



UNIVERSIDADE ESTADUAL DE CAMPINAS  
Faculdade de Engenharia Elétrica e de Computação

Gustavo Retuci Pinheiro

**Aplicação de imagens de difusão por  
ressonância magnética e redes convolucionais  
para avanços em segmentação do tálamo e seus  
núcleos**

Campinas

2024

Gustavo Retuci Pinheiro

**Aplicação de imagens de difusão por ressonância magnética e redes convolucionais para avanços em segmentação do tálamo e seus núcleos**

Tese 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 Doutor em Engenharia Elétrica, na Área de Engenharia de Computação.

Orientador: Profa. Dra. Letícia Rittner

Este exemplar corresponde à versão final da tese defendida pelo aluno Gustavo Retuci Pinheiro, e orientada pela Profa. Dra. Letícia Rittner

---

Campinas

2024

Ficha catalográfica  
Universidade Estadual de Campinas (UNICAMP)  
Biblioteca da Área de Engenharia e Arquitetura  
Rose Meire da Silva - CRB 8/5974

P655a Pinheiro, Gustavo Retuci, 1989-  
Aplicação de imagens de difusão por ressonância magnética e redes convolucionais para avanços em segmentação do tálamo e seus núcleos / Gustavo Retuci Pinheiro. – Campinas, SP : [s.n.], 2024.

Orientador: Letícia Rittner.  
Tese (doutorado) – Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia Elétrica e de Computação.

1. Tálamo. 2. Segmentação de imagens médicas. 3. Redes neurais convolucionais. 4. Imagem de difusão por ressonância magnética. I. Rittner, Letícia, 1972-. II. Universidade Estadual de Campinas (UNICAMP). Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações complementares

**Título em outro idioma:** Application of diffusion magnetic resonance imaging and convolutional networks for advances in segmentation of the thalamus and its nuclei

**Palavras-chave em inglês:**

Thalamus

Medical image segmentation

Convolutional neural networks

Diffusion magnetic resonance imaging

**Área de concentração:** Engenharia de Computação

**Titulação:** Doutor em Engenharia Elétrica

**Banca examinadora:**

Letícia Rittner [Orientador]

Carina Marconi Germer

Matheus Cardoso Moraes

Mariana Pinheiro Bento

Bruno Lima Pessoa

**Data de defesa:** 28-11-2024

**Programa de Pós-Graduação:** Engenharia Elétrica

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0001-8627-9468>

- Currículo Lattes do autor: <http://lattes.cnpq.br/9257386257417266>

## COMISSÃO JULGADORA – TESE DE DOUTORADO

**Candidato:** Gustavo Retuci Pinheiro RA: 162793

**Data da Defesa:** 28 de novembro de 2024

**Título da Tese:** “Aplicação de imagens de difusão por ressonância magnética e redes convolucionais para avanços em segmentação do tálamo e seus núcleos”.

Prof. Dra. Leticia Rittner

Prof. Dr. Matheus Cardoso Moraes

Prof. Dra. Mariana Pinheiro Bento

Prof. Dra. Carina Marconi Germer

Prof. Dr. Bruno Lima Pessoa

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

Dedico minha tese de doutorado aos meus conterrâneos de Turvolândia-MG e aos meus familiares, especialmente aos meus avós paternos, Maria Antônia Pinheiro e Oswaldo Anastácio Pinheiro.

Minha avó não teve a oportunidade de ser alfabetizada e meu avô teve condições de estudar apenas até a quarta série do ensino básico, no entanto sabiam da importância do conhecimento e dedicaram ao máximo para que seus filhos e netos pudessem ter uma formação.

Apesar da pouca escolaridade, meu avô lembrava aos seus mais de 90 anos de idade claramente de tudo que havia estudado 80 anos antes. Sabia falar tupi-guarani, conhecia profundamente as regras gramaticais portuguesas, mesmo que aos termos das regras ortográficas de seu tempo, sabia história antiga, história mundial, história do Brasil, geografia mundial, nacional e regional, além de grandes conhecimentos de ciências gerais, principalmente sobre anatomia. Por vezes era convidado a dar aulas e palestras nas escolas da cidade dos mais diversos assuntos, onde fazia amizade com as crianças que posteriormente ficavam trocando conhecimentos com ele no banco da praça.

Durante seu último mês de vida, tive a felicidade de ter meu primeiro artigo científico aceito para ser apresentado em uma conferência internacional. Com isso, uma das últimas frases que ouvi de meu querido avô foi: "agora você será um cientista conhecido no mundo todo".

Espero que a conclusão do meu doutorado seja apenas uma etapa de uma grande carreira científica que honre o esforço e a expectativa dos meus familiares.

# Agradecimentos

Agradeço primeiramente a minha orientadora Letícia Rittner, alguém que me orgulho muito de ter sido orientando. Agradeço pelos conhecimentos a mim passado, agradeço pela paciência, à toda ajuda no desenvolvimento de minha pesquisa e dos artigos que publicamos em conjunto. Agradeço pela amizade, por ser uma pesquisadora tão dedicada e por ter contribuído tanto na minha formação como cientista.

Agradeço a minha esposa, Mirela Soares França, que me apoiou durante toda minha pós graduação. Agradeço também aos meu colegas de trabalho do MICLab, especialmente os colegas Livia, Diedre, Caio, Thays e Joany, que tiveram contribuições diretas em minha pesquisa.

Agradeço também a professora Simone Appeneller que colaborou de forma significativa com os dados utilizados na pesquisa e as professoras Gloria Menegaz e Lorenza Brusini que trabalharam comigo durante o período de intercâmbio na universidade de Verona.

O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil (159829/2017-8).

# Resumo

O tálamo é uma pequena estrutura subcortical que fica na região central do cérebro e que está relacionada com diversas funções como transmissão de informações sensoriais, regulação do sono e controle motor. Alterações nessa estrutura podem causar condições adversas de saúde como tremores e doença de Parkinson.

Tanto para o acompanhamento do tratamento de doenças relacionadas ao tálamo, quanto para pesquisas populacionais onde o tálamo é objeto de estudo, a segmentação e o parcelamento dessa estrutura é uma etapa de extrema importância. Assim, métodos robustos e rápidos são necessários para segmentar ou parcelar essa estrutura.

A segmentação do tálamo normalmente é feita com o emprego imagens de ressonâncias magnética ponderadas em T1 e com métodos baseados em atlas que são relativamente lentos e não necessariamente atingindo a melhor representação da estrutura. Além disso, principalmente no problema de parcelamento do tálamo, as imagens convencionais ponderadas em T1 apresentam baixo contraste, limitando ainda mais os resultados. Por outro lado, imagens de difusão têm o potencial de melhores contrastes no interior do tálamo por carregarem mais informações da microestrutura dos tecidos.

Com o objetivo de resolver a limitação de qualidade do resultado das segmentações atingidas pelos métodos atuais e o tempo de processamento das imagens, este trabalho propõe métodos que utilizam imagens ponderadas por difusão e redes neurais convolucionais para segmentar e parcelar o tálamo.

Os resultados atingidos são superiores aos métodos atuais, atingindo DICEs próximos de 0,9, estabelecendo o estado da arte, e os tempos de execução são de poucos minutos, contra horas tomadas pelos métodos de registro de atlas.

**Palavras-chaves:** Tálamo; Segmentação; CNN; Imagens de Difusão.

# Abstract

The thalamus is a small subcortical structure located in the central region of the brain and it is related to several functions such as transmitting sensory information, regulating sleep, and motor control. Changes in this structure can lead to adverse health conditions such as tremors and Parkinson's disease.

In both cases of treatment of diseases related to the thalamus and for populational studies where the thalamus is the target, the segmentation and parcellation of this structure is an essential step in conducting research or treatment. Therefore, robust and fast methods are needed to segment or parcellate this structure.

The segmentation of the thalamus is usually done using T1-weighted magnetic resonance images and employing relatively slow atlas-based methods, which do not necessarily provide the best representation of the structure. Furthermore, especially in the problem of parcellation of the thalamus, conventional T1-weighted images have low contrast, further limiting the results. On the other hand, diffusion images have the potential for better contrasts inside the thalamus due to carrying more information about the microstructure of the tissues.

With the aim of solving the limitation of the quality of the segmentation achieved by current methods and the image processing time, this work proposes methods that use diffusion-weighted images and convolutional neural networks to segment and parcellate the thalamus.

The achieved results are superior to current methods, reaching DICEs close to 0.9 and establishing the state of the art while keeping execution times to only a few minutes, compared to hours taken by atlas registration methods.

**Keywords:** Thalamus; Segmentation; CNN; Diffusion MRI.

# Lista de ilustrações

Figura 1.1 – Tálamo e sua posição no cérebro . . . . .	14
Figura 2.1 – Exemplos de imagens de RM: T1; T2; FLAIR; DWI; ADC; DTI. . . . .	29
Figura 2.2 – Exemplos de imagens de RM: T1w e mapas escalares de DTI . . . . .	30
Figura 2.3 – Mapas de Gradiente Morfológico Tensorial . . . . .	35
Figura 2.4 – Mapas de divergente proveniente de campos tensoriais . . . . .	37
Figura 2.5 – Mapa de divergente na região do Corpo Calosos . . . . .	38
Figura 2.6 – Demografia dos dados HCP . . . . .	40
Figura 2.7 – Pré-processamento dos dados . . . . .	42
Figura 2.8 – Composição do STAPLE como padrão-prata para segmentação do tálamo	47
Figura 2.9 – Parcelamento do tálamo . . . . .	48
Figura 2.10 – Arquitetura U-Net para aplicações multi modalidade . . . . .	52
Figura 3.1 – Mapa do método de segmentação do tálamo . . . . .	53
Figura 3.2 – Fluxograma do método de treinamento da CNN para segmentação do tálamo . . . . .	54
Figura 3.3 – <i>Patches</i> centralizados em diferentes núcleos do tálamo . . . . .	57
Figura 3.4 – Método de agregação de vistas . . . . .	60
Figura 3.5 – Fluxograma do processo de inferência e segmentação . . . . .	62
Figura 3.6 – Composição pelo STAPLE para segmentação com a QuickNAT . . . . .	63
Figura 3.7 – Resultados qualitativos da segmentação do tálamo . . . . .	65
Figura 3.8 – Desempenho para modelos treinados em diferentes vistas . . . . .	66
Figura 3.9 – Distribuição DICE para diferentes combinações de mapas . . . . .	67
Figura 4.1 – Fluxograma do processo de inferência e parcelamento . . . . .	74
Figura 4.2 – DICE médio do parcelamento para cada combinação de mapas de entrada	76
Figura 4.3 – Volume médio e DICE médio para cada núcleo do tálamo no conjunto de testes. . . . .	79
Figura 4.4 – Medidas de incerteza de predição na região do tálamo . . . . .	81

# Lista de tabelas

Tabela 1.1 – Métodos de segmentação do tálamo . . . . .	19
Tabela 1.2 – Métodos de parcelamento do tálamo . . . . .	23
Tabela 2.1 – Conjuntos de dados . . . . .	39
Tabela 2.2 – Comparação de máscaras entre avaliadores . . . . .	46
Tabela 2.3 – Núcleos do tálamo . . . . .	48
Tabela 3.1 – DICE médio da segmentação do tálamo no conjunto de teste para cada combinação de entrada da CNN . . . . .	64
Tabela 3.2 – Comparação entre métodos de segmentação do tálamo . . . . .	64
Tabela 3.3 – Comparação do desempenho da CNN com e sem ajuste fino . . . . .	66
Tabela 3.4 – Comparação do desempenho para diferentes combinações de imagens de entrada da CNN. . . . .	67
Tabela 4.1 – Lista dos experimentos de parcelamento: combinação de canais de en- trada. . . . .	76
Tabela 4.2 – DICE e volume médio de cada núcleo do tálamo no conjunto de testes usando o melhor modelo de CNN . . . . .	78
Tabela 4.3 – Métodos de parcelamento do tálamo . . . . .	79

# Lista de abreviações

AD	Axial Diffusivity (Difusividade Axial)
ADC	Apparent Diffusion Coefficient (Coeficiente de Difusão Aparente)
AHD	Average Hausdorff Distance (Distância Hausdorff Média)
AVC	Acidente Vascular Cerebral
DBS	Deep Brain Stimulation (Estimulação Cerebral Profunda)
CNN	Convolutional Neural Network (Rede Neural Convolucional)
DICE	Coeficiente Dice (métrica de similaridade de segmentações)
DL	Deep Learning (Aprendizado Profundo)
dMRI	Diffusion Magnetic Resonance Imaging (Imagem de Difusão por Ressonância Magnética)
DTI	Diffusion Tensor Imaging (Imagem por Tensor de Difusão)
DWI	Diffusion Weighted Imaging (Imagem Ponderada em Difusão)
EPI	Echo Planar Imaging (Imagem de Eco Planar)
FA	Fractional Anisotropy (Anisotropia Fracional)
FNR	False Negative Rate (Taxa de Falsos Negativos)
FPR	False Positive Rate (Taxa de Falsos Positivos)
HARDI	High Angular Resolution Diffusion Imaging (Imagem de Difusão de Alta Resolução Angular)
HCP	Human Connectome Project (Projeto Conectoma Humano)
HCU	Hospital de Clínicas - UNICAMP
MD	Mean Diffusivity (Difusividade Média)
MP RAGE	Magnetization Prepared - RApid Gradient Echo (Magnetização Preparada - Eco de Gradiente Rápido)
MRI	Magnetic Resonance Image (Imagem por Ressonância Magnética)
ODF	Orientation Distribution Function (Função de Distribuição de Orientação)
RD	Radial Diffusivity (Difusividade Radial)
T1w	T1 Weighted MRI (Imagem de Ressonância Magnética Ponderada em T1)
T2w	T2 Weighted MRI (Imagem de Ressonância Magnética Ponderada em T2)
TMG	Tensorial Morphological Gradient (Gradiente Morfológico Tensorial)
VS	( Similaridade Volumétrica)
WMn	White-matter-nulled MPRAGE (MPRAGE anulado na substância branca)
MPRAGE	

# Sumário

<b>1</b>	<b>Introdução</b>	<b>14</b>
1.1	Segmentação do Tálamo	14
1.2	Trabalhos relacionados	16
1.2.1	Métodos de segmentação do tálamo	16
1.2.2	Métodos de parcelamento do tálamo	20
1.3	Motivação	24
1.4	Objetivo do projeto de pesquisa	24
1.5	Principais desafios	25
1.6	Principais contribuições	27
1.7	Organização da tese	28
<b>2</b>	<b>Dados</b>	<b>29</b>
2.1	Imagem de Ressonância Magnética	29
2.1.1	Imagem de Ressonância Magnética Ponderada em T1 (T1w)	30
2.1.2	Imagem de Difusão por Ressonância Magnética (dMRI)	31
2.1.2.1	Mapas provenientes de Imagens de Tensores de Difusão	33
2.2	Conjuntos de dados	38
2.2.1	HCP	39
2.2.2	HCU	40
2.3	Pré-processamento	40
2.3.1	Normalização espacial e registro das imagens	42
2.3.2	Cálculo dos mapas de difusão	43
2.3.3	Normalização de intensidade	44
2.4	Máscaras	44
2.4.1	Anotação manual dos dados	45
2.4.1.1	Estudo intra- e inter-avaliador	46
2.4.2	Anotações automáticas	46
2.4.2.1	Segmentação	47
2.4.2.2	Parcelamento	47
2.5	Redes Neurais Convolucionais e fusão de dados	49
2.5.1	U-Net	49
2.5.2	Fusão de dados em Redes Neurais Convolucionais	50
<b>3</b>	<b>Segmentação do Tálamo</b>	<b>53</b>
3.1	Preparação dos dados	54
3.2	Técnicas para treinamento da CNN no contexto de imagens de difusão	56

3.2.1	Treinamento usando <i>Patches</i> . . . . .	56
3.2.2	Expansão de dados ( <i>data augmentation</i> ) . . . . .	56
3.2.3	Abordagem multitarefas . . . . .	57
3.2.4	Função de custo . . . . .	57
3.3	Ambiente computacional . . . . .	58
3.4	Pós-processamento . . . . .	59
3.4.1	Agregação de vistas . . . . .	59
3.4.2	Filtragem por volumetria . . . . .	60
3.5	Inferência e segmentação . . . . .	61
3.6	Resultados . . . . .	61
3.6.1	Resultados com a QuickNAT . . . . .	62
3.6.2	Resultados com a U-Net . . . . .	65
<b>4</b>	<b>Parcelamento do Tálamo</b> . . . . .	<b>69</b>
4.1	Núcleos do tálamo . . . . .	69
4.2	Arquitetura da CNN para parcelamento . . . . .	70
4.3	Técnicas de treinamento de CNN 3D . . . . .	71
4.3.1	Treinamento usando <i>Patches</i> . . . . .	71
4.3.2	Abordagem multitarefas ou multi estrutura . . . . .	72
4.3.3	Função de custo . . . . .	72
4.4	Pós-processamento . . . . .	72
4.5	Inferência e parcelamento . . . . .	73
4.6	Incertezas de um modelo de CNN . . . . .	73
4.6.1	Incerteza epistêmica . . . . .	74
4.6.2	Incerteza estatística . . . . .	75
4.7	Resultados . . . . .	75
<b>5</b>	<b>Discussão</b> . . . . .	<b>82</b>
<b>6</b>	<b>Conclusão</b> . . . . .	<b>86</b>
6.1	Trabalhos futuros . . . . .	87
6.2	Trabalhos apresentados, publicações e participação em eventos . . . . .	89
	<b>Referências</b> . . . . .	<b>91</b>
	<b>APÊNDICE A Protocolo de segmentação manual</b> . . . . .	<b>104</b>

# 1 Introdução

## 1.1 Segmentação do Tálamo

O tálamo é uma estrutura cerebral subcortical localizada na região central do cérebro, especificamente na parte superior do tronco cerebral (Fig. 1.1), abaixo do ventrículo lateral e do fórnix (MACHADO, 1974). Ele desempenha um papel crucial no processamento e na transmissão de informações sensoriais para o córtex cerebral, que é a camada externa do cérebro responsável por funções cognitivas complexas. Adicionalmente, o tálamo desempenha um papel vital na regulação do sono, da consciência e do estado de alerta, além de estar envolvido em funções motoras, controle da memória e processamento emocional (SHERMAN; GUILLERY, 2001).

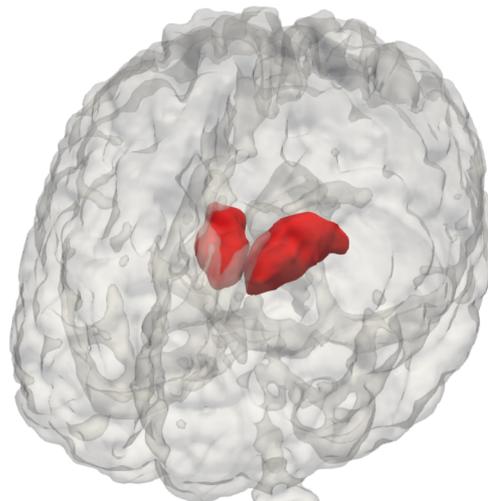


Figura 1.1 – Tálamo e sua posição no cérebro: renderização 3D da superfície cortical a de uma imagem de ressonância magnética, com os tálamos destacados (em vermelho).

Além de relação com diversas funções vitais, anomalias no tálamo estão relacionado à uma série de condições de saúde, como tremor essencial (AKRAM *et al.*, 2018). De fato, alguns dos possíveis tratamentos para tremor essencial são procedimentos cirúrgicos envolvendo o tálamo (BERK *et al.*, 2004; HARIZ *et al.*, 2008), onde a principal região alvo no interior do tálamo é o núcleo ViM (ventral intermédio) (MIDDLEBROOKS *et al.*, 2018b; NAJDENOVSKA *et al.*, 2019).

Exemplos desse tipo de intervenção são a estimulação cerebral profunda (DBS, do inglês *Deep Brain Stimulation*) (BRICE; MCLELLAN, 1980; CASTRO *et al.*, 2006; XIAO *et al.*, 2012; BERNEY *et al.*, 2002) e a talamotomia (SPEAKMAN, 1963; SIEG-

FRIED; REA, 1988; TASKER *et al.*, 1983; TASKER, 1990). No procedimento de DBS, um eletrodo é inserido no cérebro, atingindo o tálamo, e estímulos elétricos controlados por um dispositivo externo são administrados em sua extremidade, em uma região precisa do tálamo. Diferente do DBS, a talamotomia é irreversível, já que ao invés de estímulos controlados, o tecido do tálamo é cauterizado em um ponto específico.

Em ambos os procedimentos, para que haja sucesso na intervenção e principalmente para que não haja prejuízo aos pacientes, existe a necessidade do uso de imagens médicas, particularmente Imagens de Ressonância Magnética (MRI, do inglês *Magnetic Resonance Imaging*), e uma definição exata da região alvo do tálamo (MIDDLEBROOKS *et al.*, 2018a; NAJDENOVSKA *et al.*, 2019; MALAGA *et al.*, 2023). Essa região bem determinada é definida por critérios funcionais e estruturais do tálamo, sendo cada uma dessas regiões um dos núcleos que compõem o tálamo.

Não somente para planejamento cirúrgico e apontamento de alvos para intervenção, uma definição exata das bordas e da posição dos núcleos dos tálamos são de grande importância para estudos que relacionam condições patológicas à essa estrutura. Nesses contextos, surge a necessidade de métodos para segmentação e parcelamento do tálamo de forma precisa e robusta (FADDA *et al.*, 2019; BURGGRAAFF *et al.*, 2021). Entende-se por segmentação a delimitação da estrutura de interesse, no caso o tálamo, destacando-a do restante do cérebro. Já o termo parcelamento é utilizado neste trabalho como a divisão do tálamo em partes menores, no caso, em seus núcleos.

O primeiro e mais fundamental dos métodos de segmentação utilizados é a segmentação manual. Porém, devido ao tempo despendido para essa tarefa e a necessidade de especialistas em anatomia específica, é basicamente impraticável o uso dessa técnica em estudos populacionais, que normalmente envolvem um grande número de sujeitos.

Sendo a segmentação manual já impraticável nas condições mencionadas, o parcelamento do tálamo também é uma atividade inviável quando feito de forma manual. Além do grande nível de detalhamento adicionado para a definição das bordas dos núcleos, aumentando exponencialmente o trabalho, as imagens de ressonância normalmente adquiridas no ambiente clínico e em grande parte dos centros de pesquisa não apresentam contrastes necessários para a identificação das fronteiras entre os núcleos.

Devido às limitações dos métodos manuais e à impossibilidade de utilizá-los para conduzir estudos em larga escala (DEELEY *et al.*, 2011), é necessário o emprego de métodos automáticos que desempenham as tarefas de segmentação e parcelamento do tálamo. Esses métodos automáticos podem ser classificados em dois grupos principais: baseados em atlas e baseados em dados.

Os métodos baseados em atlas são aqueles que ajustam um mapa pré definido à

imagem do sujeito (BLEZEK; MILLER, 2006; WANG *et al.*, 2012). Esse é uma técnica de segmentação que delimita as bordas das estruturas de modo puramente geométrico, não considerando o valor de cada um dos *voxels* internos. Normalmente essa técnica é aplicada via registro da imagem do sujeito a um atlas definido em um espaço padrão, seguido por uma transformação nas máscaras para transportá-las para o espaço do sujeito.

Já os métodos baseados em dados são aqueles que classificam os *voxels* como pertencentes ou não às estruturas de acordo com seu valor individual ou em relação aos *voxels* vizinhos. Essa classe de métodos é mais abrangente que a anterior já que pode ser aplicada de inúmeras formas. Alguns exemplos de métodos dessa classe são os baseados em limiarização, baseados na transformada de *watershed* (RITTNER *et al.*, 2010), baseados em filtros e os métodos modernos baseados em aprendizado profundo (DL, do inglês *Deep Learning*).

## 1.2 Trabalhos relacionados

Devido à necessidade de aplicações como DBS e pesquisas que buscam estudar doenças relacionadas ao tálamo, como o tremor essencial, métodos de segmentação e de parcelamento do tálamo em MRI têm sido propostos nas últimas décadas.

### 1.2.1 Métodos de segmentação do tálamo

Existem diversos métodos e algumas ferramentas disponíveis que são capazes de desempenhar a tarefa de segmentação do tálamo em MRI (Tab. 1.1). Vamos considerar 3 grupos distintos de trabalhos que fazem a segmentação do tálamo, independente do tipo de imagem de entrada do algoritmo: o primeiro é o grupo dos métodos por registro de atlas, o segundo é o grupo dos que aplicam as técnicas mais modernas de DL, especificamente os métodos com redes convolucionais (CNN, do inglês *Convolutional Neural Network*), e por fim, o grupo dos demais métodos que usam técnicas de classificação dos *voxels* por suas características, seja valor absoluto ou relativos aos vizinhos.

Os métodos do primeiro grupo, os baseados em atlas, são os mais consolidados. Esses métodos apresentam bons resultados com DICEs (medida de concordância entre segmentações) variando de 0,74 a 0,88, como reportado em Hannoun *et al.* (2019) que compara as principais ferramentas que usam esses métodos. Apesar dos bons resultados, esses métodos são os que mais sofrem com tempo de execução já que precisam de pré-processamentos pesados para regularização das imagens e de algoritmos robustos para registrar o atlas às imagens a serem segmentadas. Outra desvantagem desses métodos é a dificuldade ou inviabilidade de se trabalhar com dados multi contraste, como MRI de múltiplas sequências de aquisição. De fato, praticamente todos os métodos que fazem

registro de atlas são exclusivos ou especializados somente em imagens T1w (T1w, do inglês *T1-weighted*) (Seç. 2.1.1), o que limita as possibilidades de aproveitamento de contrastes diferentes, como os gerados por imagens de difusão (Seç. 2.1.2).

Algumas das ferramentas mais populares que utilizam métodos de registro em atlas são o FreeSurfer (FISCHL *et al.*, 2002), volBrain (MANJÓN; COUPÉ, 2016) e o FSL (JENKINSON *et al.*, 2012), que fazem a segmentação de múltiplas estruturas cerebrais, incluindo o tálamo.

Apesar de segmentarem o tálamo, existem duas considerações importantes relacionada às tarefas desempenhadas por essas e outras ferramentas e métodos baseados em registro de atlas. O primeiro ponto é que não são necessariamente especializados, portanto, não têm o objetivo específico de segmentar da melhor maneira possível cada uma das estruturas individualmente. Outra consideração é que todas essas ferramentas requerem muito poder computacional e tempo de processamento para concluir a tarefa de segmentação. Como exemplo, o FreeSurfer e o FSL, apesar de suportarem processamento paralelo, chegam a levar horas para fazer a segmentação de um único sujeito. Similarmente, o volBrain entrega os resultados algumas horas após ser feito o *upload* das imagens em sua plataforma.

Dos trabalhos proponentes de métodos baseados em atlas, o que apresenta melhor resultado é o MALF (*Multi-Atlas likelihood fusion*) (TANG *et al.*, 2015), reportando DICE médio de 0,96 para um dos conjuntos de dados que é utilizado. No entanto, existem alguns pontos relevantes a serem considerados neste trabalho. O primeiro ponto é que ele envolve 30 sujeitos no método que segmenta o tálamo, sendo que 29 são empregados na criação do atlas e a avaliação é feita em um único sujeito por vez, de forma que o resultado reportado é a média desse processo sendo executado para cada um dos sujeitos. Outra consideração é que nem o atlas resultante, nem o protocolo de segmentação manual e nem algum tipo de implementação são disponibilizados, de forma que não é possível reproduzir os experimentos, seja para avaliação do método ou para o emprego como uma ferramenta.

O segundo grupo é o dos métodos que aplicam CNNs. Esses métodos apresentam bons resultados tanto no que se refere a qualidade da segmentação quanto ao tempo de execução dos algoritmos de segmentação.

Um dos exemplos mais relevantes dos trabalhos utilizando CNN é a QuickNAT (ROY *et al.*, 2018a) que reporta uma medida de similaridade DICE de 0,87 e é capaz de executar a segmentação em apenas 20 segundos, dado que a imagem de entrada seja T1w e já esteja padronizada. Outra vantagem importante deste trabalho é que os códigos e os modelos treinados são disponibilizados, de forma que outros pesquisadores possam empregar o método, apesar de não ser tão amigável ao uso.

Outro trabalho relevante que emprega CNN na tarefa de segmentação do tálamo é Opfer *et al.* (2023). Este trabalho usa uma variação 3D da U-Net e a treina com um conjunto de dados privado, tomando a segmentação feita pelo FSL-FIRST como padrão-ouro. Para a avaliação do desempenho do método, imagens de 18 sujeitos manualmente segmentadas são empregadas, atingindo uma DICE médio de 0,86 e ficando tecnicamente empatado ou levemente inferior a outros métodos, como o FSL.

Os trabalhos de Li *et al.* (2021) e Dolz *et al.* (2018a) também utilizam CNNs e imagens T1w para segmentar o tálamo. Inclusive, Li *et al.* (2021) aplica uma abordagem interessante onde a imagem é primeiramente segmentada em uma resolução menor, salvando recursos computacionais, e depois é aplicada a resolução total para atingir o melhor resultado. Esse trabalhos reportam DICEs acima de 0,9, porém são utilizadas somente máscaras geradas por outros métodos automáticos e os resultados são avaliados em no máximo 3 sujeitos. Portanto, os resultados não são necessariamente representativos.

O trabalho de Wang *et al.* (2021), assim como os anteriores, utiliza a U-Net para segmentar o tálamo, porem com um diferencial de usar imagens MPRAGE (III; BROOKEMAN, 1991; BRANT-ZAWADZKI *et al.*, 1992; PLANCHE *et al.*, 2020) e mapas de difusão, especificamente FA e MD. O DICE atingido é de 0,84, porém não é especificado o tipo de anotação dos dados.

O último grupo de métodos é o dos que usam classificação dos *voxels* (LIU *et al.*, 2016; GLAISTER *et al.*, 2017). Esses são os menos empregados atualmente, principalmente por terem desempenhos inferiores quando comparados com os métodos de redes convolucionais. Enquanto Liu *et al.* (2016) reporta um DICE de 0,81, Glaister *et al.* (2017) reporta um DICE de 0,87. Porém, neste último, foram utilizados apenas 7 sujeitos, sendo que para cada sujeito, os outros 6 eram utilizados para o desenvolvimento do modelo, o que fragiliza a confiabilidade do resultado. O ponto forte desses métodos é que eles podem ser facilmente aplicados simultaneamente a múltiplas sequências de MRI. De fato, isso é necessário já que os valores dos *voxels* precisam de mais especificidade por não contarem com tanta informação de vizinhança e de textura como os métodos de CNN.

Podemos notar que, independentemente do método empregado, seja por registro de atlas, CNN ou classificação de *voxels*, a vasta maioria dos trabalhos de segmentação do tálamo faz isso utilizando somente imagens T1w. Apenas alguns poucos trabalho empregam imagens de difusão (LIU *et al.*, 2016; WANG *et al.*, 2021; GLAISTER *et al.*, 2017), que apresentam melhores contrastes para o tálamo. Vale ressaltar que estes ainda consideram somente a informação não direcional dos dados, como mapas escalares derivados dos autovalores (FA, MD, AD e RD) (Seç. 2.1.2.1). Sendo assim, uma deficiência dos métodos atuais de segmentação do tálamo é o não aproveitamento da informação de direcionalidade das fibras para uma definição mais acurada dos limites da estrutura.

Tabela 1.1 – Métodos de segmentação automatizada do tálamo encontrados na literatura

Trabalho	Tipo de dado	Base de dados	Método	Dice	n imagens treino/val./teste	resolução [mm]	Padrão-Ouro
(JENKINSON <i>et al.</i> , 2012)	T1w	Privado	Atlas (FSL-Anat)	0,859†	126	0,937x0,937x1	manual
(JENKINSON <i>et al.</i> , 2012)	T1w	Privado	Atlas (FSL FIRST)	0,86†	126	0,937x0,937x1	manual
(FISCHL <i>et al.</i> , 2002)	T1w	Privado	Atlas (FreeSurfer)	0,838†	126	0,937x0,937x1	manual
(MORI <i>et al.</i> , 2016)	T1w	Privado	Atlas (MRICloud)	0,736†	126	0,937x0,937x1	manual
(MANJÓN; COUPÉ, 2016)	T1w	Privado	Atlas (volBrain)	0,882†	126	0,937x0,937x1	manual
(POHL <i>et al.</i> , 2006)	T1w	Privado	Atlas	0,88	22	0,9375x0,9375x1,5	-
(TANG <i>et al.</i> , 2015)	T1w	Pediátrico (privado) MICCAI Challenge	Multi-Atlas (MALF)	0,96	30 ( <i>cross validation</i> 29/1)	1x1x1	manual
(ZHANG <i>et al.</i> , 2017)	T1w	(2013)	Multi-Atlas (MALSF)	0,92	20/-/15	1x1x1	manual
(ROY <i>et al.</i> , 2018a)	T1w	IBSR	CNN (QuickNAT)	0,87	18	0,8x0,8x1,5	-
(OPFER <i>et al.</i> , 2023)	T1w	Privado	CNN 3D	0,86	1975/-/18	-	FSL-FIRST manual
(LI <i>et al.</i> , 2021)	T1w	IBSR PREDICT-HD	CNN (U-Nets em cascata)	0,918 0,97	18/-/3 20/-/2	0,8x0,8x1,5-1x1x1,5	automático inspecionado
(DOLZ <i>et al.</i> , 2018a)	T1w	ABIDE IBSR	CNN 3D	0,91 0,83	12/3/3	0,8x0,8x1,5	-
(WANG <i>et al.</i> , 2021)	T1 MPRAGE e DTI	Privado	CNN (U-Net)	0,84	20/2/10	-	-
(LIU <i>et al.</i> , 2016)	T1w e DTI	Privado	Classificador (Dicionário)	0,81	23	-	manual
(GLAISTER <i>et al.</i> , 2017)	DTI (FA e MD)	Privado	Classificador ( <i>Random Forest</i> )	0,87	7 ( <i>cross validation</i> 6/1)	2,2x2,2x2,2	manual

†Resultados reportados por (HANNOUN *et al.*, 2019)

### 1.2.2 Métodos de parcelamento do tálamo

O parcelamento do tálamo é um processo no qual o objetivo é delinear os núcleos do tálamo, ou seja, o parcelamento é a segmentação das partes que constituem o tálamo. Isso é ainda mais desafiador que a segmentação da estrutura como um todo já que essas estruturas são muito menores, dificultando a obtenção de máscaras padrão-ouro, sendo consideravelmente mais influenciáveis por volume parcial e apresentando contrastes muito inferiores entre cada uma das regiões.

Devido o maior grau de complexidade para encontrar os núcleos do tálamo, os métodos que têm esse objetivo normalmente precisam de imagens com mais especificidade para cada *voxel*, como imagens de maior resolução espacial, maior contraste (como imagens de máquinas de 7 tesla) (IGLESIAS *et al.*, 2018; SU *et al.*, 2019) e imagens de difusão, que trazem informações mais detalhadas de micro estrutura e das fibras dos tecidos (LIU *et al.*, 2015; SARANATHAN *et al.*, 2021; NAJDENOVSKA *et al.*, 2018).

Diferentemente do problema de segmentação do tálamo, não existem muitos métodos capazes de parcelar o tálamo (Tab. 1.2). Apenas algumas poucas ferramentas conseguem fazer o parcelamento do tálamo, como o FreeSurfer (FISCHL *et al.*, 2002) e o volBrain (MANJÓN; COUPÉ, 2016), que possuem módulos específicos para essa tarefa. Normalmente esses métodos quem tem o objetivo de parcelamento do tálamo se baseiam em registro de atlas, que é o caso do FreeSurfer, ou clusterização de *voxels*.

A ferramenta FreeSurfer faz a segmentação do tálamo com o método proposto por Iglesias *et al.* (2018). Esse método aplica o registro de um atlas proposto por ele mesmo, criado a partir de estudos histológicos de 6 sujeitos e composto por 26 núcleos em cada hemisfério do tálamo. Esse método é possivelmente o mais aplicado atualmente para a tarefa de parcelamento do tálamo devido sua integração com o FreeSurfer, porém, ele é dependente da segmentação prévia do tálamo e de imagens e parâmetros intermediários do processo feito pelo FreeSurfer. Portanto, o tempo de execução deste método é de algumas horas.

Outro método relevante, denominado THOMAS (SU *et al.*, 2019), parcela o tálamo em 12 núcleos distintos. Este método também é baseado em registro de atlas e atinge resultados de DICE entre 0,7 e 0,85 em imagens WMn-MPRAGE 7T. O artigo relata tempos de 10 minutos para completar a segmentação dos 12 núcleos do tálamo de cada sujeito. Algumas das desvantagens desse método é a necessidade de uma sequência específica de aquisição e a dependência de *scanner* com grande intensidade de campo para conseguir atingir resultado ótimo.

Além dos métodos que são baseados em registro de atlas, temos uma classe dos métodos que aplicam clusterização (NAJDENOVSKA *et al.*, 2018; BATTISTELLA

*et al.*, 2017; ZIYAN *et al.*, 2006). Em resumo, o que os métodos de clusterização fazem é agrupar os *voxels* que possuem propriedades similares, portanto, tecidos biológicos com o mesmo tipo de microestrutura.

Como as imagens de ressonância ponderadas por difusão são bons descritores da micro estrutura, os métodos baseados em clusterização empregam exatamente esse tipo de imagem. Importante ressaltar que parte dos métodos que fazem clusterização com imagens de difusão aplica em modelos que consideram informações de direção das fibras do cérebro, como é o modelo ODF (do inglês *orientation distribution function*). Apesar dessas imagens realmente apresentarem informações relevantes, principalmente para gerar contraste no interior do tálamo, as aquisições de dMRI que permitem a reconstrução desses modelos complexos é consideravelmente cara do ponto de vista de tempo de aquisição e recursos computacionais.

Apesar de relativamente raros nesse contexto, outro grupo que deve ser considerado é o dos métodos que aplicam redes convolucionais para a tarefa de parcelamento. Um dos principais obstáculos no desenvolvimento desses métodos é a ausência de conjuntos de dados com anotações padrão-ouro para treinamento das CNN visto que esses métodos requerem grandes quantidades de dados para a fase de treinamento.

Um exemplo de método de parcelamento do tálamo baseado em CNN foi proposto por Majdi *et al.* (2020) que utiliza um conjunto de dados privado de 52 imagens WMn-MPRAGE 7T, que é um dos melhores tipos de imagem não ponderada por difusão para parcelamento do tálamo (PLANCHE *et al.*, 2020), e atinge valores de DICE entre 0,55 e 0,84, dependendo do núcleo.

Outro trabalho que emprega CNN para o parcelamento do tálamo é o de Umapathy *et al.* (2022). Esse trabalho sintetiza imagens WMn-MPRAGE a partir de imagens MPRAGE e faz o parcelamento nas imagens sintetizadas. O método atinge valores de DICE variando de 0,7 até 0,84, com resultado sendo avaliados em 2 sujeitos, considerando o padrão-ouro a segmentação feitas pelo método THOMAS.

Por fim, temos o método proposto por Ruiz-Perez *et al.* (2024), atualmente sendo utilizado na ferramenta volBrain. O diferencial desse método é empregar simultaneamente imagens ponderada em T1 e imagens WMn-MPRAGE sintetizadas, que juntamente com uma quantidade massiva de dados de treinamento, conseguem valores de DICE de 0,87 até 0,97. Entretanto, os resultados foram avaliados em um grupo de 20 sujeitos, sendo que o padrão-ouro para a avaliação é proveniente de um método automático de registro de atlas, THOMAS, executado nas própria imagens WMn-MPRAGE sintetizadas. Isso significa que, mesmo apresentando ótimos resultados de reprodução do atlas, o método ainda está limitado ao desempenho das máscaras automáticas, que por

sua vez são limitadas ao desempenho do método THOMAS.

Existem ainda métodos que parcelam o tálamo via tractografia (SEMEDO *et al.*, 2017; SEMEDO *et al.*, 2018; MIDDLEBROOKS *et al.*, 2018a), que é uma técnica onde se estima as quantidades e posições dos feixes de fibras dos cérebro. Porém, esses métodos não serão considerados aqui já que a tractografia possui uma série de limitações que restringe a aplicação para o problema de parcelamento de grande quantidades de dados ou para processamento rápido. De fato, a literatura apresenta uma quantidade ínfima de trabalhos de parcelamento empregando tractografia. Exemplo das limitações é a natureza não determinística de alguns dos métodos de tractografia e o tempo e o processamento necessário para reconstrução dos feixes de fibras.

Os problemas comuns para todos os métodos de parcelamento do tálamo são, principalmente, o baixo contraste entre os núcleos e a falta de anotações de padrão-ouro. Para tentar contornar o primeiro problema, de contraste, os métodos empregam imagens de ressonância com sequências específicas, como WMn-MPRAGE e imagens de difusão. Já o problema de falta de anotação é abordado de duas formas: anotação manual em conjuntos de dados pequenos e avaliação dos modelos em conjuntos de dados com múltiplas aquisições para o mesmo sujeito. Esse artifício, chamado de teste-reteste ou *scan-rescan*, avalia principalmente a repetibilidade do método. Em resumo, o método de parcelamento é executado para duas aquisições distintas de um mesmo sujeito, os parcelamentos são registrados entre si (normalmente a matriz de registro computada com as imagens de entrada do método) e a métrica de desempenho é calculada entre os parcelamentos registrados. Assumindo que não há diferença anatômica entre as aquisições, essa medida irá indicar se o método consegue reproduzir fielmente o parcelamento feito em imagens de um mesmo sujeito que foram adquiridas momentos distintos.

Tabela 1.2 – Métodos de parcelamento do tálamo

Trabalho	Tipo de dado	Base de dados	Método	N núcleos	Dice	N imagens teste	Resolução [mm]	Avaliação (padrão-ouro)
(IGLESIAS <i>et al.</i> , 2018)	ME-MPRAGE	Privado	Registro de atlas	26	0,65 - 0,85	31	1x1x1	test-retest
(SU <i>et al.</i> , 2019)	WMn MPRAGE 7T	Privado	THOMAS (Registro de atlas)	12	0,7 - 0,85	5	1x1x1	manual
(DATTA <i>et al.</i> , 2021)	MP2RAGE 7T	Privado	THOMAS modificado	12	0,7 - 0,89	13	1x1x1	WMn-MPRAGE THOMAS
(NAJDENOVSKA <i>et al.</i> , 2018)	DW-MRI (ODF)	HCP	Clusterização (k-means)	7	0,5 - 0,72	34	2x2x2,5	atlas Morel
(BATTISTELLA <i>et al.</i> , 2017)	DW-MRI (ODF)	Privado	Clusterização (k-means)	7	0,83 - 0,93	8	2x2x2	scan-rescan
(ZIYAN <i>et al.</i> , 2006)	DTI	Privado	Clusterização	7	0,6 - 0,83	10	2x2x2	manual
(TREGIDGO <i>et al.</i> , 2023)	DTI	HCP	Inferência bayesiana	26 (5 grupos)	0,84 - 0,94 0,7 - 0,89	21 10	1,25x1,25x1,25	test-retest manual
(MAJDI <i>et al.</i> , 2020)	WMn-MPRAGE 7T	Privado	CNNs em cascata	11 (4 grupos)	0,55 - 0,84	52 ( <i>cross validation</i> 80/20)	1x1x1	manual
(UMAPATHY <i>et al.</i> , 2022)	WMn-MPRAGE sintetizado	Privado	CNNs em cascata	11	0,7 - 0,84	33/-/2	-	WMn-MPRAGE THOMAS
(RUIZ-PEREZ <i>et al.</i> , 2024)	T1w + WMn-MP. sintetizado	Vários públicos	CNN	11	0,87 - 0,97	4000+/-/20	-	WMn-MPRAGE THOMAS

## 1.3 Motivação

A segmentação e o parcelamento do tálamo são extremamente úteis em quaisquer procedimentos, científicos ou clínicos, que têm como objeto essa estrutura. Em casos específicos como em tratamentos intervencionistas de doenças tais quais tremor essencial, a localização exata da região alvo é indispensável para o sucesso da intervenção.

Idealmente, a localização das regiões deve ser feita com segmentações em imagens de grande resolução, multimodalidade e múltiplos contrastes, garantindo máxima especificidade de cada tecido. Porém existem limitações práticas que inviabilizam a aplicação de métodos ideais.

No ambiente clínico, especialmente em hospitais com recursos limitados, não há disponibilidade de imagens de ressonância de alta resolução, licença de software para processamento imediato das imagens e nem equipe dedicada. Em pesquisas, onde grandes quantidades de imagens são consideradas, existem limitações de armazenamento, já que uma aquisição de MRI de um sujeito contendo imagens T1 e difusão em formato DICOM tem a ordem de 2GB de dados, e processamento de todos os dados, que quando tratados com os algoritmos convencionais chegam a consumir mais de 5 horas de processamento para cada sujeito.

Particularmente no caso de imagens de difusão, as restrições são agravadas já que o tempo de aquisição é estendido em relação às imagens estruturais mais usuais, principalmente nos casos de aquisições de grande resolução espacial e direcional. Apesar do tempo extra de aquisição, as imagens de difusão são capazes de medir diferenças microestruturais que outras sequências de MRI não conseguem, portanto são umas das poucas opções interessantes para a tarefa de segmentação do tálamo, mesmo que sendo coletadas em menores resoluções.

Com a necessidade de segmentação rápida e robusta do tálamo e seus núcleos, atendendo tanto requisitos clínicos quanto de pesquisa, surge a carência de métodos eficientes que consigam segmentar precisamente essas estruturas.

## 1.4 Objetivo do projeto de pesquisa

O principal objetivo desse trabalho é o desenvolvimento de métodos de segmentação e parcelamento do tálamo utilizando imagens de ressonância magnética e redes convolucionais atendendo os seguintes requisitos: realizar a segmentação a partir de imagens T1w e de difusão; atingir níveis de qualidade estado da arte de métodos automáticos; apresentar tempo de execução da ordem de segundos por imagem processada; possuir códigos abertos com fácil execução.

Os objetivos específicos são:

**Método com CNN para segmentação e parcelamento do tálamo:** desenvolvimento de arquitetura, treinamento e avaliação de redes convolucionais para desempenharem as tarefas de segmentação e parcelamento do tálamo.

**Harmonização dos dados:** desenvolver métodos de pré-processamento para harmonizar os múltiplos tipos de imagens de entrada, sendo estas as imagens T1w, mapas escalares de difusão (Sec 2.1.2.1), mapas de gradiente morfológico tensorial e mapas de divergente tensorial. A harmonização resolverá problemas de diferenças de domínio, tanto intensidade quando de escala e posição.

**Testes com combinações dos mapas:** avaliar e comparar o desempenho das CNNs com diferentes combinações de imagens de entradas, T1w e mapas de difusão.

**Estudar as limitações dos dados:** estudar as limitações das imagens T1w e imagens de difusão para segmentação com CNN. Estudar também as limitações impostas pelas máscaras de origens automáticas e manuais.

**Ciência aberta:** disponibilizar conjunto de dados com imagens T1w e mapas difusão já prontos para treinamento de CNNs e disponibilizar os códigos para treinamento e predição com os modelos de redes neurais treinadas para as tarefas de segmentação e parcelamento.

## 1.5 Principais desafios

**Baixo contrastes da estrutura:** Um dos principais desafios para a segmentação do tálamo é a falta de contraste da estrutura, seja em imagens T1w ou em outros tipos de imagem. A falta de contraste nessa estrutura é um desafio tanto para os métodos automáticos quanto para profissionais, já que as fronteiras de seus núcleos são praticamente invisíveis.

**Tamanho das estruturas:** Além do baixo contraste, existe um agravante que é o tamanho das estruturas. O tálamo é uma estrutura relativamente pequena quando comparada com o cérebro, medindo aproximadamente  $5000mm^3$  frente à  $1300cm^3$  do cérebro, o que representa um valor inferior a 5% do volume cerebral. Quando consideramos os núcleos do tálamo, que podem ser menores que  $10mm^3$ , fica evidente que a inclusão ou não de alguns *voxels* são relevantes para a qualidade da segmentação.

**Volume parcial:** Outro desafio encontrado na tarefa de segmentação do tálamo é o volume parcial. Esse problema é causado principalmente nas bordas das estruturas, onde um único *voxel* acaba sendo representante de propriedades de múltiplos tecidos já que em seu espaço físico couberam duas ou mais estruturas. Normalmente, em estruturas maiores, o problema do volume parcial pode ser negligenciado já que os tamanhos dos *voxels* são pequenos em relação a estrutura. Porém, como mencionado, a inclusão ou não de um único *voxel* pode afetar significativamente a segmentação, principalmente dos núcleos do tálamo.

**Resoluções das imagens:** Outro desafio é a resoluções das imagens, que é determinada pela quantidade de *voxels* em cada uma das três direções, tamanho dos *voxels* e tamanho do campo de visão. Por questões de limitações de intensidade de campo das máquinas quanto por tempo de aquisição, as imagens de ressonância são, normalmente, limitadas em resolução. Contextualizando, tanto em ambientes clínicos quanto de pesquisa, o tempo de aquisição em sujeitos vivos é um fator crítico, limitando a quantidade de *voxels* coletados. No caso específico de dMRI, as imagens de difusão normalmente possuem resolução inferiores àquelas de T1 ou T2. Enquanto uma imagem T2 tipicamente possui resolução de  $256 \times 256 \times 256$ , as imagens de difusão têm resolução de  $128 \times 128 \times 70$ , para um mesmo campo de visão. Isto influencia na qualidade da segmentação pois os efeitos de volume parcial (TOHKA, 2014; LEEMPUT *et al.*, 2003) são agravados, principalmente em estruturas menores.

**Dimensionalidade dos dados:** Apesar de normalmente possuírem resolução espacial inferior, as imagens de difusão possuem alta dimensionalidade, o que também é algo desafiador, tanto para métodos convencionais quanto para as CNNs. Esta dimensionalidade vem das diversas aquisições para diferentes gradientes de campo, tanto em direções quanto em intensidades. O problema causado neste caso é a grande quantidade de dados a ser tratada e a interpretação do significado desses dados, exigindo grande poder computacional, ou a necessidade de simplificação dos dados por ajuste em modelos pré-determinados de difusão e que façam sentido para a aplicação específica.

**Aplicação de dados multi contraste:** No contexto de segmentação em dados multimodalidade, existem diversos desafios a serem superados para que se possa aproveitar do potencial dos dados de diferentes origens, por exemplo, a compatibilização dos dados multimodais entre si e com os métodos atuais de CNNs.

Quando utilizando imagens de difusão, um destes desafios é a harmonização espacial e radial dos dados de difusão com imagens de outras sequências. O principal motivo que torna esta etapa mandatória é que, via de regra, as anotações manuais dos

especialistas são geradas em cima de imagens como T1 ou T2, que têm uma resolução e referências espaciais diferentes, e as anotações precisam estar no mesmo espaço que as imagens de difusão para que a CNN funcione corretamente. Além disso, em certos casos, as imagens de outras sequências são usadas simultaneamente, de modo que não só as anotações, mas as entradas distintas também precisem estar no mesmo espaço.

Existe também um desafio em relação ao processamento mínimo necessário para a entrada dos dados multimodalidades nas redes neurais convolucionais. Enquanto que na maioria dos casos de aplicação de CNN utilizando imagens escalares se pode aplicar os dados sem qualquer tipo de processamento, ou com apenas uma normalização simples, quando se trata de imagens de difusão, a mesma estratégia tende a não funcionar. De fato, quase todos os trabalhos com imagens de difusão, mesmo aqueles que não se baseiam em redes neurais, utilizam ajustes dos dados em modelos pré-determinados para reduzir a complexidade e extrair as informações úteis e necessárias para cada aplicação.

**Definição dos mapas adequados para o problema:** Considerando a entrada das CNNs, como as imagens T1w e os mapas computados a partir das imagens de difusão, ainda existe o desafio da seleção de quais mapas são efetivos para resolver os problemas de segmentação do tálamo, dada a grande quantidade de combinações possíveis. Além da seleção destes mapas, é necessária a definição da técnica de compatibilização e fusão destes para a entrada na rede.

**Anotação dos dados:** Por fim, existem as dificuldades relacionadas a quantidade de dados anotados disponíveis. Conjunto de dados públicos, sejam com difusão ou apenas T1w, anotados manualmente por especialistas, configurando máscaras padrão-ouro do tálamo, são basicamente inexistentes. Portanto, dois desafios relevantes aparecem: criação de máscaras para os conjuntos de dados empregados e garantia da qualidade das máscaras criadas.

## 1.6 Principais contribuições

- Método estado da arte para segmentação do tálamo usando imagens estruturais de ressonância magnética: T1w e mapas derivados de imagens de tensores de difusão.
- Método para parcelamento do tálamo usando imagens estruturais de ressonância magnética: T1w e mapas derivados de imagens de tensores de difusão.
- Protocolo de segmentação manual do tálamo com auxílio de imagens de difusão.

- Conjunto de dados público com máscaras manuais e automáticas.
- Desenvolvimento de método de adequação e harmonização de mapas de difusão para uso em CNNs.

## 1.7 Organização da tese

O próximo capítulo (Cap. 2) apresenta os dados que foram utilizados durante o desenvolvimento deste trabalho. Além de apresentar os dados, introduz alguns conceitos básicos que são fundamentais para o entendimento do trabalho e descreve os pré-processamentos sofridos por esses dados. O capítulo 3 apresenta os experimentos e os resultados atingidos em relação a tarefa de segmentação do tálamo, enquanto o capítulo 4 foca na tarefa do parcelamento do tálamo. Por fim, o capítulo 5 apresenta discussões e o último capítulo (Cap. 6) apresenta as conclusões do trabalho.

## 2 Dados

Este trabalho emprega técnicas de inteligência artificial baseadas em aprendizado de máquina, portanto, precisa de dados para o treinamento, validação e avaliação dos resultados atingidos pelos métodos propostos. Os dados necessários para essas tarefas são imagens nas quais seja possível identificar o tálamo para a segmentação e o parcelamento. Por atender esses requisitos, as imagens de ressonância magnéticas são as melhores candidatas para essa tarefa.

### 2.1 Imagem de Ressonância Magnética

Imagem de ressonância magnética vem de uma técnica utilizada para visualização de tecidos biológicos internos de maneira não invasiva. O contraste nestas imagens é gerado pela interação entre a estrutura biológica com sequências específicas de pulsos de rádio frequência, que, por sua vez, expõe propriedades diferentes de cada tecido. Como cada tipo de imagem consegue mostrar características diferentes dos tecidos, para cada tipo de aplicação existem sequências de aquisição mais adequadas. Por exemplo, a sequência FLAIR (Fig. 2.1-c) é muito utilizada para detecção de lesões (HAJNAL *et al.*, 1992; ROVARIS *et al.*, 2000).

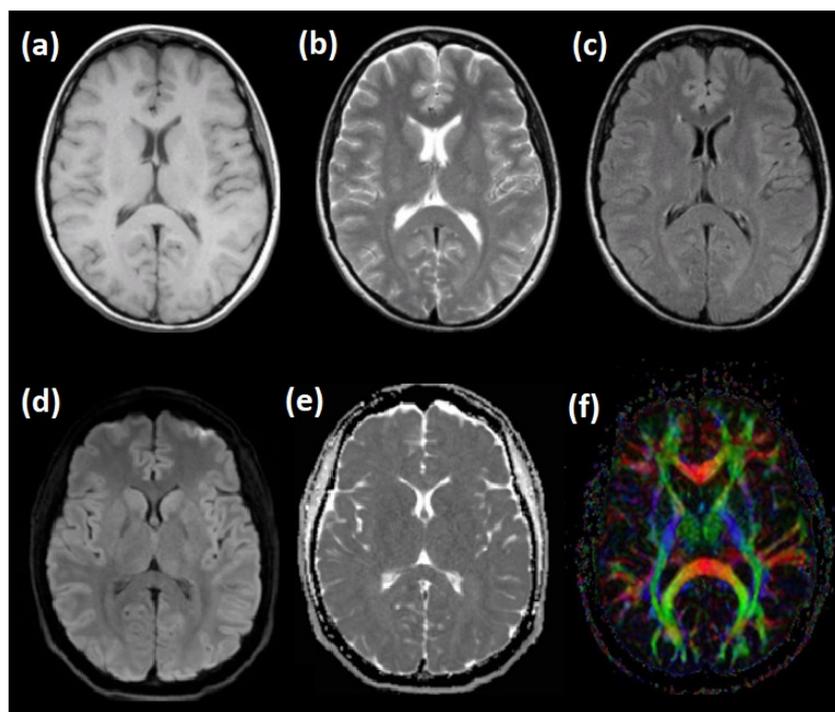


Figura 2.1 – Exemplos de imagens de RM: (a) T1; (b) T2; (c) FLAIR; (d) DWI; (e) ADC; (f) DTI.

Existem também aplicações que se beneficiam da informação complementar de diferentes sequências. Um exemplo é a segmentação de AVC (KAMNITSAS *et al.*, 2017; CHEN *et al.*, 2017; ZHAO *et al.*, 2019; RAINA *et al.*, 2019; BAGHER-EBADIAN *et al.*, 2011; ZHANG *et al.*, 2018; WINZECK *et al.*, 2019; SATHISH *et al.*, 2019), que apresenta melhores resultados quando utiliza simultaneamente imagens de diferentes sequências como imagens ponderadas em T1, em T2, e imagens ponderadas em difusão (Fig. 2.1-a,b,d,e).

Neste trabalho, foram aplicadas imagens provenientes de duas sequências de aquisição (Fig. 2.2): imagens ponderadas em T1 (T1w) (Sec. 2.1.1) e imagens ponderadas por difusão (dMRI) (Sec. 2.1.2). Essa escolha foi tomada por serem as imagens de ressonância populares presentes em bases dados públicas, que melhor conseguem representar o tecido do tálamo. Além de serem complementares por apresentarem contrastes diferentes em regiões distintas do tálamo, as imagens de difusão contêm informações de direcionalidade das fibras do tecido biológico.

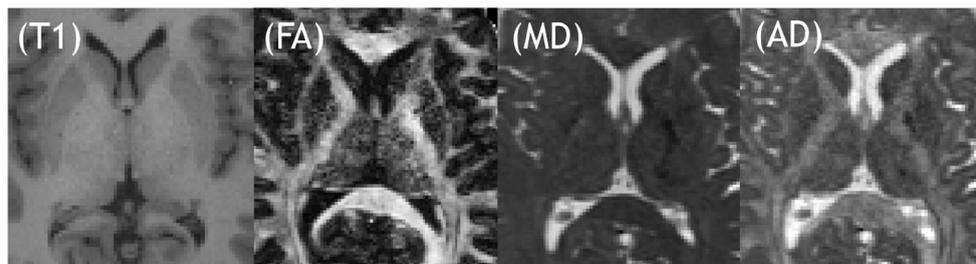


Figura 2.2 – Exemplos de imagens de RM: T1w e mapas escalares de DTI (FA, MD e AD)

### 2.1.1 Imagem de Ressonância Magnética Ponderada em T1 (T1w)

As imagens de ressonância ponderadas em T1 são possivelmente as imagens mais populares das imagens de ressonância. Esse tipo de imagem é altamente difundido no meio acadêmico e clínico, dentre outros motivos, por suas características de fácil interpretação pela similaridade com fatias anatômicas dos tecidos, por sua rápida aquisição e pela relativa boa resolução, mesmo em máquinas mais simples.

As imagens T1w são capazes de apresentar contrastes diferentes para diferentes tipos de tecido devido ao tempo de relaxação magnético T1 para diferentes composições de tecidos contido em cada *voxel*. De modo extremamente resumido, o contraste das imagens T1w é gerado pela diferença no tempo em que cada composição de tecido leva para retornar ao seu estado de precessão após serem atingidos por um pulso de rádio em sua frequência de ressonância (MITCHELL; COHEN, 2004; WESTBROOK, 2021).

Com a diferença em contraste de acordo com a composição dos tecidos em cada *voxel*, existe uma infinidade de aplicações que podem se beneficiar do emprego desse tipo

de imagem, como auxílio em diagnóstico clínico, segmentação de estruturas, identificação de alguns tipos de lesão, etc.

Apesar da grande variedade de aplicações para as imagens T1w, esse tipo de imagem apresenta algumas limitações que restringem seu uso em determinadas aplicações. Um exemplo relevante é a segmentação de determinadas estruturas onde o contraste é muito baixo ou inexistente, como nos núcleos do tálamo. Outro exemplo é a aplicação de imagens de ressonância para identificação de acidente vascular cerebral durante a fase aguda, onde as imagens T1w não apresentam variação enquanto as imagens ponderadas por difusão ou as imagens de perfusão conseguem apresentar diferenças de intensidade nas regiões afetadas (MAIER *et al.*, 2017; PINHEIRO *et al.*, 2019; DEMEESTERE *et al.*, 2020; CLÈRIGUES *et al.*, 2020).

### 2.1.2 Imagem de Difusão por Ressonância Magnética (dMRI)

Imagem de difusão por ressonância magnética (dMRI) se baseia na medida da difusão de água dentro dos tecidos para gerar contraste. Isto é alcançado por sequências de aquisição que, basicamente, fazem amostragens indiretas da difusão para diversas direções e para intensidades determinadas de gradiente de campo (BASSER *et al.*, 1994; BASSER; PIERPAOLI, 2011). Como o interior dos tecidos fibrosos cerebrais é um ambiente de difusividade restrita, a difusão para cada uma das direções é diferente, de modo que é possível inferir sobre a organização dos tecidos com a utilização das informações presentes em imagens de difusão (DOUEK *et al.*, 1991).

Após a aquisição dos dados brutos de difusão, são necessárias diversas etapas de correções para que o dado coletado seja coerente (IRFANOGLU *et al.*, 2019; WU *et al.*, 2008). Algumas dessas etapas são as correções de distorções geométricas, causadas pela própria magnetização do *scanner* pelas variações dos campos magnéticos, e o registro das imagens provenientes de cada uma das direções de aquisição. Essas correções podem ser feitas no ato da aquisição das imagens ou com processamentos posteriores, o que normalmente é vantajoso quando se é necessário ter domínio de todas as etapas do processamento.

Posteriormente às correções, as medidas de difusão são ajustadas a modelos físico-matemáticos para tentar explicar numericamente a organização ou para que se possa extrair informações microestruturais dos tecidos. Dentre os possíveis modelos matemáticos utilizados para ajustar as imagens de difusão, o mais popular é o tensorial, que é a base das Imagens de Tensores de Difusão (DTI) (2.1-f) (BASSER *et al.*, 1994; BASSER; PIERPAOLI, 2011). Está é uma modalidade construída a partir das Imagens Ponderadas por Difusão (DWI) (2.1-d) e que exprime as informações de cada ponto do espaço, ou

*voxel*, em um tensor que é a medida de difusão para cada uma das direções. Em regiões em que há uma grande organização das fibras, o tensor tende a se alongar no eixo paralelo ao eixo do feixe de fibras, onde há movimento praticamente irrestrito, os tensores tendem a ser esféricos e de maior norma, e assim por diante.

Apesar de existirem modelos mais complexos e representativos das imagens de difusão, esse trabalho se restringe ao emprego do modelo de tensores de difusão, principalmente pela maior facilidade de aquisição de imagens que atendam os requisitos mínimos para o ajuste nesse modelo. Com essa possibilidade de aquisição facilitada, mais conjuntos de dados são disponíveis para o desenvolvimento dos métodos e maior a chance dos métodos poderem ser utilizados em casos reais de aplicações clínicas e científicas.

Para se ajustar os dados de difusão ao modelo tensorial é necessária uma quantidade mínima de 6 aquisições de imagens ponderadas em difusão, cada uma em uma direção, e um medida com vetor direcional nulo (BASSER *et al.*, 1994; BASSER; PIERPAOLI, 2011). Esta quantidade mínima de direções é matematicamente definida pela quantidade de graus de liberdade de um tensor simétrico de ordem 3, utilizado como modelo físico-matemático para DTI. Porém, para diminuir as incertezas e influência de ruídos da aquisição, as aquisições dos dados brutos são normalmente em quantidades superiores a 30, podendo chegar a mais de 200 dependendo do modelo ao qual se pretende ajustar as medidas.

Como DTI é um mapa tensorial, que portanto não pode ser representado tão facilmente como as imagens escalares, é necessária alguma forma para exibir os dados de maneira inteligível. Uma dessas formas é um mapa de cores com a composição do grau de organização dos feixes de fibras com a informação de direcionalidade destes feixes (Fig. 2.1-f). Nesta representação, a intensidade assinalada para cada *voxel* é o valor de Anisotropia Fracional (FA), e a cor representa a direção dos feixes; vermelho - lateral; verde - anteroposterior; azul - craniocaudal.

DTI é apenas um dos modelos utilizados, e um dos mais antigos, em imagens de difusão. Outros exemplos de modelos que são construídos sobre DWI são HARDI (TUCH, 1999) (High Angular Resolution Diffusion Imaging), ODF (DESCOTEAUX *et al.*, 2006) (Orientation Density Function) e NODDI (ZHANG *et al.*, 2012) (Neurite Orientation Dispersion and Density Imaging). Esses modelos de maior complexidade, como HARDI, costumam apresentar mais informações da microestrutura, sendo úteis inclusive para melhor explicar o cruzamento de fibras em métodos de tractografia, por exemplo. Porém, como requerem grande resolução angular, e portanto, consideravelmente mais tempo de aquisição, acabam por serem impraticáveis no ambiente clínico.

### 2.1.2.1 Mapas provenientes de Imagens de Tensores de Difusão

Considerando as Imagens de Tensores de Difusão para segmentação de estruturas cerebrais, a forma mais comum de uso dessas imagens é através de uma simplificação dos tensores em medidas mais simples. Essa simplificação dos tensores tem o objetivo de manter a parcela relevante para a aplicação enquanto descarta a parcela desnecessária, de forma que esses dados tensoriais complexos possam ser compatibilizados para serem entradas dos algoritmos, como métodos de CNN no caso desse trabalho. Essa simplificação dos tensores é feita com o cálculo de mapas escalares que consideram as relações entre as dimensões e direções dos tensores bem como as relações com seus vizinhos.

Podemos separar os mapas em três grupos principais. O primeiro grupo contém mapas escalares que levam em consideração apenas os autovalores (Mapas Escalares *intravoxel*) e cada valor é proveniente de apenas o seu respectivo *voxel*. O segundo grupo, Mapas de Gradiente Morfológico Tensorial, é composto por mapas de dissimilaridade entre os tensores vizinhos, portanto, levam em conta as direções relativas dos tensores adjacentes (RITTNER; LOTUFO, 2008; RITTNER *et al.*, 2010). Por fim temos os mapas de medidas de divergência e rotação (Rotacional e Divergente) dos campos vetoriais provenientes dos tensores. Esses mapas também levam em conta a direção principal dos tensores e a relação de vizinhança entre eles.

Cada um desses grupos de mapas apresentam informações diferentes dos tensores que são definidos pela difusividade no meio restrito dos tecidos biológicos. Portanto, cada um desses mapas carrega um tipo de informação da microestrutura dos tecidos, permitindo que aplicações se beneficiem das informações mútuas e complementares de cada uma dessas medidas.

#### Mapas Escalares *intravoxel*

A partir dos modelos de tensores de difusão, diversas métricas podem ser extraídas para que determinadas características do tecido dentro do espaço de cada *voxel* sejam expostas em valores escalares. Algumas dessas métricas, por exemplo, são a Anisotropia Fracional (FA), a Difusividade Média (MD), Difusividade Radial (RD), Difusividade Axial (AD) e Difusividade Aparente (ADC) (KINGSLEY, 2006) (Fig. 2.1-e).

Além de expressarem informações de interesse, como a medida da organização de um feixe de fibras indicada pelo alto valor de FA, ou a presença de líquido não restrito em regiões com grandes valores de MD, esses mapas escalares costumam ser boas representações gráficas da difusão já que a visualização simples de um campo tensorial é demasiadamente complexa.

Neste trabalho, foram usados principalmente os mapas de FA, MD, AD e RD.

Esses mapas foram escolhidos não só por serem os mais populares, mas por apresentarem informações complementares entre si, de modo que praticamente toda informação microstrutural representável pelos autovalores possam ser expressadas e combinadas para auxiliar na tarefa de segmentação do tálamo.

O mapa de FA traduz para cada *voxel* o quanto o tensor é excêntrico, ou seja, essa medida exprime a diferença entre os comprimentos dos eixos do elipsoide. Com um valor de FA igual a 0, o tensor é isotrópico, enquanto um FA igual a 1 corresponde a um tensor completamente anisotrópico (Eq. 2.1). Logo, em estruturas fibrosas, o valor de FA tenderá a ser mais próximo de 1, e mais próximo de 0 onde não houver restrições de movimento.

$$FA = \sqrt{\frac{1}{2} \frac{(\lambda_1 - \lambda_2)^2 (\lambda_2 - \lambda_3)^2 (\lambda_3 - \lambda_1)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (2.1)$$

O mapa de MD está mais relacionado com o tamanho absoluto de cada elipsoide. Um valor de MD próximo de 1 indica uma grande difusão naquele *voxel*, enquanto MD próximo de 0 indica que quase não há difusão naquele ponto (Eq. 2.2). Assim, em estruturas fibrosas, o valor de MD tenderá a ser mais baixo, e mais próximo de 1 onde não houver restrições de movimento.

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \quad (2.2)$$

O mapa das medidas de difusividade axial (AD), representa a medida de difusividade aparente no eixo de maior difusão. Isso significa que essa medida só considera a difusividade na direção de menor restrição à difusão (Eq. 2.3), ou seja, a difusividade que está alinhada com o autovetor principal do tensor. Essa medida é coincidente com o valor do maior autovalor do tensor, justamente por ser a medida na direção mais alongada do tensor, portanto a medida AD muitas vezes é apelidada de autovalor 1.

$$AD = \lambda_1 \quad (2.3)$$

Já o mapa de difusividade radial (RD) considera somente a difusividade perpendicular ao eixo da direção principal. Assim, só leva em conta os valores de difusividade ortogonais ao autovetor principal (Eq. 2.4).

$$RD = \frac{\lambda_2 + \lambda_3}{2} \quad (2.4)$$

Um ponto muito importante comum a todos esses mapas escalares é que eles consideram somente os autovalores dos tensores em seu cálculo (Eq. 2.1, Eq. 2.2, Eq. 2.3).

e Eq 2.4). Isso significa que nenhuma informação direcional de difusão, portanto sobre direcionalidade das fibras, é considerada. Assim, podemos dizer que esses mapas somente expressam informação referentes às intensidades de difusividades, mas não sobre às direções em que elas ocorrem.

### Mapas de Gradiente Morfológico Tensorial

O Gradiente Morfológico Tensorial (TMG) (RITTNER *et al.*, 2007) é uma representação da dissimilaridade entre *voxels* vizinhos, sendo cada *voxel* representado por um tensor. Assim, o TMG tem como resultado um mapa escalar das medidas de dissimilaridade em cada ponto do campo tensorial (Fig. 2.3).

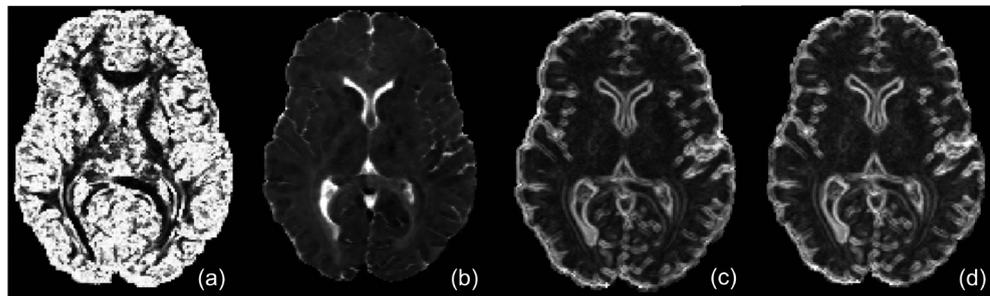


Figura 2.3 – Mapas de Gradiente Morfológico Tensorial: (a) produto interno dos autovetores principais; (b) produto interno tensorial; (c) norma de Frobenius; (d) distância Log-Euclidiana.

A dissimilaridade entre os *voxels* pode ser calculada através de diversas funções matemáticas. No caso do trabalho atual, as funções matemáticas para cálculo da dissimilaridade dos tensores são: produto interno dos autovetores principais (prod) (Eq. 2.5), produto interno tensorial (pten) (Eq. 2.6), norma de Frobenius (frob) (Eq. 2.7) e a distância Log-Euclidiana (logE) (Eq. 2.8).

$$d_1(T_i, T_j) = |e_{1,i} \cdot e_{1,j}| \quad (2.5)$$

$$d_2(T_i, T_j) = \lambda_{1,i}\lambda_{1,j}(e_{1,i} \cdot e_{1,j})^2 + \lambda_{2,i}\lambda_{2,j}(e_{2,i} \cdot e_{2,j})^2 \quad (2.6)$$

$$d_2(T_i, T_j) = \lambda_{1,i}\lambda_{1,j}(e_{1,i} \cdot e_{1,j})^2 + \lambda_{2,i}\lambda_{2,j}(e_{2,i} \cdot e_{2,j})^2 \quad (2.7)$$

$$d_3(T_i, T_j) = \sqrt{\text{Trace}((T_1 - T_2)^2)} \quad (2.8)$$

onde  $e_{1,i}$  e  $e_{1,j}$  são os autovetores principais dos tensores  $T_i$  e  $T_j$ .

Importante ressaltar que, apesar do resultado das métricas entregues por essas funções matemáticas serem mapas escalares, uma parcela da informação direcional originalmente contida no dado tensorial é mantida. Isso pode ser verificado pelo fato de que as funções matemáticas levam em conta não só os autovalores dos tensores mas também os autovetores.

Devido à natureza dos mapas TMG conseguirem carregar informações direcionais relativas, esses mapas são bons candidatos para serem aplicados como mapas de entrada para métodos de segmentação de estruturas de baixo contraste escalar e compostas por tecidos fibrosos.

### Rotacional e Divergente

Rotacional e divergente são conceitos aplicados em campos vetoriais ou tensoriais para estudo do comportamento regional do campo sobre determinadas condições. Alguns exemplos de suas aplicações são medida de fluxo magnético, determinação de pontos de fontes ou drenos em fluidos e análise de vorticidade. Tanto o rotacional quanto o divergente são computados pelo produto do operador diferencial Nabla (Eq 2.9) com o campo em estudo, vetorial ou tensorial. No entanto, a diferença entre o cálculo do rotacional e do divergente é que no primeiro é feito o produto interno enquanto no segundo se calcula o produto vetorial.

$$\nabla = \hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \quad (2.9)$$

Analogamente à operação entre dois vetores, em que o produto interno diminui um grau na ordem do vetor e que o produto vetorial mantém a ordem, o divergente é uma ordem menor do que o campo que o gerou enquanto o campo vetorial tem a mesma ordem. Exemplificando, o divergente (Eq 2.10) de um campo vetorial (tensor de ordem 1) é um campo escalar (tensor de ordem 0) e o rotacional (Eq 2.11) de um campo vetorial é também um campo vetorial.

$$\text{div}(\vec{F}) = \nabla \cdot \vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \quad (2.10)$$

$$\text{rot}(\vec{F}) = \nabla \times \vec{F} = \begin{pmatrix} \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \\ \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \\ \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \end{pmatrix} \quad (2.11)$$

Diferentemente dos mapas escalares computados com o emprego de apenas os autovalores, os mapas de rotacional e divergente consideram os autovetores dos tensores (PINHEIRO *et al.*, 2016). Portanto, esses são mapas influenciáveis pela direcionalidade e pela variação de direcionalidade regional das fibras dos tecidos cerebrais.

Para utilização desses mapas em imagens de tensores de difusão, existem algumas possibilidades de configurações. A primeira delas é a aplicação do operador diferencial

diretamente no campo tensorial, gerando um campo vetorial, no caso do divergente, e um outro campo tensorial, no caso do rotacional. No entanto, neste trabalho foi utilizado uma segunda abordagem, onde o campo tensorial é convertido em campo vetorial antes da computação dos mapas. Essa abordagem é vantajosa para aplicações que utilizam CNNs ou outras técnicas de processamento de imagens por reduzir a dimensionalidade dos dados enquanto preserva a maior parte da informação direcional presentes dos dados de difusão.

A conversão de campo tensorial em campo vetorial é feita em duas etapas (Fig. 2.4). A primeira é a exclusão das direções secundárias dos tensores, mantendo somente a direção principal do tensor de difusão. A segunda etapa é a aplicação de um sentido de interesse ( $\vec{u}$ ) já que a etapa anterior resulta em um campo vetorial de direções definidas mas de sentidos inespecíficos. O resultado dessa conversão é um campo vetorial de direções altamente correlacionadas com os feixes de fibras dos tecidos e sentidos alinhados com um vetor de interesse.

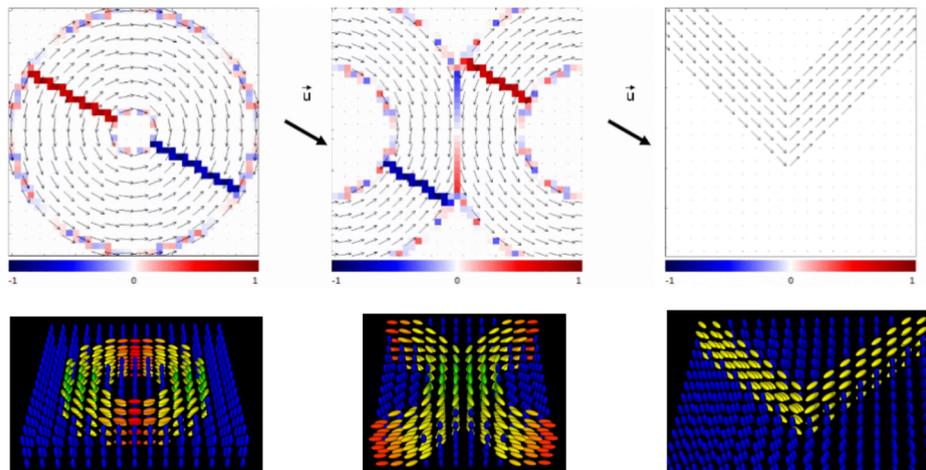


Figura 2.4 – Exemplos de mapas de divergente proveniente de campos tensoriais dado um sentido de interesse ( $\vec{u}$ ): campos vetoriais sobrepostos aos mapas de divergente (linha superior) e campos tensoriais sintéticos (linha inferior).

Em certas aplicações, como na definição de regiões de interesse do corpo caloso (PINHEIRO *et al.*, 2016), o vetor de interesse é definido se baseando na geometria em que se quer caracterizar (Fig. 2.5). Em outros casos, como na aplicação de segmentação do tálamo com CNNs, uma das possíveis abordagens é a utilização de múltiplos vetores de interesse, gerando uma coleção de campos vetoriais nos quais se pode calcular as medidas de divergência, por exemplo.

Após a obtenção do campo, ou campos vetoriais, é aplicada a função de divergente ou rotacional para a obtenção dos novos mapas com as medidas que caracterizam pontualmente os campos vetoriais. Nesse trabalho foram aplicados 32 vetores de interesse sendo todos eles aleatórios e não coplanares, e em seguida foram computados os mapas

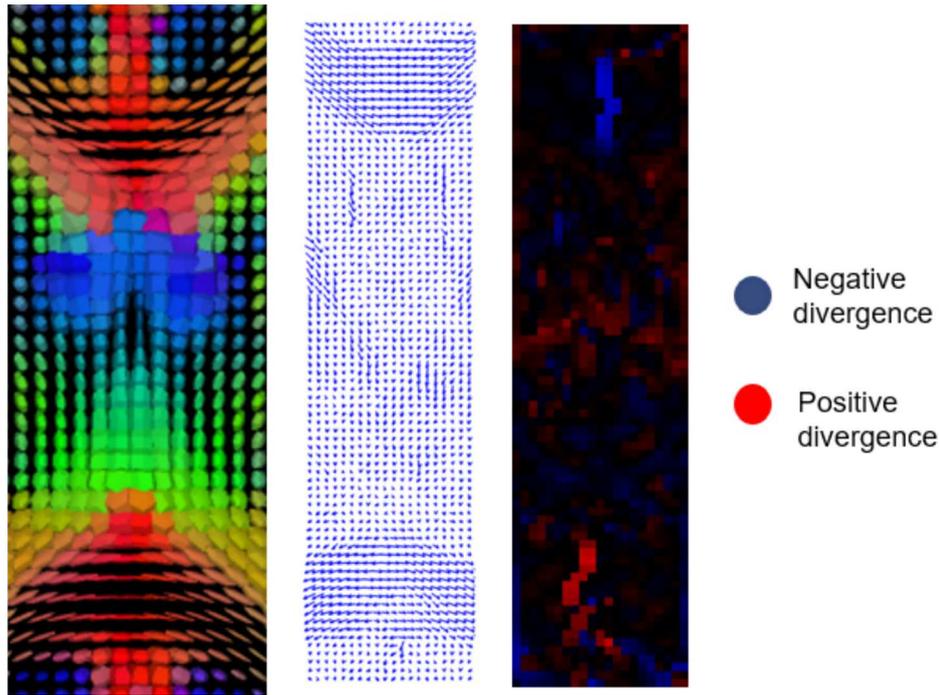


Figura 2.5 – Vista axial de mapa de divergente proveniente de campo tensorial na região do corpo caloso para direção de interesse anteroposterior: destaque de mínimos e máximos nas regiões do joelho e esplênio do corpo caloso.

de divergente para serem tomados como um dos possíveis grupos de entradas das CNNs.

## 2.2 Conjuntos de dados

Apesar de uma grande disponibilidade de conjuntos de dados públicos relacionadas a imagens médicas, a maioria delas não possui os requisitos necessários para conduzir estudos de segmentação do tálamo.

Um dos principais requisitos está relacionado com a disponibilidade de imagens de resolução razoável para o problema, principalmente para o parcelamento do tálamo que contém estruturas com tamanho inferior a  $10mm^3$ . No caso de imagens de baixa resolução, as estruturas ficam compreendidas em uma quantidade muito pequenas de *voxels*, diminuindo a precisão da segmentação e intensificando o efeito de volume parcial.

Outro requisito, não mandatário mas com grande diferencial, é a presença de imagens ponderadas em difusão, além das imagens ponderadas em T1. Trabalhos recentes (IGLESIAS *et al.*, 2018; BATTISTELLA *et al.*, 2017; MAJDI *et al.*, 2020; UMAPATHY *et al.*, 2022) mostram que o parcelamento do tálamo apresenta resultados mais precisos quando feito em imagens específicas como T1 em 7T (MPRAGE) ou em imagens de difusão, que carregam mais informações da microestrutura dos tecidos.

Apresentado os requisitos considerados relevantes para a aplicação proposta,

Tabela 2.1 – Conjuntos de dados: quantidade de imagens, tipo de dados e resoluções.

Base de dados	N sujeitos	Qtd usada	Tipos de imagem	N direções de difusão	Resolução
HCP	1200	1063	T1 e difusão	90 direções 3 <i>shells</i>	1,25x1,25x1,25mm
HCU	134	121	T1 e difusão	32 <i>single shell</i>	1x1x2mm

temos um número reduzido de conjuntos de imagens capazes de atender as especificações. Alguns exemplos de conjuntos de dados candidatos são o HCP (Human Connectome Project) e HCU (Hospital de Clínicas - UNICAMP).

Por ser o maior conjunto de dados e com melhor resolução, o HCP foi adotado como a principal coleção de imagens. Outro conjunto de dados utilizado nos experimentos é o HCU, tanto por atender os requisitos mínimos para o tipo de aplicação quanto pela disponibilidade local e interesse mútuo de outros pesquisadores do grupo de pesquisa e de outros departamentos da UNICAMP.

Nos experimentos preliminares, onde foram testadas as técnicas de treinamento de CNNs, foi utilizado o conjunto de dados HCU, porém no contexto de segmentação de diversas estruturas sub-corticais. Já nos experimentos específicos de segmentação e parcelamento do tálamo, foi utilizado os conjunto de dados HCP (Tab. 2.1). Vale ressaltar que uma pequena parte dos sujeitos de cada conjunto não foi utilizada por problemas ocorridos durante o pré-processamento desses dados.

### 2.2.1 HCP

O Human Connectome Project (HCP) (ESSEN *et al.*, 2013) é um consórcio que estuda a conectividade cerebral em adultos saudáveis, ao mesmo tempo que permite acesso público a todos os dados relevantes. Ele fornece imagens de RM coletadas em diversas modalidades, incluindo imagens estruturais T1w, dMRI e imagens funcionais. Porém, como os objetos desse trabalho são imagens T1w e dMRI, outras imagens como T2w e imagens funcionais serão ignoradas.

O HCP fornece dados minimamente pré-processados para ambas as modalidades (GLASSER *et al.*, 2013). Os dados de ressonância magnética T1w foram corrigidos para distorções de gradiente, leitura e campo de polarização, enquanto os dados de dMRI incluem normalização de intensidade  $b_0$  e correção para campos  $b_0$  induzidos por suscetibilidade, corrente parasita, movimento do sujeito e distorções de gradiente.

Essa base de dados conta com mais de 1200 sujeitos, sendo 1065 atendendo o requisito de coexistência de imagem T1w e dMRI. Entretanto, esse trabalho utiliza imagens de 1063 sujeitos já que o algoritmo de registro entre as imagens T1w e dMRI não convergiu em dois sujeitos.

O conjunto de dados é relativamente equilibrado em relação ao sexo dos sujeitos

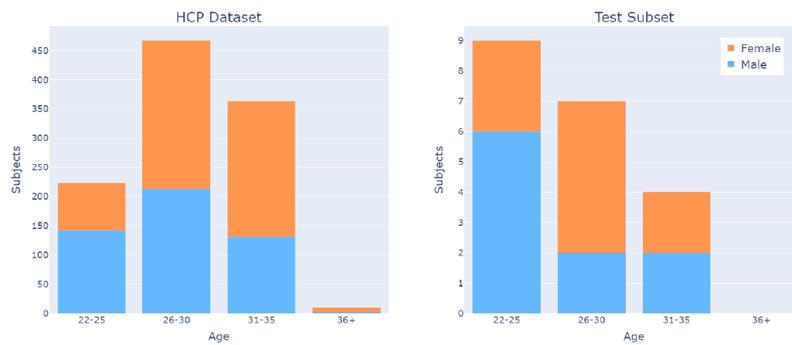


Figura 2.6 – Distribuições de faixa etária e sexo para os subconjuntos de dados do HCP utilizados durante as fases de treinamento e de teste das redes neurais.

e tem uma distribuição etária de 22 a 35 anos (Fig. 2.6).

Com essas especificações, boa resolução espacial para imagens escalares e excelente resolução direcional para dMRI, o conjunto de dados do HCP se tornou a principal coleção de imagens para o trabalho de segmentação e parcelamento do tálamo.

## 2.2.2 HCU

O conjunto de dados do Hospital de Clínicas da UNICAMP (HCU) foi desenvolvido na Faculdade de Ciências Médicas da UNICAMP em parceria com o Hospital de Clínicas e o Cepid BRAINN.

Esse conjunto de dados é privado, com algumas exceções de dados publicados para propósitos específicos, como no conjunto de dados Calgary-Campinas-359 (CC-359) (SOUZA *et al.*, 2018). Portanto, apenas grupos de pesquisa associados às entidades detentoras dos dados possuem acesso e para aplicações determinadas, inclusive previamente aprovadas pelo comitê de ética.

A coleção completa de imagens do conjunto de dados do HCU é composta por milhares de pacientes e sujeitos saudáveis. Grande parte dessas imagens foram coletadas com o propósito primário de estudo de Lupus e Epilepsia, utilizando um protocolo de aquisição com diversas sequências. Com isso esse conjunto de dados é composto por aquisições de imagens funcionais e estruturais, dentre elas T1w, FLAIR, T2w e dMRI.

## 2.3 Pré-processamento

Para a utilização dos dados em redes convolucionais, ou quaisquer outros métodos de processamento ou segmentação das imagens, são necessários vários passos para

que os dados sejam adequados e padronizados, independentemente de qual conjunto de dados as imagens são originárias.

Principalmente no caso onde imagens de múltiplas origens ou de múltiplas sequências de RM são empregadas, como é o caso do trabalho em questão, a harmonização entre as diferentes imagens é fundamental para o bom desempenho dos algoritmos. Além da necessidade de diversos passos para a compatibilização, essas etapas não são triviais e nem suficientemente documentadas na literatura. Essas etapas que manipulam os dados, garantindo que estes sejam corretamente aplicados às redes neurais convolucionais são definidas nessa tese como pré-processamento e compõe parte das contribuições dessa pesquisa.

No caso específico deste trabalho, o pré-processamento dos dados é feito com múltiplos propósitos:

- Compatibilizar as imagens de diferentes origens, sejam elas de sequências de aquisição diferentes ou de métricas distintas dentro de uma mesma sequência;
- Padronizar as imagens dentro de escalas de intensidade consistentes;
- Assegurar consistência da resolução das imagens;
- Filtrar ruídos e artefatos de aquisição das imagens;
- Gerar máscaras de padrão-prata para o pré-treinamento;

O pré-processamento segue um fluxo que parte das imagens provenientes da máquina de ressonância (Fig. 2.7-verde) e vai até a entrega de imagens e máscaras apropriadas para treinamento e aplicação de CNNs. O fluxo dos dados passa por dois caminhos paralelos, sendo o primeiro para compatibilizar as imagens de entrada e o segundo focado em preparação das máscaras padrão-prata.

O tratamento das imagens de entrada é, por sua vez, sub-dividido em processamento das imagens de difusão, onde são calculados os tensores do modelo DTI e os mapas de difusão a partir dos tensores, e processamento das imagens T1w, onde estas são registradas para o espaço de difusão (Fig. 2.7-azul).

O outro ramo do pré-processamento (Fig. 2.7-amarelo) foca na preparação das máscaras, onde essas são computadas pelos métodos dos FSL, FreeSurfer e QuickNAT (partindo do espaço original das imagens), e em seguida, registradas para o espaço de difusão. Por fim são combinadas entre si com o método STAPLE (WARFIELD *et al.*, 2004), gerando uma quarta máscara que é considerada o padrão-prata.

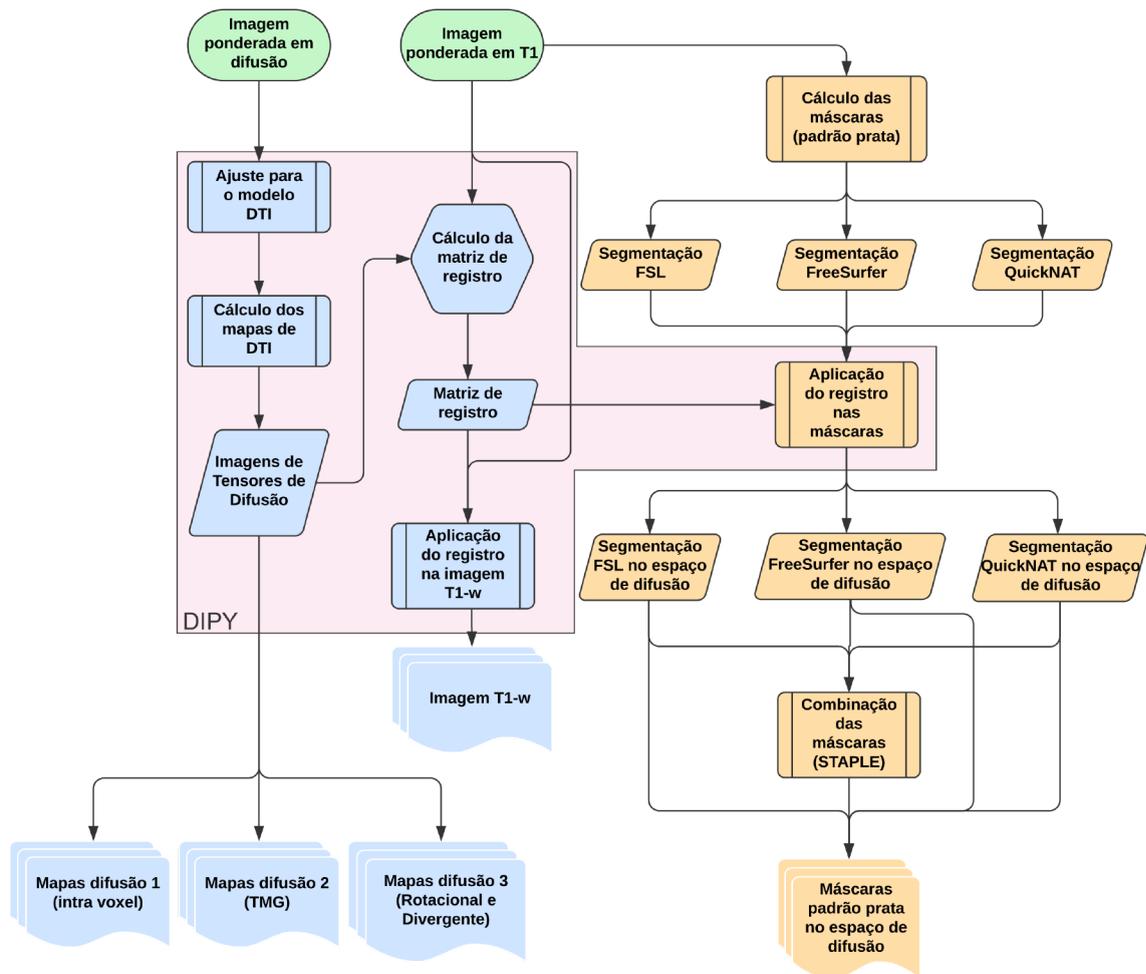


Figura 2.7 – Pré-processamento dos dados: imagens originais (verde); etapas de processamento das imagens para entrada das CNNs (azul); preparação das máscaras de padrão-prata (amarelo). Cálculo do modelo DTI e registros entre sequências de MRI feitas com a biblioteca DIPY (rosa) (GARYFALLIDIS *et al.*, 2014).

### 2.3.1 Normalização espacial e registro das imagens

Uma técnica importante para assegurar o desempenho da rede convolucional quando trabalhamos com imagens médicas é a normalização da resolução espacial, ou tamanho do *voxel*, garantido que seja consistente durante as etapas de treino e predição. Este efeito ocorre porque a rede aprende a identificar as características da imagem para uma determinada escala, que está diretamente relacionada com o tamanho físico do *voxel*, e a presença de imagens com *voxel* de outros tamanhos no conjunto de treinamento pode atrapalhar a convergência da rede. Isso ocorre pois a identificação da textura de dada região da imagem ocorre pela caracterização feita pela convolução dos *kernels* com as imagens. Portanto, com a proporção entre os tamanhos dos *kernels* e *voxels* se mantendo fixa, a CNN não precisa lidar com possibilidades provenientes de razões diversas.

Além dos tamanhos compatíveis, é importante que os *voxels* referentes ao mesmo lugar no espaço estejam alinhados. Isto significa que, principalmente quando estamos lidando com dados multimodais, todas as imagens de entrada e as máscaras precisam estar co-registradas em um espaço de referência. Este espaço pode ser o espaço nativo de cada sujeito ou um espaço padrão, como o MNI (DARKNER, 2013) por exemplo.

### 2.3.2 Cálculo dos mapas de difusão

Os primeiros passos em métodos que utilizam imagens de tensores de difusão são as correções necessárias para viabilizar o ajuste dos dados em algum modelo (IRFANOGLU *et al.*, 2019; WU *et al.*, 2008). Algumas das etapas necessárias são a correção geométrica das distorções causadas por *eddy current* na máquina de ressonância, correções de intensidade e o registro entre as imagens provenientes de cada uma das direções de gradientes. No caso dos conjuntos de dados HCP e HCU, os dados são previamente corrigidos, sendo entregues adequados para as próximas etapas.

Os modelos de difusão são construídos sobre premissas que restringem os graus de liberdade, viabilizando uma modelagem matemática. No caso do modelo DTI, as principais premissas são de que a difusão existe em todo o espaço, segue uma distribuição probabilística de formato tensorial e que não há fluxo, ou seja, cada *voxel* contém tensores e estes são simétricos (BASSER *et al.*, 1994; BASSER; PIERPAOLI, 2011). Essas premissas implicam que não faz sentido computar o modelo em regiões onde não há difusão, como na calota craniana e em seu exterior, pois isso iria gerar medidas instáveis. Portanto, logo após as correções supracitadas, é executado um processamento que delimita o cérebro (*skull-stripping*), mascarando todas as regiões exteriores ao cérebro.

Com a região onde serão computados os tensores delimitada, é feito o ajuste das medições ao modelo DTI. Para isso, são necessárias computações para a remoção da parcela que não se refere a difusão, ajuste por mínimos quadrados, e outras etapas intermediárias que não serão detalhadas aqui. Nos experimentos iniciais, todas as etapas foram feitas por implementações próprias, porém, para acelerar o processamento e para que os métodos possam ser popularizados mais facilmente, a biblioteca DIPY (GARYFALLIDIS *et al.*, 2014) passou a ser utilizada para computar o modelo DTI a partir das imagens DWI.

Com os modelos DTI em mãos, podemos computar os mapas derivados desse modelo, chegando a medidas que expressam diferentes informações dos dados, e indiretamente dos tecidos biológicos. Como mencionado em Seção 2.1.2.1, são definidos 3 grupos de mapas, sendo que os mapas escalares *intravoxel* são computados com a biblioteca DIPY e os mapas dos outros 2 grupos são computados com ferramentas próprias, já que são mapas propostos pelo nosso grupo de pesquisa.

### 2.3.3 Normalização de intensidade

Esse trabalho utiliza imagens oriundas de múltiplas sequencias de MRI e mapas que possuem diferentes escalas de contrastes. Portanto, se faz necessária a padronização das intensidades de cada uma das imagens, que são as amplitudes dos canais de entrada das CNNs. Um dos principais fatores motivadores dessa necessidade de padronização das intensidades entre os canais é que se as escalas forem diferentes, os canais de maior amplitude se sobressaem em relação aos demais, excluindo de forma indireta a contribuição possível das outras imagens de menor amplitude. Como as CNNs e suas funções de ativação são adequadas para uma determinada faixa de valores, as imagens precisam estar em uma faixa restrita de valores. Nos experimentos conduzidos nesse estudo, todas as imagens tiveram seus valores restritos entre 0 e 1.

Devido à natureza da aquisição que insere ruídos nas imagens, a melhor estratégia para normalização é a eliminação dos valores mais baixos e mais altos (*clipping*), referentes aos picos superiores e inferiores, seguidos pelo ajuste dos valores restantes entre 0 e 1. O valores de pico são, como no caso dos experimentos conduzidos nesse trabalho, são aqueles definidos na faixa dos percentis 0,02 a 0,98.

Essa filtragem por percentil é fundamental, principalmente nos mapas de imagens de difusão. Por conta dos ruídos anteriores ao ajuste no modelo de difusão, acabam aparecendo valores espúrios que são fisicamente impossíveis (como uma difusão negativa) ou muito destoantes dos demais valores. Esses problemas normalmente ocorrem em regiões de baixa difusão ou difusão nula, como nas regiões da calota craniana que eventualmente remanesçam parcialmente após a etapa de remoção.

## 2.4 Máscaras

O treinamento de CNNs é um processo supervisionado, portanto dependente de dados anotados de modo que cada dado de entrada possua a respectiva máscara de segmentação da estrutura de interesse. Idealmente, a máscara de segmentação de cada uma das imagens deve ser fornecida por especialistas, estabelecendo o padrão-ouro, porém isso é inviável em grande parte dos casos, principalmente naqueles de grandes volumes de dados. Nesses casos, uma alternativa é utilizar o artifício de máscaras geradas automaticamente por outros métodos ou algoritmos, mesmo que essas não possuam a mesma qualidade de uma segmentação manual.

Como nesse trabalho são usadas diversas técnicas de treinamento de CNNs, grandes quantidade de dados, e parte dos resultados avaliados contra o padrão-ouro, foram utilizadas tanto máscaras feitas por especialista quanto máscaras geradas por algoritmos.

### 2.4.1 Anotação manual dos dados

A segmentação manual é atualmente considerada como o padrão-ouro para a geração de máscaras. Tal consideração vem do fato de um especialista, experiente com a atividade, ser capaz de distinguir não só pela intensidade do *voxel*, mas também pelo contexto, as bordas das estruturas. Para que o processo de segmentação seja executado de forma consistente, os especialistas adotam protocolos que padronizam as etapas da segmentação, aumentando a precisão das segmentações e mitigando divergências entre segmentações feitas em momentos distintos ou por especialistas diferentes.

No caso da segmentação do tálamo, foram construídas segmentações manuais das estruturas para uma parcela dos volumes do conjunto de dados HCP. Devido à carência de protocolos de segmentação específicos para o tálamo e limitações dos atuais protocolos, as segmentações foram feitas com o auxílio de um protocolo desenvolvido juntamente ao trabalho atual. O protocolo de segmentação manual proposto é uma das contribuições originais deste trabalho (Apêndice A).

O protocolo desenvolvido em conjunto com especialistas, considera não somente o uso de imagens T1w quanto o emprego de imagens de tensores de difusão para esclarecer dúvidas de classificação dos *voxels*, principalmente nas bordas inferiores e laterais externas do tálamo. Além de considerar dois tipos de imagem simultaneamente, o protocolo proposto recomenda a utilização das três vistas (coronal, sagital e axial) para que o avaliador consiga perceber o volume da estrutura e tomar a decisão de pertencimento ou não de um *voxel* com mais informação de contexto no entorno do ponto avaliado.

Tal protocolo, além de facilitar o processo de segmentação e reduzir o tempo tomado pela tarefa, visa mitigar as variabilidades e incertezas que podem ser inseridas pelo anotador especialista. Vale ressaltar que os resultados do protocolo de segmentação manual proposto não devem ser diferentes da segmentação feita somente em imagens T1w já que ambos se referem à mesma estrutura e o objetivo é apenas facilitar o processo e melhorar a acurácia da segmentação.

Essas imprecisões, intrínsecas à tarefa de segmentação, são ocasionadas por diversos fatores como volume parcial, vieses dos especialistas, limitações humanas de atenção, qualidade e configuração do monitor, recursos da ferramenta de segmentação, entre outros. Por consequência, mesmo o padrão-ouro possui limitações que acabam impondo um limite a quaisquer algoritmos que levam em conta as máscaras geradas pela segmentação manual.

Tabela 2.2 – Comparação de máscaras entre avaliadores: DC - Coeficiente DICE; FNR - Taxa de Falsos Negativos; FPR - Taxa de Falsos Positivos; VS - Similaridade volumétrica; AHD - Distancia Hausdorff média.

Comparação	DC (SD)	FNR	FPR	VS	AHD (SD)
Especialista vs Amador	0.8022 (0.0804)	0.1228	0.2472	0.1555	0.2501 (0.0573)
Especialista vs STAPLE	0.7880 (0.0310)	0.0070	0.3459	0.4128	0.2190 (0.0340)
Amador vs STAPLE	0.7108 (0.0655)	0.0129	0.4397	0.5577	0.3685 (0.0971)

#### 2.4.1.1 Estudo intra- e inter-avaliador

O protocolo de segmentação manual empregado foi projetado para facilitar essa tarefa. Outra função do protocolo é estabilizar e dar mais precisão às segmentações, principalmente considerando a variabilidade relativamente alta intra- e inter-avaliadores na segmentação manual do tálamo apenas para imagens T1w, conforme demonstrado em Makowski *et al.* (2018).

Dos 60 sujeitos considerados para a tarefa de segmentação manual, 16 foram realizados por um avaliador não especialista, sendo os 44 restantes conduzidos por um médico com 15 anos de experiência em estrutura cerebral e segmentação manual em ressonância magnética. Todas as segmentações manuais realizadas pelo avaliador amador também seguiram o protocolo e foram utilizadas apenas para ajuste fino da CNN. Para avaliação dos modelos de CNN são consideradas apenas segmentações manuais, totalizando 20 sujeitos.

Para avaliar a variabilidade da segmentação entre os avaliadores especialistas e não especialistas, assim como na comparação com o padrão-prata (STAPLE), cinco sujeitos adicionais foram segmentados e algumas métricas foram calculadas. A comparação (Tab. 2.2) indica que as segmentações manuais dos dois avaliadores estão mais próximas uma da outra do que do padrão-prata proposto, especialmente quando se considera as métricas de sobreposição (coeficiente DICE) e a similaridade volumétrica (TAHA; HANBURY, 2015).

#### 2.4.2 Anotações automáticas

Diferentemente do padrão-ouro, segmentações geradas automaticamente por algoritmos e não inspecionadas por um especialista são consideradas padrão-prata. Isso significa que são segmentações ou máscaras com qualidades e limitações caracterizadas, mas sem a chancela de um especialista, e assim, não necessariamente possuem validade para avaliações oficiais dos resultados. Entretanto, o uso dessas máscaras é um bom artifício para auxiliar no treinamento das redes neurais, especialmente com grandes conjuntos de dados sem anotação manual.

### 2.4.2.1 Segmentação

Para a geração automática das máscaras do tálamo, foram utilizadas as ferramentas FreeSurfer, FSL e QuickNAT. Essas três máscaras foram fundidas com o método STAPLE (WARFIELD *et al.*, 2004) (Fig. 2.8), que gera uma mapa probabilístico da segmentação da estrutura. Após uma limiarização, a fusão das máscaras foram utilizadas no processo supervisionado de treino das CNNs.

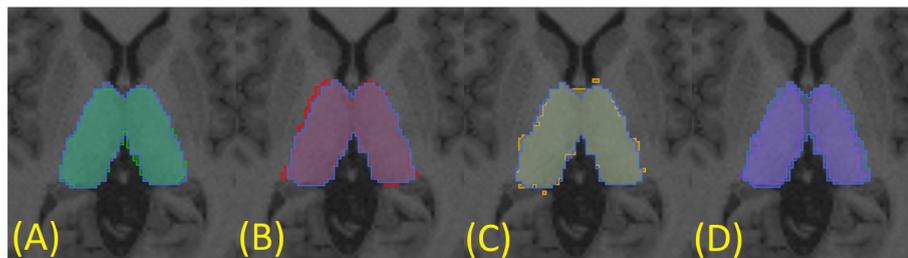


Figura 2.8 – Visão axial das máscaras do tálamo calculada pelo STAPLE (em azul) sobreposta à máscara calculada por cada método: (A) FSL—verde; (B) FreeSurfer—vermelho; (C) QuickNAT—amarelo; (D) segmentação manual—roxo.

Vale ressaltar que o padrão-prata é um comitê que combina métodos baseado em atlas com máscaras geradas por CNN por uma questão de diversidade, e por consequência, melhor qualidade nas definições das bordas das estruturas.

### 2.4.2.2 Parcelamento

Para geração de máscaras para os experimentos de parcelamento do tálamo, somente a ferramenta FreeSurfer foi utilizada. Essa ferramenta é a única que possui um módulo que, após a segmentação do tálamo juntamente com as outras estruturas pelo processamento padrão da ferramenta, consegue gerar o parcelamento do tálamo através de um módulo especial para essa finalidade específica (IGLESIAS *et al.*, 2018).

Apesar de relativamente pequeno, o tálamo pode ainda ser subdividido em núcleos, chegando a apresentar 25 regiões distintas para cada metade do tálamo (Fig. 2.9). Entretanto, em parte das aplicações e estudos, os núcleos podem ser agrupados em regiões anatômicas ou funcionais, constituindo tipicamente entre 5 e 10 grupos (Tab. 2.3).

Esse agrupamento viabiliza aplicações e métodos já que na escala de resolução das MRI típicas, alguns dos núcleos seriam de dimensões *sub-voxel* ou assumiriam apenas alguns poucos *voxels*.

Um detalhe importante é que os núcleos equivalentes oriundos de cada hemisfério levam rótulos diferentes. No entanto, essas estruturas são consideradas como sendo algo único já que as propriedades dos tecidos de cada estrutura são as mesmas, fazendo com que a rede aprendesse que objetos com mesmas propriedades são diferentes, piorando

o desempenho do método. A diferenciação entre núcleos de cada hemisfério é feita a posteriori com um mascaramento de uma segmentação do tálamo que considera a lateralidade na definição dos rótulos.

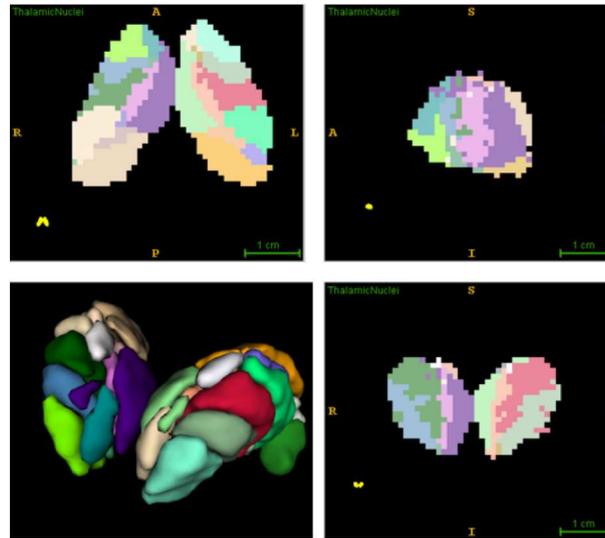


Figura 2.9 – Exemplo de parcelamento do tálamo: parcelamento em 25 regiões (IGLESIAS *et al.*, 2018).

Tabela 2.3 – Núcleos do tálamo e seu agrupamentos em 5 e 10 regiões (IGLESIAS *et al.*, 2018).

Agrupamento (5)	Agrupamento (10)	Abreviação	Núcleo
Antero-Lateral	Anterior	AV	Anteroventral
	Dorsal	LD	Laterodorsal
		LP	Lateral posterior
	Lateral-Rostral	VA	Ventral anterior
		VAmc	Ventral anterior magnocellular
		VLa	Ventral lateral anterior
		VLP	Ventral lateral posterior
VM		Ventromedial	
Lateral-Caudal	Lateral-Caudal	VPL	Ventral posterolateral
Intralaminar	Intralaminar	CeM	Central medial
		Pc	Paracentral
		Pf	Parafascicular
		MV(Re)	Reuniens (medial ventral)
		Pi	Paratenial
	Intralaminar-Posterior	CM	Centromedian
Medial	Medial	CL	Central lateral
		MDI	Mediodorsal lateral parvocellular
		MDm	Mediodorsal medial magnocellular
Posterior	Pulvinar	L-Sg	Limitans (suprageniculate)
		PuA	Pulvinar anterior
		PuI	Pulvinar inferior
		PuL	Pulvinar lateral
		PuM	Pulvinar medial
	LGN	Lateral geniculate	
	MGN	Medial Geniculate	

## 2.5 Redes Neurais Convolucionais e fusão de dados

Apesar das propostas frequentes de arquiteturas, os métodos atuais baseados em aprendizagem profunda, incluindo QuickNAT (ROY *et al.*, 2018a) e nnU-Net (ISENSEE *et al.*, 2021), são baseados na arquitetura U-Net (RONNEBERGER *et al.*, 2015), uma das arquiteturas mais populares de CNN para problemas de imagens médicas. Por esse motivo, a U-Net foi tomada como a arquitetura principal para os experimentos conduzidos para a segmentação do tálamo.

Além de experimentos conduzidos com a U-Net, neste trabalho, utilizamos também a QuickNAT e a nnU-Net em experimentos secundários e preliminares a fim de comparação do desempenho das diferentes arquiteturas e métodos.

### 2.5.1 U-Net

A U-Net é uma arquitetura de rede totalmente convolucional que foi desenvolvida para resolver problemas de segmentação de imagens biomédicas. Essa arquitetura é composta por duas partes, sendo a primeira metade denominada codificadora (*encoder*) e a segunda o decodificadora (*decoder*), ambas compostas por blocos convolucionais. O *encoder* cria representações cada vez mais abstratas das imagens de entrada enquanto a imagem é internalizada, enquanto o *decoder* faz a reconstrução das informações das abstrações, levando à segmentação.

A U-Net é atualmente uma das redes neurais mais conhecidas para segmentação de imagens médicas, e tem sido amplamente utilizada, tanto em sua forma direta ou como base para outras arquiteturas, como a QuickNAT (ROY *et al.*, 2018a; ROY *et al.*, 2018b). Uma das características mais importantes da U-Net é ser totalmente convolucional (FCNN - *Fully Convolutional Neural Network*). Isto significa que todos seus blocos são convolucionais, de modo que ela é capaz de trabalhar com qualquer tamanho de imagem e que este tamanho pode ser diferente na fase de treinamento e na fase de teste. Esta é uma vantagem quando se trabalha com quantidade reduzida de dados, já que permite o uso de *patches* ao invés de imagens inteiras no treinamento.

Como tanto a saída quanto a entrada da U-Net são camadas convolucionais, é possível definir qualquer quantidade de canais, tanto de entrada quanto de saída. Isto é particularmente interessante para aplicações multimodalidade por permitir que cada modalidade entre em um canal (fusão de dados na entrada). É também uma característica interessante para aplicações multitarefas, já que permite múltiplas saídas simultaneamente.

## 2.5.2 Fusão de dados em Redes Neurais Convolucionais

Parte dos trabalhos empregando CNNs visam resolver problemas em imagens coloridas, onde as imagens são compostas por múltiplas bandas, como R, G e B, por exemplo. Analogamente, no contexto de imagens de ressonância magnética, podemos entender as imagens de diferentes sequências ou diferentes mapas de uma mesma sequência como bandas de uma única imagem.

Neste contexto de tratamento de imagens multibanda, é necessária a aplicação de técnicas de fusão de dados. A maioria dos trabalhos de tratamento de imagens multibanda com CNN, especialmente no contexto de MRI, utiliza basicamente três abordagens de fusão (ZHOU *et al.*, 2019): fusão na saída; fusão nas camadas internas; fusão na entrada.

A técnica de fusão de dados na saída, ou na camada de decisão (NIE *et al.*, 2016; AYGÜN *et al.*, 2018; KAMNITSAS *et al.*, 2018) é uma técnica onde a fusão ocorre após a execução completa de CNNs distintas para cada uma das imagens de entrada, e a última camada apenas combina os resultados de cada uma das CNNs. Essa técnica também é conhecida como comitê de máquinas, onde o resultado final é a combinação dos resultados parciais oriundos de CNNs distintas. Um exemplo de uso desta técnica são os trabalhos que fazem segmentação em múltiplas vistas e que possuem uma rede para cada vista, combinando os resultados de cada CNN por maioria de votos ou uma pequena rede neural dedicada para a fusão.

Na técnica de fusão de dados nas camadas internas (DOLZ *et al.*, 2018c; DOLZ *et al.*, 2018b; CHEN *et al.*, 2018), a rede convolucional possui uma entrada distinta para cada imagem. Depois de uma ou mais camadas convolucionais, os dados são combinados, podendo esta combinação ser feita apenas em uma camada ou ao longo de todas as camadas convolucionais. Este tipo de fusão de dados apresenta a vantagem de conseguir extrair informações de relações complexas entre os diferentes mapas de entrada.

Por fim, a técnica de fusão de dados na entrada da CNN é, normalmente, a mais trivial e empregada na vasta maioria dos trabalhos (PEREIRA *et al.*, 2016; ISENSEE *et al.*, 2018a; ISENSEE *et al.*, 2018b; WANG *et al.*, 2017; CLÈRIGUES *et al.*, 2020) de segmentação multimodalidade. Nesta técnica, as imagens provenientes de diferentes sequências de aquisição são tratadas como canais de uma única imagem, análogo ao tratamento de imagens RGB. Porém, a abordagem típica de fusão dos dados na camada de entrada é aplicável quando cada uma das imagens de entrada são de mesma origem ou compatíveis entre si.

Em aplicações onde imagens com propriedades muito distintas, ou modalidades diferentes, como é o caso de imagens T1 e mapas de divergente tensorial, é necessário

o emprego de alguma técnica mais elaborada para compatibilizar os dados na entrada da CNN. No caso particular deste trabalho, quando é necessária a utilização de alguma técnica de fusão de dados incompatíveis, ou múltiplas modalidades, uma camada convolucional adicional é aplicada para que dados de origens distintas gerem quantidades e propriedades parecidas na entrada da CNN (Fig. 2.10).

Essa camada convolucional adicional tem a função de permitir a entrada de múltiplas imagens, ou grupos de imagens, simultaneamente enquanto compatibiliza e trata de maneira igualitária e equilibrada cada uma das entradas. Basicamente, essa camada aplica filtros convolucionais nas imagens de entrada, gerando a mesma quantidade de mapas de *features* para serem concatenados e alimentados em uma arquitetura de CNN convencional.

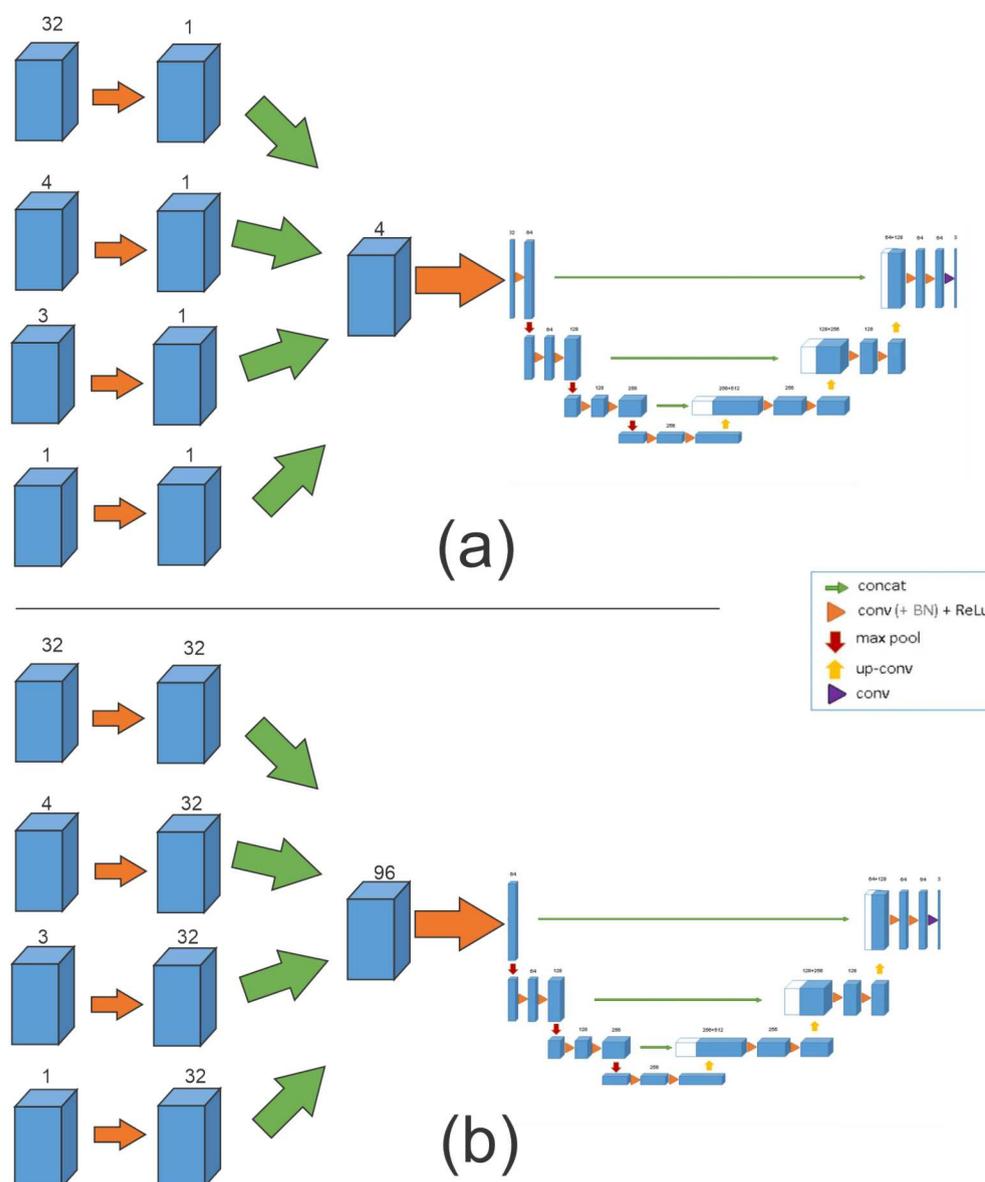


Figura 2.10 – Arquitetura U-Net para aplicações multi modalidade com fusão de dados nas camadas de entrada: (a) fusão de dados com redução de dimensão para uma entrada por modalidade; (b) fusão de dados com aumento de dimensão para múltiplas entrada por modalidade.

### 3 Segmentação do Tálamo

A segmentação do tálamo é fundamental para diversas aplicações como estudos populacionais e intervenções cirúrgicas. Em ambos os casos, é necessário que o processamento seja feito de forma robusta e rápida, já que nesses cenários, longos tempos de processamento podem inviabilizar as aplicações. Com o propósito de desenvolvimento de um mecanismo para a segmentação rápida, robusta e simples do tálamo, esse trabalho de doutorado propõe um método utilizando redes neurais totalmente convolucionais de segmentação do tálamo em imagens de ressonância.

O método de segmentação proposto compreende todas as etapas desde a preparação dos dados dos dados, treinamento de redes neurais 2D com agregação de vistas e emprego de múltiplas sequências de MRI (Fig. 3.1), inferências para a segmentação e pós-processamento.

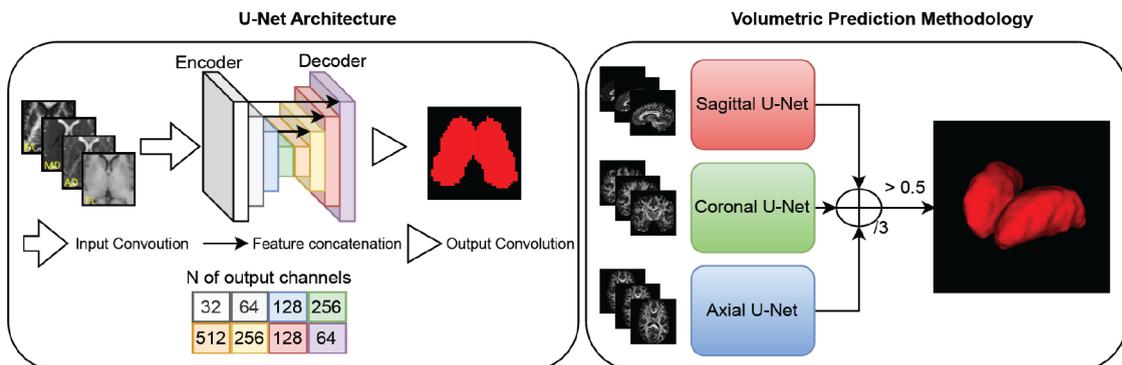


Figura 3.1 – Mapa do método de segmentação do tálamo: arquitetura da CNN 2D (U-Net como exemplo), empilhamento e agregação de vistas.

Além da proposta do método, foram feitos experimentos utilizando diferentes redes neurais convolucionais a fim de identificar a melhor arquitetura e as suas limitações no desempenho da tarefa particular de segmentação do tálamo, tanto com imagens T1w quanto com o emprego de imagens de difusão. Nesses casos, o esquema do método (Fig. 3.1) é mantido inalterado, sendo modificada apenas a arquitetura da rede neural convolucional que é empregada.

Os experimentos relatados nesse capítulo também tiveram o propósito de ser módulo de uma ferramenta integrada com outras aplicações de segmentação de estruturas subcorticais. Além da ferramenta, o método e o conjunto de dados foram publicados e disponibilizados em forma de um *benchmark*, através da plataforma CodaLab (PAVAO *et al.*, 2023), para que outros cientistas possam contribuir com o aperfeiçoamento de ferramentas para segmentação do tálamo (PINHEIRO *et al.*, 2023).

Neste capítulo, serão apresentadas as etapas do método proposto (Fig. 3.2), incluindo a preparação dos dados (Seç. 3.1), técnicas de treinamento (Seç. 3.2), ambiente computacional (Seç. 3.3), pós-processamento (Seç. 3.4), método de inferência e segmentação (Seç. 3.5), e resultados (Seç. 3.6).

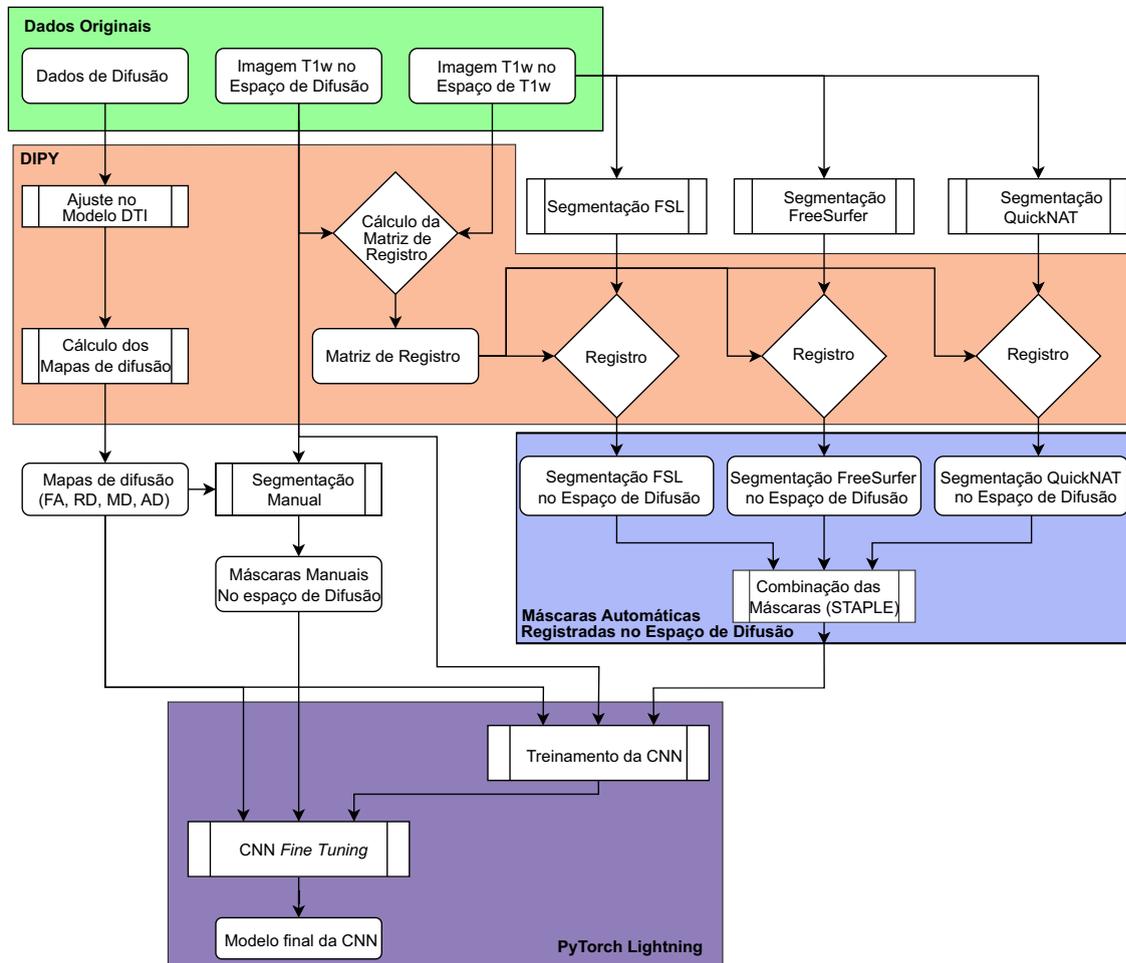


Figura 3.2 – Fluxograma do método de treinamento da CNN para segmentação do tálamo

### 3.1 Preparação dos dados

Nos experimentos preliminares desse estudo, foram utilizadas imagens do conjunto de dados HCU. Os experimentos com esse conjunto de dados tinham o objetivo de segmentar múltiplas estruturas subcorticais, incluindo o tálamo. Posteriormente, o estudo se focou particularmente no tálamo e o conjunto de dados selecionado foi o HCP.

Os conjuntos de dados foram processado para garantir sua compatibilidade com o treinamento da CNN (Seç. 2.3). Especificamente, foram tomadas medidas mínimas para padronizar e organizar os dados para utilização. O conjunto de dados de *benchmark*, vindo do HCP, foi posteriormente liberado para acesso público.

O pré-processamento compreende todas as etapas que manipulam os dados desde o primeiro tratamento dos dados originais até a disponibilização para o treinamento ou para a inferência das CNNs. É importante destacar que as imagens utilizadas para a inferência das CNNs precisam necessariamente passar pelos exatos mesmos processamento que os dados utilizados no treinamento a fim de que as imagens estejam no domínio dos canais de entrada dos modelos de CNNs treinados.

As etapas do pré-processamento dos dados são as seguintes (não necessariamente nessa ordem): cálculo da matriz de registro entre os espaços das imagens T1w e de difusão; segmentação das imagens T1w pelos métodos automáticos para a geração de máscaras padrão-prata; registro das imagens T1w e máscaras para o espaço de difusão; computação do consenso entre as máscaras automáticas registradas para o espaço de difusão; ajuste do modelo DTI; cálculo dos mapas (ou índices) do modelo DTI; segmentação manual de um subgrupo das imagens para utilização em etapa de *fine tuning*, quando necessário.

A decisão de transportar as imagens T1w para o espaço de difusão, e não o contrário, foi tomada porque as imagens T1w, por sua natureza escalar, não sofrem tanto com a interpolação de *voxel* durante o processo de registro como as imagens tensoriais sofreriam. As máscaras, que representam a classificação de cada *voxel* e são compostas por valores inteiros, são oriundas do espaço de T1w e foram também transportadas para o espaço de difusão. Essas foram interpoladas com a técnica do vizinho mais próximo, mantendo o rótulo adequado para cada *voxel* enquanto sendo praticamente imunes a erros de transporte de domínio.

Nos experimentos de segmentação do tálamo, foram utilizadas arquiteturas de CNN 2D com agregação de vistas. Experimentos preliminares foram conduzidos e ficou demonstrado que, para o problema específico de segmentação do tálamo, não há ganho de qualidade de segmentação que justifique a complexidade adicionada aos modelos e aos tratamentos dos dados, e nem o tempo e recurso computacional adicional no caso do uso de CNN 3D.

O objetivo é a segmentação de uma estrutura volumétrica e as imagens tratadas são volumes, mas as arquiteturas das CNNs são 2D, portanto, é necessária alguma forma de compatibilizar as imagens com as CNNs. A tratativa no caso desse problema é alimentar as CNNs com todas as fatias individuais do volume, de acordo com a orientação específica utilizada no treinamento, e empilhar a inferência de cada uma das fatias compondo o volume final de mesma dimensionalidade dos volumes de entrada. Por exemplo, uma CNN treinada com imagens sagitais fará a predição de cada uma das fatias sagitais e um empilhamento ao longo do eixo transversal irá compor o volume.

## 3.2 Técnicas para treinamento da CNN no contexto de imagens de difusão

Como os métodos aplicados à problemas de segmentação com CNN e imagens escalares não são diretamente compatíveis com imagens de difusão, foram testadas e adaptadas diversas técnicas de treinamento de CNN para este contexto. As técnicas exploradas foram: treinamento usando *patches*; combinação de mapas na entrada da CNN; expansão de dados (*data augmentation*); abordagem multitarefas; e diferentes funções de custo.

### 3.2.1 Treinamento usando *Patches*

Considerando que as redes empregadas são totalmente convolucionais, a CNN como um todo se comporta análoga a um único filtro convolucional, fazendo com que imagens de qualquer tamanho, maior apenas que um limite mínimo, possam ser processadas pela rede. Portanto, durante a fase de treinamento podem ser utilizados recortes das imagens (*patches*), por exemplo de tamanho  $32 \times 32$  ou  $64 \times 64$  *voxels*, enquanto a fase de teste é feita com imagens de tamanho original, por exemplo  $128 \times 128$  *voxels*.

Além de funcionar como uma espécie de técnica de expansão de dados, a utilização de *patches* ainda permite um melhor balanceamento entre as classes de *voxels* durante o treinamento. O artifício utilizado para balancear os dados é a escolha da proporção de *patches* que contenham a estrutura de interesse em relação a quantidade de *patches* aleatoriamente espalhados pelo volume da imagem. Caso o problema seja a segmentação de múltiplas estruturas simultaneamente, a escolha dos *patches* é feita para cada uma das estruturas independentemente, enquanto que para o caso de múltiplos núcleos do tálamo, os *patches* são centralizados em cada um dos núcleos (Fig. 3.3).

### 3.2.2 Expansão de dados (*data augmentation*)

Para que as redes apresentem um melhor desempenho, o ideal é treinar a CNN com imagens anotadas por especialistas, já que a rede aprende a realizar exatamente a tarefa apresentada. Porém, como este é um processo extremamente custoso, tomando horas de um especialista para fazer a segmentação de um único volume, devemos lançar mão de técnicas que consigam complementar a escassez dos dados.

Existem diversas técnicas neste sentido, e uma delas é a técnica de expansão de dados, ou *data augmentation*. Nesta técnica, aplicam-se transformações diversas às imagens rotuladas com o intuito de gerar dados verossímeis, aumentando a quantidade de dados de treino, evitando sobre-treinamento e melhorando a convergência do treinamento.

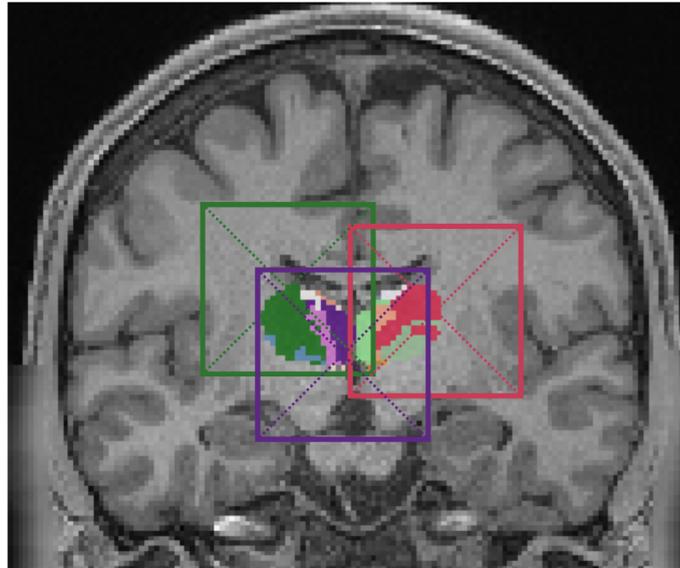


Figura 3.3 – *Patches* centralizados em diferentes núcleos do tálamo.

Neste trabalho, foram testadas e validadas as seguintes transformações: pequenas variações nas intensidades dos *voxels*; rotações aleatórias de até  $5^\circ$  nas imagens; espelhamento; adição de ruído. Vale ressaltar que, das técnicas de *data augmentation* aplicadas, nem todas são diretamente compatíveis com imagens não escalares. Por exemplo, rotações dos volumes em imagens tensoriais, como em DTI, não é uma tarefa trivial já que há necessidade de rotação e interpolação de tensores.

### 3.2.3 Abordagem multitarefas

Uma técnica promissora para treinamento de CNNs é a abordagem multitarefas. Nesta técnica, a rede aprende a tratar simultaneamente dois ou mais problemas, de modo que o treinamento é feito com a otimização de múltiplas saídas.

Em nossos experimentos, testamos esta ideia usando 3 principais abordagens. Na primeira, a rede usava somente uma saída que era otimizada para a máscara fornecida. Na segunda abordagem, a rede era treinada para aprender diversas estruturas cerebrais simultaneamente, com uma saída da rede para cada estrutura. Por fim, a CNN foi treinada para uma única estrutura, porém foram utilizadas 3 diferentes máscaras em dois contextos diferentes. No primeiro contexto cada uma proveniente de um método automático de segmentação, no segundo as máscaras foram combinadas com o métodos STAPLE.

### 3.2.4 Função de custo

A métrica mais comum para avaliar a qualidade de segmentações é a medida de similaridade DICE (ZOU *et al.*, 2004). Esta é uma medida de quanto a predição se

assemelha à máscara pela relação entre a intersecção e a união dos conjuntos de *voxels* da máscara e da predição.

Apesar desta métrica ser muito boa para validação da predição, ela não pode ser aplicada na etapa de treinamento sem algum tipo de adaptação já que é uma função discreta e não diferenciável para o otimizador do PyTorch. Por este motivo, são feitas duas alterações na definição da função da métrica. Primeiramente, cada voxel é tratado como a probabilidade de pertencimento à estrutura de interesse. Após isso, é adicionado um fator de suavização que faz o denominador da função nunca ser nulo, garantindo estabilidade numérica na função de custo.

Em redes que apresentam múltiplas saídas, é necessário expandir a definição da métrica DICE já que ela é originalmente definida para apenas duas entidades. Nestes casos, pode-se considerar uma média, simples ou ponderada, da métrica para todos os canais de saída (Eq. 3.1). Exemplos de aplicação desta métrica expandida são as redes que possuem um canal de saída para cada estrutura segmentada e que cada canal pode ser ponderado, por exemplo, de acordo com o tamanho de sua estrutura.

$$wDICE = \sum_{i=1}^n w_i \frac{2|X \cap Y| + sf}{|X| + |Y| + sf} \quad (3.1)$$

onde  $w_i$  é o peso de cada canal de saída,  $X$  é máscara,  $Y$  é a predição do modelo,  $n$  é o número de canais e  $sf$  é o fator de suavização.

Um detalhe importante é que a imagem de fundo também é tratada como um canal de saída da CNN, portanto, o fundo é ponderado na função de custo. Isso pode ser benéfico para o processo de treinamento pois a CNN aprende o que não pertence ao conjunto de regiões de interesse.

Além de uma função de custo utilizando a métrica DICE, foi testada a métrica de entropia cruzada. Inclusive foi empregada uma métrica híbrida que ponderava simultaneamente a medida DICE e a entropia. Porém, em todos os experimentos em que se considerava a entropia o desempenho da rede foi pior já que muitos *voxels* eram classificados incorretamente, tanto dentro quanto fora das estruturas de interesse.

### 3.3 Ambiente computacional

Os experimentos foram feitos utilizando principalmente a biblioteca PyTorch (PASZKE *et al.*, 2019). Esta biblioteca, que é atualmente uma das mais populares, permite uma implementação simplificada e bastante flexível das redes convolucionais. Uma das vantagens desta biblioteca é a facilidade de utilização do NVIDIA CUDA (NVIDIA Corporation, 2010), que é uma extensão para programação paralela em placas gráficas.

Para o monitoramento dos estados da máquina utilizada no treinamento, e principalmente para acompanhar o treinamento em tempo real, foram utilizadas as ferramentas PyTorch Lightning e Neptune AI. Também foi utilizado o ambiente Jupyter Notebook (PEREZ; GRANGER, 2007) em linguagem Python 3.7 (ROSSUM; JR, 1995) para elaboração dos códigos, a biblioteca NiBabel (DEVELOPERS, 2016) para leitura das imagens de ressonância e as bibliotecas Numpy (OLIPHANT *et al.*, 2006) e cc3d (SILVERSMITH, 2024) para o pré e pós-processamento das imagens.

O *hardware* utilizado nos experimentos principais segue as seguintes especificações: Processador - Intel(R) Core(TM) i9-9900K CPU @ 3.60GHz; Memória RAM - 64GB DDR4; GPU (Placa gráfica) - NVIDIA GeForce RTX 2080 Ti com 11GB VRAM e 4352 CUDA Cores; Sistema operacional - Ubuntu 20.04.6 LTS.

## 3.4 Pós-processamento

O pós-processamento é uma etapa extra que tem como objetivo o refinamento da segmentação alcançada pela CNN. No caso específico desse trabalho, foram empregadas duas técnicas de pós-processamento, sendo a primeira a agregação de vistas, que é uma espécie de comitê de máquinas e análogo à fusão de dados na saída da CNN, e a segunda uma filtragem da segmentação por componentes conexos.

### 3.4.1 Agregação de vistas

Um dos artifícios mais eficientes para melhorar a qualidade da segmentação 3D a partir de predições 2D é o uso de agregação de vistas. Essa técnica elimina *voxels* que foram classificados de forma errada em uma determinada vista, mas foram classificadas de maneira correta em predições feitas em outras vistas. Por exemplo, um *voxel* que foi classificado errado na vista axial pode ter sua classificação correta nas predições das demais vistas, portanto, sendo corrigido na segmentação final.

Os erros de classificação são normalmente falsos positivos e ocorrem frequentemente nas regiões distantes da estrutura ou nas fatias em que a estrutura não aparece para aquela vista específica. É comum também ter falsos positivos e falsos negativos em fatias nas quais a aparição do tálamo é muito pequena, como nas extremidades, por exemplo. Em todos esses casos, a técnica de agregação de vistas resolve ou atenua consideravelmente a taxa de falhas.

Uma consideração importante é que cada uma das vistas necessita, obrigatoriamente, de um modelo de CNN dedicado. Isso significa, exemplificando, que um modelo treinado com imagens axiais só fará segmentações em imagens axiais, já que no caso de

predições em CNNs e imagens não correspondentes a qualidade da segmentação é completamente deteriorada. Portanto, isso implica que o método, quando aplicando a técnica de agregação de vistas, precisa de pelo menos dois modelos de CNN treinados, cada um para uma vista distinta.

Os pontos negativos dessa técnica são a necessidade de treinamento de outras CNNs para as demais vistas e a adição de passos extras no processo de inferência (Seç. 3.5), onde é necessário se atentar às transposições dos volumes para garantir o alinhamento das predições. No entanto, esses processamentos são simples e trazem ganhos significativos para o desempenho do método.

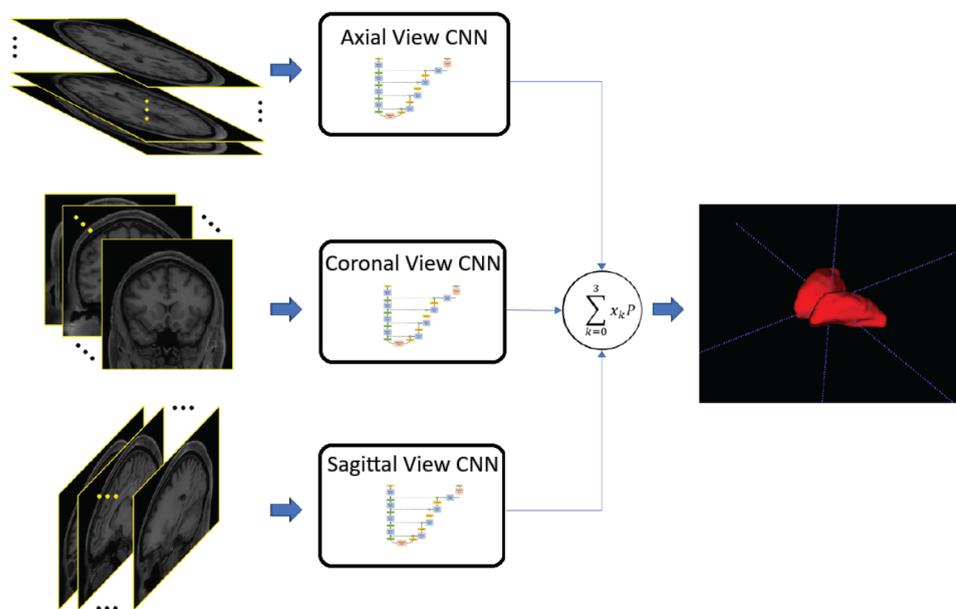


Figura 3.4 – Método de agregação de vistas: decisão dos *voxels* por maioria de votos de predições feitas em CNNs em múltiplas vistas.

### 3.4.2 Filtragem por volumetria

Mesmo obtendo bons resultados de segmentação com a rede individual ou com múltiplas redes no modo de vistas agregadas, existe a possibilidade de *voxels* esparsos que são erroneamente classificados como pertencentes à estrutura de interesse. Sendo assim, é comum se utilizar uma técnica que elimina estes *voxels*, melhorando significativamente o resultado da classificação. Neste trabalho, utilizamos como pós-processamento a aplicação de um filtro de componente conexo. Este filtro seleciona o maior, ou os maiores, componentes conexos para cada uma das estruturas, e elimina os pequenos artefatos que aparecem na predição.

No caso de predições feitas em imagens 2D, portanto com segmentações planares, o cálculo dos componentes conexos é necessariamente feito após o empilhamento das fatias. Com isso, a filtragem dos elementos espúrios é feita por um limiar de volume dos

componentes monolíticos, sobrando apenas aqueles que possuem volumes, ou quantidades de *voxels*, maiores ou iguais aos da estrutura de interesse. Nesse trabalho foi empregada a biblioteca cc3d (SILVERSMITH, 2024) para calcular os componentes conexos.

### 3.5 Inferência e segmentação

Para computar a segmentação, as imagens necessariamente sofrem as mesmas etapas do pré-processamento da etapa de treino, de modo que elas estejam padronizadas tais quais as imagens de treinamento.

Dentre todos os passos de pré-processamento (Fig. 2.7), ficam desconsideradas apenas as etapas que geram as máscaras padrão-prata que são utilizadas exclusivamente na etapa de treinamento das CNNs. Assim, o fluxograma do processo de inferência (Fig. 3.5) se difere do fluxograma de treinamento por utilizar apenas o ramo que padroniza as imagens, por utilizar apenas modelos de CNNs já treinados e por adicionar a etapa de pós-processamento.

Após a inferência da CNN com as imagens de entrada, existe uma etapa de pós-processamento que gera a segmentação final. Essa etapa conta com dois processamentos, sendo o primeiro a combinação de múltiplas inferências quando se tem o resultado de múltiplas CNNs (seja por multiplicidade de saídas ou por predição para cada uma das vistas) e o segundo por filtragem de *voxels* desconexos, ambos após o empilhamento das fatias de cada predição quando a CNN é 2D.

### 3.6 Resultados

A proposta desse trabalho é de definir métodos de segmentação do tálamo baseados em CNN que sejam tão bons quanto ou que superem a qualidades dos métodos já estabelecidos, como os baseados em atlas, com o diferencial de serem de execução rápida e com capacidade de lidarem com imagens de difusão. Portanto, diversos experimentos foram feitos para comparar o desempenho de outros métodos contra os métodos propostos e para medir o desempenho dos métodos conforme os tipos de imagens empregadas.

Apesar de outras arquiteturas terem sido experimentadas, como a V-Net (MILLETARI *et al.*, 2016), nnU-Net (ISENSEE *et al.*, 2021), e outras, essas não serão reportadas aqui pois são consideradas parte dos experimentos preliminares. Experimentos esses que ajudaram em diversas decisões de projeto como definição de tamanho de *patches*, qual otimizador usar, nível de pré-processamento, arquiteturas de CNN relevantes para o projeto, entre outras. Sendo assim, os resultados reportados nesse trabalho são restritos aos experimentos feitos com as arquiteturas QuickNAT e U-Net.

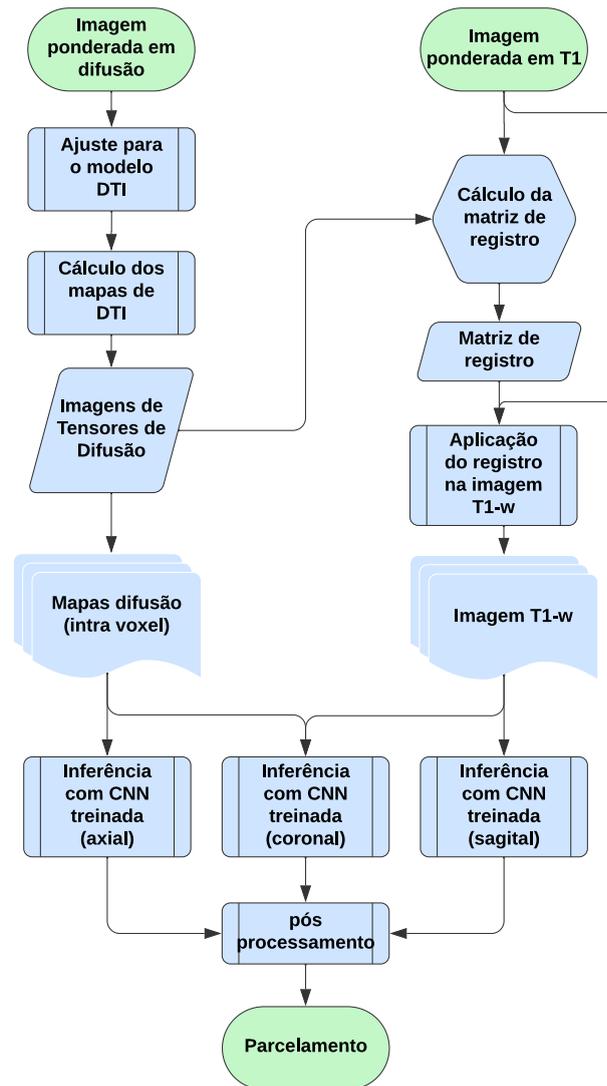


Figura 3.5 – Fluxograma do processo de inferência e segmentação do tálamo: imagens de entrada e saída (interfaces do método) em verde e etapas do método em azul.

Importante ressaltar que os resultados de desempenho das redes são computados com um grupo separado de imagens de testes, que não foram utilizados no treinamento.

### 3.6.1 Resultados com a QuickNAT

O primeiro grupo de experimentos foi conduzido tomando a QuickNAT como CNN, e tiveram o objetivo de avaliar o desempenho da arquitetura dadas diferentes combinações de mapas de entrada. Para isso, foi usado não somente a arquitetura da QuickNAT, mas também todo o ambiente de treinamento e inferência, que tem código aberto em repositório GitHub.

A QuickNAT é originalmente desenvolvida para segmentar diversas estruturas do cérebro em imagens T1w. Portanto, foram necessárias modificações para especializar a CNN no tálamo e comportar os múltiplos canais de entrada, permitindo o uso simultâneo

de imagens de difusão. Essas modificações foram feitas com a adição de canais nas primeira camada convolucional da CNN, de acordo com a quantidade de mapas que estavam sendo testados em cada um dos experimentos, salvo os casos de um único canal de entrada onde a arquitetura se manteve original.

Os dados utilizados nesses experimentos são um grupo de 180 sujeitos do conjunto de dados HCP. Todas as imagens foram processadas conforme os passos apresentados em Seç. 2.3, e as máscaras usada para o treinamento são as máscaras geradas pelo STAPLE, como apresentado em Seç. 2.4.2.1. As máscaras de métodos automáticos que pertencem a composição do padrão-prata são as segmentações do tálamo feitas pelo FreeSurfer, FSL (fsl anat) e volBrain (Fig.3.6).

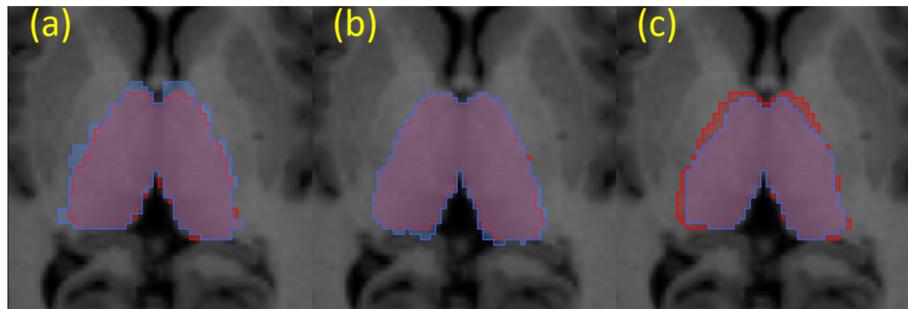


Figura 3.6 – Máscara de tálamo calculada por STAPLE (vermelho) sobre as máscaras de cada método individual (azul): (a) FreeSurfer; (b) FSL; (c) volBrain.

Para a avaliação do desempenho dos modelos e a comparação entre eles, foram utilizados 8 sujeitos, também do conjunto de dados HCP. Como máscaras de referência, para medir o desempenho, foram utilizadas as segmentações obtidas pelo FreeSurfer e o resultado do STAPLE.

O pós-processamento usou agregação de vistas, como descrito em Seç. 3.4, e foram consideradas as três vistas possíveis (axial, coronal e sagital). Porém não foi feita a filtragem por componente conexo, já que a agregação de vistas removeu praticamente toda classificação errada feita pelas CNNs, e esse é o procedimento padrão da QuickNAT.

Os resultados destes experimentos (Tab. 3.1), reportados em Pinheiro *et al.* (2021), mostraram, primeiramente, que a máscara STAPLE foi melhor representante do tálamo do que a máscara do FreeSurfer. Esse já era o resultado esperado visto que as redes foram treinadas para imitar o STAPLE, mas mostra que de fato as máscaras são relativamente diferentes. Outra conclusão importante que podemos tirar desses resultados é que os mapas de difusão (AD, FA, RD e MD) realmente são bons candidatos para resultados de segmentação, superando o modelo que só utilizava imagens T1w.

Além da comparação dos resultados obtidos pelas redes treinadas com o método proposto, podemos analisar os resultados relativos àqueles reportados pela Quick-

Tabela 3.1 – DICE médio da segmentação do tálamo no conjunto de teste para cada combinação de entrada da CNN (linhas) e *Máscara de referência* (colunas). A segmentação é a combinação de 3 modelos CNN, um para cada orientação de volume de MRI (agregação de 3 vistas).

Mapas de entrada	Máscara de referência	
	FreeSurfer	STAPLE
T1	0.7456	0.8643
AD, FA, RD, MD	0.7670	<b>0.8749</b>
AD, FA, RD, MD, B0, T1	0.7711	0.8693

NAT para a estrutura específica do tálamo (Tab.3.2).

Tabela 3.2 – Comparação entre métodos de segmentação do tálamo

Método	Dado de entrada	N imagens	DICE
QuickNAT (reportado)	T1w	581	0.87
QuickNAT (retreinada)	T1w	180	0.75
QuickNAT multimodal (retreinada)	T1w e difusão	180	0.87
(GLAISTER <i>et al.</i> , 2017) (random forest)	difusão	9	0.88

O primeiro detalhe a ser verificado é que na versão da QuickNAT que foi retreinada com os dados específicos desse trabalho (180 sujeitos do HCP), o resultado ficou consideravelmente inferior, sendo 0,87 contra 0,75 de DICE. Podemos atribuir essa diferença principalmente ao deficit de imagens no treinamento, já que a QuickNAT original foi treinada com mais do que o triplo de imagens, e talvez a qualidade de máscara usada, já que o trabalho original faz um ajuste fino para máscaras padrão-ouro.

Outro ponto importante é que garantidamente as imagens de difusão trouxeram ganhos na qualidade da segmentação. Especificamente, a adição dos mapas de difusão fez com que o DICE saltasse de 0,75 para 0,87 nos modelos retreinados, ou seja, um ganho de 0,12 (ou 16%) no DICE. Com certeza esse ganho tão expressivo não se reproduziria exatamente no trabalho original da QuickNAT, porém, garantidamente teríamos um DICE superior se mapas de difusão fossem aplicados.

Por fim, podemos observar (Fig. 3.7) que a segmentação do tálamo feita pela CNN é muito próxima do padrão-prata, diferindo apenas em alguns *voxels* na borda da estrutura. Esse resultado mostra que, de fato, a máscara usada durante o treinamento (padrão-prata) e os dados de difusão têm informações para explicar muito bem o formato do tálamo.

Além dos resultados quando se usa imagens de difusão se apresentarem superiores, podemos notar que, de fato, as bordas do tálamo são mais contrastantes nas imagens de difusão do que são nas imagens T1w. Isso é bastante evidente na fronteira do tálamo com a capsula interna (Fig. 3.7 - T1 e FA), onde, em imagens T1w, a capsula interna é

quase indiferenciável do tálamo, mas em FA apresenta valores bem distintos daqueles no interior do tálamo.

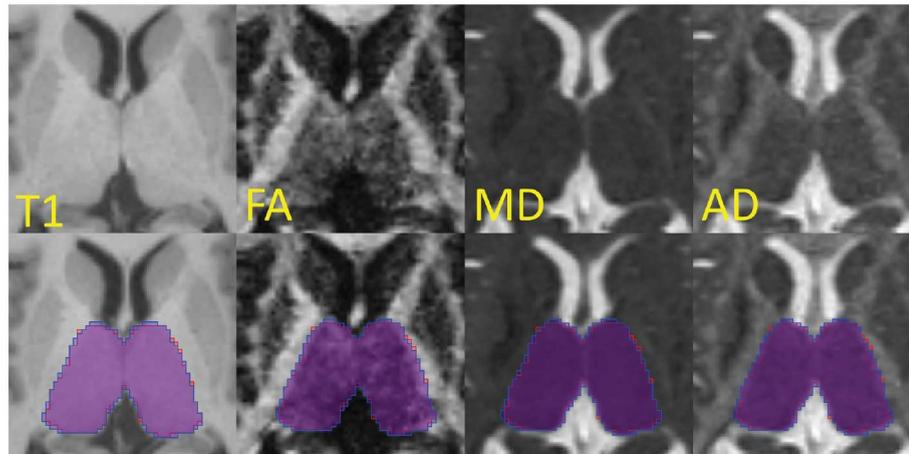


Figura 3.7 – Resultados qualitativos da segmentação do tálamo: Máscara STAPLE (magenta) e predição (azul) pelo modelo treinado com dados de difusão. Máscara e segmentação apresentadas sobre fatias axiais de alguns mapas de entrada.

### 3.6.2 Resultados com a U-Net

A U-Net, como já mencionado anteriormente, é arquitetura de CNN mais popular para imagens médicas, sendo inclusive base para diversos outros trabalhos como a QuickNAT e nnU-Net. Portanto, essa arquitetura é tomada como a principal para os experimentos aqui reportados.

Os experimentos foram conduzidos com o conjunto de dados do HCP, contendo imagens de 1063 sujeitos. Desses, 963 foram usados para treinamento, 40 foram usados para ajuste fino, 20 usados para avaliar o desempenho e 40 ficaram reservados para uso futuro. Tanto os 20 sujeitos de teste quanto os 40 destinados ao ajuste fino foram segmentados manualmente, sendo considerados padrão-ouro.

O primeiro experimento é a análise do desempenho do método quando se considera ou não a agregação de vistas (Fig.3.8). Esse experimento consiste na medida de desempenho na segmentação do tálamo em modelos treinados para cada uma das vistas e com a agregação de todas as vistas em uma predição, seguindo o método descrito em Seq. 3.4.1. Podemos verificar que das vistas individuais, a vista sagital apresenta o pior desempenho na medida DICE e a axial apresenta o pior em distância Hausdorff, enquanto a coronal apresenta o melhor em ambas as medidas. No entanto, a agregação das três vistas é o melhor dos casos, sendo essa etapa definida como padrão no método de segmentação do tálamo.

O próximo experimento avalia o desempenho do método para modelos com ajuste fino para o padrão-ouro e sem esse ajuste (Tab. 3.3). A variação de desempenho foi

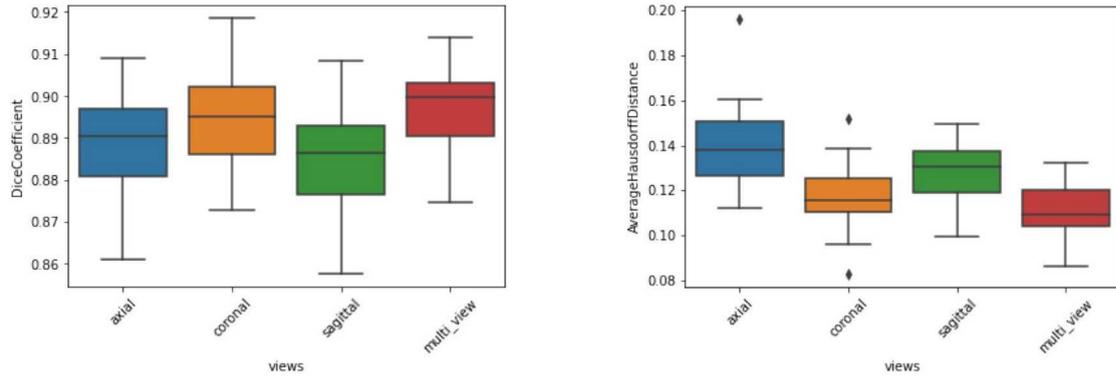


Figura 3.8 – Desempenho para modelos treinados em diferentes vistas: resultados de DICE e distância Hausdorff média para vistas individuais e para o modelo com agregação de todas as vistas simultaneamente.

Tabela 3.3 – Resultados no conjunto de teste antes e depois do ajuste fino: coeficiente DICE (DC), distância média de Hausdorff (AHD) no tamanho do *voxel* e seus respectivos desvios padrões (dp).

Mapas de entrada	Sem ajuste fino		Com ajuste fino	
	DC (dp)	AHD (dp)	DC (dp)	AHD (dp)
FA, MD, RD, AD, T1	0.7775 (0.022)	0.2320 (0.025)	0.8970 (0.012)	0.1102 (0.013)
FA, MD, RD, AD	0.7721 (0.023)	0.2389 (0.026)	0.8904 (0.010)	0.1158 (0.010)
T1	0.7676 (0.024)	0.2440 (0.027)	0.8904 (0.015)	0.1158 (0.016)

avaliada para diferentes combinações de mapas de entrada das CNNs. Como era esperado, os resultados depois do ajuste fino para as máscaras manuais realmente foram superiores, indo de um DICE de 0,777 para 0,897. A distância Hausdorff, que é uma medida das distâncias entre as cascas da segmentações, também teve uma melhora significativa, sendo praticamente reduzida à metade.

Importante ressaltar que o ajuste fino altera somente a segunda metade da CNN, de forma que apenas as camadas que possuem função de reconstrução da imagem de saída sejam alteradas, e a parcela que extrai as informações dos dados de entrada se mantém intacta. Experimentos preliminares mostraram que quando toda a rede é retreinada nessa etapa, a otimização do ajuste fino fica prejudicada, demorando mais para atingir a convergência, e que os resultados são inferiores.

Tomando os modelos com ajuste fino e agregação das três vistas como padrão, foi feita uma extensiva comparação entre diversas combinações de mapas de entrada (Tab. 3.4 e Fig.3.9). Essas comparações foram feitas no conjunto de testes com 20 sujeitos de anotações feitas por especialista.

Os resultados apontam que o modelo treinado em imagens ponderadas em T1 produziu os piores resultados para cada métrica computada. Embora ainda comparável com o estado da arte, os piores coeficientes DICE foram obtidos enquanto usando imagens

Tabela 3.4 – Métricas de avaliação para cada combinação de mapa de entrada CNN no conjunto de teste: DC—coeficiente DICE; FNR—razão de falsos negativos; FPR—razão de falsos positivos; JAC—coeficiente de Jaccard; VS—similaridade volumétrica; AHD—distância média de Hausdorff (em tamanho de *voxel*). Os melhores e piores resultados de DC e AHD em negrito.

Mapas de entrada	DC (SD)	FNR	FPR	JAC	VS	AHD (SD)
T1	<b>0.8837 (0.0154)</b>	<b>0.1649</b>	<b>0.0596</b>	<b>0.7920</b>	<b>-0.1193</b>	<b>0.1230 (0.0157)</b>
FA	0.8945 (0.0089)	0.1390	0.0677	0.8092	-0.0795	0.1116 (0.0097)
MD	0.8931 (0.0122)	0.1326	0.0776	0.8071	-0.0615	0.1143 (0.0133)
RD	0.8935 (0.0115)	0.1364	0.0724	0.8077	-0.0715	0.1137 (0.0134)
AD	0.8925 (0.0131)	0.1419	0.0680	0.8061	-0.0829	0.1143 (0.0142)
FA, T1	0.8943 (0.0107)	0.1373	0.0700	0.8090	-0.0754	0.1119 (0.0112)
FA, MD	0.8943 (0.0113)	0.1327	0.0754	0.8089	-0.0638	0.1126 (0.0136)
MD, T1	0.8882 (0.0136)	0.1462	0.0716	0.7992	-0.0842	0.1185 (0.0132)
FA, MD, RD, AD	0.8904 (0.0100)	0.1509	0.0623	0.8026	-0.0994	0.1158 (0.0105)
FA, MD, RD, AD, T1	<b>0.8970 (0.0116)</b>	<b>0.1182</b>	<b>0.0853</b>	<b>0.8135</b>	<b>-0.0367</b>	<b>0.1102 (0.0129)</b>

T1w como o único canal de entrada (DC de 0,8837). A mesma consideração é válida para a distância média de Hausdorff (AHD de 0,1230).

Por outro lado, os modelos treinados usando mapas de DTI apresentaram melhorias na métrica DICE. De fato, o pior resultado apresentado por um modelo treinado com um único mapa de difusão foi um DICE de 0,8925 (AD), e o único modelo de canal combinado que produziu um coeficiente DICE abaixo de 0,89 foi aquele treinado em MD e T1. Além disso, todos os modelos combinados produziram resultados superiores aos do modelo usando apenas imagens T1w.

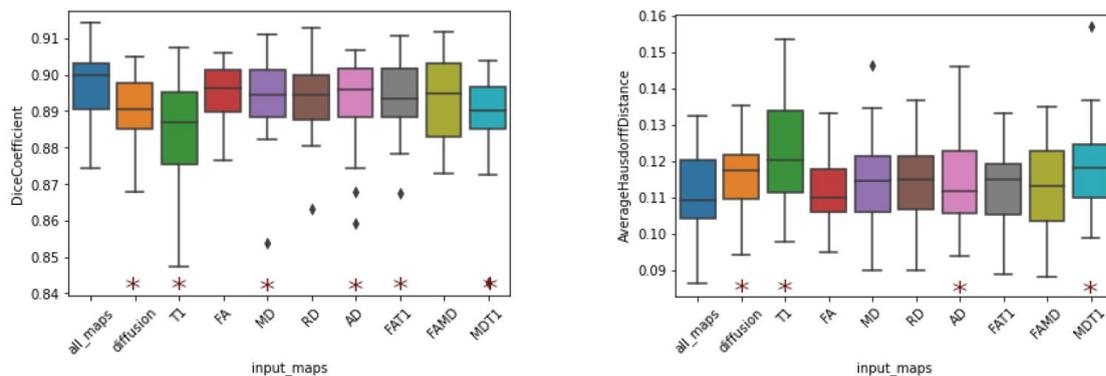


Figura 3.9 – Distribuição Dice e distância Hausdorff média para cada combinação de mapa de entrada avaliada no conjunto de teste. Estrela (\*) denota significância estatística ( $p < 0.05$ ) quando comparado com o melhor modelo (all\_maps). Todos os valores de  $p$  foram estimados usando um teste de postos sinalizados de Wilcoxon bilateral e apenas comparando com o melhor modelo.

Os resultados que mais se destacam são os obtidos pelo modelo treinado na combinação de todos os mapas disponíveis, T1w e mapas escalares de DTI (FA, MD, RD, AD e T1). Essa combinação, que pode ser considerada como a adição de T1 à

combinação de todos os mapas de difusão, melhorou o DICE de 0,8904 para 0,8970. No entanto, se considerarmos a adição dos mapas de difusão à T1, o DICE aumenta de 0,8837 para 0,8970, representando uma melhoria muito mais significativa. Independentemente da interpretação, essa combinação produziu o melhor desempenho em termos de métricas de avaliação, indicando que a segmentação que usa todos os canais de entrada apresenta uma concordância maior com o padrão-ouro.

Outra consideração importante é sobre o desvio padrão nos resultados. Essa medida reflete a estabilidade da segmentação para cada mapa e para cada combinação. Novamente, notamos que os modelos treinados em imagens ponderadas em T1 exibiram a maior dispersão nas métricas DICE e distância Hausdorff. Já sobre os mapas de difusão, entre os modelos treinados com um único canal de entrada, aquele treinado com FA produziu os melhores resultados de média e dispersão. Esse resultado faz sentido quando analisamos qualitativamente as imagens de cada canal e podemos constatar que o mapa de FA é o que mais apresenta contraste em todas as bordas do tálamo simultaneamente.

Os resultados apresentados nessa sessão foram publicados na revista "*Applied Sciences*" (PINHEIRO *et al.*, 2023). Além disso, com a proposta de ciência aberta, existe o código disponível para reprodução e propostas de melhorias em um repositório GitHub <[https://github.com/MICLab-Unicamp/thalamus\\_benchmark\\_diffusion](https://github.com/MICLab-Unicamp/thalamus_benchmark_diffusion)>. Os dados pré-processados, inclusive com todas as máscaras, estão disponíveis em <<https://www.ccdataset.com/thalamus-benchmark>>, e um *challenge* está aberto para que competidores possam tentar superar os resultados atingidos em <<https://codalab.lisn.upsaclay.fr/competitions/8329>>.

## 4 Parcelamento do Tálamo

Diferentemente do problema de segmentação do tálamo, onde o interesse é segmentar a estrutura como um todo, no problema do parcelamento o propósito é delinear as bordas dos núcleos que constituem o tálamo.

Um dos maiores motivadores para o desenvolvimento de um método capaz de segmentar os núcleos do tálamo é a necessidade de localização exata de regiões que podem ser alvos de intervenções cirúrgicas, como DBS e talamotomia.

Esse problema é ainda mais desafiador que a segmentação já que todos os problemas anteriores, como tamanho das estruturas, interferência por volume parcial, contraste reduzido e limitação das máscaras, são intensificados no caso do parcelamento do tálamo.

Apesar do objetivo ser distinto do problema de segmentação do tálamo (Cap. 3), muitas das decisões de projeto, elementos de arquitetura, padronizações dos dados e técnicas de treinamento podem ser reaproveitadas para a tarefa de parcelamento do tálamo. Alguns exemplos de processos reaproveitados são: pré-processamento e normalização espacial com registro das imagens. Sendo assim, esse capítulo será focado principalmente nos detalhes de implementação e técnicas aplicadas que são distintas daquelas empregadas na tarefa de segmentação.

### 4.1 Núcleos do tálamo

Na literatura, são encontradas diversas formas de se fazer o parcelamento do tálamo levando em conta quantidades diferentes de núcleos e agrupamentos desses núcleos. A maioria dos trabalhos divide cada lado do tálamo entre 5 e 10 conjuntos, sendo os conjuntos compostos de 1 a 10 núcleos. Tomando em conta os núcleos individualizados, são contabilizados um total de 25 em cada hemisfério do tálamo (Fig. 2.9). Essa quantidade de núcleos implica que as estruturas são consideravelmente pequenas, especialmente para a escala típica de resolução das imagens de ressonância largamente utilizadas nos conjuntos de dados.

Devido ao tamanho pequeno, muitas vezes nem sendo perceptíveis nas MRIs, e pela dificuldade de se estabelecer um padrão-ouro para a segmentação desses pequenos núcleos, as máscaras são definidas em conjuntos de núcleos. Alguns dos trabalhos mais relevantes de segmentação das regiões internas do tálamos agrupam os núcleos em 5 ou 10 regiões (Tab. 2.3) (IGLESIAS *et al.*, 2018).

## 4.2 Arquitetura da CNN para parcelamento

Devido ao tamanho reduzido das estruturas de interesse na tarefa de parcelamento do tálamo, surgiu a necessidade de uso exclusivo de uma arquitetura de CNN tridimensional. São dois os pontos principais que resolvem ou atenuam os problemas mencionados: o aumento da quantidade de *voxels* por estrutura que são considerados durante o treinamento e a predição; relação de vizinhança espacial, melhorando a diferenciação entre os tecidos das estruturas reduzidas.

Com essas necessidades, foi feita uma reimplementação da U-Net utilizada no capítulo 3, aumentando a dimensão de cada um dos canais de entrada da arquitetura da rede de forma que volumes tridimensionais pudessem alimentar a CNN.

Outra modificação relevante é a adição de grupos de canais de entrada (Fig 2.10). Isso se deu pela aplicação de mapas mais complexos que consideram uma parcela da informação direcional dos *voxels* das imagens de difusão, o que não era considerado nos experimentos da tarefa de segmentação do tálamo como um todo.

Nos experimentos de parcelamento do tálamo, foram utilizados 4 grupos de canais de entrada na CNN. O primeiro é de imagens escalares simples, onde apenas uma imagem T1w foi adicionada. O segundo grupo é composto por mapas escalares de imagens de difusão (AD, RD, FA e MD). O terceiro grupo é o conjunto das imagens de gradiente tensorial, onde os canais são o produto escalar, a norma de Frobenius e o Log Euclidiano. Por fim, o último grupo é composto por uma coleção de mapas de divergente tensorial computados para 32 direções de interesse não correlacionadas.

Foram feitos experimentos para determinar a quantidade de convoluções na camada de fusão da arquitetura proposta. Os casos extremos testados são com apenas um filtro convolucional e com 32 canais convolucionais. Vale ressaltar que a quantidade de imagens de saída dessa primeira camada convolucional é equivalente para cada um dos grupos de imagens, de modo que as informações inerentes a cada classe de imagem de entrada seja tratada de modo igualitário para as camadas seguintes da CNN. No primeiro caso, com apenas um canal de saída por grupo, a informação é muito reduzida, limitando a quantidade de informações que passa para a parte seguinte da arquitetura. Já no segundo caso, quando foram utilizados 32 filtros convolucionais, que é justamente a quantidade de canais no maior grupo de entrada, a rede recebe mais informações das imagens de entrada e, portanto, obtendo mais aproveitamento dos dados utilizados.

Assim, a arquitetura tem um desenho onde as imagens são alimentadas na rede de forma tridimensional, os canais de entradas são fundidos por classes e inseridos em uma U-Net 3D, e o final da rede tem um conjunto de múltiplas saídas, sendo cada uma delas destinada para a classificação de núcleos específicos. Por fim, outro detalhe

relevante de implementação é que um dos canais de saída é destinado a tudo aquilo que não é o tálamo. Principalmente no tratamento de estruturas pequenas, é necessário que a CNN seja alimentada com a informação do que não é a estrutura de interesse. Caso contrário, a CNN poderia classificar elementos de fundo como pertencentes à alguma das estruturas.

### 4.3 Técnicas de treinamento de CNN 3D

Diferente da tarefa de segmentação do tálamo, o parcelamento do tálamo requer que as arquiteturas de rede neural identifiquem padrões de dimensão consideravelmente menores. Como a relação de vizinhança dos filtros convolucionais bidimensionais são menores do que aquelas de mesmo dimensão volumétrica, informações valiosas são desperdiçadas ou nem consideradas quando se segmenta estruturas tão pequenas com restes 2D.

Outra consideração importante para o emprego de arquiteturas 3D para a resolução desse problema é o fato de estruturas pequenas raramente aparecerem nas imagens 2D quando se utiliza o volume completo. Por exemplo, se um núcleo tem dimensão de 3mm e estamos trabalhando com fatias de 1mm, esse núcleo tem chance de aparecer em somente 3 fatias (ou 4 quando se considera o volume parcial), o que gera escassez de amostras dessa estrutura específica na etapa de treinamento.

#### 4.3.1 Treinamento usando *Patches*

Diferente do problema de segmentação do tálamo, no parcelamento não precisamos apenas considerar amostras de uma única estrutura, preocupando em balancear somente a proporção entre objeto de interesse e imagem de fundo. No caso de múltiplas estruturas de tamanho reduzido, é necessário também se preocupar com a quantidade de amostras que contenham cada um dos núcleos do tálamo.

Considerando a necessidade dessas proporções, os treinamentos passam a ser feitos com lotes de imagens que contenham a mesma quantidade de *patches* centralizados em cada um dos núcleos do tálamo, e claro, com uma certa quantidade de amostras de imagem de fundo para que a CNN tome conhecimento do que não é uma estrutura de interesse.

Outra diferença em relação à técnica de *patches* aplicadas no problema de segmentação é que agora os *patches* necessariamente são tridimensionais devida a natureza da arquitetura 3D da CNN. Portanto, no problema de parcelamento, os *patches* tridimensionais são selecionados com o critério de que, dadas as proporções por cada núcleo, os centros são pertencentes a cada um dos núcleos ou combinação deles.

### 4.3.2 Abordagem multitarefas ou multi estrutura

No contexto de parcelamento do tálamo, tratamos a abordagem como multi tarefa já que a CNN aprende simultaneamente o que é cada um dos núcleos e o que é a imagem de fundo. Por exemplo, quando estamos considerando o agrupamento dos núcleos em 10 conjuntos, a CNN possui 11 canais de saída, sendo o primeiro dedicado para o que não é o tálamo e os demais dedicados para cada um dos núcleos ou conjunto de núcleos.

### 4.3.3 Função de custo

No problema de parcelamento, foi utilizada a mesma função de custo definida no problema de segmentação (Eq. 3.1). Isso foi possível, pois essa função de custo estava generalizada para múltiplas saídas.

No caso específico do parcelamento, cada núcleo (ou agrupamento de núcleos) é tratado como um canal de saída, sendo que núcleos dos dois hemisférios são tratados como um único canal antes da etapa do pós-processamento. Além disso, o fundo também é tido como um canal independente, pois é importante que a CNN aprenda diferenciar o que não se encaixa em nenhum dos núcleos. Com isso, a função de custo é uma ponderação entre o DICE de cada um dos núcleos e o plano de fundo.

## 4.4 Pós-processamento

Como a arquitetura da CNN aplicada no problema do parcelamento é tridimensional, não faz sentido o uso de agregação de vistas, como é feito no caso da segmentação do tálamo. Com a arquitetura sendo 3D, os dados entram na CNN de modo volumétrico, e não por fatias.

Entretanto, assim como no caso de predições 2D, existem *voxels* que são classificados erroneamente na região exterior ao tálamo. Para lidar com esses falsos positivos, é aplicado um filtro por componente conexo no volume da predição.

A filtragem é feita independentemente para cada um dos canais de saída, que são referentes a cada um dos núcleos. Como cada uma das sub-regiões do tálamo é distinta das demais em relação a tamanho, a filtragem é condicionada ao volume típico de cada núcleo, sendo eliminados aqueles agrupamentos que possuem valores fora da faixa (acima ou abaixo) de um determinado tamanho típico para aquela estrutura.

Um caso particular seria o grupo medial que por vezes faz conexão com sua parte oposta no outro hemisfério devida sua localização central. Quando conectada, essa estrutura teria um único rótulo, porém o mascaramento pela segmentação prévia do tá-

lamo faz com que seja possível e assertiva a separação por hemisférios desse grupo, viabilizando a filtragem por faixa de volumes.

## 4.5 Inferência e parcelamento

Assim como no caso da segmentação, as imagens necessariamente sofrem as mesmas etapas do pré-processamento da etapa de treino, de modo que elas estejam padronizadas. Porém existem algumas diferenças importantes entre o processos de segmentação e parcelamento, principalmente na etapa de pós-processamento.

O fluxograma de parcelamento (Fig. 4.1) se diferencia do fluxograma de segmentação (Fig. 3.5) por conter um ramo a mais, justamente o que contém a CNN que faz o parcelamento. O bloco da segmentação é mantido, já que a segmentação do tálamo, que diferencia os dois hemisférios, é usada com máscara para distinguir os núcleos de cada lado do tálamo.

Sendo assim, o bloco de pós-processamento, faz combinação das duas inferências e a filtragem, entregando a segmentação final com um rótulo diferente para cada um dos núcleos (ou agrupamentos) de cada hemisfério. Vale ressaltar que não há a necessidade de empilhamento ou a combinação da predição de múltiplas vistas já que as imagens de entrada e saída da CNN são 3D.

## 4.6 Incertezas de um modelo de CNN

Na tarefa de parcelamento do tálamo, sofremos ainda mais pelo problema de falta de dados rotulados com máscaras padrão-ouro quando comparado com o problema de segmentação dessa estrutura. Isso se dá não só pelo fato do contraste entre os núcleos serem ainda menor mas também pelo fato da segmentação manual ser extremamente dificultada já que as estruturas possuem bordas menores e o problema do volume parcial é significativamente aumentado.

Devido a falta do padrão-ouro, a avaliação dos resultados atingidos pelo modelo é pouco significante, de modo que não podemos avaliar de forma direta se o modelo é capaz de segmentar de forma correta a estrutura biológica.

Para contornar essa situação, uma das possibilidades é a avaliação das incertezas que estão embutida nos modelos e nas imagens de entradas das redes neurais. No caso do trabalho atual, foram avaliadas duas formas de incerteza: incerteza epistêmica e incerteza estatística.

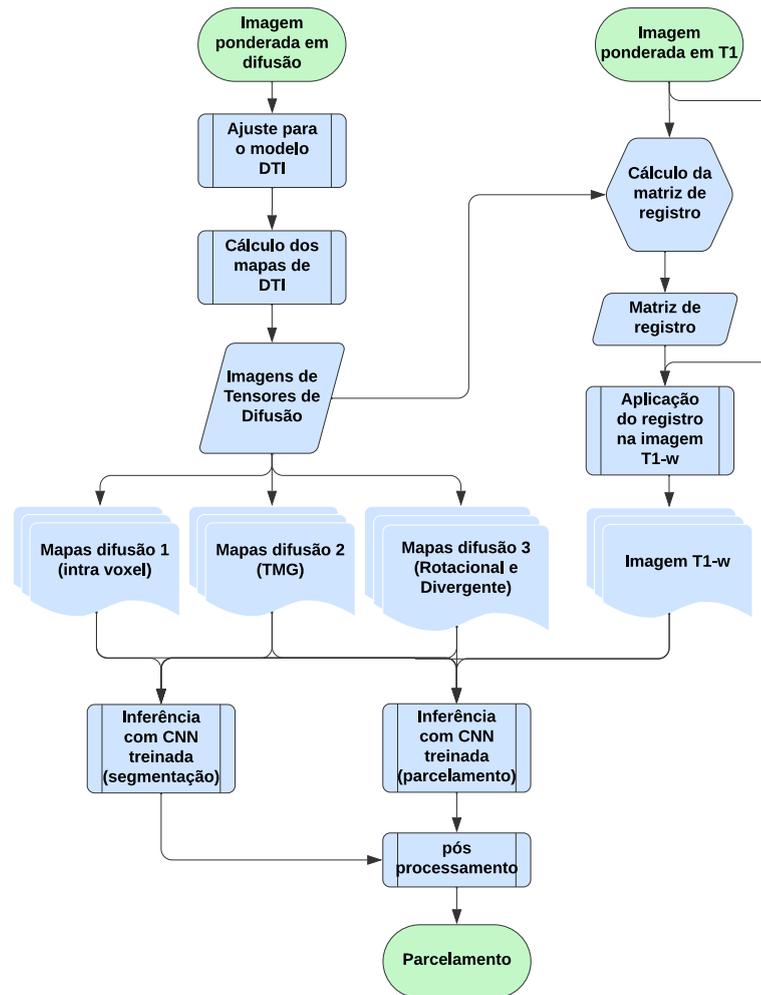


Figura 4.1 – Fluxograma do processo de inferência e parcelamento do tálamo: imagens de entrada e saída (interfaces do método) em verde e etapas do método em azul.

#### 4.6.1 Incerteza epistêmica

A incerteza epistêmica é definida pela incerteza intrínseca ao modelo treinado, portanto, essa é uma medida de quanto o modelo é confiável do ponto de vista de resultado atingido (KENDALL; GAL, 2017).

Uma das formas de calcular essa incerteza é com ciclos de perturbação aleatórias do modelo e predições. Essa técnica é chamada de predição de Monte Carlo pois é inspirada em jogos de cassino em que cada jogada contém elementos de incerteza (GAL; GHARAMANI, 2016; KENDALL *et al.*, 2015; WANG *et al.*, 2019).

Na prática, o modelo é perturbado pela técnica de *dropout*, onde uma porcentagem dos pesos do modelo é aleatoriamente levado a zero e os resultados das predições perturbadas são guardados para análises estatísticas posteriores.

No caso específico desse trabalho, foram utilizadas predições com *dropout* va-

riando de 0,1% a 10% nas camadas convolucionais dos blocos da U-Net e as análises estatísticas, executadas para cada um dos *voxels*, foram a média, desvio padrão e variância.

#### 4.6.2 Incerteza estatística

A incerteza estatística é definida pela confiança do modelo em relação à variações das imagens que alimentam o algoritmo (KENDALL; GAL, 2017). Assim, a incerteza estatística é uma medida da imunidade da rede à variações ou distorções (naturais ou artificiais) das imagens apresentadas em suas entradas.

Para calcular a incerteza, similarmente ao que é feito na incerteza epistêmica, são feitas diversas predições com elementos aleatórios e os resultados são armazenados para o cálculo das estatísticas para cada um do *voxels*. Porém, nesse caso, os elementos de incerteza são inseridos nas imagens de entrada da CNN (WANG *et al.*, 2019).

Este trabalho utilizou variações de intensidade de *voxel* com ruídos Gaussianos, distorções rígidas e distorções elásticas. Para cada predição, as distorções são constituídas a partir da imagem original, sendo que o ruído (média 0 e desvio padrão de 0,01) tem 50% de probabilidade de acontecer e as distorções geométricas têm 70% de chance de serem aplicadas, sendo que metade das transformações geométricas são rígidas e metade são elásticas. Outro detalhe relevante é que as perturbações geométricas possuem transformações aleatórias (distribuição gaussiana) limitadas a distâncias de 2 *voxels*, pois isso é mais do que suficiente para contemplar qualquer incerteza referente ao volume parcial, ao mesmo tempo que não impede o empilhamento da coleção de predições que são consideradas nas estatísticas da incerteza.

### 4.7 Resultados

O núcleos foram agrupados tanto na etapa de treinamento quanto para os testes e avaliação dos resultados. Para a resolução e tipo dos dados disponíveis, e considerando o tamanho de alguns dos menores núcleos, é impraticável o parcelamento em todos os possíveis 25 núcleos. De fato, até mesmo a ferramenta que registra o atlas falha para encontrar os menores núcleos. Foram feitos experimentos com agrupamento de 5 e de 10 núcleos, e os resultados foram similares entre si. Como a maioria dos trabalhos reporta pelo menos 7 núcleos (Tab. 1.2), os experimentos foram mais concentrados em 10 regiões (Tab. 2.3), com os seguintes grupos: Anterior, Dorsal, Lateral-Rostral, Lateral-Caudal, Intralaminar, Intralaminar-Posterior, Medial, Pulvinar, Lateral Geniculado e Medial Geniculado.

Tabela 4.1 – Lista dos experimentos de parcelamento: combinação de canais de entrada para cada um dos experimentos; DICE médio entre todos os sujeitos do conjunto de testes e todos os núcleos do tálamo.

Experimento	Canais de entrada da CNN	N filtros fusão	DICE médio	DICE médio pós-proc.
1	DTI divergente, DTI escalares, DTI TMG, T1	16	0,675075	0,872855
2	DTI divergente, DTI escalares, DTI TMG, T1	32	0,786669	0,874059
3	DTI divergente, T1	32	0,751830	0,871654
4	DTI divergente	32	0,515117	0,814254
5	DTI escalares, T1	32	0,733735	0,835424
6	DTI escalares	32	0,602957	0,878069
7	FA	32	0,680983	0,875463
8	T1	32	0,730710	0,828739
9	DTI TMG, T1	32	0,575022	0,877548
10	DTI TMG	32	0,731767	0,832800

DTI escalares	evaluel, FA, RD, MD
DTI TMG	TMG_frob, TMG_logE, TMG_prod

O primeiro grupo de experimentos reportado é um comparativo do desempenho das arquiteturas de CNN para cada combinação de mapas de entrada (Tab. 4.1). Foram feitas 10 combinações de mapas de entrada, sendo essas combinações envolvendo imagens T1w, escalares *intravoxel* de DTI, mapas de gradiente morfológico tensorial (TMG), e o mapa de divergente tensorial computado para 32 direções pré determinadas. Esses experimentos também foram utilizados para avaliar a diferença do desempenho das CNNs com a adição do pós-processamento de filtragem por componentes conexo (Fig. 4.2).

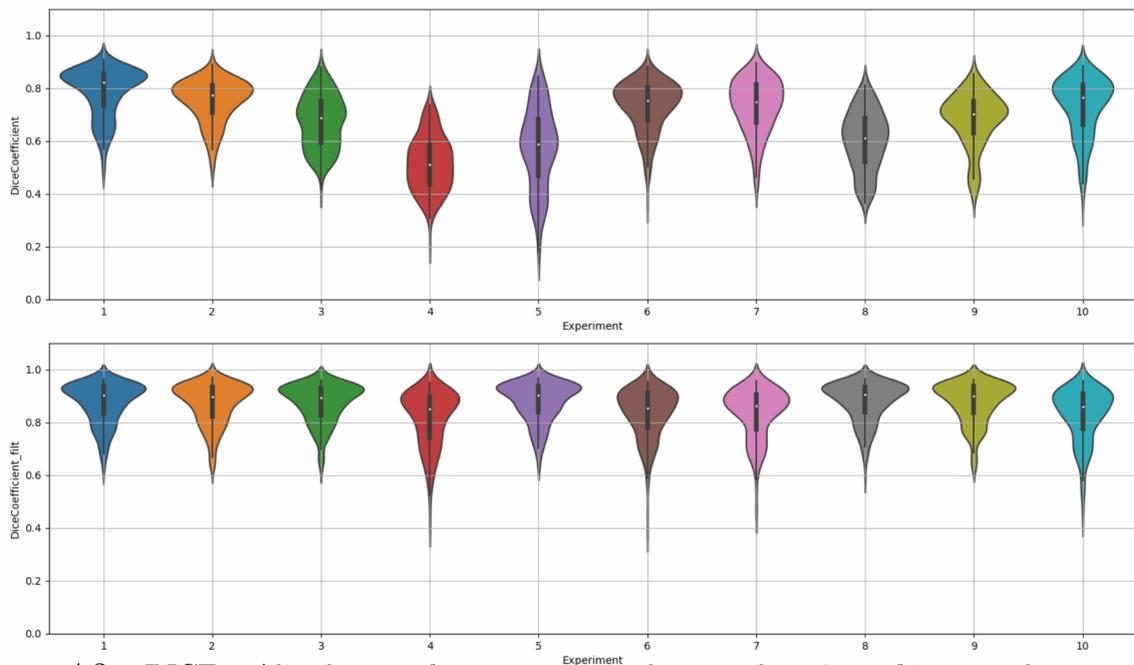


Figura 4.2 – DICE médio do parcelamento para cada uma das 10 combinações de mapas de entrada (Tab 4.1): sem o pós-processamento por componente conexo (superior) e com aplicação do pós-processamento (inferior).

Os resultados mostram, primeiramente, que a quantidade de filtros da camada

de fusão de dados é um fator importante a ser considerado. Quando comparamos os experimentos 1 e 2, onde a única diferença são os filtros convolucionais, podemos notar que há uma diferença significativa dos resultados entre o modelo que tem 16 e o que tem 32 filtros convolucionais, mostrando que a CNN com mais filtros na etapa de fusão consegue segmentar melhor as estruturas. Isso vem, provavelmente, do fato de não termos uma redução drástica, logo nas primeiras camadas, das informações de microestrutura presentes em cada uma das imagens de entrada da CNN.

O próximo ponto a ser considerado é sobre o pós-processamento do método. A filtragem de falsos positivos elevou consideravelmente o valor médio do DICE em todos os casos. Por exemplo, nos experimentos 4 e 9, os valores eram inferiores a 0,6 e passaram para acima de 0,8 após a filtragem. Como os núcleos são pequenos, poucos *voxels* classificados de forma errada fazem grande diferença na métrica DICE. Portanto, o pós processamento, que elimina espúrios por componente conexo, é indispensável para o método de parcelamento do tálamo. Além de melhorar a média do DICE no conjunto de testes, melhora também a dispersão (Fig. 4.2), indicando que o resultado tem melhor qualidade e mais robustez.

O último ponto a ser considerado nesses experimentos é o comparativo de desempenho do método para cada combinação de imagens de entrada. É nítido que independentemente de quais imagens são empregadas, o resultado do DICE médio de todas as estruturas no conjunto de testes é superior a 0.814, o que pode ser considerado um excelente resultado para esse problema específico. O pior dos casos foi o experimento 4, que usou apenas as imagens de divergente derivadas de DTI, atingindo um DICE médio de 0.814. Esse resultado deve-se ao fato de que esse mapa específico contém exclusivamente informações de direcionalidade relativa das fibras do cérebro, ignorando qualquer valor escalar absoluto da microestrutura. Já o segundo pior resultado veio do modelo que usa apenas imagens T1w, com DICE de 0.829. Nesse caso, o que justifica o valor é esse mapa ter características opostas ao mapa de divergente, ignorando completamente as informações relacionadas as fibras. Em relação aos melhores resultados, podemos verificar que os mapas escalares de DTI apresentam o melhor desempenho, chegando em uma média de 0.878. Além desse, as combinações de DTI TMG com T1, Divergente com T1, e todos os mapas simultaneamente atingiram valores de DICE médio acima de 0,871, indicando que o método se beneficia da informação mutua proveniente de imagens de origens distintas.

Outro experimento avaliou o desempenho do método para cada um dos núcleos considerados (Tab. 4.2). Foi selecionado o melhor modelo, que teve o melhor desempenho nos experimentos que fizeram o comparativo (Experimento 6 - DTI escalares), e usado para extrair as medidas dos núcleos. A primeira constatação é que, mais uma vez, o pós-processamento é relevante nessa aplicação, já que todos os núcleos apresentaram melhoria

Tabela 4.2 – DICE e volume médio de cada núcleo do tálamo no conjunto de testes usando o melhor modelo de CNN. Volumes em  $mm^3$ .

Estrutura	Volume médio Esquerda	Volume médio Direita	DICE médio sem filtro	DICE médio filtrado	DICE médio esquerdo sem filtro	DICE médio direito filtrado
1 - Anterior	131,35	156,15	0,7042	0,7895	0,7864	0,7903
2 - Dorsal	146,68	143,75	0,6118	0,6855	0,6926	0,6771
3 - Lateral-Rostral	2480,27	2422,46	0,7817	0,9105	0,9147	0,9062
4 - Lateral-Caudal	1362,60	1312,99	0,8009	0,8845	0,8844	0,8840
5 - Intralaminar	125,00	153,71	0,5734	0,6855	0,6936	0,6782
6 - Intralaminar-Post.	317,97	306,25	0,7835	0,8355	0,8368	0,8338
7 - Medial	1222,85	1183,50	0,8298	0,8840	0,8880	0,8797
8 - Pulvinar	2240,23	1991,60	0,7545	0,9145	0,9167	0,9120
9 - Lateral Geniculado	441,89	432,32	0,7767	0,9135	0,9128	0,9136
10 - Medial Geniculado	120,51	141,70	0,7013	0,8250	0,8204	0,8300

no DICE após a filtragem. Outro fato interessante é que existe alguma correlação entre o DICE e o tamanho do núcleo (Fig. 4.3). Não é sempre o caso, mas existe uma tendência de melhores valores de DICE para núcleos maiores, com é o caso do Pulvinar e Lateral-Rostral, que são os maiores núcleos e apresentam alguns dos maiores valores de DICE. Algumas hipóteses podem ser atribuídas a esse fenômeno (MAIER-HEIN *et al.*, 2024). A primeira é que erros de classificação de poucos *voxels* causam grande interferência proporcionalmente na métrica, e a segunda é que com tão poucos *voxels* para treinar e classificar, a CNN tem dificuldade em categorizar de forma assertiva.

Os resultados desses experimentos supracitados colocam o método desenvolvido como um dos melhores desempenhos quando comparados com outros métodos presentes na literatura (Tab. 4.3). O método proposto atinge DICES de 0,6855 a 0,9145, enquanto a maioria dos outros métodos reportam valores entre 0,7 e 0,89.

Como não existem conjuntos de dados públicos com anotação manual e nem padronização na quantidade de núcleos, a comparação entre os métodos fica prejudicada. Porém, podemos tentar comparar de forma indireta os resultados. Os métodos que reportaram os resultados em conjuntos de dados privados com máscaras padrão-ouro (SU *et al.*, 2019; ZIYAN *et al.*, 2006; MAJDI *et al.*, 2020) obtiveram valores de DICE de 0,6 a 0,89. Já os métodos que comparam o parcelamento com o atlas (DATTA *et al.*, 2021; NAJDENOVSKA *et al.*, 2018; UMAPATHY *et al.*, 2022), assim como esse trabalho, apresentam DICES compreendidos entre 0,5 e 0,89, com exceção de (RUIZ-PEREZ *et al.*, 2024) que apresenta valores de 0,87 a 0,97, porém esses últimos resultados ocorrem em imagens WMn-MPRAGE sintetizadas sobre as imagens originais. Os outros resultados reportados são medidas do tipo *test-retest* ou *scan-rescan*, onde imagens de um mesmo sujeito, em aquisições distintas, são segmentadas e comparadas (BATTISTELLA *et al.*, 2017; TREGIDGO *et al.*, 2023). Esses trabalhos reportam DICES de 0,83 a 0,94.

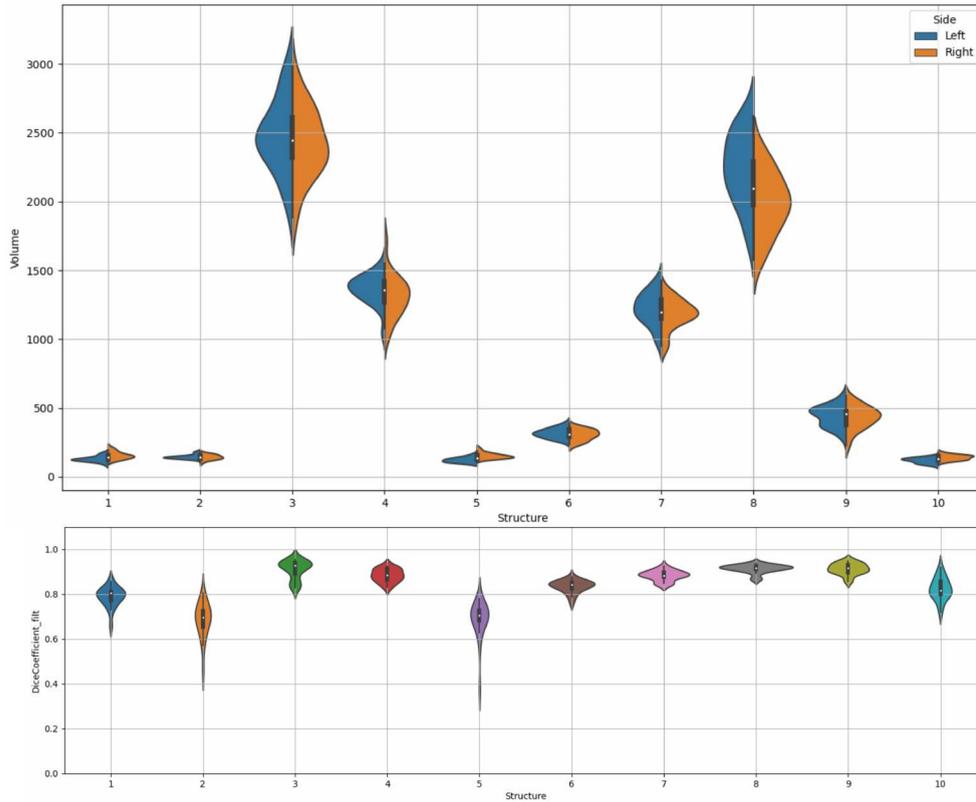


Figura 4.3 – Distribuição do volume médio (superior) e DICE médio (inferior) para cada um dos 10 agrupamentos de núcleos do tálamo no conjunto de testes. Volumes em  $mm^3$ .

Tabela 4.3 – Métodos de parcelamento do tálamo

Trabalho	Tipo de dado	Método	N núcleos (grupos)	DICE	n img. teste	Avaliação
(SU <i>et al.</i> , 2019)	MPRAGE	Atlas	12	0,7 - 0,85	5	manual
(DATTA <i>et al.</i> , 2021)	MP2RAGE	Atlas	12	0,7 - 0,89	13	THOMAS
(NAJDENOVSKA <i>et al.</i> , 2018)	DW-MRI (ODF)	Cluster.	7	0,5 - 0,72	34	Morel
(BATTISTELLA <i>et al.</i> , 2017)	DW-MRI (ODF)	Cluster.	7	0,83 - 0,93	8	scan-rescan
(ZIYAN <i>et al.</i> , 2006)	DTI	Cluster.	7	0,6 - 0,83	10	manual
(TREGIDGO <i>et al.</i> , 2023)	DTI	Inferência bayesiana	26 (5 grupos)	0,84 - 0,94 0,7 - 0,89	21 10	test-retest manual
(MAJDI <i>et al.</i> , 2020)	MPRAGE	CNN	11 (4 grupos)	0,55 - 0,84	10	manual
(UMAPATHY <i>et al.</i> , 2022)	MPRAGE	CNN	11	0,7 - 0,84	2	THOMAS
(RUIZ-PEREZ <i>et al.</i> , 2024)	MPRAGE	CNN	11	0,87 - 0,97	20	THOMAS
<b>NOSSO</b>	DTI+T1	CNN	10	0,68 - 0,91	20	FreeSurfer

Devido a falta de anotações manuais, os resultados foram avaliados contra as máscaras geradas pelo método Iglesias *et al.* (2018), obtidas pelo processamento feito com o FreeSurfer. Portanto, os resultados são medidas do quanto a CNN treinada consegue reproduzir o atlas.

Apesar de não ser o ideal, a reprodução desse atlas é um bom indicativo do poder de parcelamento do tálamo. Essa afirmativa é válida por duas principais considerações. A primeira é que o atlas em questão é atualmente o estado da arte no que se refere a núcleos do tálamo, portanto, uma concordância com ele é o que mais se aproxima do

parcelamento real. O segundo ponto é que a CNN obrigatoriamente precisa saber extrair informações de cada *voxel* para compor o parcelamento, e isso implica que, em atingindo um bom resultado, a primeira metade da CNN garantidamente é capaz de extrair com precisão de cada um dos *voxels* tudo que necessário para a CNN conseguir identificar suas classe. Nesse último caso, uma etapa de ajuste fino com imagens padrão-ouro seria o que basta para atingir o estado da arte, análogo à solução empregada no problema de segmentação.

Felizmente, existem artifícios para tentar de medir o quanto o modelos de CNN são assertivos em suas predições que não sejam a comparação direta com o padrão-ouro. Uma dessas formas é as medidas de incerteza de um modelo (Fig. 4.4).

A primeira consideração sobre as medidas de incerteza é que todas elas possuem valores muito baixos, da ordem de  $10^{-8}$  para as interferências de ruído e *dropout*, e  $10^{-1}$  para perturbações geométricas. Como essas medidas são o desvio padrão de múltiplas predições, podemos afirmar que a variabilidade da predição é mínima, mesmo com perturbações relativamente grandes, indicando robustez do modelo treinado e boa tolerância a perturbações e variações nas imagens de entrada.

No caso da incerteza epistêmica, os *voxels* que apresentam alguma incerteza estão espalhados por toda a imagem. Isso significa que algumas das predições o modelo pode classificar falsos positivos fora das regiões de interesse. Entretanto, todos os falsos positivos são removidos na etapa do pós-processamento, corrigindo qualquer eventual fragilidade do modelo.

Em relação a incerteza estatística, podemos afirmar que o modelo é praticamente imune ao ruído dos níveis que foram aplicados nas imagens. Apesar de qualitativamente ser possível identificar o formato do tálamo no mapa de incertezas gerado por adição do ruído, todos os valores ficam distantes do limiar que alteraria a classificação do *voxel*. Já em relação às variações geométricas, podemos notar que as regiões internas dos núcleos são constantes, ou seja, o modelo tem total convicção da predição nas regiões internas de todos os núcleos, apresentando algum valor de incerteza somente nas bordas de cada um dos agrupamentos de núcleos. Esse comportamento é, de fato, o que se espera de um modelo de CNN em uma tarefa de segmentação, já que são nas bordas em que efeitos como volume parcial apresentam algum tipo de interferência. Além disso, as transformações geométricas aplicadas nesse experimento podem ter um deslocamento de até dois *voxels*, possibilitando que predições diferentes tenham classificações distintas para um mesmo *voxels*.

Em resumo, podemos considerar que qualitativamente, pelas formas em que as incertezas aparecem nos mapas, e quantitativamente, pelos valores de incertezas de

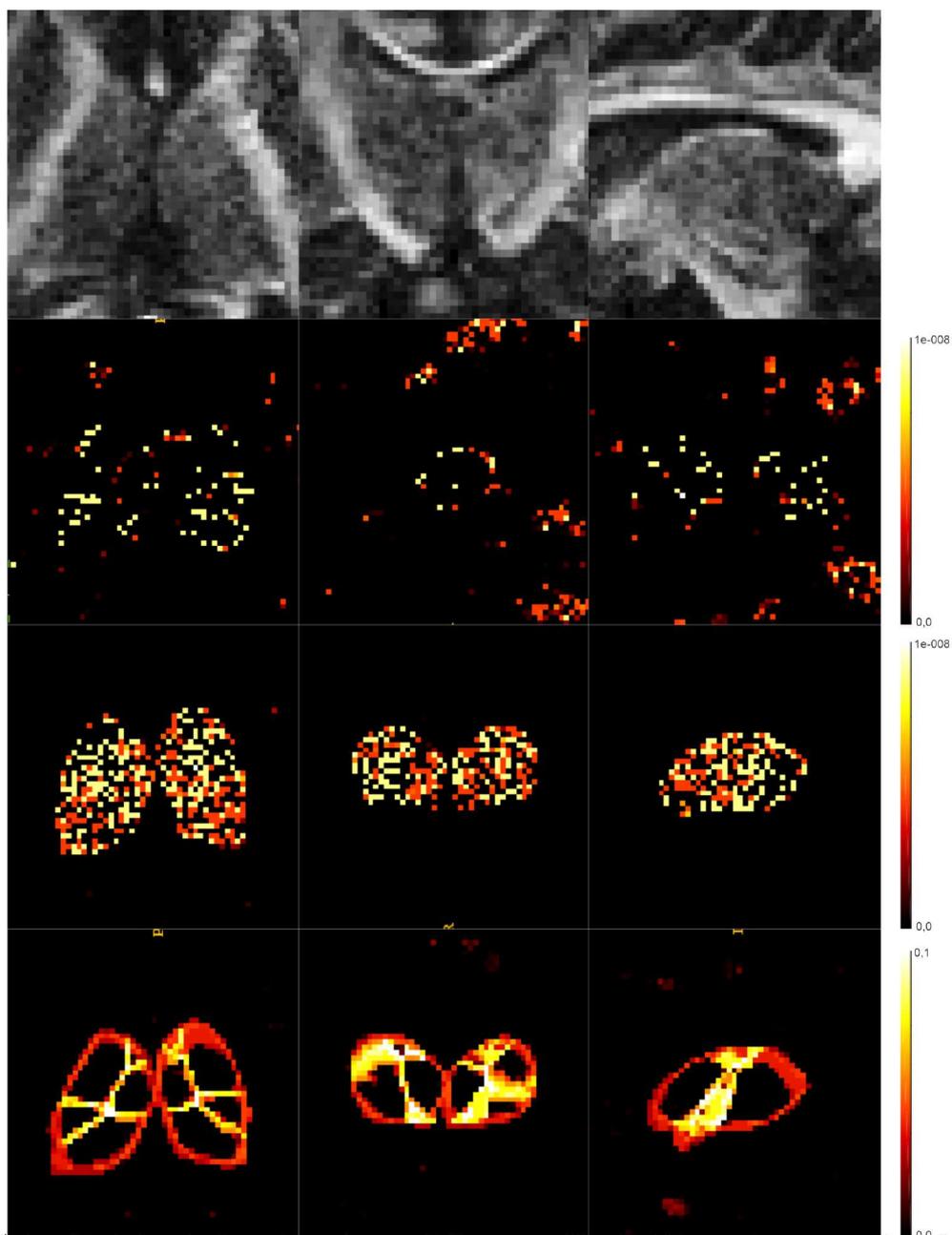


Figura 4.4 – Medidas de incerteza de predição na região do tálamo: imagem original em T1w (primeira linha), incerteza epistêmica, gerada por *dropout* na predição (segunda linha), incerteza gerada pela adição de ruído (terceira linha) e incerteza gerada por deformações (quarta linha), nas vistas axial (esq.), coronal (centro) e sagital (dir.)

cada *voxel*, o método proposto é capaz de parcelar o tálamo com precisão e robustez, apresentando repetibilidade e sendo imune a variações das imagens de entrada.

## 5 Discussão

O principal problema para o desenvolvimento de métodos para segmentação e parcelamento do tálamo é a escassez de máscaras padrão-ouro para o treinamento e para a validação dos algoritmos. Isso se deve ao custo temporal elevado de especialistas com experiência em anatomia de estrutura específica e habilidade em segmentação. Além disso, existem casos onde a segmentação manual em MRI nem é possível, como nos casos dos pequenos núcleos do tálamo que não apresentam contraste suficiente para delimitação.

Para tentar contornar a falta de máscaras para a etapa de treinamento, a técnica de ajuste fino, ou *fine tuning*, é uma boa alternativa, como constatado nos experimentos de segmentação do tálamo (Seq. 3.6.2). Porém ainda apresenta limitações já que mesmo assim são necessárias quantidades relevantes de segmentações padrão-ouro para serem consideradas nessa etapa.

A técnica de ajuste fino reduz a necessidade de máscaras manuais mas não reduz a necessidade da quantidade total de máscaras. Portanto, máscaras automáticas ou provenientes de outros métodos continuam sendo necessárias para a primeira etapa do treinamento das CNNs. Como utilizar máscaras oriundas de um único método por embutir vieses nas CNNs, o uso de múltiplas máscaras é benéfico para o treinamento (Tab. 3.1 e 3.3). Assim, foram testadas algumas formas com o uso de múltiplas máscaras simultaneamente, seja cada uma como uma saída da CNN seja como a combinação via STAPLE, sendo que ambas se provaram boas técnicas para emprego durante a fase de treinamento das redes.

Mesmo com grandes quantidades de dados rotulados com máscaras padrão-ouro, o desempenho das CNNs ainda seria limitado à qualidade das máscaras disponíveis já que a CNN apenas reproduz aquilo para qual foi treinada. Podemos notar a limitação das máscaras manuais quando fazemos a comparação de múltiplos especialistas anotadores e existe um grau de divergência considerável entre eles (Tab. 2.2), sendo o acordo da ordem de 0,8 na métrica DICE, ainda que seguindo o mesmo protocolo de segmentação manual. É tecnicamente impossível algum método de segmentação automática apresentar métricas superiores do que aquelas possíveis de se chegar manualmente. Isso é verdade, já que se considerarmos um método capaz de segmentar perfeitamente alguma estrutura sendo avaliado contra uma máscara manual, o resultado apresentado será análogo à avaliação da própria máscara manual. Sendo assim, DICEs acima de 0,9 podem ser considerados como o limite superior para esse tipo de problema de segmentação de estruturas pequenas e em imagens com baixo contraste.

Além das limitações dos métodos impostas pela da qualidade das máscaras padrão-ouro, existem limitações da avaliação do resultado pela própria métrica DICE. Principalmente nas estruturas de menor volume, onde a proporção do tamanho do *voxel* em relação à estrutura é maior, pequenos erros penalizam consideravelmente essa métrica (MAIER-HEIN *et al.*, 2024). Apesar dessas limitações, a métrica DICE ainda é a mais utilizada para avaliação dos resultados, sendo necessária a computação dessa métrica para avaliação do método e para a comparação com outros trabalhos. Dadas essas limitações, outras métricas como distância entre centros dos objetos segmentados e similaridade volumétrica são boas alternativas para avaliação do desempenho do métodos, especialmente quando o objetivo é identificar a posição correta da estrutura ou a medição dos volumes estruturais.

Apesar de bons resultados de segmentação, as imagens de ressonância T1w e DTI possuem características intrínsecas que limitam a qualidade da segmentação e do parcelamento. Algumas dessas características são a baixa resolução espacial dessas imagens e baixa resolução direcional das imagens de difusão. A baixa resolução espacial faz com que o volume parcial tenha grande influência no resultado da segmentação, especialmente para os núcleos do tálamo que não passam do tamanho de poucos *voxels*. Já a baixa resolução direcional dos dados de difusão limitam os modelos que podem ser aplicados aos dados, e, por sua vez, a variabilidade de mapas que exprimem as propriedades micro estruturais dos tecidos, portanto, apresentado baixos contrastes.

Outra consideração importante é sobre o benefício trazido pela adição de imagens de difusão ao método. Os resultados dos experimentos de segmentação (Seç. 3.6) e parcelamento (Seç. 4.7) demonstram que a adição dos mapas de difusão refinam a segmentação, trazendo mais qualidade e confiabilidade para o resultado final. No caso da segmentação, o DICE foi de 0,88 para 0,90 quando se adicionou os mapas de difusão, e no caso do parcelamento o DICE médio entre todos os núcleos foi de 0,83 para 0,88. Porém, esse ganho vem as custas de uma aquisição extra e mais passos de pré processamento, como os passos de registro entre as sequências. Portanto, a adição de imagens de difusão se justifica em casos onde a melhor qualidade possível de segmentação é desejada e não há prejuízo pela adição de alguns poucos minutos no tempo total que vai desde a aquisição até a segmentação final da imagem do sujeito. Ainda assim, o tempo de execução do método completo, incluindo o registro das imagens, é apenas uma fração do tempo despendido por um método baseado em atlas.

Importante considerar que no caso do processamento das imagens do HCP, foram necessários meses para que pudéssemos ter todos os dados processados, mesmo em uma máquina com excelentes configurações. Como referência, um processamento do FreSurfer pode levar mais de 6 horas para fazer todas as etapas, incluindo o parcelamento.

Como processamos mais de 1000 sujeitos, foram necessárias algo em torno de 6000 horas, mas que foram contadas em paralelo por 12 núcleos de processamento, totalizando praticamente um mês apenas para computação das máscaras do FreeSurfer. Além disso, são necessárias as máscaras provenientes dos demais métodos, registro das imagens de difusão e T1w, computação dos mapas de DTI, entre outros, que também levam tempos consideráveis. Sendo assim, devemos saber que a preparação dos dados para o treinamento de uma CNN para esse objetivo é bastante custosa. Porém, com a disponibilização dos dados totalmente processados para *download*, como esse trabalho fez, quaisquer outros pesquisadores podem pular essa etapa de pré-processamento e começar direto no treinamento das CNNs.

Além de melhores resultados apresentados com a métrica DICE, tanto para a tarefa de segmentação quanto para o parcelamento, o emprego dos mapas de difusão usados nesse trabalho possuem outras vantagens. A principal delas é a imunidade aos parâmetros de aquisição e particularidades do *scanner* utilizado. Isso é válido já que por usar modelos matemáticos que ajustam as medidas de difusão, os mapas intra e inter *voxel* são quantitativos e determinísticos. Portanto, um diferencial relevante dos métodos propostos, tanto para segmentação quanto para o parcelamento, é a capacidade de aproveitamento dos modelos treinados em outros conjuntos de dados e a validade dos resultados com imagens de qualquer origem, desde que atentam os requisitos mínimos para o modelo DTI.

Um problema comum dos métodos baseados em CNN é a necessidade de adaptação para outros conjuntos de dados. Como os modelos são treinados em domínios específicos, o uso adequado em outros conjuntos de dados que não participaram do treinamento requerem um treinamento novo do modelos, ou ao menos algum tipo de ajuste fino para que a CNN consiga ter o melhor desempenho possível. Portanto, para atingir o potencial caracterizado do método aqui apresentado, em outros conjuntos de dados que não o HCP, as etapas de pré-processamento do novo conjunto de dados e o retreinamento ou o ajuste fino das CNNs são necessários. O único caso que teoricamente não exige algum tipo de adaptação de domínio, devido a imunidade ao parâmetros de aquisição, é aquele que emprega exclusivamente imagens provenientes do modelos DTI, desde que a resolução espacial seja a mesma daquela dos modelos já treinados.

Por fim, sobre as limitações do método, podemos afirmar que apesar de trazer benefícios para a qualidade da segmentação, as imagens DTI possuem suas limitações de resolução angular e espacial, o que restringe a quantidade de informação da microestrutura que as imagens podem carregar. Para que essa barreira possa ser superada, é necessário o uso de aquisições de alta resolução angular (HARDI) e *multi shell*, que permitam reconstruções mais elaboradas como o ODF. De fato, alguns dos métodos con-

correntes mais relevantes que também utilizam DWI (Tab. 4.3) empregam modelos de difusão que exigem aquisições de maiores resoluções.

Os modelos de CNN foram treinados exclusivamente com sujeitos controle, sem qualquer imagem de pacientes ou com qualquer tipo de anomalia identificada presente. Sendo assim, não há medições de desempenho em conjuntos de dados ou em imagens que contenham alguma patologia. Apesar dos conjuntos de dados conterem dados demográficos, como sexo e idade, também não foram avaliados os resultados de acordo com categorias dos sujeitos, de modo o desempenho dos métodos por faixa de idade ou sexo dos sujeitos não foi caracterizado.

## 6 Conclusão

Esse trabalho desenvolveu um método de segmentação do tálamo com emprego de imagens de ressonância magnética e redes totalmente convolucionais. Esse método é capaz de segmentar as imagens T1w e imagens difusão, atingindo níveis de qualidade estado da arte de métodos automáticos, enquanto apresenta tempo de execução consideravelmente menor dos que os métodos mais difundidos. Além da segmentação, esse trabalho apresenta um método de parcelamento do tálamo com as mesmas características de desempenho, qualidade e tempo de execução que o método de segmentação.

Os métodos segmentam o tálamo e seus núcleos, atingindo níveis de qualidade estado da arte (DICE médio próximos de 0,9), com imagens T1w e mapas de DTI (FA, MD, AD, RD, TMG e divergente). Essas imagens podem ser usadas de forma individual ou combinadas, mas é recomendada a combinação já que as CNNs se beneficiam da informação mutua, apresentando melhores resultados.

No caso específico de parcelamento, esse trabalho é provavelmente a primeira proposta com emprego de mapas de DTI e CNNs para a segmentação. Outros trabalhos apresentam propostas de CNN com imagens MP-RAGE e T1w, ou apresentam classificadores e clusterização em imagens de difusão. Portanto, esse trabalho é pioneiro no emprego não só de dados DTI (imagens puramente escalares ou com informação direcional), mas imagens T1w compatibilizadas com as demais.

Para viabilizar a segmentação e parcelamento do tálamo com múltiplas sequências de MRI, foram propostos formas de emprego de imagens T1w e imagens de difusão simultaneamente, considerando as particularidades de cada umas das sequências de imagens, incluindo seus diferenciais e suas limitações. Para compatibilizar e harmonizar as imagens de diferentes origens, foram apresentados os passos de um pré-processamento que coloca as imagens na mesma resolução, alinha as estruturas, ajusta em modelos, filtra ruídos e ajusta as intensidades.

Foram propostas modificações nas arquiteturas de CNNs para que pudessem comportar imagens de múltiplas origens, de forma a ponderar de forma equivalente e harmonizada imagens de naturezas distintas. No caso do parcelamento, onde são empregados mapas com muitos canais, o formato proposto usa os primeiros filtros convolucionais para tornar equivalentes as múltiplas entradas, fazendo com que a CNN trate as imagens de origens distintas com a mesma importância. Vale ressaltar que as modificações propostas podem ser facilmente aplicadas a qualquer arquitetura de CNN, já que as modificações interferem exclusivamente nas extremidades da arquitetura de rede.

Os experimentos apresentados nesse trabalho testaram diversas combinações de imagens de entrada, incluindo imagens T1w e mapas escalares de DTI para avaliar e comparar o desempenho de cada uma delas. Ficou constatado que imagens T1w sozinhas conseguem bons resultados, mas que a adição das imagens de DTI garantidamente contribuem com a qualidade da segmentação. Como exemplo, no caso da segmentação do tálamo, a adição de imagens de DTI à imagem T1w no método fez com o que DICE fosse de 0,8837 para 0,8970. Já no caso do parcelamento, o DICE foi de 0,8287 com imagens T1w para 0,8775 com a adição de mapas de difusão.

O objetivo de praticar ciência aberta com os métodos desenvolvidos ao longo desse trabalho também foi atingido. Temos os códigos para todas as etapas do método, desde o pré-processamento, até o pós-processamento e a computação das métricas dos resultados abertos à comunidade. Lançamos também um desafio em uma plataforma online onde os competidores podem partir da versão atual do método de segmentação para aprimorá-lo ou até mesmo usar em outros contextos.

Além do compartilhamento do códigos, métodos e resultados, foram compartilhados todos os dados pré-processados. Esses dados incluem as imagens registradas, máscaras automáticas, segmentação manual, mapas de difusão já computados, e tudo mais necessário para aplicação direta no contexto de CNNs. Também foi desenvolvido e compartilhado um protocolo de segmentação manual do tálamo, que orienta como tirar proveito da imagem T1w e dMRI para identificação das bordas da estrutura.

Sobre o tempo de execução dos métodos, tivemos o objetivo atingido de superar significativamente os competidores diretos e indiretos. O tempo de execução da segmentação de um sujeito, considerando a agregação das três vistas e do pós-processamento é sempre inferior a 5 segundos no método proposto, sendo esse tempo válido também para a arquitetura 3D do parcelamento. Mesmo com a adição de algo próximo de 10 minutos para o registro das imagens e a execução de todos os passos do pré-processamento, o método proposto toma apenas uma fração do tempo que os métodos clássicos, como os baseados em atlas, levam para executar essa tarefa.

## 6.1 Trabalhos futuros

Os experimentos apresentados nesse trabalho demonstraram comparações entre os desempenhos das CNNs para cada combinação dos canais de entrada, e foi demonstrado que as imagens de difusão são benéficas para a segmentação e para o parcelamento do tálamo. Porém, os mecanismos pelos quais as imagens de difusão objetivamente contribuem para a melhora da qualidade da segmentação não foi investigado. Portanto, a investigação desses mecanismos através de técnicas de *Explainable AI*, onde se tenta en-

tender o papel de cada bloco e partes dos blocos das rede convolucionais, é uma linha de pesquisa que pode, inclusive, ajudar a entender como as informações microestruturais dos tecidos se manifestam.

É notável que, apesar de apresentarem ganhos, os mapas de DTI ainda são limitados em relação a informação que carregam para os problemas de segmentação e parcelamento. Isso se deve principalmente ao fato do modelo DTI ser um dos mais simples e não apresentar especificidade suficiente em determinadas regiões, como em cruzamento de fibras por exemplo. Sendo assim, o emprego de modelos mais complexos e completos de imagens de difusão, como modelos de alta resolução angular (HARDI), ODF e dados *multi shell*, certamente trariam melhorias nos resultados das segmentações, especialmente para o problema do parcelamento do tálamo.

O principal motivador para o uso de imagens de difusão mais simples (*single shell* e poucas direções), é a possibilidade de aplicar o método em conjuntos de dados menos graduados, como o conjunto de dados de difusão do HCU, e até mesmo em imagens clínicas, que naturalmente são mais simples. Porém, devido as limitações das imagens DTI, e com a possibilidade de melhora na segmentação com imagens e modelos mais complexos, a aplicação de métodos de melhoramento sobre as imagens de difusão antes de entrarem no fluxo de processamento pelo método apresentado pode ser uma opção viável, especialmente para a segmentação dos núcleos do tálamo.

Outra motivação para o desenvolvimento de métodos para segmentação e parcelamento de tálamo é a possibilidade de aplicação em pacientes para auxílio em tratamento de doenças e planejamento cirúrgico . No entanto, com a disponibilidade apenas de dados de sujeitos controle, não foi possível avaliar o desempenho dos métodos em imagens de pacientes. Portanto, existe a possibilidade aplicar os algoritmos desenvolvidos nesse trabalho em casos clínicos reais.

Para o tratamento de tremor essencial, o núcleo ventral intermédio é o principal alvo cirúrgico. Como o método de parcelamento se mostrou eficaz para a delimitação das bordas dos núcleos e dos agrupamentos dos núcleos, um aprofundamento no problema específico da segmentação do núcleo ventral intermédio seria de grande utilidade para cirurgiões. Portanto, uma das principais possibilidades de avanço desse estudo é justamente no aprimoramento dos métodos para focar nesse núcleo específico.

## 6.2 Trabalhos apresentados, publicações e participação em eventos

Trabalhos como primeiro autor, relacionados ao tema do doutorado:

(PINHEIRO *et al.*, 2018) Automatic Callosal Fiber Convergence Plane Computation through DTI-based Divergence Map. SPIE Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging.

(PINHEIRO *et al.*, 2019) V-Net and U-Net for Ischemic Stroke Lesion Segmentation in a Small Dataset of Perfusion Data. 4th International BrainLes Workshop. MICCAI 2018.

(PINHEIRO *et al.*, 2020) Convolutional Neural Network on DTI Data for Sub-cortical Brain Structure Segmentation. Computational Diffusion MRI Workshop. MICCAI 2019.

(PINHEIRO *et al.*, 2021) Diffusion MRI and silver standard masks to improve CNN-based thalamus segmentation. SPIE Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging.

(PINHEIRO *et al.*, 2023) Thalamus Segmentation Using Deep Learning with Diffusion MRI Data - An Open Benchmark. Applied Sciences. 2023 Apr 23;13(9):5284.

Bolsas de intercâmbio e para participação em eventos:

**COOPERPRINT (Itália 2019):** 3 meses como pesquisador visitante na universidade de Verona.

**IEEE-SPS Summer School in Biomedical Imaging (Colômbia 2023):** Participação como palestrante.

Participação em eventos

**MICCAI/CDMRI (China 2019):** Apresentação de trabalho.

**BRAINN (Campinas 2017, 2019 e 2020):** Apresentação de trabalho.

**Winter School on Brain Connectomics (Itália 2017):** Participação como aluno, com apresentação de trabalho.

**Winter School on Brain Connectomics (Itália 2019):** Participação como aluno e como voluntário na organização do evento.

**IEEE-SPS Summer School in Biomedical Imaging (Colômbia 2023):** Participação como aluno, com apresentação de trabalho.

Trabalhos em colaboração com outros pesquisadores (co-autor):

- (**PHELLAN *et al.*, 2019**) Automatic detection of age- and sex-related differences in human brain morphology. ISMRM 27th ANNUAL MEETING & EXHIBITION.
- (**CRUCIANI *et al.*, 2021a**) Explainable 3D-CNN for Multiple Sclerosis Patients Stratification. International Conference on Pattern Recognition. 2021.
- (**RODRIGUES *et al.*, 2021**) Volumetric segmentation of the corpus callosum: training a deep learning model on diffusion MRI. SPIE Medical Imaging 2021: International Symposium on Medical Information Processing and Analysis.
- (**CRUCIANI *et al.*, 2021b**) Interpretable deep learning as a means for decrypting disease signature in multiple sclerosis. Journal of Neural Engineering. 2021, v.18, n.4.
- (**CARMO *et al.*, 2021**) Rapidly deploying a COVID-19 decision support system in one of the largest Brazilian hospitals. Health Informatics Journal. 2021 v.27, n.3.
- (**CRUCIANI *et al.*, 2023**) Explainable deep learning for decrypting disease signatures in multiple sclerosis. Explainable Deep Learning AI. 2023. p. 97–123.
- (**CARMO *et al.*, 2023**) Automated computed tomography and magnetic resonance imaging segmentation using deep learning a beginner's guide. arXiv preprint arXiv:2304.05901, 2023
- (**JULIO *et al.*, 2023**) Microstructural Changes in the Corpus Callosum in Systemic Lupus Erythematosus. Cells, MDPI, v.12, n.3, p. 355.

## Referências

- AKRAM, H.; DAYAL, V.; MAHLKNECHT, P.; GEORGIEV, D.; HYAM, J.; FOLTYNIE, T.; LIMOUSIN, P.; VITA, E. D.; JAHANSHAHI, M.; ASHBURNER, J. *et al.* Connectivity derived thalamic segmentation in deep brain stimulation for tremor. *NeuroImage: Clinical*, Elsevier, v. 18, p. 130–142, 2018. Citado na página 14.
- AYGÜN, M.; ŞAHİN, Y. H.; ÜNAL, G. Multi modal convolutional neural networks for brain tumor segmentation. *Computing Research Repository*, 2018. Citado na página 50.
- BAGHER-EBADIAN, H.; JAFARI-KHOZANI, K.; MITSIAS, P. D.; LU, M.; SOLTANIAN-ZADEH, H.; CHOPP, M.; EWING, J. R. Predicting final extent of ischemic infarction using artificial neural network analysis of multi-parametric MRI in patients with stroke. In: . [S.l.]: Public Library of Science San Francisco, USA, 2011. v. 6, n. 8, p. e22626. Citado na página 30.
- BASSER, P. J.; MATTIELLO, J.; LEBIHAN, D. Mr diffusion tensor spectroscopy and imaging. *Biophysical journal*, Elsevier, v. 66, n. 1, p. 259–267, 1994. Citado 3 vezes nas páginas 31, 32 e 43.
- BASSER, P. J.; PIERPAOLI, C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of magnetic resonance*, Elsevier, v. 213, n. 2, p. 560–570, 2011. Citado 3 vezes nas páginas 31, 32 e 43.
- BATTISTELLA, G.; NAJDENOVSKA, E.; MAEDER, P.; GHAZALEH, N.; DADUCCI, A.; THIRAN, J.-P.; JACQUEMONT, S.; TULEASCA, C.; LEVIVIER, M.; CUADRA, M. B. *et al.* Robust thalamic nuclei segmentation method based on local diffusion magnetic resonance properties. *Brain Structure and Function*, Springer, v. 222, p. 2203–2216, 2017. Citado 6 vezes nas páginas 20, 21, 23, 38, 78 e 79.
- BERK, C.; CARR, J.; SINDEN, M.; MARTZKE, J.; HONEY, C. Assessing tremor reduction and quality of life following thalamic deep brain stimulation for the treatment of tremor in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, BMJ Publishing Group Ltd, v. 75, n. 8, p. 1210–1211, 2004. Citado na página 14.
- BERNEY, A.; VINGERHOETS, F.; PERRIN, A.; GUEX, P.; VILLEMURE, J.-G.; BURKHARD, P.; BENKELFAT, C.; GHIKA, J. Effect on mood of subthalamic dbs for parkinson’s disease a consecutive series of 24 patients. *Neurology*, AAN Enterprises, v. 59, n. 9, p. 1427–1429, 2002. Citado na página 14.
- BLEZEK, D. J.; MILLER, J. V. Atlas stratification. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.], 2006. p. 712–719. Citado na página 16.
- BRANT-ZAWADZKI, M.; GILLAN, G. D.; NITZ, W. R. Mr rage: a three-dimensional, t1-weighted, gradient-echo sequence—initial experience in the brain. *Radiology*, v. 182, n. 3, p. 769–775, 1992. Citado na página 18.
- BRICE, J.; MCLELLAN, L. Suppression of intention tremor by contingent deep-brain stimulation. *The Lancet*, Elsevier, v. 315, n. 8180, p. 1221–1222, 1980. Citado na página 14.

- BURGGRAAFF, J.; LIU, Y.; PRIETO, J. C.; SIMOES, J.; SITTER, A. de; RUGGIERI, S.; BROUWER, I.; LISSENBERG-WITTE, B. I.; ROCCA, M. A.; VALSASINA, P. *et al.* Manual and automated tissue segmentation confirm the impact of thalamus atrophy on cognition in multiple sclerosis: A multicenter study. *NeuroImage: Clinical*, Elsevier, v. 29, p. 102549, 2021. Citado na página 15.
- CARMO, D.; CAMPIOTTI, I.; RODRIGUES, L.; FANTINI, I.; PINHEIRO, G.; MORAES, D.; NOGUEIRA, R.; RITTNER, L.; LOTUFO, R. Rapidly deploying a covid-19 decision support system in one of the largest brazilian hospitals. *Health Informatics Journal*, SAGE Publications Sage UK: London, England, v. 27, n. 3, p. 14604582211033017, 2021. Citado na página 90.
- CARMO, D.; PINHEIRO, G.; RODRIGUES, L.; ABREU, T.; LOTUFO, R.; RITTNER, L. Automated computed tomography and magnetic resonance imaging segmentation using deep learning: a beginner's guide. *arXiv preprint arXiv:2304.05901*, 2023. Citado na página 90.
- CASTRO, F. S.; POLLO, C.; MEULI, R.; MAEDER, P.; CUISENAIRE, O.; CUADRA, M. B.; VILLEMURE, J.-G.; THIRAN, J.-P. A cross validation study of deep brain stimulation targeting: from experts to atlas-based, segmentation-based and automatic registration algorithms. *IEEE transactions on medical imaging*, IEEE, v. 25, n. 11, p. 1440–1450, 2006. Citado na página 14.
- CHEN, L.; BENTLEY, P.; RUECKERT, D. Fully automatic acute ischemic lesion segmentation in dwi using convolutional neural networks. *neuroimage clin* 15: 633–643. 2017. Citado na página 30.
- CHEN, L.; WU, Y.; DSOUZA, A. M.; ABIDIN, A. Z.; WISMÜLLER, A.; XU, C. MRI tumor segmentation with densely connected 3D CNN. In: SPIE. *Medical imaging 2018: image processing*. [S.l.], 2018. v. 10574, p. 357–364. Citado na página 50.
- CLÈRIGUES, A.; VALVERDE, S.; BERNAL, J.; FREIXENET, J.; OLIVER, A.; LLADÓ, X. Acute and sub-acute stroke lesion segmentation from multimodal MRI. *Computer Methods and Programs in Biomedicine*, p. 105521, 2020. Citado 2 vezes nas páginas 31 e 50.
- CRUCIANI, F.; BRUSINI, L.; ZUCHELLI, M.; PINHEIRO, G. R.; SETTI, F.; GALAZZO, I. B.; DERICHE, R.; RITTNER, L.; CALABRESE, M.; MENEGAZ, G. Explainable 3d-cnn for multiple sclerosis patients stratification. In: SPRINGER. *International Conference on Pattern Recognition*. [S.l.], 2021. p. 103–114. Citado na página 90.
- CRUCIANI, F.; BRUSINI, L.; ZUCHELLI, M.; PINHEIRO, G. R.; SETTI, F.; GALAZZO, I. B.; DERICHE, R.; RITTNER, L.; CALABRESE, M.; MENEGAZ, G. Interpretable deep learning as a means for decrypting disease signature in multiple sclerosis. *Journal of Neural Engineering*, IOP Publishing, v. 18, n. 4, p. 0460a6, 2021. Citado na página 90.
- CRUCIANI, F.; BRUSINI, L.; ZUCHELLI, M.; PINHEIRO, G. R.; SETTI, F.; DERICHE, R.; RITTNER, L.; CALABRESE, M.; GALAZZO, I. B.; MENEGAZ, G. Explainable deep learning for decrypting disease signatures in multiple sclerosis. In: *Explainable Deep Learning AI*. [S.l.]: Elsevier, 2023. p. 97–123. Citado na página 90.

- DARKNER, S. *FDG-PET template MNI152 1mm*. 2013. Citado na página 43.
- DATTA, R.; BACCHUS, M. K.; KUMAR, D.; ELLIOTT, M. A.; RAO, A.; DOLUI, S.; REDDY, R.; BANWELL, B. L.; SARANATHAN, M. Fast automatic segmentation of thalamic nuclei from mp2rage acquisition at 7 tesla. *Magnetic Resonance in Medicine*, Wiley Online Library, v. 85, n. 5, p. 2781–2790, 2021. Citado 3 vezes nas páginas 23, 78 e 79.
- DEELEY, M.; CHEN, A.; DATTERI, R.; NOBLE, J.; CMELAK, A.; DONNELLY, E.; MALCOLM, A.; MORETTI, L.; JABOIN, J.; NIERMANN, K. *et al.* Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study. *Physics in Medicine & Biology*, IOP Publishing, v. 56, n. 14, p. 4557, 2011. Citado na página 15.
- DEMEESTERE, J.; WOUTERS, A.; CHRISTENSEN, S.; LEMMENS, R.; LANSBERG, M. G. Review of perfusion imaging in acute ischemic stroke: from time to tissue. *Stroke*, Am Heart Assoc, v. 51, n. 3, p. 1017–1024, 2020. Citado na página 31.
- DESCOTEAUX, M.; ANGELINO, E.; FITZGIBBONS, S.; DERICHE, R. A fast and robust odf estimation algorithm in q-ball imaging. In: IEEE. *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006*. [S.l.], 2006. p. 81–84. Citado na página 32.
- DEVELOPERS, N. *nibabel 1.2.0*. 2016. Disponível em: <<https://doi.org/10.5281/zenodo.60858>>. Citado na página 59.
- DOLZ, J.; DESROSIERS, C.; AYED, I. B. 3d fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *NeuroImage*, Elsevier, v. 170, p. 456–470, 2018. Citado 2 vezes nas páginas 18 e 19.
- DOLZ, J.; DESROSIERS, C.; AYED, I. B. IVD-Net: Intervertebral disc localization and segmentation in MRI with a multi-modal UNet. In: SPRINGER. *International workshop and challenge on computational methods and clinical applications for spine imaging*. [S.l.], 2018. p. 130–143. Citado na página 50.
- DOLZ, J.; GOPINATH, K.; YUAN, J.; LOMBAERT, H.; DESROSIERS, C.; AYED, I. B. Hyperdense-net: a hyper-densely connected cnn for multi-modal image segmentation. *IEEE transactions on medical imaging*, IEEE, v. 38, n. 5, p. 1116–1126, 2018. Citado na página 50.
- DOUEK, P.; TURNER, R.; PEKAR, J.; PATRONAS, N.; BIHAN, D. L. MR color mapping of myelin fiber orientation. *Journal of Computer Assisted Tomography*, v. 15, n. 6, p. 923–929, 1991. ISSN 0363-8715. Disponível em: <<https://doi.org/10.1097/00004728-199111000-00003>>. Citado na página 31.
- ESSEN, D. C. V.; SMITH, S. M.; BARCH, D. M.; BEHRENS, T. E.; YACOUB, E.; UGURBIL, K.; CONSORTIUM, W.-M. H. *et al.* The wu-minn human connectome project: an overview. *Neuroimage*, Elsevier, v. 80, p. 62–79, 2013. Citado na página 39.
- FADDA, G.; BROWN, R. A.; MAGLIOZZI, R.; AUBERT-BROCHE, B.; O'MAHONY, J.; SHINOHARA, R. T.; BANWELL, B.; MARRIE, R. A.; YEH, E. A.; COLLINS, D. L. *et al.* A surface-in gradient of thalamic damage evolves in pediatric multiple sclerosis. *Annals of neurology*, Wiley Online Library, v. 85, n. 3, p. 340–351, 2019. Citado na página 15.

- FISCHL, B.; SALAT, D. H.; BUSA, E.; ALBERT, M.; DIETERICH, M.; HASELGROVE, C.; KOUWE, A. V. D.; KILLIANY, R.; KENNEDY, D.; KLAVENESS, S. *et al.* Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, Elsevier, v. 33, n. 3, p. 341–355, 2002. Citado 3 vezes nas páginas 17, 19 e 20.
- GAL, Y.; GHAHRAMANI, Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In: PMLR. *international conference on machine learning*. [S.l.], 2016. p. 1050–1059. Citado na página 74.
- GARYFALLIDIS, E.; BRETT, M.; AMIRBEKIAN, B.; ROKEM, A.; WALT, S. V. D.; DESCOTEAUX, M.; NIMMO-SMITH, I.; CONTRIBUTORS, D. Dipy, a library for the analysis of diffusion MRI data. *Frontiers in neuroinformatics*, Frontiers Media SA, v. 8, p. 8, 2014. Citado 2 vezes nas páginas 42 e 43.
- GLAISTER, J.; CARASS, A.; NESSAIVER, T.; STOUGH, J. V.; SAIDHA, S.; CALABRESI, P. A.; PRINCE, J. L. Thalamus segmentation using multi-modal feature classification: Validation and pilot study of an age-matched cohort. *NeuroImage*, v. 158, p. 430–440, 2017. ISSN 1053-8119. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1053811917305189>>. Citado 3 vezes nas páginas 18, 19 e 64.
- GLASSER, M. F.; SOTIROPOULOS, S. N.; WILSON, J. A.; COALSON, T. S.; FISCHL, B.; ANDERSSON, J. L.; XU, J.; JBABDI, S.; WEBSTER, M.; POLIMENI, J. R. *et al.* The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, Elsevier, v. 80, p. 105–124, 2013. Citado na página 39.
- HAJNAL, J. V.; BRYANT, D. J.; KASUBOSKI, L.; PATTANY, P. M.; COENE, B. D.; LEWIS, P. D.; PENNOCK, J. M.; OATRIDGE, A.; YOUNG, I. R.; BYDDER, G. M. Use of fluid attenuated inversion recovery (flair) pulse sequences in MRI of the brain. *Journal of computer assisted tomography*, LWW, v. 16, n. 6, p. 841–844, 1992. Citado na página 29.
- HANNOUN, S.; TUTUNJI, R.; HOMSI, M. E.; SAAYBI, S.; HOURANI, R. Automatic thalamus segmentation on unenhanced 3d t1 weighted images: Comparison of publicly available segmentation methods in a pediatric population. *Neuroinformatics*, Springer, v. 17, p. 443–450, 2019. Citado 2 vezes nas páginas 16 e 19.
- HARIZ, M. I.; KRACK, P.; ALESCH, F.; AUGUSTINSSON, L.; BOSCH, A.; EKBERG, R.; JOHANSSON, F.; JOHNELS, B.; MEYERSON, B.; N’GUYEN, J. *et al.* Multicentre european study of thalamic stimulation for parkinsonian tremor: a 6 year follow-up. *Journal of Neurology, Neurosurgery & Psychiatry*, BMJ Publishing Group Ltd, v. 79, n. 6, p. 694–699, 2008. Citado na página 14.
- IGLESIAS, J. E.; INSAUSTI, R.; LERMA-USABIAGA, G.; BOCCHETTA, M.; LEEMPOT, K. V.; GREVE, D. N.; KOUWE, A. Van der; FISCHL, B.; CABALLERO-GAUDES, C.; PAZ-ALONSO, P. M. *et al.* A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. *Neuroimage*, Elsevier, v. 183, p. 314–326, 2018. Citado 7 vezes nas páginas 20, 23, 38, 47, 48, 69 e 79.
- III, J. P. M.; BROOKEMAN, J. R. Rapid three-dimensional t1-weighted mr imaging with the mp-rage sequence. *Journal of Magnetic Resonance Imaging*, Wiley Online Library, v. 1, n. 5, p. 561–567, 1991. Citado na página 18.

- IRFANOGLU, M. O.; SARLLS, J.; NAYAK, A.; PIERPAOLI, C. Evaluating corrections for eddy-currents and other epi distortions in diffusion MRI: methodology and a dataset for benchmarking. *Magnetic resonance in medicine*, Wiley Online Library, v. 81, n. 4, p. 2774–2787, 2019. Citado 2 vezes nas páginas 31 e 43.
- ISENSEE, F.; JAEGER, P. F.; KOHL, S. A.; PETERSEN, J.; MAIER-HEIN, K. H. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, Nature Publishing Group, v. 18, n. 2, p. 203–211, 2021. Citado 2 vezes nas páginas 49 e 61.
- ISENSEE, F.; KICKINGEREDER, P.; WICK, W.; BENDSZUS, M.; MAIER-HEIN, K. H. Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In: SPRINGER. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3*. [S.l.], 2018. p. 287–297. Citado na página 50.
- ISENSEE, F.; KICKINGEREDER, P.; WICK, W.; BENDSZUS, M.; MAIER-HEIN, K. H. No new-net. In: *International MICCAI Brainlesion Workshop*. [S.l.: s.n.], 2018. p. 234–244. Citado na página 50.
- JENKINSON, M.; BECKMANN, C. F.; BEHRENS, T. E.; WOOLRICH, M. W.; SMITH, S. M. Fsl. *Neuroimage*, Elsevier, v. 62, n. 2, p. 782–790, 2012. Citado 2 vezes nas páginas 17 e 19.
- JULIO, P. R.; CALDEIRA, T.; PINHEIRO, G. R.; CAPELLO, C. H.; FRITOLLI, R. B.; MARINI, R.; CENDES, F.; FERNANDES, P. T.; COSTALLAT, L. T.; RITTNER, L. *et al.* Microstructural changes in the corpus callosum in systemic lupus erythematosus. *Cells*, MDPI, v. 12, n. 3, p. 355, 2023. Citado na página 90.
- KAMNITSAS, K.; BAI, W.; FERRANTE, E.; MCDONAGH, S.; SINCLAIR, M.; PAWLOWSKI, N.; RAJCHL, M.; LEE, M.; KAINZ, B.; RUECKERT, D. *et al.* Ensembles of multiple models and architectures for robust brain tumour segmentation. In: SPRINGER. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3*. [S.l.], 2018. p. 450–462. Citado na página 50.
- KAMNITSAS, K.; LEDIG, C.; NEWCOMBE, V. F.; SIMPSON, J. P.; KANE, A. D.; MENON, D. K.; RUECKERT, D.; GLOCKER, B. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical image analysis*, Elsevier, v. 36, p. 61–78, 2017. Citado na página 30.
- KENDALL, A.; BADRINARAYANAN, V.; CIPOLLA, R. Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding. *arXiv preprint arXiv:1511.02680*, 2015. Citado na página 74.
- KENDALL, A.; GAL, Y. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems*, v. 30, 2017. Citado 2 vezes nas páginas 74 e 75.

- KINGSLEY, P. B. Introduction to diffusion tensor imaging mathematics: Part II. anisotropy, diffusion-weighting factors, and gradient encoding schemes. *Concepts in Magnetic Resonance Part A*, v. 28A, n. 2, p. 123–154, 2006. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/cmr.a.20049>>. Citado na página 33.
- LEEMPUT, K. V.; MAES, F.; VANDERMEULEN, D.; SUETENS, P. A unifying framework for partial volume segmentation of brain mr images. *IEEE transactions on medical imaging*, IEEE, v. 22, n. 1, p. 105–119, 2003. Citado na página 26.
- LI, H.; ZHANG, H.; JOHNSON, H.; LONG, J. D.; PAULSEN, J. S.; OGUZ, I. MRI subcortical segmentation in neurodegeneration with cascaded 3D CNNs. In: SPIE. *Medical Imaging 2021: Image Processing*. [S.l.], 2021. v. 11596, p. 236–243. Citado 2 vezes nas páginas 18 e 19.
- LIU, L.; GLAISTER, J.; SUN, X.; CARASS, A.; TRAN, T. D.; PRINCE, J. L. Segmentation of thalamus from MR images via task-driven dictionary learning. In: SPIE. *Medical Imaging 2016: Image Processing*. [S.l.], 2016. v. 9784, p. 923–929. Citado 2 vezes nas páginas 18 e 19.
- LIU, Y.; D’HAESE, P.-F.; NEWTON, A. T.; DAWANT, B. M. Thalamic nuclei segmentation in clinical 3t t1-weighted images using high-resolution 7t shape models. In: SPIE. *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*. [S.l.], 2015. v. 9415, p. 91–99. Citado na página 20.
- MACHADO, Â. *Neuroanatomia funcional*. Livraria Atheneu, 1974. (Biblioteca biomédica : Série morfológica). Disponível em: <<https://books.google.com.br/books?id=yXJgAAAAMAAJ>>. Citado na página 14.
- MAIER-HEIN, L.; REINKE, A.; GODAU, P.; TIZABI, M. D.; BUETTNER, F.; CHRISTODOULOU, E.; GLOCKER, B.; ISENSEE, F.; KLEESIEK, J.; KOZUBEK, M. *et al.* Metrics reloaded: recommendations for image analysis validation. *Nature methods*, Nature Publishing Group US New York, v. 21, n. 2, p. 195–212, 2024. Citado 2 vezes nas páginas 78 e 83.
- MAIER, O.; MENZE, B. H.; GABLENTZ, J. Von der; HÄNI, L.; HEINRICH, M. P.; LIEBRAND, M.; WINZECK, S.; BASIT, A.; BENTLEY, P.; CHEN, L. *et al.* Isles 2015-a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Medical image analysis*, Elsevier, v. 35, p. 250–269, 2017. Citado na página 31.
- MAJDI, M. S.; KEERTHIVASAN, M. B.; RUTT, B. K.; ZAHR, N. M.; RODRIGUEZ, J. J.; SARANATHAN, M. Automated thalamic nuclei segmentation using multi-planar cascaded convolutional neural networks. *Magnetic resonance imaging*, Elsevier, v. 73, p. 45–54, 2020. Citado 5 vezes nas páginas 21, 23, 38, 78 e 79.
- MAKOWSKI, C.; BÉLAND, S.; KOSTOPOULOS, P.; BHAGWAT, N.; DEVENYI, G. A.; MALLA, A. K.; JOOBER, R.; LEPAGE, M.; CHAKRAVARTY, M. M. Evaluating accuracy of striatal, pallidal, and thalamic segmentation methods: Comparing automated approaches to manual delineation. *Neuroimage*, Elsevier, v. 170, p. 182–198, 2018. Citado na página 46.
- MALAGA, K. A.; HOUSHMAND, L.; COSTELLO, J. T.; CHANDRASEKARAN, J.; CHOU, K. L.; PATIL, P. G. Thalamic segmentation and neural activation modeling based on individual tissue microstructure in deep brain stimulation for essential tremor.

- Neuromodulation: Technology at the Neural Interface*, Elsevier, v. 26, n. 8, p. 1689–1698, 2023. Citado na página 15.
- MANJÓN, J. V.; COUPÉ, P. volBrain: an online MRI brain volumetry system. *Frontiers in Neuroinformatics*, v. 10, p. 30, 2016. Citado 3 vezes nas páginas 17, 19 e 20.
- MIDDLEBROOKS, E. H.; HOLANDA, V. M.; TUNA, I. S.; DESHPANDE, H. D.; BREDEL, M.; ALMEIDA, L.; WALKER, H. C.; GUTHRIE, B. L.; FOOTE, K. D.; OKUN, M. S. A method for pre-operative single-subject thalamic segmentation based on probabilistic tractography for essential tremor deep brain stimulation. *Neuroradiology*, Springer, v. 60, p. 303–309, 2018. Citado 2 vezes nas páginas 15 e 22.
- MIDDLEBROOKS, E. H.; TUNA, I. S.; ALMEIDA, L.; GREWAL, S. S.; WONG, J.; HECKMAN, M. G.; LESSER, E. R.; BREDEL, M.; FOOTE, K. D.; OKUN, M. S. *et al.* Structural connectivity-based segmentation of the thalamus and prediction of tremor improvement following thalamic deep brain stimulation of the ventral intermediate nucleus. *NeuroImage: Clinical*, Elsevier, v. 20, p. 1266–1273, 2018. Citado na página 14.
- MILLETARI, F.; NAVAB, N.; AHMADI, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: IEEE. *2016 fourth international conference on 3D vision (3DV)*. [S.l.], 2016. p. 565–571. Citado na página 61.
- MITCHELL, D.; COHEN, M. *MRI Principles*. Saunders, 2004. ISBN 9780721600246. Disponível em: <<https://books.google.com.br/books?id=Ii1sQgAACAAJ>>. Citado na página 30.
- MORI, S.; WU, D.; CERITOGU, C.; LI, Y.; KOLASNY, A.; VAILLANT, M. A.; FARIA, A. V.; OISHI, K.; MILLER, M. I. MRICloud: delivering high-throughput MRI neuroinformatics as cloud-based software as a service. *Computing in Science & Engineering*, IEEE, v. 18, n. 5, p. 21–35, 2016. Citado na página 19.
- NAJDENOVSKA, E.; ALEMÁN-GÓMEZ, Y.; BATTISTELLA, G.; DESCOTEAUX, M.; HAGMANN, P.; JACQUEMONT, S.; MAEDER, P.; THIRAN, J.-P.; FORNARI, E.; CUADRA, M. B. In-vivo probabilistic atlas of human thalamic nuclei based on diffusion-weighted magnetic resonance imaging. *Scientific data*, Nature Publishing Group, v. 5, n. 1, p. 1–11, 2018. Citado 5 vezes nas páginas 20, 21, 23, 78 e 79.
- NAJDENOVSKA, E.; TULEASCA, C.; JORGE, J.; MAEDER, P.; MARQUES, J. P.; ROINE, T.; GALLICHAN, D.; THIRAN, J.-P.; LEVIVIER, M.; CUADRA, M. B. Comparison of MRI-based automated segmentation methods and functional neurosurgery targeting with direct visualization of the ventro-intermediate thalamic nucleus at 7T. *Scientific reports*, Nature Publishing Group UK London, v. 9, n. 1, p. 1119, 2019. Citado 2 vezes nas páginas 14 e 15.
- NIE, D.; WANG, L.; GAO, Y.; SHEN, D. Fully convolutional networks for multi-modality iso-intense infant brain image segmentation. In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. [S.l.: s.n.], 2016. p. 1342–1345. Citado na página 50.
- NVIDIA Corporation. *NVIDIA CUDA C Programming Guide*. 2010. Version 3.2. Citado na página 58.
- OLIPHANT, T. E. *et al. Guide to numpy*. [S.l.]: Trelgol Publishing USA, 2006. v. 1. Citado na página 59.

- OPFER, R.; KRÜGER, J.; SPIES, L.; OSTWALDT, A.-C.; KITZLER, H. H.; SCHIPPLING, S.; BUCHERT, R. Automatic segmentation of the thalamus using a massively trained 3d convolutional neural network: higher sensitivity for the detection of reduced thalamus volume by improved inter-scanner stability. *European radiology*, Springer, v. 33, n. 3, p. 1852–1861, 2023. Citado 2 vezes nas páginas 18 e 19.
- PASZKE, A.; GROSS, S.; MASSA, F.; LERER, A.; BRADBURY, J.; CHANAN, G.; KILLEEN, T.; LIN, Z.; GIMELSHEIN, N.; ANTIGA, L. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, v. 32, 2019. Citado na página 58.
- PAVAO, A.; GUYON, I.; LETOURNEL, A.-C.; TRAN, D.-T.; BARO, X.; ESCALANTE, H. J.; ESCALERA, S.; THOMAS, T.; XU, Z. Codalab competitions: An open source platform to organize scientific challenges. *Journal of Machine Learning Research*, v. 24, n. 198, p. 1–6, 2023. Citado na página 53.
- PEREIRA, S.; PINTO, A.; ALVES, V.; SILVA, C. A. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE transactions on medical imaging*, IEEE, v. 35, n. 5, p. 1240–1251, 2016. Citado na página 50.
- PEREZ, F.; GRANGER, B. E. IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, v. 9, n. 3, p. 21–29, maio 2007. ISSN 1521-9615. Disponível em: <<https://ipython.org>>. Citado na página 59.
- PELLAN, R.; RODRIGUES, L.; PINHEIRO, G. R.; SOTO, A. Q.; RODRIGUES, I. D.; RITTNER, L.; FERRARI, R.; BROWN, M. R.; FORKERT, N. D.; MEDEIROS, R. *et al.* Automatic detection of age-and sex-related differences in human brain morphology. In: *Proceedings of International Society for Magnetic Resonance in Medicine (ISMRM) 27th ANNUAL MEETING & EXHIBITION*. [S.l.: s.n.], 2019. Citado na página 90.
- PINHEIRO, G.; CARMO, D.; YASUDA, C.; LOTUFO, R.; RITTNER, L. Convolutional neural network on DTI data for sub-cortical brain structure segmentation. In: SPRINGER. *Computational Diffusion MRI: MICCAI Workshop, Shenzhen, China, October 2019*. [S.l.], 2020. p. 135–146. Citado na página 89.
- PINHEIRO, G. R.; BRUSINI, L.; BAJRAMI, A.; PIZZINI, F. B.; CALABRESE, M.; REIS, F.; APPENZELLER, S.; MENEGAZ, G.; RITTNER, L. Diffusion MRI and silver standard masks to improve CNN-based thalamus segmentation. In: SPIE. *Medical Imaging 2021: Image Processing*. [S.l.], 2021. v. 11596, p. 692–698. Citado 2 vezes nas páginas 63 e 89.
- PINHEIRO, G. R.; BRUSINI, L.; CARMO, D.; PRÔA, R.; ABREU, T.; APPENZELLER, S.; MENEGAZ, G.; RITTNER, L. Thalamus segmentation using deep learning with diffusion MRI data: An open benchmark. *Applied Sciences*, MDPI, v. 13, n. 9, p. 5284, 2023. Citado 3 vezes nas páginas 53, 68 e 89.
- PINHEIRO, G. R.; COSTA, G. S. S. A. L.; LOTUFO, R. A.; RITTNER, L. Divergence map from diffusion tensor imaging: Concepts and application to corpus callosum. In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.: s.n.], 2016. p. 1120–1123. Citado 2 vezes nas páginas 36 e 37.

- PINHEIRO, G. R.; COVER, G. S.; BENTO, M. P.; RITTNER, L. Automatic callosal fiber convergence plane computation through dti-based divergence map. In: SPIE. *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*. [S.l.], 2018. v. 10578, p. 256–264. Citado na página 89.
- PINHEIRO, G. R.; VOLTOLINE, R.; BENTO, M.; RITTNER, L. V-Net and U-Net for ischemic stroke lesion segmentation in a small dataset of perfusion data. In: SPRINGER. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part I 4*. [S.l.], 2019. p. 301–309. Citado 2 vezes nas páginas 31 e 89.
- PLANCHE, V.; SU, J. H.; MOURNET, S.; SARANATHAN, M.; DOUSSET, V.; HAN, M.; RUTT, B. K.; TOURDIAS, T. White-matter-nulled mprage at 7t reveals thalamic lesions and atrophy of specific thalamic nuclei in multiple sclerosis. *Multiple Sclerosis Journal*, SAGE Publications Sage UK: London, England, v. 26, n. 8, p. 987–992, 2020. Citado 2 vezes nas páginas 18 e 21.
- POHL, K. M.; FISHER, J.; GRIMSON, W. E. L.; KIKINIS, R.; WELLS, W. M. A bayesian model for joint segmentation and registration. *NeuroImage*, Elsevier, v. 31, n. 1, p. 228–239, 2006. Citado na página 19.
- RAINA, K.; YAHORAU, U.; SCHMAH, T. Exploiting bilateral symmetry in brain lesion segmentation. *arXiv preprint arXiv:1907.08196*, 2019. Citado na página 30.
- RITTNER, L.; FLORES, F.; LOTUFO, R. New tensorial representation of color images: Tensorial morphological gradient applied to color image segmentation. In: *XX Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2007)*. [S.l.: s.n.], 2007. p. 45–52. Citado na página 35.
- RITTNER, L.; LOTUFO, R. Diffusion tensor imaging segmentation by watershed transform on tensorial morphological gradient. In: IEEE. *2008 XXI Brazilian Symposium on Computer Graphics and Image Processing*. [S.l.], 2008. p. 196–203. Citado na página 33.
- RITTNER, L.; LOTUFO, R. A.; CAMPBELL, J.; PIKE, G. B. Segmentation of thalamic nuclei based on tensorial morphological gradient of diffusion tensor fields. In: IEEE. *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. [S.l.], 2010. p. 1173–1176. Citado 2 vezes nas páginas 16 e 33.
- RODRIGUES, J.; PINHEIRO, G.; CARMO, D.; RITTNER, L. Volumetric segmentation of the corpus callosum: training a deep learning model on diffusion mri. In: SPIE. *17th International Symposium on Medical Information Processing and Analysis*. [S.l.], 2021. v. 12088, p. 198–207. Citado na página 90.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. [s.n.], 2015. (LNCS, v. 9351), p. 234–241. Disponível em: <<http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a>>. Citado na página 49.
- ROSSUM, G. V.; JR, F. L. D. *Python tutorial*. [S.l.: s.n.], 1995. Citado na página 59.

- ROVARIS, M.; INGLESE, M.; VITI, B.; CIBODDO, G.; GEREVINI, S.; CAPRA, R.; FILIPPI, M. The contribution of fast-FLAIR MRI for lesion detection in the brain of patients with systemic autoimmune diseases. *Journal of neurology*, Springer, v. 247, p. 29–33, 2000. Citado na página 29.
- ROY, A. G.; CONJETI, S.; NAVAB, N.; WACHINGER, C. Quicknat: Segmenting MRI neuroanatomy in 20 seconds. *ArXiv*, abs/1801.04161, 2018. Disponível em: <<https://api.semanticscholar.org/CorpusID:195346495>>. Citado 3 vezes nas páginas 17, 19 e 49.
- ROY, A. G.; CONJETI, S.; NAVAB, N.; WACHINGER, C. Quicknat: segmenting MRI neuroanatomy in 20 seconds. *arXiv preprint arXiv:1801.04161*, CoRR, 2018. Citado na página 49.
- RUIZ-PEREZ, M.; MORELL-ORTEGA, S.; GADEA, M.; VIVO-HERNANDO, R.; RUBIO, G.; APARICI, F.; IGLESIA-VAYA, M. de la; TOURDIAS, T.; COUPÉ, P.; MANJÓN, J. V. DeepThalamus: A novel deep learning method for automatic segmentation of brain thalamic nuclei from multimodal ultra-high resolution MRI. *arXiv preprint arXiv:2401.07751*, 2024. Citado 4 vezes nas páginas 21, 23, 78 e 79.
- SARANATHAN, M.; IGLEHART, C.; MONTI, M.; TOURDIAS, T.; RUTT, B. In vivo high-resolution structural mri-based atlas of human thalamic nuclei. *Scientific data*, Nature Publishing Group UK London, v. 8, n. 1, p. 275, 2021. Citado na página 20.
- SATHISH, R.; RAJAN, R.; VUPPUTURI, A.; GHOSH, N.; SHEET, D. Adversarially trained convolutional neural networks for semantic segmentation of ischaemic stroke lesion using multisequence magnetic resonance imaging. In: IEEE. *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. [S.l.], 2019. p. 1010–1013. Citado na página 30.
- SEMEDO, C.; CARDOSO, M. J.; VOS, S. B.; MENDELSON, A. F.; RIBBENS, A.; SMEETS, D.; ROHRER, J. D.; OURSELIN, S. Improved tractography-based segmentation of the human thalamus. In: *Proceedings of the International Society for Magnetic Resonance in Medicine*. [S.l.: s.n.], 2017. Citado na página 22.
- SEMEDO, C.; CARDOSO, M. J.; VOS, S. B.; SUDRE, C. H.; BOCCHETTA, M.; RIBBENS, A.; SMEETS, D.; ROHRER, J.; OURSELIN, S. Thalamic nuclei segmentation on dementia using tractography and population-specific priors. In: *Proceedings of the International Society for Magnetic Resonance in Medicine*. [S.l.: s.n.], 2018. Citado na página 22.
- SHERMAN, S. M.; GUILLERY, R. W. *Exploring the thalamus*. [S.l.]: Elsevier, 2001. Citado na página 14.
- SIEGFRIED, J.; REA, G. L. Thalamotomy for parkinson's disease. In: \_\_\_\_\_. *Modern Stereotactic Neurosurgery*. Boston, MA: Springer US, 1988. p. 333–340. ISBN 978-1-4613-1081-5. Disponível em: <[https://doi.org/10.1007/978-1-4613-1081-5\\_25](https://doi.org/10.1007/978-1-4613-1081-5_25)>. Citado 2 vezes nas páginas 14 e 15.
- SILVERSMITH, W. connected-components-3d 1.1. 0. 2024. Citado 2 vezes nas páginas 59 e 61.

- SOUZA, R.; LUCENA, O.; GARRAFA, J.; GOBBI, D.; SALUZZI, M.; APPENZELLER, S.; RITTNER, L.; FRAYNE, R.; LOTUFO, R. An open, multi-vendor, multi-field-strength brain mr dataset and analysis of publicly available skull stripping methods agreement. *NeuroImage*, v. 170, p. 482–494, 2018. ISSN 1053-8119. Segmenting the Brain. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1053811917306687>>. Citado na página 40.
- SPEAKMAN, T. Results of thalamotomy for parkinson's disease. *Canadian Medical Association Journal*, Canadian Medical Association, v. 89, n. 13, p. 652, 1963. Citado 2 vezes nas páginas 14 e 15.
- SU, J. H.; THOMAS, F. T.; KASOFF, W. S.; TOURDIAS, T.; CHOI, E. Y.; RUTT, B. K.; SARANATHAN, M. Thalamus optimized multi atlas segmentation (THOMAS): fast, fully automated segmentation of thalamic nuclei from structural MRI. *Neuroimage*, Elsevier, v. 194, p. 272–282, 2019. Citado 4 vezes nas páginas 20, 23, 78 e 79.
- TAHA, A. A.; HANBURY, A. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, Springer, v. 15, p. 1–28, 2015. Citado na página 46.
- TANG, X.; CROCETTI, D.; KUTTEN, K.; CERITOGU, C.; ALBERT, M. S.; MORI, S.; MOSTOFKY, S. H.; MILLER, M. I. Segmentation of brain magnetic resonance images based on multi-atlas likelihood fusion: testing using data with a broad range of anatomical and photometric profiles. *Frontiers in neuroscience*, Frontiers Media SA, v. 9, p. 61, 2015. Citado 2 vezes nas páginas 17 e 19.
- TASKER, R.; SIQUEIRA, J.; HAWRYLYSHYN, P.; ORGAN, L. What happened to VIM thalamotomy for parkinson's disease? *Stereotactic and Functional Neurosurgery*, S. Karger AG Basel, Switzerland, v. 46, n. 1-4, p. 68–83, 1983. Citado 2 vezes nas páginas 14 e 15.
- TASKER, R. R. Thalamotomy. *Neurosurgery Clinics of North America*, Elsevier, v. 1, n. 4, p. 841–864, 1990. Citado 2 vezes nas páginas 14 e 15.
- TOHKA, J. Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review. *World journal of radiology*, Baishideng Publishing Group Inc, v. 6, n. 11, p. 855, 2014. Citado na página 26.
- TREGIDGO, H. F.; SOSKIC, S.; ALTHONAYAN, J.; MAFFEI, C.; LEEMPUT, K. V.; GOLLAND, P.; INSAUSTI, R.; LERMA-USABIAGA, G.; CABALLERO-GAUDES, C.; PAZ-ALONSO, P. M. *et al.* Accurate bayesian segmentation of thalamic nuclei using diffusion MRI and an improved histological atlas. *NeuroImage*, Elsevier, v. 274, p. 120129, 2023. Citado 3 vezes nas páginas 23, 78 e 79.
- TUCH, D. S. High angular resolution diffusion imaging of the human brain. In: *Proc. of the International Society for the Magnetic Resonance in Medicine*. [S.l.: s.n.], 1999. Citado na página 32.
- UMAPATHY, L.; KEERTHIVASAN, M. B.; ZAHR, N. M.; BILGIN, A.; SARANATHAN, M. Convolutional neural network based frameworks for fast automatic segmentation of thalamic nuclei from native and synthesized contrast structural MRI. *Neuroinformatics*, Springer, p. 1–14, 2022. Citado 5 vezes nas páginas 21, 23, 38, 78 e 79.

- WANG, G.; LI, W.; AERTSEN, M.; DEPREST, J.; OURSELIN, S.; VERCAUTEREN, T. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. *Neurocomputing*, Elsevier, v. 338, p. 34–45, 2019. Citado 2 vezes nas páginas 74 e 75.
- WANG, G.; LI, W.; OURSELIN, S.; VERCAUTEREN, T. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In: *International MICCAI Brainlesion Workshop*. [S.l.: s.n.], 2017. p. 178–190. Citado na página 50.
- WANG, H.; SUH, J. W.; DAS, S. R.; PLUTA, J. B.; CRAIGE, C.; YUSHKEVICH, P. A. Multi-atlas segmentation with joint label fusion. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 35, n. 3, p. 611–623, 2012. Citado na página 16.
- WANG, S. L.; HAN, S.; CARASS, A.; ZHUO, J.; ROYS, S.; GULLAPALLI, R. P.; JIANG, L.; PRINCE, J. L. Thalamus segmentation using convolutional neural networks. In: *SPIE. Medical Imaging 2021: Image Processing*. [S.l.], 2021. v. 11596, p. 1159634. Citado 2 vezes nas páginas 18 e 19.
- WARFIELD, S. K.; ZOU, K. H.; WELLS, W. M. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE transactions on medical imaging*, IEEE, v. 23, n. 7, p. 903–921, 2004. Citado 2 vezes nas páginas 41 e 47.
- WESTBROOK, C. *Handbook of MRI technique*. [S.l.]: John Wiley & Sons, 2021. Citado na página 30.
- WINZECK, S.; MOCKING, S. J.; BEZERRA, R.; BOUTS, M. J.; MCINTOSH, E. C.; DIWAN, I.; GARG, P.; CHUTINET, A.; KIMBERLY, W. T.; COPEN, W. A. *et al.* Ensemble of convolutional neural networks improves automated segmentation of acute ischemic lesions using multiparametric diffusion-weighted MRI. *American Journal of Neuroradiology*, Am Soc Neuroradiology, v. 40, n. 6, p. 938–945, 2019. Citado na página 30.
- WU, M.; CHANG, L.-C.; WALKER, L.; LEMAITRE, H.; BARNETT, A. S.; MARENCO, S.; PIERPAOLI, C. Comparison of EPI distortion correction methods in diffusion tensor MRI using a novel framework. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008*, Springer Berlin Heidelberg, p. 321–329, 2008. Citado 2 vezes nas páginas 31 e 43.
- XIAO, Y.; BAILEY, L.; CHAKRAVARTY, M. M.; BERIAULT, S.; SADIKOT, A. F.; PIKE, G. B.; COLLINS, D. L. Atlas-based segmentation of the subthalamic nucleus, red nucleus, and substantia nigra for deep brain stimulation by incorporating multiple MRI contrasts. In: *SPRINGER. Information Processing in Computer-Assisted Interventions: Third International Conference, IPCAI 2012, Pisa, Italy, June 27, 2012. Proceedings 3*. [S.l.], 2012. p. 135–145. Citado na página 14.
- ZHANG, H.; SCHNEIDER, T.; WHEELER-KINGSHOTT, C. A.; ALEXANDER, D. C. Noddi: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, Elsevier, v. 61, n. 4, p. 1000–1016, 2012. Citado na página 32.
- ZHANG, M.; LU, Z.; FENG, Q.; ZHANG, Y. Automatic thalamus segmentation from magnetic resonance images using multiple atlases level set framework (malsf). *Scientific*

*Reports*, Nature Publishing Group UK London, v. 7, n. 1, p. 4274, 2017. Citado na página 19.

ZHANG, R.; ZHAO, L.; LOU, W.; ABRIGO, J. M.; MOK, V. C.; CHU, W. C.; WANG, D.; SHI, L. Automatic segmentation of acute ischemic stroke from dwi using 3-d fully convolutional densenets. *IEEE transactions on medical imaging*, IEEE, v. 37, n. 9, p. 2149–2160, 2018. Citado na página 30.

ZHAO, B.; DING, S.; WU, H.; LIU, G.; CAO, C.; JIN, S.; LIU, Z. Automatic acute ischemic stroke lesion segmentation using semi-supervised learning. *arXiv preprint arXiv:1908.03735*, 2019. Citado na página 30.

ZHOU, T.; RUAN, S.; CANU, S. A review: Deep learning for medical image segmentation using multi-modality fusion. *Array*, v. 3-4, p. 100004, 2019. ISSN 2590-0056. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2590005619300049>>. Citado na página 50.

ZIYAN, U.; TUCH, D.; WESTIN, C.-F. Segmentation of thalamic nuclei from dti using spectral clustering. In: SPRINGER. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part II 9*. [S.l.], 2006. p. 807–814. Citado 5 vezes nas páginas 20, 21, 23, 78 e 79.

ZOU, K. H.; WARFIELD, S. K.; BHARATHA, A.; TEMPANY, C. M.; KAUS, M. R.; HAKER, S. J.; III, W. M. W.; JOLESZ, F. A.; KIKINIS, R. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, Elsevier, v. 11, n. 2, p. 178–189, 2004. Citado na página 57.

# APÊNDICE A – Protocolo de segmentação manual

# Supplementary Materials: Thalamus Manual Segmentation Protocol

Gustavo Retuci Pinheiro<sup>1\*</sup>, Lorenza Brusini<sup>2</sup>, Diedre Carmo<sup>1</sup>, Renata Prôa<sup>1,3</sup>, Thays Abreu<sup>1</sup>, Simone Appenzeller<sup>4</sup>, Gloria Menegaz<sup>2</sup>, and Leticia Rittner<sup>1</sup>

## 1. Introduction

This document is supplementary material to the article "Thalamus Segmentation Using Deep Learning with Diffusion MRI Data: An Open Benchmark". It describes a thalamus segmentation protocol. The protocol considers both T1w and diffusion images, thus, taking advantage of diffusion properties of the brain tissue to distinguish the structures. In particular, the diffusion is considered through the FA (Fractional Anisotropy), however, other diffusion indices could be used and even combined.

For the protocol to work properly, the images from different MRI sequences must be aligned. If the acquisition sequences do not deliver the images in the same space, the alignment could be done by registering the images. Some popular tools for this task are the DIPY [1] and FSL-FLIRT [2]. In this tutorial, we used the ITK-SNAP [3] as a tool to draw the structures, but any other tool that has the option of opening multiple images simultaneously and drawing over them could be used for this purpose. For standardization, the left and right thalamus were assigned as label 10 and 49, respectively, following the look up table employed for FreeSurfer [4] and FSL.

Besides using diffusion combined with T1 to enhance the thalamus borders, another key element of this protocol is the use of multiple views to help the specialist to decide the level of contrast belonging to the thalamus.

## 2. Procedure

- **Step 1:** Load the images T1w (Fig. S1) and FA (Fig. S2) overlaying them with 50% transparency (Fig. S3). We can clearly see, for example, that where the thalamus neighbors the internal capsule the contrast in the T1w image is very dim, but it is clear in the FA image. When the two images are combined, the contrast is enhanced for the whole thalamus border.
- **Step 2:** Using the axial view, draw the borders of the thalamus in the slices where the borders are the easiest to identify (Figs. S4 and S5). The following landmarks were taken into consideration following the publication by [5]:
  - Medial border: the upper part of the lateral wall of the third ventricle and by the middle commissure that connects both thalami;
  - Lateral border: genu and the posterior limb of the internal capsule;
  - Anterior border: anterior commissure;
  - Posterior border: atrium of the lateral ventricle or the tail of the hippocampus, or both structures;
  - Superior border: lateral ventricles in anterior slices and the fornix in posterior slices.
  - Inferior border: *zona incerta* and the internal capsule.

This procedure should be done for at least one slice at the top of the thalamus, one in the mid-axial slice, and one near the bottom of the structure. For each segmented slice, the projection of the partial segmentation should appear on the other views, helping on the definition on what contrast levels should be considered as the borders. Conventionally, use the label 49 for the left thalamus.

- **Step 3:** With some slice defined on the axial view, the threshold of the thalamus borders on the sagittal view is considerably easier to spot, thus, the borders can now be drawn. With at least one slice drawn on the sagittal view, it is possible to proceed to the coronal view (Fig. S6).

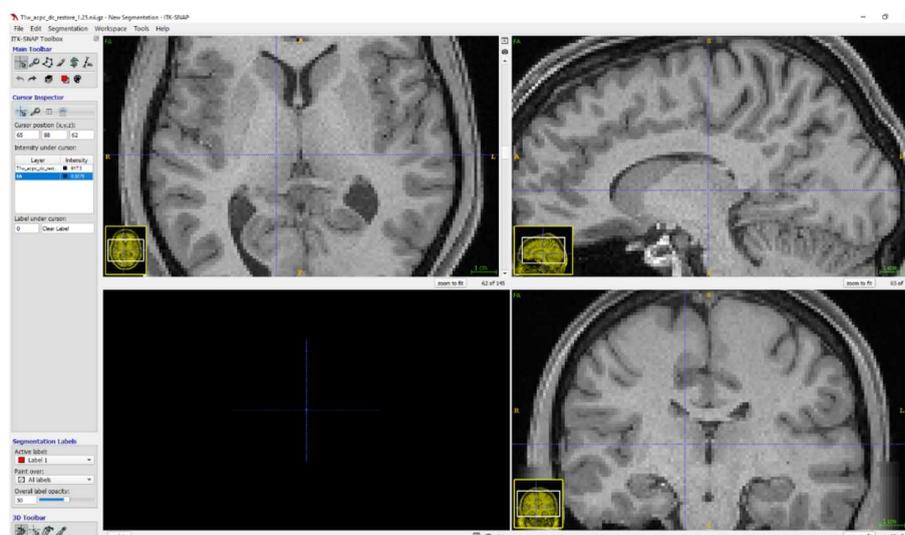


Figure S1. T1w image.

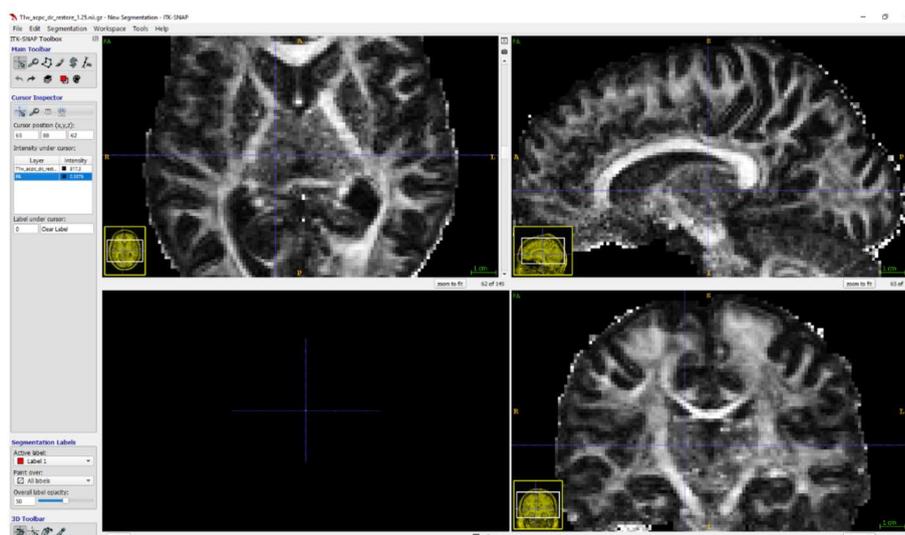


Figure S2. FA image.

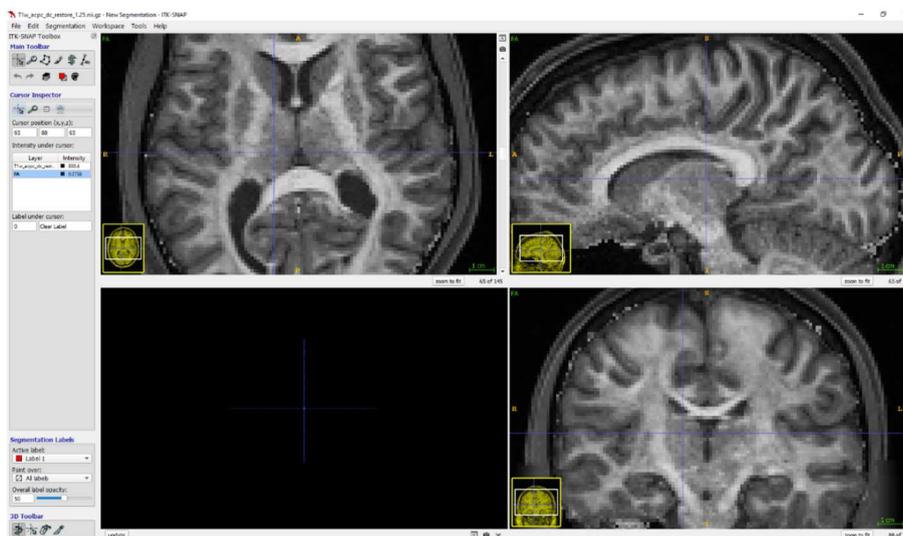


Figure S3. T1w and FA images combined.

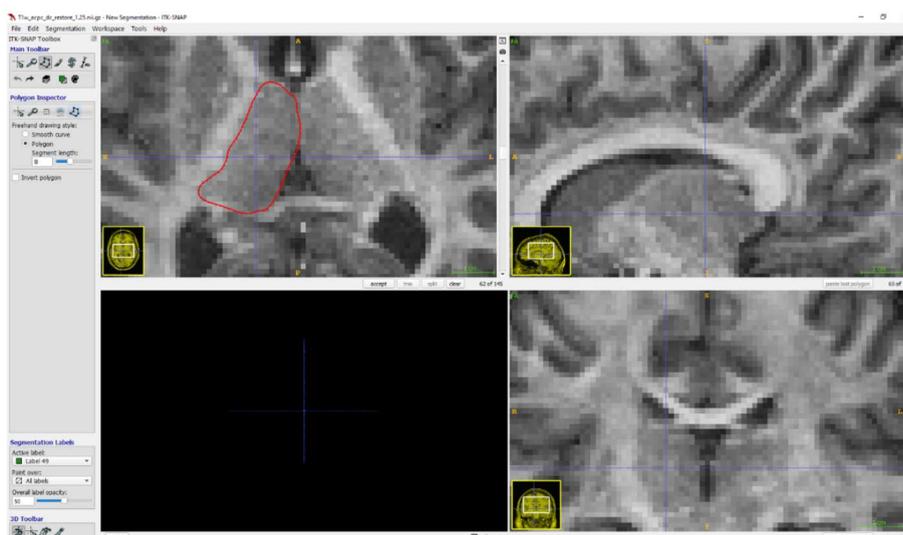


Figure S4. Outline of right thalamus on axial view.

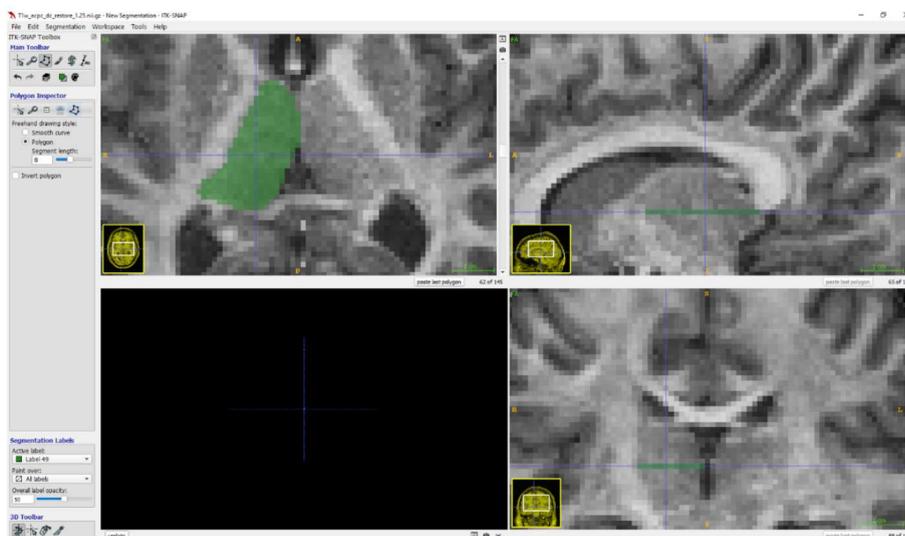


Figure S5. Filled thalamus slice on axial view and projection on coronal and sagittal views.

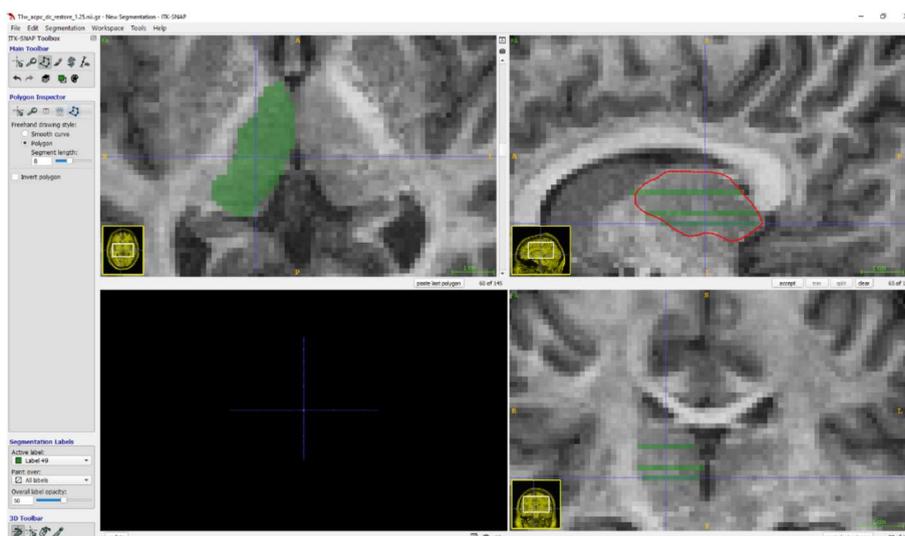
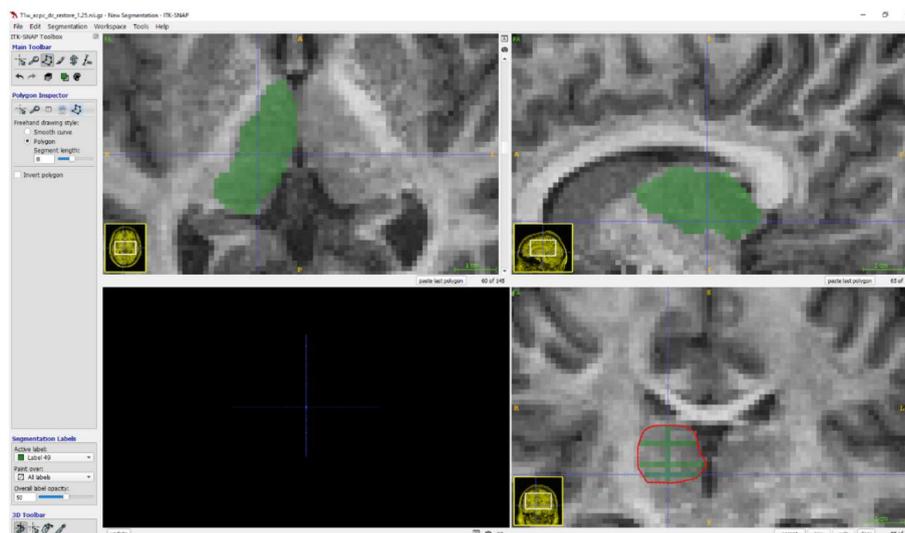
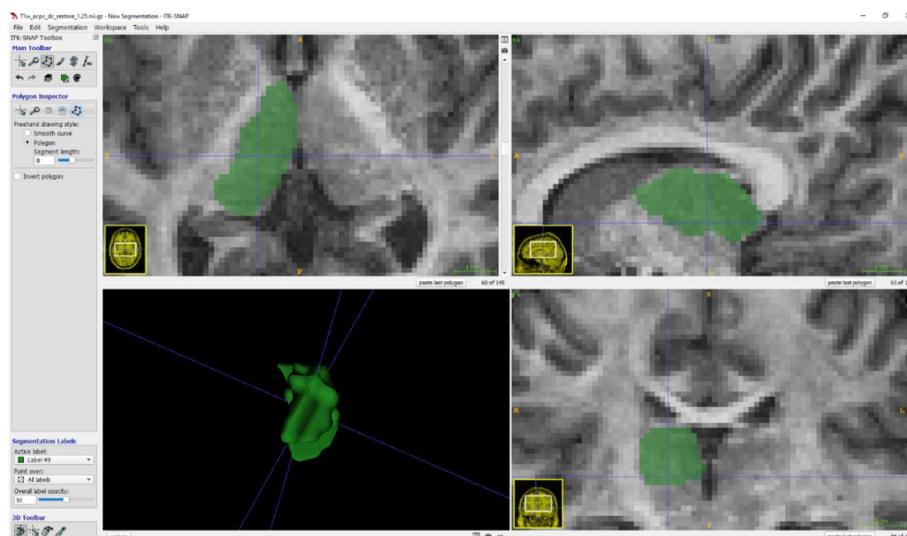


Figure S6. Filled thalamus slice on axial view, projection of thalamus filled slices on coronal and sagittal views, and thalamus outline on sagittal view.

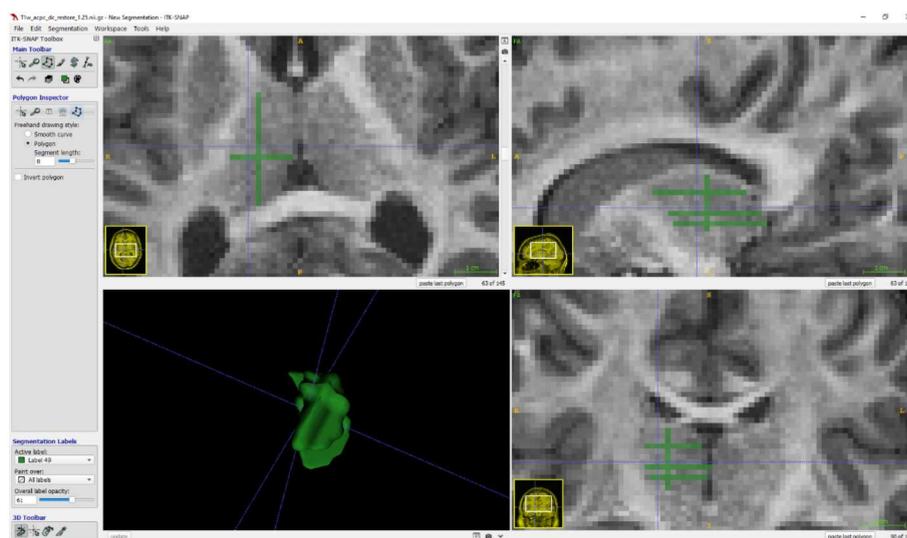


**Figure S7.** Filled thalamus slice on axial and sagittal views, projection of thalamus filled slice on coronal view, and thalamus outline on coronal view.

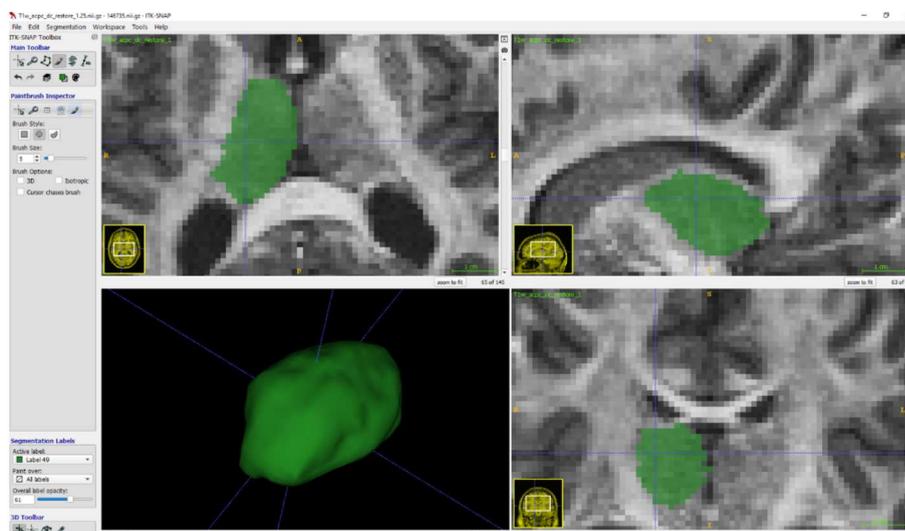
- **Step 4:** After defining slices on axial and sagittal views, it is possible to define the thalamus borders on the coronal view, specially the top and bottom borders that have the lowest contrasts (Fig. S7).
  - **Step 5:** At this point, a rough segmentation of the thalamus is achieved, but containing gaps among the segmented slices. To fill the gaps, the specialist should change the slices on each view and draw the outlines based on the projections of the slices drawn on other views. A 3D visualization of the current state of the segmentation is useful to spot gaps and to evaluate the shape of the segmentation (Figs. S8 and S9).
  - **Step 6:** Finally, it is recommended to check all slices, one by one and for all views, looking for voxels that were mistakenly marked, deleting them (Fig. S10). Again, the 3D visualization can help to spot those points as they usually appear as spikes.
  - **Step 7:** The same procedure described above should be repeated for the right thalamus (Figs. S11, S12, and S13). As a standard, use the label 10 assigned to the right thalamus.
1. Garyfallidis, E.; Brett, M.; Amirbekian, B.; Rokem, A.; Van Der Walt, S.; Descoteaux, M.; Nimmo-Smith, I. Dipy, a library for the analysis of diffusion MRI data. *Front. in Neuroinform.* **2014**, *8*, 8. <https://doi.org/10.3389/fninf.2014.00008>.
  2. Patenaude, B.; Smith, S.M.; Kennedy, D.N.; Jenkinson, M. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* **2011**, *56*, 907–922.
  3. Yushkevich, P.A.; Piven, J.; Cody Hazlett, H.; Gimpel Smith, R.; Ho, S.; Gee, J.C.; Gerig, G. User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability. *Neuroimage* **2006**, *31*, 1116–1128.
  4. Bruce, F. Freesurfer. *NeuroImage* **2012**, *62*, 774–781.
  5. Burggraaff, J.; Liu, Y.; Prieto, J.C.; Simoes, J.; de Sitter, A.; Ruggieri, S.; Brouwer, I.; Lissenberg-Witte, B.I.; Rocca, M.A.; Valsasina, P.; et al. Manual and automated tissue segmentation confirm the impact of thalamus atrophy on cognition in multiple sclerosis: A multicenter study. *NeuroImage: Clinical* **2021**, *29*, 102549.



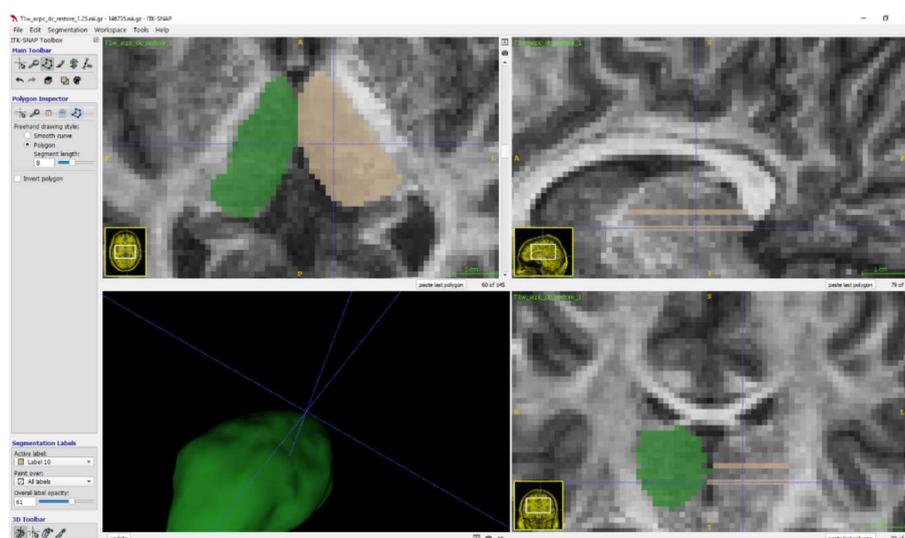
**Figure S8.** Filled thalamus slices on axial, coronal, and sagittal views, and 3D render of thalamus partial segmentation.



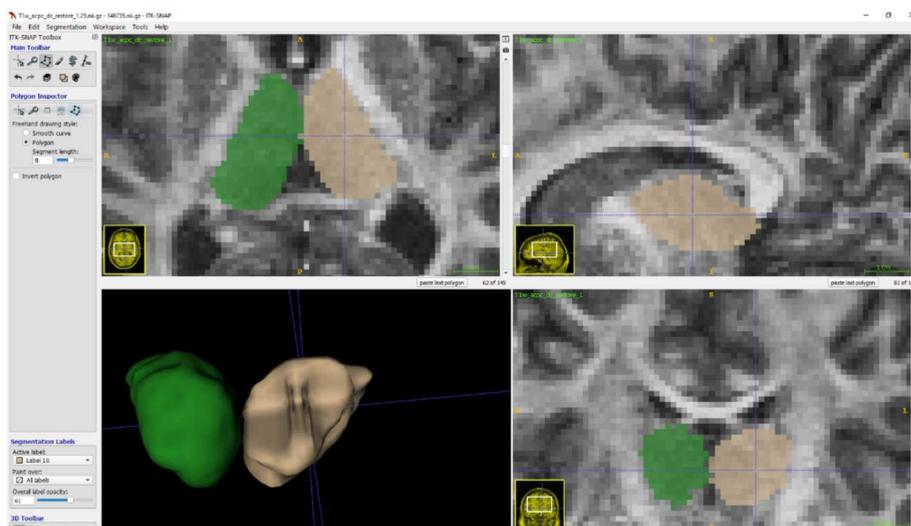
**Figure S9.** Projection of thalamus filled slices on coronal, axial, and sagittal views, and a 3D visualization of the current state of the thalamus segmentation.



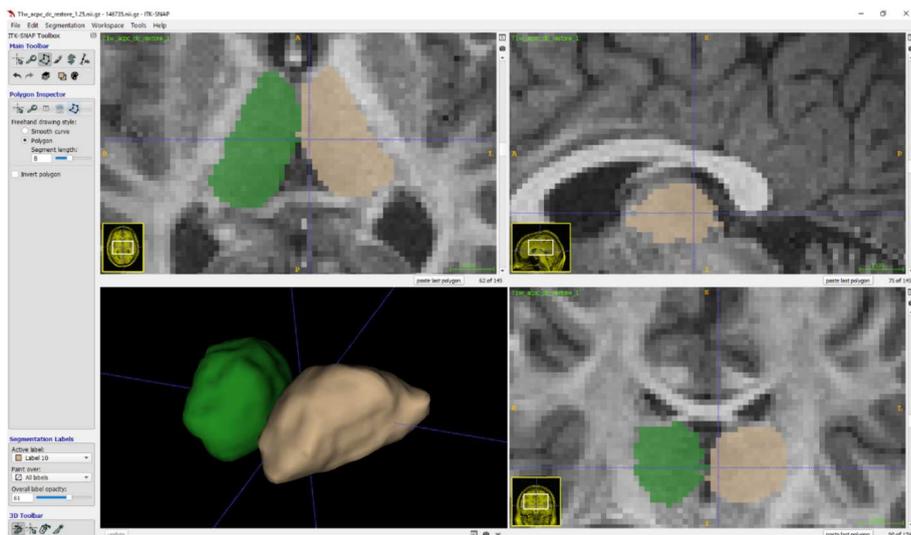
**Figure S10.** Segmentation of the left thalamus on 3 views (axial, sagittal, and coronal) and a 3D representation of the segmentation.



**Figure S11.** Segmentation of the left thalamus on 3 views, a 3D representation of the left thalamus, and two axial slices of the right thalamus.



**Figure S12.** Segmentation of the left thalamus on 3 views, a 3D representation of the left thalamus, and partial segmentation of the right thalamus.



**Figure S13.** Segmentation of both Thalami on 3 views and a 3D representation of both segmentations.