truth*.

No Capítulo 3 selecionou-se um conjunto com 4 medidas de avaliação: Dice, distância de Mahalanobis (MAH), sensibilidade (SEN) e especificidade (ESP). Dentre essas medidas, o coeficiente de Dice caracteriza-se como a medida mais importante na avaliação, pois pode ser considerada uma medida global bastante utilizada na literatura de imagens médicas (QIU et al., 2014; GUBERN-MÉRIDA et al., 2015; YOUSEFI; KEHTARNAVAZ; GHOLIPOUR, 2012; TAMEZ-PEÑA et al., 2012) e uma das medidas menos sensíveis ao desbalanceamento de classes (Ver Tabela 4). As demais medidas servem como complemento para entender melhor o desempenho dos métodos.

7.2 Análise das características

Uma das etapas da abordagem proposta ISECO é a extração de características, como visto no Capítulo 6. Três extratores de textura a nível de pixel bastante utilizados na literatura foram detalhados: Gabor, LBP e LDP. Outra característica utilizada neste

² As imagens fractais foram criadas utilizando a função em Matlab: https://www.mathworks.com/matlabcentral/fileexchange/29181-universal-mutli-fractal-random-field-generation/content/gen_UMF2d.m

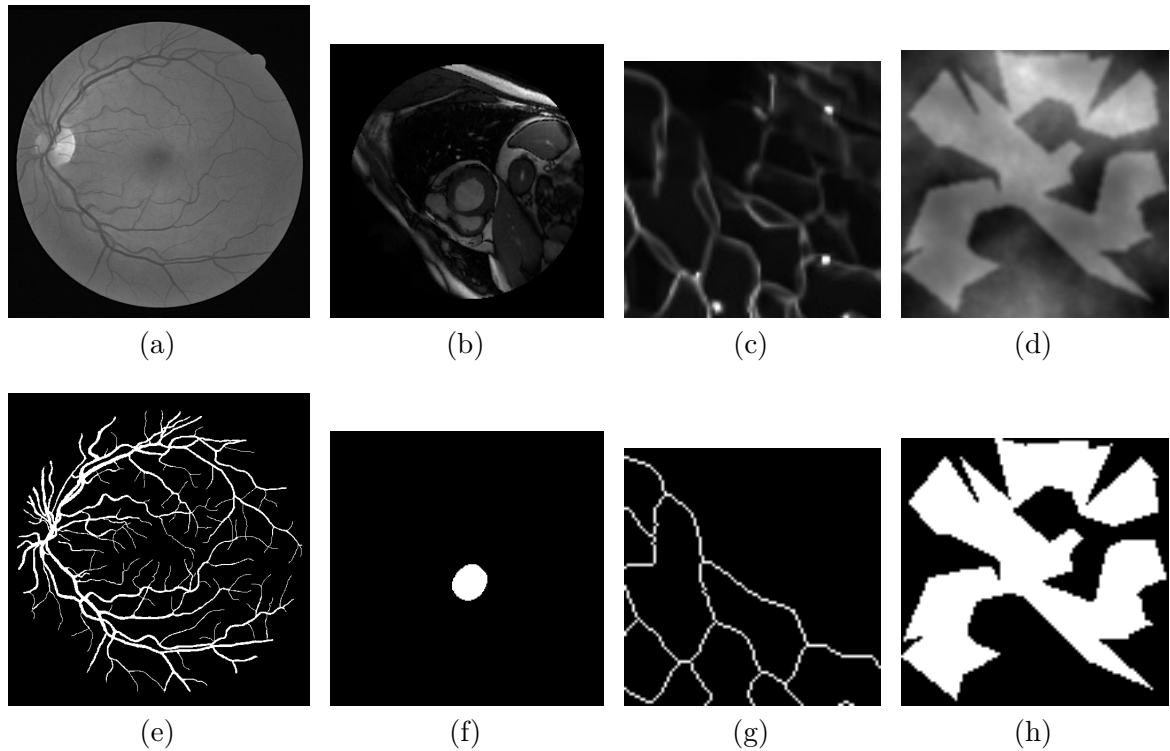


Figura 41: Exemplos de imagens das bases de dados com seus respectivos *ground-truths*: (a) imagem da retina em nível de cinza; (b) imagem do ventrículo esquerdo; (c) retalho da sépala em nível de cinza; (d) imagem da base sintética; (e) *ground-truth* da Figura 41a; (f) *ground-truth* da Figura 41b; (g) *ground-truth* da Figura 41c; (h) *ground-truth* da Figura 41d.

trabalho é o nível de cinza dos pixels na imagem. Um sub-conjunto da base de dados da retina contendo 5 imagens é utilizado para analisar o desempenho da abordagem ISECO com as características descritas. O objetivo dos próximos experimentos é analisar e definir com quais parâmetros e características o ISECO apresenta melhor desempenho para a base de dados da retina. Espera-se que os parâmetros, mesmo ajustados para essa base, possuam um alto grau de generalização, evitando assim a necessidade de ajuste dos parâmetros a cada nova base de imagens.

7.2.1 Seleção dos parâmetros dos extratores

Antes de aferir o desempenho do ISECO utilizando a combinação das características, investigamos qual o melhor conjunto de parâmetros para cada extrator de características: LBP, Gabor e LDP.

A abordagem escolhida para o LBP utiliza um conjunto de raios R com tamanhos distintos para extrair características em diferentes escalas. O experimento consiste em utilizar o LBP como único extrator de características do ISECO e variar os raios de 1 a 10. A análise dos raios é incremental. Por exemplo, primeiro será verificado o desempenho ao

utilizar apenas o raio $R = \{1\}$, em seguida o raios $R = \{1, 2\}$, e assim por diante. No final, será analisado o ganho com o incremento dos raios e qual o melhor conjunto de raios para o extrator de características LBP. A Figura 42 ilustra o comportamento do coeficiente de Dice à medida que novos raios vão sendo adicionados. A curva mostra que o coeficiente de Dice é constante com a adição de novos raios. Portanto, será utilizado um único raio R igual a 1 por ser menos custoso computacionalmente.

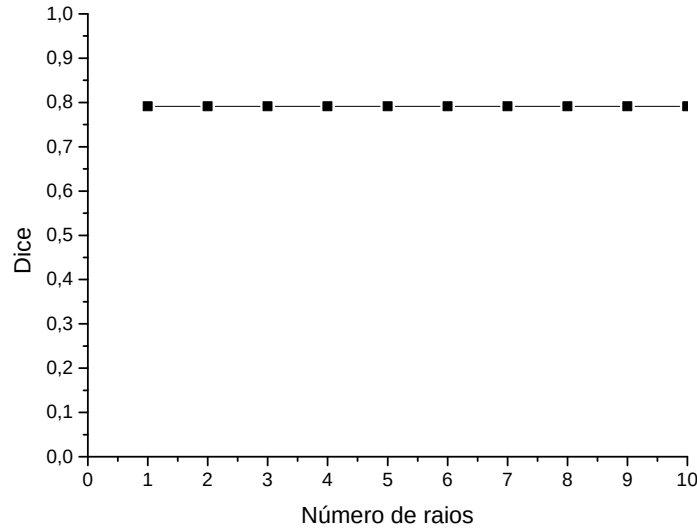


Figura 42: Comportamento do coeficiente de Dice à medida que adiciona-se novos raios.

Os filtros de Gabor apresentam os seguintes parâmetros: número de frequências da função sinusoidal f_m , número de orientações θ_n e a variância ao longo dos eixos σ_x e σ_y . Para facilitar a análise, assumimos que σ_x e σ_y são iguais. Neste experimento o número de orientações θ_n é definido como 4, 6 e 8 variando em passos 90° , 60° e 45° , respectivamente. O valor da variância ao longo dos eixos é 0.5, 1 e 1.5. Já o número de frequências da função sinusoidal varia em 4, 5 e 6. Para calcular as frequências f utiliza-se a seguinte função (JAIN; FARROKHNI, 1991): $2^i \times \sqrt{2}$, $i = 1, 2, \dots, f_n$. A Figura 43 mostra o coeficiente de Dice variando esses parâmetros para o ISECO utilizando somente os filtros de Gabor como extrator de características. De acordo com a figura, há pouca variação no valor do coeficiente de Dice. No entanto, o melhor valor com o menor custo computacional é obtido com $\sigma_x = \sigma_y = 0.5$, f_4 e θ_4 .

O extrator LDP apresenta como único parâmetro o limiar Z responsável por limitar o extrator pelos valores mais significativos. A Tabela 6 mostra o coeficiente de Dice dos resultado do ISECO usando somente o LDP e variando Z entre 1 e 8. De acordo com a tabela, qualquer raio de 1 a 7 pode ser uma boa escolha. Neste trabalho escolheu-se $Z = 4$ por ser a média do números de raios.

$\sigma_x = \sigma_y = 0.5$				$\sigma_x = \sigma_y = 1.0$				$\sigma_x = \sigma_y = 1.5$			
	θ_4	θ_6	θ_8		θ_4	θ_6	θ_8		θ_4	θ_6	θ_8
f_4	0.7855	0.7854	0.7854	f_4	0.7853	0.7852	0.785	f_4	0.7853	0.7852	0.7852
f_5	0.7855	0.7855	0.7854	f_5	0.7854	0.7852	0.7852	f_5	0.7854	0.7853	0.7852
f_6	0.7855	0.7855	0.7854	f_6	0.7854	0.7853	0.7853	f_6	0.7855	0.7853	0.7853

Figura 43: Coeficiente de Dice para todas as combinações das configurações descritas. As tabelas presentes na imagem são organizadas de acordo com o valor da variância ao longo dos eixos σ_x e σ_y . As colunas significam o número de orientações $\theta_n, n = 4, 6, 8$ e as linhas representam o número de frequências $f_m, m = 4, 5, 6$.

Tabela 6: Coeficiente de Dice variando o parâmetro Z entre 1 e 8.

Z	1	2	3	4	5	6	7	8
Dice	0,791	0,791	0,791	0,791	0,791	0,791	0,791	0,782

7.3 Seleção dos parâmetros do ISECO

O algoritmo ISECO apresenta uma restrição espacial (Equação (6.6)) que possui um parâmetro λ ($\lambda > 0$) que controla o raio de ação dessa restrição. Um valor de λ pequeno significa que pixels localizados na mesma região do pixel confuso serão tão importantes quanto os pixels de outras regiões da imagem, enquanto que um valor de λ grande pode prejudicar pixels de uma região próxima e diferente da que envolve o pixel confuso. O gráfico da Figura 44 mostra o comportamento do coeficiente de Dice para algoritmo ISECO de acordo com a variação do valor de λ . Nesse experimento utilizou-se o sub-conjunto de imagens da base de dados da retina e os parâmetros selecionados nos experimentos anteriores para cada extrator de características. De acordo com a Figura 44, a diferença do coeficiente de Dice entre λ igual a 1 e 100 é 0,0004. Apesar da diferença ser pequena, o melhor desempenho para essa base de dados foi obtido com $\lambda = 1$ ou $\lambda = 2$. O baixo valor de λ pode ser uma indicação de que a restrição espacial não tem influência no resultado final. No entanto, caso não se utilize a restrição espacial, o valor de Dice para esse conjunto de imagens diminui em torno de 0,0005. Portanto, mesmo que o ganho obtido com a restrição espacial seja pequeno, optamos por utilizá-la com o intuito de aumentar a assertividade do método proposto.

Outro parâmetro importante na abordagem proposta é o valor de K do algoritmo de pesquisa *KD-tree*. O valor de K determina o número de pixels que serão utilizados na Equação (6.5) para o *foreground* e *background*. Para avaliar o valor de K , o ISECO utiliza os parâmetros selecionados para cada extrator de característica e $\lambda = 1$. De acordo com a Figura 45, o melhor coeficiente de Dice obtido pelo ISECO para a base de dados da retina foi com $K = 15$. Esse valor será utilizado nos demais experimentos deste trabalho. Em relação ao tempo de processamento envolvendo o valor de K , a Figura 46 mostra que a

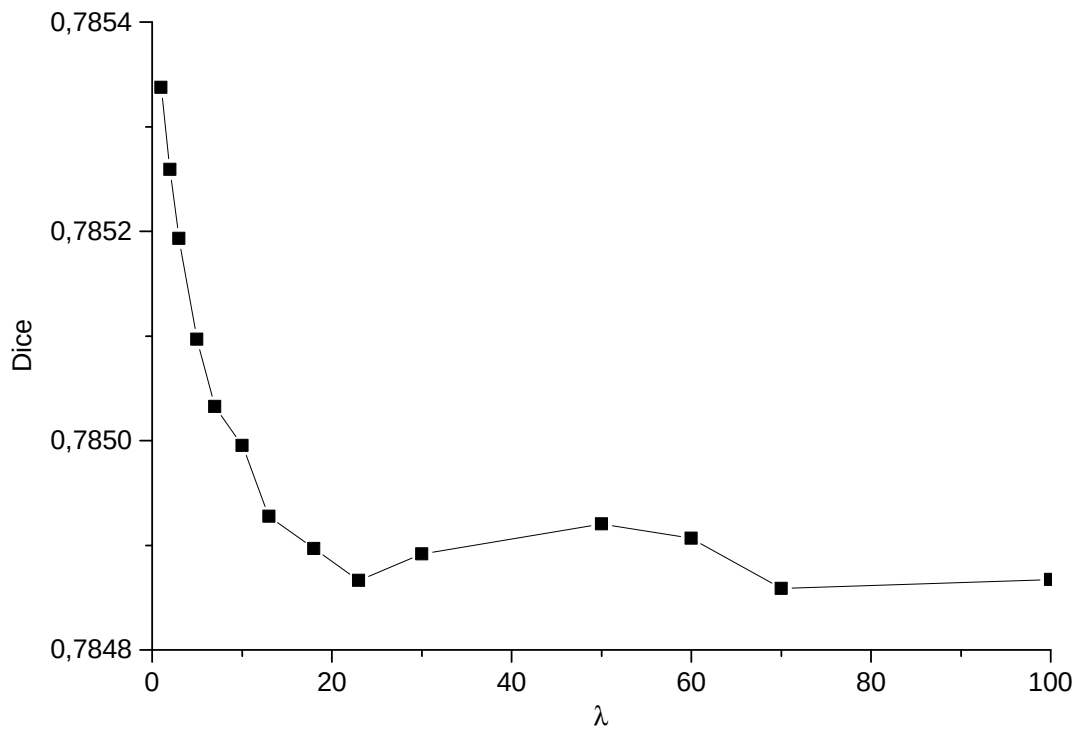


Figura 44: comportamento da taxa Dice para algoritmo ISECO de acordo com a variação do valor de λ .

medida que o valor de K aumenta, o tempo de processamento também aumenta.

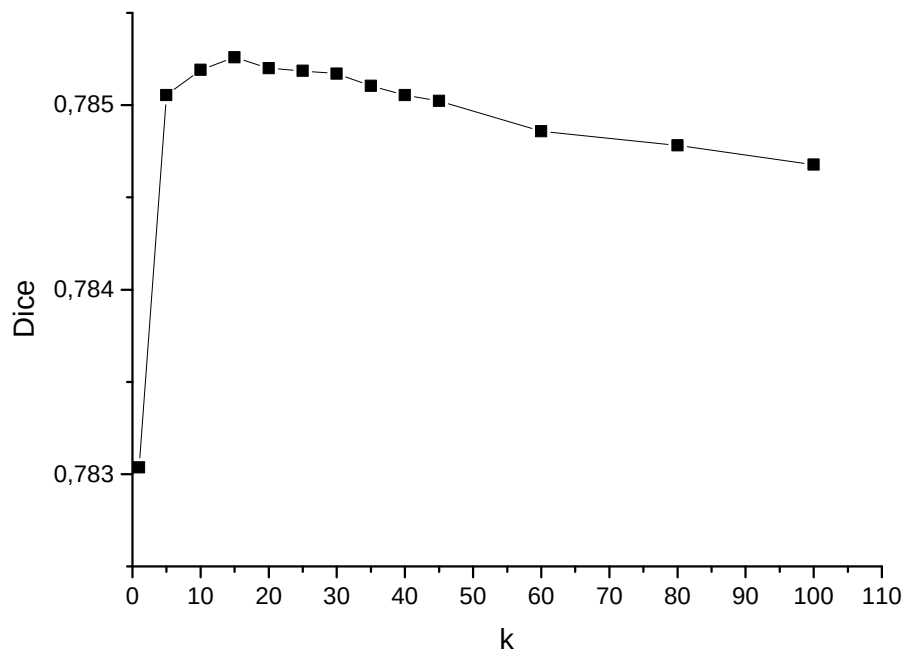


Figura 45: Desempenho do ISECO para a base de dados da retina variando o valor de k .

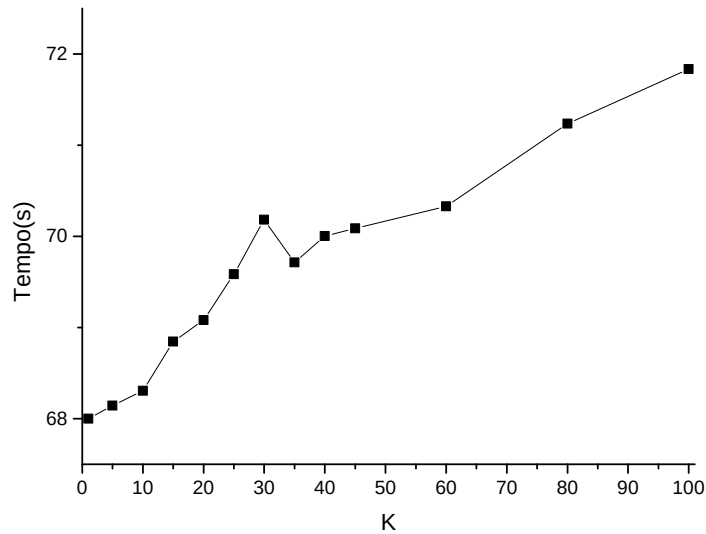


Figura 46: Tempo de processamento médio para cada imagem da base de dados da retina variando o valor de k .

7.4 Avaliação dos métodos para o consenso de segmentações

Neste experimento os métodos de consenso Voto Majoritário (VM), SIMPLE, STAPLE, COLLATE e as abordagens propostas *Distance Transform Merge* (DTM) e *Intelligent SEgmentation COnsensus* (ISECO) são avaliados de acordo com uma base contendo imagens sintéticas e três bases de imagens reais. As implementações dos métodos SIMPLE, STAPLE e COLLATE foram obtidas com os autores, já os métodos do voto majoritário e DTM foram implementados durante o desenvolvimento desse trabalho de tese.

7.4.1 Desempenho com imagens sintéticas

A Tabela 7 apresenta o desempenho dos métodos de consenso utilizando a base de dados com imagens sintéticas. As melhores taxas para cada medida estão em negrito. De acordo com a tabela, a abordagem proposta possui as melhores taxas de Dice, distância de Mahalanobis (MAH) e especificidade (ESP) que os demais métodos. O valor entre parênteses corresponde ao desvio padrão. Considerando a medida especificidade (ESP), pode-se perceber que todos os métodos fazem uma boa segmentação dos pixels do *background* nas imagens dessa base. No entanto, a diferença de desempenho entre os métodos pode ser justificada observando-se a medida SEN. Nesse caso, o algoritmo ISECO apresenta desempenho melhor que os outros. Apesar das imagens dessa base serem fáceis de serem segmentadas, como indicam as altas taxas, o método o ISECO se sobressai sobre os demais métodos.

A Figura 47 mostra um exemplo de segmentação para cada método de consenso utilizando a base de dados sintética.

Tabela 7: Desempenho médio dos métodos de consenso na base de imagens sintéticas. O valor entre parênteses corresponde ao desvio padrão para cada medida.

Métodos	Dice	MAH	SEN	ESP
VM	0,9688 (0,005)	0,0197 (0,0082)	0,9423 (0,0094)	0,9978 (0,0009)
SIMPLE	0,9688 (0,005)	0,0197 (0,0082)	0,9423 (0,0094)	0,9978 (0,0009)
STAPLE	0,9688 (0,005)	0,0197 (0,0082)	0,9423 (0,0094)	0,9978 (0,0009)
COLLATE	0,9688 (0,005)	0,0197 (0,0082)	0,9423 (0,0094)	0,9978 (0,0009)
DTM	0,9455 (0,0081)	0,0257 (0,0132)	0,8970 (0,0145)	0,9998 (0,0002)
ISECO	0,9751 (0,0038)	0,0163 (0,0083)	0,9584 (0,0064)	0,9946 (0,0017)

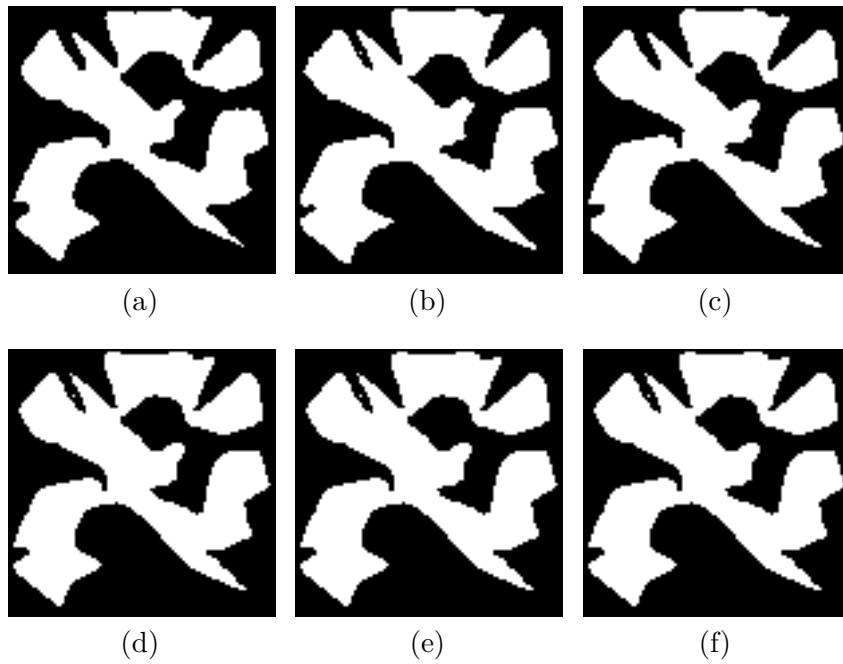


Figura 47: Exemplos de segmentações da imagem 41d obtidos pelos métodos de consenso para a base de imagens sintéticas: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.

7.4.2 Desempenho utilizando base de dados reais

As Tabelas 8 e 9 apresentam o desempenho dos métodos para as bases de imagens da retina e do ventrículo esquerdo, respectivamente. As melhores taxas estão em negrito. Como se pode observar, o método proposto apresenta maior coeficiente de Dice médio para a base da retina. Portanto, as imagens segmentadas pelo ISECO apresentam melhor similaridade em relação às imagens do *ground-truth* e dentre os algoritmos, o método proposto e o STAPLE segmentam melhor os pixels do *foreground*, pois apresentam maiores taxas de sensibilidade (SEN). Por outro lado, considerando a medida de distância de Mahalanobis (MAH), a abordagem proposta é a mais distante do *ground-truth* de que o VM, SIMPLE, COLLATE e STAPLE. Na Tabela 9 vemos que a abordagem proposta apresenta maior coeficiente de Dice e sensibilidade (SEN) para a base de imagens do ventrículo esquerdo, enquanto que os métodos VM, SIMPLE e COLLATE apresentam melhor distância de Mahalanobis.

A Tabela 10 apresenta o desempenho utilizando a base de imagens da sépala. Nessa tabela, percebe-se que o melhor coeficiente de Dice foi obtido pelo algoritmo do Voto Majoritário (VM). Note que, apesar do ISECO utilizar o voto majoritário, o método proposto apresenta coeficiente de Dice inferior. A justificativa está na taxa e sensibilidade (SEN), a qual informa que o ISECO rotulou erradamente mais pixels do *foreground* do que o voto majoritário. Isso pode acontecer em alguns casos, pois o ISECO sempre vai ter pelo menos o mesmo número de pixels do *foreground* que o VM. O fato dos valores de sensibilidade (SEN) na base da sépala serem baixos para todos os métodos indica que os segmentadores do COSE e os especialistas que geraram o *ground-truth* (segmentação manual) não concordam com a segmentação dos pixels do *foreground*. A Figura 49 ilustra um exemplo de retalho da sépala segmentado pelo COSE utilizando o ISECO e o *ground-truth* rotulado manualmente por um especialista. Percebe-se que visualmente o resultado do ISECO é melhor que o *ground-truth*. Além disso, as medidas selecionadas podem não ser ideais para avaliar problemas de segmentação de bordas. Considerando que a espessura da borda do *ground-truth* seja de 1 pixel, pontos segmentados fora dessa borda serão considerados como erros, mesmo estando próximos. A Figura 48 exemplifica erros dessa natureza. Quanto mais indefinida for a borda real do objeto, maiores são as chances desse tipo de erro acontecer. Dentre as quatro medidas descritas a que menos sofre com esse problema é a distância de Mahalanobis, pois considera a distância das bordas e não a sobreposição delas. Portanto, considerando essa medida os métodos VM, SIMPLE e COLLATE apresentam melhores resultados. Por outro lado, a segmentação obtida pelo ISECO de acordo com a distância de Mahalanobis está mais próxima do *ground-truth* que os métodos STAPLE e DTM.

Com o objetivo de avaliar estatisticamente o método proposto em relação aos demais, usou-se o teste pareado de Wilcoxon (GIBBONS; CHAKRABORTI, 2003). A

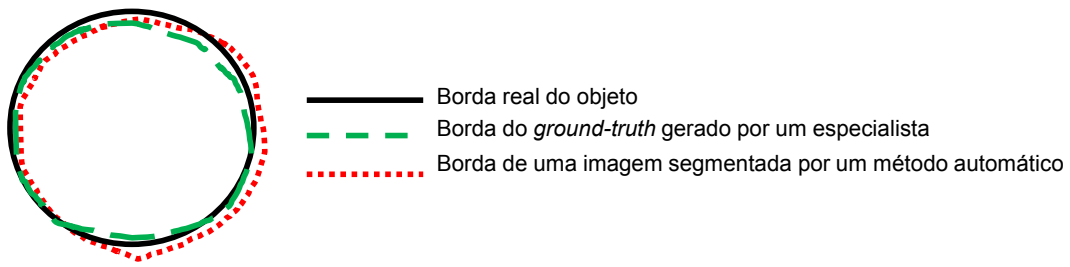


Figura 48: Exemplo de erros que podem acontecer na segmentação de bordas em imagens. A borda real do objeto é representada por uma linha sólida preta, enquanto que o *ground-truth* e a segmentação obtida são representadas por uma linha tracejada verde e uma linha pontilhada vermelha, respectivamente.

Tabela 8: Desempenho dos métodos de consenso na base de imagens da retina. O valor entre parênteses corresponde ao desvio padrão para cada medida.

Método	Dice	MAH	SEN	ESP
VM	0,7811 (0,0295)	0,0709 (0,036)	0,7227 (0,0657)	0,9883 (0,0043)
SIMPLE	0,7811 (0,0295)	0,0709 (0,0360)	0,7227 (0,0657)	0,9883 (0,0043)
STAPLE	0,7885 (0,0261)	0,0712 (0,0437)	0,7836 (0,0588)	0,9809 (0,007)
COLLATE	0,7811 (0,0295)	0,0709 (0,0360)	0,7227 (0,0657)	0,9883 (0,0043)
DTM	0,7698 (0,0313)	0,0812 (0,0399)	0,687 (0,0615)	0,9911 (0,0035)
ISECO	0,7887 (0,0259)	0,0718 (0,0429)	0,7835 (0,0589)	0,981 (0,0071)

Tabela 9: Desempenho dos métodos de consenso na base de imagens do ventrículo esquerdo. O valor entre parênteses corresponde ao desvio padrão para cada medida.

Método	Dice	MAH	SEN	ESP
VM	0,9220 (0,049)	0,1238 (0,077)	0,8685 (0,083)	0,9998 (0,004)
SIMPLE	0,9220 (0,049)	0,1238 (0,077)	0,8685 (0,083)	0,9998 (0,001)
STAPLE	0,9496 (0,034)	0,0976 (0,08)	0,9393 (0,047)	0,9993 (0,001)
COLLATE	0,9487 (0,032)	0,0757 (0,063)	0,9151 (0,055)	0,9998 (0,001)
DTM	0,9144 (0,039)	0,1152 (0,072)	0,8534 (0,068)	0,9998 (0,004)
ISECO	0,9513 (0,034)	0,0955 (0,082)	0,9445 (0,044)	0,9992 (0,001)

Tabela 10: Desempenho dos métodos de consenso na base de retalhos da sépala de uma planta. O valor entre parênteses corresponde ao desvio padrão para cada medida.

Método	Dice	MAH	SEN	ESP
VM	0,4596 (0,09)	0,1747 (0,33)	0,8384 (0,107)	0,8969 (0,025)
SIMPLE	0,4596 (0,09)	0,1747 (0,33)	0,8384 (0,107)	0,8969 (0,025)
STAPLE	0,4538 (0,097)	0,2103 (0,761)	0,851 (0,099)	0,8917 (0,025)
COLLATE	0,4596 (0,09)	0,174 (0,3302)	0,8384 (0,107)	0,8969 (0,025)
DTM	0,4517 (0,106)	0,1942 (0,381)	0,8509 (0,085)	0,8913 (0,022)
ISECO	0,4561 (0,094)	0,1814 (0,342)	0,8475 (0,101)	0,8929 (0,026)

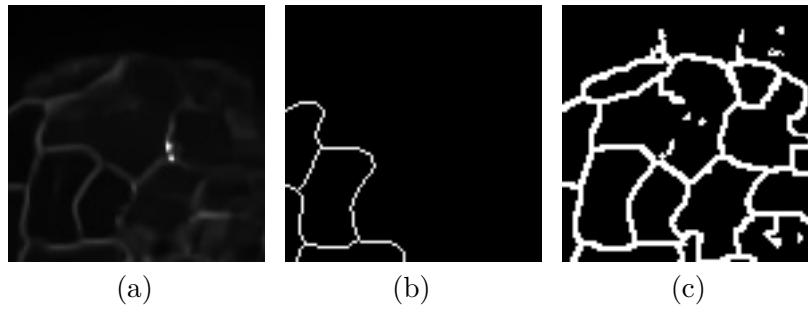


Figura 49: Exemplo de um retalho da sépala rotulado erroneamente: (a) retalho da sépala; (b) *ground-truth* rotulado manualmente por um especialista; (c) resultado utilizando o ISECO.

hipótese nula é que não há diferença no valor do coeficiente de Dice entre o método proposto e qualquer um dos métodos, enquanto que a hipótese alternativa é que o método proposto apresenta coeficiente de Dice maior que qualquer um dos métodos. Nesse experimento, pelos problemas descritos na base da sépala, ela não foi utilizada. Dado o nível de significância $\alpha = 0.05$, a Tabela 11 mostra o *p-value* do teste obtido entre o método proposto e cada um dos métodos para as bases da retina e ventrículo esquerdo. De acordo com a tabela, o método proposto apresenta coeficiente de Dice significativamente maior que os demais métodos para a base de imagens do ventrículo esquerdo ($p\text{-value} < \alpha$). O mesmo acontece para a base da retina, exceto quando comparado com o STAPLE, pois não há diferença estatística entre o coeficiente de Dice dos dois métodos para essa base.

Tabela 11: *P-values* obtidos utilizando o teste pareado de Wilcoxon entre a abordagem proposta e os demais métodos para as bases da retina e do ventrículo esquerdo.

Base de imagens	STAPLE	COLLATE	SIMPLE	DTM	VM
Retina	0,152	0,009	0,009	0,003	0,009
Vetrículo esquerdo	0,001	0,001	0,001	0,001	0,001

As Figuras 50, 51 e 52 ilustram o resultado dos métodos de consenso para as bases da retina, ventrículo esquerdo e para a base de imagens da sépala, respectivamente.

7.5 Análise da influência na quantidade de imagens segmentadas

No desenvolvimento deste trabalho, foi feita a suposição de que quanto mais segmentações de uma mesma imagem existirem melhor será o resultado de um método de consenso de segmentações. Com o objetivo de evidenciar essa suposição, a Figura 53 apresenta o coeficiente de Dice para os métodos de consenso utilizando a base de dados da retina segmentada por 25 algoritmos distintos. Primeiramente, os métodos são aferidos utilizando apenas 3 segmentações por imagem. Em seguida, esse número vai aumentando

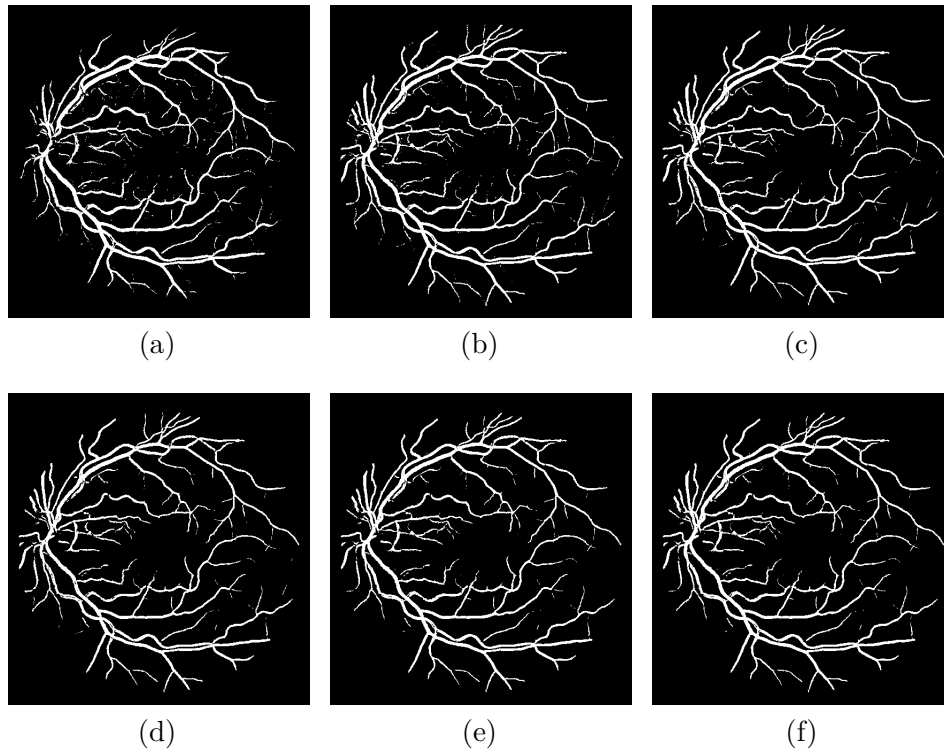


Figura 50: Exemplos de segmentações da Figura 41a obtidos pelos métodos de consenso para a base de imagens da retina: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.

em passos de 1 até chegar a 25 segmentações por imagem. Olhando o gráfico, podemos observar que o desempenho dos métodos aumenta à medida que o número de segmentações por imagem cresce. Em particular, o desempenho do método ISECO apresenta melhores resultados na maioria dos pontos da curva, com exceção dos últimos 5 pontos. Vale ressaltar também a melhora do desempenho do STAPLE à medida que o número de segmentações aumenta. As oscilações no gráfico são devido à existência de segmentações melhores que outras. No entanto, mesmo com essas segmentações "ruins", a tendência é de crescimento do coeficiente de Dice para todos os métodos.

7.5.1 Avaliação do ISECO usando SVM

A abordagem proposta pode ser considerada um classificador de padrões. Nesse sentido, o método de aprendizagem SVM, como descrita no Capítulo 6, pode ser utilizada para classificar os pixels confusos³. O *kernel* utilizado é o linear e o fluxo do classificador utilizando SVM é o mesmo do ISECO mostrado no Capítulo 6 até a etapa de extração de características. A Tabela 12 apresenta o resultado do consenso de segmentações baseado

³ Um pixel é considerado confuso se ele não foi classificado como *foreground* pelo voto majoritário e pelo menos um segmentador o classificou como *foreground*.

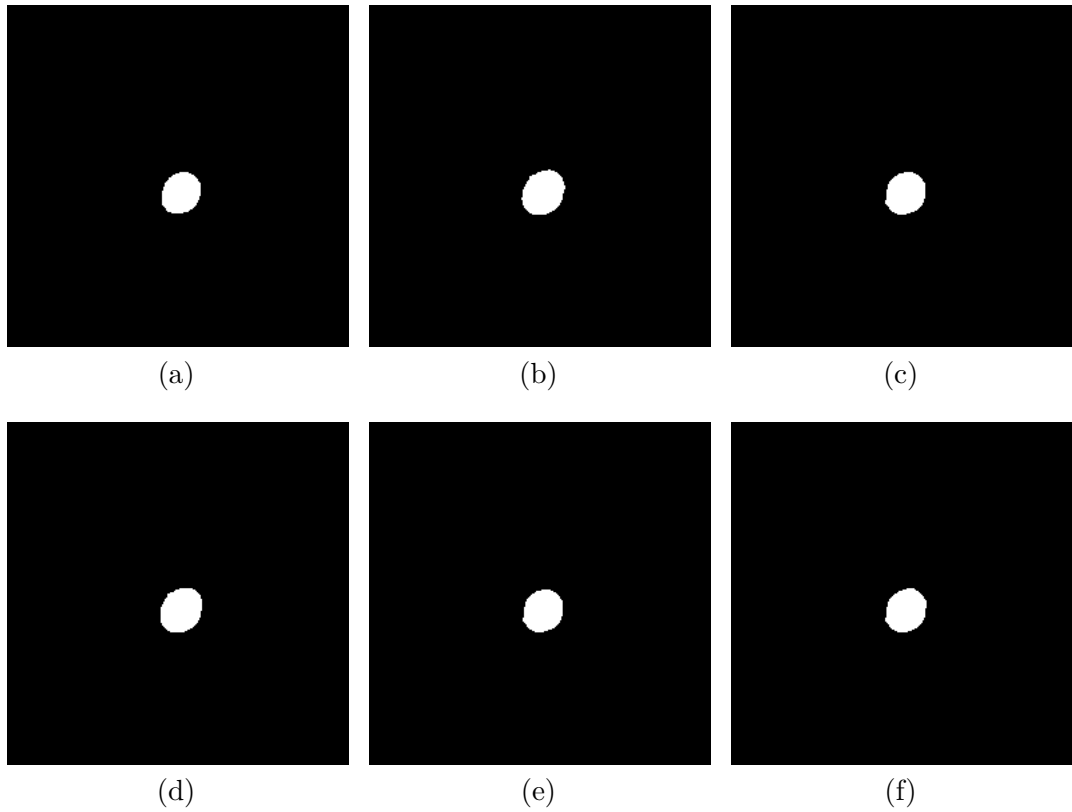


Figura 51: Exemplos de segmentações da Figura 41b obtidos pelos métodos de consenso para a base de imagens do ventrículo esquerdo: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.

nesse classificador para as bases da retina, ventrículo esquerdo e sépala de acordo com o coeficiente de Dice, escolhido anteriormente como a principal medida de desempenho neste trabalho. Observando a tabela, pode-se perceber que a abordagem utilizando SVM possui taxas semelhantes ao STAPLE, mas inferiores à abordagem proposta utilizando *KNN*. Para a abordagem com SVM não foi utilizada qualquer etapa de otimização dos parâmetros. No futuro, esperamos otimizar esses parâmetros, bem como testar o ISECO com outros modelos de aprendizagem.

Tabela 12: Resultado da abordagem usando ISECO com SVM em comparação com os métodos ISECO usando *KNN* e STAPLE.

	ISECO usando <i>KNN</i>	STAPLE	ISECO usando SVM
Ventrículo esquerdo	0,951	0,949	0,949
Retina	0,789	0,789	0,789
Sépala	0,456	0,449	0,449

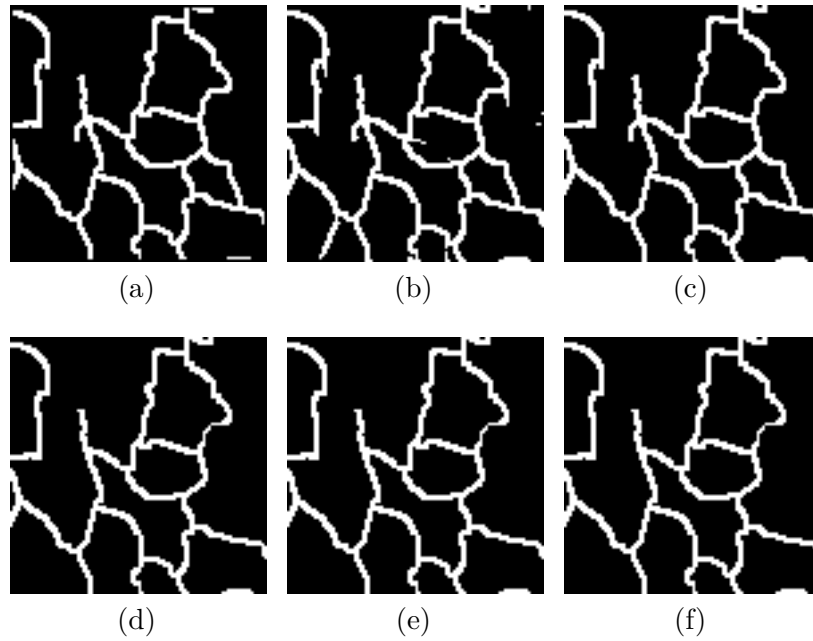


Figura 52: Exemplos de segmentações da imagem 41c obtidos pelos métodos de consenso para a base de imagens de retalhos da sépala: (a) resultado do DTM; (b) resultado do ISECO; (c) resultado do SIMPLE; (d) resultado do STAPLE; (e) resultado do VM; (f) resultado do COLLATE.

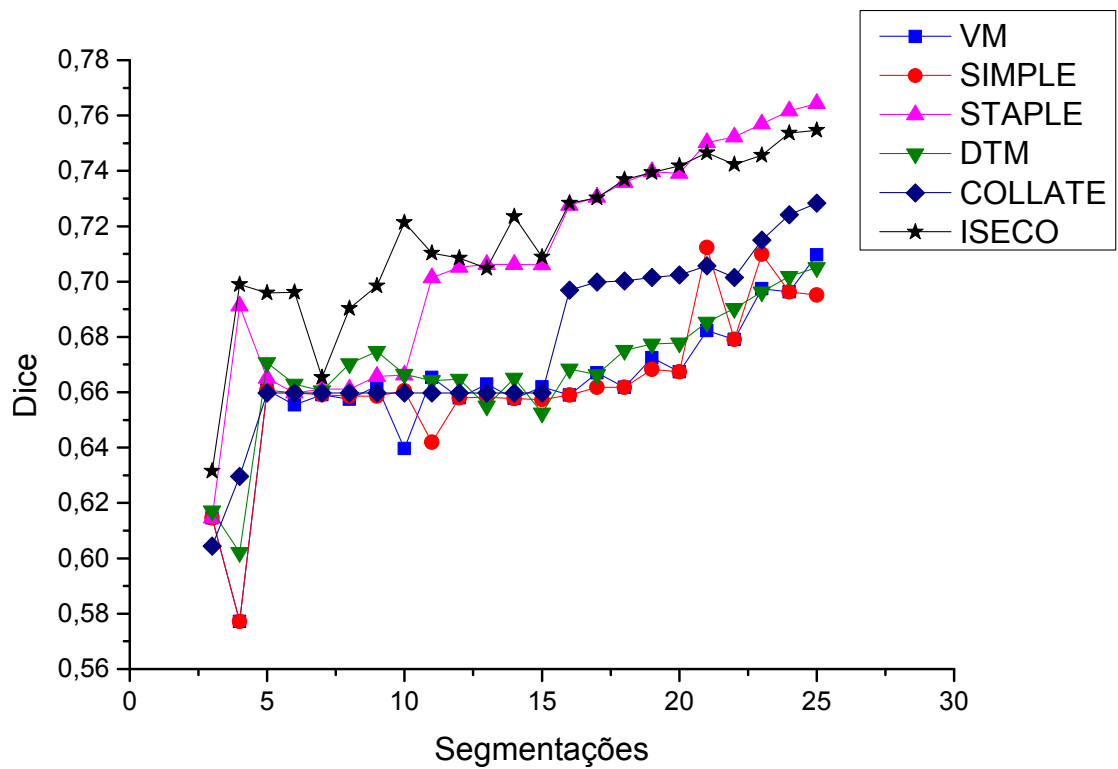


Figura 53: Coeficiente de Dice dos métodos de consenso à medida que o número de segmentações aumenta.

7.6 Tempo de processamento

Este experimento tem o propósito de avaliar os métodos de consenso quanto à eficiência no tempo de processamento. Utilizando o ambiente de desenvolvimento Matlab-2014a, o tempo de execução é medido em segundos. A máquina utilizada nos testes possui um processador Intel CORE i7-5500U @ 2.40GHz, com 8GB de memória RAM, disco rígido de 1TB com 8GB de SSD e sistema operacional de 64 bits Windows 8. A Tabela 13 mostra o tempo médio de processamento dos métodos de consenso para processar a base de dados da retina. De acordo com a Tabela 13, o algoritmo mais eficiente é o voto majoritário, seguido de perto pelo DTM. Entre os dois algoritmos que apresentam melhores coeficientes de Dice para essa base, o método ISECO é mais rápido que o STAPLE em torno de 20%. No entanto, em aplicações médicas, na maioria das vezes, é mais importante ter uma boa assertividade do que um tempo de processamento baixo, principalmente para evitar resultados falsos negativos. Além disso, o tempo de processamento pode não ser um problema, haja visto que muitos resultados de exames demorem horas ou até mesmo dias.

Tabela 13: Tempo médio de processamento dos métodos de consenso para processar a base de dados da retina.

Método	Tempo de execução (s)
VM	0,13
DTM	0,14
COLLATE	4,99
SIMPLE	2,12
ISECO	72,26
STAPLE	87,94

A Figura 54 mostra o tempo de processamento à medida que o número de segmentações aumenta. Como pode-se observar o algoritmo ISECO apresenta tempo de processamento menor que do STAPLE com o aumento no número de segmentações. Já o DTM, SIMPLE e voto majoritário apresentam tempo de processamento baixo mesmo com mais segmentações.

7.7 Avaliação do COSE como ferramenta para estimativa de *ground-truths*

Com o objetivo de propor um método para a estimação de *ground-truths*, pode-se supor que a inclusão de algoritmos de consenso ao COSE, torna essa ferramenta ideal para a estimativa de *ground-truths*. Ideal no sentido de diminuir o fator de subjetividade inerente a qualquer segmentação feita por um especialista manualmente. Portanto, o objetivo desse experimento é avaliar o quão boa é a estimativa de *ground-truth* do COSE em comparação com as segmentações individuais. A base de dados sintética foi utilizada nesse experimento por ter imagens rotuladas para avaliação sem qualquer erro. Ao contrário de outras bases,

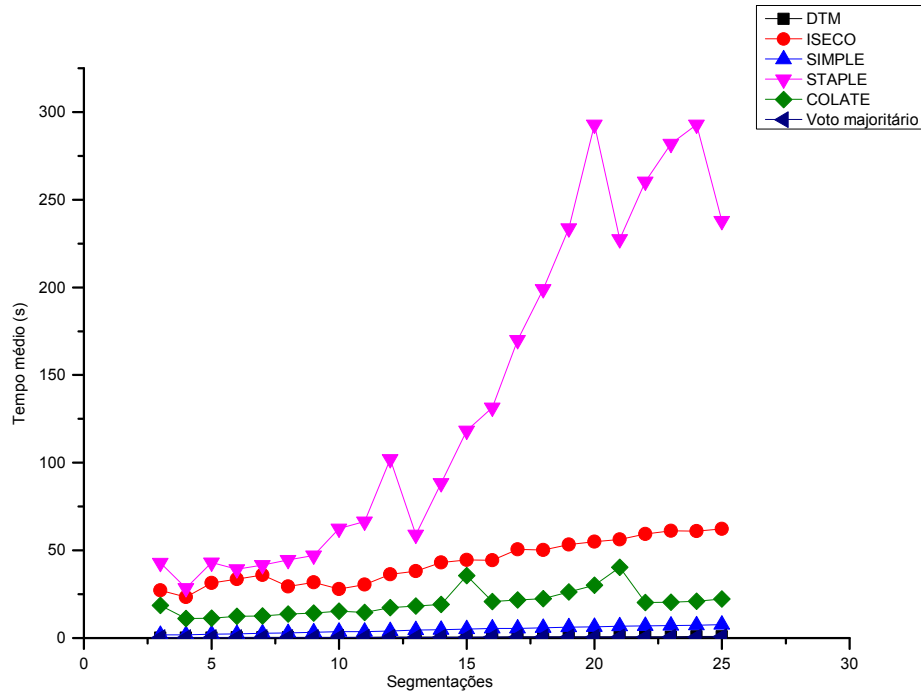


Figura 54: Tempo de processamento à medida que o número de segmentações aumenta.

como a da retina, ventrículo e sépala, os rótulos das imagens sintéticas não contêm viés de subjetividade dos especialistas, e consequentemente, os erros provocados por esses. As medidas Dice, distância de Mahalanobis (MAH), especificidade (ESP) e sensibilidade (SEN) foram utilizadas na avaliação.

A Tabela 14 apresenta o desempenho de cada segmentador em comparação com o resultado obtido pelo COSE. Como pode-se observar a segmentação obtida pela combinação do COSE mais o algoritmo ISECO é superior aos resultados obtidos por qualquer um dos segmentadores. Para afirmar que o resultado do COSE é estatisticamente superior aos dos segmentadores foi usado o teste de hipótese pareado de Wilcoxon considerando cada uma das 10 imagens da base de dados. A hipótese nula é que não há diferença entre as segmentações do COSE e a dos segmentadores, enquanto que a hipótese alternativa é que a segmentação obtida pelo COSE é melhor. Dado um nível de significância $\alpha = 0,05$, a Tabela 15 mostra os *p-values* associados à medida global Dice do COSE em comparação com cada segmentador. De acordo com a Tabela 15, pode-se rejeitar a hipótese nula ($p < \alpha$) com nível de confiança de 95%, constatando que estatisticamente o COSE apresenta melhor coeficiente de Dice que os segmentadores.

A segmentação utilizando o COSE pode ser realizada por qualquer pessoa. A assertividade na segmentação depende do nível de complexidade da imagem. A Figura 55 mostra o coeficiente de Dice para 3 usuários distintos na segmentação dos 10 primeiros retalhos da sépala apresentados pelo COSE. Como pode-se notar, a primeira imagem segmentada tende a ter uma baixa assertividade e de modo geral, a segmentação de um

Tabela 14: Desempenho dos segmentadores comparado com o resultado obtido pelo COSE.

Método	DICE	MAH	ESP	SEN
Segmentador 1	0,964	0,019	0,996	0,936
Segmentador 2	0,962	0,023	0,995	0,933
Segmentador 3	0,964	0,024	0,994	0,938
Segmentador 4	0,963	0,022	0,995	0,936
Segmentador 5	0,960	0,03	0,985	0,940
Consenso do COSE	0,975	0,016	0,956	0,995

Tabela 15: Teste de hipótese de Wilcoxon entre os segmentadores e o COSE

Método	p-value
Segmentador 1	0,001
Segmentador 2	0,001
Segmentador 3	0,001
Segmentador 4	0,001
Segmentador 5	0,001

usuário depende do nível de complexidade da imagem. Portanto, um usuário para ser considerado especialista para segmentação um determinado problema, depende do quão complexas são as imagens da base de dados. Outros fatores também podem influenciar para que um usuário possa se tornar um especialista, como por exemplo, a experiência em informática e a ferramenta interativa utilizada.

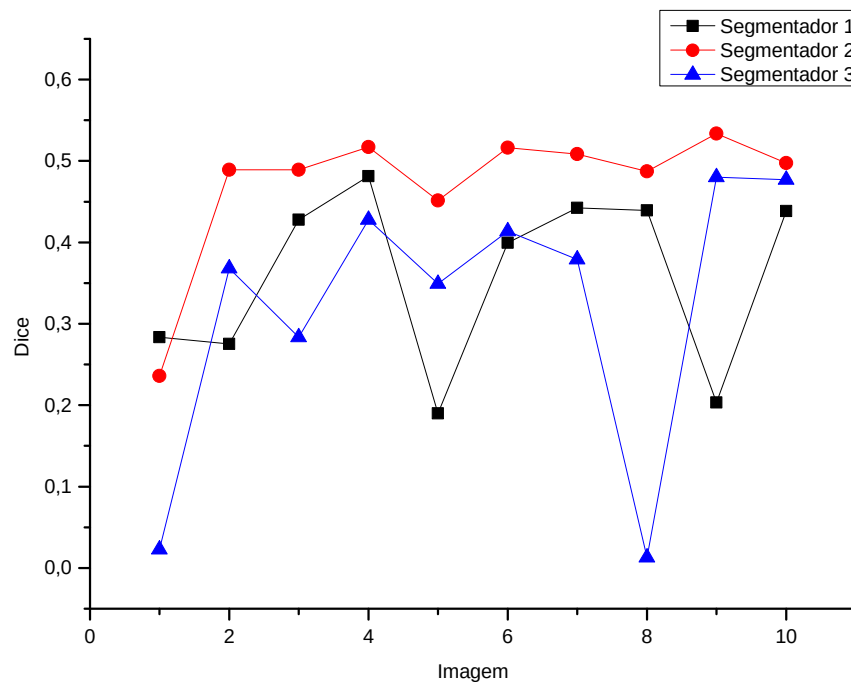


Figura 55: Coeficiente de Dice para 3 usuários distintos na segmentação dos 10 primeiros retalhos da sépala apresentados pelo COSE.

8 CONCLUSÕES E DISCUSSÕES

Nesse trabalho foi apresentada uma abordagem para a estimativa de *ground-truths* em imagens médicas que combina consenso de segmentações e segmentação colaborativa. A segmentação é uma das etapas do processamento de imagens digitais muito utilizada na área médica e pode fornecer informações para quantificar parâmetros clínicos importantes para o diagnóstico de doenças.

O sistema de segmentação colaborativa COSE é um *framework* que possibilita que vários usuários segmentem remotamente uma imagem. O objetivo neste trabalho foi usar o COSE para criar um conjunto com R segmentações da mesma imagem. A partir desse conjunto, um método de consenso de segmentações é usado para obter um resultado melhor que cada uma das segmentações.

Os métodos de consenso STAPLE, SIMPLE, COLLATE, voto majoritário e DTM são métodos que utilizam unicamente informações das imagens binárias. No método proposto de consenso, denominado ISECO, extrapolamos essa ideia e adicionamos informações extraídas da imagem em nível de cinza. A extração de características é baseada em três extratores de textura: filtros de Gabor, LBP e LDP. Dado o resultado do algoritmo baseado no voto majoritário, o ISECO tem como objetivo analisar os pixels marcados como *foreground* por pelo menos um segmentador. No entanto, quanto mais votos um pixel possuir, maior é a confiança de que o pixel realmente pertença ao *foreground*. O algoritmo analisa cada pixel confuso e fundamentado na similaridade deste com os pixels representantes do *foreground* e *background* decide qual a verdadeira classe do pixel. Os representantes das classes são escolhidos utilizando os k -vizinhos mais próximos nos dois conjuntos de acordo com o algoritmo *KD-tree*.

Na literatura existem várias abordagens para mensurar o desempenho de um método de segmentação. No entanto, não há medidas globais universalmente aceitas para todos os tipos de problema de segmentação. Cada problema de segmentação de imagens médicas possui características intrínsecas e, portanto, as medidas devem levá-las em consideração. Dessa forma, com o intuito de mostrar a generalidade dos métodos de consenso apresentados neste trabalho, propusemos a utilização de 4 medidas para a avaliação de desempenho. As medidas foram agrupadas em medidas baseadas em similaridade e distância. Dentre as medidas baseadas em similaridade, escolheu-se o coeficiente de Dice por ser uma medida bastante conhecida e menos sensível ao problema do desbalanceamento de classes, comum em imagens médicas. A medida baseada na distância de Mahalanobis foi selecionada como a representante do grupo das medidas baseadas em distância. Como mostrado, a distância de Mahalanobis é uma medida menos sensível a ruídos que podem aparecer em imagens segmentadas. A sensibilidade e especificidade também foram selecionadas por

serem medidas que descrevem o erro no *foreground* e no *background*, respectivamente.

Nos experimentos, a abordagem proposta para o consenso de segmentações ISECO apresentou melhores resultados em duas bases de dados: ventrículo esquerdo e sintética. Na base de imagens da retina o algoritmo proposto apresenta coeficiente de Dice estatisticamente igual ao algoritmo STAPLE. Na avaliação utilizando a base de retalhos de uma imagem sépala, o algoritmo apresentou resultado inferior ao voto majoritário, COLLATE e SIMPLE. No entanto, de acordo com as baixas taxas obtidas por todos os métodos, não se pode garantir a precisão do *ground-truth* criado manualmente pelo especialista. Esse tipo de erro, provocado pela subjetividade dos especialistas, pode ser considerado um problema típico da criação de *ground-truths* utilizando segmentação manual. Neste caso, um método de consenso pode ser bastante útil ao fornecer a possibilidade de diminuir essa subjetividade. Como mostrado nos experimentos, o resultado de um método de consenso tende a melhorar à medida que o número de segmentações aumenta. Em particular o desempenho do método proposto apresentou melhor coeficiente de Dice para diferentes quantidades de segmentações ao longo do experimento.

Se considerarmos o tempo de processamento, métodos mais simples como o DTM e o voto majoritário são os mais indicados. No entanto, a abordagem proposta apresenta tempo de processamento menor que o STAPLE, que é a abordagem sofisticada mais utilizada para o consenso de segmentações (JOMIER; LEDIGARCHER; AYLWARD, 2005; SUINESIAPUTRA et al., 2014). Além disso, o tempo de processamento do ISECO, no que diz respeito a classificação, está exclusivamente ligado ao número de pixels confusos e à quantidade de pixels das classes de *foreground* e *background*. Ao contrário do STAPLE, que apresenta tempo de processamento maior à medida que o número de segmentações para uma mesma imagem aumenta.

A abordagem apresentada para o consenso de segmentações utiliza o KNN classificar os pixels confusos. O método difere das abordagens clássicas por construir a base de treinamento de modo *online*. Os pixels das classes *foreground* e *background* usados no treinamento são selecionados em tempo de execução após a pré-segmentação. Nessa área de aprendizagem de máquina, nos últimos anos vem crescendo o número de trabalhos em segmentação de imagens que utilizam redes *deep learning* (XING; XIE; YANG, 2016; MANSOOR et al., 2016). A rede *deep learning* mais utilizada em processamento de imagens, conhecida como *Convolutional Neural Network* (CNN), utiliza um conjunto de camadas de convolução e é capaz de aprender filtros para extração de características no momento do treinamento da rede. O uso de *deep learning* no problema de consenso de segmentações é uma abordagem ainda não investigada, mas que pode proporcionar bons resultados. No entanto, redes *deep learning* apresentam como característica o fato de possuírem arquiteturas enormes, com milhares de nós, o que torna o seu treinamento

bastante custoso, mesmo utilizando GPU¹. Uma abordagem menos custosa seria treinar a rede de modo *offline* para aprender os filtros de extração de características e utilizar alguns desses filtros no treinamento *online* de outros métodos de aprendizagem com tempo de treinamento menor.

Uma limitação da abordagem proposta ISECO é a dificuldade de segmentar imagens cujo o objeto de interesse apresenta estrutura similar a outras partes da imagem. Se os usuários marcarem erroneamente essas partes, o algoritmo irá analisar os pixels dessas estruturas e caso sejam realmente similares, o ISECO pode classificá-los também erroneamente como *foreground*. Além disso, o algoritmo não é hábil para segmentações com múltiplos rótulos. A abordagem se restringe apenas a problemas de segmentações binárias. Por apresentar tempo de processamento elevado, o ISECO não deve ser utilizado em problemas de consenso de segmentação em tempo real. Em trabalhos futuros pretendemos diminuir esse tempo de processamento através da otimização do código e a otimização no número características utilizadas. Além disso, pretendemos torná-lo um método para o consenso de segmentações com várias classes.

A outra abordagem desenvolvida neste trabalho, denominada *Distance Transform Merge* (DTM), realiza a combinação de segmentações através da criação de mapas de confiança baseados na transformada de distância. A média dos mapas de confiança é confrontada com um valor de limiar obtido através do algoritmo de Otsu. Essa abordagem tem como principal vantagem o baixo tempo de processamento. No entanto, apresenta precisão inferior ao ISECO. Como trabalho futuro pode-se investigar outras abordagens para construir o mapa de confiança e técnicas de limiarização.

O sistema colaborativo de segmentação de imagens COSE é um *framework* que possibilita aumentar a precisão na segmentação de imagens, tornando-se, dessa forma, uma alternativa interessante para a criação de *ground-truths*. No entanto, a ferramenta ainda precisa de algumas melhorias, como a criação de um mecanismo que estimule os usuários a continuarem utilizando o sistema. Nesse sentido, uma ideia para trabalhos futuros é *gamificar* o sistema de modo a torná-lo mais empolgante para o usuário que participa do processo. O sistema atualmente conta com uma ferramenta de segmentação manual e 3 métodos interativos de segmentação de imagens: *Live-Wire*, *Live-Vessel* e *Watershed*. Esses algoritmos de segmentação interativa podem aumentar a eficiência da segmentação dependendo do tipo da imagem. No entanto, podem existir casos onde a melhor alternativa continue sendo a segmentação manual. Nesse sentido, no futuro pretende-se implementar mais métodos de segmentação interativos ou automáticos para diminuir a necessidade da segmentação manual.

¹ Processador especializado em processar dados em paralelo.

REFERÊNCIAS

- ABDALLAH, S. A.; NOLAND, K.; SANDLER, M.; CASEY, M.; RHODES, C. Theory and evaluation of a bayesian music structure extractor. In: **Proceedings of the 6th International Conference on Music Information Retrieval**. London, United Kingdom: University of London, 2005. p. 420–425.
- AMIS, E. S. Anatomy and physiology of the prostate. In: LANG, E. K. (Ed.). **Radiology of the Lower Urinary Tract**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994. p. 167–169.
- ASMAN, A. J.; LANDMAN, B. A. Robust statistical label fusion through consensus level, labeler accuracy, and truth estimation (COLLATE). **IEEE Transactions on Medical Imaging**, IEEE, v. 30, n. 10, p. 1779–1794, out. 2011.
- BAILEY, D.; TOWNSEND, D.; VALK, P.; MAISEY, M. **Positron Emission Tomography: Basic Sciences**. London, UK: Springer London, 2004.
- BALAFAR, M. A. Gaussian mixture model based segmentation methods for brain mri images. **Artificial Intelligence Review**, v. 41, n. 3, p. 429–439, 2012.
- BALAFAR, M. A.; RAMLI, A. R.; SARIPAN, M. I.; MASHOHOR, S. Review of brain mri image segmentation methods. **Artificial Intelligence Review**, Kluwer Academic Publishers, v. 33, n. 3, p. 261–274, 2010.
- BALDI, P.; BRUNAK, S.; CHAUVIN, Y.; ANDERSEN, C. A. F.; NIELSEN, H. Assessing the accuracy of prediction algorithms for classification: an overview. **Bioinformatics**, v. 16, n. 5, p. 412–424, 2000.
- BANKMAN, I. N. (Ed.). **Handbook of Medical Imaging**. Orlando, FL, USA: Academic Press, Inc., 2000. ISBN 0-12-077790-8.
- BAR, A.; MAHESWARAN, M. **Confidentiality and Integrity in Crowdsourcing Systems**. New York: Springer, 2013.
- BEAUCHEMIN, M.; THOMSON, K.; EDWARDS, G. On the hausdorff distance used for the evaluation of segmentation results. **Canadian Journal of Remote Sensing**, v. 24, n. 1, p. 3–8, 1998.
- BEGA, G.; LEV-TOAFF, A.; KUHLMAN, K.; KURTZ, A.; GOLDBERG, B.; WAPNER, R. Three-dimensional ultrasonographic imaging in obstetrics: present and future applications. **Journal of Ultrasound in Medicine**, American Institute of Ultrasound in Medicine, v. 20, n. 4, p. 391–408, 2001.
- BENTLEY, J. L. Multidimensional binary search trees used for associative searching. **Communications of the ACM**, v. 18, n. 9, p. 509–517, set. 1975.
- BLAND, J.; ALTMAN, D. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. **Computers in Biology and Medicine**, v. 20, n. 5, p. 337 – 340, 1990.

- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: **Proceedings of the Fifth Annual Workshop on Computational Learning Theory**. New York, NY, USA: ACM, 1992. p. 144–152.
- BOURAOUI, B.; RONSE, C.; BARUTHIO, J.; PASSAT, N.; GERMAIN, P. Fully automatic 3d segmentation of coronary arteries based on mathematical morphology. In: **ISBI**. Paris, France: IEEE, 2008. p. 1059–1062.
- BUADES, A.; COLL, B.; MOREL, J. M. A non-local algorithm for image denoising. In: **Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on**. San Diego, CA: IEEE Computer Society, 2005. v. 2, p. 60–65.
- CHALANA, V.; KIM, Y. A methodology for evaluation of boundary detection algorithms on medical images. **IEEE Transaction on Medical Imaging**, v. 16, n. 5, p. 642–652, 1997.
- CHAUDHURI, S.; CHATTERJEE, S.; KATZ, N.; NELSON, M.; GOLDBAUM, M. Detection of blood vessels in retinal images using two-dimensional matched filters. **IEEE Transactions on Medical Imaging**, v. 8, n. 3, p. 263–269, 1989.
- CHENG, Y.; HU, X.; WANG, J.; WANG, Y.; TAMURA, S. Accurate vessel segmentation with constrained b-snake. **IEEE Transactions on Image Processing**, v. 24, n. 8, p. 2440–2455, 2015.
- CHODOROWSKI, A.; MATTSSON, U.; LANGILLE, M.; HAMARNEH, G. Color lesion boundary detection using live wire. In: . San Diego, CA: IEEE Computer Society, 2005. v. 5747, p. 1589–1596.
- CHORAS, R. S. Feature extraction for cbir and biometrics applications. In: **Proceedings of the 7th Conference on 7th WSEAS International Conference on Applied Computer Science**. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2007. v. 7, p. 1–9.
- COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37, 1960.
- CORMEN, T. H.; STEIN, C.; RIVEST, R. L.; LEISERSON, C. E. **Introduction to Algorithms**. 2nd. ed. London: McGraw-Hill Higher Education, 2001.
- COTE, M.; ALBU, A. B. Robust texture classification by aggregating pixel-based LBP statistics. **IEEE Signal Processing Letters**, v. 22, n. 11, p. 2102–2106, 2015.
- DAINTY, J. C.; SHAW, R. **Image science: principles, analysis and evaluation of photographic-type imaging processes**. London: Academic Press London ; New York, 1974. 402 p.
- DAWOUD, A. Lung segmentation in chest radiographs by fusing shape information in iterative thresholding. **IET Computer Vision**, v. 5, n. 3, p. 185–190, 2011.
- DICE, L. R. Measures of the Amount of Ecologic Association Between Species. **Ecology**, v. 26, n. 3, p. 297–302, jul. 1945.

- DICKIE, R.; HAMARNEH, G.; ABUGHARBIEH, R. **Live-vessel: Interactive vascular image segmentation with simultaneous extraction of optimal medial and boundary paths**. School of Computing Science, Simon Fraser University, Burnaby, BC, Canada, 2009. 1-30 p.
- DIGABEL, H.; LANTUÉJOUL, C. Iterative algorithms. In: VERLAG, R. (Ed.). **Actes du Second Symposium Europeen d'Analyse Quantitative des Microstructures en Sciences des Materiaux, Biologie et Medecine**. Caen, França: ACM, 1978. p. 85–99.
- DOAN, A.; RAMAKRISHNAN, R.; HALEVY, A. Y. Crowdsourcing systems on the world-wide web. **Commun. ACM**, ACM, v. 54, n. 4, p. 86–96, abr. 2011.
- DODIN, P.; PELLETIER, J. P.; MARTEL-PELLETIER, J.; ABRAM, F. Automatic human knee cartilage segmentation from 3-d magnetic resonance images. **IEEE Transactions on Biomedical Engineering**, v. 57, n. 11, p. 2699–2711, 2010.
- DUNN, D. F.; HIGGINS, W. E. Optimal gabor filters for texture segmentation. **IEEE Transactions on Image Processing**, v. 5, n. 7, p. 947–964, 1995.
- DUNN, D. F.; HIGGINS, W. E.; WAKELEY, J. Texture segmentation using 2-d gabor elementary functions. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 16, n. 2, p. 130–149, 1994.
- ESTRADA, F. J.; JEPSON, A. D. Benchmarking image segmentation algorithms. **International Journal of Computer Vision**, v. 85, n. 2, p. 167–181, 2009.
- FALCÃO, A. X.; STOLFI, J.; LOTUFO, R. A. The image foresting transform: Theory, algorithms, and applications. **IEEE Transaction on Pattern Analysis and Machine Intelligence**, v. 26, n. 1, p. 19–29, 2004.
- FALCÃO, R. X.; UDUPA, J. K.; MIYAZAWA, F. K. An ultra-fast user-steered image segmentation paradigm: live-wire-on-the-fly. **IEEE Transactions on Medical Imaging**, p. 55–62, 2000.
- FAWCETT, T. An introduction to roc analysis. **Pattern Recognition Letters**, Elsevier Science Inc., v. 27, n. 8, p. 861–874, jun. 2006.
- FELSON, D. T. Osteoarthritis of the knee. **New England Journal of Medicine**, v. 354, n. 8, p. 841–848, 2006.
- FIX, E.; HODGES, J. L. Discriminatory analysis: Nonparametric discrimination: Consistency properties. **US Air Force School of Aviation Medicine**, Technical Report 4, n. 3, p. 261–279, jan. 1951.
- FRANGI, R. F.; NIESSEN, W. J.; VINCKEN, K. L.; VIERGEVER, M. A. Multiscale vessel enhancement filtering. In: . Miami: Springer-Verlag, 1998. p. 130–137.
- GABOR, D. Theory of communication. **Journal of the Institution of Electrical Engineering**, v. 93, n. 26, p. 429–457, nov. 1946.
- GAO, Z.; WANG, L.; ZHOU, L.; ZHANG, J. Hep-2 cell image classification with deep convolutional neural networks. **IEEE Journal of Biomedical and Health Informatics**, PP, n. 99, p. 1–1, 2016.

- GAVAGHAN, K. A.; PETERHANS, M.; OLIVEIRA-SANTOS, T.; WEBER, S. A portable image overlay projection device for computer-aided open liver surgery. **IEEE Transactions on Biomedical Engineering**, v. 58, n. 6, p. 1855–1864, 2011.
- GHOSE, S.; OLIVER, A.; MARTÍ, R.; LLADÓ, X.; VILANOVA, J. C.; FREIXENET, J.; MITRA, J.; SIDIBÉ, D.; MERIAUDEAU, F. A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images. **Computer Methods and Programs in Biomedicine**, v. 108, n. 1, p. 262 – 287, 2012.
- GIBBONS, J.; CHAKRABORTI, S. **Nonparametric Statistical Inference**. Berlin, Heidelberg: Marcel Dekker Incorporated, 2003. (Statistics, textbooks and monographs).
- GIBBS, P.; BUCKLEY, D.; BLACKBAND, S. J.; HORSMAN, A. Tumour volume determination from mr images by morphological segmentation. **Physics in Medicine and Biology**, v. 41, n. 11, p. 2437, 1996.
- GILLES, B.; MOCCOZET, L.; MAGNENAT-THALMANN, N. Anatomical modelling of the musculoskeletal system from mri. In: LARSEN, R.; NIELSEN, M.; SPORRING, J. (Ed.). **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part I**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 289–296.
- GINNEKEN, B. v.; KERKSTRA, S.; SHNEIDER, C. **Grand Challenges in Biomedical Image Analysis**. 2016. Disponível em: <<https://grand-challenge.org/>>.
- GONG, L.; PATHAK, S. D.; HAYNOR, D. R.; CHO, P. S.; KIM, Y. Parametric shape modeling using deformable superellipses for prostate segmentation. **IEEE Transactions on Medical Imaging**, v. 23, n. 3, p. 340–349, 2004.
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing (3rd Edition)**. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.
- GUBERN-MÉRIDA, A.; KALLENBERG, M.; MANN, R. M.; MARTÍ, R.; KARSSEMEIJER, N. Breast segmentation and density estimation in breast mri: A fully automatic framework. **IEEE Journal of Biomedical and Health Informatics**, v. 19, n. 1, p. 349–357, Jan 2015.
- HAN, S.; NIJDAM, N. A.; SCHMID, J.; KIM, J.; MAGNENAT-THALMANN, N. Collaborative telemedicine for interactive multiuser segmentation of volumetric medical images. **The Visual Computer**, Springer-Verlag New York, Inc., v. 26, p. 639–648, June 2010.
- HAN, X.; FISCHL, B. Atlas renormalization for improved brain mr image segmentation across scanner platforms. **IEEE Transaction on Medical Imaging**, v. 26, n. 4, p. 479–486, 2007.
- HARALICK, R. M.; SHAPIRO, L. G. **Computer and Robot Vision**. Michigan: Addison Wesley, 1992. v. 1.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation (3rd Edition)**. Nova Jersey: Prentice-Hall, Inc., 2007.

- HE, J.; KIM, C.-S.; KUO, C.-C. J. **Interactive Segmentation Techniques: Algorithms and Performance Evaluation**. New York, London: Springer Publishing Company, Incorporated, 2013.
- HOSSEINZADEH, M.; KHOSHVAGHT, P. A comparative study of image segmentation algorithms. **International Journal of Computer, Electrical, Automation, Control and Information Engineering**, v. 9, n. 8, p. 1966 – 1971, 2015.
- HUBERT, L.; ARABIE, P. Comparing partitions. **Journal of classification**, Springer, v. 2, n. 1, p. 193–218, 1985.
- JABID, T.; KABIR, M. H.; CHAE, O. Local directional pattern (LDP) for face recognition. In: **2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE)**. Oxford, England: IEEE, 2010. p. 329–330.
- JACCARD, P. The distribution of the flora in the alpine zone. **New Phytologist**, v. 11, n. 2, p. 37–50, fev. 1912.
- JAIN, A. K.; FARROKHNI, F. Unsupervised texture segmentation using gabor filters. **Pattern Recognition**, v. 24, n. 12, p. 1167–1186, dez. 1991.
- JOMIER, J.; LEDIGARCHER, V.; AYLWARD, S. R. Comparison of vessel segmentations using staple. In: DUNCAN, J. S.; GERIG, G. (Ed.). **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005: 8th International Conference, Palm Springs, CA, USA, October 26-29, 2005, Proceedings, Part I**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. p. 523–530.
- JONES, A. **Advanced Data Analytics for Business Leaders Explained**. 2014. Disponível em: <http://www.kdnuggets.com/2014/09/advanced-data-analytics-business-leaders-explained.html>.
- KAMARAINEN, J.-K.; KYRKI, V.; KÄLVIÄINEN, H. Invariance properties of Gabor filter based features - overview and applications. **IEEE Transactions on Image Processing**, v. 15, n. 5, p. 1088–1099, 2006.
- KAPUR, T.; GRIMSON, W.; KIKINIS, R.; III, W. W. Enhanced spatial priors for segmentation of magnetic resonance imagery. **Medical Image Computing and Computer-Assisted Intervention. MICCAI 1998**, v. 1, p. 457–68, 10 1998.
- KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: Active contour models. **INTERNATIONAL JOURNAL OF COMPUTER VISION**, v. 1, n. 4, p. 321–331, 1988.
- KHOOSHABI, G. S. **Segmentation Validation Framework**. 2013. 53 p. Dissertação (Mestrado) — Linköping UniversityLinköping UniversityLinköping University, Department of Biomedical Engineering, Center for Medical Image Science and Visualization, CMIV, The Institute of Technology, 2013.
- KITTLER, J.; HATEF, M.; DUIN, R. P. W.; MATAS, J. On combining classifiers. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE Computer Society, Washington, DC, USA, v. 20, n. 3, p. 226–239, mar. 1998.

KLAVA, B.; HIRATA, N. S. T. Interactive image segmentation with integrated use of the markers and the hierarchical watershed approaches. In: **VISAPP 2009 - Proceedings of the Fourth International Conference on Computer Vision Theory and Applications**. Lisboa, Portugal: Springer Berlin Heidelberg, 2009. v. 1, p. 186–193.

LAM, B. S. Y.; YAN, H. A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields. **IEEE Transactions on Medical Imaging**, v. 27, n. 2, p. 237–246, Feb 2008.

LANDIS, R. J.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159–74, 3 1977.

LANGERAK, R.; HEIDE, U. A. van der; KOTTE, A. N. T. J.; VIERGEVER, M. A.; VULPEN, M. van; PLUIM, J. P. W. Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (simple). **IEEE Transaction on Medical Imaging**, v. 29, n. 12, p. 2000–2008, 2010.

LEBENBERG, J.; LALANDE, A.; CLARYSSE, P.; BUVAT, I.; CASTA, C.; COCHET, A.; CONSTANTINIDÈS, C.; COUSTY, J.; CESARE, A. de; JEHAN-BESSON, S.; LEFORT, M.; NAJMAN, L.; ROULLOT, E.; SARRY, L.; TILMANT, C.; FROUIN, F.; GARREAU, M. Improved estimation of cardiac function parameters using a combination of independent automated segmentation results in cardiovascular magnetic resonance imaging. **PLOS ONE**, Public Library of Science, v. 10, p. 1–16, 08 2015.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.

LEI, C.; XIAOJIAN, L.; JIE, Z.; WUFAN, C. Automated lung segmentation algorithm for cad system of thoracic ct. **Journal of Medical Colleges of PLA**, v. 23, n. 4, p. 215–222, 2008.

LI, Q.; NISHIKAWA, R. **Computer-Aided Detection and Diagnosis in Medical Imaging**. New York, USA: CRC Press, 2015. (Imaging in medical diagnosis and therapy).

LIU, S.; LI, J. Automatic medical image segmentation using gradient and intensity combined level set method. In: **2006 International Conference of the IEEE Engineering in Medicine and Biology Society**. New York, NY, USA: IEEE, 2006. p. 3118–3121.

MAGESWARI, S. U.; SRIDEVI, M.; MALA, C. An experimental study and analysis of different image segmentation techniques. **Procedia Engineering**, v. 64, n. 0, p. 36 – 45, 2013.

MAHALANOBIS, P. C. On the generalized distance in statistics. **Proceedings of the National Institute of Sciences (Calcutta)**, v. 2, p. 49–55, 1936.

MANOUSAKAS, I.; UNDRILL, P.; CAMERON, G.; REDPATH, T. Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions. **Computers and biomedical research, an international journal**, v. 31, n. 6, 1998.

MANSOOR, A.; CERROLAZA, J. J.; IDREES, R.; BIGGS, E.; ALSHARID, M. A.; AVERY, R. A.; LINGURARU, M. G. Deep learning guided partitioned shape model for anterior visual pathway segmentation. **IEEE Transactions on Medical Imaging**, v. 35, n. 8, p. 1856–1865, 2016.

MARTELLI, A. Edge detection using heuristic search methods. **Computer Graphics and Image Processing**, v. 1, n. 2, p. 169 – 182, 1972.

MARTIN, D.; FOWLKES, C.; TAL, D.; MALIK, J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: **Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001**. Vancouver, Canada: IEEE, 2001. v. 2, p. 416–423.

MARTÍNEZ-PÉREZ, M. E.; HUGHES, A. D.; STANTON, A. V.; THOM, S. A.; BHARATH, A. A.; PARKER, K. H. Scale-space analysis for the characterisation of retinal blood vessels. In: **Proceedings of the Conference on Medical Image Understanding and Analysis (MIUA'99)**. Oxford, England: C. Taylor and A. Colchester, eds, 1999. p. 57–60.

MATTHEWS, B. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. **Biochimica et Biophysica Acta (BBA) - Protein Structure**, v. 405, n. 2, p. 442 – 451, 1975.

MAURER, C. R.; QI, R.; RAGHAVAN, V.; MEMBER, S. A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, p. 265–270, 2003.

MEILA, M. Comparing clusterings-an information based distance. **Comparing clusterings**, v. 98, n. 5, p. 873–895, 2007.

MITCHELL, H. **Data Fusion: Concepts and Ideas**. London: Springer Berlin Heidelberg, 2012.

MOESKOPS, P.; VIERGEVER, M. A.; MENDRIK, A. M.; VRIES, L. S. de; BENDERS, M. J. N. L.; IŞGUM, I. Automatic segmentation of mr brain images with a convolutional neural network. **IEEE Transactions on Medical Imaging**, v. 35, n. 5, p. 1252–1261, 2016.

MORTENSEN, E.; MORSE, B.; BARRETT, W.; UDUPA, J. Adaptive boundary detection using ‘live-wire’ two-dimensional dynamic programming. In: **Proceedings Computers in Cardiology**. Durham, NC: Brigham Young University. Department of Computer Science., 1992. p. 635–638.

MORTENSEN, E. N.; BARRETT, W. A. Intelligent scissors for image composition. In: **Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques**. New York, NY, USA: ACM, 1995. (SIGGRAPH '95), p. 191–198. ISBN 0-89791-701-4.

OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE Computer Society, v. 24, n. 7, p. 971–987, jul. 2002.

- OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, v. 29, n. 1, p. 51–59, 1996.
- OLABARRIAGA, S. D.; SMEULDERS, A. W. M. Interaction in the segmentation of medical images: A survey. **Medical Image Analysis**, v. 5, n. 2, p. 127–142, 2001.
- ORGANIZATION, W. H. **Global Status Report on Noncommunicable Diseases 2014**. Geneva, Switzerland: World Health Organization, 2015.
- ORTIZ, A.; GORRIZ, J.; RAMIRES, J.; SALAS-GONZALEZ, D. Mr brain image segmentation by growing hierarchical som and probability clustering. **Electronics Letters**, v. 47, n. 10, p. 585–586, 2011.
- OTSU, N. A threshold selection method from gray-level histograms. **IEEE Transactions on Systems, Man and Cybernetics**, v. 9, n. 1, p. 62–66, jan. 1979.
- PAUMARD, J. Robust comparison of binary images. **Pattern Recognition Letters**, v. 18, n. 10, p. 1057–1063, 1997.
- PENG, P.; LEKADIR, K.; GOOYA, A.; SHAO, L.; PETERSEN, S. E.; FRANGI, A. F. A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. **Magnetic Resonance Materials in Physics, Biology and Medicine**, v. 29, n. 2, p. 155–195, 2016.
- PEREIRA, S.; PINTO, A.; ALVES, V.; SILVA, C. A. Brain tumor segmentation using convolutional neural networks in mri images. **IEEE Transactions on Medical Imaging**, v. 35, n. 5, p. 1240–1251, May 2016.
- PETER, Z.; BOUSSON, V.; BERGOT, C.; PEYRIN, F. A constrained region growing approach based on watershed for the segmentation of low contrast structures in bone micro-ct images. **Pattern Recognition**, v. 41, n. 7, p. 2358–2368, 2008.
- PETROU, M.; PETROU, C. **Image Processing: The Fundamentals**. United Kingdom: John Wiley & Sons, 2010.
- PHAM, D. L.; XU, C.; PRINCE, J. L. A survey of current methods in medical image segmentation. **Annual Review of Biomedical Engineering**, v. 2, p. 315–337, 2000.
- QIU, W.; YUAN, J.; UKWATTA, E.; SUN, Y.; RAJCHL, M.; FENSTER, A. Prostate segmentation: An efficient convex optimization approach with axial symmetry using 3-d trus and mr images. **IEEE Transactions on Medical Imaging**, v. 33, n. 4, p. 947–960, April 2014.
- RADAU, P.; LU, Y.; CONNELLY, K.; PAUL, G.; DICK, A.; WRIGHT, G. Evaluation framework for algorithms segmenting short axis cardiac mri. **The MIDAS Journal**, 07 2009.
- RAND, W. Objective criteria for the evaluation of clustering methods. **Journal of the American Statistical Association**, v. 66, n. 336, p. 846–850, 1971.
- RAYLMAN, R. R.; STOLIN, A. V.; SOMPALLI, P.; RANDALL, N. B.; MARTONE, P. F.; CLINTHORNE, N. H. Investigation of a dedicated, high resolution pet/ct scanner for staging and treatment planning of head and neck cancer. **IEEE Transactions on Nuclear Science**, v. 62, n. 5, p. 2057–2066, Oct 2015.

- RODRÍGUEZ, J. H.; CONDE, M. J. R.; FRAILE, F. J. C. Artificial neural networks applications in computer aided diagnosis: System design and use as an educational tool. In: **Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality**. New York, NY, USA: ACM, 2016. (TEEM '16), p. 1201–1208.
- ROERDINK, J. B.; MEIJSTER, A. The watershed transform: Definitions, algorithms and parallelization strategies. **Fundamenta Informaticae**, IOS Press, v. 41, n. 1,2, p. 187–228, abr. 2000.
- ROGOWSKA, J. Overview and fundamentals of medical image segmentation. In: BANKMAN, I. N. (Ed.). **Handbook of Medical Image Processing and Analysis**. Second edition. Burlington: Academic Press, 2009. p. 73 – 90.
- ROSENFELD, A.; KAK, A. C. **Digital Picture Processing**. 2nd. ed. New York: Academic Press, Inc., 1982. ISBN 0125973020.
- ROYCHOWDHURY, S.; KOOZEKANANI, D. D.; PARHI, K. K. Iterative vessel segmentation of fundus images. **IEEE Transactions on Biomedical Engineering**, v. 62, n. 7, p. 1738–1749, 2015.
- SABUNCU, M. R.; YEO, B. T. T.; LEEMPUT, K. V.; FISCHL, B.; GOLLAND, P. A generative model for image segmentation based on label fusion. **IEEE Transaction on Medical Imaging**, v. 29, n. 10, p. 1714–1729, 2010.
- SCHERTZER, D.; LOVEJOY, S. Nonlinear variability in geophysics: Multifractal simulations and analysis. In: PIETRONERO, L. (Ed.). **Fractals Physical Origin and Properties**. Boston, MA: Springer US, 1989. p. 49–79.
- SCHMID, J.; NIJDAM, N.; HAN, S.; KIM, J.; MAGNENAT-THALMANN, N. Modelling the physiological human: 3d physiological human workshop. In: MAGNENAT-THALMANN, N. (Ed.). Zermatt, Switzerland: Springer Berlin Heidelberg, 2009. cap. Interactive Segmentation of Volumetric Medical Images for Collaborative Telemedicine, p. 13–24.
- SEABORN, K.; FELS, D. I. Gamification in theory and action. **Int. J. Hum.-Comput. Stud.**, v. 74, n. C, p. 14–31, fev. 2015.
- SEUNG, S. **EyeWire**. 2016. Disponível em: <http://eyewire.org/>.
- SKOUNAKIS, E.; BANITSAS, K.; BADII, A.; TZOULAKIS, S.; MARAVELAKIS, E.; KONSTANTARAS, A. Atd: A multiplatform for semiautomatic 3-d detection of kidneys and their pathology in real time. **IEEE Transactions on Human-Machine Systems**, v. 44, n. 1, p. 146–153, 2014.
- SMITH, N.; WEBB, A. **Introduction to Medical Imaging: Physics, Engineering and Clinical Applications**. Cambridge: Cambridge University Press, 2010. (Cambridge Texts in Biomedical Engineering).
- SOARES, J. V. B.; LE, J. J. G.; CESAR, R. M.; JELINEK, H. F.; CREE, M. J.; MEMBER, S. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. **IEEE Transactions on Medical Imaging**, v. 25, p. 1214–1222, 2006.

- SOARES, J. V. B.; LEANDRO, J. J. G.; CESAR, R. M.; JELINEK, H. F.; CREE, M. J. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. **IEEE Transactions on Medical Imaging**, v. 25, p. 1214–1222, 2006.
- SOARES, J. V. B.; LEANDRO, J. J. G.; JÚNIOR, R. M. C.; JELINEK, H. F.; CREE, M. J. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. **IEEE Transactions on Medical Imaging**, v. 25, n. 9, p. 1214 –1222, 2006.
- STAAL, J.; ABRAMOFF, M.; NIEMEIJER, M.; VIERGEVER, M.; GINNEKEN, B. van. Ridge based vessel segmentation in color images of the retina. **IEEE Transactions on Medical Imaging**, v. 23, n. 4, p. 501–509, 2004.
- SUINESIAPUTRA, A.; COWAN, B. R.; AL-AGAMY, A. O.; ELATTAR, M. A.; AYACHE, N.; FAHMY, A. S.; KHALIFA, A. M.; MEDRANO-GRACIA, P.; JOLLY, M.; KADISH, A. H.; LEE, D. C.; MARGETA, J.; WARFIELD, S. K.; YOUNG, A. A. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. **Medical Image Analysis**, v. 18, n. 1, p. 50–62, 2014.
- SUN, S.; BAUER, C.; BEICHEL, R. Automated 3-d segmentation of lungs with lung cancer in ct data using a novel robust active shape model approach. **IEEE Transactions on Medical Imaging**, v. 31, n. 2, p. 449–460, 2012.
- SURULIANDI, A.; MEENA, K.; ROSE, R. R. Local binary pattern and its derivatives for face recognition. **IET Computer Vision**, v. 6, n. 5, p. 480–488, 2012.
- TAMEZ-PEÑA, J. G.; FARBER, J.; GONZÁLEZ, P. C.; SCHREYER, E.; SCHNEIDER, E.; TOTTERMAN, S. Unsupervised segmentation and quantification of anatomical knee features: Data from the osteoarthritis initiative. **IEEE Transactions on Biomedical Engineering**, v. 59, n. 4, p. 1177–1186, April 2012.
- TOMASHEFSKI, J. F.; FARVER, C. F. Anatomy and histology of the lung. In: TOMASHEFSKI, J. F.; CAGLE, P. T.; FARVER, C. F.; FRAIRE, A. E. (Ed.). **Dail and Hammar's Pulmonary Pathology: Volume I: Nonneoplastic Lung Disease**. New York, NY: Springer New York, 2008. p. 20–48.
- TORBATI, N.; AYATOLLAHI, A.; KERMANI, A. An efficient neural network based method for medical image segmentation. **Computers in Biology and Medicine**, v. 44, n. 5, p. 76 – 87, 2014.
- UDUPA, J. K.; SAMARASEKERA, S.; BARRETT, W. A. Boundary detection via dynamic programming. **Proc. SPIE**, v. 1808, p. 33–39, 1992.
- VEREDAS, F.; MESA, H.; MORENTE, L. Binary tissue classification on wound images with neural networks and bayesian classifiers. **IEEE Transactions on Medical Imaging**, v. 29, n. 2, p. 410–427, Feb 2010.
- VERGÉS-LLAHÍ, J.; SANFELIU, A. Evaluation of distances between color image segmentations. In: **Proceedings of the Second Iberian Conference on Pattern Recognition and Image Analysis - Volume Part II**. Berlin, Heidelberg: Springer-Verlag, 2005. (IbPRIA'05), p. 263–270.
- WANG, Y. **Principles of Magnetic Resonance Imaging: Physics Concepts, Pulse Sequences, & Biomedical Applications**. New York, USA: CreateSpace Independent Publishing Platform, 2012. ISBN 9781479350414.

-
- WARFIELD, S.; ZOU, K.; WELLS, W. M. Simultaneous truth and performance level estimation (staple): An algorithm for the validation of image segmentation. **IEEE Transaction on Medical Imaging**, v. 23, n. 7, p. 903–921, 07 2004.
- WU, H. S.; BARBA, J.; GIL, J. Region growing segmentation of textured cell images. **Electronics Letters**, v. 32, n. 12, p. 1084–1085, 1996.
- XIE, J.; JIANG, Y.; TSUI, H. Segmentation of kidney from ultrasound images based on texture and shape priors. **IEEE Transactions on Medical Imaging**, v. 24, n. 1, p. 45–57, 2005.
- XING, F.; XIE, Y.; YANG, L. An automatic learning-based framework for robust nucleus segmentation. **IEEE Transactions on Medical Imaging**, v. 35, n. 2, p. 550–566, 2016.
- YAZDANPANA, A.; HAMARNEH, G.; SMITH, B. R.; SARUNIC, M. V. Segmentation of intra-retinal layers from optical coherence tomography images using an active contour approach. **IEEE Transactions on Medical Imaging**, v. 30, n. 2, p. 484–496, 2011.
- YI, X.; ERAMIAN, M. LBP-based segmentation of defocus blur. **IEEE Transactions on Image Processing**, v. 25, n. 4, p. 1626–1638, 2016.
- YOUSEFI, S.; KEHTARNAVAZ, N.; GHOLIPOUR, A. Improved labeling of subcortical brain structures in atlas-based segmentation of magnetic resonance images. **IEEE Transactions on Biomedical Engineering**, v. 59, n. 7, p. 1808–1817, July 2012.
- ZANA, F.; KLEIN, J. C. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. **IEEE Transactions on Image Processing**, v. 10, n. 7, p. 1010–1019, 2001.
- ZHANG, B.; GAO, Y.; ZHAO, S.; LIU, J. Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. **IEEE Transactions on Image Processing**, v. 19, n. 2, p. 533–544, 2010.
- ZHANG, H.; FRITTS, J. E.; GOLDMAN, S. A. Image segmentation evaluation: A survey of unsupervised methods. **Computer Vision and Image Understanding**, v. 110, n. 2, p. 260–280, maio 2008.
- ZHAO, F.; XIE, X. An overview of interactive medical image segmentation. In: **Annals of the BMVA**. Swansea, United Kingdom: IEEE, 2013. p. 1–22.
- ZIJDENBOS, A. P.; DAWANT, B. M.; MARGOLIN, R. A.; PALMER, A. C. Morphometric analysis of white matter lesions in MR images: method and validation. **IEEE Transactions on Medical Imaging**, v. 13, n. 4, dez. 1994.

APÊNDICES

APÊNDICE A – ÁREA DE TRABALHO DO COSE

A interface de usuário é simples e intuitiva. Após se cadastrar e logar no sistema, o segmentador visualiza uma imagem a ser segmentada e um painel de controle no lado esquerdo da tela, como pode ser visto na Figura 56. O segmentador pode escolher entre a ferramenta inteligente de segmentação (chamada de *smart tool*) ou pode optar por realizar a segmentação manualmente.

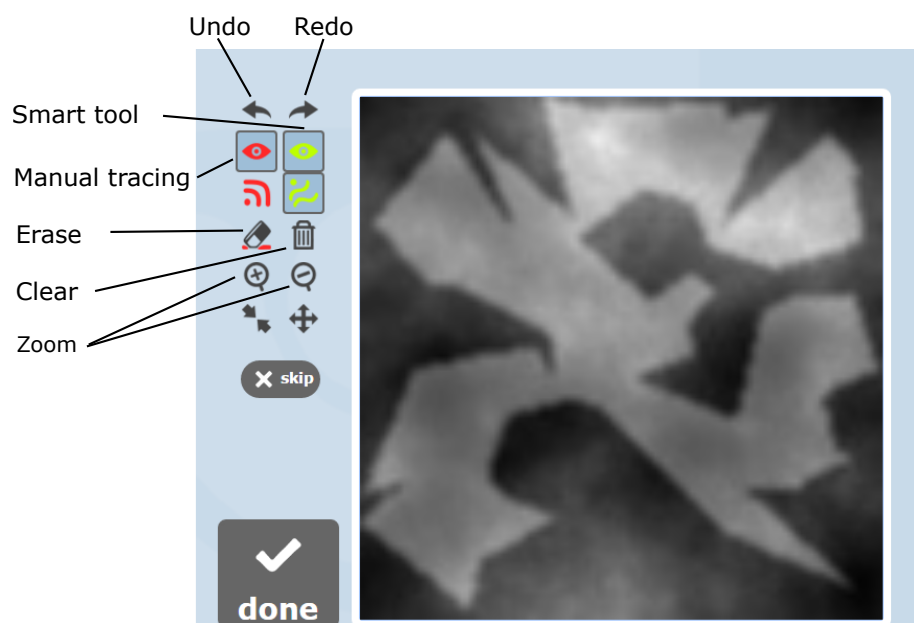


Figura 56: Visualização da área de trabalho com suas ferramentas: *undo*, *redo*, *smart tool*, *manual tracing*, *erase*, *clear* e *zoom*. A imagem a ser segmentada corresponde a maior parte da área de trabalho.

Os botões *done* e *skip* fornecem ao usuário a possibilidade de trabalhar no próximo retalho. Após analisar e concluir o trabalho de segmentação no retalho atualmente visível na tela, o usuário pode clicar no botão *done*, fazendo com que o sistema armazene o resultado da segmentação e apresente um novo retalho a ser segmentado. O usuário pode utilizar o botão *skip* quando julgar não ser capaz de segmentar a imagem, seja por considerá-la difícil ou por achar que a imagem não possui informações suficientes. Uma imagem pode ser considerada com informações insuficientes quando as bordas não estão bem definidas, ou quando o baixo contraste entre o objeto e o *background* dificulta a visualização do segmentador, ou ainda quando a imagem apresenta bastante ruído. Ao pressionar o botão *skip*, o usuário abdica de resolver o retalho e o sistema apresenta uma nova imagem. Para realizar uma alteração na segmentação da imagem atualmente visível no dispositivo, o usuário conta com 4 botões: *undo*, *redo*, *clear* e *erase*. Os botões *undo* e *redo* são responsáveis por desfazer e refazer uma ação no processo de segmentação,

respectivamente. O botão *clear* apaga totalmente a segmentação feita na imagem. Se o segmentador selecionar a ferramenta *erase*, ele poderá apagar partes específicas da segmentação. O usuário tem ainda a opção de *zoom*, que aumenta ou diminui a imagem, facilitando a visualização.

APÊNDICE B – ARQUITETURA DO COSE

O sistema foi desenvolvido utilizando Django, um *framework* da linguagem de programação Python para o desenvolvimento *web*. Os algoritmos de segmentação interativa foram implementados em JavaScript. Desta forma, os algoritmos são executados no *browser* do usuário do sistema descentralizando o processamento.

A Figura 57 exibe o diagrama de classes que compõem o sistema. Problemas de segmentação (*SegmentationProblem*) são agrupados em uma tarefa (*Task*) que, por sua vez, são agrupadas em uma categoria de tarefas (*TaskCategory*). A segmentação do meristema da Arabdopsis Thaliana, organismo modelo no estudo de plantas, é um exemplo de categoria de tarefas. Diferentes conjuntos de imagens de meristema podem corresponder a diferentes tarefas. Cada tarefa pode conter um conjunto de problemas de segmentação que correspondem a diferentes seções transversais de um mesmo meristema. Os parâmetros de um problema de segmentação são representados pela classe *SegmentationProblemDetails*.

Ao cadastrar um problema de segmentação, o administrador escolhe um conjunto de parâmetros incluindo o tamanho do retalho. Em seguida, deve criar os retalhos através da opção *Create* como pode ser visto na Figura 57. Cada instância da classe *Assignment* possui um retalho (*tile*) distinto da imagem, informações como a posição do retalho na imagem original (*tile_bbox_x0*, *tile_bbox_y0*, etc.), uma indicação se o retalho possui ou não conteúdo significativo (*workable*) e uma indicação se o retalho foi resolvido (*concluded*). O atributo *concluded* é verdadeiro quando o retalho é segmentado por n usuários. O valor de n é definido pelo atributo *min_results_per_assignment* do problema de segmentação. O atributo *merge* representa a combinação das segmentações dos n usuários.

Cada instância da classe *AssignmentSession* representa o estado da sessão de um usuário. O atributo *result* é o resultado da segmentação produzido pelo usuário. Este resultado será combinado a outros resultados para compor uma combinação de segmentações (atributo *merge* da classe *Assignment*). A classe *User* representa tanto o usuário como o administrador do sistema. Os atributos *is_staff* e *is_superuser* definem se um usuário tem permissão de administrador e, portanto, permissão para criar tarefas e cadastrar problemas de segmentação. Um usuário sem permissão de administrador não pode cadastrar problemas de segmentação.

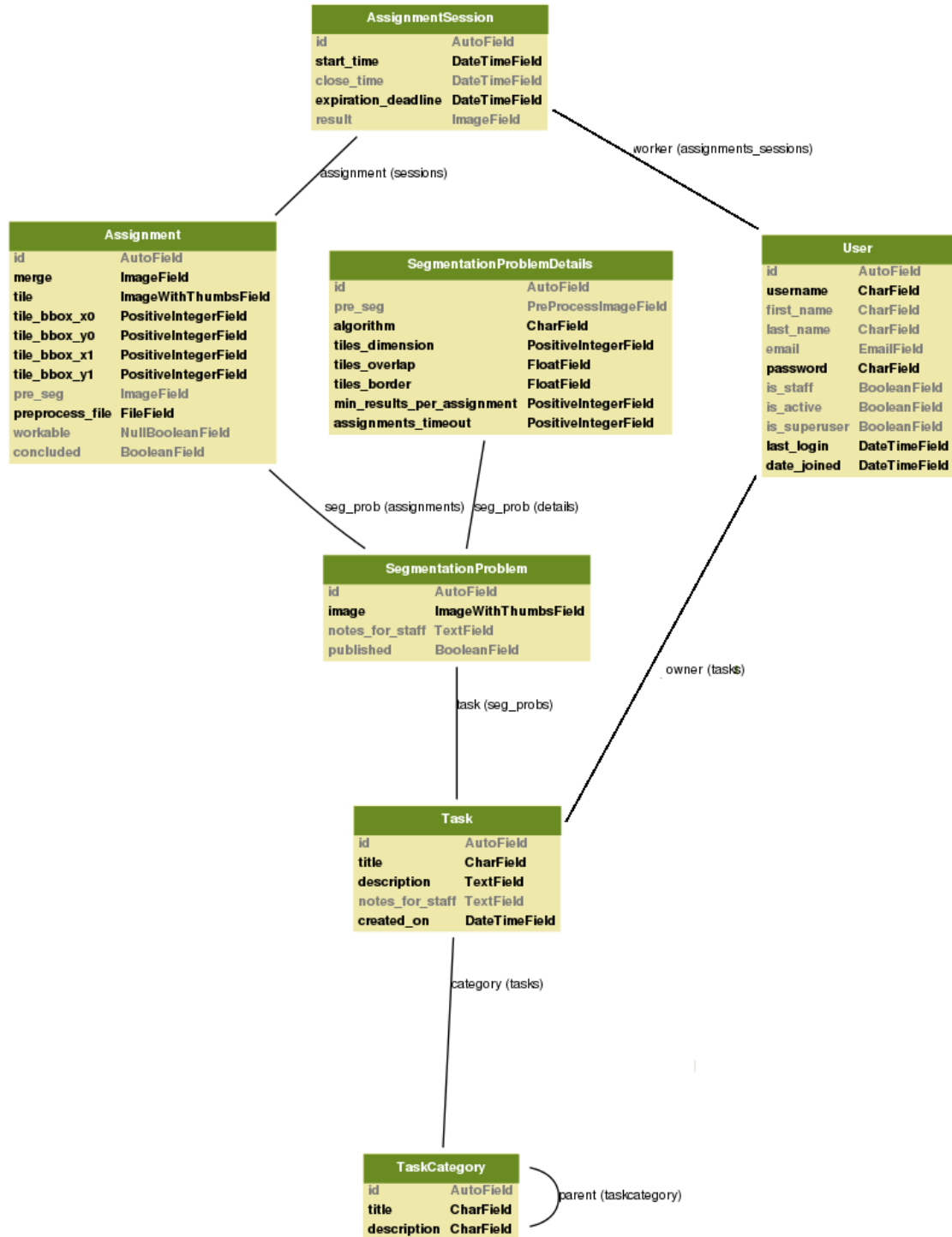


Figura 57: Diagrama de classes do COSE.

APÊNDICE C – FLUXO DE EXECUÇÃO DO COSE

A Figura 56 mostra a sequência de execução em telas desde o cadastro e uma tarefa até a segmentação final de um retalho. O processo inicia com o administrador cadastrando uma tarefa (58b) e um problema de segmentação de uma imagem da retina (58c e 58d). Em seguida, o administrador pode criar retalhos dessa imagem (57e). Os parâmetros para a criação dos retalhos são:

- *Algorithm*: O algoritmo iterativo que será utilizado para auxiliar na segmentação dos retalhos;
- *Tiles dimension*: O tamanho quadrado dos retalhos;
- *Tiles Overlap*: Porcentagem da área de sobreposição de retalhos em relação ao tamanho dos retalhos;
- *Tiles border*: Porcentagem da borda dos retalhos em relação ao tamanho dos retalhos;
- *Min. results per assignments*: Número mínimos de segmentações para que o trabalho com um determinado retalho seja dado como concluído;
- *Assignments timeout*: Tempo máximo para que um usuário possa segmentar o retalho.

A Figura 57f mostra os retalhos criados para o problema de segmentação. Nessa tela é possível ver: se o retalho está como *workable* (ativo para segmentação dos usuários), a quantidade de segmentações cada retalho possui (*sessions*), se o retalho está concluído (*concluded?*) e o resultado do consenso usando o algoritmo DTM (*result*). A Figura 57g mostra um exemplo de retalho sem vasos sanguíneos para segmentar. Nesse caso, o usuário pode clicar no botão *skip*. Já a Figura 56h apresenta o início da segmentação em um retalho. O resultado da segmentação desse retalho é mostrado na Figura 56i.

Collaborative Segmentation

Site administration

Auth	
Groups	+ Add Change
Users	+ Add Change
Crowd	
Assignment sessions	Change
Assignments	Change
Segmentation problems	Change
Task categories	+ Add Change
Tasks	+ Add Change
Registration	
Registration profiles	+ Add Change
Sites	
Sites	+ Add Change

(a)

Collaborative Segmentation

Home > Crowd > Tasks > Add task

Add task

Title:

Description:

Category: [+](#)

Notes for staff:

Created on: (None)

Owner: (None)

(b)

Collaborative Segmentation

Home > Crowd > Tasks > Retinal image

Welcome, **Rafael**. [Change password](#) / [Log out](#)

Change task

Title: Retinal image

Description:

Category: Retinal Vessels

Notes for staff:

Created on: April 23, 2017, 6:32 p.m.

Owner: barreto

Segmentation problems	Notes for staff	Has assignments?	Published?	Delete?
Segmentation problems Image Escolher arquivo Nenhum arquivo selecionado				

[+ Add another Segmentation Problem](#)

[Delete](#) [Save and add another](#) [Save and continue editing](#) [Save](#)

(c)

Collaborative Segmentation

Home > Crowd > Tasks > Retinal image

Welcome, **Rafael**. [Change password](#) / [Log out](#)

Change task

Title: Retinal image

Description:

Category: Retinal Vessels

Notes for staff:

Created on: April 23, 2017, 6:32 p.m.

Owner: barreto

Segmentation problems	Notes for staff	Has assignments?	Published?	Delete?
Segmentation problems Image Escolher arquivo Nenhum arquivo selecionado				

[+ Add another Segmentation Problem](#)

[Delete](#) [Save and add another](#) [Save and continue editing](#) [Save](#)

(d)

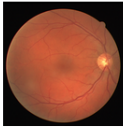
Collaborative Segmentation Welcome, Rafael. Change password / Log out

Home > Crowd > Segmentation problems > segmentation problem 1073 (task: Retinal image)

Change segmentation problem

[Delete](#) [Save and continue editing](#) [Save](#)

ID: 1073

Preview: 

Task: Retinal image

Status

Has details?: ●

Assignments: None

Result: N/A

☐ Published?

Notes for staff (Show)

Segmentation Problems Details

Not specified yet

Algorithm: LiveVessel

Pre seg.: Esoterio arguio Nenhum arguio... selecionado

Tiles dimension: 100 Tiles overlap: 0.15 Tiles border: 0.25

Min. results per assignment: 5

Assignments timeout: 300

[Delete](#) [Save and continue editing](#) [Save](#)

(e)

Collaborative Segmentation Welcome, Rafael. Change password / Log out

Home > Crowd > Assignments

Select assignment to change

Action: All Go 0 of 100 selected

ID	Preview	Workable?	Sessions	Concluded?	Result	Segmentation problem	Task
324		Yes	N/A	●	N/A		Retinal image
323		Yes	N/A	●	N/A		Retinal image
322		Yes	N/A	●	N/A		Retinal image
321		Yes	N/A	●	N/A		Retinal image
320		Yes	N/A	●	N/A		Retinal image
319		Yes	N/A	●	N/A		Retinal image
318		Yes	N/A	●	N/A		Retinal image
317		Yes	N/A	●	N/A		Retinal image

Filter

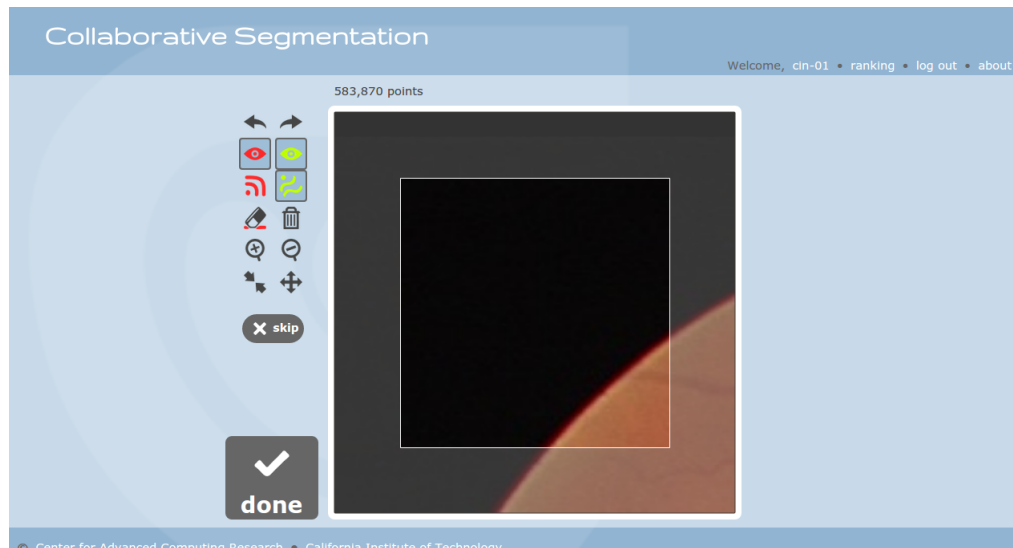
By workable? All

By concluded? All

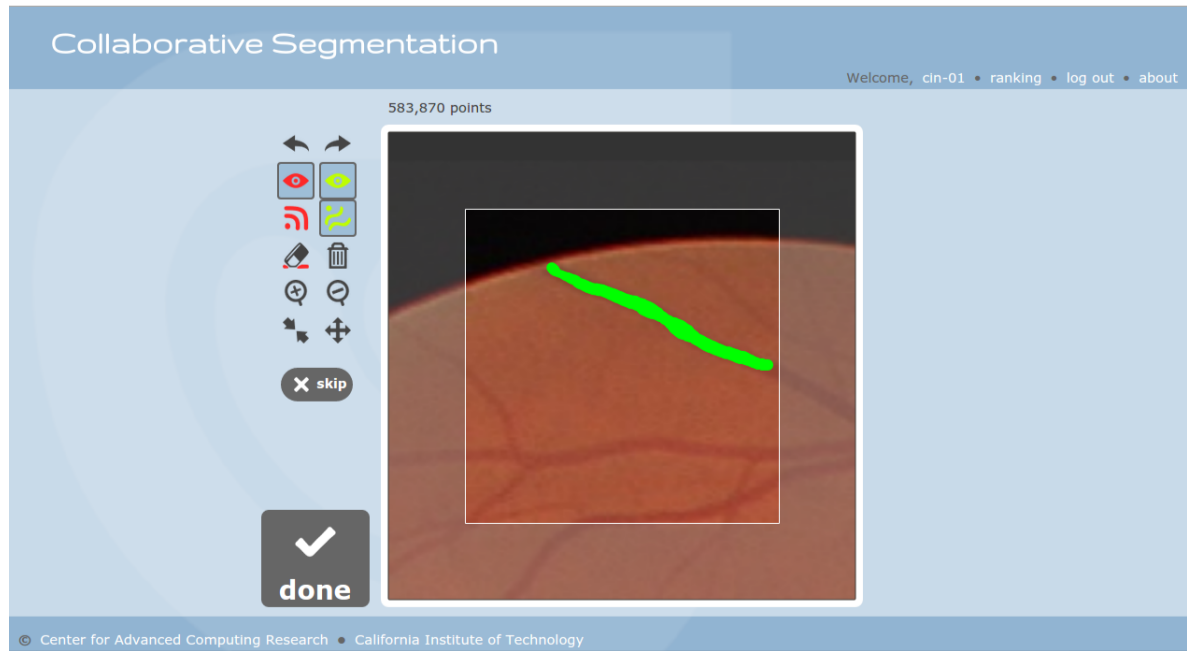
By task All

By seg. problem All

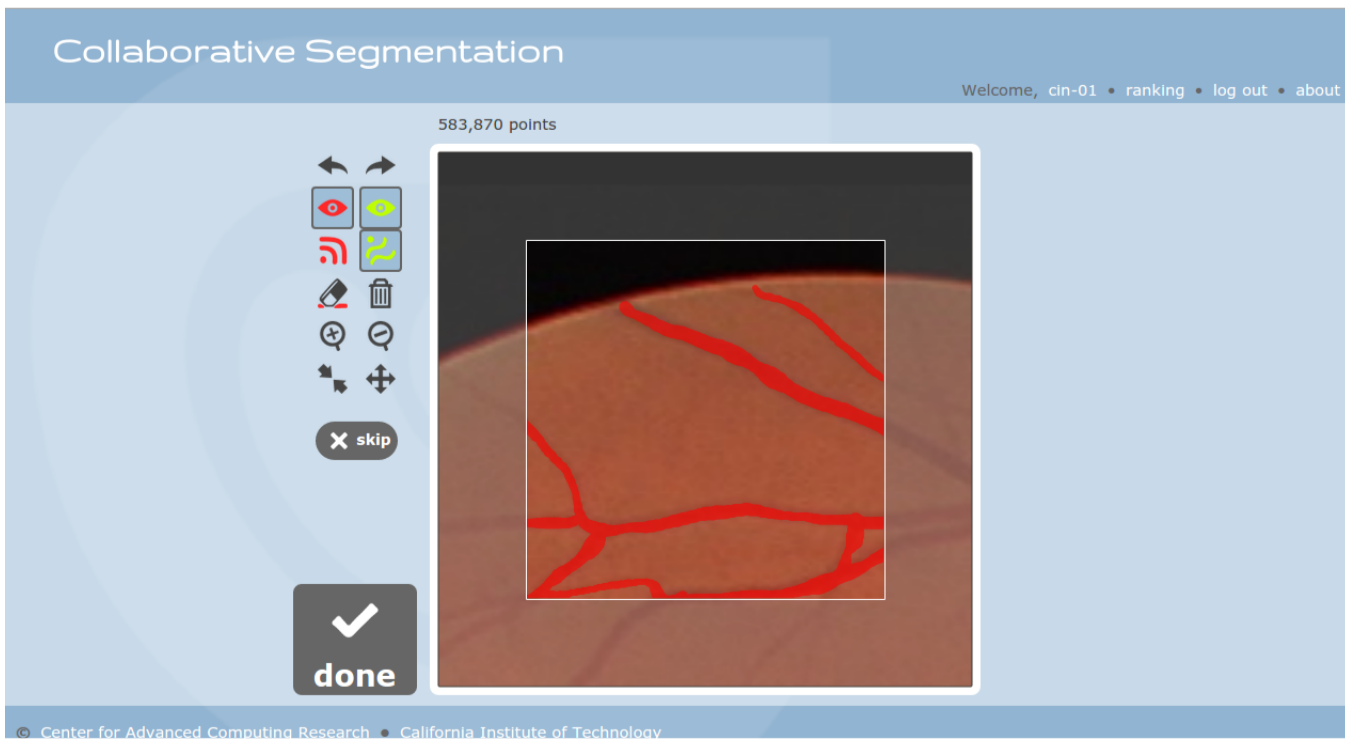
(f)



(g)



(h)



(i)

Figura 56: Fluxo de execução do COSE em telas: (a) Tela administrativa inicial; (b) Cadastro de tarefas; (c) Inserção de um problema de segmentação a uma tarefa; (d) Problema de segmentação após a inserção; (e) Criação dos retalhos; (f) Retalhos criados; (g) Retalho com pouca informações; (h) Início da segmentação utilizando o live-vessel; (h) Segmentação final para o retalho.