



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

William Javier Garcia Herrera

**METODOLOGIA DE AVALIAÇÃO DE
SEGMENTAÇÕES 2D DO CORPO CALOSO
EM IMAGENS DE RESSONÂNCIA
MAGNÉTICA DE DIFUSÃO ATRAVÉS DO
PERFIL DE CURVATURA**

Campinas

2017



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

William Javier Garcia Herrera

METODOLOGIA DE AVALIAÇÃO DE SEGMENTAÇÕES 2D DO CORPO CALOSO EM IMAGENS DE RESSONÂNCIA MAGNÉTICA DE DIFUSÃO ATRAVÉS DO PERFIL DE CURVATURA

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica, na Área de Engenharia de Computação.

Orientadora: Prof^ª Dra. Leticia Rittner

Este exemplar corresponde à versão final da tese defendida pelo aluno William Javier Garcia Herrera, e orientada pela Prof^ª Dra. Leticia Rittner

Campinas

2017

Agência(s) de fomento e nº(s) de processo(s): CNPq, 190557/2014-1

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Luciana Pietrosanto Milla - CRB 8/8129

G165m Garcia Herrera, William Javier, 1985-
Metodologia de avaliação de segmentações 2D do corpo caloso em
imagens de ressonância magnética de difusão através do perfil de curvatura /
William Javier Garcia Herrera. – Campinas, SP : [s.n.], 2017.

Orientador: Leticia Rittner.

Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade
de Engenharia Elétrica e de Computação.

1. Imagem de difusão por ressonância magnética. 2. Curvatura. 3.
Segmentação de imagens. 4. Padrão-ouro. 5. Corpo caloso. I. Rittner, Leticia.
II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de
Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: 2D Segmentation evaluation method of the corpus callosum in
diffusion MRI using corpus callosum signature

Palavras-chave em inglês:

Magnetic resonance diffusion imaging

Curvature

Image segmentation

Golden pattern

Corpus callosum

Área de concentração: Engenharia de Computação

Titulação: Mestre em Engenharia Elétrica

Banca examinadora:

Leticia Rittner [Orientador]

Helio Pedrini

Rafael Ferrari

Data de defesa: 23-05-2017

Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - DISSERTAÇÃO DE MESTRADO

Candidato: William Javier Garcia Herrera, **RA:** 162642

Data da defesa: 23 de maio de 2017.

Título da dissertação: Metodologia de avaliação de segmentações 2D do corpo caloso em imagens de ressonância magnética de difusão através do perfil de curvatura

Comissão Julgadora:

Prof^a Dra. Leticia Rittner (Presidente, FEEC/UNICAMP)

Prof^a Dr. Helio Pedrini (IC/UNICAMP)

Dr. Rafael Ferrari (FEEC/UNICAMP)

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no processo de vida acadêmica do aluno.

Agradecimentos

Quem já cursou um programa de mestrado sabe quão difícil é o caminho: noites sem dormir, leituras e escritas intermináveis, mais perguntas que respostas, pressão e estresse. Não é vanidade se reconhecer como responsável de um trabalho como este. Mas também deve se reconhecer a contribuição, tanto acadêmica quanto espiritual, das pessoas que fazem parte desta vitória.

Em primeiro lugar claro está, a minha família, meu suporte vital: minha mãe, lutadora incansável e maior contribuinte de todo quanto tenha feito e farei. Aos meus irmãos, amigos de vida, lhes devo admiração e um carinho imenso; seus papos, ideias e aspirações alimentam meus sonhos e trouxeram me até este ponto.

Agradeço a minha orientadora, Leticia Rittner, mulher à que admiro muito, por ser a minha guia na pesquisa. Agradeço sua paciência, sua prolixidade na excelência e seus conselhos. Também para meus colegas de laboratório um agradecimento por me aceitar como parte de seu time, me ensinar e apoiar nas tarefas da vida acadêmica e me deixar compartilhar experiências maravilhosas.

Finalmente, todos os amigos que tenho feito no Brasil, de várias nacionalidades, naturezas e saberes. Cada um deles tem aumentado meu carinho pelo Brasil e tem me dado razões demais para acreditar que o mundo vai para uma condição melhor.

Não posso deixar de reconhecer o apoio econômico do Programa de Estudantes-Convênio de Pós-Graduação (PEC-PG), da CAPES/CNPq-Brasil.

Resumo

O corpo caloso (CC) é a maior estrutura de substância branca no cérebro, está localizada sob o córtex cerebral e conecta os dois hemisférios cerebrais, servindo de ponte de comunicação entre eles. É, portanto, uma estrutura de grande interesse no âmbito médico e de pesquisa. Sua forma e tamanho estão associadas com algumas características do sujeito e alterações na sua estrutura apresentam correlação com várias doenças e condições médicas.

Nas imagens de difusão por ressonância magnética, a segmentação desta estrutura é importante já que a informação contida neste tipo de imagem permite estudar a microestrutura das fibras neuronais e os tecidos usando o modelo de difusão da água. Na literatura existem poucos métodos de segmentação do CC baseados em imagens de difusão por ressonância magnética e não existem estudos sobre avaliação quantitativa de segmentações neste espaço.

A avaliação de segmentações em difusão é feita normalmente usando um padrão-ouro obtido manualmente sobre imagens de ressonância ponderadas em T_1 e registrado no espaço de difusão. Porém, o registro é computacionalmente custoso e introduz erros no padrão final. Outros padrões podem ser construídos diretamente no espaço de difusão, porém também apresentam desvantagens. A avaliação quantitativa, neste caso, é feita usando uma métrica de sobreposição.

Com o propósito de melhorar o esquema de avaliação usual por sobreposição, neste trabalho é proposto um método de avaliação que permite usar diretamente o padrão obtido manualmente em T_1 sem ser necessário realizar o registro para o espaço de difusão. Este método está baseado no perfil de curvatura, um descritor que permite comparar segmentações através da forma, sem necessidade de sobreposição ou registro de imagens.

O método proposto foi usado para avaliar segmentações em difusão obtidas através de três métodos distintos, em 145 sujeitos. A raiz do erro médio quadrático (RMSE), calculado a partir da comparação entre os perfis de curvatura, mostrou-se uma métrica complementar ao coeficiente Dice e apresentou capacidade para discriminar segmentações. Para exploração em trabalhos futuros, o perfil de curvatura pode ser usado para identificação automática de segmentações incorretas em grandes bases de dados, estudos populacionais e longitudinais e caracterização de outras formas e estruturas.

Palavras-chave: Imagem de difusão por ressonância magnética, Curvatura, Segmentação de imagens, Padrão-ouro, Corpo caloso

Abstract

Corpus callosum (CC) is the greatest white matter structure in brain. It is located beneath the cortex and connects both of two hemispheres, making possible their communication. Therefore, CC is important in medical and academic scene. CC's shape and size are associated with some subject's characteristics and alterations in its structure have correlation with some diseases and medical conditions.

Diffusion MRI makes possible study of the neuronal fibers and tissues micro structure using the water diffusion model. In the literature, there are few CC segmentation methods based-on Diffusion MRI and there are not studies related with segmentation quantitative evaluation in this model.

Segmentation evaluation in diffusion is commonly performed using registered gold-standard or any standard draw directly in this modality. However, both of the two standards have problems because of the registering process itself, that introduce error in the final standard, or the drawing process on the diffusion images. Quantitative evaluation is done using overlap metric.

In this work, an evaluation method is proposed making possible direct use of gold-standard without any register process. This methodology is based on CC signature, a descriptor for comparing segmentations using shape, without overlap requirement with standard.

The segmentation evaluation method proposed was used in diffusion for quantitative assessment of 145 subjects using the gold-standard through CC signature. The RMSE metric, based on CC signature, showed to be complementary with Dice coefficient and capability for differentiating segmentations. For future work, the CC signature, used as characterization tool of the CC shape, would make possible automatic identification of incorrect segmentations in large databases, longitudinal studies, classification of populations and characterization of other structures.

Key-words: Magnetic resonance diffusion imaging, Curvature, Image segmentation, Golden pattern, Corpus callosum

Lista de ilustrações

1.1	Corpo caloso	14
2.1	Planos usados para descrição do cérebro	19
2.2	Partes do corpo caloso	20
2.3	Cérebro em várias modalidades de MRI	22
2.4	Exemplo de processamento morfológico	25
2.5	Exemplos de elemento estruturante	25
2.6	Erosão e dilatação morfológica	26
2.7	Abertura e fechamento morfológico	27
2.8	Esquema transformada <i>Watershed</i>	28
2.9	FA meio por fatia ao longo do plano sagital	31
3.1	Desafios segmentação CC	34
3.2	Segmentação do CC para um sujeito em MRI	34
3.3	Método usual de avaliação de segmentações	36
3.4	Matriz de confusão	39
3.5	Distância Hausdorff	40
4.1	Método proposto de classificação por pixel	41
4.2	Intensidade DWI para diferentes gradientes	42
5.1	Método proposto de avaliação de segmentações	44
5.2	Esquema de curvatura	45
5.3	Construção do perfil de curvatura	47
5.4	Cálculo RMSE	48
5.5	Perfil de curvatura do CC	51
5.6	RMSE em função das resoluções de casamento e comparação	52
5.7	Soma dos valores de gradiente em função de ls de casamento	53
5.8	Comparação segmentações em função de ls de comparação	55
5.9	Diferença de RMSE em função de ls de comparação	55
6.1	Exemplo de dados usados	58
6.2	Segmentação baseada em <i>Watershed</i>	59
6.3	Erros segmentação baseada em <i>Watershed</i>	59
6.4	Segmentação do ROQS	61
6.5	Erros segmentação ROQS	61
6.6	Segmentação baseada em classificação por pixel	63
6.7	Erros da segmentação baseada em classificação por pixel	63
6.8	Padrão manual registrado no espaço de difusão	65
6.9	Padrão semi-automático no espaço de difusão	65
6.10	Construção padrão STAPLE	66

6.11	Avaliação qualitativa dos padrões-prata em difusão	67
6.12	Avaliação quantitativa dos padrões-prata em difusão	67
6.13	Esquema proposto de avaliação de segmentações	68
6.14	Avaliação segmentações com padrão-ouro	69
6.15	Escala (ampliação)	71
6.16	RMSE sob alteração de escala	71
6.17	RMSE sob alteração de rotação	72
6.18	RMSE sob alteração da estrutura	73
6.19	Comparação Dice e RMSE normalizado	75
6.20	Comparação Dice e RMSE	76
6.21	Exemplos de faixa de operação da RMSE	77
1.1	Ferramenta interativa do perfil de curvatura	91
1.2	ls para padrões manual e semi-automático	93
1.3	Perfil de curvatura para diferentes valores de ls	95
2.1	Esquema de avaliação usual de segmentação	96
2.2	Avaliação segmentações com padrão-prata registrado	97
2.3	Métrica de área em segmentação errada	98
2.4	Avaliação segmentações com padrões-prata semi-automático e STAPLE .	99

Lista de tabelas

3.1	Métricas usadas na literatura	39
5.1	Correlação métricas alternativas com RMSE	49
6.1	Hiper-parâmetros classificador SVM	62

Lista de Abreviações

CC	Corpo caloso
DH	Distância Hausdorff
DTI	Diffusion Tensor Imaging (Imagem por tensor de difusão)
DWI	Diffusion Weighted Imaging (Imagem ponderada em difusão)
FA	Fractional Anisotropy (Anisotropia fracionada)
FMRIB	Oxford Centre for Functional MRI of the Brain (Centro Oxford para MRI funcional do cérebro)
FN	Falso negativo
FP	Falso positivo
FSL	FMRIB Software Library (Biblioteca de software do FMRIB)
GT	Ground-Truth (Padrão-absoluto)
LBP	Local Binary Patterns (Padrões binários locais)
MM	Matriz de magnificação
MRI	Magnetic Resonance Imaging (Imagem por ressonância magnética)
RBF	Radial Basis Function (Função de base radial)
ROQS	Reproducible Objective Quantification Scheme (Esquema de quantificação de objetivo reproduzível)
RMSE	Root Mean Square Error (Raiz do erro médio quadrático)
RSL	Revisão sistemática da literatura
STAPLE	Simultaneous Truth and Performance Level Estimation (Estimador de nível de desempenho e verosimilitude)
SVM	Support Vector Machine (Máquina de vetores suporte)
TR	Time Repetition (Tempo de repetição)
TE	Time Echo (Tempo de eco)
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
k	Curvatura
l_s	Largura de segmento
nCC	Classe não corpo caloso
pCC	Classe corpo caloso

Sumário

Sumário	
1 Introdução	14
1.1 Motivação	15
1.2 Objetivos	16
1.2.1 Objetivo geral	16
1.2.2 Objetivos específicos	17
1.3 Principais contribuições	17
1.4 Organização do trabalho	18
2 Conceitos teóricos	19
2.1 Corpo Caloso (CC)	19
2.2 Imagens por Ressonância Magnética	20
2.2.1 MRI ponderada em T_1	21
2.2.2 MRI ponderada em difusão (DWI)	22
2.2.3 Imagem por tensor de difusão (DTI)	23
2.2.4 Mapa de anisotropia fracionada (FA)	24
2.3 Processamento morfológico de imagens	25
2.3.1 Erosão e dilatação	25
2.3.2 Abertura e fechamento	26
2.3.3 Transformada <i>Watershed</i>	27
2.4 Modelos de aprendizado	27
2.4.1 Máquina de vetores suporte (SVM)	28
2.4.2 <i>k-means</i>	29
2.5 Métodos de segmentação do corpo caloso em MRI	30
2.5.1 Método baseado em <i>Watershed</i>	31
2.5.2 Método de quantificação de objetivo reproduzível	32
3 Método usual de avaliação de segmentações	33
3.1 Segmentação em difusão	36
3.2 Padrão em difusão	36
3.3 Avaliação	38
4 Método proposto de segmentação baseado em classificação por pixel	41
5 Método proposto de avaliação de segmentações	44
5.1 Perfil de curvatura	45
5.2 Construção do perfil de curvatura	46
5.3 Cálculo do RMSE	47
5.4 Definição de parâmetros do perfil de curvatura	49
5.4.1 Definição dos parâmetros do perfil de curvatura	50

5.4.2	Definição do valor de ls para casamento de perfis	51
5.4.3	Definição do valor de ls para comparação de perfis	53
6	Resultados e Discussão	57
6.1	Dados e ferramentas	57
6.2	Segmentações em difusão	58
6.2.1	Segmentação baseada em <i>Watershed</i>	58
6.2.2	Segmentação por ROQS	60
6.2.3	Segmentação usando Classificação por pixel	60
6.3	Padrões	63
6.3.1	Padrão manual em $T1$	64
6.3.2	Padrão manual registrado no espaço de difusão	64
6.3.3	Padrão semi-automático	64
6.3.4	Padrão STAPLE	65
6.3.5	Avaliação dos padrões-prata em difusão	66
6.4	Avaliação de segmentações usando o método proposto	68
6.5	Caracterização do perfil de curvatura e do RMSE	69
6.5.1	Caracterização do perfil de curvatura como descritor	69
6.5.1.1	Escala (Ampliação)	70
6.5.1.2	Rotação	72
6.5.1.3	Ruído (Sal e pimenta)	72
6.5.2	Caracterização do perfil de curvatura e RMSE como métrica	74
7	Conclusões	78
7.1	Trabalhos futuros	79
	Referências	81
	Apêndices	90
	A Análise qualitativa do perfil de curvatura	91
	B Avaliação de segmentações usando o método usual	96

1 Introdução

O corpo caloso (CC) é a maior estrutura de substância branca no cérebro, está localizada sob o córtex cerebral e conecta os dois hemisférios cerebrais, permitindo o intercâmbio de informação entre eles [1] (Fig. 1.1). É, portanto, uma estrutura de grande interesse no âmbito de pesquisa, e sua forma e tamanho estão associadas com algumas características do sujeito como: gênero [100], idade [51], habilidades numéricas e matemáticas [76], lateralidade (uso preferencial de uma mão) [112], entre outras. Além disso, do ponto de vista clínico, alterações do corpo caloso apresentam correlação com várias doenças e condições médicas, entre as quais podemos citar: Alzheimer [35], autismo [101], esquizofrenia [99], dislexia [109], epilepsia [106], esclerose múltipla [24], depressão [110], tabaquismo [34], alcoolismo [80] e obesidade [58].

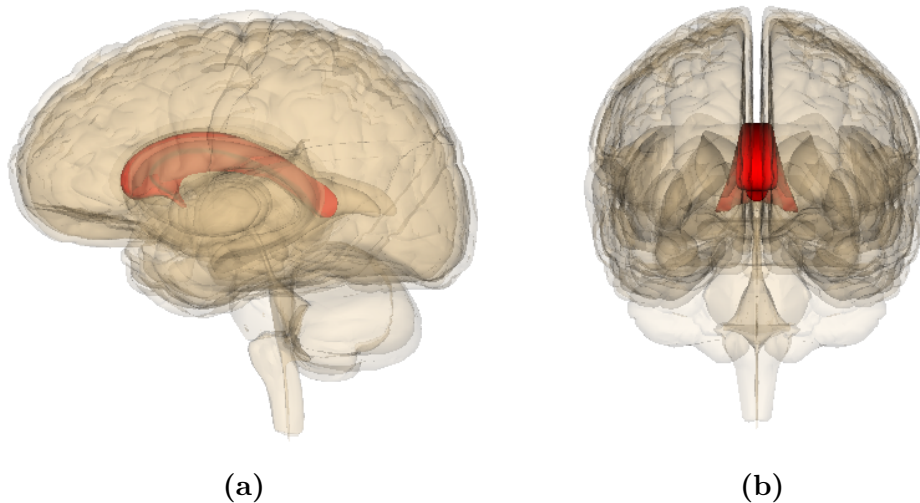


Figura 1.1 – Corpo caloso (em vermelho) visto em a) Corte sagital e b) Corte coronal (Fonte: Life Science Databases(LSDB), 2009)

O estudo do CC, e em geral do cérebro, é melhor feito em imagens de ressonância magnética (MRI), uma vez que elas fornecem melhor contraste de tecidos moles do que outras tecnologias de imageamento como radiografia ou tomografia computadorizada [25]. Como parte do estudo do CC, um passo necessário é a sua segmentação que permite qualquer análise posterior, tanto clínica quanto de pesquisa. A segmentação é entendida como a subdivisão da imagem em regiões intra-homogêneas com sentido semântico e sem sobreposição; e, em condições ideais, separa a estrutura de interesse do resto da imagem. A segmentação do corpo caloso permite a extração de características morfológicas e fisiológicas da estrutura, possibilitando avaliação tanto de indivíduos quanto de populações e estudos clínicos e de pesquisa.

1.1 Motivação

A segmentação do CC é um tema bastante estudado, porém não está resolvido. Pelo contrário, a segmentação do CC em MRI é um processo desafiador pelas seguintes razões: variabilidade da forma do CC e da intensidade da imagem do CC, entre sujeitos e tomadas em diferentes *scanners*; efeito de volume parcial, que dificulta a definição da fronteira da estrutura; proximidade e similaridade de outras estruturas como o fórnix, que dificulta a separação do CC; zonas estreitas do CC, especialmente na parte central, que causam divisão da segmentação final.

A maioria dos métodos de segmentação do CC são implementados em modalidades estruturais como imagens ponderadas em T_1 . Na literatura existem poucos métodos de segmentação do CC baseados em imagens de difusão por ressonância magnética, que abrangem imagens ponderadas em difusão (DWI), imagens por tensor de difusão (DTI), mapas de anisotropia fracionada (FA) e outros mapas escalares, já que este tipo de imagem é ainda pouco explorada. O processo de segmentação em difusão é ainda mais desafiador do que em T_1 dado que estas imagens apresentam baixa resolução, definição e contraste. Porém, como este tipo de imagem oferece informação adicional, não presente em outras modalidades, e permite estudar a microestrutura das fibras neuronais usando um modelo de difusão da água, a segmentação neste espaço é ainda desejável.

A avaliação de uma segmentação normalmente é feita usando um padrão base com o qual comparar. O *Ground-Truth* (GT), o padrão ideal para avaliar segmentações, quase sempre não está disponível, já que deve ser obtido através de medidas diretas da estrutura. No caso do CC, o GT seria obtido através de disseção e medida direta da estrutura real. Por isso, é bastante comum o uso de um padrão-ouro (*gold standard*), que é uma boa aproximação ao GT. Na área de neuroimagem, o padrão-ouro é construído a partir de uma segmentação manual feita por um ou mais especialistas, e é o padrão normalmente usado para avaliar os resultados de segmentação. Dado o interesse que temos em segmentação do CC em imagens de difusão, é preciso ter um padrão-ouro neste espaço. Porém, a obtenção de um padrão-ouro é difícil pelas características já vistas nas imagens deste espaço. Por isso, um padrão-prata (*silver-standard*) torna-se uma aproximação razoável ao padrão-ouro. O padrão-prata pode ser construído usando métodos computacionais e atualmente é motivo de estudo tanto do ponto de vista da construção quanto do seu impacto na avaliação de segmentações.

A partir de nossa Revisão Sistemática da Literatura (RSL) em métodos de segmentação do CC em MRI [18], foram identificados 36 estudos, considerados relevantes, que incluem métodos automáticos e semi-automáticos, 2D e 3D e as diferentes modalidades de MRI. Para chegar nos 36 estudos finais, fez-se uma etapa prévia de revisão detalhada dos estudos primários, descartando aqueles que não apresentavam resultados quantitativos. A partir desta revisão, fazemos algumas observações que consideramos

importantes para o presente trabalho:

- Os métodos em imagens 2D estão melhor estabelecidos (maior número de métodos usam este tipo de imagem e são mais robustos) do que em 3D.
- Dos 36 estudos finais da RSL, 27 apresentam resultados quantitativos de segmentação 2D do CC, entre os quais só 3 baseiam-se em imagens de difusão por ressonância magnética; os restantes usam imagens ponderadas em T_1 .
- Durante a etapa de revisão detalhada, 65% dos artigos primários foram descartados por não terem resultados quantitativos.
- Em geral, há uma grande variabilidade na forma de avaliar os resultados de segmentação e são usadas diferentes bases de dados. A grande maioria das abordagens compara o resultado obtido com uma segmentação manual feita sob condições específicas. Não existe uma base de dados pública nem um único padrão para avaliar os resultados.
- Os métodos computacionais de segmentação do CC em MRI, analisados durante a revisão, não têm conseguido resolver todos os desafios existentes, como por exemplo, não intervenção humana, boa acurácia, reprodutibilidade, boa capacidade de generalização.

Considerando que é desejável uma boa segmentação do CC no espaço das imagens de difusão, percebemos a necessidade de propor melhorias no método usual por sobreposição para avaliação de segmentações, dando ênfase ao padrão e às métricas usadas. É importante deixar claro que neste trabalho, a avaliação de segmentações refere à avaliação da máscara binária resultado da execução de um método de segmentação.

Para estudar a avaliação de segmentações 2D do CC, é necessário abordar questões importantes tais como: como escolher/construir o melhor padrão; como comparar segmentações em espaços diferentes (diferentes modalidades de MRI); quais são as métricas apropriadas para comparação dos resultados.

1.2 Objetivos

1.2.1 Objetivo geral

Propor um método de avaliação de segmentações 2D do corpo caloso em imagens de difusão por ressonância magnética usando diretamente o padrão-ouro em T_1 junto com uma métrica baseada no perfil de curvatura da estrutura.

1.2.2 Objetivos específicos

- Implementar os métodos mais relevantes de segmentação 2D do CC em difusão.
- Propor um método de avaliação de segmentações baseado no perfil de curvatura da estrutura.
- Caracterizar a métrica usada para comparar perfis de curvatura.
- Avaliar diferentes segmentações com o método proposto em 145 sujeitos.

1.3 Principais contribuições

Enquanto, os estudos da literatura têm se focado quase exclusivamente em propor métodos de segmentação, a importância do padrão e da métrica, dentro do processo de avaliação, melhora e comparação de métodos de segmentação do corpo caloso em MRI, tem sido negligenciada. A principal contribuição deste trabalho é estudar essa importância, analisando o padrão a ser usado e propondo um método que permita melhorar a avaliação de segmentações. Além desta contribuição principal, este trabalho apresenta as seguintes contribuições:

- Realização de uma RSL em métodos computacionais focados na segmentação do corpo caloso em MRI [18].
- Proposta de um método de segmentação do corpo caloso baseado em classificação por pixel em DWI [42, 43] levando em consideração que este tipo de imagem não tem sido usado, até agora, para segmentar o CC (os métodos em difusão tem sido implementados sobre imagens DTI e FA).
- Comparação dos métodos do estado da arte em segmentação do CC no espaço de difusão para uma base de dados de 145 sujeitos, sendo que na literatura mais de 90% dos estudos não usam mais de 100 sujeitos.
- Formulação do perfil de curvatura como método para a obtenção da assinatura do CC [41] e criação de uma ferramenta interativa para explorá-lo (Apêndice A).
- Formulação de um método para comparar resultados de segmentação em diferentes modalidades baseado em descrição da forma, que dispensa o uso de registro de imagens e comparação pixel a pixel (sobreposição). Ainda que métodos de comparação de forma já tenham sido propostos, nenhum deles foi usado para permitir comparação de resultados de segmentação em diferentes espaços.

1.4 Organização do trabalho

O presente trabalho de dissertação vem organizado em 7 capítulos como detalha-se a seguir. O Cap. 1, descreve a motivação, objetivos e contribuições do trabalho. O Cap. 2, apresenta a base teórica relacionada ao CC, e às diferentes modalidades de MRI. Já o Cap. 3 descreve o método usual por sobreposição (definido assim por ser o usado normalmente na literatura) para avaliação de segmentações, abordando em detalhe os componentes: métodos de segmentação encontrados na literatura, os padrões e as métricas.

O Cap. 4 apresenta o método de segmentação do CC em DWI proposto no presente trabalho, baseado em classificação por pixel. Este método é usado junto com os outros métodos encontrados na literatura para estudar a avaliação de segmentação do CC em difusão. No Cap. 5 é proposto um novo método de avaliação de resultados de segmentações baseado no perfil de curvatura.

Posteriormente, no Cap. 6, são apresentados os resultados dos experimentos numa população de 145 sujeitos, e uma discussão com relação aos resultados obtidos. Finalmente, o Cap. 7 apresenta as conclusões mais importantes do trabalho.

2 Conceitos teóricos

O presente capítulo aborda os conceitos teóricos necessários para entender os capítulos posteriores: o CC como estrutura de interesse do presente trabalho (Sec. 2.1) e as modalidades de MRI usadas no seu estudo (Sec. 2.2).

2.1 Corpo Caloso (CC)

O plano que divide o cérebro em duas metades simétricas ao longo da fissura inter-hemisférica é chamado plano meio sagital. Planos que são paralelos a este plano meio sagital são chamados de planos sagitais indo em direção direita-esquerda (Fig. 2.1a). Os planos coronais (Fig. 2.1b) são perpendiculares ao plano sagital e vão da parte traseira da cabeça (parte posterior) até a zona dos olhos (parte anterior). Perpendiculares aos planos mencionados anteriormente, estão os planos axiais (Fig. 2.1c) que vão desde a parte inferior da estrutura até a parte superior [28].

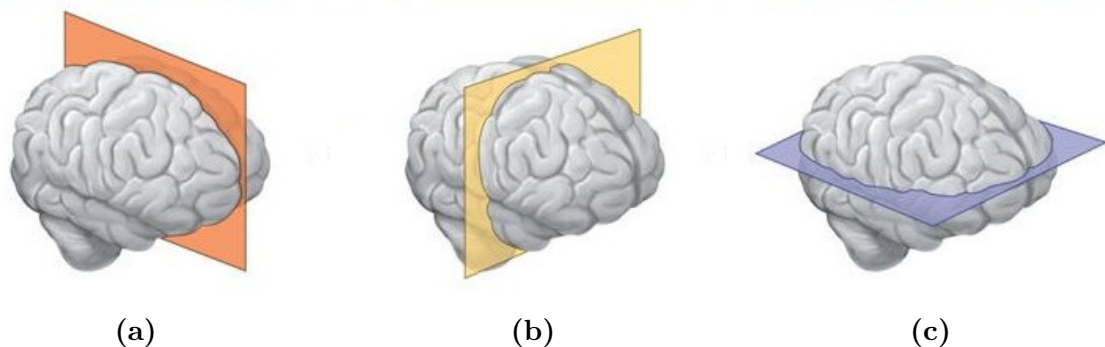


Figura 2.1 – Planos usados para descrição do cérebro: a) Plano sagital, b) Plano coronal, c) Plano axial (Fonte: <http://biology-forums.com>, 2016)

CC (corpo duro em latim) é uma estrutura localizada abaixo do córtex cerebral e é o motivo deste estudo. Com aproximadamente 200 a 250 milhões de projeções axonais perpendiculares ao plano meio sagital, o CC é a maior estrutura de substância branca no sistema nervoso central.

No plano sagital, começando pela parte anterior do CC, a porção externa é conhecida como **joelho** (Fig. 2.2a) e a curva inferior que descreve o CC partindo do joelho, é denominada **rostró** (Fig. 2.2b). Entre a parte anterior e a posterior, está primeiro o **corpo** (Fig. 2.2c) e depois o **tronco** (Fig. 2.2d). Finalmente o extremo posterior é referido como esplênio (Fig. 2.2e). Adicionalmente, o fórnix (Fig. 2.2f) é um feixe de fibras que não faz parte do CC mas que tem importância no nosso estudo dada sua proximidade com

o CC; o fórnix fica localizado na parte inferior deste último e leva sinais do hipocampo para o hipotálamo.

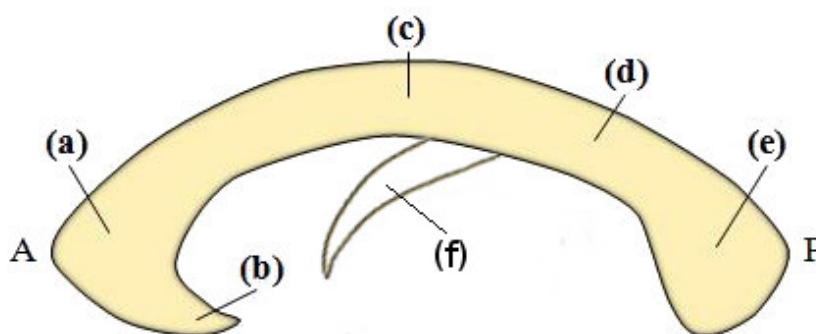


Figura 2.2 – Partes do corpo caloso no plano sagital ordenadas desde a parte anterior (A) até a parte posterior (P): a) joelho, b) rostro, c) corpo, d) tronco, e) esplênio; e f) fórnix (Fonte: Segmentação e parcelamento do corpo caloso em imagens de tensor de difusão, 2012 [31])

Enquanto as funções do hemisfério esquerdo diferem das do lado direito do cérebro, o CC facilita a conexão entre eles, atuando como uma ponte e transmitindo informação de um hemisfério para o outro, possibilitando o funcionamento apropriado do sistema nervoso como um todo.

Atualmente existe uma grande quantidade de estudos, sobre a relação existente entre o tamanho do CC e características da pessoa tais como gênero [100], idade [51], habilidades numéricas e matemáticas [76], lateralidade (uso preferencial de uma mão) [112]. Mais importante ainda, do ponto de vista clínico, o corpo caloso é afetado por doenças como Alzheimer [35], autismo [101], esquizofrenia [99], dislexia [109], epilepsia [106], esclerose múltipla [24], depressão [110]. Os autores também o relacionam com alcoolismo [80], obesidade [58] e consumo de cigarro [34].

2.2 Imagens por Ressonância Magnética

Imagem por Ressonância Magnética (MRI) possibilita o estudo *in vivo* das estruturas cerebrais e suas funções. Na atualidade, esta é a melhor técnica para obter informação sobre o CC uma vez que utiliza radiação não ionizante, facilita detecção de tumores e, quando comparada com Raio X ou Tomografia computadorizada, é rápida e provê melhor contraste para tecidos moles [25].

Os estudos de ressonância magnética começaram formalmente em 1939, com a descoberta da técnica para medir momentos magnéticos nucleares [94]. A partir dali uma série de experimentos e teoria foram desenvolvidos em torno da ressonância magnética nuclear. Porém foi ao longo dos anos 70 quando Raymond Damadian, ajudado pelo trabalho de Paul Lauterbur, desenvolveu as bases de MRI tanto teóricas quanto técnicas: gerou as primeiras imagens de MRI, usou o atraso nas constantes de tempo de relaxação

(T_1 e T_2) para detecção de tecidos cancerosos e produziu os primeiros *scanners* para aquisição de imagens por ressonância magnética [21, 20, 60].

Desde então, MRI tem sido aplicado com sucesso para obter imagens do corpo humano devido ao seu alto conteúdo de água (átomos de hidrogênio na água e tecido adiposo correspondem a 60% do peso total do corpo). Já que há um próton em cada núcleo de hidrogênio e devido ao *spin* do núcleo, um momento magnético é criado. Quando um átomo de hidrogênio é posicionado em um campo magnético estático, o momento magnético dele rotaciona em volta do campo magnético estático como um giroscópio à frequência Larmor, que é única para o elemento [28].

Para adquirir as imagens, é necessário um pulso de radio-frequência em um plano perpendicular ao campo magnético estático. A frequência deste pulso é idêntica à frequência de Larmor e, uma vez o átomo é excitado, o campo magnético é desligado e o momento magnético original decai ao seu estado de equilíbrio emitindo um sinal. Este sinal é detectado e duas constantes de tempo de relaxamento podem ser medidas: a constante de magnetização longitudinal T_1 , que é sensível às propriedades térmicas do tecido, e a constante de magnetização transversal T_2 , que é afetada pelas variações locais do campo. MRI engloba várias modalidades, dentre elas descreveremos os que serão utilizados neste trabalho: MRI ponderada em T_1 (na literatura é usado simplesmente T_1) (Fig. 2.3a), MRI ponderada em difusão (DWI) (Fig. 2.3b), Imagem por tensor de difusão (DTI) (Fig. 2.3c) e Mapa de anisotropia fracionada (FA) (Fig. 2.3d).

2.2.1 MRI ponderada em T_1

A imagem ponderada em T_1 (Fig. 2.3a), a modalidade de MRI mais amplamente usada, deriva diretamente da medição da constante de magnetização longitudinal T_1 formulada a partir do trabalho de Felix Bloch, Nicolaas Bloembergen e Edward Purcell em MRI [8, 9]. Esta modalidade de MRI é considerada estrutural, já que descreve qualitativa e quantitativamente a forma, tamanho e integridade das estruturas de substância branca e cinza do cérebro.

T_1 representa a constante de magnetização longitudinal e é gerada por repetições curtas de tempo entre pulsos de radio frequência e um período de recuperação de sinal curto [55]. Como T_1 é uma constante de tempo de crescimento exponencial, tecidos com curta taxa T_1 são visualizados como áreas claras, enquanto tecidos com taxas longas são vistos como regiões mais escuras em imagens de ponderação T_1 [95]. Desta forma, o CC é visualizado como uma estrutura brilhante em imagens sagitais de ponderação T_1 , permitindo melhor contraste com as áreas circundantes [78].

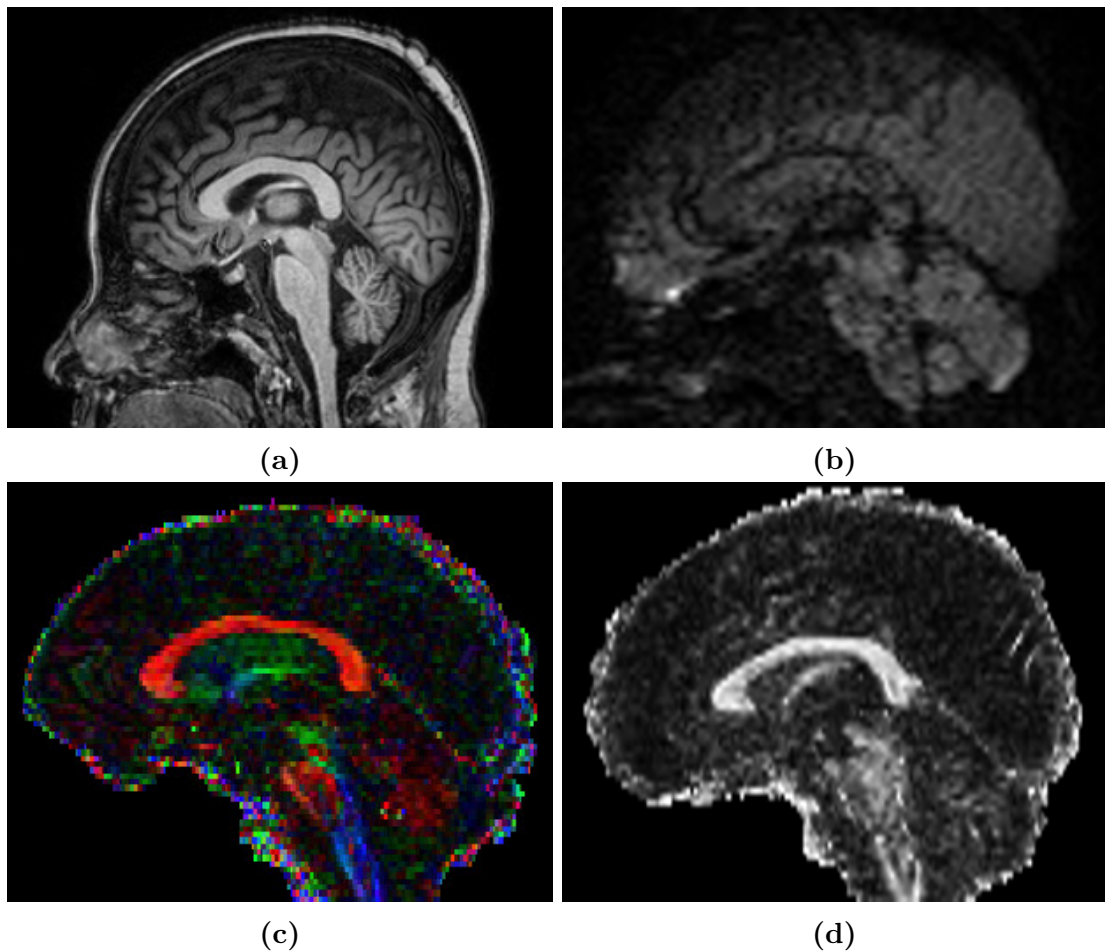


Figura 2.3 – Cérebro em plano sagital como visto em várias modalidades de MRI: a) T_1 , b) DWI (Resposta à aplicação de um gradiente), c) DTI (Visualizado mediante um mapa de cores), d) Mapa de Anisotropia fracional (FA).

2.2.2 MRI ponderada em difusão (DWI)

As imagens ponderadas em difusão (DWI) provêm contraste baseado nas diferenças de difusão de moléculas de água dentro do cérebro. As primeiras sequencias de pulso ponderado em difusão foram descritas em meados dos anos 60 [102] mas só em 1986 foram clinicamente disponibilizadas para rotinas médicas [61]. A difusão representa o movimento térmico aleatório das moléculas, mais conhecido como movimento Browniano, e depende de uma variedade de fatores como o tipo de molécula sob estudo, a temperatura e o micro-entorno em que ocorre a difusão [5]. Em conjuntos de fibras, a difusão molecular efetiva é maior na direção ao longo da fibra do que na direção perpendicular. Esta classe de difusão é conhecida como difusão anisotrópica em contraste à difusão isotrópica [45].

A obtenção de DWI é feita aplicando dois gradientes de campo magnético: o primeiro gradiente faz com que as moléculas adquiram deslocamentos de fase, enquanto o segundo gradiente cancela estes deslocamentos reorientando os *spins* estacionários. A reorientação para o estado inicial não será perfeita para moléculas que tenham se movimentado durante o intervalo de tempo entre os gradientes. Isto causa uma perda de

sinal. O grau de perda de sinal é determinado pelo grau de movimento aleatório (dado pelo coeficiente de difusão) e pela intensidade e duração dos gradientes. Para o caso do CC, a perda de sinal é baixa devido ao baixo movimento aleatório, em virtude da sua estrutura fibrosa (altamente orientada). O grau de perda de sinal (Eq. 2.1) pode ser melhorado incrementando-se a intensidade e duração dos gradientes de codificação da difusão.

$$A \simeq e^{-bD} \quad (2.1)$$

onde A representa a perda de sinal ou atenuação, D é o coeficiente de difusão e b é uma variável que depende da razão giromagnética (γ) da amplitude (G), duração (δ) e intervalo (Δ) dos gradientes aplicados (Eq. 2.2).

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \quad (2.2)$$

Assim, em DWI, a intensidade de sinal (IS) de um voxel é igual à intensidade do sinal sem gradiente aplicado ($IS_0 : b \approx 0 \text{ s/mm}^2$), reduzido pela atenuação (A) relacionada ao coeficiente de difusão do tecido (D) e às características do gradiente aplicado b (Eq. 2.3).

$$IS = IS_0 \times e^{-bD} \quad (2.3)$$

Porém, em presença de anisotropia, a difusão não pode ser mais caracterizada por um coeficiente D . Assim, para descrever completamente o tecido mediante DWI, gradientes em várias direções são aplicados obtendo um mapa de intensidade 3D em cada direção IS_d (Fig. 2.3b).

2.2.3 Imagem por tensor de difusão (DTI)

Introduzida em meados dos 90's [6], o modelo DTI (Fig. 2.3c) é calculado a partir de DWI, aplicando um modelo tensorial para combinar os valores de difusão em cada direção, onde cada voxel é representado por um tensor de segunda ordem \underline{D} (Eq. 2.4) que descreve completamente a mobilidade molecular ao longo de cada direção e a correlação entre elas [62, 63].

$$\underline{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \quad (2.4)$$

Os elementos diagonais ($D_{ii} > 0 : i = x, y, z$) representam a variância ao longo das direções x, y, z e os elementos fora da diagonal ($D_{ij} : i, j = x, y, z \forall i \neq j$) representam a

covariância entre duas direções. Esta matriz é simétrica ($D_{ij} = D_{ji} : i, j = x, y, z$) devido ao fato da probabilidade de difusão ao longo de uma determinada direção ser a mesma em sentido contrário ao longo da mesma direção. A atenuação do sinal medido é expressa na Eq. 2.5.

$$A \simeq \exp \left(- \sum_{i=x,y,z} \sum_{j=x,y,z} b_{ij} D_{ij} \right) \quad (2.5)$$

Para determinar completamente os tensores de difusão do volume, é necessário coletar DWI ao longo de várias direções de gradiente (mínimo 6 direções mais a imagem adquirida com $b=0$). Os valores D_{ij} são calculados a partir da DWI por regressão linear usando a equação 2.5. Para cada voxel, os autovalores e autovetores do tensor podem ser calculados, o que fornece, respectivamente, as principais direções de difusão e suas difusividades associadas (autovalores).

2.2.4 Mapa de anisotropia fracionada (FA)

A partir de DTI, alguns mapas escalares podem ser derivados para prover, de forma mais simplificada, informação sobre a microestrutura do tecido e a arquitetura de cada voxel. Entre eles, o grau de anisotropia tem grande importância já que, ainda que simplifique o modelo tensorial, provê informação valiosa sobre a estrutura do cérebro [62].

A anisotropia descreve quanto o deslocamento molecular varia no espaço e é relacionada à presença de tecidos altamente organizados. Muitos índices tem sido propostos para caracterizar a anisotropia da difusão, porém o mais aceito é o mapa de anisotropia fracionada (FA) (Fig. 2.3d) que pode ser calculado a partir dos autovalores do tensor \underline{D} do modelo DTI com ajuda da equação 2.6, proposta em [90]. FA indica indiretamente o nível de organização do tecido, sendo 0 para meios isotrópicos (difusão ocorre igualmente em todas as direções) e 1 para meios completamente anisotrópicos (difusão só ocorre numa direção).

$$FA = \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{2 \cdot (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \quad (2.6)$$

Devido às fibras do CC serem bem orientadas no plano sagital e à sua alta anisotropia, DTI e particularmente FA, é capaz de gerar um bom contraste desta estrutura [1].

2.3 Processamento morfológico de imagens

O processamento morfológico tem como objetivo filtrar ou quantificar uma imagem de acordo à forma como um elemento estruturante encaixa (ou não encaixa) dentro da imagem. O elemento estruturante pode ser visto como o operador básico do processamento morfológico e é usualmente uma imagem binária, simétrica e mais pequena do que a imagem sobre a que opera.

Na Fig. 2.4, um elemento estruturante circular é posicionado em dois lugares sobre uma imagem binária. Na posição da esquerda o elemento encaixa, entanto que na posição da direita não encaixa. Marcando as posições onde o elemento encaixa, pode ser extraída informação estrutural da imagem. Esta informação depende basicamente do tamanho e forma do elemento estruturante. Este conceito simples é usado nas operações morfológicas que abordaremos ao longo deste trabalho: erosão, dilatação, fechamento e abertura. Serão usados dois tipos de elemento estruturante: caixa (Fig. 2.5a) e cruz (Fig. 2.5b).

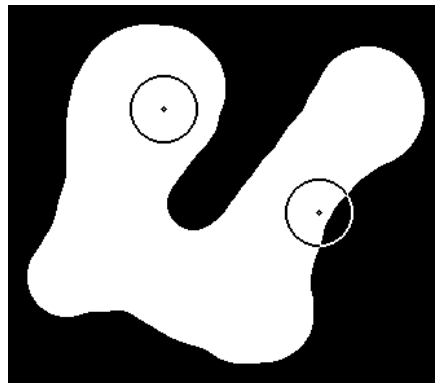


Figura 2.4 – Processamento morfológico sobre uma imagem binária com um elemento estruturante circular em duas posições. Fonte: [69].

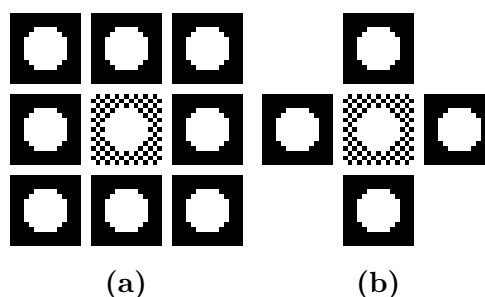


Figura 2.5 – Elementos estruturantes de tamanho 1 usados ao longo deste trabalho: a) caixa b) cruz. Fonte: [69].

2.3.1 Erosão e dilatação

A operação básica de processamento morfológico é a erosão por um elemento estruturante. A erosão de uma imagem é calculada provando o elemento estruturante

em todas as posições da imagem. Onde o elemento estruturante encaixa completamente dentro do objeto da imagem (região diferente de zero), o centro do elemento estruturante é marcado. A erosão consiste em manter como objeto a região marcada. Frequentemente, a erosão é uma versão encolhida da imagem original (Fig. 2.6a). A erosão de uma imagem A por um elemento estruturante b é definida pela Eq. 2.7.

$$A \ominus b = \{x : b_x \subset A\} \quad (2.7)$$

onde b_x representa o deslocamento do elemento b para um ponto x que está contido em A .

No caso da dilatação, esta pode ser vista como o operador dual da erosão, aplicada sobre a imagem invertida (imagem negada). Enquanto a erosão opera sobre o objeto da imagem, a dilatação opera sobre o fundo da imagem. Como resultado, a dilatação é uma versão expandida da imagem original (Fig. 2.6b). A dilatação de uma imagem A por um elemento estruturante b é definida pela Eq. 2.8.

$$A \oplus b = (A^c \ominus b)^c \quad (2.8)$$

onde A^c denota o complemento de A e $(A^c \ominus b)^c$ denota o complemento de $(A^c \ominus b)$.

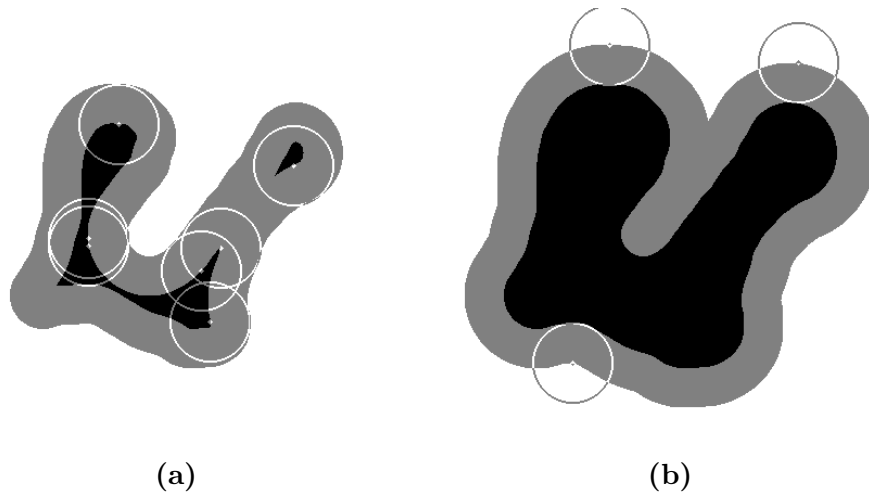


Figura 2.6 – Erosão e dilatação onde a imagem original aparece em preto e a operação morfológica aparece em cinza: a) Erosão morfológica b) Dilatação morfológica. Fonte: [69].

2.3.2 Abertura e fechamento

A abertura é definida como a união de todos os elementos estruturantes que encaixam dentro da imagem (Fig. 2.7a). A abertura é dada pela Eq. 2.9.

$$A \circ b = \cup \{x : b_x \subset A\} \quad (2.9)$$

A versão dual da abertura é o fechamento (Fig. 2.7b) dado pela Eq. 2.10

$$A \bullet b = (A^c \circ b)^c \quad (2.10)$$

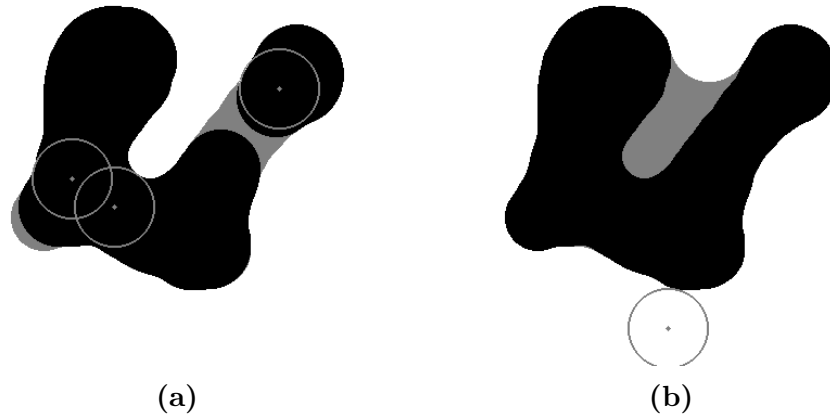


Figura 2.7 – Abertura e fechamento onde a imagem original aparece em preto e a operação morfológica aparece em cinza: a) Abertura morfológica b) Fechamento morfológico. Fonte: [69].

2.3.3 Transformada *Watershed*

Um algoritmo amplamente usado para segmentação morfológica é a transformada *Watershed* [7] que é baseada no conceito de inundação onde a imagem, sendo considerada como uma superfície (Fig. 2.8a), é submergida a velocidade vertical constante, em um lago. A água entra na superfície através de furos (Fig. 2.8b); e fluxos de água que provêm de diferentes furos podem se misturar quando chegam aos pontos de encontro. Para evitar essa mistura são construídos diques verticais nos pontos de encontro, antes de eles ocorrerem, mantendo separadas as zonas de inundação de cada furo (Fig. 2.8c). No final, quando a superfície é completamente submersa, os diques resultantes estão expostos e são as linhas de segmentação *Watershed* resultante (Fig. 2.8d).

2.4 Modelos de aprendizado

Existem dois tipos de classificadores: supervisionado e não supervisionado. Os classificadores supervisionados precisam de exemplos rotulados para aprender; este estágio de aprendizado é chamado treinamento do modelo. No caso dos classificadores não supervisionados, a informação passada para treinar o modelo não é rotulada e o classificador aprende baseado em padrões existentes na informação. Uma vez treinado o modelo, novas amostras podem ser passadas e o classificador irá predizer a classe; este segundo estágio é chamado teste do modelo [39]. A seguir, são apresentados dois tipos

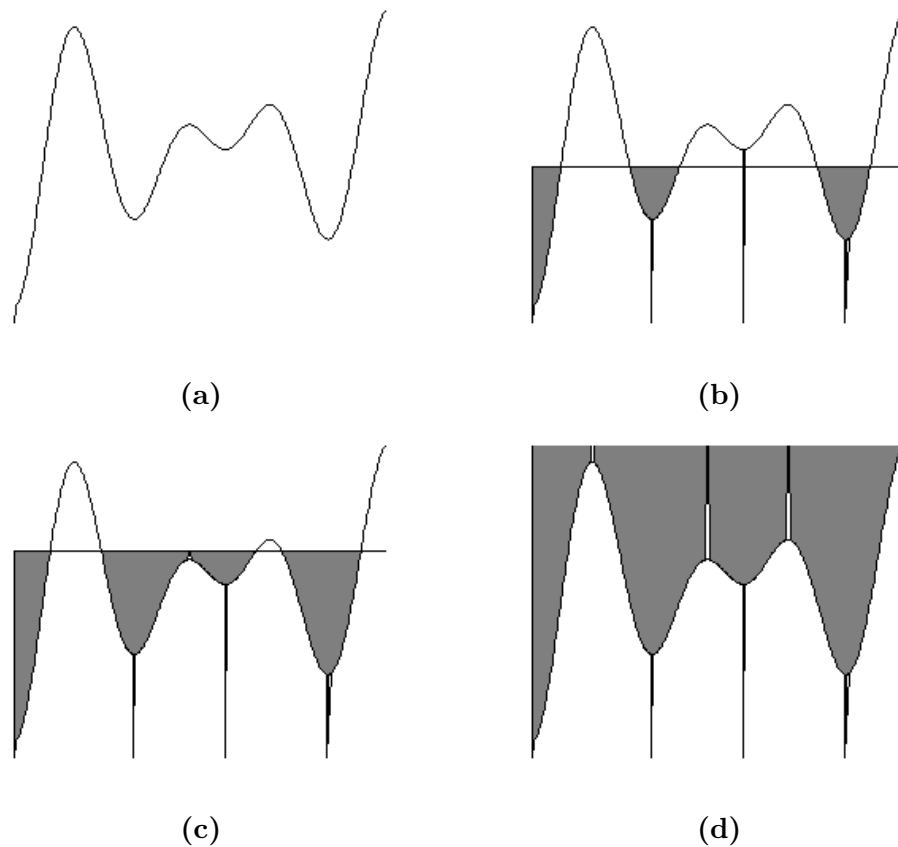


Figura 2.8 – Esquema transformada *Watershed* (representação 1D): a) Sinal de entrada, b) Entrada de água através de furos (mínimos locais), c) Dique construído no ponto de encontro de duas bacias, d) Diques expostos na segmentação *Watershed* resultante. Fonte: [69].

comuns de classificador que serão usados depois: Máquina de vetores suporte como técnica supervisionada e *k-means* como técnica não supervisionada.

2.4.1 Máquina de vetores suporte (SVM)

Dentro das técnicas de aprendizado de máquina supervisionada, as SVM são reconhecidas por terem bons resultados em múltiplas aplicações e serem efetivas em espaços de alta dimensionalidade e cenários desbalanceados [14, 103]. A técnica foi proposta por Cortes e Vapnik em 1995 para classificar amostras de treinamento não linearmente separáveis [17].

Dado um conjunto de treinamento de N amostras ($\{(x_i, y_i)\}$), onde $x_i \in R^n$ são observações com n características e $y_i \in \{0, 1\}$ são seus correspondentes rótulos, SVM encontra um hiperplano com máxima margem de separação entre classes resolvendo a função de classificação (Eq. 2.11).

$$f(x) = \sum_{i=1}^N a_i y_i K(x_j, x_i) + b \quad (2.11)$$

$K(x_j, x_i) = \langle \varphi(x_j) \cdot \varphi(x_i) \rangle$ é o kernel que leva a função φ do espaço original para um espaço de maior dimensão ($(R_n \mapsto R_m) : m \geq n$) onde as classes são linearmente separáveis. No espaço de maior dimensão (m), o vetor de pesos (w) que define o hiperplano de separação, é a combinação linear de $\varphi(x_i)$ (Eq. 2.12).

$$w = \sum_{i=1}^N a_i y_i \varphi(x_i) \quad (2.12)$$

No caso em que o hiperplano de separação não exista, uma variável de folga (ξ_i), que mede o desvio de uma amostra da condição ideal de separabilidade, é introduzida (Eq. 2.13).

$$y_i[(w \cdot x_i) + b] \geq 1 - \xi_i, \quad \xi_i > 0 \quad (2.13)$$

Encontrar um hiperplano de separação para o qual o erro de classificação, na média sobre as amostras de treinamento, é minimizado é equivalente a minimizar a função Φ na Eq. 2.14, sujeita à Eq. 2.13.

$$\Phi(w, \xi) = \frac{1}{2} w^T w + C \sum_{i=1}^N \xi \quad (2.14)$$

onde C é um parâmetro de penalidade que controla a relação de benefício entre complexidade do modelo e o número de pontos não separáveis.

2.4.2 *k-means*

O algoritmo de *k-means* é um método não supervisionado iterativo para particionar um conjunto de dados em k clusters, escolhendo o número inicial de conjuntos e agrupando cada amostra seguindo algum critério de proximidade [48]. O algoritmo de *k-means* é descrito a seguir:

1. São escolhidos aleatoriamente k amostras para ser os centros de cada *cluster*.
2. Cada amostra é atribuída ao centro mais próximo (usando alguma métrica de distância).
3. São re-calculados os centros de cada *cluster*, como sendo o centro de massa das amostras desse *cluster*.
4. Os passos 2. e 3. são repetidos até lograr convergência (usa-se algum critério de parada).

2.5 Métodos de segmentação do corpo caloso em MRI

Um método de segmentação é um algoritmo que rotula cada pixel da imagem como fundo (Falso) e objeto de interesse (Verdadeiro), sendo o objeto de interesse a estrutura a ser segmentada. O resultado da execução do método de segmentação é uma máscara binária.

Dado que, no caso de MRI, as imagens são volumétricas, o primeiro passo para fazer a segmentação 2D do CC é a escolha da fatia dentro do volume do sujeito. Idealmente, a segmentação é feita na fatia meio sagital (fatia localizada no plano meio sagital que divide o cérebro em duas metades simétricas) onde o CC é melhor visualizado. Porém, obter a fatia meio sagital não é uma tarefa trivial dado que, na prática, o sujeito não é posicionado exatamente com o plano de referência de aquisição dos dados. Isto quer dizer que o plano de simetria do sujeito não coincide com o plano de referência de aquisição dos dados. Para lograr esta coincidência, o volume do sujeito teria que ser transformado para o plano de referência.

Adamson et al. [3] propõem um método para encontrar a fatia meio sagital em T_1 usando duas opções: alinhamento com o eixo das comissuras anterior e posterior (feixes de fibras na região próxima do fórnix) ou aplicando uma transformada de registro para um atlas. Estas operações de registro são complexas e precisam de interpolação, e portanto os métodos na literatura costumam realizar a segmentação na fatia mais próxima à fatia meio sagital, denominada fatia central. Para a maioria dos casos, esta simplificação é suficientemente boa, porém, a influência da escolha da fatia na segmentação final e nas análises clínicas posteriores (baseados na segmentação do CC) não tem sido estudada.

Neste estudo estamos interessados em comparar segmentações, portanto independe da escolha da fatia, desde que a comparação seja feita sobre a mesma fatia. Utilizaremos portanto o método para a escolha da fatia central proposto por Freitas et al. [30]. Tal método usa o fato de que a fatia meio sagital (fissura inter-hemisférica do cérebro) é principalmente composta de fluido cérebro-espinal. Este alto conteúdo de fluido cérebro-espinal está associado a valores baixos de FA. Assim, o valor médio de FA é calculado para cada fatia ao longo do plano sagital. A fatia com o valor médio de FA mais baixo é a fatia central. O método leva em conta que fatias dos extremos do crânio devem ser desconsideradas apesar de apresentarem valores baixos de FA médio (Fig. 2.9).

Dentre os estudos de segmentação do CC em difusão, anteriormente apresentados, só o método de limiarização [82] e os baseados em *Watershed* [30, 97] são de interesse deste projeto, por serem implementados em imagens 2D e, portanto, serão descritos a seguir (Subseções 2.5.1 e 2.5.2).

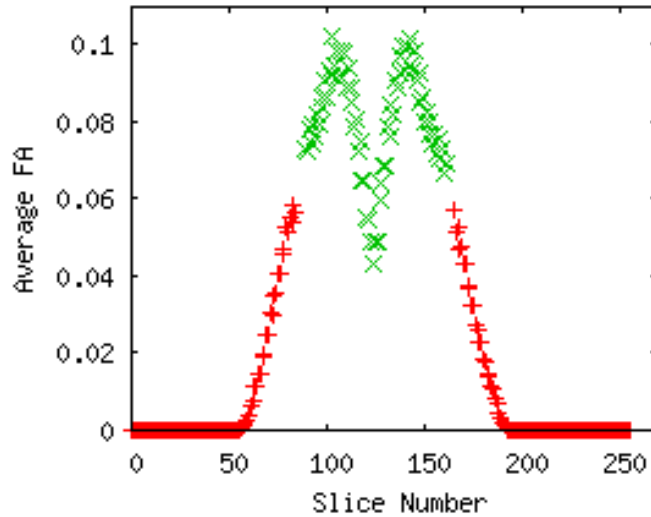


Figura 2.9 – Valor médio de FA por fatia ao longo do plano sagital, onde os pontos em vermelho são descartados por não serem candidatos. O mínimo em verde corresponde à fatia central. Fonte: [30].

2.5.1 Método baseado em *Watershed*

O método de segmentação baseado em *Watershed* está baseado nos conceitos da Seção 2.3.3. Na implementação clássica da segmentação baseada em *Watershed*, a superfície é o gradiente da imagem e os furos, denominados marcadores, são os mínimos locais da imagem. Esta implementação leva na prática à sobre-segmentação, e portanto, os marcadores podem ser escolhidos com um filtro ou devem ser posicionados manualmente.

Neste trabalho foi utilizada a implementação de Freitas et al. [30], onde a transformada *Watershed* é executada sobre o gradiente morfológico externo do mapa de FA ponderado. O gradiente morfológico é obtido calculando a diferença entre o mapa de FA ponderado e uma versão dilatada dele (Eq. 2.15).

$$G_E = FA_W \oplus b - FA_W \quad (2.15)$$

$FA_W = FA * e_{1x}$ é o mapa ponderado de FA e e_{1x} é a projeção do autovetor principal e_1 na direção x (perpendicular ao plano sagital).

Para lidar com o problema de sobre-segmentação, os marcadores são posicionados nos mínimos locais da imagem do gradiente do mapa ponderado de FA com os maiores valores de extinção.

Como resultado do *Watershed*, a imagem é dividida nas zonas de influência de cada marcador. Para obter a segmentação do CC, é necessário filtrar e agrupar as regiões que fazem parte desta estrutura. Portanto, são escolhidas as regiões cujo valor de FA ponderado médio está acima de um limiar, usando o conhecimento a priori que o CC apresenta alto valor de FA. Finalmente, entre os candidatos acima do limiar, é escolhido

como segmentação final do CC, o maior componente conexo.

2.5.2 Método de quantificação de objetivo reproduzível

O segundo método, denominado Esquema de Quantificação de Objetivo Reproduzível (ROQS), baseado em limiarização e proposto por Niogi et al. [82], usa um esquema que segmenta em três etapas as regiões com características de difusão similares às de uma semente manualmente posicionada. Primeiro, são selecionados como candidatos aqueles pixels com a mesma orientação da semente (escolhendo aqueles pixels cujo maior componente do autovetor principal seja o mesmo que o do pixel semente). Segundo, multiplicam-se cada um dos componentes do autovetor principal (nos pixels selecionados no primeiro passo), com uma matriz de magnificação (MM) baseada no FA (Eq. 2.16), com o intuito de destacar os pixels com grande anisotropia.

$$MM = \frac{FA - \max(FA) * \alpha}{\max(FA) * \beta} + \gamma \quad (2.16)$$

onde α , β e γ são constantes definidas empiricamente pelo autor, testando combinações em vários sujeitos. Logo após, são novamente selecionados aqueles pixels cuja anisotropia esteja próxima à da semente; assim, são selecionados aqueles pixels cujo maior componente do autovetor principal esteja abaixo de certo limiar com respeito ao valor do maior componente do autovetor principal da semente. O limiar é inferido empiricamente calculando a diferença entre a anisotropia da semente e a dos pixels em volta do corpo caloso. Finalmente, o terceiro passo consiste em selecionar todos aqueles pixels que sejam conexos e que passaram pelos dois critérios anteriores.

3 Método usual de avaliação de segmentações do CC em difusão por sobreposição

MRI é a ferramenta comumente usada para levar a cabo a segmentação do CC que permita fazer estudos quantitativos e qualitativos da estrutura [25, 97]. Em aplicações clínicas, a segmentação permite o monitoramento de doenças ou remissões [96], avaliação pré-operatória e planejamento cirúrgico [44] e análise estatístico de variabilidade anatômica [105]. A segmentação consiste em particionar a imagem em regiões independentes, mas que são espacialmente contíguas, cada uma contendo pixels com características homogêneas segundo alguns critérios [37]. A segmentação do CC se torna desafiadora pelas seguintes razões [40]:

- Variabilidade da forma do CC e da intensidade da imagem do CC entre sujeitos, o que dificulta aplicar técnicas padrão de segmentação como deformação de contorno e limiarização por intensidade (Fig. 3.1). Imagens vindas de diferentes *scanners* costumam ter variação na intensidade, ainda que sejam usados protocolos de aquisição similares.
- Em imagens de difusão, o efeito de volume parcial, causado pela resolução de aquisição da imagem e pelas imperfeições do perfil de pulso do equipamento, dificulta definir a fronteira como parte, ou não, de uma estrutura (Fig. 3.1a).
- Proximidade e similaridade de estruturas próximas ao CC, especialmente o fórnix (estrutura na parte inferior do CC), o que dificulta sua separação (Fig. 3.1b).
- Zonas estreitas do CC, especialmente no tronco, que podem fazer com que a estrutura seja separada em dois componentes (Fig. 3.1c).

A maioria dos métodos de segmentação do CC são implementados em imagens ponderadas em T_1 (Fig. 3.2a). Ao contrário das imagens ponderadas em T_1 , as imagens baseadas em difusão apresentam baixa resolução, definição e contraste, o que torna a segmentação um processo ainda mais desafiador, como pode se ver na imagem de DWI (Fig. 3.2b) e mapa de FA (Fig. 3.2c). Até mesmo, a segmentação manual feita por especialista a partir das imagens de difusão é propensa a erros.

A segmentação manual do CC ainda é amplamente usada como padrão absoluto, ou GT, para avaliar algoritmos de segmentação, semi ou completamente automáticos [37].

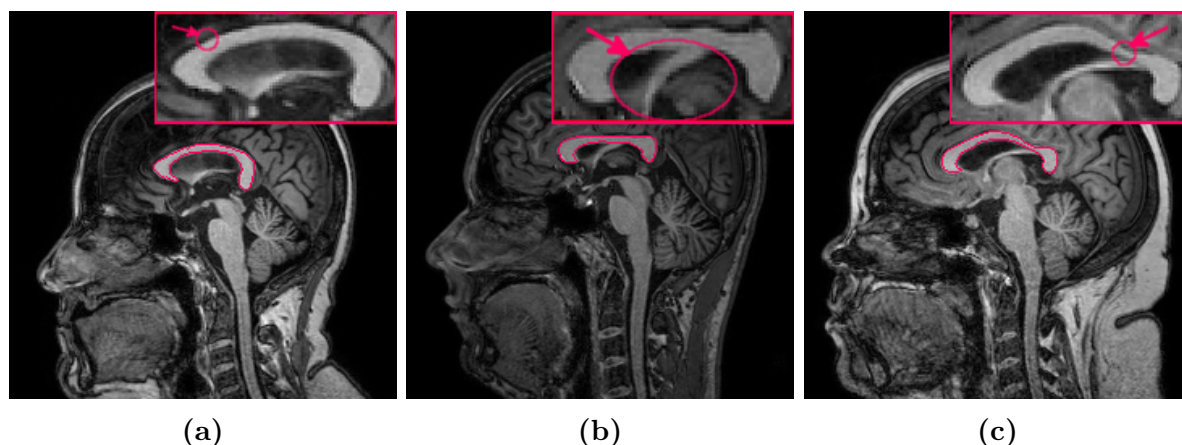


Figura 3.1 – Desafios segmentação do CC, ilustrados na fatia central de imagens ponderadas em T_1 de diferentes sujeitos. Além da variabilidade da forma do CC e da intensidade dos pixels, apresentam-se os seguintes desafios: a) Efeito de volume parcial, b) Proximidade do fórnix e c) Zonas estreitas

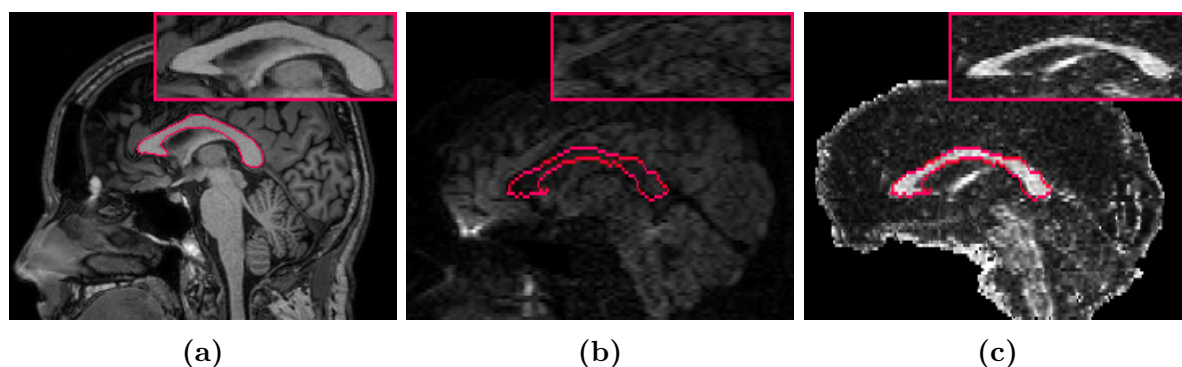


Figura 3.2 – Ilustração, para o mesmo sujeito, da segmentação do CC sobre imagens em diferentes modalidades de MRI na fatia central: a) Imagem ponderada em T_1 , b) Imagem re-escalada DWI (resposta para um gradiente) e c) Mapa de FA ponderado re-escalado

Porém, este processo tem grandes desvantagens como a dependência de especialistas treinados e experientes, o desgaste visual associado à tarefa, gasto de tempo já que a segmentação manual é normalmente feita fatia a fatia (o CC caloso abarca na média 20 fatias num volume) e tem uma grande intra e inter-variabilidade entre especialistas [38]. Com o propósito de minimizar estes problemas, métodos de segmentação automáticos e semi-automáticos são desejáveis e estão sendo continuamente desenvolvidos e melhorados como será visto a seguir.

Métodos semi-automáticos tentam solucionar parcialmente as lacunas dos métodos manuais conservando, ainda que em menor grau, alguns dos seus inconvenientes como a variabilidade entre especialistas. A intervenção humana ainda é necessária para inicializar o método, conferir o resultado e em alguns casos, corrigi-lo [29].

Por outro lado, métodos completamente automáticos são os mais desejáveis devido a sua total independência do fator humano; porém, estes ainda não têm conseguido o grau de acurácia e confiabilidade necessário para substituir os métodos manuais [85].

A RSL permitiu identificar os métodos de segmentação do CC em MRI que fazem parte do estado da arte [18]. A maioria dos estudos são feitos em imagem estrutural T_1 [93, 26, 91, 65, 104, 74, 4, 16, 47, 3, 92, 13, 78, 27, 46, 88]. Outro grupo de estudos usa imagens cuja modalidade não foi especificada, mas que aparenta ser imagens estruturais em T_1 ou T_2 [33, 40, 108, 22, 77, 10, 64, 54, 107, 66, 11, 114, 50, 89].

Finalmente em difusão, sendo uma modalidade pouco explorada, encontraram-se apenas 6 estudos usando diferentes abordagens: Niogi et al. [82] propuseram um método usando limiarização, Freitas et al. [30] e Rittner et al. [97] propuseram métodos de região baseados em *Watershed*, Nazem-Zadeh et al. [81] implementaram um método baseado em superfícies de nível, Kong et al. [57] implementaram um algoritmo de clusterização e finalmente, Garcia et al. [71] propuseram um método híbrido baseado em regiões ativas geodésicas e um modelo de curvas de nível para a implementação da evolução da curva.

Como indicado na Seção 1.1), temos interesse em estudar métodos de segmentação no espaço de difusão, já que este espaço oferece informação importante que permite estudar a microestrutura das fibras nervosas com base na facilidade ou dificuldade que a água tem para se deslocar. Estudos anteriores têm se focado principalmente em modalidades T_1 , já que a informação neste espaço é fácil de manipular, é mais intuitiva e seu uso está melhor estabelecido. É importante agora avançar no estudo de segmentação do CC no espaço de difusão, principalmente em temas de padronização e avaliação das técnicas disponíveis.

Além disso, focamos nosso estudo em métodos de segmentação 2D, que operam na fatia central do CC, uma vez que neste tipo de imagens os métodos são mais consistentes e estão melhor estabelecidos. A fatia central é extraída a partir do volume do cérebro e é normalmente usada na literatura para estudo do CC devido a que ali é melhor visualizada esta estrutura.

Avaliar segmentações para uma aplicação específica tem grande importância tanto para seleção de melhores algoritmos quanto para ajuste de parâmetros em procura do melhor desempenho. Porém, este problema não é trivial e, em geral, para a mesma imagem, a segmentação ótima pode ser diferente dependendo da aplicação final [15]. Na Fig. 3.3 é apresentado o método de avaliação de segmentação do CC em difusão normalmente usado na literatura.

Zhang [115] denomina este método de avaliação apresentado na Fig. 3.3, método de avaliação empírica de discrepância. Este tipo de avaliação é feita comparando a segmentação com uma referência, que neste trabalho denominamos padrão, usando uma métrica que traduz quantitativamente a proximidade entre as duas estruturas: segmentação e padrão. É importante dizer que o padrão é uma segmentação, idealmente com uma qualidade melhor, cujo objetivo é servir como referência para avaliar outras segmentações.



Figura 3.3 – Método usual de avaliação de segmentações do CC em difusão por sobreposição

O esquema da Fig. 3.3 tem três componentes: a segmentação a ser avaliada (vermelho), o padrão usado como referência para fazer a avaliação (azul) e a avaliação usando alguma métrica quantitativa (amarelo).

A seguir, detalhamos os componentes da metodologia de avaliação da Fig. 3.3: métodos de segmentação (Sec. 3.1), padrão (Sec. 3.2) e métricas de avaliação (Sec. 3.3).

3.1 Segmentação em difusão

Os métodos de segmentação que fazem parte do estado da arte, apresentados na seção 2.5, serão implementados para obter as segmentações em difusão. Neste trabalho serão avaliadas as máscaras binárias (segmentação em difusão) obtidas executando os métodos de segmentação. Porém, os resultados obtidos são válidos quando aplicados em qualquer segmentação em difusão. Os resultados também podem ser extrapolados para outras modalidades e para outros domínios de imagens.

3.2 Padrão em difusão

Neste trabalho denominamos padrão à referência com a qual uma segmentação é comparada. Ainda que conceitualmente o GT é a referência ideal para avaliar segmentações, esta é difícil de conseguir na prática, já que só existe em condições ideais como no caso de imagens sintéticas e *phantoms*. Portanto, na literatura, o padrão em segmentação usado é comumente uma segmentação manual chamada padrão-ouro [115], como sendo o melhor padrão possível de se obter na prática. O desenho manual de um padrão pode ser uma tarefa extremamente árdua e levar bastante tempo, mesmo para uma pessoa experiente [75]. Além disso, é um processo propenso a erros, especialmente em imagens de difusão, devido à baixa resolução e pouco contraste entre as estruturas.

Seguindo as abordagens da literatura, e considerando o fato que as imagens em T_1 costumam ter melhor resolução, definição e contraste do que as imagens de difusão, neste estudo é considerado como padrão-ouro a segmentação manual em T_1 . Porém, as

segmentações que estamos interessados em avaliar estão no espaço de difusão, o que traz um problema: a comparação e avaliação não podem ser feitas diretamente, por sobreposição, já que as segmentações e o padrão-ouro estão em espaços (modalidades) diferentes.

Para superar este problema, uma alternativa é o uso direto do padrão-ouro junto com uma métrica que não seja aplicada pixel a pixel (que não faça uso de sobreposição). Uma segunda alternativa é o registro do padrão-ouro para o espaço de difusão. O registro é um processo de transformação de um conjunto de pontos em um dado sistema de coordenadas para um outro conjunto de coordenadas 3.1 [12]. No caso de uma transformação linear, o registro é normalmente feito usando uma matriz de transformação que permite mapear cada ponto de um domínio para um codomínio. Normalmente este mapeamento requer interpolação já que nem todo ponto no domínio tem correspondente no codomínio.

$$I_T(x, y) = I(f(x, y)) \quad (3.1)$$

onde a imagem I é registrada para imagem I_T a través da função de transformação f .

O registro é um dos temas mais ativos na comunidade de imagem médica por ser um procedimento complexo, que apresenta muitas deficiências e desafios e que ainda demanda melhoras significativas em termos de eficiência computacional e acurácia final [84]. As ferramentas disponíveis costumam acrescentar erros na estrutura registrada, são difíceis de configurar e estas configurações não generalizam bem.

Dito isso, ao registrar o padrão-ouro para o espaço de difusão, este perde suas características ideais e passa a ser um padrão aproximado, que é chamado padrão-prata. O padrão-prata em difusão é uma boa aproximação ao padrão-ouro, cujo objetivo é permitir uma avaliação direta, por sobreposição, das estruturas já que se encontra no espaço de difusão. A segmentação manual em T_1 e logo registrada no espaço de difusão é então o nosso primeiro padrão-prata.

Devido aos problemas de registro, padrões-prata alternativos podem ser obtidos por outros métodos: desenho da segmentação diretamente no espaço de difusão, com auxílio de ferramentas semi-automáticas de segmentação, e construção do padrão prata a partir de várias segmentações base usando estratégias de acordo entre anotadores.

O desenho direto no espaço de difusão costuma ser um processo mais difícil e propenso a erros, quando comparado com o desenho sobre imagens de T_1 , devido à menor resolução e contraste das imagens de difusão. Neste caso, o desenho é auxiliado por ferramentas semi-automáticas. A segmentação em difusão usando uma ferramenta semi-automática é nosso segundo padrão-prata.

Estratégias de consenso entre anotadores costumam ter bons resultados, mas trazem problemas de bias (tendência do modelo a preferir um anotador) e variância (variação natural de um anotador para o outro) [59]. Dentre as estratégias de consenso entre anotadores, as mais simples são de interseção (pixels comuns marcados pela totalidade de anotadores) [56], voto simples e voto ponderado baseados em informação local, global e semi-local [59].

Estas estratégias de consenso entre anotadores ainda que simples resultam insuficientes para gerar bons padrões, e por tanto os modelos estatísticos são mais robustos e geram melhores anotações. O mais conhecido dos modelos estatísticos para acordo entre anotadores é o STAPLE (Simultaneous Truth and Performance Level Estimation), um algoritmo robusto e amplamente usado para construção de padrão-prata [111]. O algoritmo STAPLE toma uma coleção de segmentações e calcula um mapa de probabilidade da segmentação real e uma avaliação, usando sensibilidade e especificidade, de cada segmentação. A partir das segmentações de entrada, e uma possível estimativa a priori do desempenho de cada uma, estima-se a densidade de probabilidade condicional do desconhecido padrão-prata. Com essa densidade, o algoritmo calcula o valor esperado de probabilidade do modelo (segmentações e padrão-prata). Logo após, usando este valor estimado, os parâmetros de desempenho das segmentações são achados por estimativa da máxima verossemelhança. Neste ponto é aplicado um algoritmo iterativo de estimativa da probabilidade condicional do padrão-prata e os parâmetros de desempenho até lograr convergência. O padrão-prata construído a partir de STAPLE é o nosso terceiro padrão-prata.

3.3 Avaliação

As métricas permitem a avaliação quantitativa dos métodos de segmentação quando comparados com um padrão. A RSL, permitiu identificar as métricas mais usadas na literatura para avaliação de segmentações do CC em MRI, sendo classificadas em 5 grupos: métricas de sobreposição, métricas de classificação, métricas de distância, métricas básicas e métricas específicas de cada estudo (Tab. 3.1).

A Tabela 3.1 mostra que as métricas preferidas na literatura para avaliar métodos de segmentação do CC são de sobreposição, seguidas pelas de classificação e, em menor grau, métricas de distância. Estas três classes de métricas necessitam da sobreposição da segmentação com o padrão. As métricas básicas têm pouco uso e já métricas mais específicas foram propostas como parte de um estudo individual.

As métricas de sobreposição permitem medir o grau de sobreposição de duas estruturas: a segmentação (que é a saída do método) e o padrão. Este tipo de métrica pode ser descrita usando a matriz de confusão [32] (Fig. 3.4). A matriz de confusão

Tabela 3.1 – Métricas usadas na literatura (Tabela baseada em 36 estudos da RSL) para avaliação da segmentação do CC em MRI, junto com a quantidade de estudos em que foi implementada

Tipo de métrica	Métricas implementadas (Número de vezes implementada)
Sobreposição	Dice (11), Overlap (7), Jaccard (4), Kappa (2) Coeficiente de Rogers e Tanimoto (1), Similaridade Hamann (1)
Classificação	Sensibilidade (7), Acurácia (7), Especificidade (4) Fração de falsos positivos (3), <i>Simple Matching</i> (2) Fração de falsos negativos (1)
Distância	Distância Hausdorff (7) ¹ , Erro de distância (5) Raiz quadrática média do erro (2), Erro máximo (2)
Básicas	Área (2), Volume (1)
Específicas	Relação de fibras (1), Erro de sobreposição (1)

permite quantificar os pixels corretamente classificados: Verdadeiros Positivos (VP) e Verdadeiros Negativos (VN), e pixels erroneamente classificados: Falsos Positivos FP e Falsos Negativos (FN).

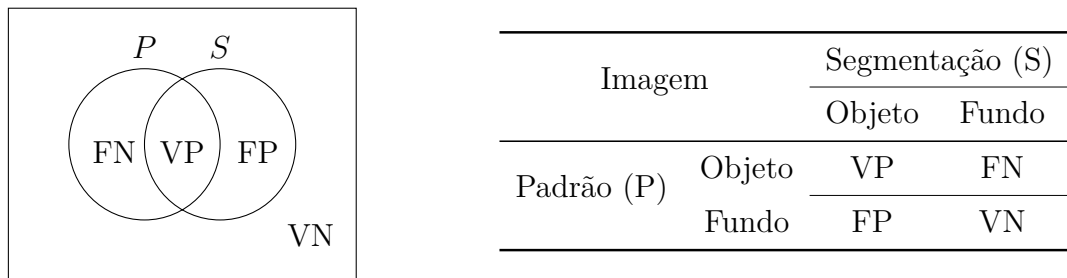


Figura 3.4 – Matriz de confusão que permite a visualização do desempenho do método de segmentação (S) em relação a um padrão (P) para uma imagem dada

Entre as métricas de segmentação, o Dice (Eq. 3.2) é a métrica mais usada, em grande parte, devido a sua versatilidade. Esta métrica é sensível a mudanças na matriz de confusão, uma característica desejável, dando ênfase na detecção da variação no número de VP (apropriada para avaliar resultados com interesse sobre a classe positiva).

$$Dice = \frac{2 \times VP}{2 \times VP + FP + FN} \quad (3.2)$$

Levando-se em conta que um tipo específico de métrica não consegue expressar todas as características de uma segmentação pode ser interessante utilizar métricas de diferentes tipos.

Serão usadas, portanto, duas métricas de classificação, também amplamente usadas na literatura: a sensibilidade (Eq. 3.3), que mede a fração de pixels corretamente classificados como objeto (classe positiva), e a especificidade (Eq. 3.4), que é complementar

à métrica anterior e mede a fração de pixels corretamente classificados como fundo (classe negativa).

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (3.3)$$

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (3.4)$$

Como métrica de distância, a Distância Hausdorff (DH) mede a proximidade entre os contornos da segmentação (S) e o padrão (P) (Eq. 3.5) [66].

$$DH = \max \left(\max_{s \in S} |d_s|, \max_{p \in P} |d_p| \right) \quad (3.5)$$

onde as variáveis aparecem representadas na Fig. 3.5. Esta equação representa a máxima distância espacial (usando distância euclidiana) entre P e S . O problema deste tipo de métrica é a sua sensibilidade aos valores atípicos (*outliers*).

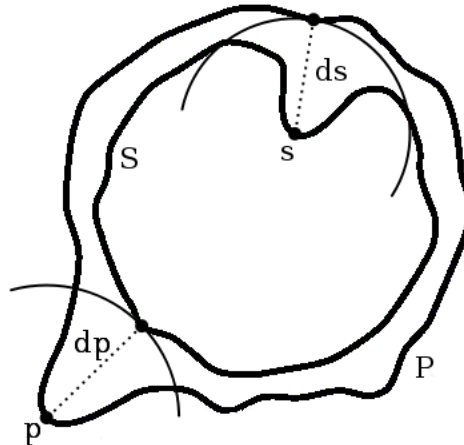


Figura 3.5 – Cálculo da Distância Hausdorff entre os contornos da segmentação e o padrão, onde p e s são pontos ao longo dos contornos P e S , respectivamente. d_p e d_s são as distâncias do ponto p e s aos pontos mais próximos nos contornos S e P , respectivamente.

Finalmente, a área, sendo uma métrica básica, mede a relação da quantidade de pixels entre duas estruturas (Eq. 3.6).

$$A = \frac{A_s}{A_p} = \frac{VP + FP}{VP + FN} \quad (3.6)$$

onde A_s e A_p representam as áreas da segmentação e do padrão, respectivamente, sendo calculadas como a soma de pixels contidas nas respectivas estruturas.

4 Método proposto de segmentação do corpo caloso baseado em classificação por pixel

Um método baseado em classificação por pixel foi proposto [42, 43]. O método consiste de duas etapas (Fig. 4.1): uma etapa de **Classificação** com dois componentes, um algoritmo não supervisionado *k-means* para sub-amostragem da classe majoritária e um modelo supervisionado, usando uma Máquina de Vetores Suporte (SVM), que classifica cada pixel da imagem como pertencente à classe corpo caloso (pCC) ou à classe não corpo caloso (nCC); e uma etapa de **Segmentação** que consiste em um pós-processamento baseado em tratamento morfológico de imagens.

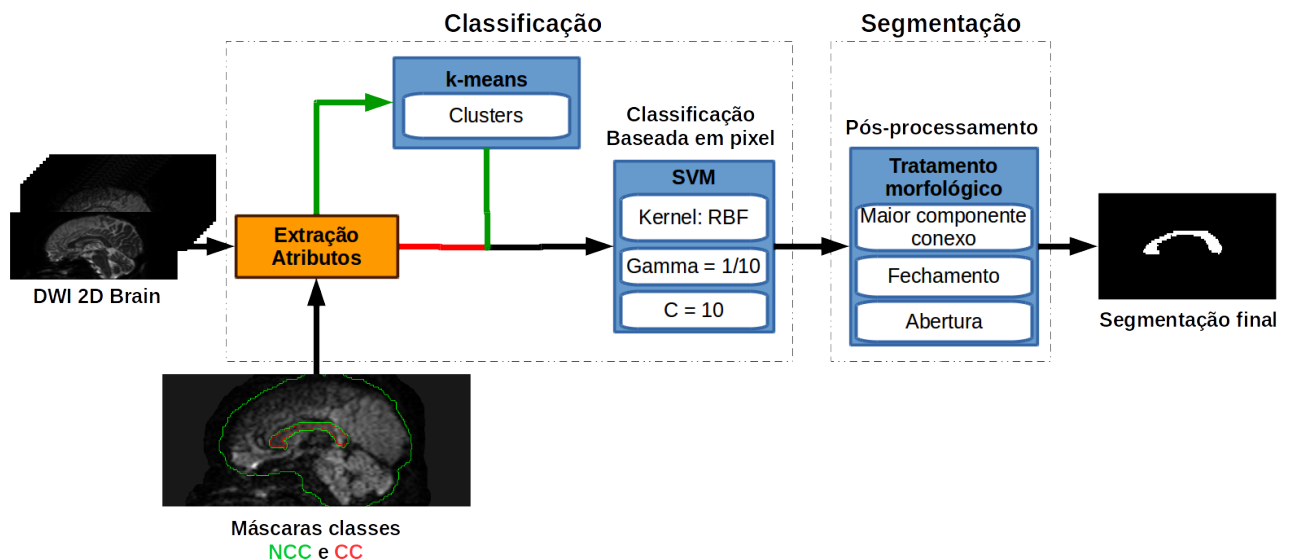


Figura 4.1 – Esquema do método proposto de segmentação do corpo caloso baseado em classificação por pixel denotando as etapas de Classificação e Segmentação

A grande vantagem deste método em comparação com os métodos de segmentação do CC em difusão existentes na literatura é que este é implementado diretamente em DWI. Isto reduz o processamento já que não é necessário calcular tensores ou mapas escalares.

O modelo de classificação é baseado em pixels da imagem como amostras, sendo seus atributos as respostas ao gradiente de campo magnético aplicado em diferentes direções (valor de intensidade em DWI) (Fig. 4.2). A fatia usada para treino e teste é a fatia central, extraída usando o método da Seção 2.5.

Outros atributos foram adicionados e testados como LBP (Padrões locais binários) das imagens e coordenadas x, y dos pixels sem lograr melhora na classificação final.

Para treinar o classificador SVM, as amostras são rotuladas em duas classes: pCC

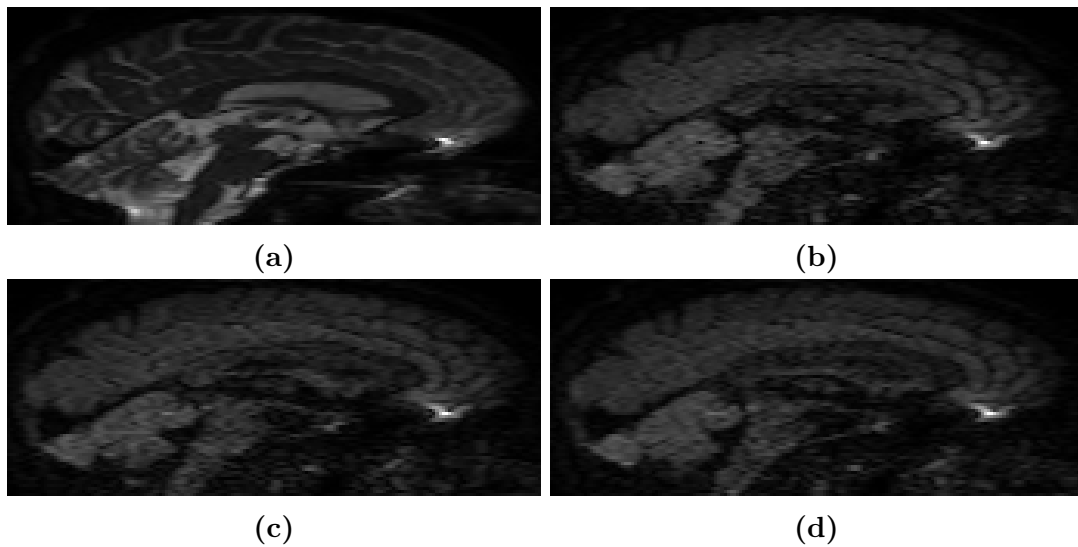


Figura 4.2 – Exemplos dos valores de intensidade DWI para diferentes gradientes de campo magnético usados como atributos: a) Imagem sem campo aplicado, b) Imagem para gradiente 10, c) Imagem para gradiente 20, d) Imagem para gradiente 30

e nCC. Esta rotulação é feita usando as máscaras do CC (pCC) e do cérebro (nCC), respectivamente. Uma vez que o tamanho do CC é bem menor em relação ao resto do cérebro, o número de amostras classe pCC é bem menor que o número de pixels da classe nCC resultando em um problema desbalanceado. Para lidar com este problema, a classe nCC deve ser sub-amostrada usando um método que garanta que as amostras escolhidas sejam representativas dessa classe.

Com essa finalidade, primeiro é usado o algoritmo de *k-means* para agrupar as amostras da classe negativa em *clusters* e depois é aplicada uma sub-amostragem aleatória por *cluster*. Desta forma, a classe nCC é reduzida, garantindo, ao mesmo tempo, que o conjunto de treino final contém amostras representativas dos tecidos nCC do cérebro.

Os *clusters* serão definidos usando *k-means* como visto na Seção 2.4.2. Uma desvantagem do *k-means* é a necessidade de definir a priori o número k de *clusters* em que as amostras serão particionadas. k será definido empiricamente, aumentando iterativamente o número k de *cluster* até chegar num valor em que o *cluster* adicional contenha poucas amostras em comparação com os outros. Este valor indica que as sub-classes possíveis já foram definidas e que *clusters* adicionais fazem parte dos *clusters* já definidos.

Uma vez definido o conjunto de treino final, uma SVM pode ser treinada para segmentar novas imagens usando classificação por pixel, como visto na Seção 2.4.1. A configuração da SVM, que consiste principalmente na escolha e configuração do *kernel* e do parâmetro de penalidade, é um processo complexo que é melhor feito fazendo uma busca de hiper-parâmetros. Para garantir uma boa generalização do modelo, o treinamento da SVM é feito usando uma estratégia de validação cruzada com k -pastas.

Uma vez escolhidos os hiper-parâmetros, a SVM é treinada e assim, todos os pixels da fatia central de uma nova imagem podem ser classificados como pCC ou nCC usando a função de decisão (Eq. 2.11).

O resultado da etapa de classificação usando a SVM é uma matriz de amostras com o rótulo predito pelo classificador. Essas amostras são alocadas novamente na imagem usando as coordenadas x e y associadas a cada uma. Dessa forma, uma imagem binária é obtida.

A imagem binária ainda não é a segmentação final já que contém componentes classificados como CC que devem ser eliminados. Para obter uma segmentação homogênea com um único elemento, a imagem binária é passada pela etapa de segmentação usando tratamento morfológico em três passos: o maior componente conexo é escolhido (selecionando o componente com maior área), uma operação de fechamento morfológico é aplicada para resolver possíveis buracos na estrutura e uma operação de abertura é aplicada para limpar as bordas da segmentação final.

5 Método proposto de avaliação de segmentações baseado no perfil de curvatura

O método usual de avaliação por sobreposição, no caso de avaliação de segmentações no espaço de difusão, apresenta o problema de uso de um padrão-prata, referência que não é ideal. Para resolver esta questão, é proposto um método de avaliação cuja maior contribuição é permitir o uso do padrão-ouro em T_1 usando um descritor de forma denominado perfil de curvatura. A avaliação de segmentações usando este perfil evita o uso de registro de imagens já que permite a comparação de perfis sem necessidade de sobreposição.

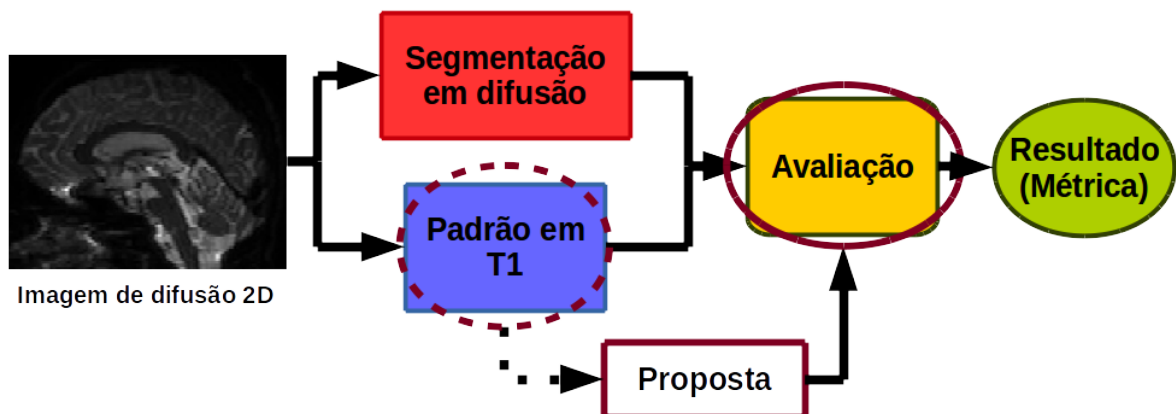


Figura 5.1 – Etapas do método proposto de avaliação de segmentações do CC em difusão. O método proposto é similar ao método usual (Fig. 3.3); porém, a diferença neste caso é que o padrão usado está em T_1 e é usado diretamente sem necessidade de registro

O método proposto (Fig. 5.1), é similar ao método usual de avaliação por sobreposição (Fig. 3.3). Porém, o fato de mudar o padrão-prata em difusão para o padrão-ouro em T_1 , leva à melhoria do método e implica mudar a métrica de avaliação já que não é possível sobrepor as estruturas (padrão e segmentação). Dado que o padrão-ouro em T_1 já foi descrito na Seção 3.2, centraremos a discussão deste capítulo no perfil de curvatura proposto e na métrica utilizada para comparação de perfis. Primeiro, é descrito o método de perfil de curvatura de forma geral (Sec. 5.1); depois, é apresentado o método de construção do perfil (Sec. 5.2); na sequência, a métrica RMSE baseada no perfil é apresentada (Sec. 5.3); e finalmente, é descrita a definição de parâmetros do perfil de curvatura para avaliação de segmentações (Sec. 5.4).

5.1 Perfil de curvatura

O perfil de curvatura é um descritor de forma que mede a taxa de variação ao longo do contorno do objeto [87]. A curvatura k no ponto p , com coordenadas (x_p, y_p) , pode ser calculada usando-se a Eq. 5.1 e representa o ângulo que existe entre os segmentos $\overline{(x_{p-ls}, y_{p-ls})(x_p, y_p)}$ e $\overline{(x_p, y_p)(x_{p+ls}, y_{p+ls})}$, localizados em uma vizinhança $ls > 0$. Dado que os pontos anterior e posterior, localizados em (x_{p-ls}, y_{p-ls}) e (x_{p+ls}, y_{p+ls}) respectivamente, estão posicionados sobre o contorno, o ângulo resultante mede a curvatura no ponto p (x_p, y_p) (Fig. 5.2). Quando calculada a curvatura para todo ponto p ao longo do contorno da estrutura, é obtido seu perfil de curvatura: um vetor onde cada componente é o ângulo de curvatura no ponto p .

$$k(x_p, y_p) = \arctan\left(\frac{y_{p+ls} - y_p}{x_{p+ls} - x_p}\right) - \arctan\left(\frac{y_p - y_{p-ls}}{x_p - x_{p-ls}}\right) \quad (5.1)$$

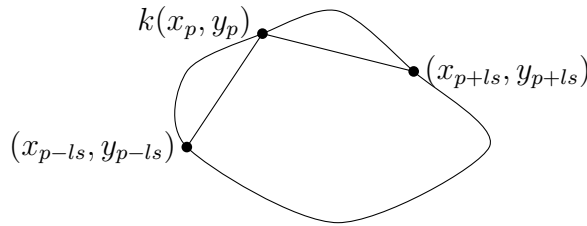


Figura 5.2 – Esquema da curvatura k para um contorno no ponto i com coordenadas (x_p, y_p)

É possível encontrar na literatura algumas propostas de representação de objetos planares utilizando perfil de curvatura [79, 2, 52]. Em todos estes casos, o perfil de curvatura foi usado como descritor de forma para aplicações de recuperação de imagem por conteúdo. Até agora, nenhum descritor de forma tem sido usado para comparar segmentações de MRI em diferentes espaços.

Neste caso, o objeto de interesse é a segmentação do CC (sendo segmentação ou padrão). O perfil de curvatura permite a comparação de segmentações a través da forma. As vantagens deste métodos são várias:

- Permite a comparação de segmentações em espaços (modalidades de imagens) diferentes, dado que a forma se mantém sempre que as segmentações provenham da mesma estrutura
- É um método versátil já que, como veremos nas próximas seções, o perfil de curvatura pode ser levantado em várias resoluções (capturando mais ou menos detalhes da estrutura)
- Já que o perfil de curvatura é uma curva 2D, podemos aproveitar métricas já estabelecidas na literatura para comparar quantitativamente curvas 2D.

- Sob algumas condições, o perfil de curvatura é invariante a perturbações tais como rotação, traslação e escala.

5.2 Construção do perfil de curvatura

A construção do perfil de curvatura é feita a partir do contorno da segmentação do CC (Fig. 5.3a). Este contorno é achado aplicando a operação lógica XOR, pixel a pixel, entre a segmentação original (S) e a segmentação erodida por um elemento estruturante b com forma de cruz de tamanho 1, usando a Eq. 5.2 (Fig. 5.3b).

$$G_E = XOR(S, S \ominus b) \quad (5.2)$$

A operação lógica XOR compara as duas segmentações (original e erodida) e retorna aqueles pixels que mudaram durante a erosão, que são precisamente aqueles da borda. Depois, é calculada a *spline* (Fig. 5.3c) do contorno com o objetivo de suavizá-lo e obter uma representação mais conveniente para o cálculo da curvatura mediante funções polinomiais.

Mediante *spline*, um contorno C pode ser representado como combinação linear de funções polinomiais \mathbf{p} de grau g que dependem de um parâmetro t [72]:

$$C(t) = \sum_{i=g+1}^n \mathbf{p}_{i,g}(t) B_{i,0}(t), \quad t \in [t_{g+1}, t_{n+1}] \quad (5.3)$$

$$\text{onde, } B_{i,0}(t) = \begin{cases} 1, & t_i \leq t \leq t_{i+1}, \\ 0, & \text{de outro modo,} \end{cases} \quad (5.4)$$

Neste caso, temos n pontos de controle e $n + d - 1$ nós para as *splines* de grau g que conformam o contorno. Qualquer ponto ao longo do contorno pode ser acessado de forma paramétrica o que facilita o cálculo do ângulo entre dois pontos. Assim, a curvatura é calculada em um ponto p usando a Eq. 5.5 onde ls determina a resolução do perfil de curvatura, já que define a distância paramétrica entre o ponto p (que denominaremos ponto pivô) e os pontos anterior e posterior. O perfil de curvatura é obtido quando calculada a curvatura para todos os pontos ao longo da *spline* (Fig. 5.3d).

$$k(p) = \arctan \left(\frac{C(p+ls) - C(p)}{C(p+ls) - C(p)} \right) - \arctan \left(\frac{C(p) - C(p-ls)}{C(p) - C(p-ls)} \right) \quad (5.5)$$

O parâmetro ls tem grande importância no perfil de curvatura e será denominado: **largura de segmento**. Por simplicidade, ls será dada em porcentagem da quantidade de pontos de construção da *spline*; assim por exemplo, se a *spline* é construída com quinhentos

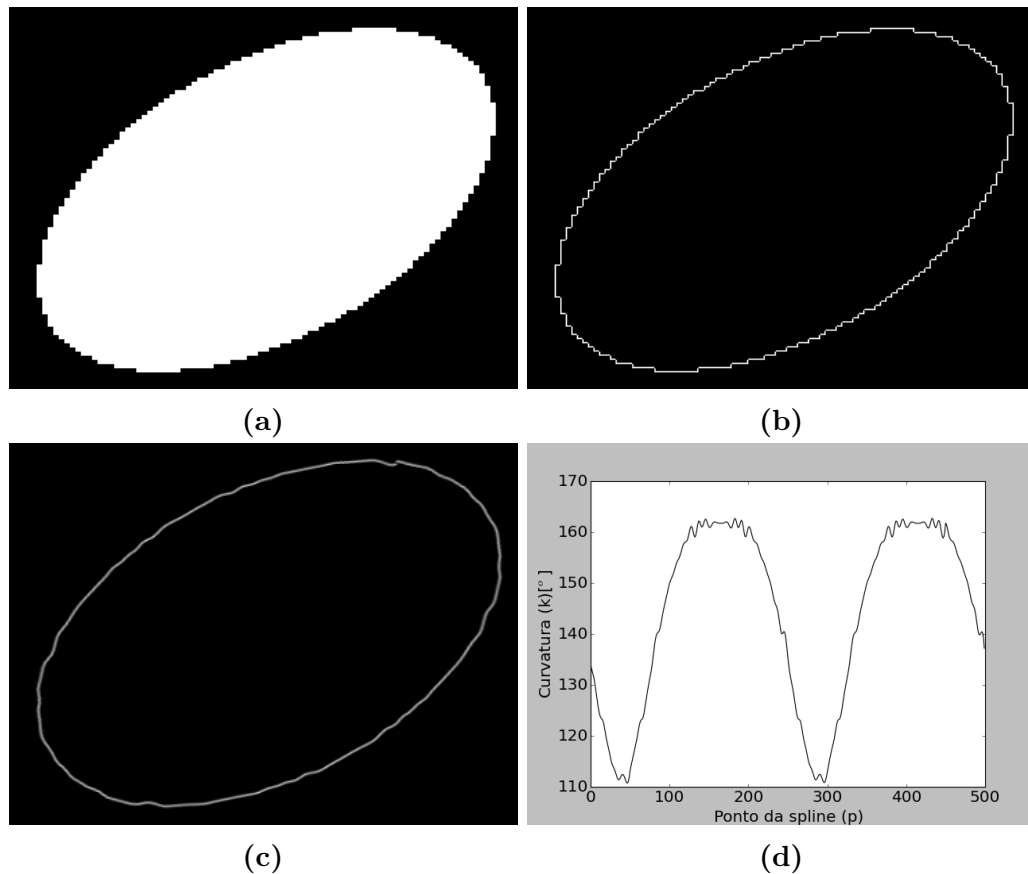


Figura 5.3 – Etapas de construção do perfil de curvatura para uma figura sintética: a) Figura binária de um elipse, b) Cálculo do contorno da máscara binária, c) Cálculo da spline a partir do contorno, d) Perfil de curvatura ao longo da spline com 500 pontos ($P = 500$)

pontos ($P = 500$), uma ls de 0.1, significa que o ponto anterior está 50 pontos atrás (em sentido anti-horário) do ponto pivô e o ponto posterior está 50 pontos na frente (em sentido horário) do ponto pivô. Assim, para uma quantidade P de pontos de construção da *spline*, ls pode variar segundo a Equação 5.6.

$$\frac{1}{P} < ls < \frac{1}{2} \quad (5.6)$$

onde $1/P$ equivale a um ponto na *spline*, 0.5 equivale a $P/2$ pontos da *spline* e P é o número total de pontos da *spline*. Uma ls de 0.5 significa um perfil de curvatura igual a zero para todos os pontos, já que os segmentos correspondentes ao ponto anterior e o ponto posterior, encontraram-se na metade da curva dando como resultado um ângulo zero entre eles para qualquer ponto pivô.

5.3 Cálculo do RMSE

O perfil de curvatura serve para descrever a forma da estrutura segmentada. Porém, para fazer a avaliação quantitativa de uma segmentação a partir do perfil de curvatura

é necessário comparar o perfil de curvatura associado à segmentação com o perfil de curvatura do padrão através de uma métrica para comparação de curvas 2D. A métrica escolhida foi a raiz do erro médio quadrático (RMSE) entre os perfis de curvatura a e b (Eq. 5.7) ao longo de todos os pontos p dos perfis.

$$RMSE = \sqrt{\frac{1}{P} \sum_{p=1}^P (k_{ap} - k_{bp})^2} \quad (5.7)$$

A partir de um padrão (Fig. 5.4a) é calculado o perfil de curvatura (Fig. 5.4b) como indicado na seção 5.2. Dado que o cálculo do perfil de curvatura pode começar em um ponto arbitrário ao longo da *spline*, é provável que as curvas não estejam casadas e portanto seja impossível compará-las diretamente (Fig. 5.4c). Para solucionar esta questão, as curvas devem ser casadas: mantém-se uma das curvas fixa e a outra é deslocada ao longo do eixo x ; para cada deslocamento, o RMSE entre os dois perfis é medido. O ponto de casamento é onde o RMSE é mínimo (Fig. 5.4d). Por convenção a curva fixa é o padrão e a curva deslocada é a segmentação. O processo de casamento é descrito com detalhe na Sec. 5.4.2. Uma vez casadas as curvas, pode-se calcular o RMSE, que permite avaliar de forma quantitativa a similaridade na forma das duas estruturas segmentadas.

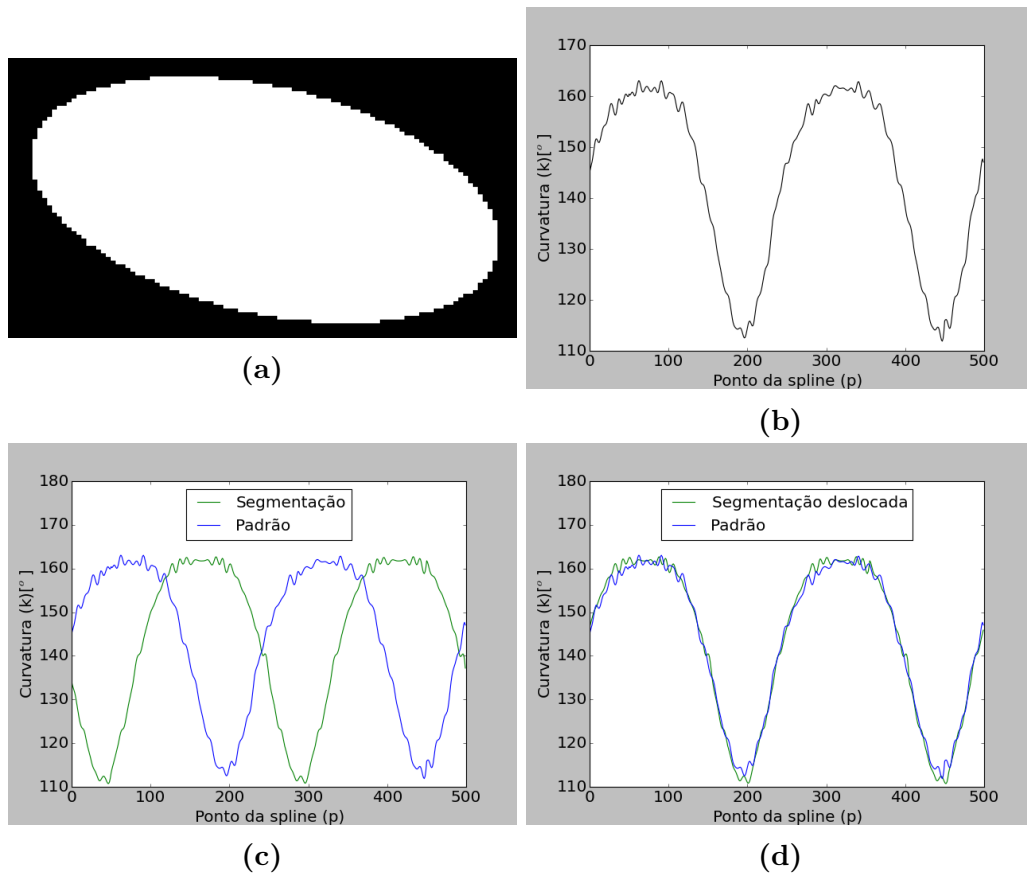


Figura 5.4 – Etapas para cálculo de RMSE a partir do perfil de curvatura: a) Figura binária da elipse padrão, b) Perfil de curvatura do padrão, c) Comparação dos perfis de curvatura da segmentação e do padrão, d) Alinhamento dos perfis de curvatura para cálculo do RMSE

Existem outras métricas para comparar perfis de curvatura. Neste caso, foram testadas as seguintes métricas:

- Distância Minkowski (Eq. 5.8), que mede a distância normalizada, ponto a ponto, entre duas curvas [113]. Neste caso foi testada a distância Minkowski com $n = 1$ e $n = 3$. O RMSE é um caso especial da distância Minkowski quando $n = 2$.

$$MK = \sqrt[n]{\frac{1}{P} \sum_{p=1}^P |k_{ap} - k_{bp}|^n} \quad (5.8)$$

- Entropia cruzada (Eq. 5.9), usada em redes neurais como função de custo [36].

$$CE = \frac{1}{P} \sum_{p=1}^P (k_{ap} \ln(k_{bp}) + (1 - k_{ap}) \ln(1 - k_{bp})) \quad (5.9)$$

- Relação da energia de curvatura (Eq. 5.10), que é a relação entre as funções de energia de duas curvas. As curvas representam a curvatura de uma estrutura [67].

$$BER = \frac{BE_b}{BE_a} \quad (5.10)$$

onde $BE_n = \frac{1}{P} \int_{p=1}^P (k_{np})^2 dp$, é a energia de curvatura da curva n

Estas três métricas mostraram uma correlação alta com o RMSE quando usadas para avaliar 3 segmentações sobre 145 sujeitos (Tab. 5.1). Além de estarem fortemente correlacionadas, algumas das métricas apresentam restrições matemáticas como o cálculo do \ln para a entropia cruzada. Portanto, não vimos razão para implementá-las.

Tabela 5.1 – Correlação de métricas alternativas (Distância Minkowski, Entropia cruzada e Relação de energia de curvatura) com o RMSE usando o valor absoluto do coeficiente de Pearson.

Distância Minkowski ($n = 1$)	Distância Minkowski ($n = 3$)	Entropia cruzada	Relação energia de curvatura
0.99	0.99	0.96	0.87

5.4 Definição de parâmetros do perfil de curvatura adequados para avaliação de segmentações

O perfil de curvatura é um descritor de forma versátil, o que pode ser visto a partir de uma análise qualitativa (Apêndice A). A versatilidade deste descritor está baseada no fato de poder ser calculado em diferentes resoluções, através da escolha de diferentes valores de ls . Valores pequenos de ls (maior resolução) descrevem o detalhe da borda da

estrutura, enquanto valores altos de ls (menor resolução) descrevem a forma global da estrutura.

A definição de parâmetros para construção do perfil de curvatura foi feita visando a comparação de diferentes segmentações do mesmo sujeito. O perfil de curvatura para outras aplicações, como por exemplo, para estudos populacionais, pode requerer diferentes valores de parâmetros.

Três tipos de parâmetros devem ser definidos antes de usar o RMSE para avaliação de segmentações a partir do perfil de curvatura: parâmetros do perfil de curvatura em si, resolução para casamento dos perfis e resolução para comparação dos perfis propriamente dita. Os parâmetros do perfil de curvatura vêm diretamente da implementação computacional da *spline* e do perfil de curvatura e serão descritos na Seção 5.4.1. Na sequência, a resolução para casamento permite posicionar os perfis de curvatura, um em relação ao outro, de forma que sejam comparáveis. Espera-se que o casamento seja melhor feito em resoluções menores dos perfis onde as curvas apresentam pontos comuns de casamento e onde o casamento seja robusto (pouca ou nenhuma variação no RMSE diante de pequenas variações no ls de casamento). Esta discussão será apresentada na Seção 5.4.2.

Finalmente, uma vez casados os perfis, deve ser definida a resolução para comparação de perfis, conforme será apresentado na Sec. 5.4.3. Esta resolução é a que maior importância tem no processo de avaliação de segmentações. Dado que o objetivo é comparar segmentações do mesmo sujeito, onde a forma global é a mesma e as diferenças existem na borda, é desejável que a resolução de comparação seja alta o suficiente para diferenciar segmentações, mas não alta demais a ponto de comparar ruído de bordas.

5.4.1 Definição dos parâmetros do perfil de curvatura

Para definir os parâmetros do perfil de curvatura, tomaremos como base um exemplo: uma segmentação baseada em *Watershed* (Fig. 5.5a). A partir desta segmentação, é calculado o contorno da estrutura usando a Eq. 5.2 (Fig. 5.5b). A partir do contorno, é calculada a *spline*. Neste caso, a implementação é uma *spline básica de Boor* ou *B-spline* [23], que tem dois parâmetros importantes: o grau g , que representa o grau dos polinômios que conformam a *spline*, e a suavidade que é a relação de compromisso entre proximidade e suavidade no ajuste da *spline*. Os parâmetros foram escolhidos empiricamente, usando a implementação do Scipy[53]: o grau da curva deve ter um valor entre um e cinco ($1 < g < 5$); neste caso, o grau escolhido foi cinco ($g = 5$) permitindo obter uma boa representação do contorno. Da mesma forma, a suavidade foi fixada em setecentos ($suavidade = 700$), valor baseado na quantidade média de pixels que formam o contorno das máscaras de CC e que são passadas para o cálculo da *spline*. A *spline* é reconstruída usando quinhentos pontos, valor suficiente para ter uma representação

adequada da estrutura (Fig. 5.5c).

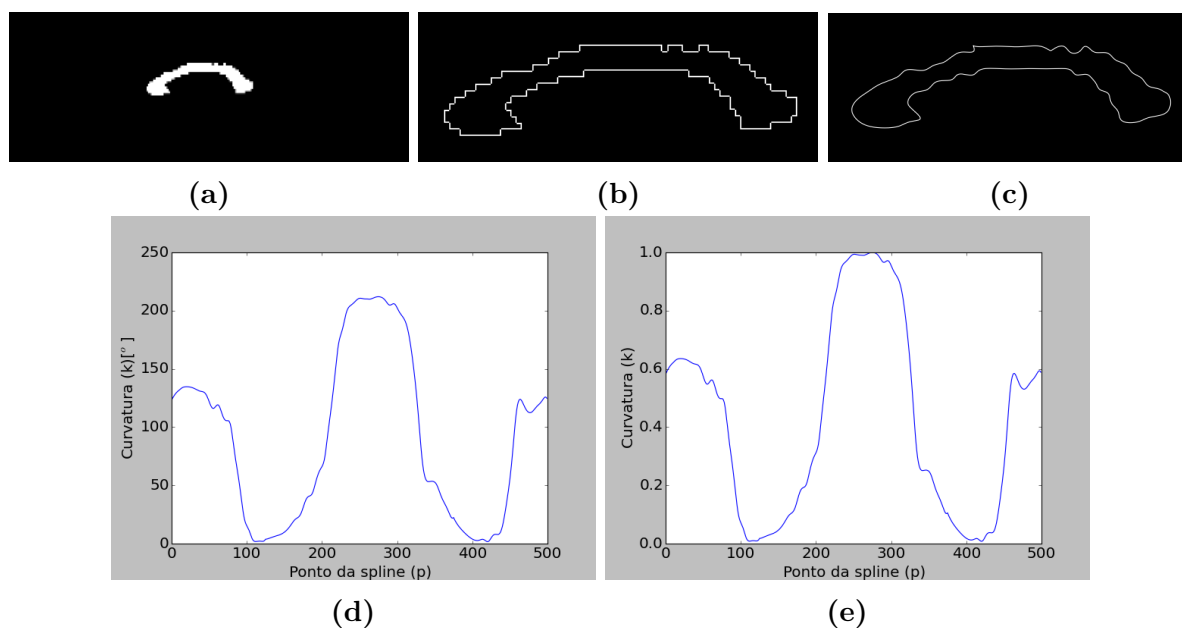


Figura 5.5 – Perfil de curvatura do CC com 500 pontos para um sujeito: a) Segmentação baseada em *Watershed* do CC, b) Contorno do CC, c) *Spline* do CC com *grau* = 5 e *suavidade* = 700, d) Perfil de curvatura ($ls = 0.25$), e) Perfil de curvatura normalizado ($ls = 0.25$)

Uma vez construída a *spline*, são calculados os valores de curvatura, para um ls dado, sobre os 500 pontos p ao longo da curva (Eq. 5.1) chegando no perfil de curvatura da estrutura. Até aqui, o perfil de curvatura pode ser construído em qualquer resolução, porém, este parâmetro será fixado nas próximas seções. Neste exemplo escolhemos $ls = 0.25$ (Fig. 5.5d). Finalmente, o perfil é normalizando, dividindo cada valor pelo maior valor de curvatura do perfil (Fig. 5.5e). Desta forma, como será visto na seção de caracterização do perfil como descritor (Sec. 6.5.1), garantimos que o perfil seja invariante à escala, uma propriedade importante num descritor de forma.

5.4.2 Definição do valor de ls para casamento de perfis

Uma vez obtidos os perfis de curvatura a serem comparados, tanto da segmentação quanto do padrão, é necessário fazer o casamento entre eles. O desafio é escolher em qual resolução (valor de ls) este casamento deve ser realizado. É importante destacar que após o casamento dos perfis, na etapa de comparação, um novo valor de ls será escolhido. Ou seja, a melhor resolução para o casamento do perfil não obrigatoriamente é a resolução mais adequada para a comparação final entre as curvas para cálculo do RSME.

O RMSE final depende tanto da resolução para casamento quanto da resolução para comparação e portanto este valor pode ser visualizado em um gráfico 3D, com o RMSE final no eixo z , o ls para casamento no eixo x e ls para comparação no eixo y

(Fig. 5.6a). O processo de estabelecer a resolução do perfil consiste em fixar um valor y (l_s de casamento) e depois determinar um valor x (l_s de comparação).

Intuitivamente, buscamos uma faixa de valores de l_s de casamento ($l_{s_{cas}}$) para a qual, uma vez casados os perfis, o RMSE final não varia consideravelmente. Isto garante um casamento robusto, ainda ante pequenas variações no $l_{s_{cas}}$. Calculando o gradiente vertical da imagem 3D anterior (Fig. 5.6a) podemos encontrar as variações do RMSE em função do $l_{s_{cas}}$. Este gradiente é calculado aplicando um filtro detector de borda vertical (filtro de Sobel) (Fig. 5.6b).

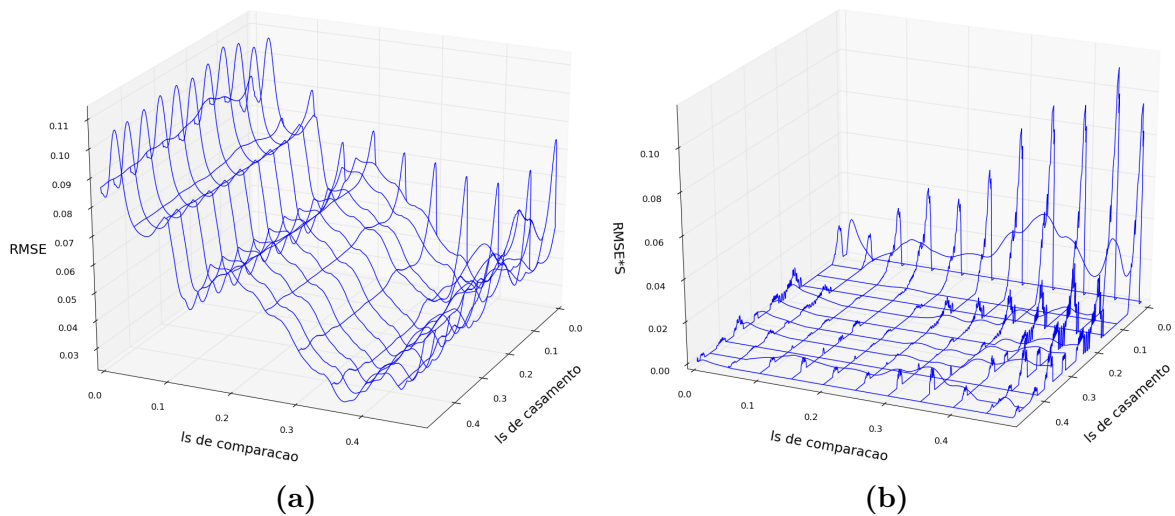


Figura 5.6 – Escolha dos valores de l_s : (a) RMSE em função de l_s de casamento e l_s de comparação, (b) Gradiente vertical da função RMSE aplicando um filtro Sobel vertical

Como esperado, valores menores de l_s para casamento ($l_{s_{cas}} < 0.2$) ainda apresentam alta variabilidade, especialmente para valores grandes de l_s para comparação. Já para valores maiores de 0.2, os valores de RMSE permanecem mais estáveis, uma vez que os perfis são casados por pontos comuns e bem estabelecidos. A Fig. 5.6b mostra uma região de pouca variação do RMSE (zona plana), que indica uma faixa de valores de l_s adequada para o casamento ($l_{s_{cas}} > 0.2$ e $0.05 < l_{s_{comp}} < 0.35$).

A forma mais direta para visualizar a melhor faixa de casamento é eliminando a influência da variável l_s de comparação do gráfico de gradiente de RMSE, já que neste momento não estamos interessados nela. Isto é feito somando-se os valores de l_s de comparação ao longo do l_s de casamento (extraídos da Fig. 5.6b), com o qual obtemos a Fig. 5.7a. Para generalizar o experimento, calculamos esta mesma curva para 16 sujeitos (aleatoriamente escolhidos da base de imagens), bem como sua média e o desvio padrão (Fig. 5.7b).

Como estamos interessados em uma faixa de valores de $l_{s_{cas}}$ que apresente pouca variação do RMSE, pode-se concluir que a faixa mais robusta para o casamento é $0.28 < l_{s_{cas}} < 0.45$ (Fig. 5.7). Escolhemos, portanto, o valor de $l_{s_{cas}} = 0.35$ para fazer

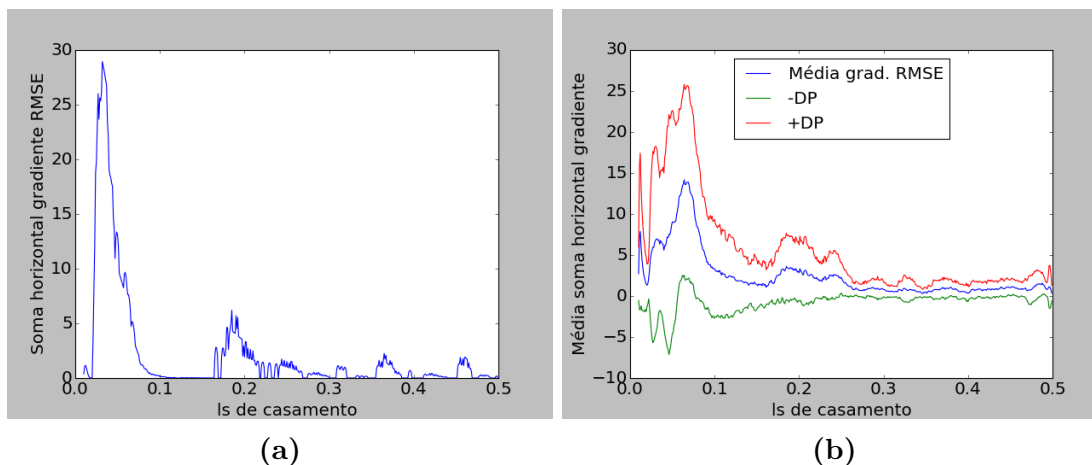


Figura 5.7 – Soma dos valores de gradiente em função de ls de casamento: a) Soma ao longo do eixo ls de casamento da imagem de gradiente (Fig. 5.6b), b) Média e desvio padrão da soma ao longo do eixo ls de casamento do gradiente para 16 sujeitos

o casamento dos perfis e, ao menos que se indique o contrário, este valor será usado ao longo deste trabalho.

5.4.3 Definição do valor de ls para comparação de perfis

Uma vez definido o valor de ls_{cas} para casamento dos perfis, precisamos definir o ls para comparação dos perfis de curvatura a partir do qual possa ser calculado o RMSE. Ressaltamos que a escolha de ls de comparação (ls_{comp}) depende do objetivo da avaliação, uma vez que diferentes resoluções permitem comparar diferentes aspectos das segmentações. Neste trabalho, a resolução escolhida deve ser sensível o suficiente para discriminar diferenças nas máscaras de segmentação e permitir comparação de diferentes segmentações do mesmo sujeito.

Para isso, foram criados dados sintéticos que permitam comparações de perfis com diversos valores de ls . Desta forma, temos controle sobre os dados criados e portanto sabemos quais comparações devem resultar em RMSE pequeno e quais devem apresentar um RMSE maior.

Escolhemos o padrão-prata, obtido a partir da segmentação manual em T1, registrado em difusão como segmentação base e comparamos seu perfil com perfis da mesma segmentação, mas que sofreram perturbações (sintéticas), listadas a seguir:

1. Segmentação rotacionada: Obtida a partir da máscara registrada rotacionada em 10°
2. Segmentação ampliada: Obtida a partir da máscara registrada ampliada por um fator de dois (tanto em altura quanto em largura). A ampliação nada mais é do que a replicação de cada pixel em uma coluna e uma linha.

3. Segmentação registrada: Obtida a partir do registro inverso da máscara registrada. Neste caso, a segmentação registrada é em realidade a máscara manual original em T_1 , que é equivalente a aplicar uma transformação inversa no padrão registrado em difusão para obter o original.

É esperado que para estas três segmentações sintéticas o RMSE seja baixo quando comparadas com a máscara base a partir da qual foram originadas. Como controle, a segmentação base será também comparada com segmentações independentes que estão no mesmo espaço (os outros dois padrões-prata em difusão, apresentados na Sec. 3.2):

1. Segmentação semi-automática em difusão: Método alternativo e independente
2. Segmentação STAPLE em difusão: Método alternativo e independente

Para estas últimas duas máscaras é esperado que o RMSE seja maior porquanto foram obtidas por métodos independentes. Para cada segmentação o perfil de curvatura é construído e comparado com o perfil da segmentação escolhida como base e o RMSE, para as 5 máscaras, é calculado em função do valor ls de comparação (Fig. 5.8a). A máscara com menor RMSE (mais similar) é a máscara ampliada já que a máscara mantém sua forma e proporção, ainda que exista algum grau de interpolação. A máscara registrada tem o segundo melhor resultado, já que é produto de uma transformação geométrica e não muda sua estrutura significativamente. Já a máscara rotacionada sofre uma transformação geométrica que muda a borda da estrutura e por isso o RMSE é relativamente maior, em relação às outras máscaras sintéticas. Ainda assim, para algumas resoluções, a máscara registrada tem um RMSE maior.

Para as duas máscaras independentes, semi-automática e STAPLE, como esperado, têm-se os maiores valores de RMSE. A segmentação semi-automática é a máscara menos similar à estrutura base.

A resolução de comparação ideal é aquela que dá a maior diferença entre as perturbações da máscara base e qualquer máscara independente já que estamos interessados em discriminar segmentações obtidas por métodos diferentes. Teoricamente, um valor de ls_{comp} adequado deveria dar a maior diferença entre o RMSE da pior das perturbações (rotacionada) e o RMSE da melhor das segmentações independentes (STAPLE) que são as curvas mais próximas e portanto mais difíceis de serem diferenciadas.

Por tanto, é calculada a diferença entre o RMSE da máscara rotacionada e o RMSE da máscara STAPLE em função de ls_{comp} . Avaliando a diferença entre os valores RMSE em função de ls_{comp} (Fig. 5.8b) concluímos que, para este sujeito em particular, $ls_{comp} = 0.1$ é o melhor valor para discriminar as segmentações.

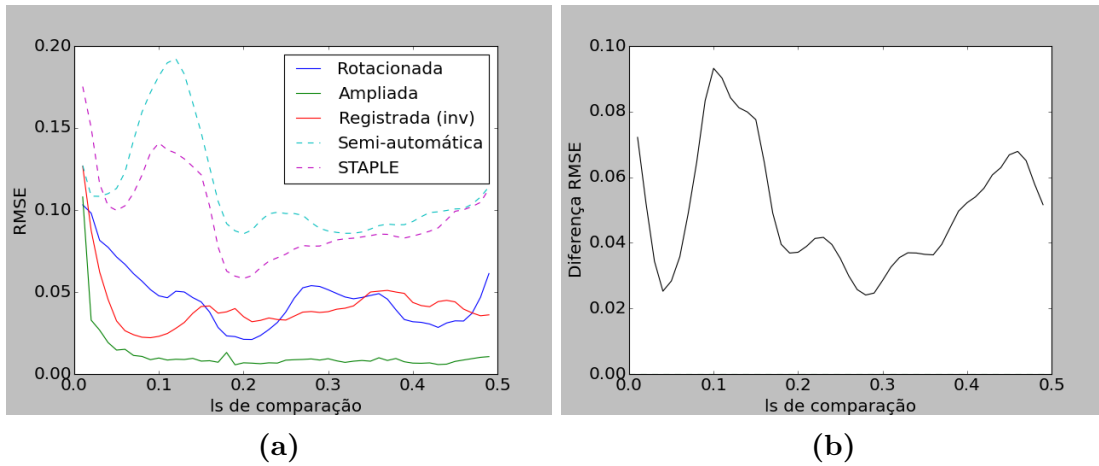


Figura 5.8 – Comparação de segmentações em função do valor de l_s de comparação: a) RMSE de 5 segmentações: 3 perturbações (registro inverso, ampliação e rotação) e 2 segmentações independentes (semi-automática e STAPLE), b) Diferença entre RMSE da máscara STAPLE e da máscara rotacionada

Com o propósito de generalizar este resultado, o mesmo experimento é reproduzido para 50 sujeitos: diferença da RMSE da máscara rotacionada e a da STAPLE em função do $l_{s_{comp}}$. Estes valores são analisados de forma estatística usando diagramas de caixa (Fig. 5.9)

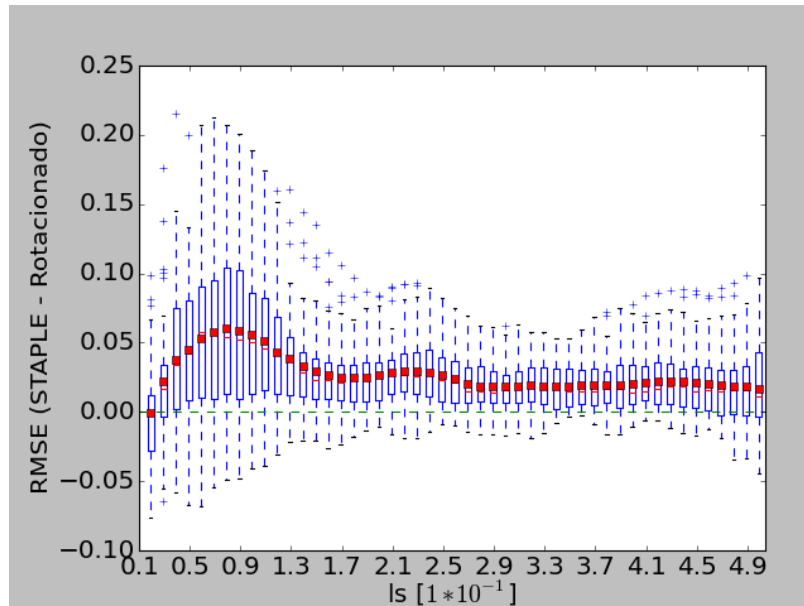


Figura 5.9 – Diferença de RMSE para máscaras rotacionada e STAPLE em função do valor de l_s de comparação

Em média, o valor de $l_{s_{comp}} = 0.08$ corresponde ao ponto onde a diferença RMSE é máxima (Fig. 5.9). Este valor está associado a dados com grande dispersão incluindo valores negativos que indicam um RMSE menor para o STAPLE em comparação com o RMSE para a segmentação rotacionada.

A análise realizada permitiu definir valores adequados para l_s , tanto para o

casamento dos perfis de curvatura ($ls_{cas} = 0.35$) quanto para a etapa de comparação de segmentações baseada no perfil ($ls_{comp} = 0.08$). Os experimentos indicaram que, utilizando os valores adequados de ls é possível utilizar o perfil de curvatura e o RMSE entre perfis para comparar segmentações do CC de um mesmo sujeito obtidas por métodos diferentes. O método permite comparar inclusive segmentações em espaços (modalidades de imagem) diferentes, no nosso caso, padrão-ouro em T1 e segmentação no espaço de difusão.

6 Resultados e Discussão

Os resultados apresentados neste capítulo estão centrados na execução do método proposto de avaliação de segmentações do CC para uma população de 145 sujeitos. Porém, para enriquecer a discussão, também serão apresentados resultados correspondentes às características das segmentações, dos padrões e da caracterização da métrica RMSE.

O capítulo está dividido em quatro seções: dados e ferramentas usados (Sec. 6.1), implementação das segmentações (Sec. 6.2), padrões (Sec. 6.3), avaliação de segmentações usando o método proposto (Sec. 6.4) e caracterização do perfil de curvatura e RMSE (Sec. 6.5).

6.1 Dados e ferramentas

Para avaliar as segmentações, dados de 145 sujeitos provenientes do Hospital das Clínicas da Universidade Estadual de Campinas (Unicamp) foram usados. Para cada sujeito foram adquiridos o volume DWI e o volume T_1 , ambos em um scanner Philips Achieva de 3T, com o qual, a base de dados está formada por:

- 145 volumes T_1 adquiridos no plano coronal com resolução espacial de $1.0 \times 1.0 \times 1.0 \text{ mm}$ ($TR = 7s$, and $TE = 3.2ms$) (Fig. 6.1a).
- 145 volumes DWI adquiridos no plano axial com resolução espacial de $1.0 \times 1.0 \times 2.0 \text{ mm}$ ao longo de 32 direções ($TR = 8.5s$, and $TE = 61ms$) (Fig. 6.1b).
- 145 máscaras automáticas do cérebro obtidas com o FSL [49] sobre o volume T_1 de cada sujeito (Fig. 6.1c).
- 145 máscaras automáticas do cérebro obtidas com o FSL sobre o volume DTI de cada sujeito (Fig. 6.1d).

Este conjunto de dados foi adquirido na Faculdade de Ciências Médicas (FCM) da Universidade Estadual de Campinas (Unicamp). Todos os sujeitos foram informados com antecedência e assinaram termos de consentimento livre e informado, aprovado pelo Comitê de Ética de Pesquisa da FCM-UNICAMP (nº CEP 920/2007; CAAE: 0669.0.146.000-07). Os experimentos foram realizados no Adessowiki, uma plataforma de programação colaborativa baseada em Python [68], usando Numpy [83], um pacote básico para gerenciamento matricial de dados e SciPy [53], uma biblioteca para tratamento científico da informação. SVM e K-means foram implementados usando Scikit-Learn [86], um módulo em Python que integra um amplo conjunto de algoritmos de aprendizado de

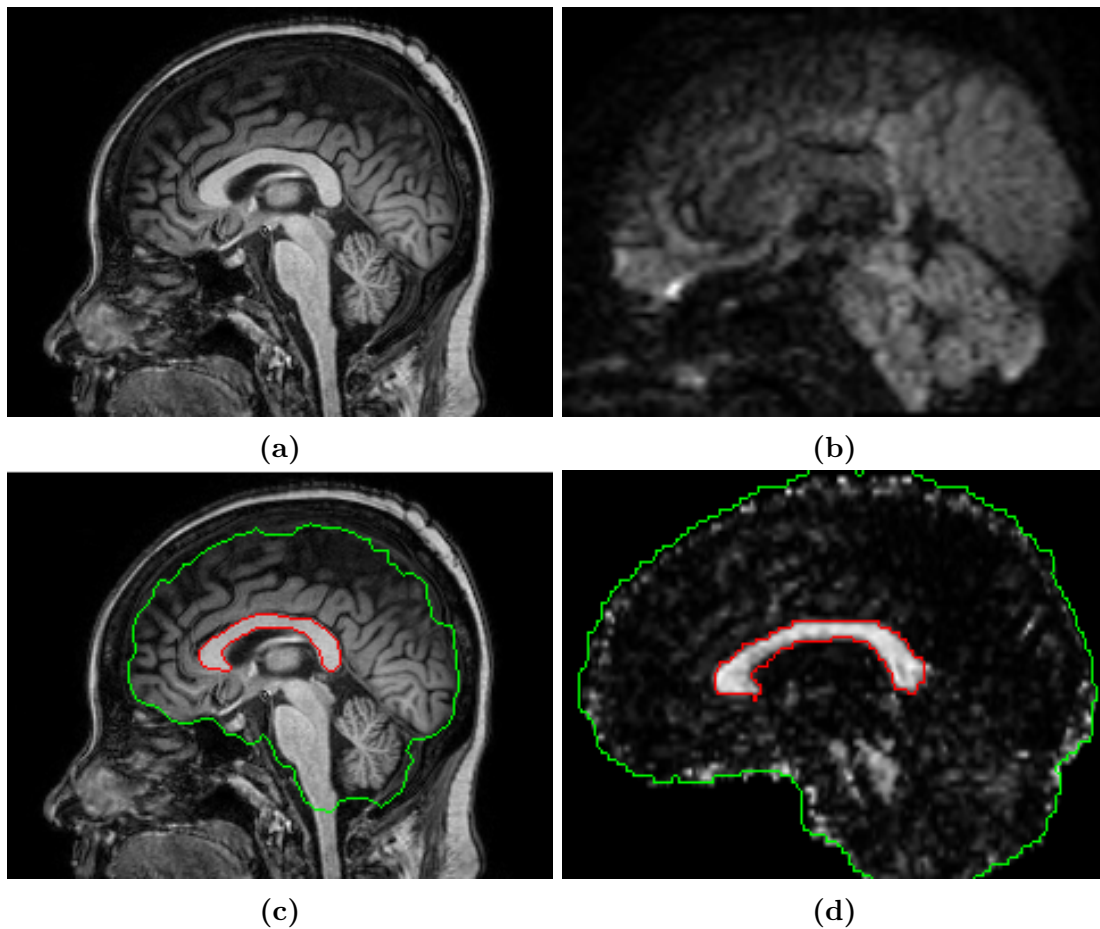


Figura 6.1 – Exemplo de dados usados: a) Fatia central do volume T_1 , b) Fatia central do volume DWI para um gradiente, c) Contorno máscara manual do CC (vermelho) e máscara automática do cérebro (verde) feitas sobre T_1 d) Contornos da máscara semi-automática do CC feita sobre mapa ponderado de FA (vermelho) e máscara automática do cérebro na fatia central feita sobre o volume DTI (verde)

máquina. STAPLE foi implementado em SimpleITK [70], uma plataforma para análise de imagem. Adicionalmente, foram usadas algumas funções e bibliotecas de propósito específico, implementadas no Adessowiki por nosso grupo de pesquisa.

6.2 Segmentações em difusão

Os três métodos de segmentação apresentados nos Caps. 3 e 4 foram implementados para serem posteriormente avaliados usando o método proposto. Os detalhes de implementação de cada segmentação estão descritos a seguir.

6.2.1 Segmentação baseada em *Watershed*

Seguindo o método de Freitas et al. [30] descrito na Sec. 2.5.1, a partir do mapa de FA ponderado (Fig. 6.2a) calculamos o gradiente morfológico externo (Fig. 6.2b) usando um disco de raio 1 como elemento estruturante. Para posicionar os marcadores, foi feita

a filtragem dos 50 mínimos com os maiores valores de extinção (Fig. 6.2c), com o qual a imagem é particionada em 50 regiões (Fig. 6.2d). Em seguida foi aplicado um limiar com valor de 0,2 (em todas as imagens) para filtrar e agrupar as regiões do CC (Fig. 6.2e). Finalmente, escolheu-se o maior componente conexo (Fig. 6.2f).

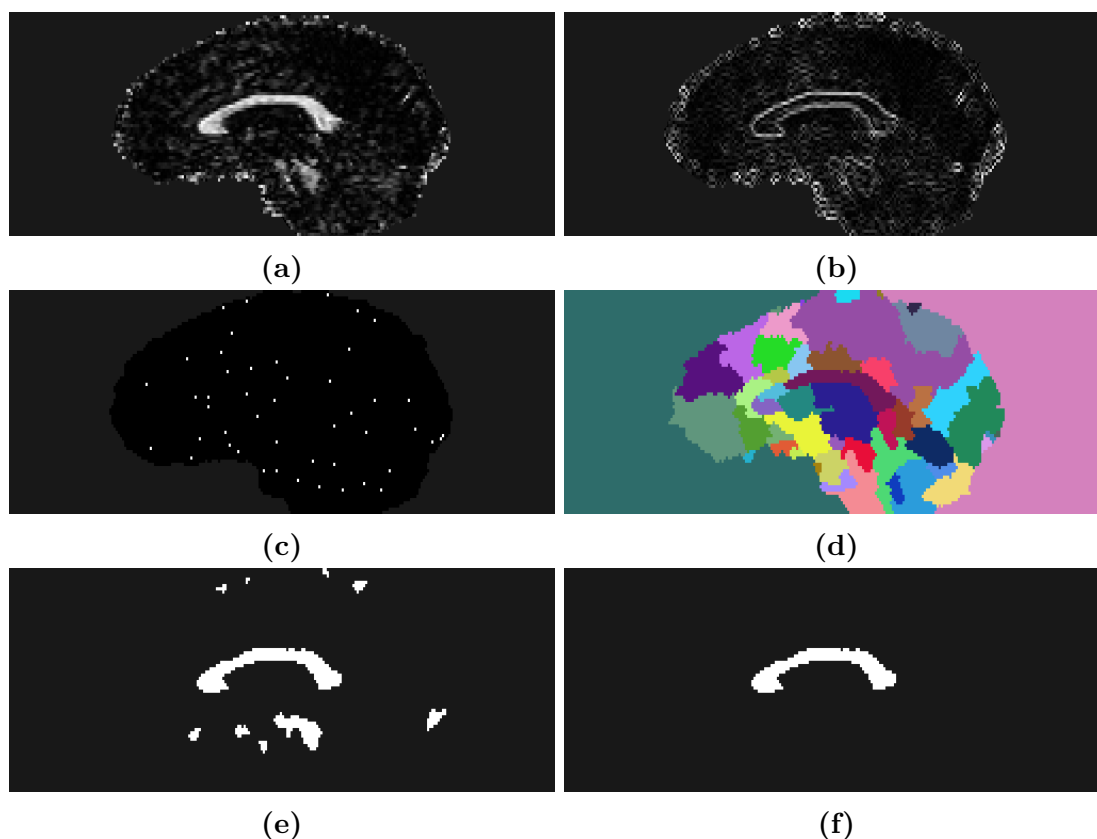


Figura 6.2 – Processo de segmentação baseada em *Watershed* para um sujeito: a) Mapa ponderado de FA, b) Gradiente morfológico externo, c) Mínimos locais filtrados por valores de extinção, d) Resultado do *Watershed*, e) Filtro das regiões com FA ponderado médio acima do limiar (0,2), f) Escolha do maior componente conexo (Segmentação final)

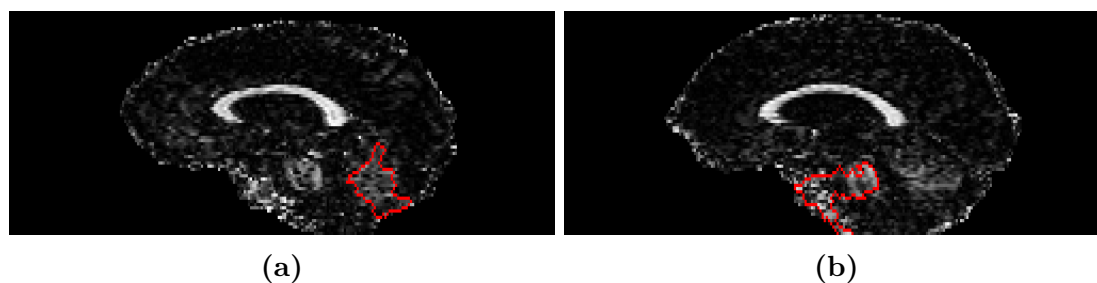


Figura 6.3 – Exemplos de erros apresentados na segmentação baseada em *Watershed* devido à escolha de um componente conexo maior do que o CC

Os parâmetros anteriores (limiar e número de marcadores) mostraram-se robustos dentro da análise feita por Freitas et al. [30]. Ainda assim, a segmentação apresenta alguns erros (Fig.6.3) devidos à escolha errada do maior componente conexo. Isto se deve ao fato que o algoritmo pressupõe que, ao aplicar o limiar por FA, as regiões que persistem e

que não são CC, não são conexas entre elas e têm uma área menor à do CC. Isso nem sempre é verdade. Nestes casos, o erro pode ser corrigido aumentando-se o limiar para filtrar os componentes remanescentes que não são CC, porém, este ajuste não pode ser generalizado.

6.2.2 Segmentação por Esquema de Quantificação de Objetivo reproduzível (ROQS)

Seguindo a implementação de Niogi et al. [82], uma semente manual foi posicionada dentro do CC. Para tal, usou-se a máscara semi-automática (Sec 6.1), a qual foi erodida por um elemento estruturante de raio 1, para descartar pixels da borda, e depois foi escolhido um pixel dentro da máscara erodida. Este pixel passou a ser a semente inicial (Fig. 6.4b). A partir dessa semente foram aplicados os dois critérios de seleção vistos na Seção 2.5.2 sobre a imagem DTI: pixels com a mesma orientação da semente (Fig. 6.4c) e pixels cuja anisotropia fosse similar à da semente (Fig. 6.4e). Para aplicação do segundo critério é usada a MM (Eq. 2.16) cujas constantes foram usadas como definidas empiricamente por Niogi et al. [82]: $\alpha = 0,3$, $\beta = 0,3$ e $\gamma = 0,5$ (Fig. 6.4d). Já o limiar para aplicação do segundo critério foi calculado achando a diferença entre o valor do máximo componente do principal autovetor encontrado no CC e o valor médio do máximo componente do principal autovetor encontrado na borda externa do CC ao longo de 10 sujeitos, dando um valor de 0,4. Finalmente, o CC foi segmentado escolhendo o maior componente depois de aplicado o segundo critério de seleção (Fig. 6.4f).

O algoritmo mostrou-se robusto à escolha das constantes da MM. Entretanto, ocorreram dois tipos de erros de segmentação. Em alguns casos a semente ficou posicionada sobre a borda do CC, (sendo afetada pelo problema do volume parcial). Devido a isso, a semente ficou com características da região próxima ao CC e portanto o algoritmo segmentou outra região (Fig. 6.5a). O segundo tipo de erro foi devido a que, em alguns casos, o limiar passou poucos pixels na zona estreita do tronco do CC e, ao escolher o maior componente conexo, o CC foi dividido (Fig. 6.5b).

6.2.3 Segmentação usando Classificação por pixel

Esta segmentação foi implementada seguindo a descrição do método no Cap. 4. Os dados de treinamento para a classificação supervisionada são tomados da fatia central de 30 sujeitos, onde temos 32 DWI (resposta ao gradiente de campo magnético em 32 direções) e uma imagem b0 (resposta sem gradiente de campo magnético aplicado) (Fig. 6.6a). Sobre a imagem b0 é calculado o gradiente morfológico (Fig. 6.6b). Para extrair e rotular os pixels (amostras) destas imagens foram passadas duas máscaras (Fig. 6.6c): a máscara do cérebro foi usada para extrair pixels de interesse dentro da imagem (todos os pixels do cérebro na fatia central), enquanto que a máscara semi-automática do CC foi

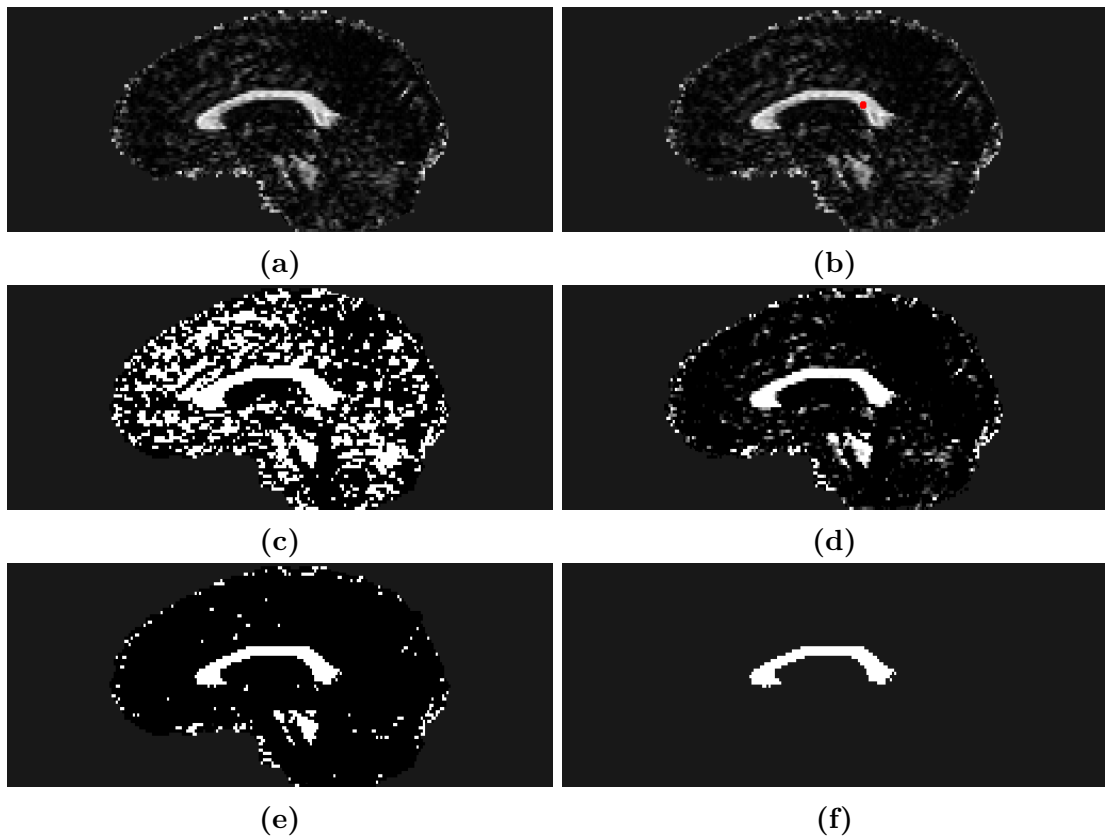


Figura 6.4 – Processo de segmentação do ROQS: a) Mapa ponderado de FA, b) Posicionamento semente (em vermelho) usando a máscara semi-automática, c) Aplicação primeiro critério (pixels com a mesma orientação da semente), d) Imagem da MM, e) Aplicação segundo critério (pixels com anisotropia similar à da semente), f) Escolha do maior componente conexo (Segmentação final)

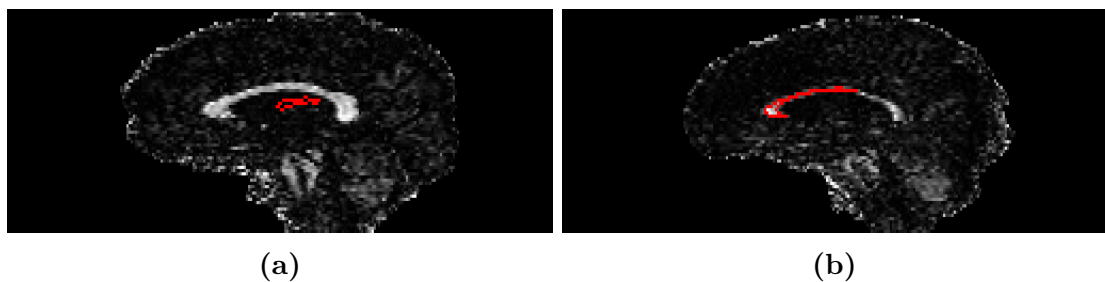


Figura 6.5 – Exemplos de erros na segmentação ROQS: a) erro devido ao posicionamento errado da semente e b) o CC é dividido em duas partes

usada para rotular as amostras de classe pCC (pixels dentro da máscara) e classe nCC (pixels fora da máscara). Para evitar o problema de volume parcial que acontece em DWI, pixels das bordas interna e externa do CC foram descartados erodindo as máscaras, tanto do CC quanto do cérebro, por um elemento estruturante com forma de cruz de tamanho 1 (Fig. 6.6d). Desta forma, uma matriz de treinamento rotulada (229010 amostras x 33 atributos) foi obtida com a proporção de amostras nCC/pCC de {37 : 1}.

Para balancear classes e poupar recursos de treinamento, 10% de amostras representativas de classe nCC foram selecionadas usando k-means da seguinte forma:

1. 7 clusters de classe negativa foram formados usando k-means, tentando representar a diversidade de tecidos que possui a classe nCC.
2. Começando pelo menor cluster, $x = N \div 70$ amostras foram aleatoriamente selecionadas de cada cluster, com N sendo o número total de amostras nCC.
3. Caso o cluster tivesse menos de x amostras, todas as amostras desse cluster seriam selecionadas e o valor faltante seria dividido entre os clusters restantes.
4. O processo foi repetido até sub-amostrar todos os clusters.

Depois da rotina de sub-amostragem, a relação de amostras nCC/pCC diminuiu para $\{4 : 1\}$, aproximadamente. Assim, a matriz final de treinamento tinha o tamanho (28288×33) . Esta matriz foi normalizada por atributos e passada ao classificador. Os hiper-parâmetros da SVM foram selecionados usando *Grid-search* onde precisão foi usada como métrica de desempenho buscando minimizar a ocorrência de falsos positivos. Assim a SVM foi treinada com a configuração de hiper-parâmetros dada pelo *Grid-search* (Tab. 6.1).

Tabela 6.1 – Hiper-parâmetros testados (coluna central) e escolhidos (coluna da direita) usando *Grid-search* para a SVM

Hiper-parâmetros	Candidatos	Valor selecionado
K	{linear, RBF}	RBF
C	{1, 10, 100}	10
γ (Caso RBF)	{0, 0.1, 0.01}	0.1

onde K (o kernel da SVM) é uma função de base radial gaussiana, C é o parâmetro de punição que controla a relação entre complexidade da máquina e o número permitido de pontos não separáveis, e γ (coeficiente do kernel RBF) é o parâmetro que controla até que ponto chega a influencia de uma amostra de treinamento. O treinamento é feito usando uma estratégia de *k-fold* com $k = 5$ (*5 folds*), com o qual, para cada época de treinamento dedicam-se 20% das amostras para teste (5658).

Uma vez treinado o modelo, a classificação é feita em novos sujeitos e a segmentação resultante é reconstruída a partir da saída do classificador usando a posição previamente salva de cada amostra (pixel) na imagem (Fig. 6.6e). Finalmente, o processo de refinamento é feito para chegar na segmentação final usando o pós-processamento morfológico com um elemento estruturante de tamanho 1 e forma de cruz (Fig. 6.6f).

O método conseguiu bons resultados nos sujeitos de teste. Porém, em alguns casos o algoritmo apresentou falsos, tanto positivos quanto negativos (Fig. 6.7a). Isto devido a que as amostras de treinamento não generalizaram o suficiente para todos os casos de teste. Em outros casos, similar ao que acontece nos erros do ROQS, a escolha do

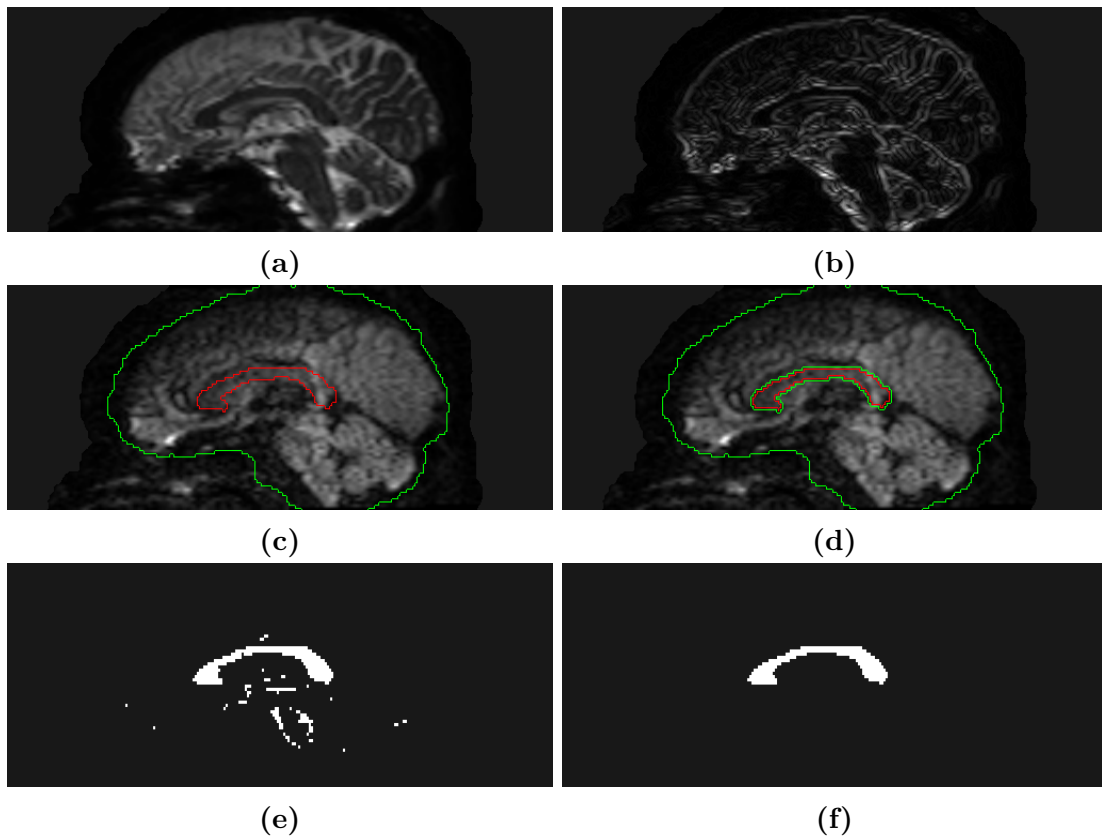


Figura 6.6 – Processo de segmentação baseada em classificação por pixel para um sujeito: a) Imagem b0 de DWI, b) Gradiente morfológico sobre a imagem b0, c) Contorno da máscara do cérebro (verde) e do CC (vermelho), d) Contornos das máscaras para extração de amostras tirando pixels da borda do CC: máscara verde (nCC) e máscara vermelha (pCC), e) Saída do classificador para um sujeito de teste, f) Aplicação do pós-processamento: maior componente conexo, fechamento e abertura (Segmentação final)

maior componente conexo divide a estrutura segmentada devido à zona estreita do CC (Fig. 6.7b).

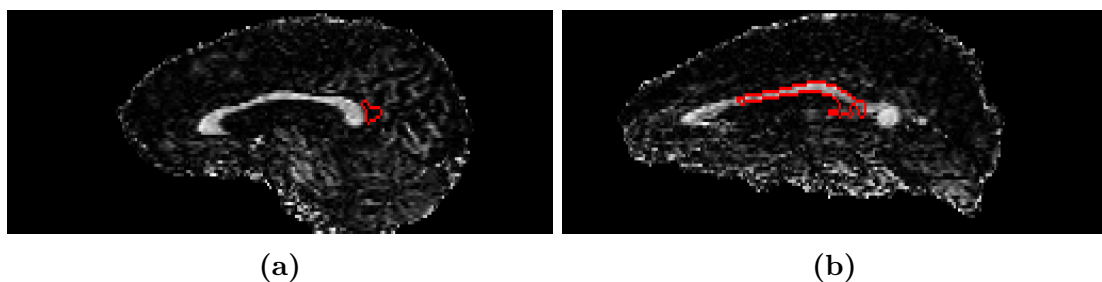


Figura 6.7 – Exemplos de erros da segmentação baseada em classificação por pixel: a) erro de generalização e b) erro na escolha do maior componente conexo

6.3 Padrões

Para fazer o estudo da avaliação de segmentações em difusão usaremos principalmente o padrão-ouro em T_1 correspondente ao método de avaliação proposto.

Porém, na Sec. 6.5.2 iremos usar os padrões-prata para caracterizar a métrica RMSE e portanto estes padrões serão apresentados a seguir. Além, provaremos que é possível usar o padrão-ouro para avaliar os padrões-prata.

Os padrões que serão descritos a seguir são o padrão-ouro (padrão manual em T_1) e os padrões-prata: padrão manual em T_1 , registrado em difusão, padrão semi-automático e padrão STAPLE.

6.3.1 Padrão manual em T_1

A partir das imagens ponderadas em T_1 de cada sujeito foram traçadas segmentações manuais do CC por um especialista do Hospital das Clínicas da Unicamp, usando a ferramenta computacional Display [73]. Estas máscaras são o padrão-ouro (Fig. 6.8a) para cada sujeito.

6.3.2 Padrão manual registrado no espaço de difusão

Um primeiro padrão-prata é obtido a partir do padrão-ouro feito manualmente em T_1 e registrado em difusão. O registro foi realizado com o FSL [49], ferramenta computacional para análise de MRI, usando a seguinte configuração:

- Transformação linear: O tipo de transformação linear vem dado pelos graus de liberdade (0-12) atribuídos à transformação. Neste caso foram usados 9 graus de liberdade que incluem operações de rotação, traslação e escalamento.
- Função de custo: A escolha da função de custo é importante pois determina o grau de similaridade que pode atingir a referência com respeito à estrutura deformada. Foi escolhida informação mútua como função de custo por ser adequada para registro multi-modalidade.
- Interpolação: Determina que valores tomam os dados que não conseguem ser mapeados diretamente pela transformação. Neste caso, a interpolação escolhida foi trilinear.

A partir da máscara manual em T_1 (Fig. 6.8a), e utilizando a configuração descrita anteriormente, foi obtido o padrão manual registrado em difusão (Fig. 6.8b).

6.3.3 Padrão semi-automático

Este padrão-prata é uma segmentação obtida com Neuroline [98], uma ferramenta de segmentação semi-automática. O usuário posiciona manualmente sementes internas e externas ao CC, sobre o mapa ponderado de FA, e a ferramenta faz a segmentação usando o algoritmo de *Watershed* (Fig. 6.9).

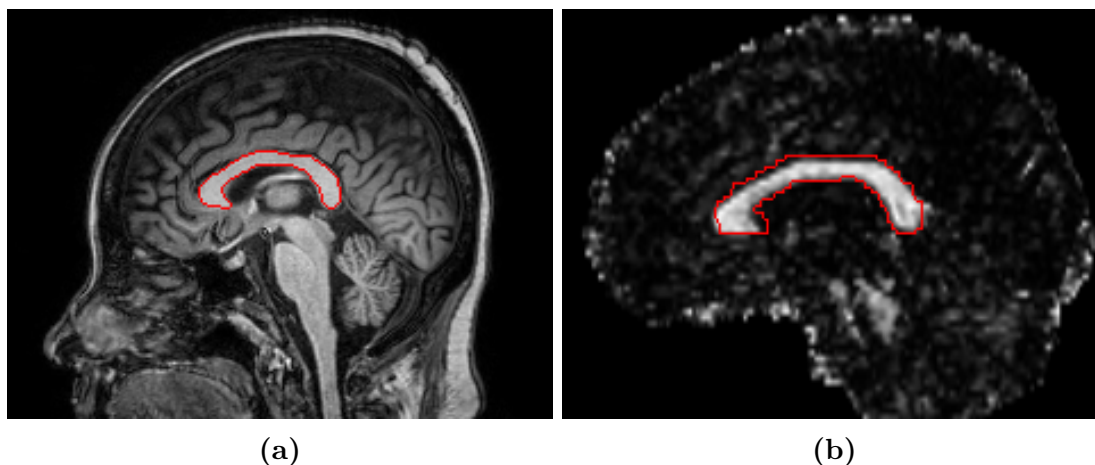


Figura 6.8 – Padrão manual registrado no espaço de difusão: a) Máscara original em T_1 b) Máscara registrada em difusão

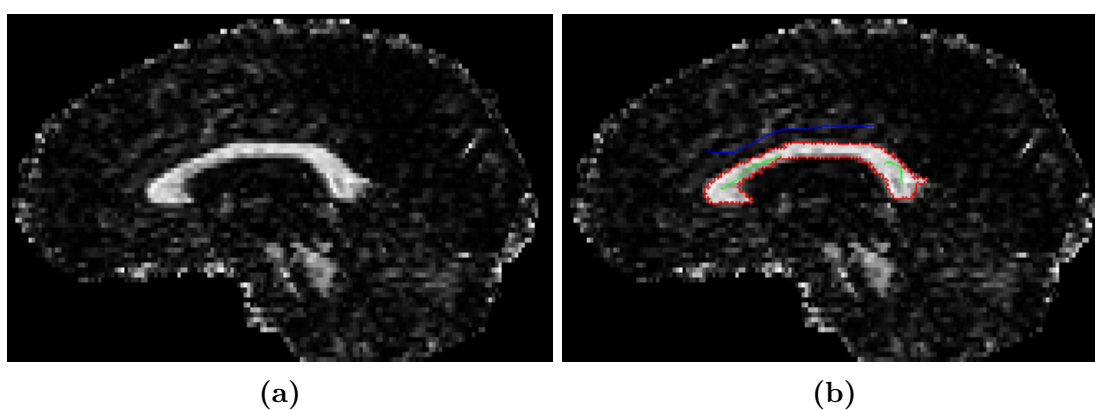


Figura 6.9 – Padrão no espaço de difusão com auxílio da ferramenta semi-automática Neuroline: a) Mapa ponderado de FA b) Posicionamento das sementes sobre o mapa ponderado de FA: semente interna (verde), semente externa (azul) e padrão final semi-automático (vermelho)

6.3.4 Padrão STAPLE

O padrão-prata, neste caso, foi construído usando o STAPLE a partir dos 3 métodos de segmentação já expostos na Sec. 6.2.

Usando como entrada as três máscaras binárias resultantes de cada um dos métodos de segmentação: baseado em *Watershed* (Fig. 6.10a), ROQS (Fig. 6.10b) e classificação por pixel (Fig. 6.10c), o STAPLE gera um mapa de probabilidade da segmentação real (Fig. 6.10d). Este mapa de probabilidade estabelece para cada pixel, valores entre 0 (probabilidade nula de pertencer ao CC) e 1 (100% de probabilidade de pertencer ao CC). Para obter o padrão-prata final, deve ser aplicado um limiar que permita binarizar o mapa de probabilidade. De modo geral, o STAPLE gera um mapa robusto na zona interna e externa do CC, porém, a borda tem valores de probabilidade intermediários e portanto esta zona é sensível à escolha do limiar.

No exemplo, são aplicados dois limiares: > 0.2 (Fig. 6.10e) e > 0.7 (Fig. 6.10f).

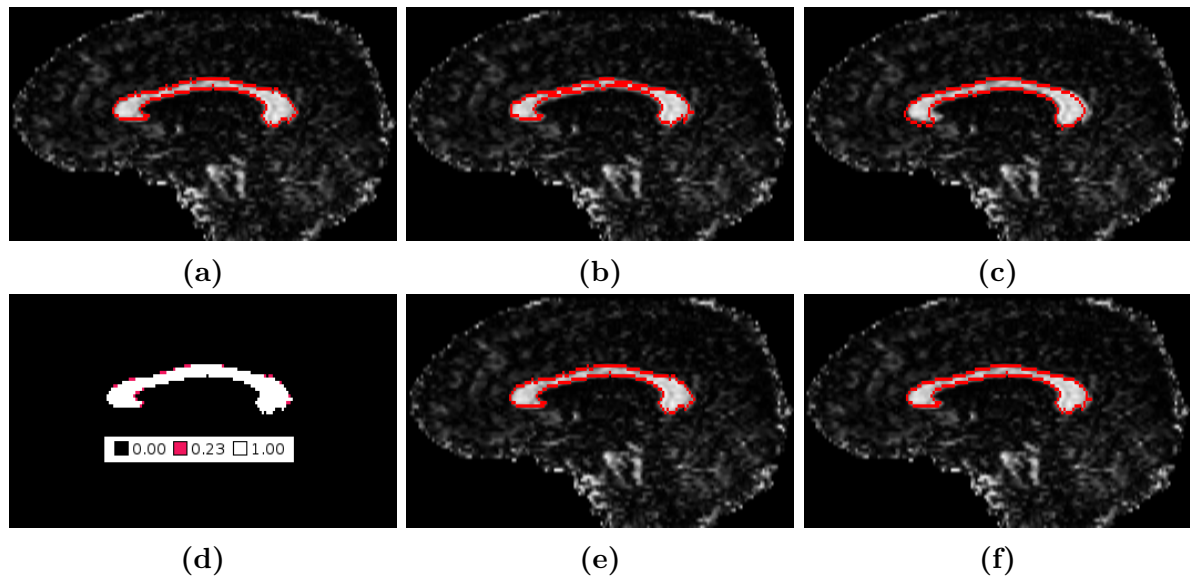


Figura 6.10 – Construção padrão STAPLE: a) Segmentação baseada em *Watershed*, b) Segmentação por ROQS, c) Segmentação usando classificação baseada em pixel, d) Mapa de probabilidade do padrão STAPLE com cores representando os diferentes valores, e) Segmentação STAPLE com limiar aplicado de 0.1, f) Segmentação STAPLE com limiar aplicado de 0.9

Neste caso, os pixels de probabilidade intermediária (0.23) são poucos e portanto a escolha do limiar não interfere muito com o padrão final. Recomenda-se aplicar limiares superiores a 0.7 já que desta forma são inclusos os pixels com alta probabilidade de pertencer ao CC.

6.3.5 Avaliação dos padrões-prata em difusão

Em primeiro lugar, fizemos uma avaliação qualitativa sobrepondo os três padrões-prata para um sujeito. Um mapa de cores foi usado para destacar as zonas de similaridade e diferença entre os padrões (Fig. 6.11).

Visualmente, a segmentação registrada (Fig. 6.11a) aparece deslocada em algumas regiões do CC, especialmente nos extremos (joelho, rostro e esplênio). Este deslocamento é devido ao processo de registro que acrescenta erros à estrutura final. O padrão semi-automático (Fig. 6.11b) coincide melhor com a estrutura da imagem, ainda que não rotula (falsos negativos) zonas do CC como o rostro e o esplênio. Este padrão-prata é questionável devido às características de baixa resolução e contraste do mapa ponderado de FA sobre o qual foi construído, ao efeito de volume parcial e à dependência do posicionamento manual das sementes. Finalmente, o padrão STAPLE (Fig. 6.11c) tem uma coincidência melhor com a estrutura do CC.

Para fazer a avaliação quantitativa dos padrões-prata em difusão usamos a segmentação manual em T_1 , que é considerada o padrão-ouro. Uma vez que o padrão-ouro e os padrões-prata encontram-se em espaços diferentes, a avaliação foi feita usando o RMSE

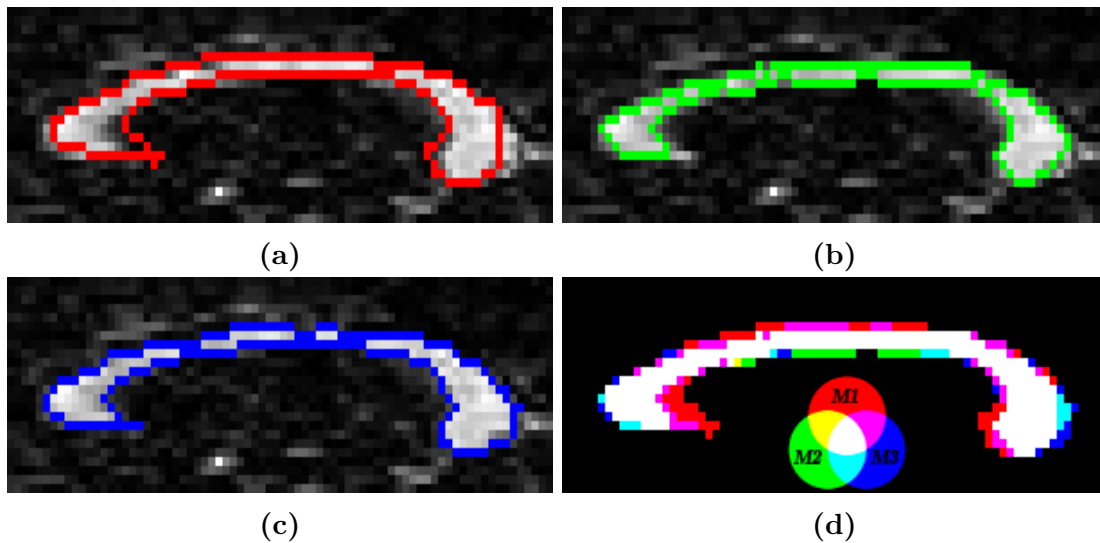


Figura 6.11 – Avaliação qualitativa dos padrões-prata em difusão (a imagem de fundo é o mapa ponderado de FA aumentado na zona do CC): a) Padrão manual registrado em difusão, b) Padrão semi-automático, c) Padrão STAPLE, d) Sobreposição dos padrões

do método de perfil de curvatura (Eq. 5.7), que como visto, não requer sobreposição das estruturas. Esta avaliação foi feita sobre os 145 sujeitos (Fig. 6.12)

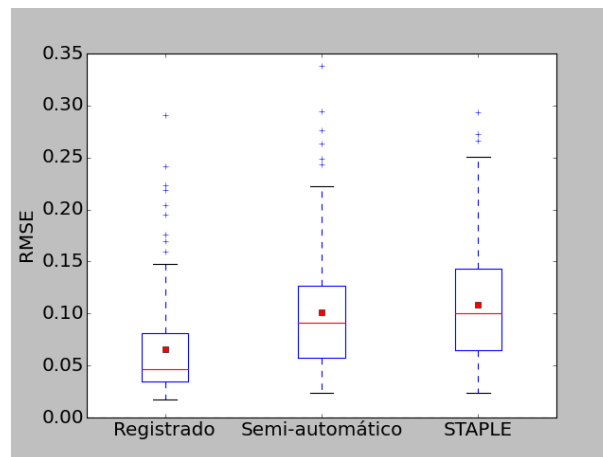


Figura 6.12 – Avaliação quantitativa dos padrões-prata em difusão: registrado, semi-automático e STAPLE, usando como padrão-ouro a segmentação manual em T_1 e como métrica RMSE do perfil de curvatura

A avaliação mostrou a vantagem do padrão registrado sobre os outros dois já que este padrão é uma forma transformada do padrão-ouro através do processo de registro. O padrão semi-automático segue em similaridade ao padrão registrado e finalmente, o STAPLE é o padrão menos similar ao padrão-ouro.

Isto não indica que o padrão registrado é o melhor padrão-prata, mas sim que tem a forma mais próxima ao padrão-ouro. O STAPLE foi indicado qualitativamente como o melhor padrão-prata, porém teve a pior avaliação quantitativa. Isto indica que, no possível, para avaliar segmentações devem ser usadas métricas globais ou métricas em

conjunto.

Como complemento, no Apêndice B foi usado o método usual de avaliação por sobreposição para comparar as segmentações em difusão usando os padrões-prata.

6.4 Avaliação de segmentações do CC usando o método proposto

Nesta seção serão apresentados os experimentos correspondentes ao método de avaliação proposto baseado no perfil de curvatura (Fig. 6.13).

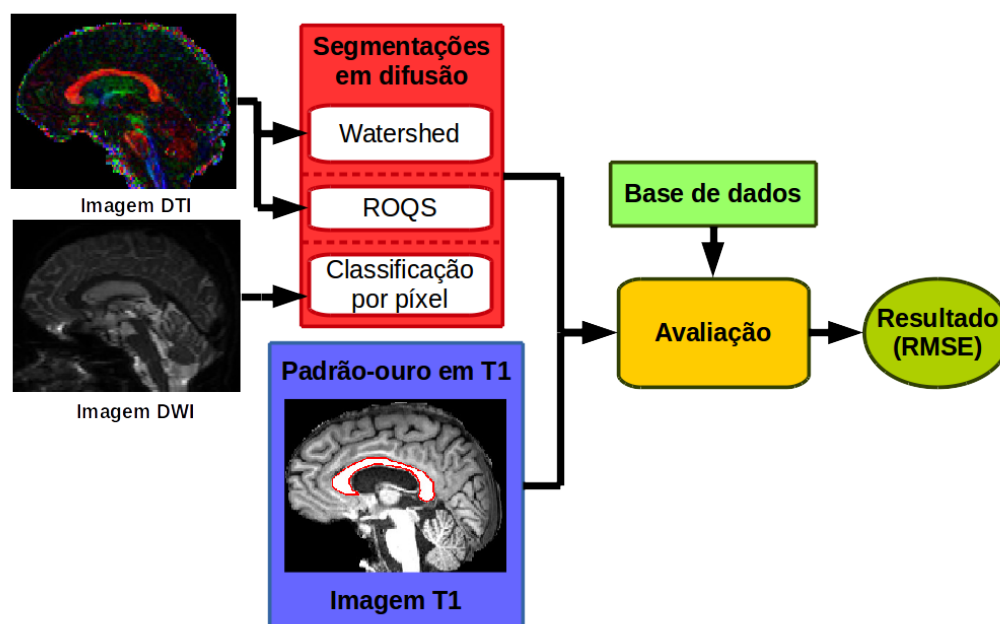


Figura 6.13 – Esquema proposto de avaliação de segmentações do CC em difusão

As três segmentações em difusão, foram diretamente avaliadas com o padrão-ouro em T_1 , usando como métrica o RMSE entre perfis de curvatura. Os resultados são apresentados usando diagramas de caixa (Fig. 6.14).

Como resultado da avaliação das segmentações temos que o ROQS é o melhor método de segmentação, além de ter menor variação em torno da média. O método baseado em *Watershed*, por outro lado, tem o maior número de segmentações erradas (20 sujeitos) e sua variação é maior. Entende-se segmentação errada como aquela que seleciona estruturas diferentes ao CC, falhando completamente.

Ainda que o método de classificação por pixel apresenta o menor número de segmentações erradas, este método tende segmentar irregularmente as bordas devido ao efeito de volume parcial (efeito mais notável em DWI). Isto prejudica a avaliação com RMSE sendo que esta é uma métrica de forma.

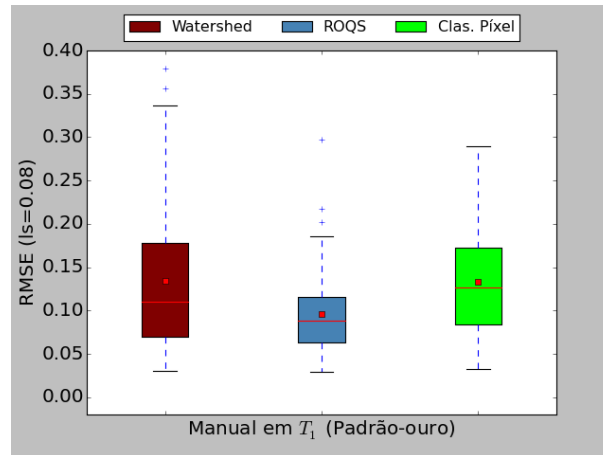


Figura 6.14 – Avaliação das segmentações em difusão para 145 sujeitos usando o padrão-ouro e RMSE ($l_s = 0.08$)

O uso do padrão-ouro é desejável, porém, é mais difícil de obter, uma vez que é uma segmentação manual. Além, imagem em T_1 , sobre a qual é traçada o padrão manual, nem sempre é disponibilizada em conjunto com a imagem de difusão.

Os padrões-prata podem ser considerados nestes casos, porém, devem ser levadas em conta as desvantagens de cada um: o padrão registrado precisa da existência do padrão manual e acrescenta erros de registro, o padrão semi-automático não tem boas características, como visto na avaliação qualitativa e quantitativa com o padrão-ouro e o padrão STAPLE tem a inconveniência de precisar de segmentações para ser construído (mínimo 3 para obter um padrão razoável).

A escolha do método de avaliação (usual ou proposto) depende tanto da disponibilidade de padrões quanto do objetivo da avaliação. O método de avaliação de segmentações proposto usando a métrica RMSE leva em conta a similaridade das formas comparadas. Avaliar outros aspectos da segmentação requer uso de outras métricas e/ou registro o que incrementa o custo na avaliação.

6.5 Caracterização do perfil de curvatura e do RMSE

O RMSE foi usado para avaliar segmentações em difusão (Sec. 6.4) e adicionalmente foi usado para avaliar os padrões-prata (Sec. 6.3.5). Nesta seção, será feita a caracterização do perfil de curvatura e do RMSE, primeiro do ponto de vista do descritor (Sec. 6.5.1) e depois do ponto de vista da métrica (Sec. 6.5.2).

6.5.1 Caracterização do perfil de curvatura como descritor

Um descritor, como o perfil de curvatura, deve garantir a representação da forma de um objeto do modo mais preciso e confiável. Assim, o nosso descritor deve garantir os

seguintes critérios [79]:

Invariância: Segmentações que têm a mesma forma devem ter a mesma representação.

Unicidade: Segmentações que não têm a mesma forma devem ter diferente representação.

Estabilidade: Segmentações com pequenas diferenças de forma devem ter representações próximas; segmentações que têm representações próximas devem ter pequenas diferenças.

O critério de invariância garante que curvas com a mesma forma têm a mesma representação ainda sob aplicação de transformações geométricas uniformes de rotação, traslação e escala. O critério de unicidade garante que curvas diferentes têm representações diferentes. Finalmente, o critério de estabilidade garante que uma pequena mudança na estrutura não causa grandes alterações no descritor.

O cumprimento dos anteriores critérios é importante porquanto permite a discriminação de segmentações, ao tempo que confere robustez ao descritor. No caso particular do nosso estudo, estes critérios apresentam duas vantagens: 1) permitem estabelecer uma métrica, baseada no perfil de curvatura, robusta e invariante e 2) possibilitam a comparação de segmentações em espaços diferentes.

Com o propósito de avaliar os critérios anteriormente mencionados, o padrão manual do CC em T_1 de um sujeito foi alterado de forma gradual e crescente. Foi calculado o RMSE (com $ls = 0.08$) entre o padrão original (sem sofrer alteração) e cada uma das alterações. Esta análise foi feita, de forma independente, para três tipos de alteração: escala, rotação e ruído.

O perfil, por definição, é invariante à traslação porquanto a medida do ângulo de curvatura é feita localmente na borda da estrutura, e em consequência seu valor é independentemente da posição do CC na imagem.

6.5.1.1 Escala (Ampliação)

O caso mais simples de alteração é a escala, que neste caso, é feita fazendo replicação uniforme dos valores de cada pixel, obtendo uma ampliação da imagem. Neste caso, ampliar a imagem por uma escala n significa replicar sequencialmente as colunas $n - 1$ vezes e depois $n - 1$ vezes as linhas; a Fig. 6.15 exemplifica o caso para $n = 2$.

Na Fig. 6.16a, mostra-se o padrão manual em T_1 do CC de um sujeito e na Fig. 6.16b o perfil associado. Como exemplo, o padrão do CC é escalado por um fator de $n = 10$, e o perfil resultante pode ser visto na Fig. 6.16c. Este experimento foi feito para

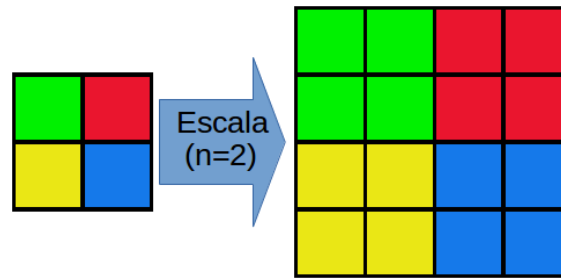


Figura 6.15 – Escala de uma imagem por um fator de 2 ($n = 2$)

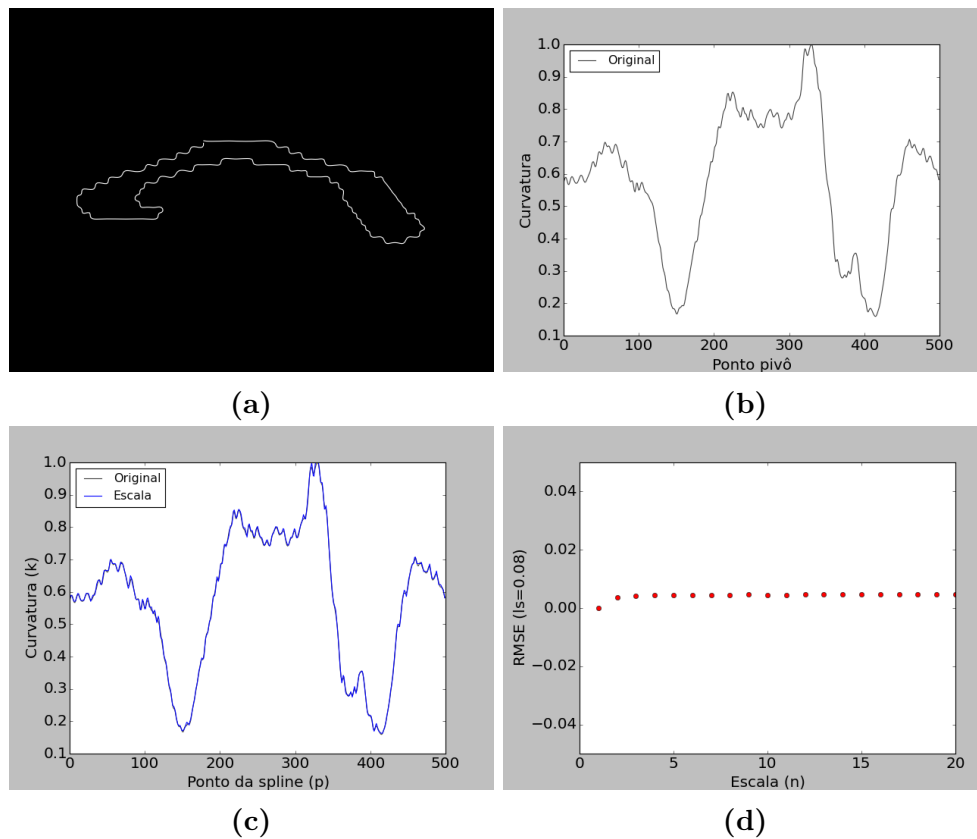


Figura 6.16 – RMSE sob alterações graduais de escala uniforme: a) CC para um sujeito (estrutura original), b) Perfil de curvatura original ($ls = 0.08$), c) Perfil de curvatura associado à estrutura com escala $n = 10$, d) RMSE medida ao longo das alterações de escala usando como padrão a estrutura original

o mesmo padrão com n variando entre 1 e 20. Para cada n o RMSE, entre a estrutura original e a escalada, foi medido (Fig. 6.16d).

Em termos práticos, podemos dizer que o perfil é invariante à escala da estrutura porquanto a RMSE está próxima de zero para qualquer n ; a diferença que registra o RMSE a partir de $n = 2$ é devida à implementação da *Spline* que muda um pouco o perfil da estrutura escalada; porém esta mudança é visualmente imperceptível como pode-se notar na Fig. 6.16c.

6.5.1.2 Rotação

Teoricamente, o descritor também é invariante à rotação já que a curvatura é medida localmente. Neste caso, a alteração é feita rotacionando a estrutura original (máscara manual em T_1) por um ângulo positivo que varia entre 0 e 45° com passos de um grau. A Fig. 6.17 apresenta o exemplo do perfil associado à rotação da estrutura original em 22°. Na Fig. 6.17c é registrado o RMSE entre a forma original e as rotações feitas ao longo das rotações entre 0° e 45°, com passo de 1°.

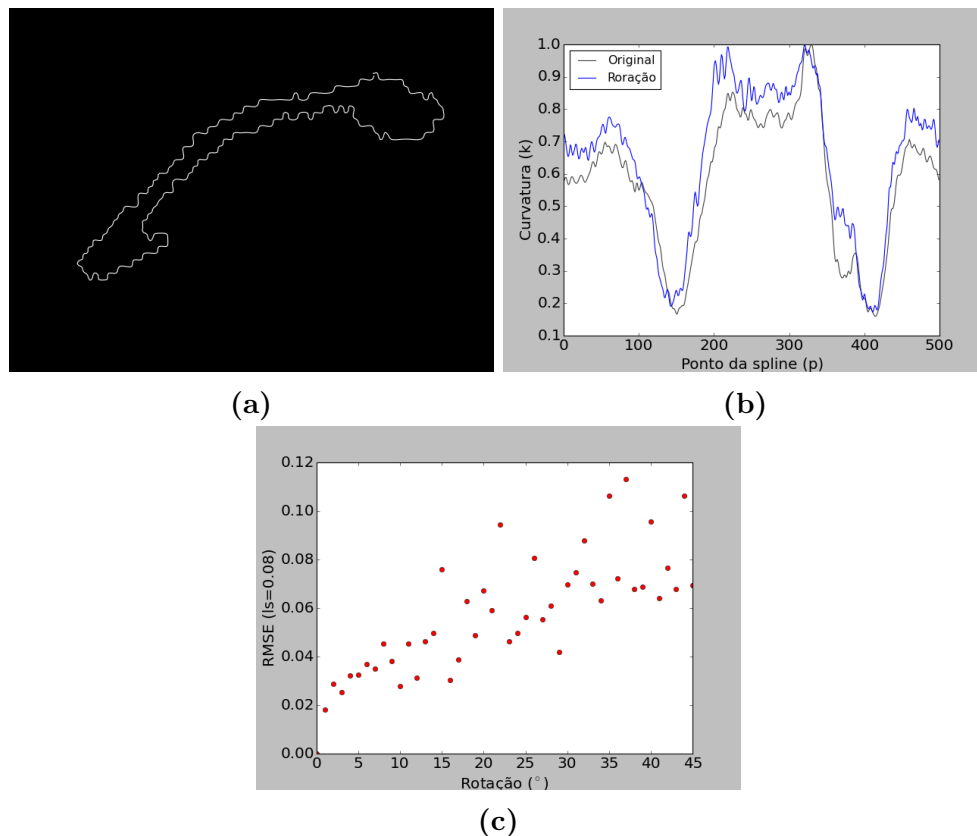


Figura 6.17 – RMSE sob alterações graduais de rotação: a) Exemplo da estrutura rotacionada 22°, b) Perfil de curvatura da segmentação rotacionada 22° c) RMSE medida ao longo das alterações usando como padrão a estrutura original

O RMSE ao longo das rotações tem um comportamento crescente ainda que variável. Esta variabilidade é devida à interpolação que acontece ao longo do processo de rotação. A interpolação é feita pelo vizinho mais próximo dado que a estrutura original é uma máscara binária. A interpolação modifica a estrutura, especialmente na borda, e em consequência o perfil sofre alterações; porém, estas alterações são consequência da implementação e não da invariância do perfil.

6.5.1.3 Ruído (Sal e pimenta)

Neste caso, acrescenta-se gradualmente ruído sal e pimenta na borda da estrutura binária de T_1 , com o qual a borda vai sendo deformada. O ruído aplica-se sobre toda

a imagem usando uma função aleatória de distribuição discreta uniforme tanto para sal (que vira pixels a Verdadeiro) quanto para pimenta (vira pixels a Falso). Depois de gerado o ruído sobre toda a imagem, e para garantir que só seja afetada a borda da estrutura, aplicam-se dois passos: 1) Função lógica OU pixel a pixel entre a imagem com ruído e uma versão erodida da máscara do CC original de forma que é preservada a parte interna do CC e 2) É escolhido o maior componente conexo de forma que permanecem todos os pixels conexos com a estrutura base do CC.

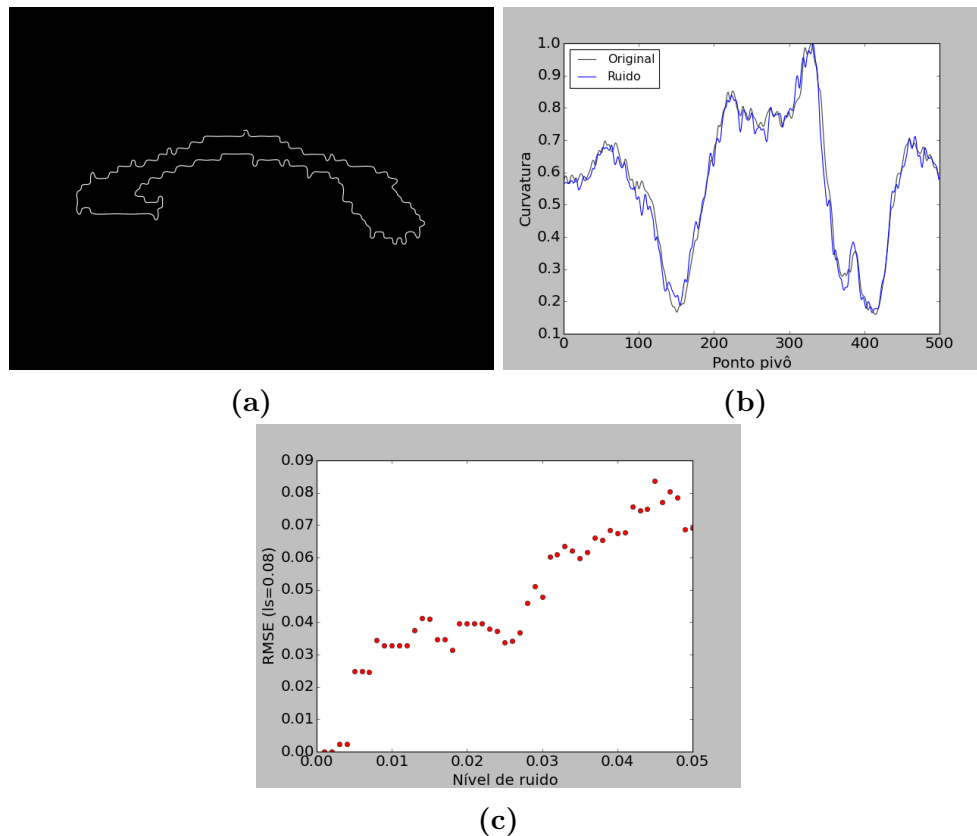


Figura 6.18 – RMSE sob alterações graduais da borda da estrutura na forma de ruído sal e pimenta: a) Estrutura alterada com nível 0.025 de ruído, b) Perfil de curvatura com nível 0.025 de ruído, c) RMSE medida ao longo das alterações usando como padrão a estrutura original

A Fig. 6.18a, apresenta um exemplo da estrutura alterada, com seu respectivo perfil (Fig. 6.18b), para nível de ruído de 0.025. Na Fig. 6.18c é apresentada a evolução do RMSE entre a estrutura original e a estrutura alterada pela adição crescente de ruído sal e pimenta para níveis entre 0.001 e 0.05. Pode-se notar que o RMSE tem um comportamento crescente. O perfil de curvatura é afetado pelas mudanças que a borda vai tendo ao longo da adição de ruído.

Como conclusão desta seção, a escala é a alteração que menos afeta a estrutura original e em termos práticos seu impacto é nulo. A rotação apresenta um comportamento crescente e desordenado devido à interpolação feita em diferentes pontos da borda da estrutura. Finalmente, o ruído vai crescendo de forma mais uniforme até superar os valores da rotação.

Os critérios de invariância, unicidade e estabilidade são atingidos pelo RMSE dentro dos limites técnicos impostos pela resolução da imagem. A invariância melhora quando usado um perfil de curvatura com uma resolução menor já que neste caso o detalhe na borda influencia menos o valor da métrica RMSE final.

As alterações graduais mostraram a estabilidade e unicidade da métrica já que alterações graduais levaram a aumentos graduais do RMSE. A invariância foi demonstrada a través das transformações geométricas de rotação e escala. Deve-se notar que a grandeza das diferenças do RMSE ao longo das alterações ($\approx 1 * 10^{-2}$) é baixa quando comparada com a faixa de medição em que opera normalmente o RMSE para avaliar segmentações, maior a ($\approx 1 * 10^{-1}$), como veremos a continuação.

6.5.2 Caracterização do perfil de curvatura e RMSE como métrica

A caracterização do perfil de curvatura e do RMSE foi feita através da comparação com uma outra métrica já estabelecida. A métrica escolhida para comparação é o coeficiente Dice que é uma métrica robusta e versátil amplamente usada na literatura.

O objetivo é conhecer melhor o RMSE e estudar a correlação entre esta métrica e o coeficiente Dice. Para isto, realizamos um experimento onde, para todas as segmentações, calculamos as duas medidas. Devido a que o Dice não pode ser usado no método de avaliação proposto (usando o padrão-manual), o experimento foi feito completamente no espaço de difusão, com os padrões-prata.

Assim, a avaliação foi feita para as diferentes combinações possíveis entre padrão-prata e segmentação, gerando uma matriz (Fig. 6.19) onde as linhas contém os diferentes padrões-prata (registrado, semi-automático e STAPLE) e as colunas as diferentes segmentações (*Watershed*, ROQS e classificação por pixel). O RMSE foi normalizado ($RMSE_{norm}$) entre 0 e 1 para facilitar sua comparação com o Dice (Dice não precisa ser normalizado já que sua faixa de operação é entre 0 e 1). Cada sub-gráfico apresenta o valor normalizado das métricas para uma combinação de padrão *versus* segmentação ao longo dos 145 sujeitos. Finalmente os sujeitos são ordenados de forma crescente em função do Dice para facilitar a interpretação dos resultados. É importante lembrar que as duas métricas são inversas: o Dice tende a crescer quanto melhor a coincidência das estruturas comparadas enquanto o RMSE tende a diminuir.

A partir destes experimentos, pudemos observar:

- Nos gráficos da coluna da segmentação *Watershed*, o Dice para os primeiros 20 sujeitos é igual a zero, já que o método segmentou regiões do cérebro que não coincidem nem em forma nem em sobreposição com o CC. Estes casos apresentam $RMSE_{norm}$ maior do que 0.5.

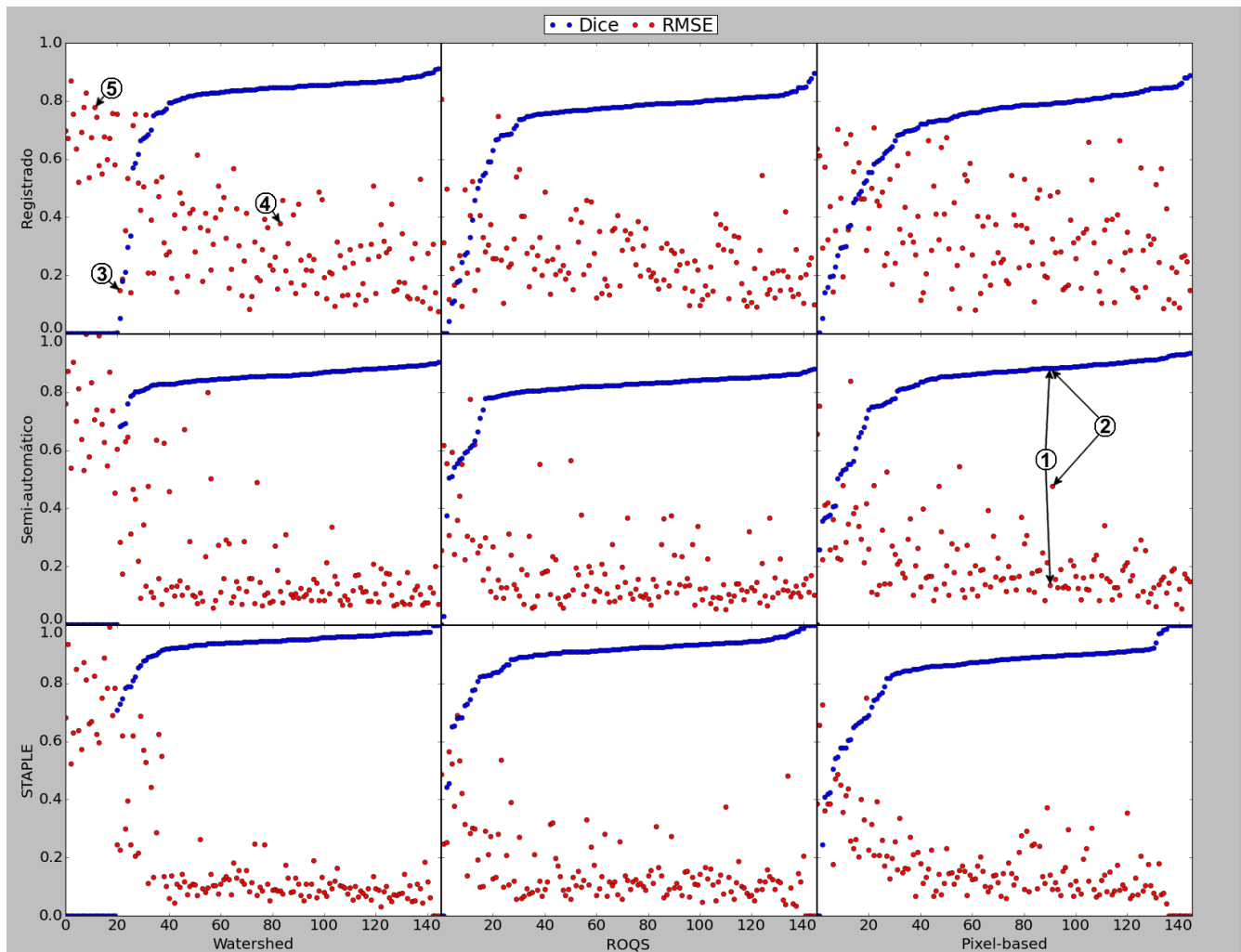


Figura 6.19 – Comparação Dice e RMSE normalizado ($RMSE_{norm}$) ao longo de 145 sujeitos em difusão para as diferentes combinações (de esquerda para direita e de cima para baixo): Registrado vs *Watershed*, Registrado vs ROQS, Registrado vs Classificação por pixel, Semi-automático vs *Watershed*, Semi-automático vs ROQS, Semi-automático vs Classificação por pixel, STAPLE vs *Watershed*, STAPLE vs ROQS, STAPLE vs Classificação por pixel

- Zonas onde o Dice é alto ($Dice > 0.8$), o $RMSE_{norm}$ está abaixo de 0.5. Esta zona, nos gráficos, normalmente corresponde a segmentações que coincidem relativamente bem com o respectivo padrão. Assim, para a linha do padrão STAPLE, as três segmentações têm menor variação do que quando comparadas com o padrão-registrado (primeira linha), o qual indica uma coincidência melhor de forma das estruturas comparadas para o STAPLE.
- A distribuição dos pontos do $RMSE_{norm}$ difere da distribuição uniforme do Dice para cada gráfico. Isto indica que as duas métricas avaliam aspectos diferentes: segmentações com Dice similares podem ter $RMSE_{norm}$ diferentes, indicando boa sobreposição mas diferença da forma nas estruturas, como pode ser visto na Fig. 6.20. Estes dois casos podem ser vistos na Fig. 6.19, no diagrama de dispersão padrão semi-automático *versus* segmentação de classificação por pixel (casos ① e

②).

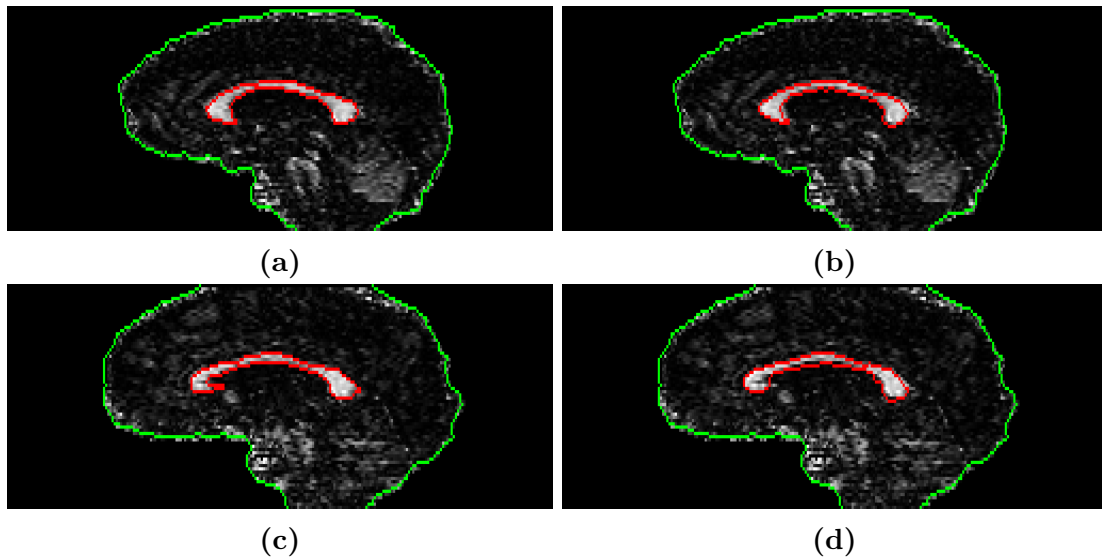


Figura 6.20 – Exemplo de comparação para dois sujeitos onde o Dice é igual e o $RMSE_{norm}$ varia consideravelmente (Os dois casos podem ser mapeados na Fig. 6.19): a) Caso ①, padrão semi-automático b) Caso ①, segmentação baseada em pixel ($Dice = 0.88$ e $RMSE_{norm} = 0.05$), c) Caso ②, padrão semi-automático d) Caso ②, segmentação baseada em pixel ($Dice = 0.88$ e $RMSE_{norm} = 0.19$)

Retomando a Fig. 6.19, podemos observar os pares padrão-segmentação a que cada ponto faz referência, com o propósito de definir as faixas numéricas de operação da métrica. Esta informação é útil, porquanto permite associar o valor do RMSE às características relativas do par padrão-segmentação que está sendo avaliado. Foram identificadas 3 faixas de operação para a avaliação de pares padrão-segmentação que, ainda que não são absolutas, pretendem é auxiliar no uso e na interpretação da métrica:

- $0 < RMSE_{norm} < 0.2$ está associado a pares similares com nenhuma ou pouca diferença de forma.
- $0.2 < RMSE_{norm} < 0.5$ está associado a pares com pequenas diferenças nas bordas (como bicos ou entradas).
- $0.5 < RMSE_{norm} < 1$ está associado a casos mais críticos, que apresentam erros na segmentação ou no padrão.

Os valores do RMSE usados ao longo deste experimento foram valores normalizados entre 0 e 1 (multiplicando por 2.5) e portanto para voltar esses valores absolutos basta dividir o valor normalizado por 2.5.

A Fig. 6.21 apresenta um exemplo em cada faixa (casos ③, ④ e ⑤). Estes casos foram extraídos aleatoriamente do sub-gráfico padrão registrado *versus* segmentação *Watershed* da Fig. 6.19.

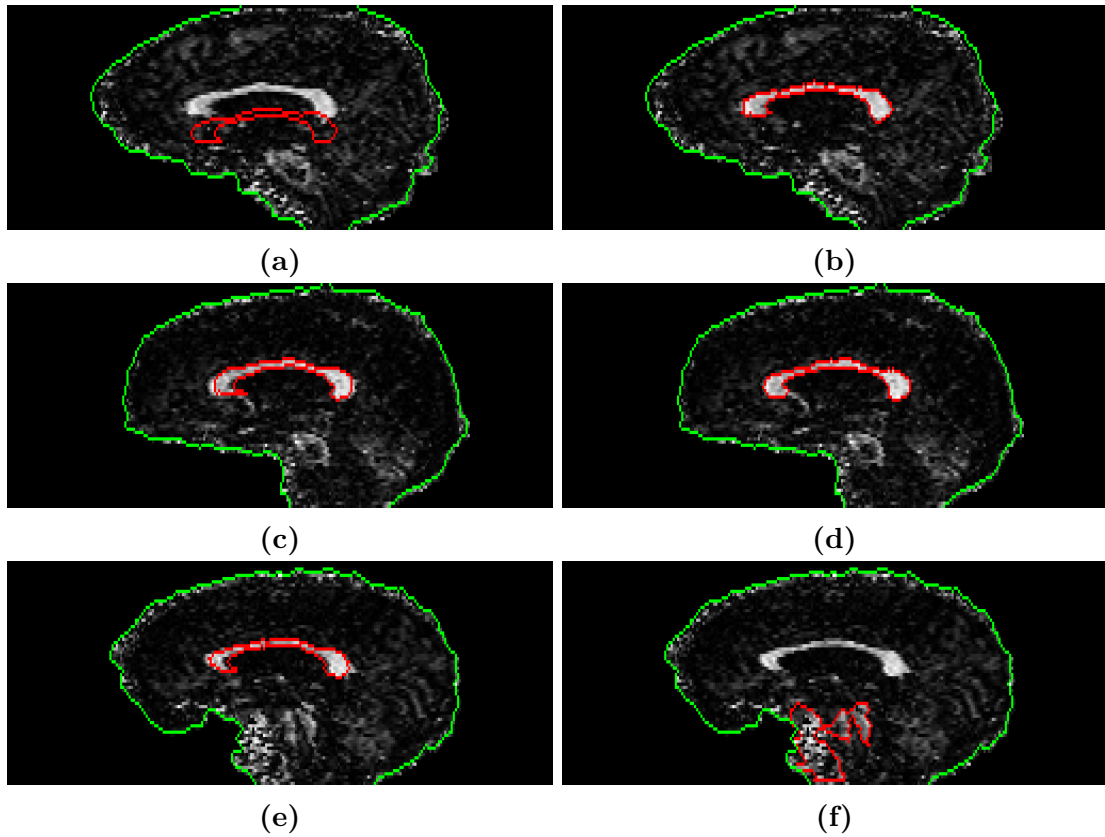


Figura 6.21 – Exemplo de faixa de operação da $RMSE_{norm}$ para o padrão registrado e segmentação Watershed (Os três casos podem ser mapeados na Fig. 6.19): a) Caso ③, padrão registrado, b) Caso ③, segmentação Watershed ($RMSE_{norm} = 0.125$), c) Caso ④, padrão registrado, d) Caso ④, segmentação Watershed ($RMSE_{norm} = 0.375$), e) Caso ⑤, padrão registrado, f) Caso ⑤, segmentação Watershed ($RMSE_{norm} = 0.775$)

Para o caso do primeiro sujeito (Figs. 6.21a e 6.21b), a forma coincide bem resultando em RMSE baixo; além disso a métrica não penaliza o fato do padrão estar mal registrado. Para o segundo caso (Figs. 6.21c e 6.21d), as formas apresentam pequenas diferenças no joelho, no rosto e em menor medida no esplênio. Já para o terceiro caso (Figs. 6.21e e 6.21f), o método baseado em Watershed segmentou outra estrutura com uma forma bem diferente e portanto o RMSE foi alto.

7 Conclusões

Neste trabalho propusemos um método de avaliação de segmentações do CC em imagens de difusão por ressonância magnética. O método de avaliação proposto permite utilizar o padrão-ouro construído manualmente em imagens ponderadas em T_1 , uma vez que não exige que o padrão e a segmentação estejam no mesmo espaço (mesma modalidade de imagem). A dissimilaridade entre segmentação e padrão é dada pelo RMSE entre perfis de curvatura extraídos de cada uma das imagens comparadas.

O perfil de curvatura proposto funciona como uma assinatura de forma da estrutura segmentada e pode ser traçado em diferentes resoluções. A escolha da resolução depende da aplicação final. Neste trabalho, a assinatura foi usada para fazer a comparação de segmentações do mesmo sujeito.

Os parâmetros do perfil de curvatura foram definidos para avaliar segmentações do mesmo sujeito, fazendo ênfase na importância da resolução no valor final do RMSE. Outras aplicações podem requerer outros valores nos parâmetros.

O RMSE entre perfis de curvatura, utilizado como métrica de dissimilaridade entre segmentações, apesar de ser uma métrica mais complexa e custosa do que as métricas tradicionais, oferece três vantagens: inclui informação de forma das estruturas envolvidas, permite o uso do padrão-ouro para avaliação sem uso de registro e cumpre com os critérios de invariância (rotação, translação e escala), unicidade e estabilidade. Além, o RMSE foi caracterizado, identificando faixas de operação da métrica, e mostrando a sua complementariedade com o coeficiente Dice.

O método proposto foi usado para avaliar as segmentações em difusão. Foram implementados dois métodos de segmentação de CC em difusão (estado-da-arte): método baseado em *Watershed* e o método ROQS. Além disso, foi proposto um terceiro método de segmentação do CC em difusão baseada em classificação por pixel. Também foram construídos três padrões em difusão (padrões-prata) os quais foram avaliados com o padrão manual em T_1 (padrão-ouro).

Os experimentos de avaliação usando o método usual por sobreposição permitiram identificar ao ROQS como a melhor segmentação quando avaliado diretamente com a máscara T_1 usando RMSE. Este resultado indica que a forma da segmentação ROQS coincide melhor com a forma estrutural do CC. Porém, o resultado da avaliação de um algoritmo depende diretamente do padrão e da métrica usados.

É difícil comparar segmentações usando diferentes métricas e padrões uma vez que a escolha da métrica e o padrão depende do objetivo da avaliação. Nos piores cenários, só

se tem disponibilidade de um padrão. A escolha da métrica é um ponto importante que determina que características serão avaliadas na segmentação; dado que não existe uma métrica absoluta que garanta uma avaliação absoluta, em alguns casos, pode ser útil usar um conjunto de métricas complementares que permitam uma avaliação mais integral.

Neste trabalho, optou-se por usar a segmentação manual em T_1 como melhor padrão disponível e avaliar a forma do CC segmentado por diferentes métodos em difusão usando o RMSE baseada no perfil de curvatura. A maior contribuição do nosso método foi permitir a avaliação direta com o padrão-ouro sem usar registro de imagens. Ainda que sejam conhecidas as desvantagens do padrão-ouro, este é ainda amplamente usado na literatura e na área clínica e, portanto, a proposta do nosso método de avaliação permite melhorar este cenário.

7.1 Trabalhos futuros

A proposta do perfil de curvatura mostrou-se útil para comparação e avaliação de segmentações do CC realizadas no espaço de difusão, principalmente quando se dispõe apenas de um padrão-ouro construído em outro espaço. Apesar dos experimentos e discussões se restringirem à aplicação escolhida, percebemos que o perfil de curvatura proposto, juntamente com o RMSE tem potencial para ser usado em diversos outros tipos de análise, dentre elas:

- Caracterização de forma do CC
- Identificação automática de segmentações incorretas em grandes bases de dados
- Estudos populacionais e longitudinais
- Caracterização de forma de outras estruturas cerebrais ou de objetos planares

Publicações em anais de eventos e em revistas

- Artigo da revisão sistemática da literatura de métodos computacionais focados em segmentação e parcelamento do CC em MRI (Cap. 1).
 - **Computational methods for Corpus Callosum Segmentation on MRI: a Systematic Literature Review.** Computer Methods and Programs in Biomedicine, 2016. Artigo submetido [18].
- Publicações sobre um método de segmentação do CC em DWI baseado em classificação por pixel (Cap. 4).

-
- **Pixel-Based Classification Method for Corpus Callosum Segmentation on Diffusion-MRI.** VipImage, 2017. Resumo extendido aceito para apresentação [43].
 - **Corpus callosum segmentation on DWI through voxel-based classification.** Journal of epilepsy and clinical neurophysiology, 2016. Resumo publicado [42].
 - Resumo sobre o uso do perfil de curvatura para comparação qualitativa de três segmentações com um padrão-ouro (Cap. 5).
 - **A method for extraction of Corpus Callosum’s Shape Signature.** Journal of epilepsy and clinical neurophysiology, 2017. Em publicação [41].
 - Publicação feita em parceria com uma colega sobre um método de segmentação em DTI usando uma rede GNG (Growing Neural Gas) para segmentação do CC.
 - **Corpus Callosum 2D Segmentation on Diffusion Tensor Imaging using Growing Neural Gas Network.** VipImage, 2017. Aceito para apresentação [19].

Referências

- [1] F. Aboitiz, A. B. Scheibel, R. S. Fisher, and E. Zaidel. Fiber composition of the human corpus callosum. *Brain Research*, 598(1–2):143 – 153, 1992.
- [2] T. Adamek and N. E. O’Connor. A multiscale representation method for nonrigid shapes with a single closed contour. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(5):742–753, 2004.
- [3] C. Adamson, R. Beare, M. Walterfang, and M. Seal. Software Pipeline for Midsagittal Corpus Callosum Thickness Profile Processing: Automated Segmentation, Manual Editor, Thickness Profile Generator, Group-Wise Statistical Comparison and Results Display. *Neuroinformatics*, 12(4):595–614, 2014.
- [4] B. A. Ardekani, T. Ikuta, A. Bachman, and P. R. Szeszko. Multi-Atlas Corpus Callosum Segmentation with Adaptive Atlas Selection. *Ismrm*, 20:2564, 2012.
- [5] R. Bammer. Basic principles of diffusion-weighted imaging. *European journal of radiology*, 45(3):169–184, 2003.
- [6] P. J. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994.
- [7] S. Beucher and F. Meyer. The morphological approach to segmentation: the watershed transformation. *Optical Engineering*, 34(12):433–481, 1992.
- [8] F. Bloch. Nuclear induction. *Physical review*, 70(7-8):460, 1946.
- [9] N. Bloembergen, E. M. Purcell, and R. V. Pound. Relaxation effects in nuclear magnetic resonance absorption. *Physical review*, 73(7):679, 1948.
- [10] M. Brejl and M. Sonka. Object localization and border detection criteria design in edge-based image segmentation: Automated learning from examples. *IEEE Transactions on Medical Imaging*, 19(10):973–985, 2000.
- [11] P. Brigger, J. Hoeg, and M. Unser. B-spline snakes: a flexible tool for parametric contour detection. *IEEE Transactions on Image Processing*, 9(9):1484–1496, 2000.
- [12] L. G. Brown. A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4):325–376, 1992.
- [13] G. Bueno, O. Musse, F. Heitz, and J. P. Armspach. Three-dimensional segmentation of anatomical structures in MR images on large data bases. *Magnetic Resonance Imaging*, 19(1):73–88, 2001.

-
- [14] C. J. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.
- [15] J. S. Cardoso and L. Corte-Real. Toward a generic evaluation of image segmentation. *IEEE Transactions on Image Processing*, 14(11):1773–1782, 2005.
- [16] M. Chen, T. Kanade, H. A. Rowley, and D. Pomerleau. Quantitative study of brain anatomy. *Workshop on Biomedical Image Analysis, Proceedings.*, pages 84–92, 1998.
- [17] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.
- [18] G. Cover, W. Herrera, M. Bento, S. Appenzeler, and L. Rittner. Computational methods for Corpus Callosum Segmentation on MRI: a Systematic Literature Review. *Journal of Computer Methods and Programs in Biomedicine, (Em revisão)*, 2016.
- [19] G. Cover, W. Herrera, M. Bento, and L. Rittner. Corpus Callosum 2D Segmentation on Diffusion Tensor Imaging using Growing Neural Gas Network. *Lecture Notes in Computational Vision and Biomechanics, (Aceito)*, 2017.
- [20] R. Damadian. Tumor detection by nuclear magnetic resonance. *Science*, 171(3976):1151–1153, 1971.
- [21] R. Damadian, L. Minkoff, M. Goldsmith, M. Stanford, and J. Koutcher. Field focusing nuclear magnetic resonance (fonar): visualization of a tumor in a live animal. *Science*, 194(4272):1430–1432, 1976.
- [22] C. Davatzikos, X. Tao, and D. Shen. Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, 22(3):414–423, 2003.
- [23] C. De Boor, C. De Boor, E.-U. Mathématicien, C. De Boor, and C. De Boor. *A practical guide to splines*, volume 27. Springer-Verlag New York, 1978.
- [24] I. de Kouchkovsky, E. Fieremans, L. Fleysher, J. Herbert, R. I. Grossman, and M. Inglese. Quantification of normal-appearing white matter tract integrity in multiple sclerosis: a diffusion kurtosis imaging study. *Journal of neurology*, pages 1–10, 2016.
- [25] R. R. Edelman and S. Warach. Magnetic resonance imaging. *New England Journal of Medicine*, 328(10):708–716, 1993.
- [26] A. El-Baz, A. Elnakib, M. F. Casanova, G. Gimel’Farb, A. E. Switala, D. Jordan, and S. Rainey. Accurate automated detection of autism related corpus callosum abnormalities. *Journal of Medical Systems*, 35(5):929–939, 2011.

- [27] A. Elnakib, M. F. Casanova, G. Gimel'Farb, A. E. Switala, and A. El-Baz. Autism diagnostics by centerline-based shape analysis of the Corpus Callosum. *Proceedings - International Symposium on Biomedical Imaging*, pages 1843–1846, 2011.
- [28] J. D. Enderle and J. D. Bronzino. *Introduction to biomedical engineering*. Academic press, 2012.
- [29] J. L. Foo. A survey of user interaction and automation in medical image segmentation methods. *Iowa State University, Human Computer Interaction Technical Report ISU-HCI*, (2), 2006.
- [30] P. Freitas, L. Rittner, S. Appenzeller, and R. Lotufo. Watershed-based segmentation of the midsagittal section of the corpus callosum in diffusion MRI. *Conference on Graphics, Patterns and Images (SIBGRAPI)*, (24):274–280, 2011.
- [31] P. F. Freitas. Segmentação e parcelamento do corpo caloso em imagens de tensor de difusão. *Universidade Estadual de Campinas, UNICAMP, Brasil*, Mestrado em Engenharia Elétrica, 2012.
- [32] V. Garcia, H. De Jesus, and B. Mederos. Analysis of discrepancy metrics used in medical image segmentation. *IEEE Latin America Transactions*, 13(1):235–240, 2015.
- [33] T. Gass, G. Székely, and O. Goksel. Simultaneous segmentation and multiresolution nonrigid atlas registration. *IEEE Transactions on Image Processing*, 23(7):2931–2943, 2014.
- [34] A. R. Gogliettino, M. N. Potenza, and S. W. Yip. White matter development and tobacco smoking in young adults: A systematic review with recommendations for future research. *Drug and alcohol dependence*, 162:26–33, 2016.
- [35] B. T. Gold, N. F. Johnson, D. K. Powell, and C. D. Smith. White matter integrity and vulnerability to Alzheimer's disease: Preliminary findings and future directions. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1822(3):416–422, 2012.
- [36] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [37] N. Gordillo, E. Montseny, and P. Sobrevilla. State of the art survey on MRI brain tumor segmentation. *Magnetic Resonance Imaging*, 31(8):1426–1438, 2013.
- [38] I. S. Gousias, A. D. Edwards, M. A. Rutherford, S. J. Counsell, J. V. Hajnal, D. Rueckert, and A. Hammers. Magnetic resonance imaging of the newborn

- brain: manual segmentation of labelled atlases in term-born and preterm infants. *Neuroimage*, 62(3):1499–1509, 2012.
- [39] J. Han and M. Kamber. *Data Mining: Concepts and Techniques*. Elsevier, 2 edition, 2006.
- [40] Q. He, Y. Duan, J. Miles, and N. Takahashi. A context-sensitive active contour for 2D corpus callosum segmentation. *International Journal of Biomedical Imaging*, 2007, 2007.
- [41] W. Herrera, A. Costa, and L. Rittner. A method for extraction of Corpus Callosum’s Shape Signature. *Journal of Epilepsy and Clinical Neurophysiology (Em publicação)*, 2017.
- [42] W. Herrera, G. Cover, R. Lotufo, and L. Rittner. Corpus Callosum Segmentation on DWI through voxel-based Classification. *Journal of Epilepsy and Clinical Neurophysiology*, 22(3):111–112, 2016.
- [43] W. Herrera, G. Cover, and L. Rittner. Pixel-Based Classification Method for Corpus Callosum Segmentation on Diffusion-MRI. *Lecture Notes in Computational Vision and Biomechanics, (Em revisão)*, 2017.
- [44] L. T. Holly and K. T. Foley. Intraoperative spinal navigation. *Spine*, 28(15S):S54–S61, 2003.
- [45] T. A. G. M. Huisman. Diffusion-weighted imaging: basic concepts and application in cerebral stroke and head trauma. *European Radiology*, 13(10):2283–2297, 2003.
- [46] S. Içer. Automatic segmentation of corpus collasum using Gaussian mixture modeling and Fuzzy C means methods. *Computer Methods and Programs in Biomedicine*, 112(1):38–46, 2013.
- [47] A. R. Jac Fredo, G. Kavitha, and S. Ramakrishnan. Segmentation and analysis of brain subcortical regions using regularized multiphase level set in autistic MR images. *International Journal of Imaging Systems and Technology*, 24(3):256–262, 2014.
- [48] A. K. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1988.
- [49] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith. FSL. *NeuroImage*, 62(2):782 – 790, 2012.
- [50] F. L. Jianfeng Xu. Bayesian co-segmentation of multiple MR images. pages 53–56, 2009.

-
- [51] S. Johnson, T. Farnworth, J. Pinkston, and E. Bigler. Corpus callosum surface area across the human adult life span: effect of age and gender. *35(4):373–377*, 1994.
- [52] H. D. Jomma and A. I. Hussein. Circle views signature: A novel shape representation for shape recognition and retrieval. *Canadian Journal of Electrical and Computer Engineering*, 39(4):274–282, 2016.
- [53] E. Jones, T. Oliphant, P. Peterson, et al. Scipy: Open source scientific tools for python, 2001. URL <http://www.scipy.org>, 73:86, 2015.
- [54] K. Karsch, Q. He, and Y. Duan. A fast, semi-automatic brain structure segmentation algorithm for magnetic resonance imaging. *2009 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2009*, pages 297–302, 2009.
- [55] G. Katti, S. A. Ara, and A. Shireen. Magnetic resonance imaging (MRI)—A review. *International Journal of Dental Clinics*, 3(1), 2011.
- [56] T. Kauppi, J.-K. Kamarainen, L. Lensu, V. Kalesnykiene, I. Sorri, H. Kälviäinen, H. Uusitalo, and J. Pietilä. Fusion of multiple expert annotations and overall score selection for medical image diagnosis. In *Scandinavian Conference on Image Analysis*, pages 760–769, 2009.
- [57] Y. Kong, D. Wang, L. Shi, S. C. N. Hui, and W. C. W. Chu. Adaptive distance metric learning for diffusion tensor image segmentation. *PLoS ONE*, 9(3):1–11, 2014.
- [58] S. Kullmann, F. Schweizer, R. Veit, A. Fritsche, and H. Preissl. Compromised white matter integrity in obesity. *Obesity Reviews*, 16(4):273–281, 2015.
- [59] T. A. Lampert, A. Stumpf, and P. Gañarski. An empirical study into annotator agreement, ground truth estimation, and algorithm evaluation. *IEEE Transactions on Image Processing*, 25(6):2557–2572, 2016.
- [60] P. C. Lauterbur. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*, 242:190–191, 1973.
- [61] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407, 1986.
- [62] D. Le Bihan, J. F. Mangin, C. Poupon, C. a. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of magnetic resonance imaging : JMRI*, 13(4):534–546, 2001.

- [63] D. Le Bihan, C. T. Moonen, P. C. van Zijl, J. Pekar, and D. DesPres. Measuring random microscopic motion of water in tissues with MR imaging: a cat brain study. *Journal of computer assisted tomography*, 15(1):19–25, 1991.
- [64] M. E. Leventon, W. E. L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. 1:316–323, 2000.
- [65] Y. Li, M. Mandal, and S. N. Ahmed. Fully automated segmentation of corpus callosum in midsagittal brain MRIs. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 5111–5114, 2013.
- [66] W. Liu and D. Ruan. A unified variational segmentation framework with a level-set based sparse composite shape prior. *Physics in Medicine and Biology*, 60(5):1865–1877, 2015.
- [67] S. Loncaric. A survey of shape analysis techniques. *Pattern recognition*, 31(8):983–1001, 1998.
- [68] R. A. Lotufo, R. C. Machado, A. Körbes, and R. G. Ramos. Adessowiki on-line collaborative scientific programming platform. In *Proceedings of the 5th international symposium on Wikis and open collaboration*, page 10, 2009.
- [69] R. A. Lotufo, L. Rittner, R. Audigier, R. C. Machado, and A. V. Saúde. Morphological image processing applied in biomedicine. In *Biomedical Image Processing*, pages 107–129, 2011.
- [70] B. C. Lowekamp, D. T. Chen, L. Ibáñez, and D. Blezek. The design of simpleitk. *Frontiers in neuroinformatics*, 7:45, 2013.
- [71] R. Luis-García, C.-F. Westin, and C. Alberola-López. Gaussian mixtures on tensor fields for segmentation: Applications to medical imaging. *Computerized Medical Imaging and Graphics*, 35(1):16–30, 2011.
- [72] T. Lyche and K. Mørken. Spline methods draft. *Department of Informatics, Center of Mathematics for Applications, University of Oslo, Oslo*, 2008.
- [73] D. MacDonald. MNI-display: Program for display and segmentation of surfaces and volumes. *McConnell Brain Imaging Center, Montreal Neurological Institute, Montreal, Canada. Available at: <http://www.bic.mcgill.ca/ServicesSoftware/HomePage>*, 1996.
- [74] V. A. Magnotta, D. Heckel, N. C. Andreasen, T. Cizadlo, P. W. Corson, J. C. Ehrhardt, W. T. C. Yuh, P. Westmoreland, P. Westmoreland-Corson, C. Ehrhardt, James, and P. Westmorland Corson. Measurement of Brain Structures with

- Artificial Neural Networks: Two- and Three-dimensional Applications. *Radiology*, 211(3):781–790, 1999.
- [75] G. Masi, R. Gaetano, G. Poggi, and G. Scarpa. A ground truth design tool for multiresolution images. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4999–5002, 2015.
- [76] A. A. Matejko and D. Ansari. Drawing connections between white matter and numerical and mathematical cognition: A literature review. *Neuroscience and Biobehavioral Reviews*, 48:35–52, 2015.
- [77] C. McIntosh and G. Hamarneh. Medial-based deformable models in nonconvex shape-spaces for medical image segmentation. *IEEE Transactions on Medical Imaging*, 31(1):33–50, 2012.
- [78] J. K. Mogali, N. Nallapareddy, C. S. Seelamantula, and M. Unser. A shape-template based two-stage corpus callosum segmentation technique for sagittal plane T1-weighted brain magnetic resonance images. *20th IEEE International Conference on Image Processing (ICIP)*, pages 1177–1181, 2013.
- [79] F. Mokhtarian and A. K. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):789–805, 1992.
- [80] M. A. Monnig, R. E. Thayer, A. Caprihan, E. D. Claus, R. A. Yeo, V. D. Calhoun, and K. E. Hutchison. White matter integrity is associated with alcohol cue reactivity in heavy drinkers. *Brain and behavior*, 4(2):158–170, 2014.
- [81] M.-R. Nazem-Zadeh, S. Saksena, A. Babajani-Fermi, Q. Jiang, H. Soltanian-Zadeh, M. Rosenblum, T. Mikkelsen, and R. Jain. Segmentation of corpus callosum using diffusion tensor imaging: validation in patients with glioblastoma. *BMC medical imaging*, 12(1):10, 2012.
- [82] S. N. Niogi, P. Mukherjee, and B. D. McCandliss. Diffusion tensor imaging segmentation of white matter structures using a Reproducible Objective Quantification Scheme (ROQS). *Neuroimage*, 35(1):166–174, 2007.
- [83] T. E. Oliphant. *A guide to NumPy*, volume 1. Trelgol Publishing USA, 2006.
- [84] F. P. Oliveira and J. M. R. Tavares. Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2):73–93, 2014.
- [85] E. Papageorgiou, P. Spyridonos, D. T. Glotsos, C. D. Stylios, P. Ravazoula, G. Nikiiforidis, and P. P. Groumpos. Brain tumor characterization using the

- soft computing technique of fuzzy cognitive maps. *Applied Soft Computing*, 8(1):820–828, 2008.
- [86] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [87] H. Pedrini and W. R. Schwartz. *Análise de imagens digitais*. Thomson, 2007.
- [88] R. Pérez de Alejo, J. Ruiz-Cabello, M. Cortijo, I. Rodriguez, I. Echave, J. Regadera, J. Arrazola, P. Avilés, P. Barreiro, D. Gargallo, and M. Graça. Computer-assisted enhanced volumetric segmentation magnetic resonance imaging data using a mixture of artificial neural networks. *Magnetic Resonance Imaging*, 21(8):901–912, 2003.
- [89] A. Pfefferbaum, M. J. Rosenbloom, T. Rohlfing, E. Adalsteinsson, C. A. Kemper, S. Deresinski, and E. V. Sullivan. Contribution of alcoholism to brain dysmorphology in HIV infection: Effects on the ventricles and corpus callosum. *NeuroImage*, 33(1):239–251, 2006.
- [90] C. Pierpaoli and P. J. Basser. Toward a quantitative assessment of diffusion anisotropy. *Magnetic resonance in Medicine*, 36(6):893–906, 1996.
- [91] A. Pitiot, H. Delingette, P. M. Thompson, and N. Ayache. Expert knowledge-guided segmentation system for brain MRI. *NeuroImage*, 23:S85–S96, 2004.
- [92] A. Pitiot, A. Toga, N. Ayache, and P. Thompson. Texture based MRI segmentation with a two-stage hybrid neural classifier. *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290)*, 3:2053–2058, 2002.
- [93] A. Pitiot, A. W. Toga, and P. M. Thompson. Adaptive elastic segmentation of brain MRI via shape-model-guided evolutionary programming. *IEEE Transactions on Medical Imaging*, 21(8):910–923, 2002.
- [94] I. I. Rabi, S. Millman, P. Kusch, and J. R. Zacharias. The molecular beam resonance method for measuring nuclear magnetic moments. *Physical review*, 55(6):526, 1939.
- [95] K. Revett. An introduction to magnetic resonance imaging: From image acquisition to clinical diagnosis. In *Innovations in Intelligent Image Analysis*, pages 127–161. 2011.
- [96] D. Rey, G. Subsol, H. Delingette, and N. Ayache. Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. *Medical image analysis*, 6(2):163–179, 2002.

- [97] L. Rittner, J. S. W. Campbell, P. F. Freitas, S. Appenzeller, G. Bruce Pike, and R. A. Lotufo. Analysis of scalar maps for the segmentation of the corpus callosum in diffusion tensor fields. *Journal of Mathematical Imaging and Vision*, 45(3):214–226, 2013.
- [98] J. M. Rondina. Segmentação interativa do ventrículo esquerdo em sequências de imagens de ressonância magnética. *Universidade Estadual de Campinas, UNICAMP, Brasil*, Mestrado em Engenharia Elétrica, 2001.
- [99] L. Samartzis, D. Dima, P. Fusar-Poli, and M. Kyriakopoulos. White matter alterations in early stages of schizophrenia: A systematic review of diffusion tensor imaging studies. *Journal of Neuroimaging*, 24(2):101–110, 2014.
- [100] G. Spasojević, S. Malobabić, D. Mikić, S. Vujnović, O. P. Spasojević, and A. Maliković. Sex differences of human corpus callosum revealed by polar coordinate system: magnetic resonance imaging study. *Folia morphologica*, 74(4):414–420, 2015.
- [101] A. C. Stanfield, A. M. McIntosh, M. D. Spencer, R. Philip, S. Gaur, and S. M. Lawrie. Towards a neuroanatomy of autism: A systematic review and meta-analysis of structural magnetic resonance imaging studies. *European Psychiatry*, 23(4):289–299, 2008.
- [102] E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, 42(1):288–292, 1965.
- [103] P. Su, Z. Xue, L. Chi, J. Yang, and S. T. Wong. Support vector machine (SVM) active learning for automated Glioblastoma segmentation. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 598–601, 2012.
- [104] E. Tan, R. Srinivasan, and R. Robb. Intensity-Based Shape Propagation for Volumetric Image Segmentation. *3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano.*, pages 738–741, 2006.
- [105] P. M. Thompson, R. P. Woods, M. S. Mega, and A. W. Toga. Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. *Human brain mapping*, 9(2):81–92, 2000.
- [106] I. Unterberger, R. Bauer, G. Walser, and G. Bauer. Corpus callosum and epilepsies. *Seizure*, 37:55–60, 2016.
- [107] B. Van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever. A non-linear gray-level appearance model improves active shape model segmentation. *IEEE Workshop on Mathematical Models in Biomedical Image Analysis*, pages 205–212, 2001.

- [108] B. Van Ginneken, A. F. Frangi, J. J. Staal, B. M. Ter Haar Romeny, and M. A. Viergever. Active shape model segmentation with optimal features. *IEEE Transactions on Medical Imaging*, 21(8):924–933, 2002.
- [109] K. Von Plessen, A. Lundervold, N. Duta, E. Heiervang, F. Klauschen, a. I. Smievoll, L. Ersland, and K. Hugdahl. Less developed corpus callosum in dyslexic subjects - A structural MRI study. *Neuropsychologia*, 40(7):1035–1044, 2002.
- [110] M. Walterfang, M. Yücel, S. Barton, D. C. Reutens, A. G. Wood, J. Chen, V. Lorenzetti, D. Velakoulis, C. Pantelis, and N. B. Allen. Corpus callosum size and shape in individuals with current and past depression. *Journal of affective disorders*, 115(3):411–420, 2009.
- [111] S. K. Warfield, K. H. Zou, and W. M. Wells. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. *IEEE Transactions on Medical Imaging*, 23(7):903–921, 2004.
- [112] S. F. Witelson and C. H. Goldsmith. The relationship of hand preference to anatomy of the corpus callosum in men. *Brain Research*, 545(1-2):175–182, 1991.
- [113] D. Zhang and G. Lu. Evaluation of similarity measurement for image retrieval. In *Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on*, volume 2, pages 928–931. IEEE, 2003.
- [114] J. Zhang, S. K. Zhou, D. Comaniciu, and L. McMillan. Conditional density learning via regression with application to deformable shape segmentation. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2, 2008.
- [115] Y. J. Zhang. A survey on evaluation methods for image segmentation. *Pattern recognition*, 29(8):1335–1346, 1996.

Apêndice A: Análise qualitativa do perfil de curvatura

Ao longo deste trabalho foi desenvolvida uma ferramenta interativa (Fig. 0.1) que permite a visualização e exploração do perfil de curvatura. A ferramenta apresenta as seguintes funcionalidades:

- Permite carregar segmentações (máximo 4), visualizar os contornos suaves e o perfil de curvatura associado a cada um deles.
- Visualização do ponto pivô sobre o contorno da estrutura e da curvatura associada a este ponto sobre o perfil de curvatura.
- Permite visualizar os segmentos que vão desde o ponto pivô aos pontos anterior (vermelho) e posterior (verde).
- Permite mudar o valor de l_s para obter o perfil de curvatura em diferentes resoluções.

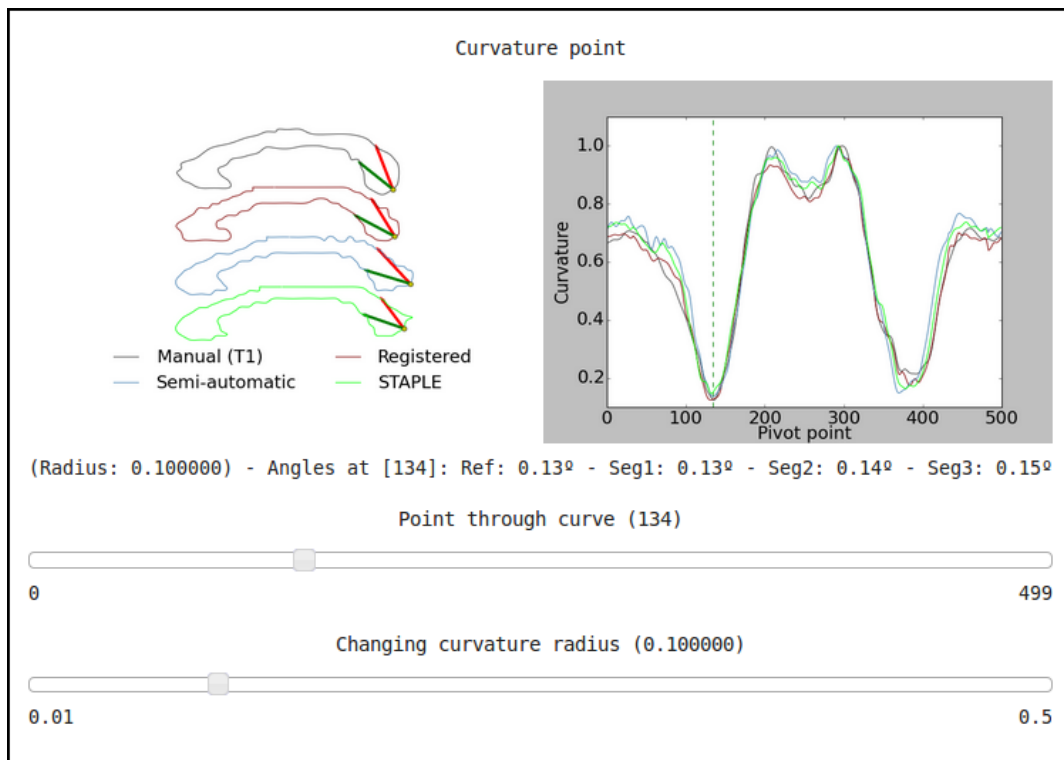


Figura 0.1 – Ferramenta interativa para exploração e estudo do perfil de curvatura: captura da tela mostrando os perfis de curvatura extraídos de quatro segmentações do CC

Com o uso desta ferramenta, é possível fazer uma caracterização qualitativa do perfil. Na Fig. 0.2 é visualizado o perfil de curvatura com uma resolução fixa de 0.08

(segmentação manual registrada em difusão) do mesmo sujeito. Esta resolução permite visualizar os detalhes do CC sem ser afetado significativamente pelo ruído da borda. O ponto pivô foi posicionado em vários pontos ao longo da *spline* (figs. 1.2a to 1.2g). Os pontos equivalentes no perfil a cada posição são apresentados na Fig. 1.2h. As setas representam os segmentos anterior (vermelho) e posterior (verde) ao ponto pivô.

Também analisamos os perfis de curvatura do padrão-ouro (segmentação manual obtida em T_1) e um padrão-prata (segmentação manual registrada em difusão) em diferentes valores de ls (diferentes resoluções) do mesmo sujeito (Fig. 1.3).

Sobre as Figs. 1.2 e 1.3 podemos observar:

- Perfis de curvatura com valores de ls muito pequenos ($ls < 0.04$) são consideravelmente afetados pelo ruído do contorno da estrutura, mesmo com a suavização da *spline* (Fig. 1.3a).
- Valores de ls baixos ($0.04 < ls < 0.1$) conseguem resoluções que permitem apreciar os detalhes do CC, e portanto são indicados para caracterização da forma desta estrutura (Fig. 1.2).
- Para $ls = 0.1$ (Fig. 1.3b), os mínimos correspondem aos extremos do CC e os máximos às bordas internas do CC.
- Para $ls = 0.25$ (Fig. 1.3c), os mínimos são os extremos do CC, entanto que o máximo é a parte inferior do corpo do CC.
- Para $ls = 0.4$ (Fig. 1.3d), apresentam-se dois máximos: o máximo absoluto é a parte superior do corpo do CC e o máximo local é a parte inferior do corpo do CC. Os mínimos correspondem aos extremos externos do joelho e do esplênio.
- A partir de 0.15 ($ls > 0.15$), o aumento de ls faz com que os máximos do perfil de curvatura fiquem mais aguçados e os mínimos se estiquem. Isto tem como consequência que a média do perfil de curvatura tende a decrescer.

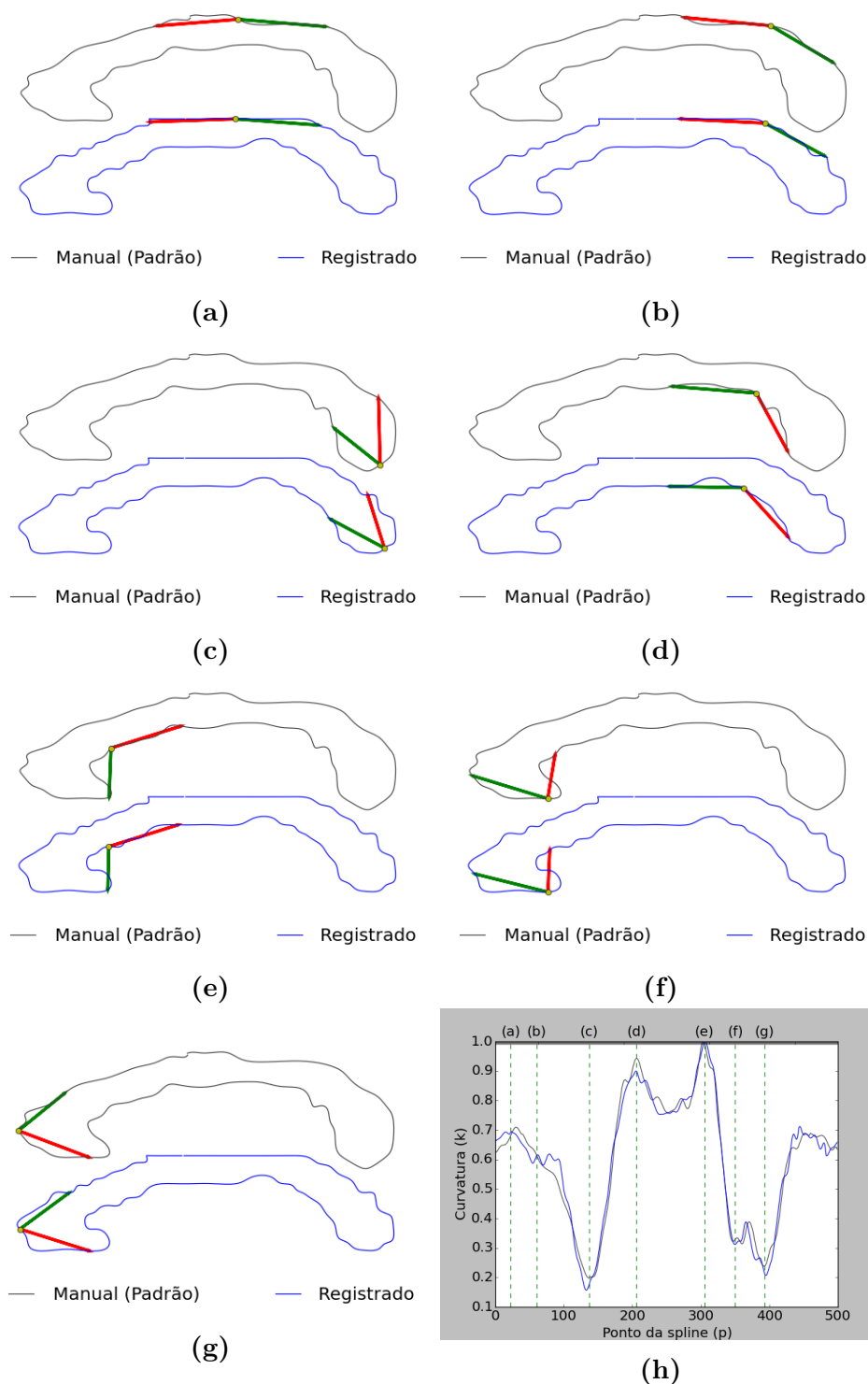
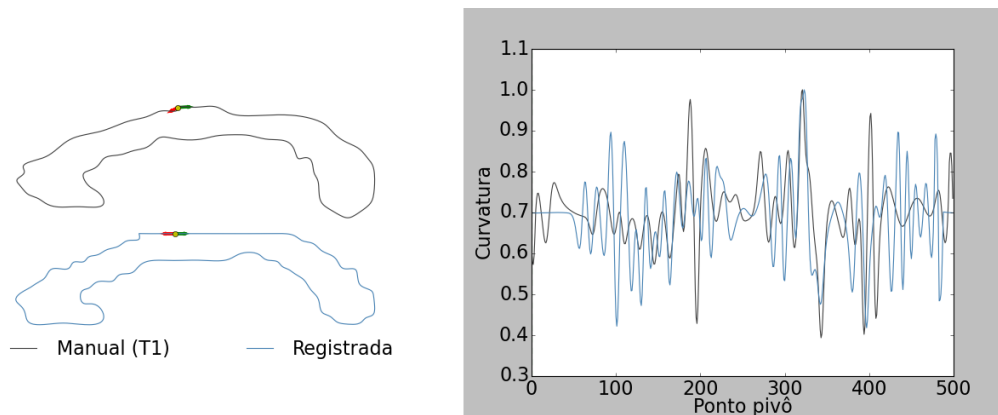
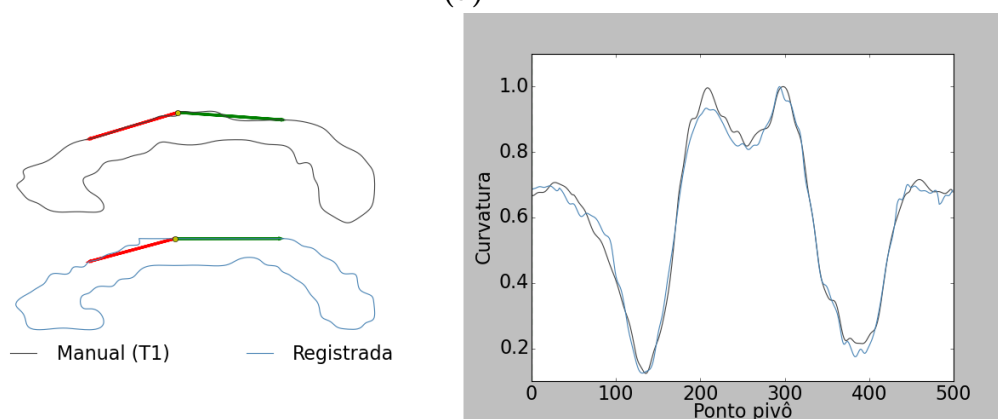


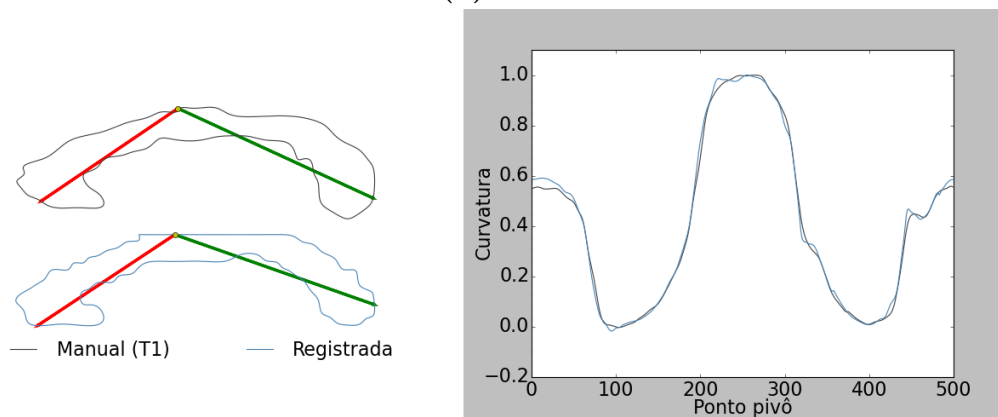
Figura 1.2 – Segmentos anterior (vermelho) e posterior (verde) com $ls = 0.08$ em diferentes pontos pivô da *spline* para padrões manual e registrado: a) $p = 22$, b) $p = 60$, c) $p = 137$, d) $p = 206$, e) $p = 305$, f) $p = 350$, g) $p = 393$. h) Localização dos pontos (a) a (g) sobre o perfil de curvatura



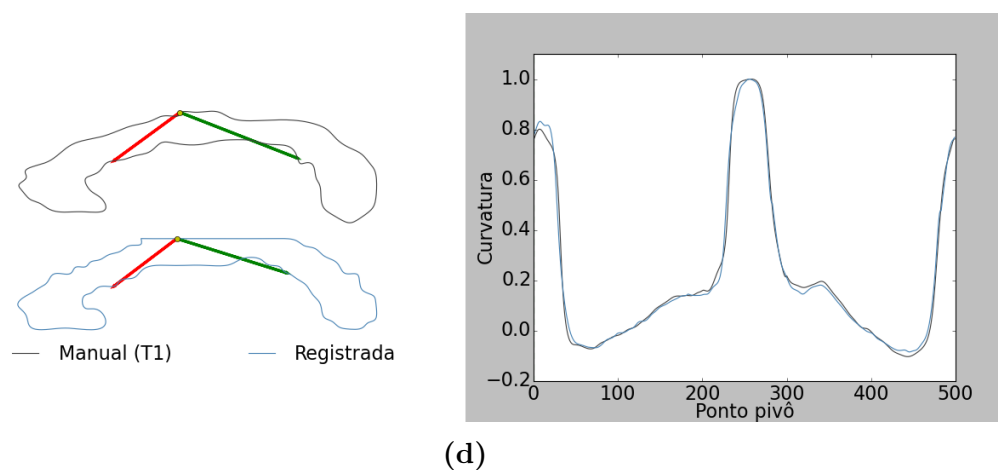
(a)



(b)



(c)



(d)

Figura 1.3 – Perfil de curvatura para diferentes valores de ls ($p = 0$): a) $ls = 0.01$, b) $ls = 0.1$, c) $ls = 0.25$, d) $ls = 0.4$

Apêndice B: Avaliação de segmentações do CC usando o método usual por sobreposição

As três segmentações apresentadas foram avaliadas na população de 145 sujeitos usando os padrões-prata em difusão junto com métricas de sobreposição (Fig. 0.4). Esta análise é abordada primeiro com o padrão registrado, como é feito normalmente na literatura, para depois usar os dois padrões-prata adicionais: semi-automático e STAPLE.

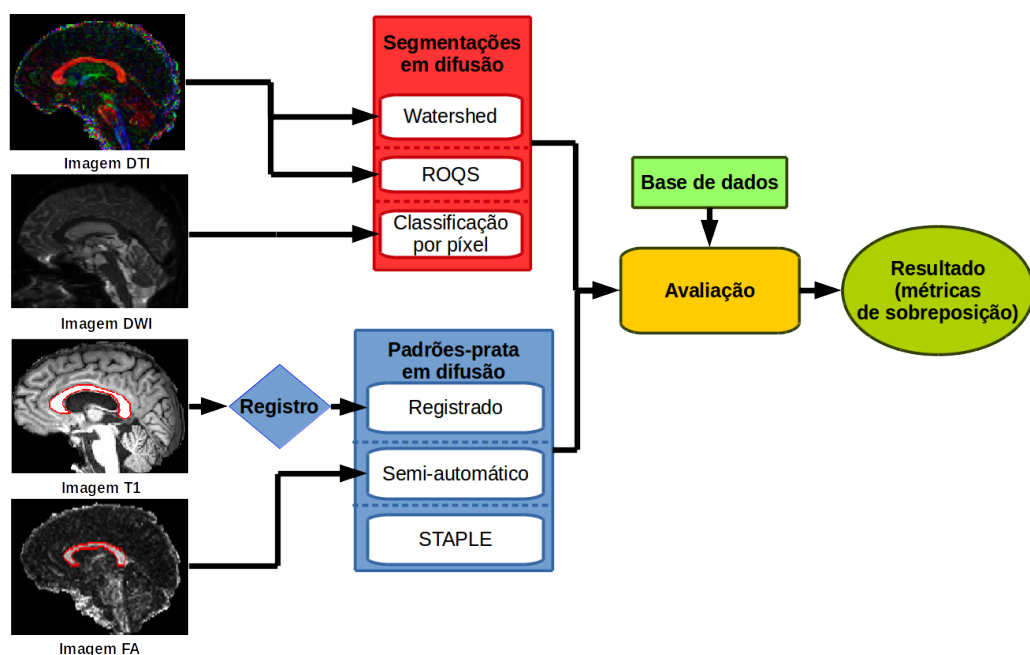


Figura 0.4 – Esquema de avaliação usual de segmentação do CC em difusão por sobreposição

Método usual de avaliação por sobreposição usando padrão registrado

Como já visto, o método de avaliação usual por sobreposição faz uso de um padrão-prata para avaliar as segmentações. Este padrão é normalmente o registrado em difusão. As três segmentações foram avaliadas usando o padrão-prata registrado junto com as métricas descritas na seção 3.3: sensibilidade, especificidades, Dice, área e distância Hausdorff para 145 sujeitos. Os resultados de avaliação são apresentados em forma de diagrama de caixa. Neste caso, apresenta-se um gráfico por métrica onde cada caixa corresponde a uma segmentação quando avaliada contra o padrão-prata registrado (Fig. 0.5).

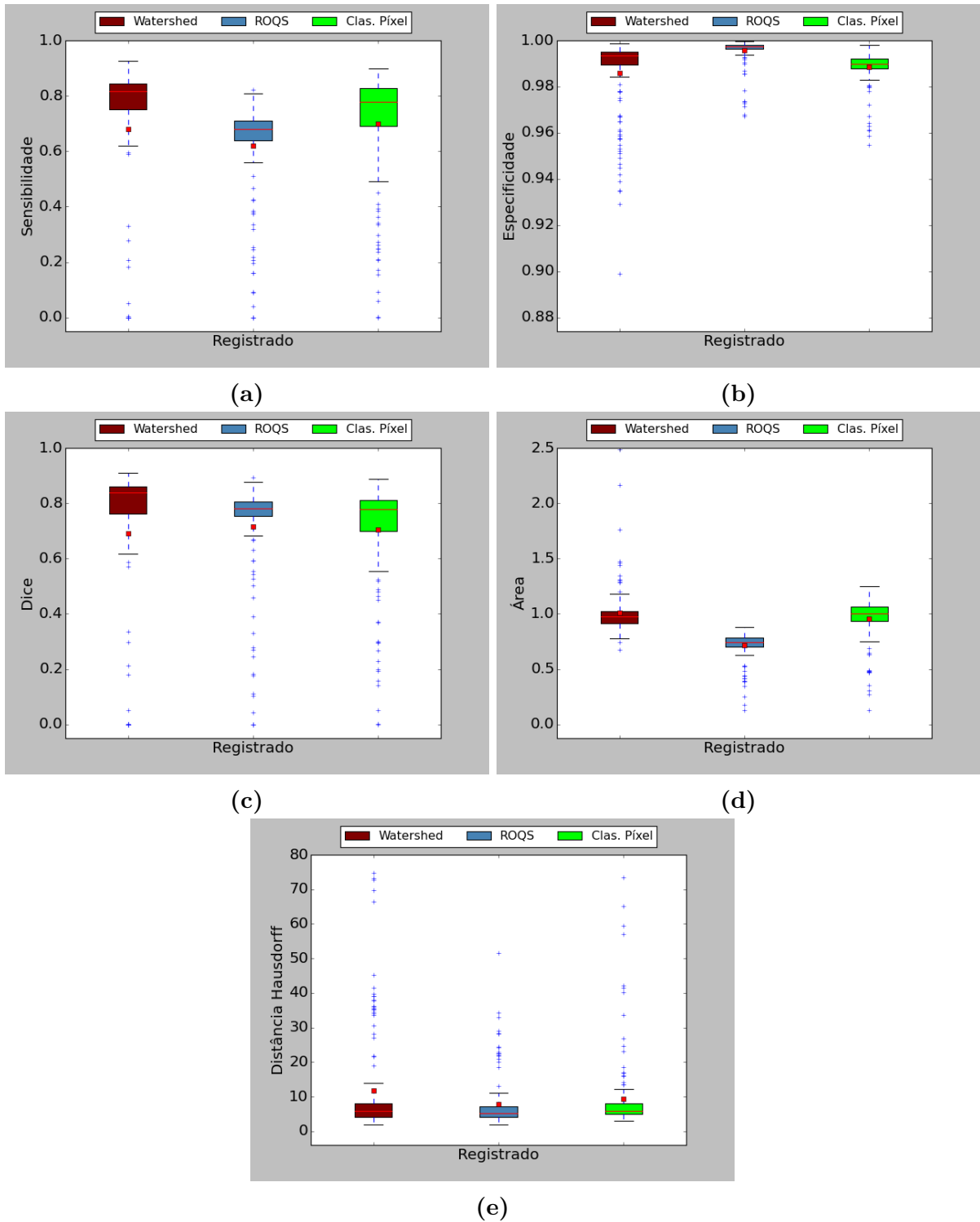


Figura 2.2 – Avaliação das segmentações em difusão para 145 sujeitos usando padrão-prata registrado: a) Sensibilidade, b) Especificidade, c) Dice, d) Área, e) Distância Hausdorff

Em primeiro lugar, pode-se notar a grande quantidade de *outliers* que apresentam os resultados. Os *outliers* se devem principalmente à segmentação *Watershed*, que apresenta a maior quantidade de erros, e ao padrão-prata registrado, que, em alguns casos, tem problemas de registro onde o padrão fica fora do CC. A segmentação baseada em *Watershed* apresenta a maior quantidade de segmentações fora do CC (14%), seguida

pelo ROQS (1,4%), e o método de classificação por pixel (0.7%).

Olhando para os diagramas de caixa (Fig. 2.2), *Watershed* é a segmentação que melhor desempenho tem, na maior parte das métricas, quando comparado com o padrão registrado, seguida pela segmentação usando classificação por pixel e finalmente ROQS. A segmentação ROQS costuma pegar menos pixels da borda do que as outras segmentações e portanto, sua probabilidade de falsos negativos é maior (sensibilidade baixa), entanto que a probabilidade de falsos positivos é baixa (especificidade alta). Isto demonstra que estas métricas (sensibilidade e especificidade) apresentam aspectos diferentes e paralelos da segmentação e por tanto, devem ser olhadas conjuntamente para obter uma avaliação mais global da segmentação. Já o Coeficiente de Dice é uma métrica que leva em conta medição de falsos positivos e negativos, porém neste experimento apresentou menos diferença entre as segmentações, tendo medianas similares. A area mostra um comportamento similar, porém, é uma métrica pobre por definição: um análise simples permite demonstrar que a relação de área pode ter valor igual a um, para segmentações totalmente erradas como aparece na Fig. 2.3; neste caso, o Dice é igual a zero evidenciando o erro de segmentação. Por fim, a distância Hausdorff é pouco discriminativa (Fig. 2.2e) e propensa a superestimar erros na borda¹.

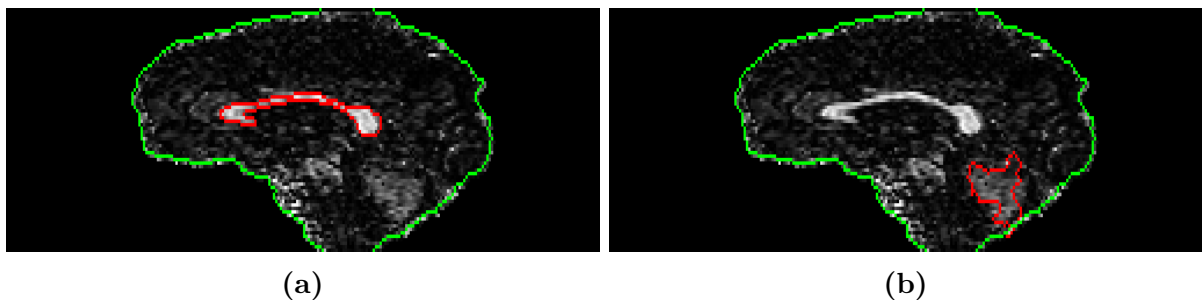


Figura 2.3 – Comparação área e Dice em segmentação errada ($Area = 1.01$, $Dice = 0$): a) Padrão semi-automático, b) Segmentação *Watershed*

Método usual de avaliação por sobreposição usando padrões-prata: semiautomático e STAPLE

Dado os problemas do padrão-prata registrado, são usados os dois padrões-prata construídos diretamente em difusão: o semi-automático e o STAPLE. Estes dois padrões-prata são usados para avaliar as três segmentações usando as mesmas métricas do experimento anterior. Da mesma forma, é apresentado um gráfico por cada métrica. Cada gráfico apresenta dois grupos de caixas: as do lado esquerdo são as avaliações feitas com o padrão semi-automático e as do lado direito são as avaliações usando o padrão STAPLE.

¹ Existe uma implementação que calcula a média da distância Hausdorff sobre todos os pontos que faz com que seja menos afetada por *outliers*, porém esta versão é pouco usada na literatura

Cada grupo possui a avaliação correspondente às três segmentações: *Watershed*, ROQS e classificação por pixel (Fig. 2.4).

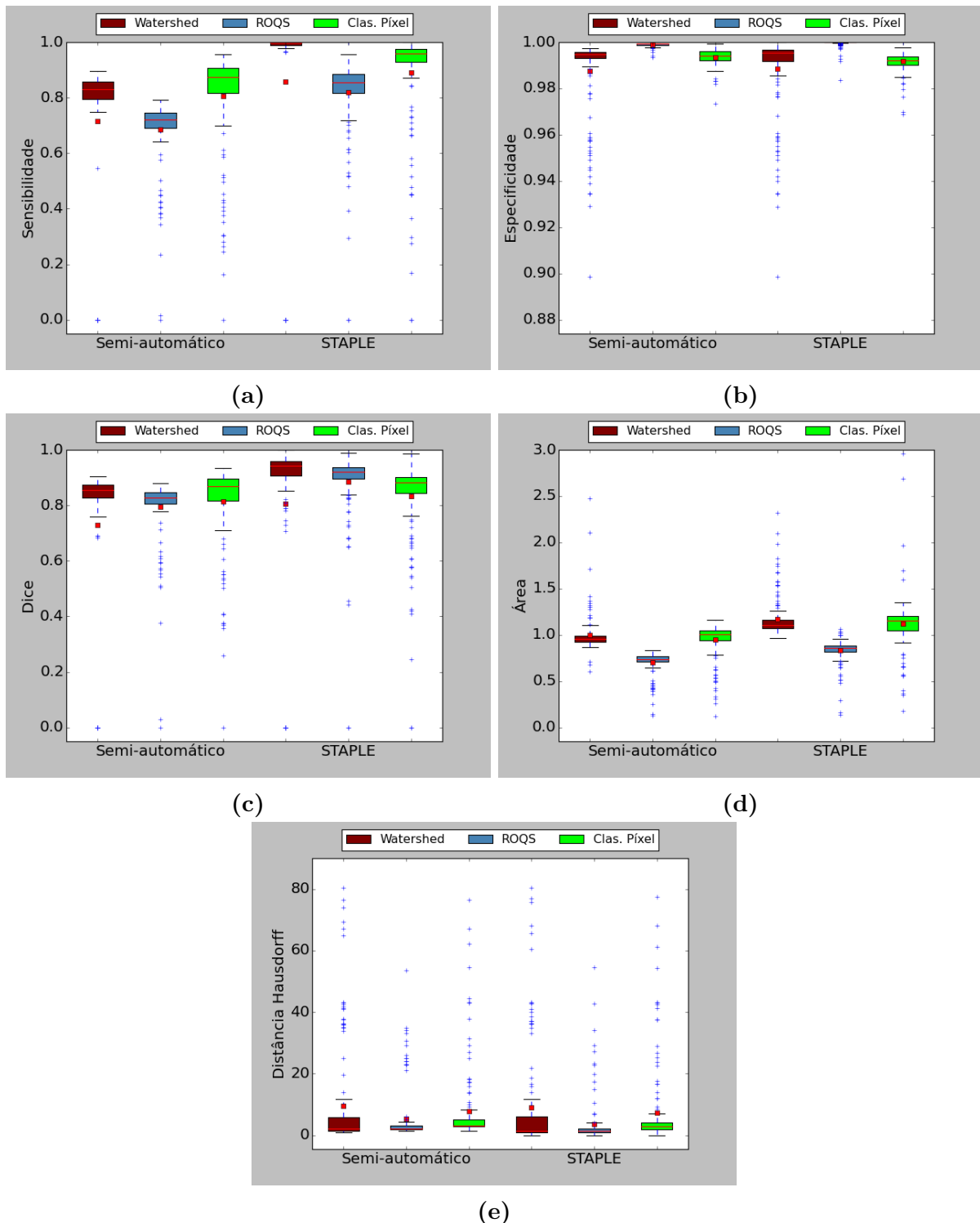


Figura 2.4 – Avaliação das segmentações em difusão para 145 sujeitos usando padrões-prata semi-automático e STAPLE: a) Sensibilidade, b) Especificidade, c) Dice, d) Área, e) Distância Hausdorff

Pode-se notar que o padrão-prata usado determina qual segmentação é considerada superior. Quando usado o padrão semi-automático, o método de classificação baseado em

pixel apresenta desempenho superior; por outro lado o padrão-prata STAPLE faz com que o método baseado em *Watershed* apresente melhor desempenho.