

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



Machine Learning para Segmentação de Lesões de Esclerose Múltipla em Imagens de Ressonância Magnética

Isabella Medeiros de Sousa

Mestrado em Engenharia Biomédica

Orientador: Prof. Dr. Jaime Cardoso

Co-orientadora: Prof. Dra Marcela de Oliveira

16 de Julho de 2021

Machine Learning para Segmentação de Lesões de Esclerose Múltipla em Imagens de Ressonância Magnética

Isabella Medeiros de Sousa

Mestrado em Engenharia Biomédica

Faculdade de Engenharia da Universidade do Porto

16 de Julho de 2021

Resumo

A Esclerose Múltipla (E.M.) é uma doença crônica e inflamatória do sistema nervoso central, que provoca lesões na substância branca no cérebro e na medula espinhal, em decorrência da desmielinização dos axônios. Essas lesões podem ser observadas na imagem de ressonância magnética (MRI do inglês *Magnetic Resonance Image*), sendo esta uma grande aliada no diagnóstico. Em adição, medidas quantitativas dessas lesões são úteis para acompanhar a progressão da doença, bem como o seu tratamento. A segmentação manual de lesões realizada por um especialista ainda é considerada o padrão ouro, embora seja um processo que consome muito tempo e está sujeito a variabilidade intra e inter observadores. Dessa forma, esforços têm sido feitos no desenvolvimento de algoritmos de Aprendizado de Máquina para segmentar lesões e poder auxiliar no diagnóstico assistido por computadores.

Esse trabalho de mestrado teve como objetivo explorar técnicas de processamento de imagens de ressonância magnética (MRI) com aplicação de métodos de Aprendizado Profundo para segmentar lesões escleróticas. Diante disso, foi realizada uma revisão bibliográfica sobre a Esclerose Múltipla, um estudo de técnicas de pré-processamento das imagens, bem como a aplicação e implementação de Redes Neurais Convolucionais para segmentação automática das lesões. Os modelos de segmentação foram testados em MRIs (FLAIR e T1) de 20 pacientes diagnosticados com Esclerose Múltipla. O conjunto de segmentações de cada modelo diferente foi comparado estatisticamente, através das métricas Dice, Valor Preditivo Positivo e Diferença Absoluta de Volume. Os resultados indicaram que os métodos de segmentação automatizada, por Redes Neurais Convolucionais, são capazes de segmentar as lesões escleróticas com boa acurácia e têm grande potencial de ser utilizados no futuro em clínicas médicas.

Abstract

Multiple Sclerosis (M.S.) is a chronic and inflammatory disease of the central nervous system, which causes white matter lesions in the brain and spinal cord, due to the demyelination of the axons. These lesions can be seen on the magnetic resonance image (MRI), being a great ally in the diagnosis. In addition, quantitative measures of these lesions are good to monitor the disease's progression and its treatment. Manual lesion segmentation made by a specialist is still considered the gold standard, although it is a time-consuming, hence an error-prone process. Thus, investments have been made in the development of machine learning algorithms to segment and support the computer-assisted diagnosis.

This master's work aims to explore techniques for processing magnetic resonance images and to analyze applications of deep learning methods to segment multiple sclerosis lesions. As a result, bibliographic review on Multiple Sclerosis was written, image pre-processing techniques was studied and application of Convolutional Neural Networks was performed to segment lesions, automatically. The segmentation models were tested on MRIs (FLAIR and T1) of 20 patients diagnosed with Multiple Sclerosis. The set of segmented images of each different model was compared statistically, through the metrics Dice index, Predictive Positive Value and Absolute Volume Difference. The results indicated that the automated segmentation methods, by Convolutional Neural Networks, are capable of segmenting sclerotic lesions with good accuracy and have great potential to be used in medical clinics in the future.

Agradecimentos

À minha orientadora *Prof.^a Dr.^a* Marcela de Oliveira por toda dedicação e paciência durante meu processo de aprendizado, sendo fundamental para a realização deste trabalho.

Ao meu orientador Prof. Dr. Jaime Cardoso pela oportunidade de realizar este projeto, por toda sua presteza, dedicação e conhecimento transmitido. Também gostaria de agradecer ao INESC TEC por possibilitar a execução deste trabalho.

Ao colega Dr. Ricardo Cruz, por ser tão solícito e ter contribuído para realização desta dissertação.

A meus pais, Silmário Sousa e Francimary Medeiros, por sempre acreditar em meu potencial e me incentivar a realizar os meus sonhos.

À minha psicóloga Karina Amorim pelo suporte emocional que tanto necessitei nesse período.

Às minhas amigas por todo apoio, força e incentivo. Em especial a Cintya Moraes, Izabela Dória, Joanna Paiva, Letícia Lessa e Roberta Escudeiro.

Aos amigos que fiz em Portugal, em especial, Julio Oliveira, Alejandra Cortez, Nathássia Marques e Ana Fonseca. Agradeço por todo apoio e cumplicidade.

À minha família, em especial, meu avô Francisco Medeiros, que sempre estará em meu coração, minhas avós Maria José de Sousa e Maria de Lourdes Maciel e à minha madrinha Lucimary Medeiros.

Isabella Medeiros de Sousa

*Someday I'll fly
Someday I'll soar
Someday I'll be so much more
Cause I'm bigger than my body gives me credit for*

John Mayer

Conteúdo

1	Introdução	1
1.1	Contextualização	1
1.2	Objetivos	2
1.2.1	Objetivos Específicos	2
1.3	Estrutura da Dissertação	2
2	Fundamentos da Esclerose Múltipla	3
2.1	Definição	3
2.2	Epidemiologia	4
2.3	Tipos de Esclerose Múltipla	5
2.4	Diagnóstico	5
2.5	Tratamento	7
3	Imagem de Ressonância Magnética	9
3.1	Formação da Imagem de Ressonância Magnética	9
3.1.1	Princípios Físicos da Ressonância Magnética	10
3.1.2	Tempos de Relaxação	11
3.1.3	Sequências das imagens	13
3.2	Características Imagiológicas das Lesões de Esclerose Múltipla	14
4	Pré-processamento e Análise de Imagens	16
4.1	Imagem Médica	16
4.2	Pré-Processamento	16
4.2.1	Remoção de Ruídos	17
4.2.2	Correção de Não Homogeneidade	17
4.2.3	Normalização de Intensidade	18
4.2.4	Extração de estruturas não cerebrais	18
4.2.5	Registro das Imagens	19
4.3	Segmentação das Lesões Encefálicas	20
5	Machine Learning	21
5.1	Introdução	21
5.2	Convolutional Neural Networks	22
5.2.1	Camadas de Convolução	22
5.2.2	Camada de Subamostragem	24
5.2.3	Camadas Totalmente Conectadas	24
5.2.4	Última camada	25
5.2.5	Treino das CNNs	25

5.3	Arquiteturas	28
5.3.1	U-Net	28
5.3.2	CNN em Cascata	29
5.3.3	CNN com Concatenação	30
5.4	Revisão sobre Trabalhos de Segmentação de Lesões de Esclerose Múltipla	30
5.4.1	HeMIS: Hetero-Modal Image Segmentation	30
5.4.2	Longitudinal Multiple Sclerosis Lesion Segmentation Using Multi-view Convolutional Neural Networks	31
5.4.3	Fast Lesion EXtraction using CONvolutional Neural Networks - FLEX-CONN	32
5.4.4	One-shot Domain Adaptation In Multiple Sclerosis Lesion Segmentation Using Convolutional Neural Networks	34
5.4.5	A Light Weighted Deep Learning Framework for Multiple Sclerosis Lesion Segme	36
5.4.6	A Dense U-Net Architecture for Multiple Sclerosis Lesion Segmentation	37
5.4.7	Multiple Sclerosis Lesion Segmentation With Tiramisu And 2.5D Stacked Slices	38
5.5	Pós-processamento da imagem segmentada	40
5.6	Considerações Finais	40
6	Metodologia	42
6.1	Banco de dados	42
6.2	Pré-processamento de imagem	43
6.2.1	Extração de Estruturas Não-cerebrais	43
6.2.2	Registro e Reamostragem das Imagens	43
6.2.3	Correção de Não-homogeneidade	43
6.2.4	Normalização das Intensidades	43
6.3	Treinamento dos modelos	44
6.3.1	FLEXCONN	44
6.3.2	Light weighted Unet	44
6.3.3	NicMSlesions	44
6.3.4	Modelo proposto: New Light Weighted Unet	45
6.4	Segmentação das Lesões e Pós-processamento	45
6.5	Métricas comparativas	46
7	Resultados e Discussão	47
7.1	Resultados do Pré-processamento das Imagens	47
7.1.1	Extração das Estruturas Não-cerebrais	47
7.1.2	Registro das imagens	48
7.1.3	Correção de não-homogeneidade	48
7.2	Segmentação das Lesões Escleróticas	49
7.2.1	Segmentação FLEXCONN	49
7.2.2	Segmentação Light Weighted Unet	53
7.2.3	NicMSlesions	57
7.3	Comparações Entre os Modelos Treinados	65
8	Conclusões	67
8.1	Conclusões Finais	67
8.2	Limitações e Trabalhos Futuros	68

Lista de Figuras

2.1	Comparação entre neurônio saudável e danificado [12].	3
2.2	Prevalência mundial da Esclerose Múltipla (Adaptada de [1]).	4
2.3	Subtipos clínicos de E.M. [2]	5
2.4	Crêterios de McDonalds para diagnóstico de E.M. (Adaptado de [17])	6
3.1	Representação da MRI em voxels ([22] citado por [20]).	9
3.2	Diferentes planos de obtenção de imagens (Adaptada de [20]).	10
3.3	Representação de um próton (Adaptada de [24]).	10
3.4	Representação das diferentes orientações dos átomos de hidrogênio [2]	11
3.5	Retorno do vetor magnetização ao equilíbrio [24]	12
3.6	Curvas de relaxação T1 (a) e T2 (b)[25]	12
3.7	Simple sequêcia de pulsos (Adaptada de [23])	13
3.8	Sequências diferentes de MRI: T1, T2 e FLAIR (Adaptada de [2]).	14
3.9	Lesões aparentes em MRI [13]	15
4.1	Correção de intensidade não-uniforme (Adaptada de [34]).	17
4.2	Extração de estruturas não cerebrais [41].	18
4.3	Diagrama do processo de registro (Adaptada de [38]).	19
5.1	Blocos de uma CNN (Adaptada de [50]).	22
5.2	Representação da operação de convolução (Adaptada de [56]).	23
5.3	Aplicação da função ReLU [57]).	24
5.4	Exemplo de operação de MaxPooling [56].	24
5.5	Divisão do banco de dados (Adaptada de [56]).	25
5.6	Visão geral do processo de treinamento de uma CNN (Adaptada de [56])	26
5.7	Influência da taxa de aprendizagem com valores diferentes (Adaptada de [63])	27
5.8	Arquitetura U-Net [66].	29
5.9	CNN em cascata [68]	29
5.10	Arquitetura de CNNs com concatenação (Adaptada de [69]).	30
5.11	Arquitetura Hetero-modal para segmentação [70].	31
5.12	Arquiteturas CNNs utilizadas [74].	32
5.13	Arquitetura CNN proposta (Adaptada de [8]).	33
5.14	Arquitetura da CNN do modelo base [39].	34
5.15	Estrutura de adaptação de domínio de intensidade supervisionada [39].	35
5.16	Arquitetura U-net adaptada [77].	37
5.17	Arquitetura U-Net adaptada com camadas densas [71]	38
5.18	Arquitetura Tiramisu proposta [78].	39
5.19	Redução de falso positivos ([81] citado por [33]).	40

6.1	Arquitetura Light Weighted Unet modificad.	45
7.1	Cortes das sequências FLAIR e T1 antes e após a extração do encéfalo	47
7.2	Comparação entre cortes das imagens FLAIR e T1.	48
7.3	Cortes das imagens FLAIR e T1 antes e após a correção da não-homogeneidade.	48
7.4	Boxplot da variação de coeficiente Dice.	49
7.5	Cortes das imagens segmentadas do Paciente 10 pelos modelos FLEXCONN.	50
7.6	Gráficos de dispersão da quantificação volumétrica das lesões.	51
7.7	Comparação da segmentação entre os dois modelos FLEXCONN treinados.	52
7.8	Boxplot da variação de coeficiente Dice.	53
7.9	Gráficos de dispersão da quantificação volumétrica das lesões pelos três modelos Light Weighted Unet implementados.	54
7.10	Cortes das imagens segmentadas do Paciente 10 pelos modelos <i>Light Weighted Unet</i>	55
7.11	Comparação da segmentação entre os três modelos <i>Light Weighted Unet</i> treinados.	56
7.12	Boxplot da variação dos coeficientes Dice dos modelos NicMSlesions de treinamento completo.	57
7.13	Gráficos de dispersão da quantificação volumétrica das lesões pelos modelos NicMSlesions implementados.	59
7.14	Cortes das imagens segmentadas do Paciente 10 pelos modelos NicMSlesions.	59
7.15	Resultados da segmentação dos dois modelos nicMSlesions treinados completamente.	60
7.16	Boxplot da variação dos coeficientes Dice dos modelos NicMSlesions por transferência de aprendizado	61
7.17	Gráficos de dispersão da quantificação volumétrica das lesões pelos modelos NicMSlesions, treinados por transferência de aprendizado.	62
7.18	Cortes das segmentações NicMSlesions, treinados por transferência de aprendizado.	63
7.19	Resultados da segmentação dos modelos nicMSlesions treinados por transferência de aprendizado.	64

Lista de Tabelas

3.1	Tempos de relaxação T1 e T2 aproximados para diversos tecidos do corpo humano a 1,5 T [24].	13
3.2	Cores e tons dos tecidos nas imagens T1, T2 e FLAIR [27]	14
5.1	Características dos bancos de filtros utilizados no trabalho [8]	33
5.2	Coeficientes Dice obtidos a partir da segmentação de imagens em banco de dados simulados, utilizando diferentes funções de perda para cada treinamento [78] . . .	39
5.3	Coeficientes Dice médio obtidos a partir dos diferentes trabalhos de segmentação de lesões escleróticas.	41
7.1	Médias e desvios padrão dos valores de DV das duas versões do modelo FLEX-CONN.	50
7.2	Medianas calculadas das amostras de VPP para cada modelo Unet adaptado. . . .	54
7.3	Medianas dos valores de DV das diferentes implementações do modelo Unet adaptado.	54
7.4	Médias, desvios padrão e medianas do VPP de cada modelo.	58
7.5	Médias, desvios padrão e medianas dos valores de DV das duas versões do modelo NicMSLesions de treinamento completo.	58
7.6	Médias, desvios padrão e medianas dos valores de VPP das duas versões do modelo NicMSLesions, treinados através da transferência de aprendizado.	62
7.7	Médias, desvios padrão e medianas dos valores de DV das duas versões do modelo NicMSLesions, treinados através da transferência de aprendizado.	62
7.8	Medianas das métricas avaliativas das segmentações dos melhores modelos implementados.	65
7.9	Teste de Dunn para análise do VPP.	65
7.10	Teste de Dunn para análise de DV.	66

Abreviaturas

BET	do inglês <i>Brain Extraction Tool</i>
BRATS	do inglês <i>Brain Tumor Segmentation Challenge</i>
CNN	do inglês <i>Convolutional Neural Networks</i>
CSF	do inglês <i>Cerebrospinal Fluid</i>
DICOM	do inglês <i>Digital Imaging and Communications in Medicine</i>
DSC	do inglês <i>Dice Similarity Coefficient</i>
DV	Diferença Absoluta de Volume
EBV	do inglês <i>Epstein-Barr Virus</i>
E.M.	Esclerose Múltipla
E.M.R.R.	Esclerose Múltipla Remissiva-Recorrente
E.M.P.P.	Esclerose Múltipla Progressiva Primária
E.M.P.R.	Esclerose Múltipla Progressiva Recorrente
E.M.P.S.	Esclerose Múltipla Secundária Progressiva
FCN	do inglês <i>Fully Convolutional Network</i>
FLAIR	do inglês <i>Fluid Attenuated Inversion Recovery</i>
FN	Falso Negativo
FP	Falso Positivo
GT	do inglês <i>Ground Truth</i>
ISBI	do inglês <i>International Symposium on Biomedical Imaging</i>
MICCAI	do inglês <i>Medical Image Computing and Computer Assisted Intervention</i>
MRI	do inglês <i>Magnetic Resonance Image</i>
NIFTI	do inglês <i>Neuroimaging Informatics Technology Initiative</i>
ReLU	do inglês <i>Rectified Linear Unit</i>
RF	Radiofrequência
SNC	Sistema Nervoso Central
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
VPP	Valor Preditivo Positivo

Capítulo 1

Introdução

Neste capítulo, apresenta-se uma contextualização sobre o tema estudado, os objetivos desse projeto e por fim a estrutura desta dissertação.

1.1 Contextualização

A Esclerose Múltipla (E.M.) é uma doença crônica e inflamatória do sistema nervoso central (SNC), que provoca lesões na substância branca no cérebro e na medula espinhal, em decorrência da desmielinização dos axônios [1]. A mielina é uma substância que rodeia o axônio, formando uma camada eletricamente isolante, tendo um fundamental papel na comunicação entre neurônios, facilitando a transmissão dos sinais elétricos [2].

A partir do processo da desmielinização, há formação de tecido fibroso ao redor das lesões (gliose) [3]. A localização do SNC afetado por essas lesões influencia as características patológicas da doença [4]. Clinicamente, a Esclerose Múltipla apresenta grande variabilidade de manifestações, podendo afetar a mobilidade, equilíbrio, visão e cognição [5]. A E.M. pode ser classificada em quatro tipos: Remissiva-recorrente (E.M.R.R.), Secundária Progressiva (E.M.P.S.), Progressiva Primária (E.M.P.P.) e Progressiva Recorrente (E.M.P.R.) [6]

Do ponto de vista radiológico, o processo de desmielinização pode ser visualizado através da imagem de ressonância magnética. A ressonância magnética é o tipo de exame por imagem mais utilizado no diagnóstico da E.M., devido a sua grande capacidade de mostrar lesões causadas pela doença mesmo na ausência de sintomas [7]. A segmentação manual de lesões em imagens de ressonância magnética realizada por um especialista é considerada o padrão ouro atualmente, embora esse método requiera bastante tempo e é suscetível a variabilidade intra e inter observadores [8] [9].

Nos últimos anos, projetos de Aprendizado de Máquina (em inglês *Machine Learning*), mais especificamente Aprendizado Profundo (em inglês *Deep Learning*) têm obtido bons desempenhos em processamento de imagens médicas e em diversas aplicações clínicas [10]. Assim, a identificação e segmentação de lesões de Esclerose Múltipla por meios automáticos têm sido uma grande

área de investigação [11]. Essas tecnologias podem ser utilizadas tanto no diagnóstico da doença como no acompanhamento da progressão da doença e avaliação do tratamento.

1.2 Objetivos

Esse trabalho de mestrado teve como objetivo explorar técnicas de processamento de imagens de ressonância magnética com aplicação de métodos de *Deep Learning* para segmentar lesões escleróticas.

1.2.1 Objetivos Específicos

Essa dissertação de mestrado teve os seguintes objetivos específicos:

- Explorar a literatura existente com modelos de segmentação das lesões escleróticas no encefalo;
- Explicar as principais etapas de pré-processamento de imagem de ressonância magnética, com a finalidade de segmentar as lesões por meios automatizados;
- Analisar técnicas de Redes Neurais Convolucionais para segmentação das lesões em MRI;
- Implementar uma metodologia de análise de imagem de ressonância magnética para segmentação automática de lesões em pacientes com Esclerose Múltipla.

1.3 Estrutura da Dissertação

Essa dissertação está dividida em 8 capítulos. No presente capítulo, é feita uma contextualização do tema e sua problemática. O capítulo 2 apresenta fundamentos da Esclerose Múltipla, abordando fisiopatologia, epidemiologia, classificação e diagnóstico. O capítulo 3 aborda princípios básicos de ressonância magnética e caracterização imagiológica das lesões. O capítulo 4 apresenta técnicas de processamento e análise de imagens de ressonância magnética. O capítulo 5 aborda conceitos fundamentais sobre Redes Neurais Convolucionais (do inglês *Convolutional Neural Networks*). O capítulo 6 descreve a metodologia computacional utilizada deste estudo. O capítulo 7 apresenta os resultados obtidos, seguido de discussão. Por fim, o capítulo 8 apresenta as conclusões e recomendações para trabalhos futuros.

Capítulo 2

Fundamentos da Esclerose Múltipla

Este capítulo aborda conceitos fundamentais sobre a Esclerose Múltipla. Inicialmente, é apresentada sua definição, seguida de uma caracterização epidemiológica, classificação e critérios de diagnóstico.

2.1 Definição

A Esclerose Múltipla é uma doença crônica, autoimune e neurodegenerativa do sistema nervoso central, caracterizada pela desmielinização dos axônios na substância branca no cérebro e na medula espinhal [5]. É uma doença onde células do próprio sistema imunológico deterioram a bainha de mielina, cuja função é envolver e isolar os axônios, permitindo que os sinais elétricos sejam conduzidos rapidamente. Assim, esse processo de degradação da bainha de mielina compromete a transmissão de sinais entre os neurônios [12]. Na Figura 2.1, é possível observar a diferença entre um neurônio saudável e um em processo de desmielinização.

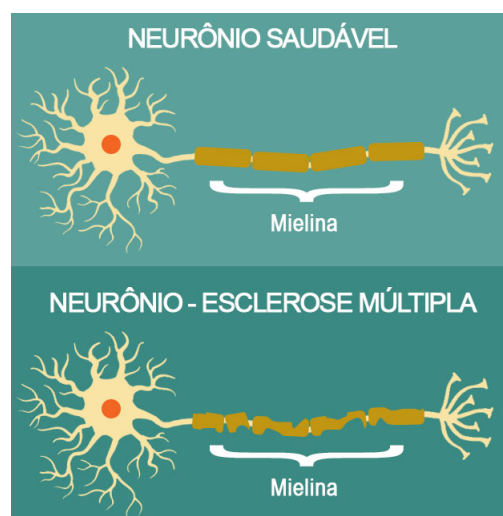


Figura 2.1: Comparação entre neurônio saudável e danificado [12].

As lesões evoluem com o tempo e podem provocar sintomas no sistema sensorial, na mobilidade, equilíbrio, visão e cognição [13]. Embora seja uma doença conhecida há bastante tempo, a E.M. tem uma origem complexa, envolvendo aspectos imunológicos, ambientais e interações genéticas [6]. Alguns fatores como baixos níveis de vitamina D, cigarro, obesidade infantil e infecção pelo vírus Epstein-Barr (um vírus da família da herpes), combinados com fatores genéticos e ambientais, estão relacionados com a causa e o desenvolvimento da doença [6].

2.2 Epidemiologia

Os estudos epidemiológicos estimam que há 2,8 milhões de pessoas com E.M. [1]. Na Figura 2.2 observa-se a distribuição de casos de E.M. a nível mundial. Nota-se que existe uma relação geográfica com a ocorrência da doença, sendo mais comum em pessoas que vivem afastadas da linha do Equador [1]. Isso pode estar relacionado com a baixa exposição a luz solar, e consequentemente baixos níveis de Vitamina D [6]. Dados apontam que a taxa de mulheres com E.M. é de duas a três vezes maior em comparação aos homens, e que a incidência é maior em jovens adultos (25 a 35 anos), sendo raros casos em crianças e idosos [2].

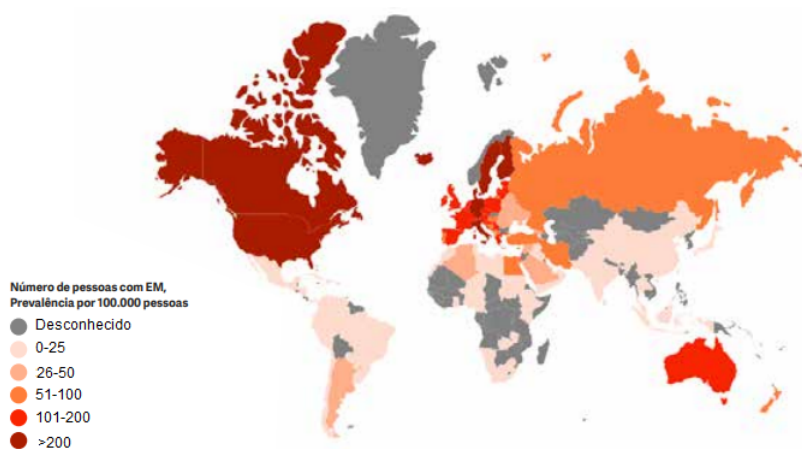


Figura 2.2: Prevalência mundial da Esclerose Múltipla (Adaptada de [1]).

A E.M. pode ser considerada uma doença de complexidade genética, onde o indivíduo que possui fatores genéticos suscetíveis à tal doença pode desenvolvê-la a partir das interações ambientais e infecções virais [6]. Um grande fator de risco para desencadeamento da E.M. é a infecção pelo vírus Epstein-Barr (EBV). Embora essa associação não seja muito clara, o risco de desenvolver E.M. é em torno de 15 vezes maior para indivíduos infectados na infância e cerca de 30 vezes maior entre os infectados na adolescência ou na vida adulta [14]. Além disso, outros fatores que podem influenciar são: baixo níveis de vitamina D (como mencionado anteriormente), cigarro e obesidade infantil [15].

2.3 Tipos de Esclerose Múltipla

A Esclerose Múltipla apresenta características clínicas bastante variáveis, podendo ser considerada uma doença com extenso espectro desde as fases inflamatórias até as fases neurodegenerativas [5]. Nos estágios iniciais, há certa preservação dos axônios, mas à medida que a doença se desenvolve, há perdas axonais irreversíveis, sendo um dos principais contribuintes para causar sequelas irreversíveis [6]. A E.M. pode ser classificada em quatro subtipos clínicos: remissiva-recorrente (E.M.R.R.), secundária-progressiva (E.M.P.S.), progressiva-primária (E.M.P.P.) e progressiva-recorrente (E.M.P.R.) [6]. A Figura 2.3 representa a evolução clínica em relação ao aumento da incapacidade do indivíduo no decorrer do tempo para os diferentes subtipos da E.M.

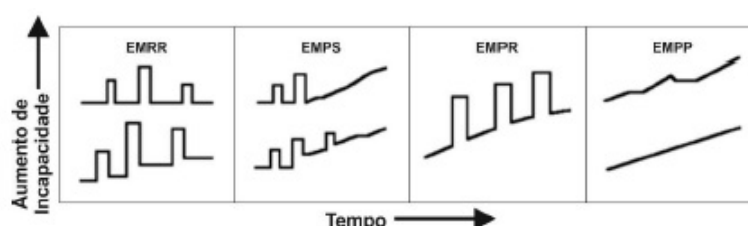


Figura 2.3: Subtipos clínicos de E.M. [2]

A E.M.R.R. é o subtipo mais frequente, afetando cerca de 85% a 90% de pacientes diagnosticados. Este subtipo está relacionado com ataques de desmielinização no SNC, consequentemente há presença de surtos definidos (episódios de disfunção neurológicas) com duração mínima de 24 horas, seguida de recuperação total ou parcial, mantendo-se estável entre ataques [16]. Enquanto a E.M.P.S. é considerada a segunda fase de E.M.R.R., geralmente se desenvolve entre 10-15 anos após início dos sintomas [6]. Nesse caso, os pacientes apresentam uma piora progressiva dos sintomas com ou sem recaídas sobrepostas [2]. A E.M.P.R. se superpõe a E.M.P.P. e E.M.S.P., sendo caracterizada por uma evolução progressiva dos sintomas, combinada com surtos ocasionais [2]. Por fim, a E.M.P.P. afeta em torno de 10% a 15% dos pacientes e destaca-se por um crescente grau de desabilidade com presença de surtos esporádicos [16].

2.4 Diagnóstico

O diagnóstico da E.M. continua predominantemente clínico, o qual é difícil de detectar precocemente, dada a variações nos intervalos de tempo em que os sintomas aparecem [2]. As manifestações iniciais de E.M. variam de acordo com a localização das lesões e se apresenta sintomas relapsos ou progressivos [13]. Os sinais e sintomas que ocorrem em pacientes diagnosticados com E.M.R.R são variados e imprevisíveis, uma vez que as lesões podem formar-se em qualquer local do SNC [15]. Entretanto, alguns sintomas iniciais comuns da doença são fraqueza em um ou mais membros, distúrbio sensorial, perda da visual monocular, visão dupla, instabilidade na marcha e ataxia [7].

Quando há suspeita de E.M., as avaliações clínicas, a análise de MRI e o estudo do líquido cefalorraquidiano (CSF) são métodos que auxiliam na concretização de um diagnóstico [6]. Com a finalidade de evitar procedimentos invasivos, exames de ressonância magnética geralmente são requisitados, pois esse tipo de exame por imagem é muito específico para visualização do processo de desmielinização. Além disso, a análise de imagens de ressonância magnética é útil para descartar outras patologias neurológicas [13].

Assim, por padrão, o diagnóstico da E.M. é realizado a partir do critério de McDonalds [6], representado na Figura 2.4. Para a definição desse critério, o neurologista Ian McDonald propôs que as análises das lesões na MRI, juntamente com as avaliações clínicas, permitem um diagnóstico mais antecipado, mesmo a partir de um único surto [17]. Dessa forma, há aumento da sensibilidade no diagnóstico sem comprometer a especificidade.

Ataques	Presença de lesões	Critério adicional necessário
Pelo menos 2 surtos	No mínimo 2 lesões	Nenhum
Pelo menos 2 surtos	No mínimo 1 lesão	<u>Disseminação no espaço:</u> No mínimo 1 lesão T2 em pelo menos 2 das seguintes regiões: periventricular, justacortical, infratentorial, medula espinal.
1 surto	No mínimo 2 lesões	<u>Disseminação no tempo:</u> Presença simultânea de lesões assintomáticas; Nova lesão em T2 em MRI seguinte; Aguardo de segundo surto;
1 surto	1 lesão	<u>Disseminação no tempo:</u> Presença simultânea de lesões assintomáticas; Nova lesão em T2 em MRI seguinte; Aguardo de segundo surto; <u>Disseminação no espaço:</u> No mínimo 1 lesão T2 em pelo menos 2 das seguintes regiões: periventricular, justacortical, infratentorial ou medula espinal; ou Aguardo de nova lesão em diferente localização no SNC.
Zero surtos	-	<u>1 ano de progressão da doença e mais 2 de 3 dos seguintes critérios:</u> a) Evidência de disseminação espacial no cérebro; b) Evidência de disseminação espacial no medula espinal; c) Evidência de bandas oligoclonais por focalização isoeétrica e/ou aumento de índice IgG.

Figura 2.4: Critérios de McDonalds para diagnóstico de E.M. (Adaptado de [17])

2.5 Tratamento

Até a presente data, não existe uma cura para a E.M. [7]. Os tratamentos têm o objetivo de desacelerar a progressão da doença e reduzir as sequelas, assim, geralmente baseiam-se em regular o sistema imunológico e reduzir a atividade inflamatória [2]. Além das medicações, é recomendado tratamento fisioterapêutico com a finalidade de melhorar a qualidade de vida do paciente. O fisioterapeuta atua tanto na fase aguda (pós-surto), como na fase remissiva. No primeiro caso, os exercícios devem ser mais passivos com pausas de recuperação mais longas. Na fase remissiva, é possível a realização de exercícios mais ativos, porém sem exigir demais do paciente [18].

Capítulo 3

Imagem de Ressonância Magnética

A imagem de ressonância magnética é uma ferramenta poderosa no auxílio ao diagnóstico médico, pois fornece informações importantes para estudar internamente o corpo e suas possíveis disfunções [19]. Este tipo de imagem apresenta uma variedade de intensidades que estão relacionadas com as estruturas anatômicas do corpo. Quanto maior as variações de intensidade, melhor o contraste na imagem, e assim melhor distinção entre regiões de interesse [20]. Este capítulo irá abordar o processo de formação da imagem de ressonância magnética, bem como os princípios físicos envolvidos nesse tipo de exame.

3.1 Formação da Imagem de Ressonância Magnética

As imagens de ressonância magnética são compostas de matrizes preenchidas por voxels (elementos de volume) e representadas em tons de cinza, com valores entre 0 (preto) a 255 (branco) [21]. A Figura 3.1(a) representa um conjunto de cortes da MRI, enquanto a Figura 3.1(b) retrata um voxel de dimensões $p \times p \times d$ e a Figura 3.1(c) representa a estrutura de uma imagem de ressonância magnética como uma matriz 3D formada por um conjunto de voxels.

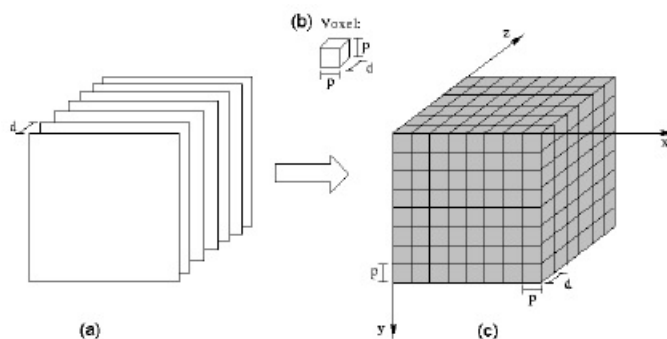


Figura 3.1: Representação da MRI em voxels. (a) Conjunto de cortes formam a imagem de ressonância magnética, (b) ilustração de um voxel e (c) estrutura da MRI como uma matriz 3D de voxels. ([22] citado por [20])

As imagens são normalmente paralelas e uniformemente espaçadas em uma região 3D, podendo ser obtidas nos planos axial, coronal e sagital, conforme observado na Figura 3.2 [20].

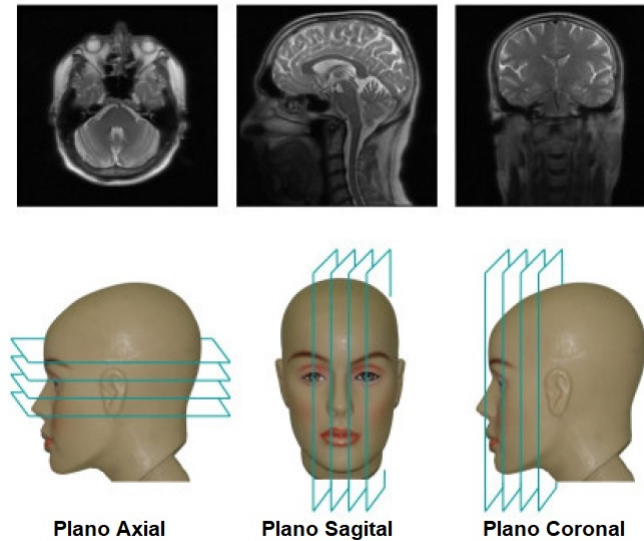


Figura 3.2: Diferentes planos de obtenção de imagens (Adaptada de [20]).

3.1.1 Princípios Físicos da Ressonância Magnética

A Imagem de ressonância magnética é resultado da interação entre campo magnético e os prótons de hidrogênio do tecido humano [2]. O hidrogênio é o elemento preferencial para a aquisição de imagens, pois está presente em grandes quantidades no corpo inteiro além de possuir grande momento magnético [23]. O próton de hidrogênio é uma partícula carregada positivamente, que possui um movimento de rotação ou spin em torno de seu próprio eixo (Figura 3.3.1). Assim, um campo magnético próprio é originado (Figura 3.3.2), podendo ser visto como um pequeno dipolo magnético ou como um ímã (Figura 3.3.3) [24].

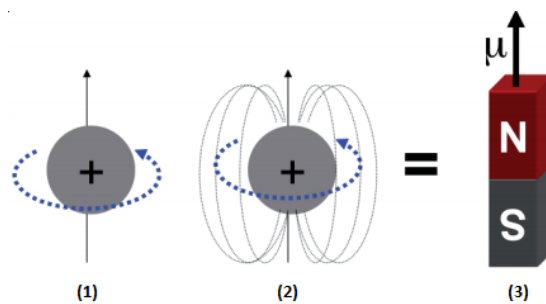


Figura 3.3: Representação de um próton (Adaptada de [24]).

Inicialmente, os prótons estão orientados aleatoriamente no corpo humano, mas quando o paciente é posicionado no interior do equipamento, essas partículas ficam sob influência do campo magnético externo [2], resultando em seu alinhamento (de maneira paralela e antiparalela), com

o campo magnético externo do equipamento [24]. Esse processo está relacionado com diferentes níveis de energia que os prótons podem ocupar: baixa energia (alinhamento paralelo) e maior energia (alinhamento antiparalelo), como observado na Figura 3.4 [25].

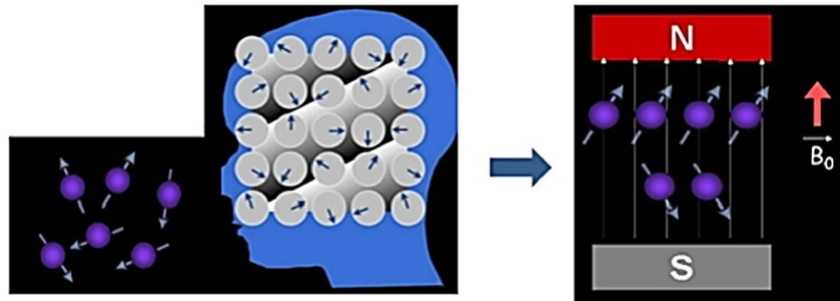


Figura 3.4: Representação das diferentes orientações dos átomos de hidrogênio [2]

Sob ação de um campo magnético externo (B_0), os prótons de hidrogênio irão se movimentar a uma específica frequência, determinada pela Equação de Larmor (Eq. 3.1), que depende da intensidade do campo [23].

$$\gamma = \omega \cdot B_0 \quad (3.1)$$

Um segundo campo magnético de curta duração (pulso) é aplicado, chamado pulso de radiofrequência (RF), perpendicular a B_0 e em fase com a frequência de precessão [24]. Então pulsos de radiofrequência perturbam os alinhamentos dos prótons, e ao encerrar a aplicação do pulso de RF, os prótons retornam ao estado alinhado através de vários processos de relaxamento, emitindo sinais que são convertidos em imagem [2], [24].

3.1.2 Tempos de Relaxação

O tempo de relaxação dos tecidos está relacionado com o processo de realinhamento dos momentos magnéticos com o campo magnético principal, após sofrer alterações pelo pulso de RF [24]. O processo de retorno do vetor de magnetização ao equilíbrio é ilustrado na Figura 3.5, onde o vetor magnetização está representado em azul e os momentos magnéticos individuais estão em amarelo.

O processo de relaxação ocorre com a liberação de energia que os prótons tinham absorvido do pulso de rádio frequência [25]. A relaxação dos spins ocorre através das trocas de energias entre spins ou entre spins e sua vizinhança, promovendo o retorno dos alinhamentos [2],[24]. Os processos de relaxação são caracterizados por duas constantes de tempo: T_1 e T_2 . Ambas constantes de tempo medem a transferência espontânea de energia por um próton excitado, porém diferem na sua disposição final [25].

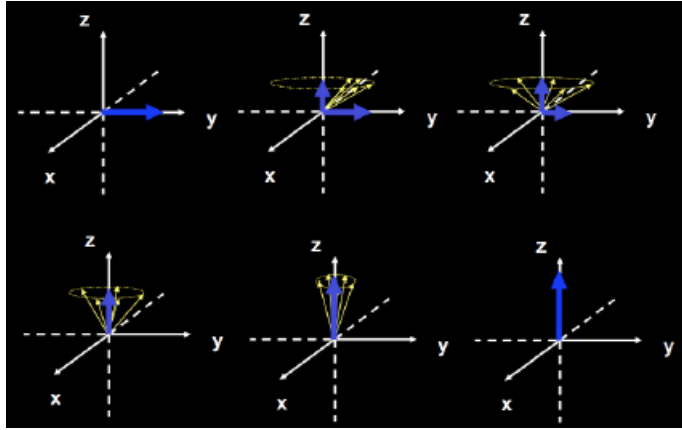


Figura 3.5: Retorno do vetor magnetização ao equilíbrio [24]

O tempo de relaxação longitudinal (T_1) está associado à transferência de energia dos prótons ao retornar sua orientação original e está relacionado com as interações dos spins com a rede [25]. Neste contexto, a constante de tempo T_1 é o tempo necessário para que a componente de magnetização no eixo Z (M_z) retorne a 63% de seu valor original (M_0). O retorno da magnetização ao eixo longitudinal é descrito pela equação 3.2 e ilustrado na Figura 3.6a [25].

$$M_z(\tau) = M_0(1 - e^{-\frac{\tau}{T_1}}) \quad (3.2)$$

Em contraste, o tempo de relaxação transversal (T_2) está relacionado à redução da magnetização no plano transversal e às interações spin-spin [24]. Nesta situação, a constante de tempo T_2 é o tempo requerido para que a componente de magnetização transversal (M_{xy}) reduza para 37% de seu valor inicial ($M_{xy\max}$); esse processo é descrito pela equação 3.3 e ilustrado na Figura 3.6b [25].

$$M_{xy}(\tau) = M_{xy\max}(e^{-\frac{\tau}{T_2}}) \quad (3.3)$$

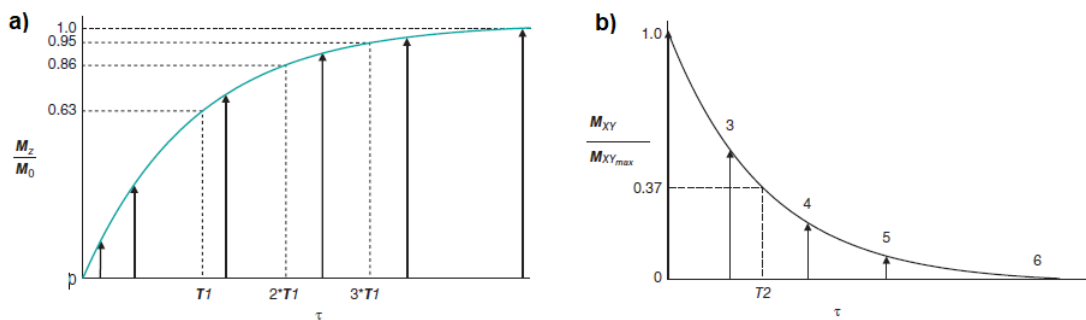


Figura 3.6: Curvas de relaxação T_1 (a) e T_2 (b)[25]

A formação da imagem de ressonância é influenciada pelos tempos de relaxação, pois estes diferentes tempos garantem excelente contraste entre tecidos distintos na MRI, sendo esta uma

grande vantagem de se utilizá-la. [23]. A Tabela 3.1 apresenta tempos de relaxação T1 e T2 para diversos tecidos a 1,5 T.

Tabela 3.1: Tempos de relaxação T1 e T2 aproximados para diversos tecidos do corpo humano a 1,5 T [24].

Tecido	T1(ms)	T2(ms)
Substância branca	790	90
Substância cinzenta	920	100
Líquido céfalo-raquidiano (líquor)	4000	2000
Sangue (arterial)	1200	50
Parênquima hepático	490	40
Miocárdio	870	60
Músculo	870	50
Lipídios (gordura)	260	80

3.1.3 Sequências das imagens

Diferentes tipos de MRI são criados, ao variar a sequência de aplicação dos pulsos de RF, variando as intensidades dos sinais em cada tipo de imagem [23]. As sequências de MRI são obtidas ao controlar os parâmetros Tempo de Repetição (TR) e Tempo de Eco (TE) [23]. O Tempo de Repetição é o intervalo de tempo entre sucessivos pulsos de RF aplicados, enquanto Tempo de Eco é o tempo entre a aplicação do pulso de RF e o máximo sinal de eco, representados na Figura 3.7 [25].

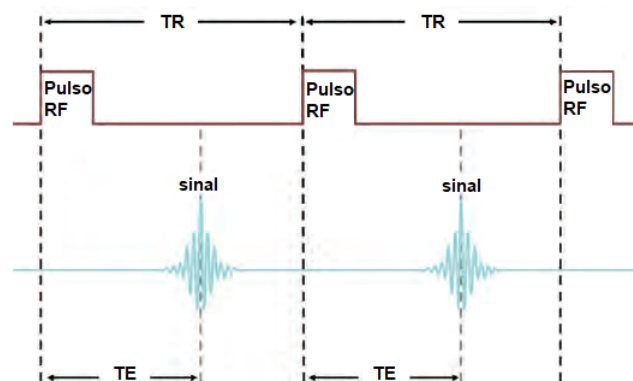


Figura 3.7: Simple sequência de pulsos (Adaptada de [23])

Alguns exemplos de sequências de MRI são as imagens T1, T2 e FLAIR (*Fluid Attenuated Inversion Recovery*). As imagens ponderadas em T1 são obtidas ao utilizar TR e TE curtos, enquanto as imagens ponderadas em T2 são resultados da aplicação de TR e TE longos [23]. Tipicamente, a sequência ponderada em T1 é adquirida ao aplicar TE inferior a 30 ms e TR inferior a 800 ms. Em oposição, a sequência ponderada em T2 é obtida utilizando TE superior a 80 ms e TR superior a 2.000 ms [26].

Conforme visualizado na Figura 3.8, as imagens ponderadas em T1 realça tecidos contendo gordura, enquanto tecidos com grande quantidade de água aparecem como hipointensidades. Já as imagens ponderadas em T2 realçam tanto gorduras como tecidos contendo majoritariamente água [23]. As imagens FLAIR são similares às T2, mas são obtidas utilizando TE e TR muito longos (TE superior a 80 ms e TR superior a 3.000 ms), dessa forma anormalidades são realçadas, enquanto o CSF é atenuado e visualizado em preto [27], [26]. Assim, essa sequência é bastante utilizada para analisar patologias [27]. Em adição, as sequências FLAIR suprimem os efeitos imagiológicos da vazão sanguínea. Dessa maneira, a detecção de lesões na substância branca e cinzenta são melhoradas [28].

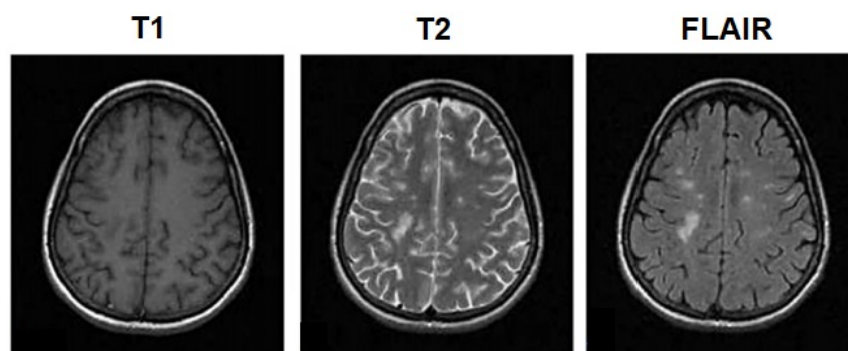


Figura 3.8: Sequências diferentes de MRI: T1, T2 e FLAIR (Adaptada de [2]).

3.2 Características Imagiológicas das Lesões de Esclerose Múltipla

As imagens de ressonância magnética apresentam as lesões escleróticas com intensidades diferentes em cada sequência, conforme indicado na Tabela 3.2, onde é apresentada uma comparação da visualização dos diferentes tecidos em cada sequência de MRI. Nas sequências T2 e FLAIR, as lesões aparecem como hiperintensidades, enquanto na sequência T1 aparecem como hipointensidades. A sequência FLAIR é a que apresenta melhor diferenciação entre as lesões de EM e os tecidos saudáveis [23].

Tabela 3.2: Cores e tons dos tecidos nas imagens T1, T2 e FLAIR [27]

Tecido / Imagem	T1	T2	FLAIR
CSF	Preto	Branco	Preto
Substância branca	Cinza claro	Cinza escuro	Cinza escuro
Substância cinzenta	Cinza escuro	Cinza claro	Cinza claro
Desmielinização	Preto	Branco	Branco

Normalmente, as lesões estão localizadas nas regiões periventricular, justacortical e infratentorial, conforme observado na Figura 3.9. A Figura 3.9.A. mostra lesões periventriculares, a Figura 3.9.B. indica uma das lesões em destaque, na Figura 3.9.C. há presença de lesões justacorticais e na 3.9.D. lesões na região infratentorial [13].

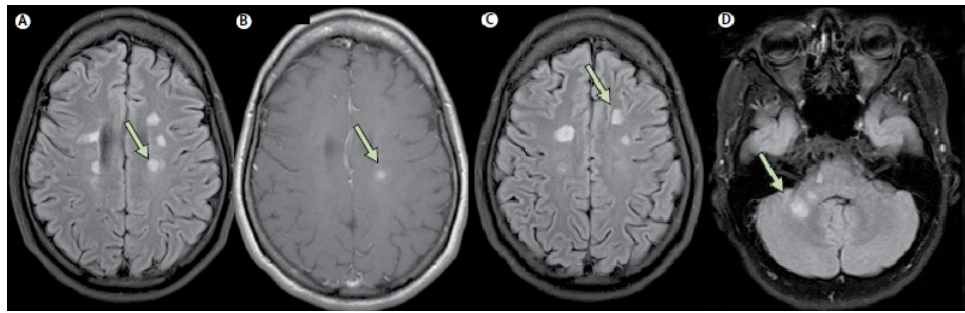


Figura 3.9: Lesões aparentes em MRI: as setas indicam, em A, lesões periventriculares, em B, uma lesão periventricular em destaque, em C, lesões justacorticais e em D, lesões na região infratentorial [13].

Capítulo 4

Pré-processamento e Análise de Imagens

Sistemas computadorizados de auxílio diagnóstico estão em desenvolvimento com a finalidade de aprimorar a interpretação de imagens médicas, promover uma avaliação prognóstica e dar suporte à decisão médica [29]. Primeiramente, é fundamental que ocorra um pré-processamento das imagens. Esse processo consiste em aplicar uma série de modificações nas imagens antes de realizar a segmentação da região de interesse [30]. Neste capítulo serão abordadas as técnicas de pré-processamento de imagem ressonância magnética, seguida da segmentação de lesões de E.M. através de métodos por *Deep Learning*.

4.1 Imagem Médica

Imagens médicas podem ser definidas como uma forma de visualização interna do corpo humano. As imagens podem ser obtidas em variadas sequências, como por exemplo tomografias, imagem de ressonância magnética e ultrassonografia [21]. Essas imagens são analisadas por especialistas em aplicações clínicas como para diagnosticar doenças [31]. Neste trabalho de mestrado, utilizou-se especificamente as imagens obtidas por ressonância magnética. As imagens médicas são normalmente imagens 2D, mas possibilitam a reconstrução e visualização em 3D e ainda podem assumir dimensões adicionais, como é o caso das sequências de MRI. As imagens podem ter formato DICOM (*Digital Imaging and Communications in Medicine*) ou formato NifTI (Neuroimaging Informations Technology Initiative), originalmente desenvolvida para análise de imagens cerebrais [31].

4.2 Pré-Processamento

O pré-processamento das imagens é a etapa responsável pelo melhoramento da qualidade da imagem e adequá-la de acordo com as especificações do método de análise ou segmentação a ser utilizado. O seu objetivo é minimizar o efeito de artefatos [30] e padronizar todas as imagens. No

caso da segmentação de imagens cerebrais em MRI são necessários alguns passos como: remoção de ruídos, extração de tecido não cerebral, registro entre as diferentes sequências e correção da não homogeneidade do campo magnético [2].

4.2.1 Remoção de Ruídos

Durante a aquisição das MRIs, podem ocorrer processos que resultem em ruídos na imagem, logo é necessário aplicar filtros para amenizar esse efeito [30]. Os filtros utilizados em imagens médicas não devem provocar degradação da imagem, seja por borrões ou artefatos [32]. Esses filtros devem ter as seguintes características:

- Preservar as bordas dos objetos e detalhes das estruturas;
- Remover os ruídos em regiões homogêneas eficientemente;
- Aperfeiçoar definições morfológicas.

Baseado na literatura, os filtros de difusão anisotrópica melhoram a qualidade das imagens e satisfazem esses critérios [32].

4.2.2 Correção de Não Homogeneidade

Um mesmo tecido pode ter variações de intensidades, em decorrência da não homogeneidade dos campos magnéticos durante a aquisição da imagem [2]. A correção da não homogeneidade reduz essas variações de intensidades e evita baixo desempenho na segmentação das lesões [33], sendo um passo importante para o pré-processamento da imagem. A Figura 4.1 apresenta uma imagem antes (Figura 4.1.a) e depois de ser corrigida (Figura 4.1.c), além de mostrar a estimativa do sinal de não-homogeneidade (Figura 4.1.b), chamado de *bias field* em inglês.

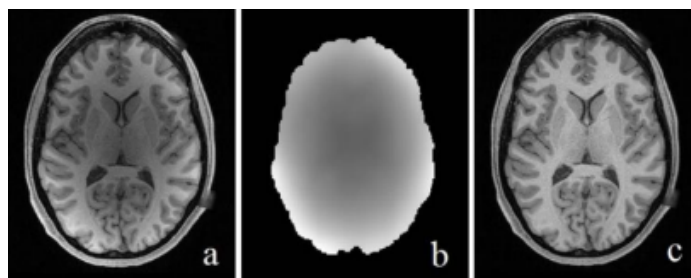


Figura 4.1: Correção de intensidade não-uniforme: imagem original em (a), o sinal de não homogeneidade (b) e imagem corrigida (c) (Adaptada de [34])

Os métodos de correção são classificados em duas categorias: Prospectivas e Retrospectivas [35]. Abordagens Prospectivas reduzem os sinais de não homogeneidade causados pelo processo de aquisição da imagem imperfeito, ajustando os equipamentos de aquisição [36]. Abordagens Retrospectivas corrigem os sinais de não homogeneidade relacionados às propriedades específicas do objeto em análise na imagem adquirida, tais métodos de correção podem ser baseados em análise de histogramas, filtragem e *surface fitting* [35].

4.2.3 Normalização de Intensidade

As MRIs obtidas por diferentes protocolos de aquisição, equipamentos ou até mesmo obtidas em diferentes ocasiões não possuem intervalo de intensidade similares [37]. Dessa forma, a falta de uma escala de intensidades padronizadas resulta na dificuldade em generalizar as características dos tecidos na presença de uma patologia [37]. A normalização de intensidades garante que todos os voxels estejam no mesmo intervalo de intensidades. Assim, há uma comparação equivalente entre imagens obtidas em diferentes condições [38].

A normalização de intensidade pode ser realizada de diferentes maneiras. Uma simples forma de realizar essa atividade é através da subtração do valor médio das intensidades e divisão pelo desvio padrão [39]. A normalização também pode ser realizada através da estimativa de densidade por *kernel*, fornecendo uma versão suave de histograma para encontrar o valor de pico associado à substância branca do cérebro, que é utilizado como denominador para normalizar as intensidades [40].

4.2.4 Extração de estruturas não cerebrais

A extração de estruturas não cerebrais é realizada, pois as lesões de interesse se encontram no encéfalo e podem apresentar intensidades similares a estas estruturas [38]. Nesse processo, o crânio (parte óssea), gordura, pele e todas estruturas não cerebrais são extraídas da MRI [41]. Dessa maneira, diminui a ocorrência de resultados falsos positivos. Observa-se na Figura 4.2, um corte de MRI antes e depois da segmentação do encéfalo.

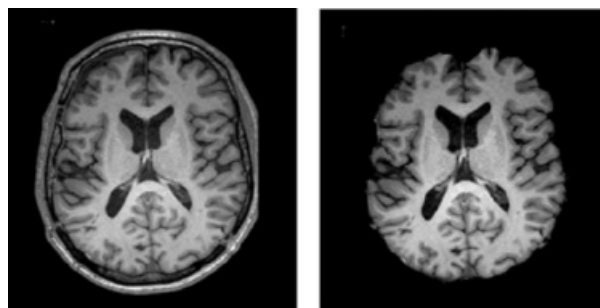


Figura 4.2: Extração de estruturas não cerebrais [41].

Segundo a literatura, as técnicas de extração do encéfalo são baseadas em diferentes categorias, como a de morfologia matemática e superfície deformável. Normalmente, os métodos de morfologia matemática utilizam detecção de bordas, *thresholding* e operações de erosão e dilatação para separar o crânio da região do cérebro [42]. As técnicas de superfície deformável, como o *Brain Extraction Tool* (BET), utilizam um modelo de superfície que se ajusta à região do cérebro até encontrar a solução ótima [43].

4.2.5 Registro das Imagens

As várias sequências de MRI apresentam informações complementares sobre o mesmo paciente. Com a finalidade de garantir a correspondência das estruturas cerebrais em todas as sequências, é necessário realizar o alinhamento das imagens [30]. O registro da MRI é o processo que realiza o alinhamento de duas ou mais sequências de imagens, garantindo que as mesmas localizem-se em um único sistema de coordenadas espaciais [44]. O registro também pode ser definido como um mapeamento espacial, com o alinhamento de uma imagem inicial para o sistema de coordenadas definido pela imagem de referência ([45] citado por [30]). A Figura 4.3 apresenta um diagrama do processo de registro das imagens, onde uma imagem móvel é alinhada à imagem de referência, se transformando na imagem registrada.

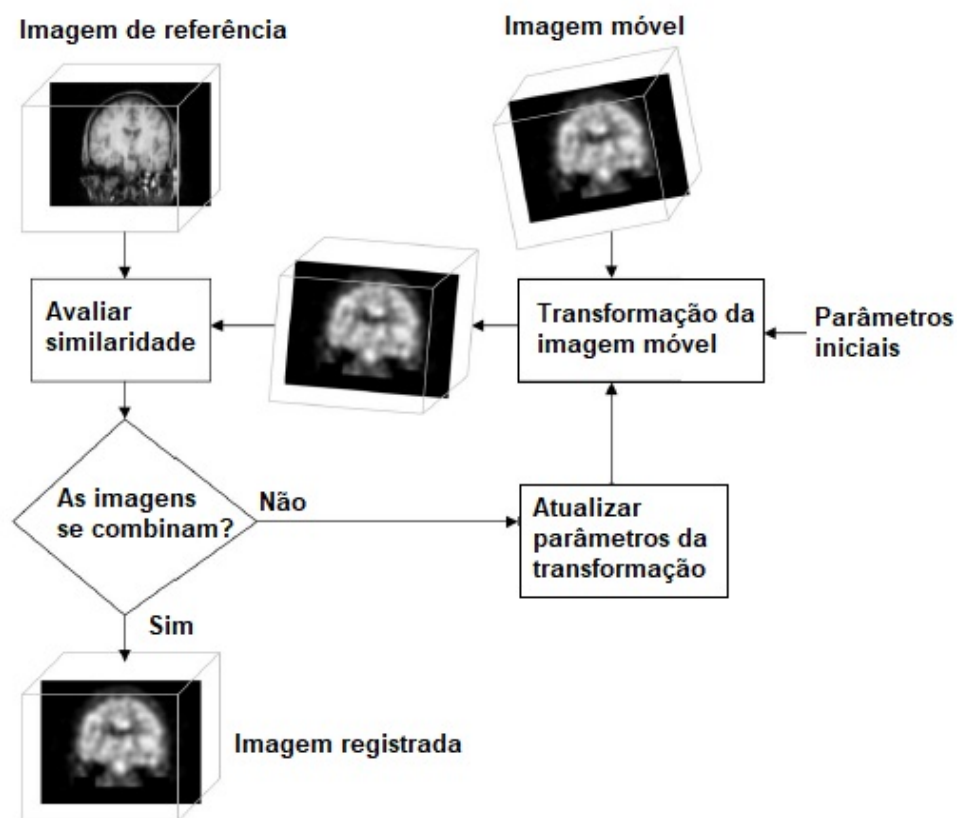


Figura 4.3: Diagrama do processo de registro (Adaptada de [38]).

O processo de registro inicialmente estima os parâmetros de transformação necessários para mapear as diferentes sequências de MRI. Em seguida, a transformação estimada é aplicada à imagem móvel para alocar no espaço de coordenadas da imagem de referência [46], esse processo é realizado repetidamente até alcançar a solução ótima. Os algoritmos de registro podem ser classificados com base em modelos de transformação em duas categorias: registro linear e não linear [47]. No registro linear, a mesma transformação é aplicada em todos os voxels da imagem móvel. No registro não linear diferentes transformações podem ser aplicadas em diferentes voxels, logo

há maior grau de elasticidade, podendo modelar uma deformação local [47], [46].

4.3 Segmentação das Lesões Encefálicas

A segmentação de uma imagem tem o objetivo de subdividi-la em regiões que se diferenciam entre si. Normalmente o conjunto de pixels de interesse é denominado de objeto, enquanto o conjunto dos outros pixels que não tem informação relevante para o objetivo é considerado "fundo"[2].

O padrão ouro de segmentação de lesões de E.M. é a segmentação manual, onde vários especialistas analisam as MRIs e identificam as regiões com intensidades diferentes que caracterizam o comprometimento encefálico decorrente do processo de desmielinização [30]. A segmentação manual é muito importante, pois produz as imagens *ground truth* utilizadas no desenvolvimento de métodos automáticos de segmentação de imagens [10]. Entretanto este método é trabalhoso e demorado, além de estar sujeito a variações inter e intra especialistas, visto que é um processo subjetivo do conhecimento do especialista [48].

Em decorrência dessas situações, processos de auxílio à segmentação manual (segmentação semiautomática) e métodos automatizados estão a ser cada vez mais aplicados para segmentações de regiões de interesse em imagens médicas [30]. A segmentação semiautomática é realizada a partir de algumas interações entre usuários e algoritmos automatizados para fornecer uma coerente segmentação [49]. A maioria dos métodos semiautomáticos utilizam técnicas de crescimento de regiões ou conexidade fuzzy [38]. Já os métodos automatizados consistem em processos realizados completamente por meio de computadores. Esta última abordagem garante uma segmentação mais rápida e diminui as variabilidades causada pela interação humana [30].

Capítulo 5

Machine Learning

Técnicas de *Machine Learning*, em especial, *Deep Learning*, têm sido utilizadas em estudos para segmentação de imagens biomédicas [50]. Dentre as técnicas de *Deep Learning*, temos uma grande aplicabilidade das *Convolutional Neural Networks* (CNN) para segmentação de estruturas de interesse. Assim, neste capítulo, conceitos fundamentais sobre CNNs serão apresentados, bem como uma revisão geral de métodos de segmentação específicos para lesões de E.M. em imagem de ressonância magnética.

5.1 Introdução

Machine Learning envolve estudos e desenvolvimento de métodos, onde os computadores adquirem habilidade de resolver problemas a partir de um aprendizado por experiências (treinamento) [50]. Essa técnica é formada por conjuntos de algoritmos para construir modelos matemáticos, onde sistemas podem extrair informações a partir de um banco de dados, e então identificar padrões e fazer inferências [51].

Os métodos de *Machine Learning* podem ser classificados como supervisionados e não supervisionados. O supervisionamento ocorre a partir de um treinamento com a utilização de uma base de dados com imagens antecipadamente segmentadas, normalmente segmentação manual [52]. Assim, o programa aprende as características do objeto em interesse, sendo capaz de diferenciá-las dos tecidos saudáveis [30]. No caso deste trabalho, o objeto de interesse é a lesão esclerótica. Por outro lado, o aprendizado não-supervisionado não utiliza bases de dados de aprendizado com rótulos [52]. A maioria dos métodos não-supervisionados utiliza técnicas de agrupamento (em inglês *clustering*) para discriminar automaticamente os voxels em categorias [38].

Existem vários subgrupos de *Machine Learning*, em especial, o *Deep learning* tem evoluído bastante nas aplicações em imagens médicas [50]. Estes métodos de aprendizado analisam os dados de forma similar ao pensamento humano [53]. *Deep learning* engloba uma técnica específica de *Neural Networks*, que utiliza um algoritmo inspirado no processo de aprendizado que ocorre

em um cérebro humano [53]. Assim, tecnologias com base em CNN tem grande potencial de aplicação nas áreas relacionadas a diagnóstico médico e análise de dados [50].

5.2 Convolutional Neural Networks

Convolutional Neural Networks são projetadas para trabalhar com dados em forma de matriz, como por exemplo imagens, onde há grande dependência espacial [53]. As CNNs são capazes de extrair características simples das imagens e criar representações complexas, através da utilização de filtros extratores de características em camadas ocultas [54]. Além disso, as CNNs mantêm as informações espaciais da imagem. Nesse sentido, as CNNs são adequadas para tarefas de reconhecimento, como classificação, localização, detecção e segmentação de objetos [55]. Estruturalmente, as CNNs têm várias camadas de convolução e ativação, normalmente intercaladas com camadas de *pooling*, conforme ilustrado na Figura 5.1 e serão descritas detalhadamente a seguir.

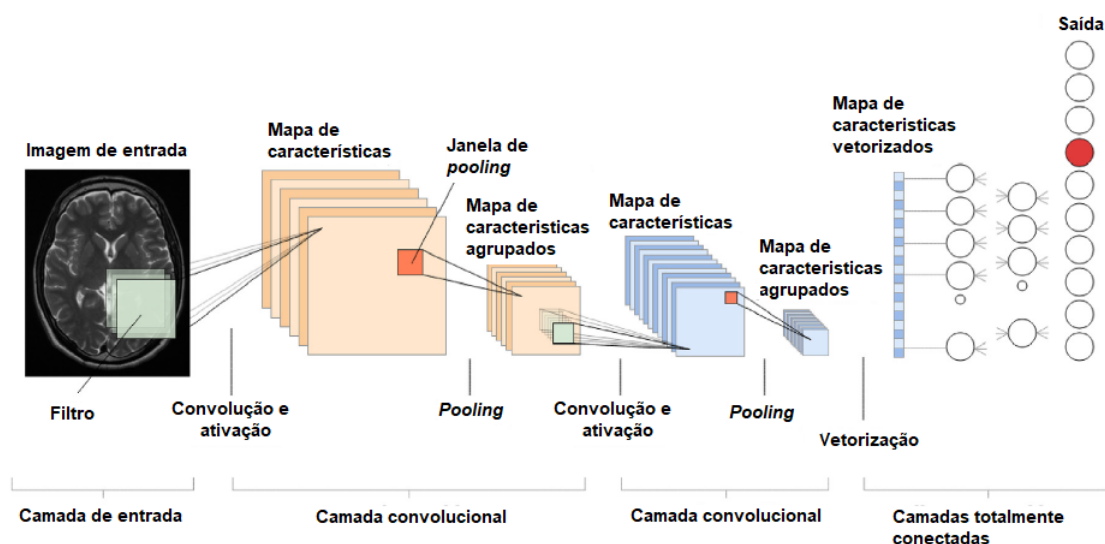


Figura 5.1: Blocos de uma CNN (Adaptada de [50]).

5.2.1 Camadas de Convolução

As camadas de convolução têm o objetivo de extrair as características da imagem de entrada, através de uma combinação de operações lineares e não-lineares, como operações de convolução e aplicação da função de ativação [56]. A convolução é uma operação linear utilizada para obter padrões, onde um filtro ou *kernel* é aplicado à imagem. Cada pixel é multiplicado por seu respectivo peso do filtro, então esses valores são adicionados e se tornam o valor final no pixel resultante [53]. Esse procedimento é realizado várias vezes até percorrer toda a imagem, resultando em um mapa de características, em inglês *feature map* [55].

O número de filtros utilizados por camada determina a quantidade de mapas de características formada na mesma. Cada filtro identifica um padrão espacial em uma pequena região da imagem, assim, uma grande quantidade de filtros é necessária para extrair diversas formas [53]. Portanto, é muito importante definir, propriamente, o número de filtros utilizados e suas dimensões [56]. Geralmente, os filtros possuem tamanhos muito menores do que a imagem analisada, tendo dimensões de 3x3 ou 5x5 pixels [53].

A distância entre as posições de dois filtros consecutivos é chamada de *stride*. Normalmente, utiliza-se como uma unidade, mas o *stride* pode assumir valores maiores, caso seja desejável diminuir a amostragem nos mapas de características [56]. Além disso, *strides* maiores ajudam o algoritmo a generalizar nos casos em que a resolução espacial é desnecessariamente grande [53].

Nesse processo, a convolução acarreta na redução das dimensões das imagens (Figura 5.2A). Uma maneira de evitar tal redução é através do *zero padding*, onde linhas e colunas de zeros são adicionadas nas bordas da imagem (Figura 5.2B) [53]. Muitas arquiteturas de Redes Neurais Convolucionais utilizam várias camadas de convolução e realizam *zero padding* para manter o tamanho das imagens, caso contrário, as dimensões dos mapas de características se tornariam cada vez menores após a operação de convolução [56].

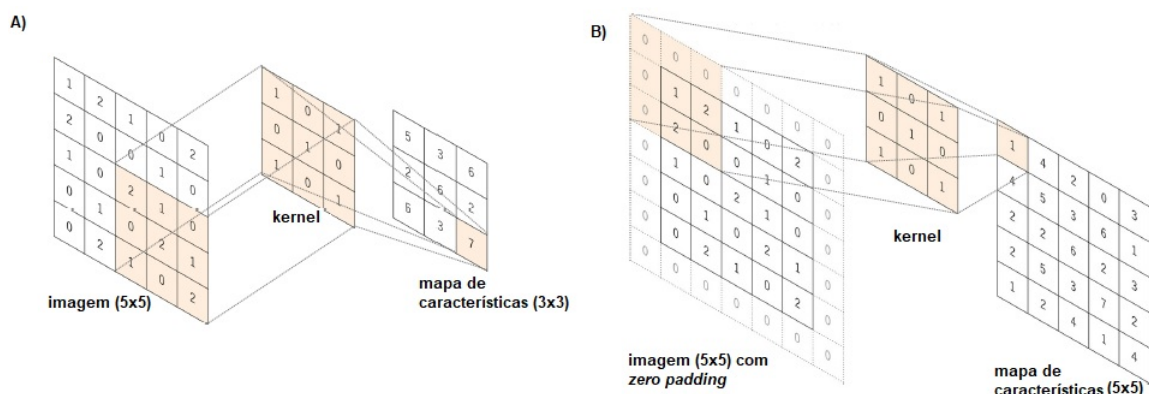


Figura 5.2: Representação da operação de convolução (Adaptada de [56]).

Em seguida, os resultados das operações de convolução passam por uma função de ativação não-linear, como a *Rectified Linear Unit* (ReLU), a qual computa a função: $f(x) = \max(0, x)$, onde os pixels negativos são convertidos a zero [55]. O uso da função ReLU melhora a velocidade de treinamento, relativamente a outras funções de ativação mais convencionais, como a sigmoid [10]. A Figura 5.3 ilustra um mapa de características (4x4) antes (5.3a) e após (5.3b) a aplicação da função de ativação ReLU.

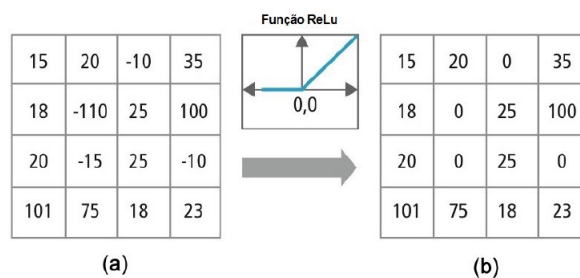


Figura 5.3: Aplicação da função ReLU [57]).

5.2.2 Camada de Subamostragem

As CNNs utilizam camadas de subamostragem, em inglês *pooling layer*, com o objetivo de diminuir o tamanho espacial das matrizes resultantes da convolução [58]. Portanto, há redução da quantidade de parâmetros aprendíveis, melhora no custo computacional e melhor generalização do algoritmo [55], [10]. Normalmente, utiliza-se filtro com dimensão (2x2) e *stride* 2, o que faz reduzir as dimensões das imagem por um fator de 2, mas a profundidade do mapa de características se mantém a mesma [56].

A operação de *pooling* computa um número específico para representar a região selecionada por um filtro [50]. A operação mais utilizada é a *Max Pooling*, onde a maior intensidade de uma região é selecionada para representá-la, conforme observado na Figura 5.4 [53]. Em adição, existe a operação de *Average Pooling* que seleciona a média dos valores da região. A escolha entre *Max* ou *Average Pooling* depende da aplicação, o primeiro faz a extração de pontos extremos, enquanto o último promove um resultado mais suave [50].

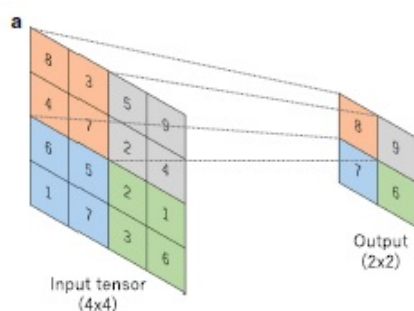


Figura 5.4: Exemplo de operação de MaxPooling [56].

5.2.3 Camadas Totalmente Conectadas

Algumas arquiteturas de CNNs utilizam Camadas Totalmente Conectadas (FCL do inglês *Fully Connected Layers*) com a finalidade de aumentar o poder computacional no fim da rede [53]. Nessas redes, o mapa de características resultante da última camada de convolução ou de agrupamento é convertido em uma tensor unidimensional e segue para uma ou mais FCL, onde cada entrada está conectada a cada saída por um peso [56].

5.2.4 Última camada

Na última camada da rede neural utiliza-se geralmente uma função de ativação diferente das outras. Em caso de multi-classificação, é aplicada a função *softmax*, que normaliza os valores de saída em probabilidades para cada classe de 0 a 1 [56]. Outra função de ativação bastante utilizada é a *sigmoid* para classificações binárias [59].

5.2.5 Treino das CNNs

O treinamento das redes neurais é o processo de encontrar os melhores pesos ou *kernels* para o aprendizado do modelo, minimalizando o erro entre a predição e o *ground truth* [56]. O treino de um algoritmo é feito a partir de um banco de dados divididos em três grupos: treino, validação e teste, conforme ilustrado na Figura 5.5. O primeiro é responsável por treinar a rede neural, o segundo é utilizado para avaliar o modelo durante o treino e o terceiro é usado no fim do processo para obter uma avaliação do modelo final [56].

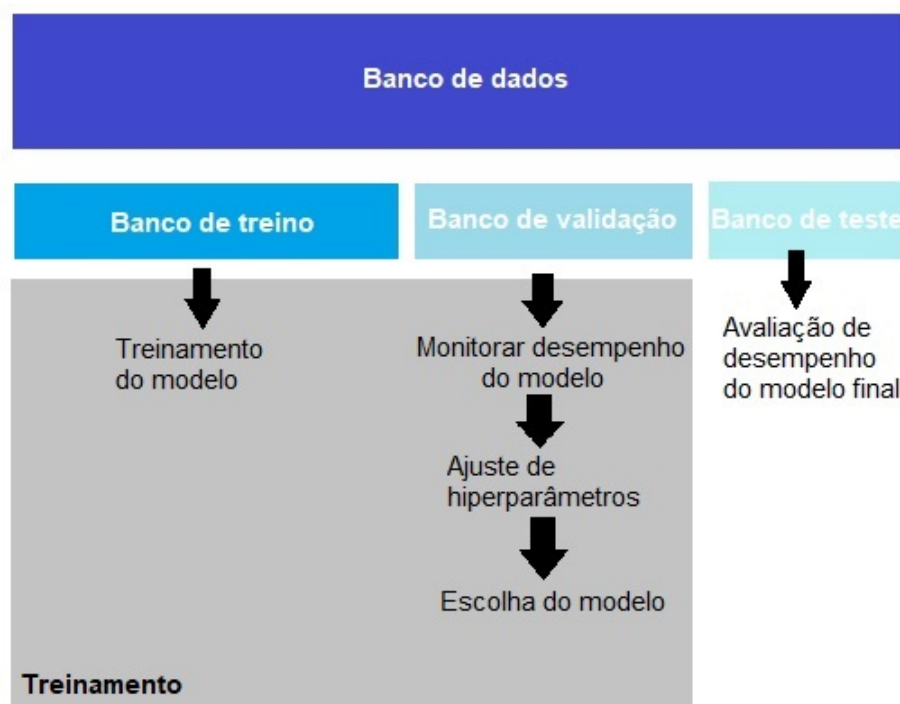


Figura 5.5: Divisão do banco de dados (Adaptada de [56]).

Geralmente, o treinamento é feito a partir de um algoritmo de otimização, chamado Gradiente Descendente. O otimizador utiliza uma função de perda para avaliar a performance do modelo ao empregar determinados filtros e pesos [50]. O cálculo do erro é feito por propagação direta e tais parâmetros aprendíveis são atualizados via retro-propagação [56]. O processo de treinamento da rede está ilustrado na Figura 5.6.

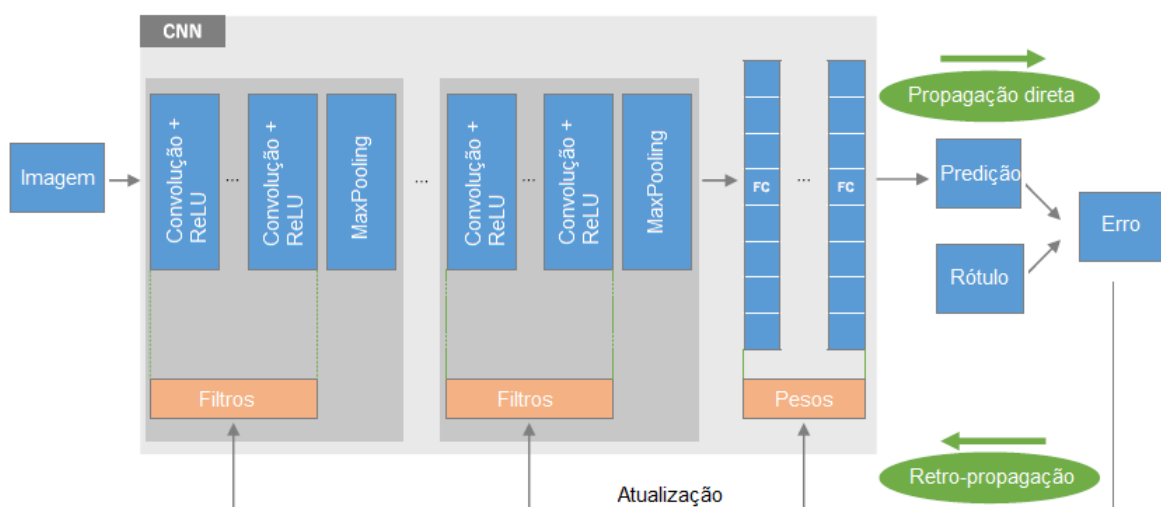


Figura 5.6: Visão geral do processo de treinamento de uma CNN (Adaptada de [56])

O Gradiente Descendente atualiza os parâmetros aprendíveis com a finalidade de minimizar o erro. Cada parâmetro é atualizado em direção ao ponto de mínimo da função de perda, de acordo com o tamanho do passo, conhecido como taxa de aprendizagem ou em inglês *learning rate* [56]. Além da importância da taxa de aprendizagem, a escolha da função de perda também apresenta um papel fundamental ao projetar um algoritmo de segmentação de imagens, pois influencia no processo de aprendizagem do modelo [60]. Tal função tem a finalidade de medir a diferença entre o *ground truth* e a previsão obtida pelo treinamento do modelo, quanto maior o valor da perda menos preciso será o modelo [57]. A atualização dos parâmetros pode ser formulada na Eq 5.1, onde ω refere-se aos parâmetros aprendíveis, α atribui-se a taxa de aprendizado e L faz referência a função de perda.

$$\omega = \omega - \alpha \cdot \frac{\partial L}{\partial \omega} \quad (5.1)$$

Entretanto, aplicar o Gradiente Descendente usando todo o conjunto de treinamento pode ser computacionalmente custoso. [61]. Alternativamente, é possível utilizar o Gradiente Descendente Estocástico (SGD do inglês *Stochastics Gradient Descent*), que consiste na atualização dos parâmetros em cada época, utilizando, aleatoriamente, apenas pequenas amostras dos dados (*mini-batches*) a cada iteração. Assim, soluciona-se problemas relacionados a memória computacional e tem-se a direção de descida estabelecida conforme tal amostra [62].

A escolha apropriada da taxa de aprendizagem é importante para garantir um treinamento eficiente da CNN [63]. Se o valor for muito baixo ou muito alto pode resultar em uma convergência pobre ou muito lenta, conforme ilustrado na Figura 5.7.

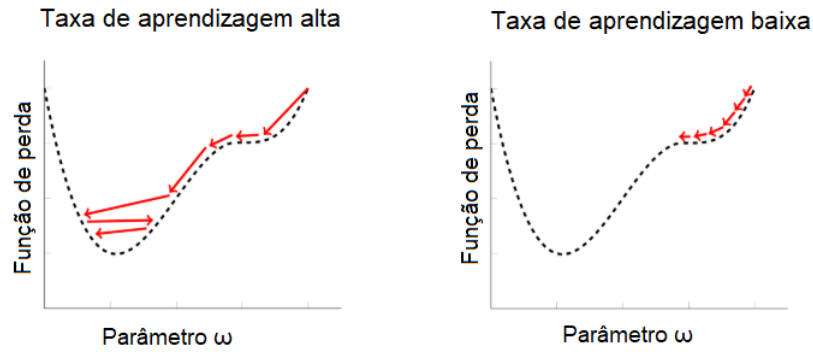


Figura 5.7: Influência da taxa de aprendizagem com valores diferentes (Adaptada de [63])

É de fundamental importância que o programa seja generalizado e capaz de funcionar com um banco de dados novo. Essa habilidade é estimada durante o treinamento, ao utilizar uma parte separada do banco de dados para validação. Ao avaliar a precisão e as limitações do programa, obtém-se um *feedback* para melhoramento do modelo [50]. Nessa situação, a acurácia é um aspecto bastante utilizado para analisar o quão próximo o resultado da segmentação automática está do *ground truth* [30]. Assim, cada voxel das imagens pode ser classificado como:

- **Verdadeiro Positivo (VP)** — Ocorre quando um voxel identificado como lesão pelo programa realmente corresponde a lesão no GT;
- **Verdadeiro Negativo (VN)** — Ocorre quando um voxel classificado como "não-lesão" pelo programa realmente corresponde a uma região não lesionada no GT;
- **Falso Positivo (FP)** — Ocorre quando uma região de tecido saudável é identificada como lesão pelo algoritmo;
- **Falso Negativo (FN)** — Ocorre quando uma região lesionada é classificada como tecido saudável pelo algoritmo.

Muitos trabalhos de segmentação de imagem biomédica utilizam o Coeficiente de Similaridade Dice (DSC) para avaliar o resultado da segmentação. Este coeficiente quantifica a similaridade entre duas imagens binárias através da sobreposição espacial [64]. O Coeficiente de Similaridade Dice é dado pela equação 5.2, variando de 0 a 1. O resultado da segmentação é considerado muito bom, quando o coeficiente for próximo de um.

$$DSC = \frac{2.TP}{FP + 2.TP + FN} \quad (5.2)$$

Um problema de falta de generalização do modelo é chamado *Overfitting*. Isto acontece quando o modelo treinado realiza excelentes inferências no banco de dados de treino, porém não é capaz de realizar boas predições em dados de teste nunca visualizados [53]. O monitoramento do erro e da acurácia é uma forma de identificar possível *overfitting* durante o treinamento. Essa situação pode ser exemplificada quando o modelo obtém resultados da segmentação do banco de validação inferiores aos do banco de treino [56].

Uma das soluções para reduzir tal problema é aumentar a quantidade de amostras do banco de dados de treinamento, dessa maneira o modelo consegue generalizar melhor [51]. Um forma de aumentar o número de amostras de treino é através de *data augmentation*, tal operação provoca transformações aleatórias, como variação de intensidade, acréscimo de ruído, rotação, recorte, etc. Assim, o modelo não vê sempre as mesmas imagens a cada interação [56]. Outra possibilidade de reduzir o *overfitting* é através de *batch normalization*, cuja função é normalizar os dados de todas as camadas da rede, e não apenas os dados iniciais. Tal procedimento resulta na adição de ruídos durante o treinamento, agindo como um regularizador [65].

5.3 Arquiteturas

Há várias arquiteturas de CNN utilizadas em problemas de visão computacional. A depender do objetivo principal do algoritmo, os componentes mencionados na seção anterior podem ser combinados, intercalados e até substituídos por outras operações para criar uma nova arquitetura.

5.3.1 U-Net

A U-Net é um tipo de rede completamente formada por camadas de convolução (*Fully Convolutional Networks*). Tal arquitetura é muito popular e foi desenvolvida por Olaf Ronneberger et al. para segmentação de imagens biomédicas [66] (Figura 5.8). A rede tem um formato em U e contém dois caminhos: contração e expansão [10]. O primeiro apresenta uma arquitetura similar a de uma típica CNN e é usado para extrair as informações das imagens por meio de operações de convoluções e *MaxPooling* [66]. Tais operações reduzem o tamanho das imagens, deixando-as com uma baixa resolução, chamado em inglês de *down sampling* [10]. Esse processo faz com que o modelo entenda melhor o que está presente na imagem, mas perde informação sobre sua localização, o que ocasiona da necessidade do caminho de expansão [67]. A seguinte etapa tem o objetivo de identificar localizações precisas da região de interesse e aumentar a resolução das imagens, através de operações de convoluções e *UpSampling* [66]. Além disso, *skip connections* são utilizadas com a finalidade de compartilhar informações espaciais do caminho de contração para o de expansão [10].

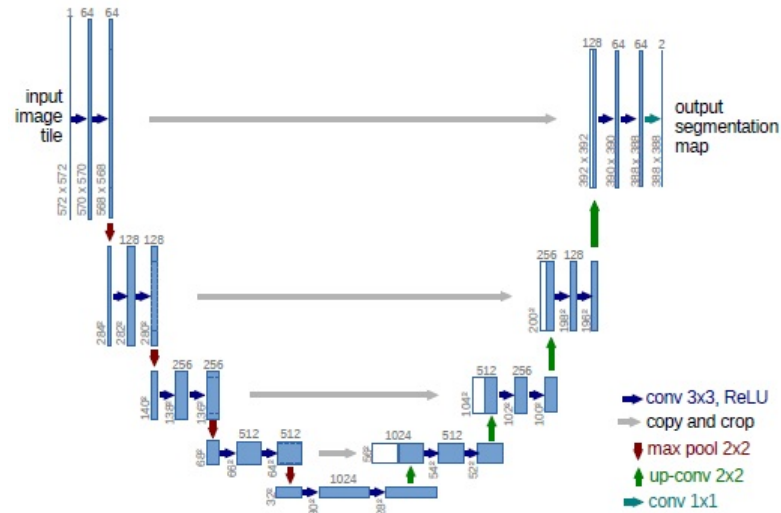


Figura 5.8: Arquitetura U-Net. Cada bloco azul representa multi-canais de mapa de características, o número de canais é indicado no topo de cada bloco. Os blocos brancos representam cópias dos mapas de características. As setas de diferentes cores representam diferentes operações [66].

5.3.2 CNN em Cascata

O modelo de CNN em cascata (Figura 5.9) utiliza a associação de duas redes idênticas para obter a segmentação das imagens médicas. A primeira rede é treinada para revelar possíveis candidatos a lesões, enquanto a segunda tem o objetivo de reduzir erros nas classificações realizadas pela primeira [68]. Como os números de voxels que representa lesões no tecido é muito menor do que o número total de voxels da imagem, há o selecionamento balanceado de voxels para o treinamento das redes.

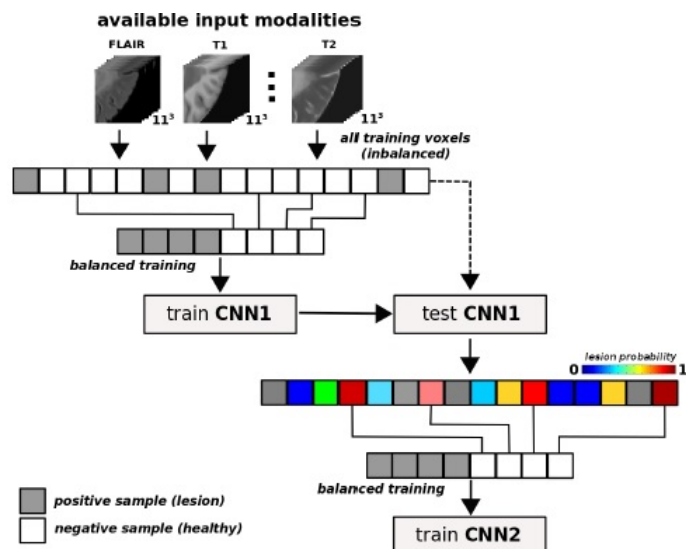


Figura 5.9: CNN em cascata [68]

5.3.3 CNN com Concatenação

A arquitetura de CNN com concatenação (Figura 5.10) aplica várias camadas de convolução e ativação para cada sequência de MRI em paralelo. Em seguida, os resultados dessas operações são concatenados e passam para outro conjunto de filtros convolucionais que realiza a predição da segmentação [8].

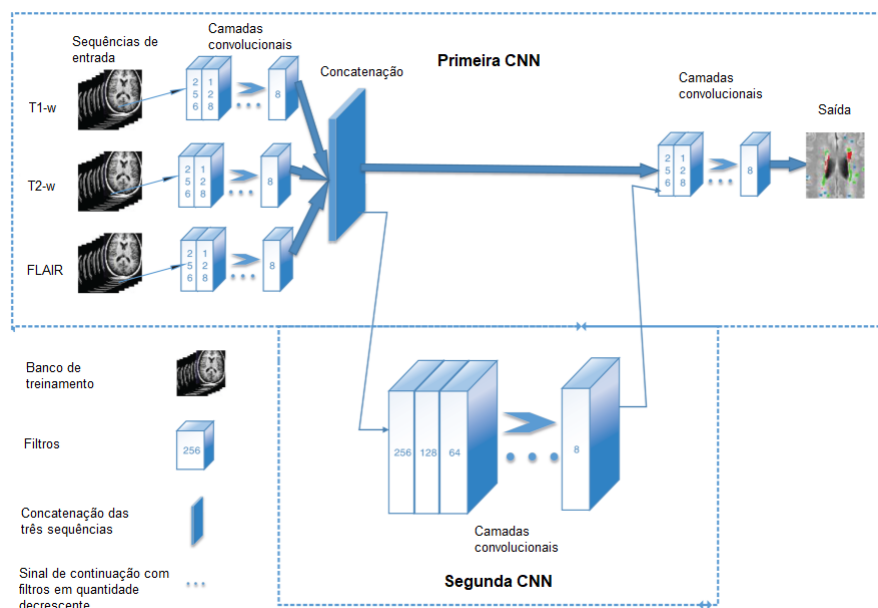


Figura 5.10: Arquitetura de CNNs com concatenação. Sequências de MRI (T1-w, T2-w e FLAIR) são entradas do modelo que seguem para a primeira parte da CNN, contendo camadas de convolução. Em seguida, os mapas de características são concatenados e seguem para segunda parte da CNN. Na arquitetura proposta, camadas convolucionais são organizadas com números decrescentes de filtros (Adaptada de [69]).

5.4 Revisão sobre Trabalhos de Segmentação de Lesões de Esclerose Múltipla

Numerosas abordagens automáticas têm sido desenvolvidas para acelerar a segmentação de imagens médicas [70]. Na última década, muitos trabalhos de segmentação de lesões cerebrais utilizaram técnicas de *Deep Learning* [71]. Nessa seção, uma revisão geral e temporal foi realizada sobre os métodos de segmentação de lesões escleróticas em MRI.

5.4.1 HeMIS: Hetero-Modal Image Segmentation

Havaei et al. [70] propuseram um modelo de segmentação de lesões cerebrais chamado de Hetero-Modal (Figura 5.11), no qual cada sequência de MRI foi processada inicialmente por um conjunto de convoluções, independentemente. Em seguida, os mapas de características, contendo valores de média e variância, de cada sequência foram computados e concatenados. Por fim, os

mapas de características seguiram para um bloco de camadas convolucionais e assim obteve-se a saída da rede. Esse algoritmo apresenta uma grande vantagem em relação a métodos anteriores: a habilidade de aprendizagem sem necessitar de todas as sequências de MRI. Como tal, qualquer subconjunto combinatório de sequências disponíveis pode ser fornecido como entrada.

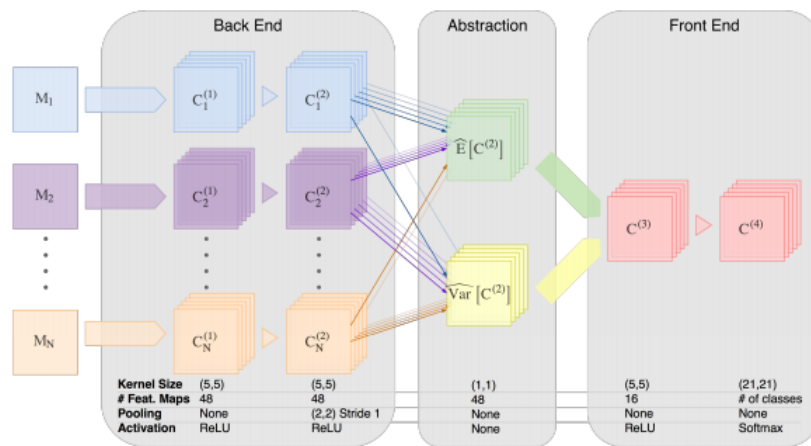


Figura 5.11: Arquitetura Hetero-modal para segmentação. M_k sequências de MRI usadas como entrada passam independentemente para camadas convolucionais no bloco *Back End*. Mapas de características estatísticas são computados no bloco *Abstraction*, após concatenação seguem para o bloco *Front End*, onde ocorre a classificação dos pixels [70].

Esse trabalho utilizou imagens de dois bancos de dados públicos: BRATS (*Brain Tumor Segmentation Challenge*) [72] e MICCAI (*Medical Image Computing and Computer Assisted Intervention*) [73]. O algoritmo foi treinado em duas etapas. Na primeira, foi utilizado um banco de dados balanceados, que permitiu aprender recursos agnósticos à distribuição de classes. Na segunda fase, realizou-se *fine-tuning* (retreinamento das últimas camadas de um modelo para execução de outra tarefa), treinando a última camada com dados distribuídos próximos ao *ground truth*. Dessa forma, o modelo foi capaz de aprender as características e manter uma boa classificação dos pixels. O método proposto utilizou dois bancos de imagens para teste, um contendo lesões de E.M. e outro com tumores cerebrais. O resultado obtido a partir da segmentação de lesões escleróticas apresentou coeficiente Dice médio de 0,388 ao utilizar as sequências FLAIR e T1. Enquanto, um coeficiente Dice médio de 0,487 foi obtido ao utilizar as sequências FLAIR, T1, T1c e T2.

5.4.2 Longitudinal Multiple Sclerosis Lesion Segmentation Using Multi-view Convolutional Neural Networks

Birenbaum and Greenspan [74] apresentaram o primeiro método de *Deep Learning* que utilizou dados longitudinais para realizar segmentação. Os autores propuseram a utilização de CNNs baseadas em multi-vistas longitudinais para segmentar as lesões de E.M.. Esse trabalho foi dividido

em dois estágios. O primeiro utilizou as intensidades da MRI e a substância branca para determinar possíveis lesões escleróticas. O segundo estágio utilizou CNN baseadas em multi-vistas para inferir as lesões e reduzir falsos positivos.

A seleção de candidatos a lesões desqualificou a maioria dos voxels das imagens, assim há redução do custo computacional na fase seguinte. A extração de possíveis lesões foi realizada levando em consideração que as mesmas aparecem como hiperintensidades nas imagens FLAIR e que tendem a estar localizadas na substância branca ou na borda entre a substância branca e cinzenta. Logo após, os candidatos a lesões seguiram como entrada da Rede Neural Convolutiva. A arquitetura proposta utilizou 4 modelos de CNNs, conforme ilustrado na Figura 5.12.

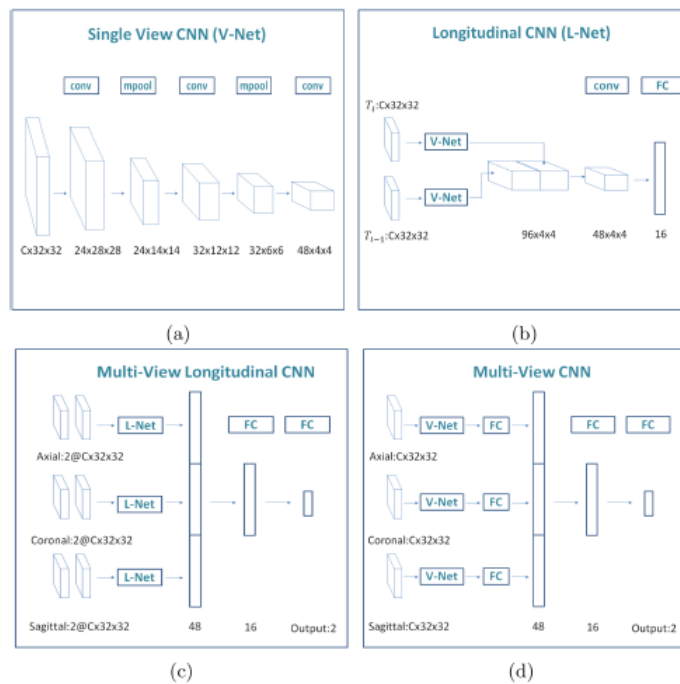


Figura 5.12: Arquiteturas CNNs utilizadas. (a) Visão única CNN (V-Net), (b) Longitudinal CNN (L-Net), (c) Multi-visão longitudinal CNN e (d) Multi-visão CNN [74].

Esse trabalho utilizou as sequências de MRI T1 MPRAGE, T2, FLAIR e PD. Essas imagens foram pré-processadas com as seguintes etapas: registro das sequências, extração do encéfalo, correção de não-homogeneidade e normalização das intensidades. O treinamento do algoritmo foi realizado com o otimizador AdaDelta e foi concluído quando os valores da acurácia se tornaram saturados durante a validação, concluindo 500 épocas. O conjunto de dados *International Symposium on Biomedical Imaging (ISBI) 2015* [75] foi utilizado para teste, e obtiveram-se resultados melhores do que as metodologias de pesquisa convencionais, obtendo um coeficiente Dice médio com valor 0,627.

5.4.3 Fast Lesion EXtraction using CONvolutional Neural Networks - FLEXCONN

Roy *et al.* [8] propuseram uma arquitetura de CNN com concatenação (Figura 5.13). *Patches* 2D foram extraídos das sequências de MRI e seguiram para a rede. Inicialmente, aplicou-se

em paralelo, conjuntos de bancos convolucionais, seguidos da função de ativação ReLU. Posteriormente, os resultados dessas operações foram concatenados e passaram para outro conjunto de filtros convolucionais que realizou a predição da segmentação. Tais conjuntos foram formados por 5 bancos de filtros de diferentes tamanhos e organizados em ordem decrescente de quantidade de filtros, conforme observado na Tabela 5.1.

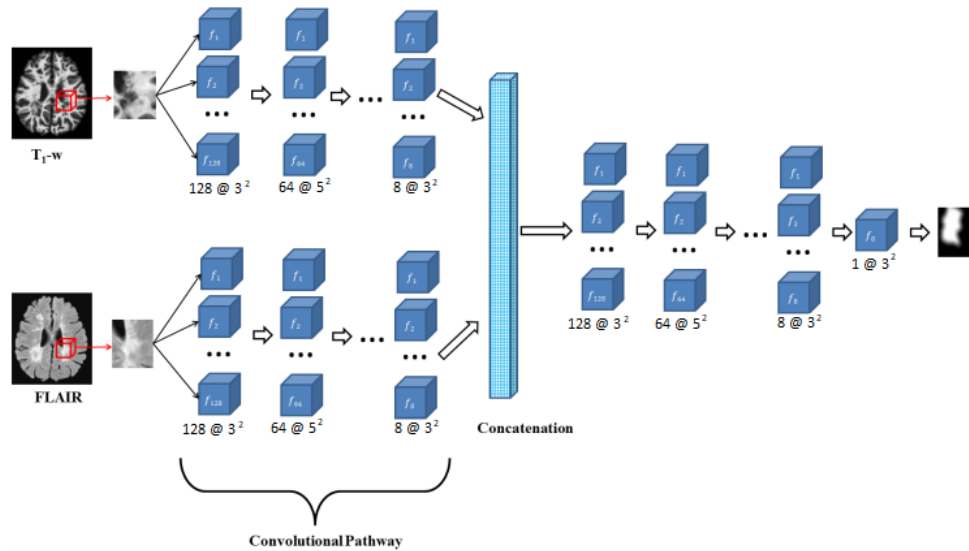


Figura 5.13: Arquitetura CNN proposta. *Patches* 2D de imagens foram usadas como entrada e blocos com filtros convolucionais foram aplicados em paralelo. Cada bloco (f_j) incluiu operações de convolução e ReLU. As saídas foram concatenadas e passadas por outro caminho de bancos convolucionais para prever a segmentação. O número e os tamanhos dos filtros foram mostrados como $x @ y^2$, indicando que o banco de filtros conteve x filtros, cada um com tamanho (y,y) pixels (Adaptada de [8]).

Tabela 5.1: Características dos bancos de filtros utilizados no trabalho [8]

Banco de filtros convolucionais	Número de filtros	Dimensões dos filtros
1º	128	3x3
2º	64	5x5
3º	32	3x3
4º	16	5x5
5º	8	3x3

Neste trabalho, o algoritmo aprimorou as segmentações de lesões ao utilizar bancos de filtros em ordem decrescente de quantidade de filtros. Segundo os autores, filtros com tamanho de 3x3 pixels e 5x5 pixels foram utilizados por garantir melhores resultados na segmentação. Esses resultados foram explicados, pois filtros com tais dimensões tendem a reconhecer melhor as bordas das lesões que não são muito bem definidas.

O banco de dados utilizados foi fornecido pelo ISBI [75], contendo as sequências T1 e FLAIR. Durante a etapa de pré-processamento, as imagens foram registradas rigidamente para que cada voxel tivesse dimensão de $1mm^3$. Em seguida, houve a extração de estruturas não-cerebrais e

correção da não homogeneidade. O otimizador Adam foi utilizado durante a implementação do algoritmo, com taxa de aprendizagem igual a 0,0001 em 20 épocas. Além disso, *batch size* de tamanho 128 foi utilizado para treino e 64 para teste.

Ao utilizar uma camada convolucional para prever os resultados, ao invés de utilizar uma camada densa, foram observadas algumas vantagens, como a redução de *overfitting*, resultados com menos falsos positivos e redução no tempo de processamento. Os resultados desse trabalho foram satisfatórios e superiores a resultados obtidos por outros algoritmos tradicionais de *Machine Learning*, resultando em DSC médio igual a 0,564.

5.4.4 One-shot Domain Adaptation In Multiple Sclerosis Lesion Segmentation Using Convolutional Neural Networks

No trabalho realizado por Valverde *et al.* [39], foi utilizado um modelo com CNN-3D em cascata, denominado NicMSLesions. A primeira rede foi treinada para revelar possíveis candidatos a lesões, enquanto a segunda teve o objetivo de reduzir erros nas classificações realizadas pela primeira. Este método apresentou bom desempenho mesmo utilizando um pequeno banco de dados, o que é considerado uma vantagem, tendo em vista as limitações de obtenção das imagens.

Esse método consistiu em treinar a rede utilizando amostras 3D das imagens, pois são dados que contêm informação contextual. Portanto, foi benéfico para segmentar as lesões na matéria branca do cérebro. A arquitetura das CNN está ilustrada na Figura 5.14. Conforme observado, as CNNs recebem 3D *patches* com dimensões (11x11x11) das sequências de MRI como entrada da rede. Logo após, foram utilizados dois blocos convolucionais e de *pooling*. O primeiro bloco conteve duas camadas convolucionais com 32 filtros de tamanhos (3x3x3), seguidas por uma camada *MaxPooling* de dimensão (2x2x2). O segundo bloco apresentou estrutura semelhante, porém com 64 filtros em cada camada convolucional. Em seguida, os dados passaram por três camadas totalmente conectadas, com 256, 128 e 64 neurônios respectivamente, finalizando com uma camada *softmax* para estimar a segmentação.

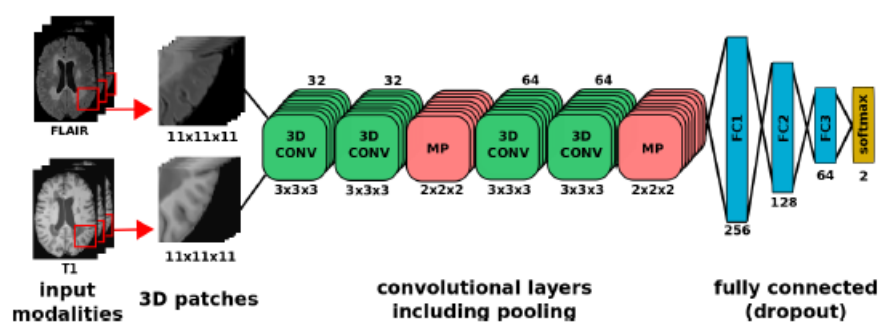


Figura 5.14: Arquitetura da CNN do modelo base [39].

Durante o treinamento de uma rede para segmentar lesões escleróticas, é importante levar em consideração que o número de voxels total de lesões é muito pequeno em relação ao número total de voxels do encéfalo, 20 ml de volume de lesão equivale por volta de 1,5% do volume

total do cérebro. Assim, a quantidade de amostras positivas é muito reduzida, podendo interferir no desempenho do algoritmo. Com o objetivo de contornar isso, o treinamento foi realizado balanceadamente, com números iguais de *patches* positivos (voxels de lesões) e negativos (tecido saudável).

Os autores também analisaram quão adaptável foi a utilização dos modelos-CNNs em cascata em imagens adquiridas de outras formas e nunca vistas antes. Além disso, o resultado do modelo base treinado com o banco de dados do MICCAI 2008 [73] e 2016 [76] foi comparado com o resultado do modelo utilizando *transfer learning*, retrainando as camadas *fully connected*, ao adicionar imagens alvo (imagens com características semelhantes as imagens do banco de teste) no re-treinamento. A Figura 5.15 ilustra esse processo.

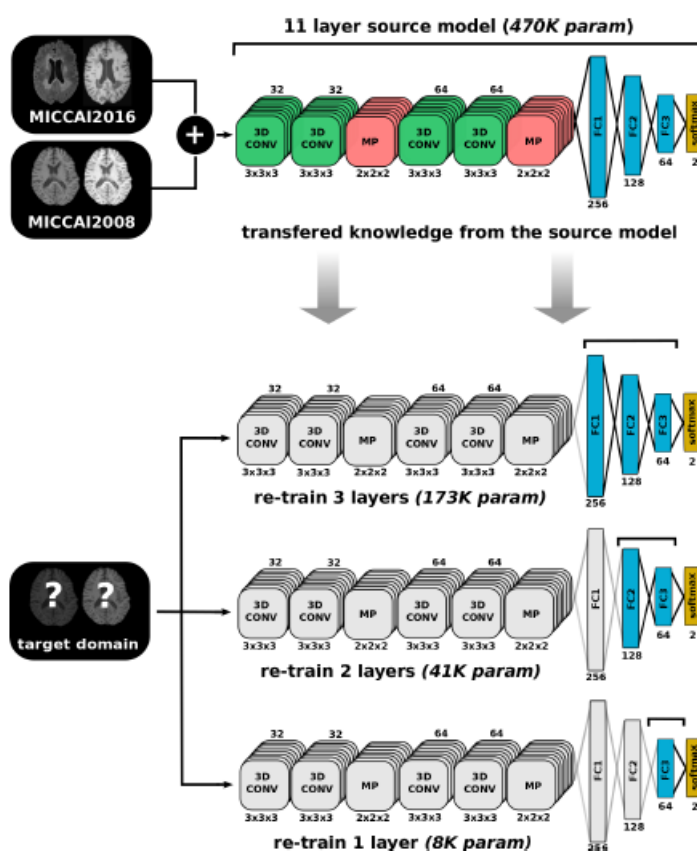


Figura 5.15: Estrutura de adaptação de domínio de intensidade supervisionada. A partir do modelo de base, transfere-se o aprendizado para um domínio de imagem de destino nunca vista. A adaptação do domínio é efetuada por meio de 3 configurações: retrainando a última camada FC, duas FC ou todas as camadas FC usando imagens e rótulos do domínio de intensidade alvo. Em todas as configurações, as camadas que não são re-treinadas estão representadas em cinza [39].

O treino foi realizado ao utilizar imagens de ressonância magnética dos tipos FLAIR e T1 pré-processadas, com as seguintes operações: extração do encéfalo, normalização das intensidades, registro no espaço T1 e re-amostragem isotrópica ($1mm^3$). O erro da primeira CNN foi calculado e em seguida, ocorreu o treinamento da segunda rede, utilizando grupos de voxels que foram

classificados erroneamente pelo primeiro modelo e voxels das lesões verdadeiras. Durante o treinamento, a quantidade de camadas que podem ser retreinadas varia de acordo com a quantidade de imagens alvo disponíveis no novo banco de dados.

Os autores avaliaram o efeito das seguintes características no treinamento do modelo:

- Variação do tamanho das lesões (0,5-18 ml) em uma única imagem de re-treino.
- Re-treinamento de 1, 2 ou 3 camadas totalmente conectadas
- Quantidade de imagens utilizadas no re-treino

Os resultados mostraram que o re-treino das FCL melhorou o desempenho da segmentação, pois o modelo base (sem adaptação de domínio) não possuiu tanta adaptabilidade para banco de dados diferentes. Além disso, os modelos obtidos utilizando *transfer learning* têm performance similar ao modelo completamente treinado, especialmente quando 2 ou 3 camadas são re-treinadas. Ainda, o algoritmo treinado apenas com uma imagem alvo, contendo pelo menos 30ml de lesões (em imagem isotrópica com voxels de dimensão 1mm^3) é eficiente para segmentar as lesões. O trabalho obteve DSC com valor médio de 0,58 e desvio padrão de 0,16.

5.4.5 A Light Weighted Deep Learning Framework for Multiple Sclerosis Lesion Segme

No modelo aplicado por Ghosal *et al.* [77], os autores apresentaram uma versão modificada da U-net, intitulada de *Light Weighted U-net*. Tal arquitetura, ilustrada na Figura 5.16, apresentou diminuição da complexidade da rede, por reduzir a quantidade de camadas convolucionais e quantidade de filtros. A rede consistiu de oito camadas convolucionais, 4 camadas de *MaxPooling*, 4 camadas *Upsampling* e 1 camada final com a função de ativação sigmoid. Além disso, *batch normalization* foi utilizado para reduzir *overfitting*. O diferencial dessa arquitetura estava na concatenação de camadas convolucionais alternadas, ao invés de concatenar todas camadas convolucionais como na tradicional rede U-net.

Inicialmente, as sequências de MRI concatenadas na terceira dimensão (256x256x5) serviram de entrada à rede. Em seguida, operações de convolução foram aplicadas às imagens e o método de *padding* foi usado para manter suas dimensões. Os filtros utilizados apresentaram tamanhos (3x3x5) e *strides* de (1x1) pixels. Inicialmente, foram utilizados 32 filtros, mas esse número aumentou exponencialmente a cada camada de convolução. Posteriormente, aplicou-se a técnica de *batch normalization* e consequentemente cada camada da rede aprendeu um pouco mais. Em seguida, aplicou-se a função de ativação ReLU e a operação de *MaxPooling* de tamanho (2x2), diminuindo as dimensões das imagens por um fator de 2. Em paralelo, as concatenações alternadas transferiram o resultado da convolução no caminho de contração para o caminho de expansão. O processo de expansão foi composto de camadas de Convolução e ReLU, assim como no caminho de contração, porém seguido da técnica de *UpSampling* para aumentar a resolução das imagens.

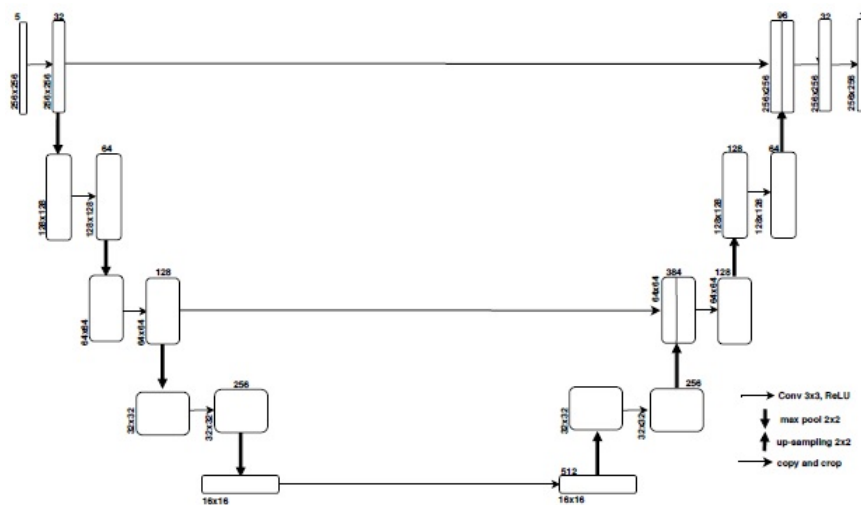


Figura 5.16: Arquitetura U-net adaptada [77].

O trabalho empregou o banco de dados do MICCAI 2016 [76] com as 5 sequências: DP, T1, T2, FLAIR e GADO (MRI com agente de contraste *Gadolinium*). Essas imagens foram pré-processadas com as seguintes técnicas: redução de ruídos, correção de inhomogeneidade das intensidades e extração de estruturas não-cerebrais. Ainda, foi realizada a concatenação do primeiro *slice* de cada sequência da MRI e assim em diante, até completar toda a imagem. Dessa maneira, o número de canais da imagem de entrada foi igual ao número de sequências utilizadas, nesse caso, cinco. Em seguida, as imagens foram redimensionadas para (256x256x5). Posteriormente, houve remoção dos *slices* nulos (sem informação), pois são dados repetitivos e inaptos. Assim, o tempo de treinamento da rede foi diminuído e o tempo de predição foi melhorado. Por fim, as imagens foram normalizadas em um intervalo de 0 a 1. Esse processo foi realizado ao subtrair, da intensidade do pixel, o valor médio de intensidade dessa imagem e dividi-lo pelo desvio padrão acrescido de um valor muito pequeno para evitar a divisão por zero.

O treinamento foi inicializado pelo algoritmo Xavier, e realizado em mini blocos (em inglês *batches*) de tamanho 32, com otimizador Adam e taxa de aprendizagem igual a 0,001, realizou 60 épocas e utilizou a *Binary Cross Entropy* como função de perda. Esse algoritmo foi treinado em um intervalo de tempo relativamente pequeno, graças a diminuição de complexidade da rede. Este método apresentou excelentes resultados de segmentação, com obtenção do Coeficiente de Similaridade Dice médio com valor 0,76.

5.4.6 A Dense U-Net Architecture for Multiple Sclerosis Lesion Segmentation

Kumar *et al.* [71] implementaram uma arquitetura U-net modificada, ao adicionar blocos de camadas densas na rede, conforme ilustrado na Figura 5.17. A utilização do bloco de camadas densas foi justificado por evitar o problema da Dissipação do Gradiente, melhorar a propagação das características extraídas e reduzir *overfitting*. Além dos procedimentos comuns de

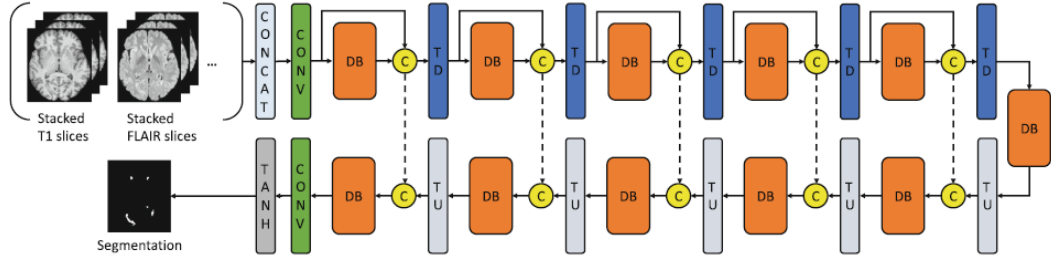


Figura 5.18: Arquitetura Tiramisu proposta. A rede consiste em caminho de contração e expansão. A última camada de Tanh traz os valores de saída para a faixa $[-1, 1]$ [78].

5 sujeitos para treino e 14 para teste. Diferentemente de outros métodos, os autores organizaram cada imagem de treinamento em blocos de três *slices* empilhados, onde ao menos o *slice* central possuiu um voxel de lesão. Dessa maneira, as informações globais e locais foram mantidas. O algoritmo foi treinado com o otimizador Adam, com a taxa de aprendizagem de 0,0002 e momento de 0,5. Assim, a taxa de aprendizagem com o valor inicial foi utilizada por 100 épocas e depois reduzida linearmente a 0 com o valor de época determinado. Além disso, o trabalho analisou o efeito de três funções de perdas diferentes em cada treinamento: L_2 loss definida na equação 5.3, $Dice$ loss (5.4) e $Focal$ loss (5.5). Nessas funções X indica os dados de entrada, Y representa as imagens *ground truth* e F representa a predição da CNN. Na função 5.5, α é uma contante utilizada para equilibrar a importância das amostras positivas e negativas dos dados e γ tem o objetivo de aumentar a foco em prováveis classificações erradas.

$$L_2 = |F(X) - Y|_2 \quad (5.3)$$

$$L_{dice} = 1 - DSC(F(X), Y) \quad (5.4)$$

$$L_{focal} = -\alpha \cdot Y(1 - F(X))^\gamma \cdot \log(F(X)) - (1 - \alpha)(1 - Y) \cdot F(X)^\gamma \cdot \log(1 - F(X)) \quad (5.5)$$

Os autores testaram os modelos em um banco de dados simulado privado e obtiveram excelentes valores de coeficiente Dice, conforme observado na Tabela 5.2.

Tabela 5.2: Coeficientes Dice obtidos a partir da segmentação de imagens em banco de dados simulados, utilizando diferentes funções de perda para cada treinamento [78]

Modelos	Coeficiente Dice
Modelo treinado com L_2 loss	0,861
Modelo treinado com $Dice$ loss	0,847
Modelo treinado com $Focal$ loss	0,865

5.5 Pós-processamento da imagem segmentada

Por fim, independente da arquitetura utilizada, o pós-processamento é a etapa que ocorre após a segmentação das imagens com o objetivo de refinar os resultados obtidos [79]. Tal processo tem o intuito de remover regiões de falso positivo, aprimorando a segmentação da região de interesse [33]. Alguns critérios para remoção de falsos positivos levam em consideração o tamanho das lesões, onde é determinado um limiar para a área da lesão [68]. Assim, considera-se lesões de E.M., as regiões que apresentam no mínimo 3mm de diâmetro [80]. Na Figura 5.19, é possível observar a diferença da imagem segmentada por um algoritmo sem alterações, contendo lesões e falsos positivos (indicadas pelas setas em vermelho) e essa mesma imagem após realizar o pós-processamento.

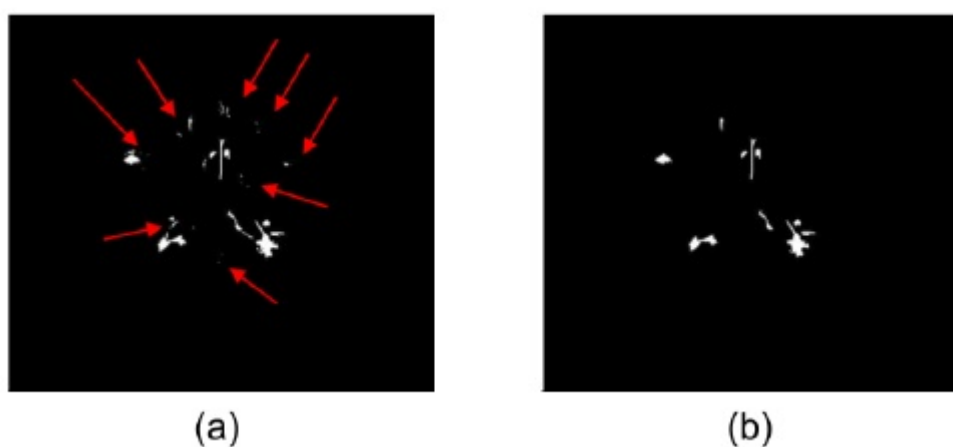


Figura 5.19: Redução de falso positivos: (a) imagem segmentada por um algoritmo; (b) imagem segmentada após a remoção de falso positivos ([81] citado por [33]).

5.6 Considerações Finais

Neste capítulo, métodos computacionais de *Machine Learning* para segmentação de lesões escleróticas foram apresentados. Um resumo dos modelos descritos, sequências de imagem de ressonância magnética utilizadas e coeficiente Dice médio, aproximado em 2 casas decimais, são indicados na Tabela 5.3. Os resultados divulgados foram obtidos utilizando banco de dados diferentes em cada trabalho, dificultando a comparação com exatidão entre os modelos. Entretanto, observou-se que as sequências de MRI do tipo FLAIR e T1 foram bastante utilizadas, garantindo uma segmentação coerente.

Tabela 5.3: Coeficientes Dice médio obtidos a partir dos diferentes trabalhos de segmentação de lesões escleróticas.

Modelos	Sequências de MRI	DSC
Hetero-modal	T1 e FLAIR	0,39
Hetero-modal	T1, T1c, T2 e FLAIR	0,49
Multi-view CNN	T1, T2, FLAIR e PD	0,63
FLEXCONN	T1 e FLAIR	0,56
NicMSlesions (domínio adaptativo)	T1 e FLAIR	0,58
<i>Light Weighted UNET</i>	T1, T2, FLAIR, PD e GADO	0,76
UNET com camadas densas (sem <i>data agumentation</i>)	T1, T1c, T2 e FLAIR	0,80
UNET com camadas densas (com <i>data agumentation</i>)	T1, T1c, T2 e FLAIR	0,87
Tiramisu treinado com L_2 loss	T1 e FLAIR	0,86
Tiramisu treinado com <i>Dice loss</i>	T1 e FLAIR	0,85
Tiramisu treinado com <i>Focal loss</i>	T1 e FLAIR	0,86

Ao longo do tempo, os trabalhos de segmentação de lesões escleróticas no encéfalo foram aprimorados. Entretanto, os trabalhos encontrados na literatura apresentaram algumas limitações, como por exemplo a necessidade de utilizar um baco de dados com muitas imagens e a diferença de domínio entre imagens clínicas e imagens utilizadas durante a pesquisa. Ainda é necessário desenvolver projetos de investigação com a finalidade de garantir a segmentação de imagens clínicas de forma relativamente simplificada e com bom desempenho ao analisar os resultados. Dessa forma, é importante realizar vários estudos para melhorar a performance dos modelos, para que seja possível a utilização de métodos automatizados em clínicas médicas, pois é uma tecnologia com grande potencial e importância no auxílio ao diagnóstico e avaliação de terapias.

Capítulo 6

Metodologia

O presente capítulo descreve a metodologia desenvolvida nesta dissertação para segmentação automática das lesões de E.M. em MRI do banco de dados da Universidade de Liubliana [82]. Inicialmente, as imagens de ressonância magnética passaram por várias etapas de pré-processamento, tendo em vista o aprimoramento da qualidade das imagens. Neste capítulo, cada etapa e módulo utilizado são descritos. Em seguida, três dos modelos de segmentação das lesões de E.M. descritos anteriormente, FLEXCONN, *Light Weighted Unet* e NicMSLesions, foram selecionados para realização do treinamento. Ao observar os resultados preliminares com as arquiteturas encontradas na literatura, algumas modificações na estrutura da rede do melhor modelo foram implementadas com a finalidade de obter resultados mais satisfatórios na segmentação das lesões escleróticas.

6.1 Banco de dados

As sequências de MRI utilizadas neste projeto foram as sequências FLAIR e T1, obtidas de quatro diferentes bancos de dados: ISBI 2015 [75], MICCAI 2016 [76], Hospital das Clínicas da Faculdade de Medicina de Botucatu (banco de dados privado) e Laboratório de Imagem da Universidade de Liubliana [82]. O banco de imagens ISBI 2015 foi composto de imagens de 5 sujeitos com 4 ou 5 imagens obtidas em diferentes períodos de tempo. As imagens foram adquiridas no equipamento Philips 3T. Os parâmetros utilizados para obtenção da sequência T1 foram TE=6 ms e TR=10,3 ms e para a sequência FLAIR, TE=68 ms e TR=11.000 ms. Referente ao banco de dados do MICCAI 2016, as imagens dos 15 sujeitos foram adquiridas em diferentes domínios: GE Discovery 3T, Siemens Aera 1.5T e Siemens Verio 3T. As imagens fornecidas pelo Hospital das Clínicas de Botucatu foram obtidas pelo equipamento Siemens 3T. Os parâmetros para obtenção das sequências T1 foram TE=9 ms e TR 465 ms, enquanto para as sequências FLAIR foram TE=80 ms e TR=9.000 ms. Por fim, as imagens do banco de dados da Universidade de Liubliana foram obtidas através do sistema 3T Siemens Magnetom Trio, em que as imagens

T1 foram obtidas utilizando os parâmetros TE=20 ms e TR=2000 ms e as imagens FLAIR foram obtidas através de TR=5.000 ms e TE=392 ms

6.2 Pré-processamento de imagem

O pré-processamento foi constituído de diferentes técnicas, que objetivaram a melhoria da qualidade das imagens de ressonância magnética. Dessa forma, as imagens tornaram-se mais adequadas para análise computacional. As imagens passaram pelas seguintes operações: extração do encéfalo, co-registro rígido, conversão para resolução isotrópica (1mm^3) e correção de não-homogeneidade. Tais atividades foram realizadas no *software 3D Slicer* [83] e estão descritas nas próximas seções.

6.2.1 Extração de Estruturas Não-cerebrais

A presença de tecidos não-cerebrais podem causar erros de classificação das lesões, dada a similaridade de intensidade com tais tecidos. Assim, a etapa de extração do encéfalo foi extremamente significativa na análise imagens médicas. Esse procedimento foi realizado através do módulo *SwissSkullStripper*, cujo funcionamento é baseado no registro entre imagens, no qual uma máscara genérica do encéfalo é modificada para executar a extração do cérebro da imagem de interesse [84].

6.2.2 Registro e Reamostragem das Imagens

A operação de registro garante que duas imagens diferentes estejam alinhadas no mesmo sistema de coordenadas. Nesse trabalho, o módulo *BRAINSFit* foi utilizado. Tal método é baseado em exemplos de registro do ITK (*Insight Segmentation and Registration Toolkit*), adicionando diferentes formas de transformação e otimização [85]. Com a finalidade de manter um padrão entre as imagens, estas foram redimensionadas a $181 \times 217 \times 181$, com resolução isotrópica de 1mm^3 .

6.2.3 Correção de Não-homogeneidade

Um mesmo tecido pode apresentar variações de intensidade, devido ao processo de aquisição de MRI. A correção foi realizada pelo módulo *N4ITK*, que é uma variação do algoritmo *nonparametric nonuniform intensity normalization* (N3), realizando a correção de não-homogeneidade através de uma aproximação B-Spline e otimização hierárquica [86].

6.2.4 Normalização das Intensidades

A normalização garante que as intensidades das imagens de ressonância magnéticas adquiridas de protocolos diferentes pertençam a um mesmo intervalo. Neste trabalho, a normalização foi realizada durante o treinamento dos modelos. Em *Light Weighted Unet*, as intensidades foram escalonadas para o intervalo de $[0,1]$. Em seguida, o valor médio das intensidades foi subtraído,

seguido da divisão pelo desvio padrão. NicMSlesions realizou o mesmo processo da *Light Weighted Unet*, porém sem o escalonamento prévio. Em FLEXCONN, as intensidades negativas foram convertidas à zero, em seguida os valores foram divididos por um valor máximo encontrado através de operações estatísticas.

6.3 Treinamento dos modelos

Três arquiteturas de CNN foram escolhidas para serem implementadas nesse trabalho, com posterior análise dos resultados. O FLEXCONN, *Light Weighted Unet* e NicMSlesions foram os modelos escolhidos e implementados em Python com Tensorflow e Keras, utilizando o servidor remoto do INESC TEC. Para cada arquitetura escolhida, foram treinadas pelo menos duas versões: a primeira contendo 33 imagens dos bancos de dados ISBI [75] e MICCAI [76], enquanto a segunda foi treinada com 46 imagens (três imagens clínicas do Hospital de Botucatu e dez imagens do Universidade de Liubiana [82] foram adicionadas). Os detalhes de cada implementação foram descritos nas seções seguintes.

6.3.1 FLEXCONN

O treinamento do modelo teve número máximo de épocas igual a 20, utilizando otimizador Adam com taxa de aprendizagem inicial de 0,0001 e função de perda *binary crossentropy*. Além disso, a taxa de aprendizagem foi reduzida a metade a cada duas épocas, nas situações em que não houve progresso na acurácia. Além disso, o modelo poderia realizar uma parada antecipada, caso a acurácia continuasse no mesmo valor durante cinco épocas. O treinamento dos modelos durou aproximadamente 12 horas, utilizando *batch size* de 32, com a finalidade de ocupar menos memória RAM.

6.3.2 Light weighted Unet

A arquitetura original [77] é formada por cinco modalidades de MRI, porém nessa dissertação, a arquitetura foi modificada para treinar o modelo com apenas as sequências FLAIR e T1. O treinamento foi realizado em 60 épocas com otimizador Adam, com taxa de aprendizagem de 0,001 e função de perda *binary crossentropy*, utilizando *batch size* de 32. O treinamento de cada versão durou aproximadamente 2 horas, sendo a arquitetura treinada mais rapidamente entre os trabalhos selecionados.

6.3.3 NicMSlesions

Inicialmente, *patches* 3D das imagens foram extraídos e reorganizados com a mesma quantidade de amostra de lesões e tecidos saudáveis, como forma de controlar o *data imbalance* das imagens médicas. Em seguida, iniciou-se o treinamento da primeira rede. Após esse treinamento, as imagens foram testadas e os resultados dessas segmentações foram reorganizados para treinar

a segunda CNN. Ambas redes foram treinadas utilizando *batch size* igual a 128, *categorical crossentropy* como função de perda e otimizador Adadelta. Além disso, os modelos foram treinados com um número máximo de 400 épocas com paciência da parada antecipada igual a 50 épocas. O modelo implementado com 33 imagens completou o treinamento após 24 horas de duração, enquanto o modelo com 46 imagens concluiu o processo em 1 dia e 7 horas.

Em adição, outras versões deste modelo foram implementadas, através da transferência de aprendizado, retreinando apenas as camadas totalmente conectadas. O modelo treinado com as imagens ISBI [75] e MICCAI [76] foi utilizado como modelo base, e então as 10 imagens da Universidade de Lubliniana [82] foram utilizadas para realizar esse novo treinamento. Alternativamente, com o intuito de promover a adaptação de domínio mais suave, um novo modelo base foi treinado adicionando cinco imagens alvo no treinamento completo. Enquanto o retreinamento foi realizado com as imagens remanescentes. Nesses casos, o número máximo de épocas foi reduzido a 50 e paciência a três.

6.3.4 Modelo proposto: New Light Weighted Unet

Além das implementações dos modelos selecionados, uma nova versão da arquitetura *Light Weighted Unet*, contendo mais *skip connections* (Figura 6.1), foi proposta, objetivando melhorar a acurácia dos resultados, visto que tais conexões permitem recuperar informações espaciais perdidas durante o caminho de contração. Neste caso, os parâmetros de treinamento (número de épocas, *batch size* e função de perda) do modelo original foram mantidos sem alteração e o tempo de treinamento do modelo foi de 2 horas.

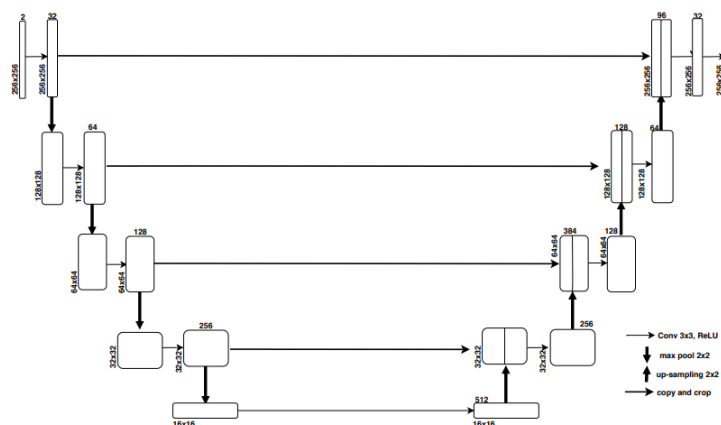


Figura 6.1: Arquitetura Light Weighted Unet modificada, utilizando 2 modalidades como entrada da rede e *skip connections* entre cada saída das camadas convolucionais no caminho de contração com a respectiva camada convolucional no caminho de expansão.

6.4 Segmentação das Lesões e Pós-processamento

A segmentação automática das lesões foi realizada após o pré-processamento de todas as MRI e aplicação dos modelos de treinamento. Em seguida, efetuou-se o pós-processamento das ima-

gens segmentadas para refinar o resultado, descartando possíveis falsos positivos. Neste trabalho, as imagens resultantes de cada modelo passaram pelos mesmos critérios de pós-processamento. No qual, a segmentação final apresentou apenas voxels classificados como lesões com probabilidade maior que 0.5 e regiões de tamanho igual ou maior a 10 voxels.

6.5 Métricas comparativas

As métricas foram essenciais para avaliar o desempenho obtido na segmentação das imagens. Neste trabalho, as seguintes métricas foram utilizadas: Coeficiente Dice, Valor Preditivo Positivo (VPP), Diferença de volumes (DV) e Quantificação volumétrica. Considera-se $|A|$ como a imagem segmentada automaticamente pelo modelo computacional e $|M|$ a segmentação manual realizada por especialista. O Coeficiente Dice indica o grau de similaridade entre as imagens comparadas e é calculado de acordo com a equação 6.1, quanto mais próximo o Coeficiente Dice for de uma unidade, mais similares as imagens foram consideradas.

$$Dice(A, M) = \frac{2(|A| \cap |M|)}{|A| + |M|} \quad (6.1)$$

O Valor Preditivo Positivo é dado pela razão entre os voxels verdadeiros positivos e o número total de voxes classificados como positivos, conforme a equação (6.2). Esta métrica indicou a porcentagem de voxels classificados como lesões, que foram corretamente classificados.

$$VPP(A, M) = \frac{|A| \cap |M|}{|A|} \quad (6.2)$$

A Quantificação Volumétrica das lesões foi realizada para comparar a quantidade total de lesões identificadas automaticamente com a quantidade real. Além disso, a Diferença de Volume Absoluta dada pela equação 6.3 é calculada, indicando que quanto mais próximo de zero, obtém-se melhor quantificação das lesões.

$$DV(A, M) = \frac{abs(|A| - |M|)}{|M|} \quad (6.3)$$

Capítulo 7

Resultados e Discussão

Neste capítulo, apresentam-se os resultados dos pré-processamento das imagens, além dos resultados quantitativos obtidos das segmentações realizadas por cada modelo, em imagens FLAIR e T1 de 20 pacientes diagnosticados com E.M. disponibilizadas pela Universidade de Liubliana [82]. Com a finalidade de avaliar a segmentação das imagens de ressonância magnética, utilizou-se diferentes métricas e testes estatísticos.

7.1 Resultados do Pré-processamento das Imagens

Nesta seção, os resultados de cada etapa de pré-processamento das imagens estão ilustrados em uma imagem clínica do Hospital de Botucatu utilizada no treino dos modelos.

7.1.1 Extração das Estruturas Não-cerebrais

A Figura 7.1 apresenta os cortes da imagem de ressonância FLAIR e T1 antes e depois a segmentação do encéfalo, diminuindo a chance do programa selecionar falsos positivos.

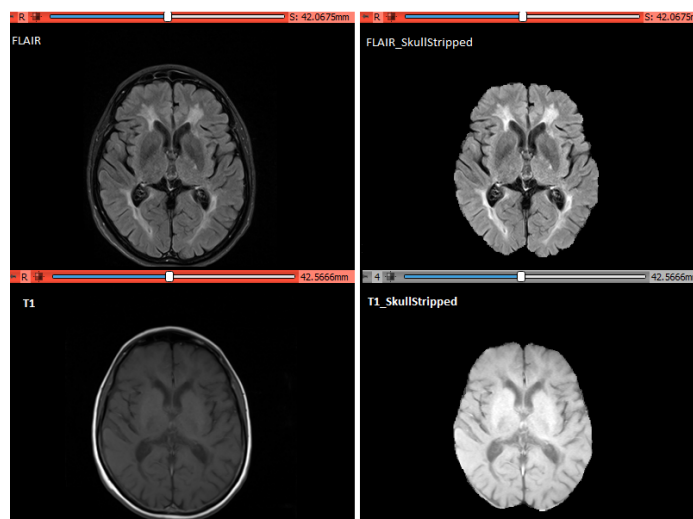


Figura 7.1: Cortes das sequências FLAIR e T1 antes e após a extração do encéfalo .

7.1.2 Registro das imagens

O registro rígido foi realizado nas imagens com a finalidade de obter as dimensões de (181,217,181) e resolução isotrópica dos voxels de $1mm^3$, conforme representado na Figura 7.2, onde a imagem do banco de dados de ISBI [75] foi utilizada como referência. Além disso, a operação foi realizada entre as modalidades T1 e FLAIR, com o objetivo de garantir que as estruturas fossem espacialmente correspondentes em ambas imagens.

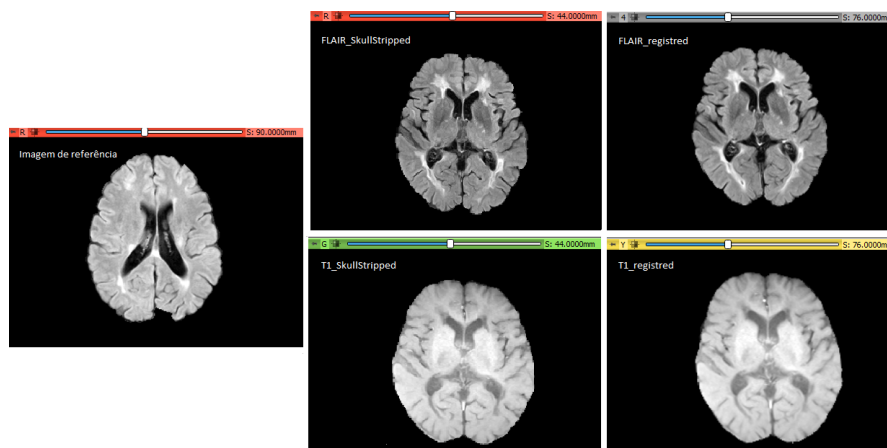


Figura 7.2: Comparação entre cortes das imagens FLAIR e T1. O primeiro ícone é uma fatia da imagem referência, enquanto os outros representam o antes e depois da operação de alinhamento da imagem móvel.

7.1.3 Correção de não-homogeneidade

A correção de não-homogeneidade foi realizada nas modalidades FLAIR e T1 para garantir que o mesmo tecido não sofresse variações de intensidades na imagem, conforme observado o antes e depois dessa etapa na Figura 7.3. Na qual, a segunda coluna representa o sinal *bias field*.

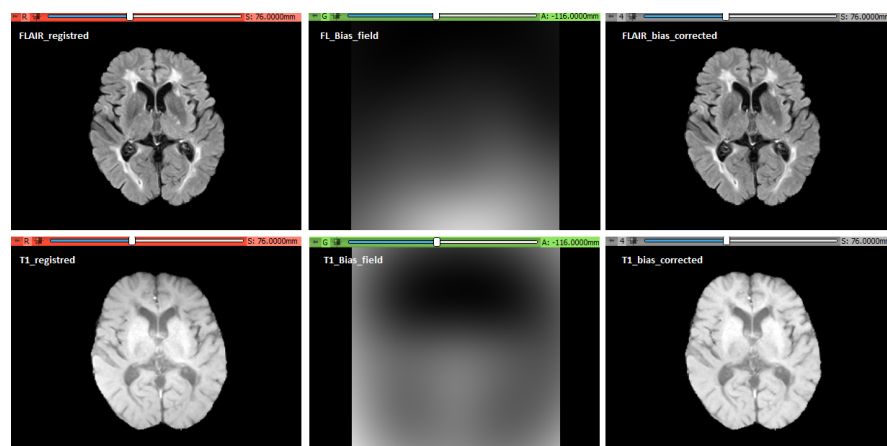


Figura 7.3: Cortes das imagens FLAIR e T1 antes e após a correção da não-homogeneidade, na coluna do meio apresenta-se o sinal de *bias field* de cada modalidade

7.2 Segmentação das Lesões Escleróticas

Nesta seção, as imagens das lesões segmentadas pelos modelos treinados foram analisadas e comparadas estatisticamente. Testes paramétricos e não-paramétricos (Teste-T, Mann-Whitney, Kruskal-Wallis e Dunn) foram realizados para verificar a existência de diferenças significativas nas métricas Dice, VPP e DV entre as imagens segmentadas pelos modelos.

7.2.1 Segmentação FLEXCONN

Os resultados das segmentações realizadas pelos dois modelos FLEXCONN, o primeiro treinado com 33 imagens e o segundo com 46 imagens, são apresentados e discutidos. Ao analisar a variação dos coeficientes Dice, na Figura 7.4, podemos observar uma melhora na segmentação das lesões pelo segundo modelo. Isso foi confirmado, ao aplicarmos os testes estatísticos para amostras independentes. Inicialmente, verificou-se a normalidade das amostras do primeiro modelo, e a não-normalidade dos dados do segundo, de acordo com o teste de Shapiro-Wilk. Em seguida, o teste de Mann-Whitney concluiu que houve diferença significativa no coeficiente Dice entre os modelos 1 e 2, com valores de medianas iguais a 0,497 e 0,688, respectivamente com $U=74$ e $p=0,0002$.

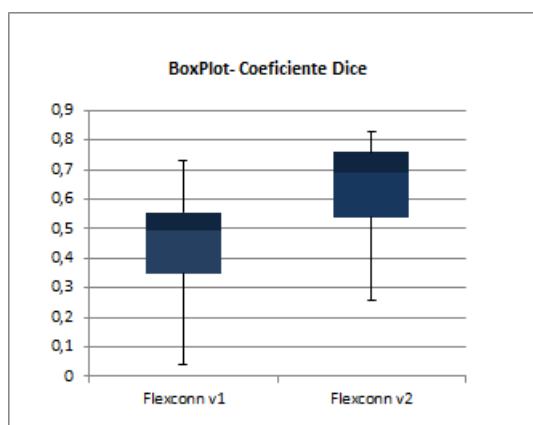


Figura 7.4: Boxplot da variação de coeficiente Dice.

Conforme observado na Figura 7.5, a segmentação da primeira versão está ilustrada em amarelo e a segmentação da segunda versão está em vermelho. Notou-se que o modelo treinado com menos imagens classificou como lesões, mais precisamente, os voxels hiper-intensos que apresentaram alto brilho, enquanto os voxels que apareceram como hiperintensidades não tão brilhosas foram classificadas erroneamente. Em contrapartida, a segunda versão do modelo foi capaz de detectar mais lesões.

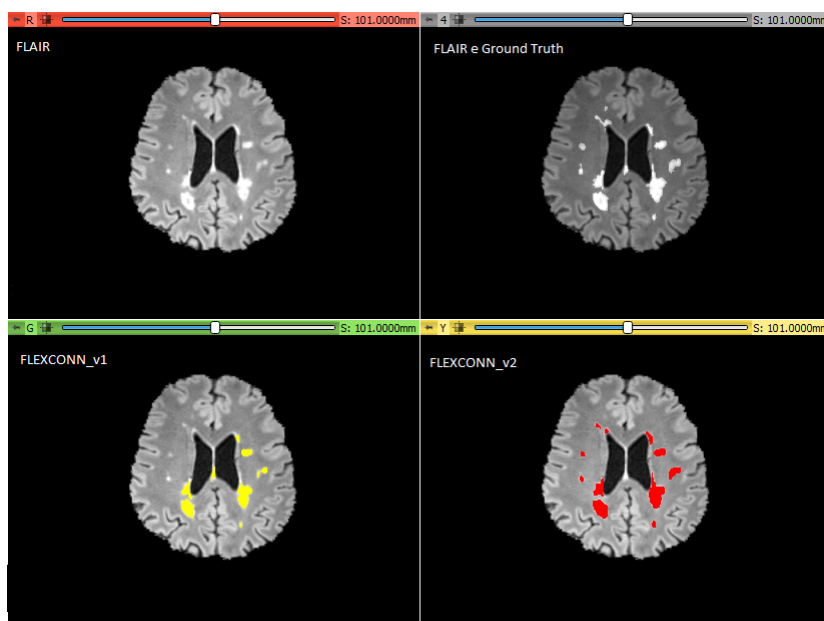


Figura 7.5: Cortes das imagens segmentadas do Paciente 10 pelos modelos FLEXCONN. Na primeira linha, há a sequência FLAIR e a segmentação manual. Enquanto na segunda linha, a segmentação da versão 1 do modelo está representada em amarelo e a segmentação da versão 2 está em vermelho.

O teste de Shapiro-Wilk também indicou a não normalidade dos dados de VPP para ambos modelos, dessa maneira o teste não paramétrico de Mann-Whitney foi realizado para verificar se houve diferença significativa dos resultados obtidos. As medianas obtidas foram iguais a 0,886 e 0,882, além de obter $U=189$ e $p=0,389$, concluindo que não houve diferença significativa no Valor Preditivo Positivo. Logo, em ambos modelos, os voxels classificados como lesões estão classificados corretamente.

Por fim, analisou-se a Diferença Absoluta de Volume (DV). Na Tabela 7.1, observa-se as médias e desvios padrão da DV para cada modelo. Na qual, o segundo modelo obteve uma média menor que o primeiro, sendo um indicativo de que a quantificação volumétrica das lesões foi melhor realizada por versão 2 do FLEXCONN. Estatisticamente, ambos grupos de amostras apresentaram distribuição normal e homogênea. Dessa maneira, foi possível realizar o teste paramétrico Teste-T, com $t=6,795$ e $p=4,67E-08$, indicando que houve diferença significativa. Portanto, o segundo modelo identificou melhor os voxels das lesões.

Tabela 7.1: Médias e desvios padrão dos valores de DV das duas versões do modelo FLEXCONN.

–	DV (versão 1)	DV (versão 2)
Média	0,624	0,340
Desvio padrão	0,028	0,031

Os gráficos ilustrados na Figura 7.6 representam a relação entre a quantificação volumétrica realizada de forma automatizada e manual. Observou-se que as retas de tendência linear representaram de maneira coerente o comportamento das variáveis analisadas, com coeficiente $R^2=0,895$ e $R^2=0,950$ para primeira e segunda versão. No primeiro caso, houve uma sub-quantificação volumétrica das lesões, isso pôde ser observado pelo coeficiente angular igual a 0,364. Em contrapartida, o modelo treinado com mais imagens foi capaz de identificar melhor as lesões, conforme indicado pelo coeficiente angular da reta igual a 0,631. Podemos observar que quanto mais próximo de uma unidade, melhor a correlação entre os volumes segmentado automaticamente com os volumes obtidos de forma manual.

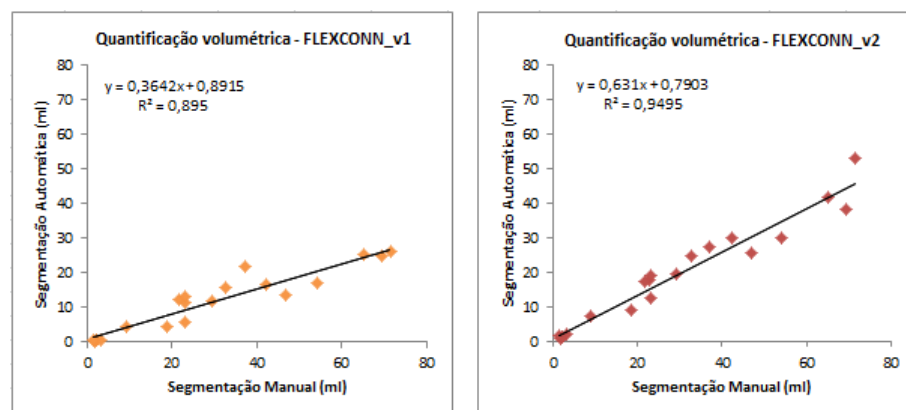


Figura 7.6: Gráficos de dispersão da quantificação volumétrica das lesões: à esquerda, relação da segmentação da primeira versão do FLEXCONN e, à direita, a segunda versão.

A diferença na identificação das lesões entre as versões dos modelos FLEXCONN ficou bastante evidente na Figura 7.7, onde vários cortes das imagens de ressonâncias magnéticas foram ilustrados.

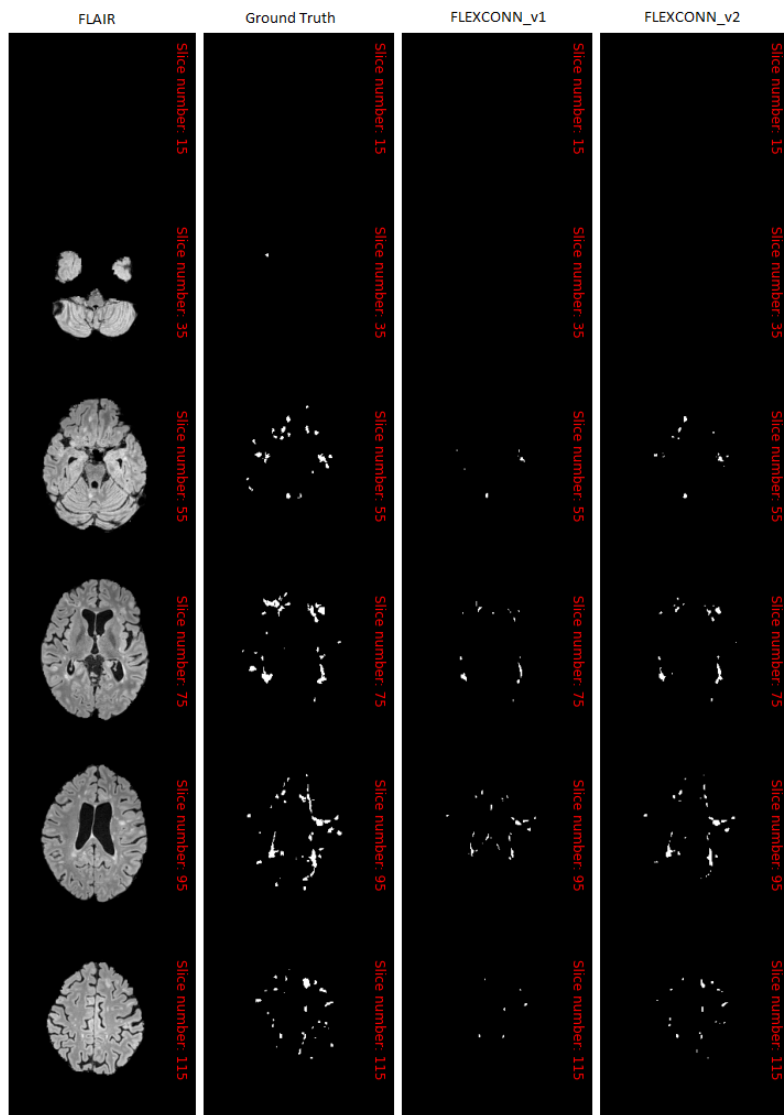


Figura 7.7: Comparação da segmentação entre os dois modelos FLEXCONN treinados. Na primeira coluna, é possível observar cortes da sequência FLAIR, na segunda coluna está o *Ground Truth* dado por um especialista e nas colunas seguintes aparecem as segmentações da versão 1 e 2 dos modelos FLEXCONN.

7.2.2 Segmentação Light Weighted Unet

Três modelos adaptados da arquitetura *Light Weighted Unet* foram treinados, onde o primeiro foi treinado com um banco de imagens menor do que os outros e além disso, o último modelo teve sua arquitetura modificada conforme descrito no capítulo anterior. Nesta seção, comparações estatísticas foram realizadas entre as métricas obtidas das segmentações das imagens, para avaliar o desempenho de cada programa. Os testes de Shapiro-Wilk concluíram que as amostras dos coeficientes Dice, VPP e DV obtidos não apresentaram distribuição normal, então os testes de Mann-Whitney foram realizados para identificar se houve diferença significativa entre os valores obtidos de cada modelo.

Ao examinar a variação dos coeficientes Dice, notou-se na Figura 7.8, um progresso nos valores calculados, a cada implementação dos diferentes modelos. Além disso, os coeficientes Dice do modelo proposto apresentaram menor intervalo de variação, e inclusive, menor diferença relativo ao *outlier*.

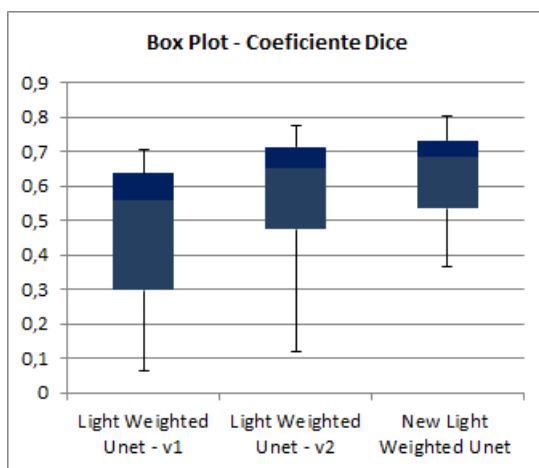


Figura 7.8: Boxplot da variação de coeficiente Dice.

Em relação ao teste de Mann-Whitney, inicialmente realizado entre a versão 1 e 2 dos modelos *Light Weighted Unet*, concluindo que houve diferença relevante, onde $U=129$ e $p=0,028$. Em seguida, o mesmo foi feito entre o modelo versão 2 e o modelo proposto *New Light Weighted Unet*. Neste caso, não houve diferença significativa entre os dados analisados, com $p=0,144$ e $U=160$. As medianas dos dados de coeficiente Dice obtidas para cada implementação foram 0,560, 0,650 e 0,687, respectivamente.

Referente ao Valor Preditivo Positivo, os testes estatísticos não indicaram diferença significativa entre os modelos. O primeiro teste resultou em $p=0,170$ e $U=164$, enquanto o segundo teste obteve $U=149$ e $p=0,087$. Embora, na Tabela 7.2, fosse possível observar um leve decaimento nas medianas do VPP nos dois últimos modelos.

Tabela 7.2: Medianas calculadas das amostras de VPP para cada modelo Unet adaptado.

Modelos	Medianas das amostras de VPP
<i>Light Weighted Unet</i> (versão 1)	0,882
<i>Light Weighted Unet</i> (versão 2)	0,852
<i>New Light Weighted Unet</i>	0,803

Por último, a Diferença Absoluta de Volume foi analisada. Observou-se na Tabela 7.3, que as medianas da DV dos modelos *Light Weighted Unet* (versão 2) e *New Light Weighted Unet* foram melhores do que o primeiro modelo, já que foram mais próximas de zero. Em complemento, os resultados dos testes não-paramétricos concluíram que houve diferença significativa na métrica estudada entre os modelos implementados. Nos quais, o primeiro teste obteve $U=103$ e $p=0,004$ e o segundo obteve $U=49$ e $p=6,67E-06$.

Tabela 7.3: Medianas dos valores de DV das diferentes implementações do modelo Unet adaptado.

Modelos	Medianas das amostras de DV
<i>Light Weighted Unet</i> (versão 1)	0,519
<i>Light Weighted Unet</i> (versão 2)	0,355
<i>New Light Weighted Unet</i>	0,252

Os gráficos de quantificação volumétrica correlacionando a segmentação realizada de forma automatizada e manual estão ilustradas na Figura 7.9. Em todos os casos, as retas de tendência linear representaram muito bem as variáveis estudadas, apresentando coeficiente $R^2=0,918$, $R^2=0,963$ e $R^2=0,970$. Ao analisar o primeiro gráfico, notou-se que a quantificação volumétrica ocorre de forma razoável, com coeficiente angular da reta igual a 0,524. Enquanto que, no segundo gráfico, houve uma melhora na identificação das lesões, cujo coeficiente angular da reta de tendência é igual a 0,663. Por fim, o último modelo apresentou melhor performance na quantificação volumétrica, conforme esperado ao analisar a Diferença de Volume Absoluta anteriormente, neste caso, o coeficiente angular foi igual a 0,756.

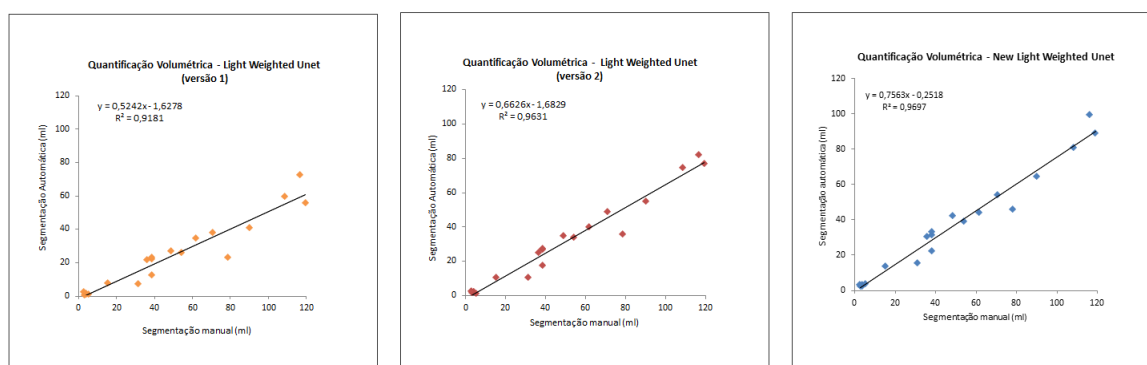


Figura 7.9: Gráficos de dispersão da quantificação volumétrica das lesões pelos três modelos Light Weighted Unet implementados.

A Figura 7.10 apresenta as diferentes segmentações das lesões escleróticas obtidas por cada modelo implementado. Observou-se que os modelos treinados com mais dados (versão 2 em vermelho) e mais *skip connections* (*New Light Weighted Unet* em azul) foram capazes de identificar mais lesões, portanto se aproximaram melhor da segmentação manual realizada pelo especialista.

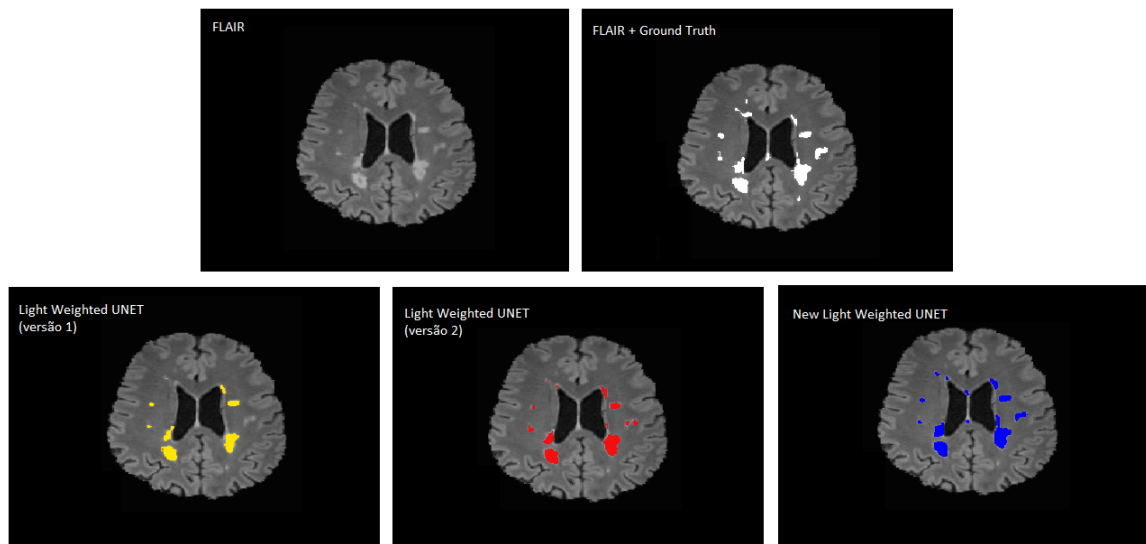


Figura 7.10: Cortes das imagens segmentadas do Paciente 10 pelos modelos *Light Weighted Unet*. Na primeira linha, há a sequência FLAIR e a segmentação manual. Enquanto na segunda linha, a segmentação da versão 1 está representada em amarelo, a segmentação da versão 2 está em vermelho e a segmentação do modelo proposto está em azul.

A Figura 7.11 apresenta vários cortes das imagens segmentadas pelos diferentes adaptações dos modelos *Light Weighted Unet* treinados. Observa-se que a identificação de lesões foi melhor realizada no modelo proposto.

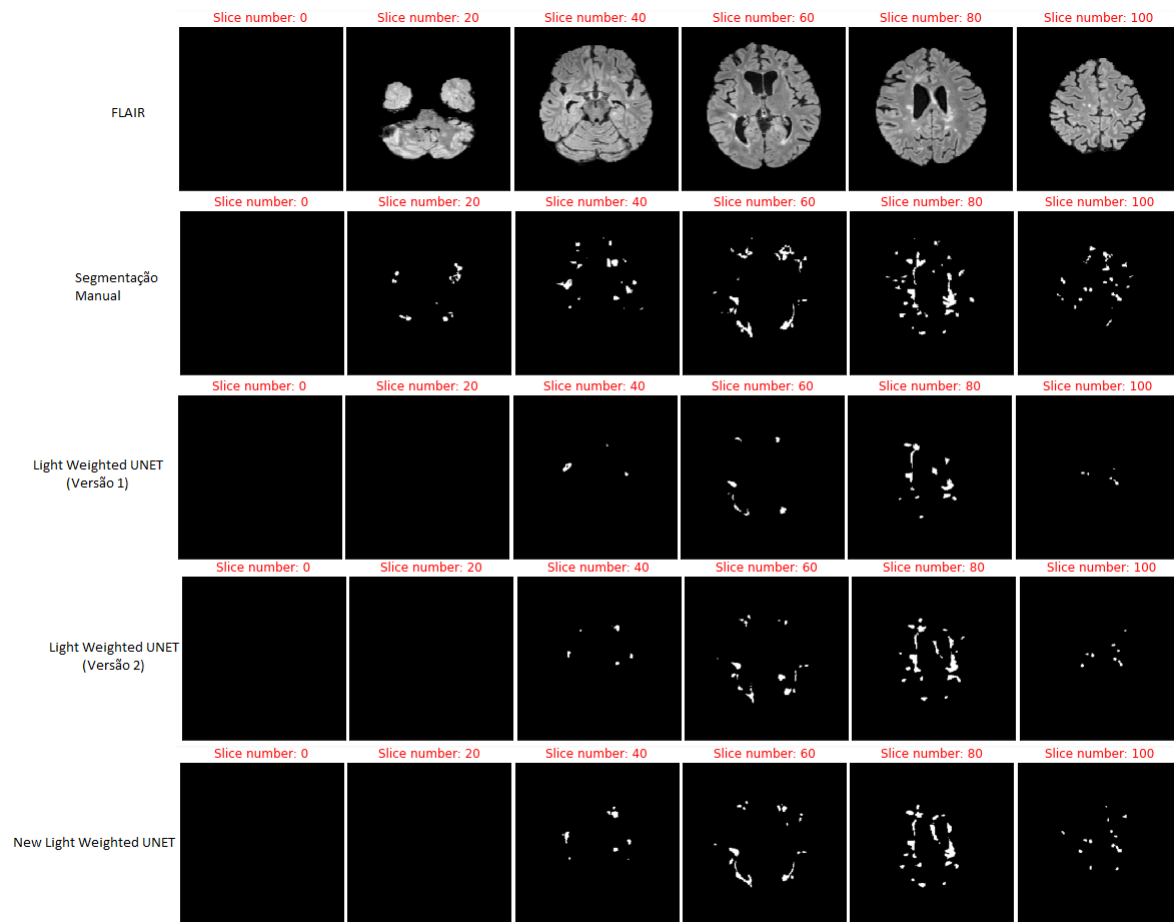


Figura 7.11: Comparação da segmentação entre os três modelos *Light Weighted Unet* treinados. Na primeira linha, as fatias da modalidade FLAIR estão ilustradas, já na segunda linha, apresentam-se fatias da segmentação manual, enquanto na terceira e quarta linha aparecem as segmentações da primeira e segunda versão dos modelos *Light Weighted Unet*. Por fim, na quarta linha, observa-se a segmentação realizada pelo modelo modificado, ao adicionar mais *skip connections*.

7.2.3 NicMSLesions

O projeto NicMSLesions possibilitou a criação de modelos de duas formas diferentes: treinamento completo de todas as camadas ou através do re-treino das camadas totalmente conectadas, por transferência de aprendizado de um modelo base. Nessa seção, analisou-se o desempenho desses modelos em categorias separadas.

7.2.3.1 Treinamento Completo

Os modelos foram treinados com 33 imagens e a segunda versão com 46 imagens. A Figura 7.12 ilustra a variação dos coeficientes Dice entre as duas versões do modelo NicMSLesions totalmente treinado. As segmentações de ambos modelos apresentaram valores notáveis dos coeficientes Dice, o conjunto de dados da primeira versão apresentou mediana igual a 0,627, enquanto o segundo apresentou valor de 0,733. O teste estatístico não indicou distribuição normal dos dados, portanto, o teste de Mann-Whitney foi utilizado para comparar o desempenho dos modelos, concluindo melhora expressiva do segundo em relação ao primeiro, com $U=102$ $p=0,004$.

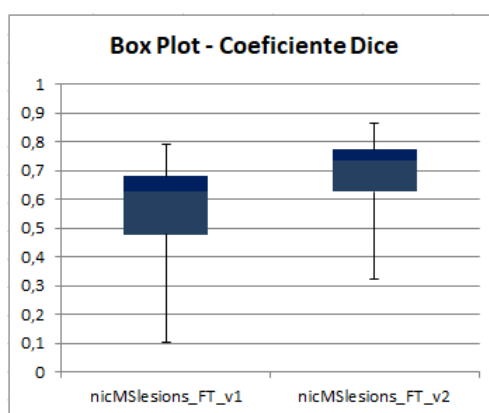


Figura 7.12: Boxplot da variação dos coeficientes Dice dos modelos NicMSLesions de treinamento completo.

Referente ao VPP, as métricas do segundo modelo foram superiores em relação ao primeiro, com médias, desvios padrão e medianas apresentados na Tabela 7.4. A distribuição dos dados de ambos modelos foram similares e condensados, já que os desvios padrão foram valores muito próximos entre si e próximos de zero. Estatisticamente, o teste de Shapiro-Wilk verificou a não-normalidade da distribuição dos dados, portanto o teste não-paramétrico de Mann-Whitney foi realizado para avaliar o desempenho dos modelos. Concluiu-se que o modelo treinado com mais imagens apresentou melhora significativa, com $U=95$ e $p=0,002$ e portanto, apresenta mais verdadeiros positivos do que o primeiro programa.

Tabela 7.4: Médias, desvios padrão e medianas do VPP de cada modelo.

-	VPP (modelo versão 1)	VPP (modelo versão 2)
Média	0,616	0,752
Desvio Padrão	0,217	0,198
Mediana	0,675	0,816

Relativo a Diferença Absoluta de Volume, a Tabela 7.5 apresenta as médias, desvios padrão e medianas da métrica em análise. A primeira versão do NicMSLesions apresentou média próxima ao dobro do valor obtido pela segunda versão. Além disso, os dados obtidos do segundo modelo estão distribuídos de maneira mais condensada, já que o desvio padrão foi 4,75 vezes menor que os dados da primeira versão. Entretanto, as medianas dos dois modelos foram valores próximos entre si. A análise estatística de Mann-Whitney foi realizada, porque os dados não apresentaram distribuição normal. Como resultado, não houve diferença significativa entre os valores de DV dos dois modelos, com $U=198$ e $p=0,484$.

Tabela 7.5: Médias, desvios padrão e medianas dos valores de DV das duas versões do modelo NicMSLesions de treinamento completo.

-	DV (versão 1)	DV (versão 2)
Média	0,447	0,240
Desvio padrão	1,165	0,245
Mediana	0,165	0,197

Os gráficos de dispersão da quantificação volumétrica associando a segmentação automática e manual foram ilustradas na Figura 7.13. Observou-se que as retas de tendência linear representaram as variáveis estudadas de maneira muito coerente, com coeficiente $R^2=0,938$ e $R^2=0,965$. Ao averiguar o primeiro gráfico, notou-se que a quantificação volumétrica foi realizada de maneira eficaz, onde o coeficiente angular da reta foi igual a 0,855. Enquanto que, no segundo caso, houve uma leve diminuição do coeficiente angular da reta de tendência, para 0,779. Apesar dessa pequena diminuição, a quantidade de lesões segmentadas pela segunda versão está mais próxima do valor manual, conforme já discutido ao analisar o coeficiente Dice, o Valor Preditivo Positivo e a Diferença de Volume Absoluto.

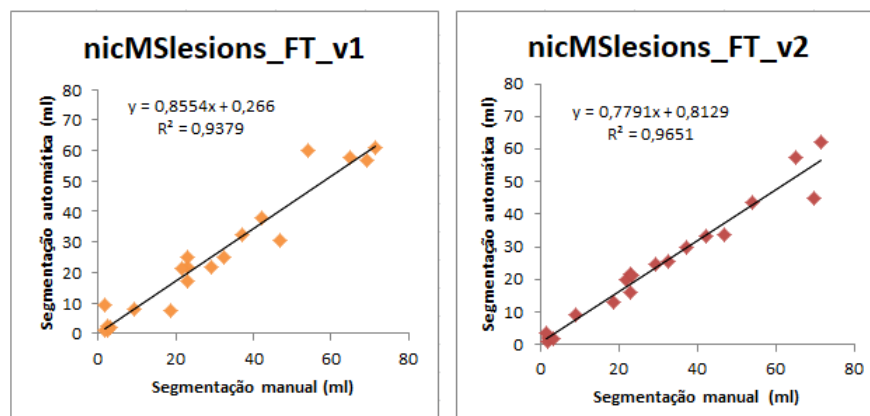


Figura 7.13: Gráficos de dispersão da quantificação volumétrica das lesões pelos modelos NicMSlesions implementados.

Foi possível observar na Figura 7.14, que os dois modelos classificaram as lesões de forma coerente, mas o segundo programa conseguiu identificar mais lesões e se aproximar melhor do *ground truth*.

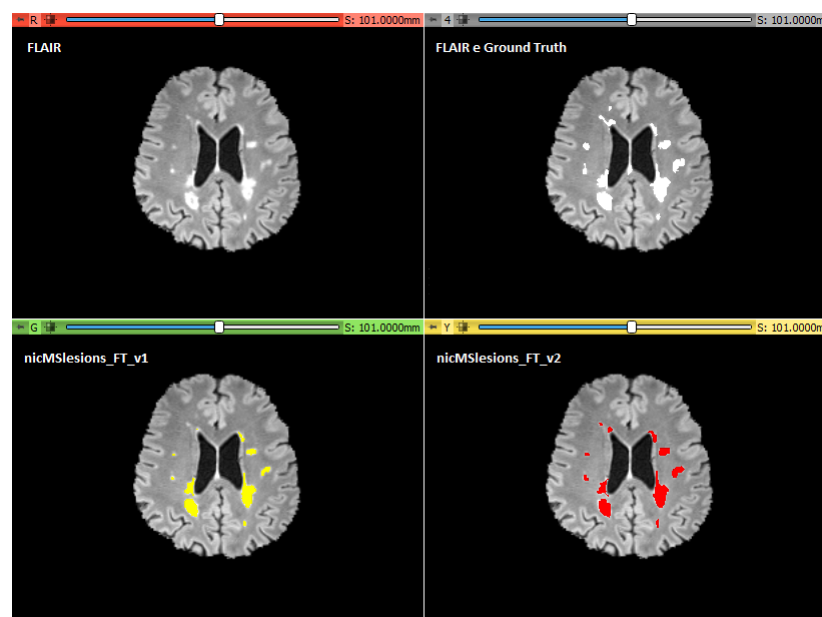


Figura 7.14: Cortes das imagens segmentadas do Paciente 10 pelos modelos NicMSlesions. Na primeira linha, há a sequência FLAIR e a segmentação manual. Enquanto na segunda linha, a segmentação da versão 1 do modelo está representada em amarelo e a segmentação da versão 2 está em vermelho.

A Figura 7.15 apresenta vários cortes das imagens segmentadas por cada modelo NicMSlesions (*Full Trained*) implementado. Observa-se que a versão com adaptação de domínio apresentou melhor desempenho na identificação das lesões, conforme já analisados pelos testes estatísticos.

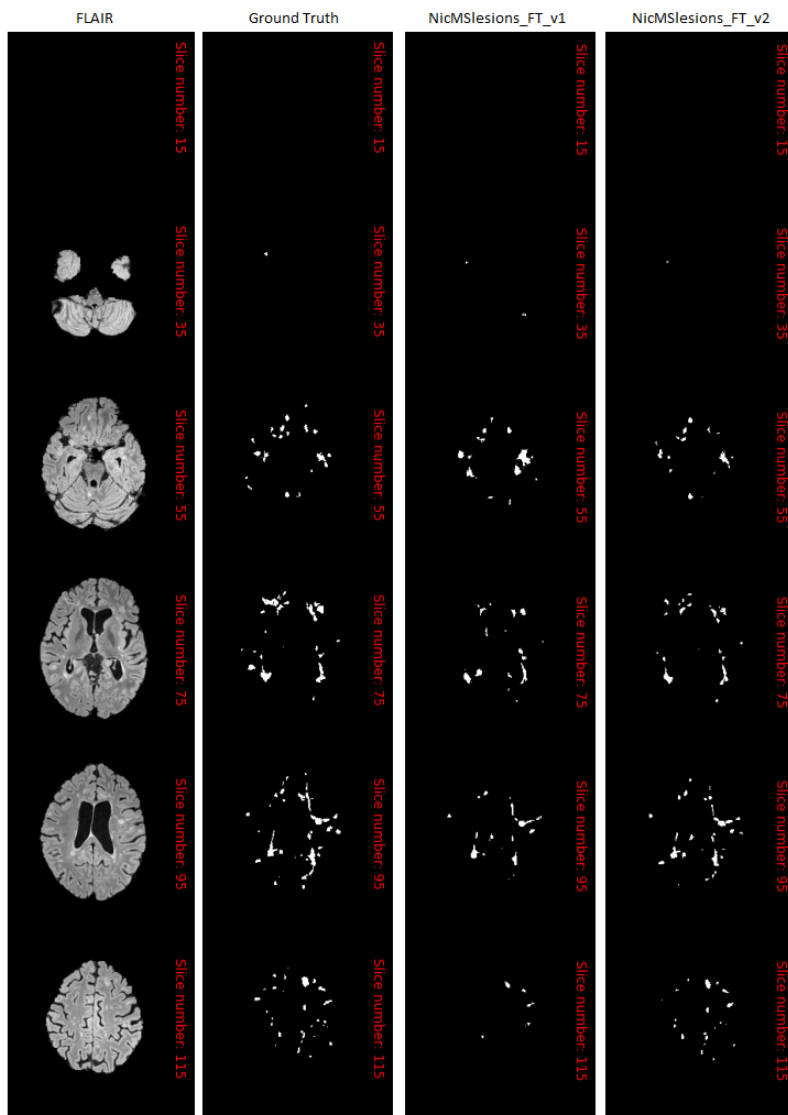


Figura 7.15: Resultados da segmentação dos dois modelos nicMSlesions treinados completamente. Na primeira coluna, é possível observar fatias da modalidade FLAIR, enquanto a segunda apresenta o *Ground Truth* e nas colunas seguintes, as segmentações da versão 1 e 2 dos modelos nicMSlesions estão indicadas.

7.2.3.2 Transferência de Aprendizado

A primeira versão desse modelo utilizou o modelo base treinado com imagens de ISBI [75] e MICCAI [76] e houve o retreinamento das FCL, com 10 imagens da Universidade de Liubliana [82]. Em contrapartida, a segunda versão foi treinada a partir de um modelo base treinado com 38 imagens (5 imagens adicionais da Universidade de Liubliana [82]), enquanto as 5 imagens remanescentes foram utilizadas para realizar o retreinamento das FCL.

A Figura 7.16 exibe as variações dos coeficientes Dice, obtidos da segmentação das imagens teste pelos dois modelos treinados a partir da transferência de aprendizado. Notou-se uma melhora nos valores do segundo modelo, em relação ao primeiro, apresentando medianas de 0,642 e 0,462, respectivamente. Tais coeficientes apresentaram distribuição normal, de acordo com o teste de Shapiro-Wilk. Logo, o Teste-T foi realizado para identificar se a melhora de uma versão para a outra foi considerável. O resultado do teste estatístico concluiu que houve diferença significativa, com $p=7,86E-05$.

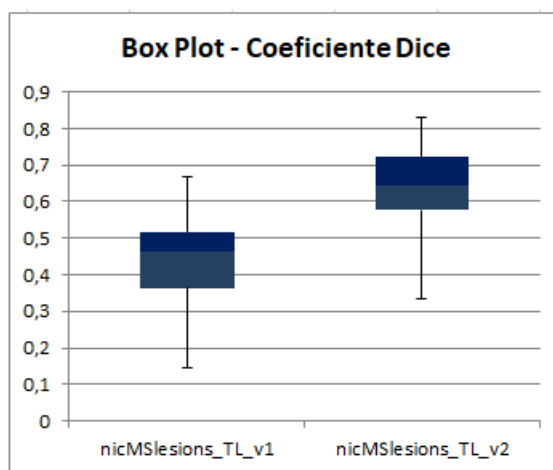


Figura 7.16: Boxplot da variação dos coeficientes Dice dos modelos NicMSlesions por transferência de aprendizado

Na Tabela 7.6, foi possível observar as métricas estatísticas de VPP. Em que, os valores médio e de desvio padrão foram relativamente próximos. Entretanto, os valores de mediana foram um pouco melhores no modelo treinado diretamente com 10 imagens, do que os da segunda versão. Durante o estudo da distribuição dos dados, o teste de Shapiro-Wilk revelou a não normalidade das amostras. Portanto, o teste de Mann-Whitney foi utilizado e indicou diferença significativa nas amostras analisadas, com $U=129$ e $p=0,028$.

Tabela 7.6: Médias, desvios padrão e medianas dos valores de VPP das duas versões do modelo NicMSLesions, treinados através da transferência de aprendizado.

–	VPP (versão 1)	VPP (versão 2)
Média	0,771	0,720
Desvio padrão	0,216	0,199
Mediana	0,872	0,797

Em relação a Diferença Absoluta de Volume, Tabela 7.7, observou-se uma superioridade no dados da segunda versão, onde os valores de média e mediana são mais próximos de zero. Indicando, que o segundo modelo identificou mais lesões do que o primeiro. Estatisticamente, as amostras não apresentaram distribuição normal, portanto o teste de Mann-Whitney foi utilizado, concluindo que houve diferença significativa, com $U=57$ e $p=2,25E-05$.

Tabela 7.7: Médias, desvios padrão e medianas dos valores de DV das duas versões do modelo NicMSLesions, treinados através da transferência de aprendizado.

–	DV (versão 1)	DV (versão 2)
Média	0,563	0,317
Desvio padrão	0,187	0,369
Mediana	0,569	0,273

A Figura 7.17 apresenta os gráficos de dispersão da quantificação volumétrica das lesões escleróticas, pelos dois modelos NicMSLesions treinados por transferência de aprendizado. A linha de regressão ajustada para o primeiro caso não representa muito bem os dados, pois R^2 foi igual a 0,672. Entretanto, no segundo gráfico, o coeficiente R^2 foi igual a 0,902, indicando ótimo ajuste. Analisando a quantificação das lesões, o segundo modelo obteve resultados mais satisfatórios, conforme esperado pela análise de DV. No primeiro gráfico, houve uma sub-quantificação de lesões, onde o coeficiente angular da reta de tendência foi igual a 0,267, enquanto o mesmo coeficiente para a segunda amostra de dados apresentou valor igual a 0,642.

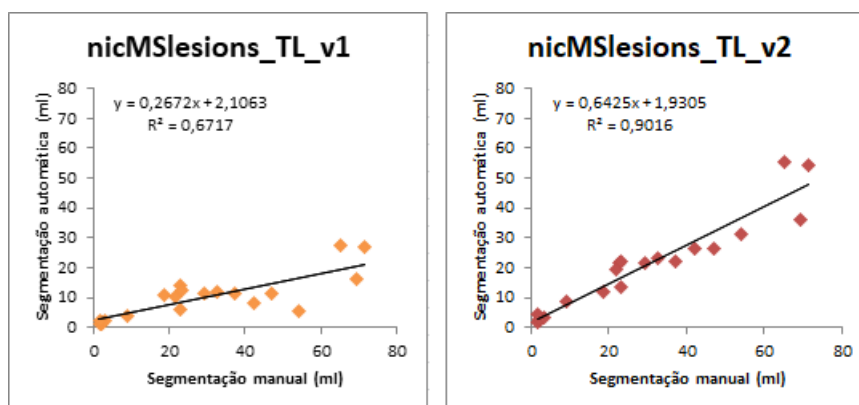


Figura 7.17: Gráficos de dispersão da quantificação volumétrica das lesões pelos modelos NicMSLesions, treinados por transferência de aprendizado.

As segmentações dos modelos NicMSLesions treinados através da transferência de aprendizagem estão ilustradas na Figura 7.18. Observou-se que a imagem segmentada pelo segundo modelo (em vermelho) se assemelha mais ao *Ground Truth*, ao comparar com a segmentação da primeira versão (em amarelo). Tal situação era esperada, conforme análise das variações de Coeficiente Dice para os dois grupos.

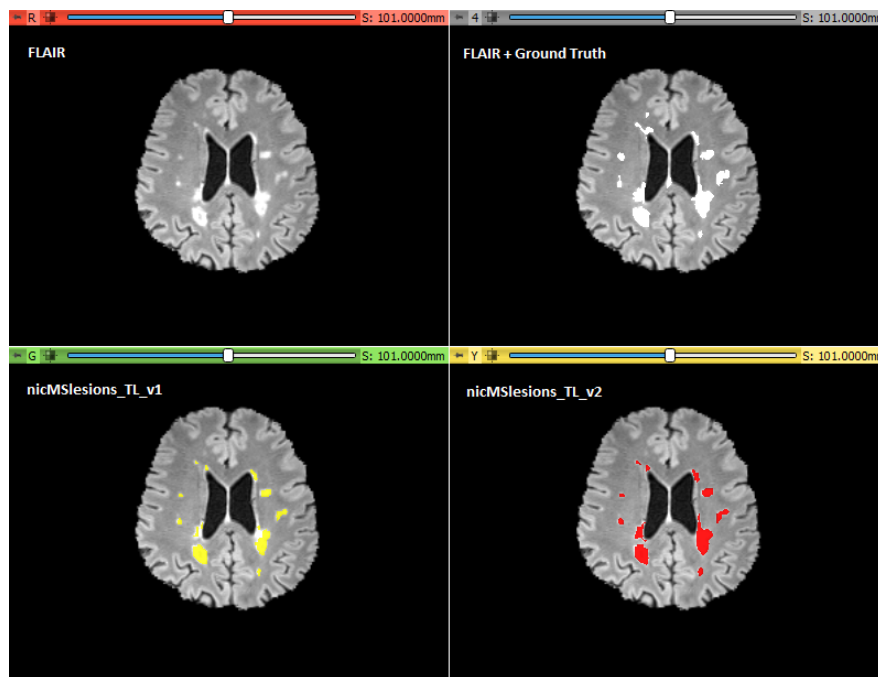


Figura 7.18: Cortes das segmentações NicMSLesions, treinados por transferência de aprendizado. Na primeira linha, há a sequência FLAIR e a segmentação manual. Já na segunda, a segmentação da versão 1 do modelo está em amarelo e a segmentação da versão 2 está em vermelho.

As segmentações das lesões em vários cortes da MRI, pelas versões dos modelos NicMSLesions (por *Transfer Learning*), podem ser visualizadas com mais detalhes na Figura 7.19. Nota-se que a segunda versão do modelo foi capaz de identificar mais lesões.

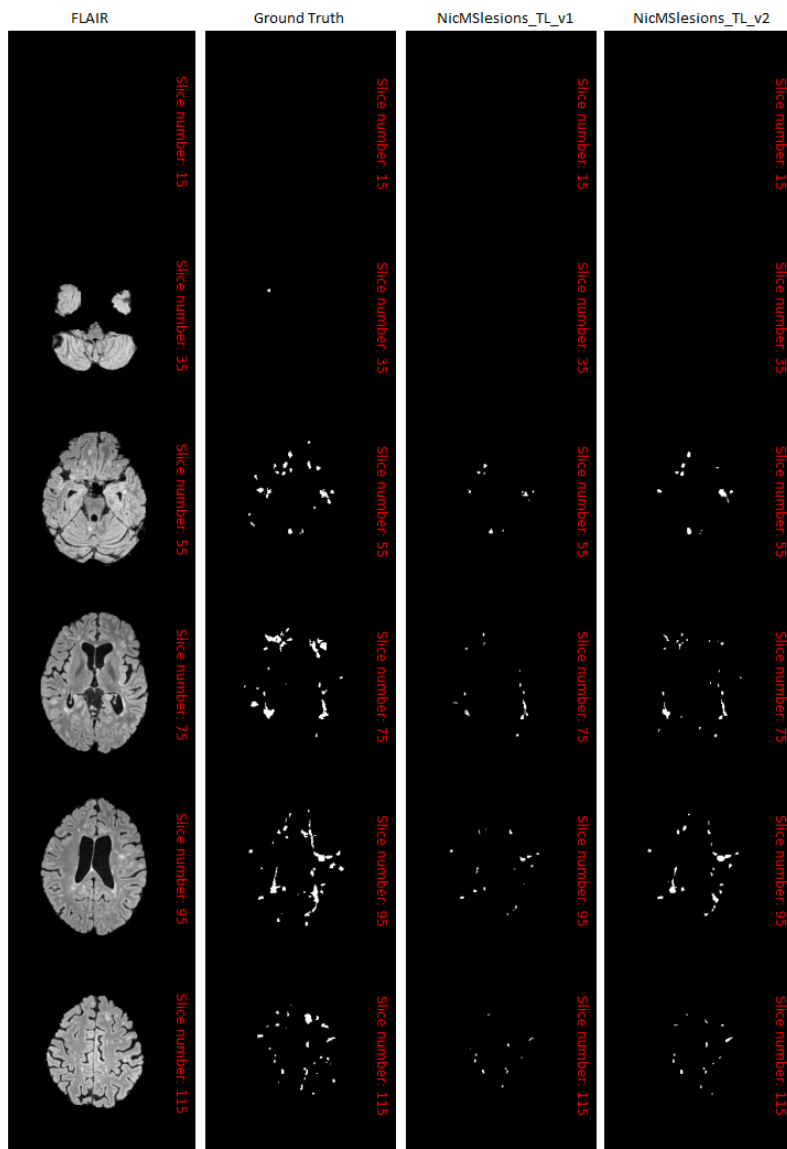


Figura 7.19: Resultados da segmentação dos modelos nicMSLesions treinados por transferência de aprendizado. Na primeira coluna, é possível observar cortes da modalidade FLAIR, enquanto a segunda apresenta o *Ground Truth* e nas colunas seguintes, as segmentações da versão 1 e 2 dos modelos nicMSLesions estão indicadas.

7.3 Comparações Entre os Modelos Treinados

Nesta dissertação, vários métodos de segmentação de lesões automatizados foram treinados, além da implementação de uma arquitetura proposta. Inicialmente, a influência de mais dados de treino e adaptação de domínio foram analisadas. Nesta seção, os melhores modelos implementados de FLEXCONN, *New Light Weighted Unet* e NicMSlesions (FT indica o modelo por treinamento completo e TL, o modelo por transferência de aprendizado) foram avaliados estatisticamente. Na Tabela 7.8, observou-se as medianas obtidas das métricas Dice, VPP e DV para segmentações de cada modelo. O teste não-paramétrico de Kruskal-Wallis foi utilizado para verificar se houve diferenças significativas das métricas entre os quatro diferentes modelos implementados. Em casos que a diferença significativa foi confirmada, o teste de Dunn foi realizado para identificar qual pares de modelos se diferenciaram.

Tabela 7.8: Medianas das métricas avaliativas das segmentações dos melhores modelos implementados.

Métricas	FLEXCONN	New Light Weighted Unet	NicMSlesions (FT)	NicMSlesions (TL)
Medianas de DICE	0,688	0,687	0,733	0,642
Medianas de VPP	0,882	0,803	0,816	0,797
Medianas de DV	0,343	0,252	0,197	0,273

Ao analisar os coeficientes Dice obtidos por cada modelo, o teste verificou que não ocorreu diferença relevante entre os modelos, com $p=0,321$. Entretanto, o modelo NicMSlesions (FT) alcançou maior mediana. Em relação ao Valor Preditivo Positivo, o teste de Kruskal-Wallis indicou diferença significativa, ao nível de confiança de 5%, onde $p=0,032$. Em seguida, o teste de Dunn, Tabela 7.9, verificou diferença entre os modelos FLEXCONN e *New Light Weighted Unet* e entre FLEXCONN e NicMSlesions (TL), treinado por transferência de aprendizado, onde o programa FLEXCONN teve desempenho superior.

Tabela 7.9: Teste de Dunn para análise do VPP.

Modelo 1	Modelo 2	Valor-p
FLEXCONN	<i>New Light Weighted Unet</i>	0,012
FLEXCONN	NicMSlesions (FT)	0,074
FLEXCONN	NicMSlesions (TL)	0,009
<i>New Light Weighted Unet</i>	NicMSlesions (FT)	0,471
<i>New Light Weighted Unet</i>	NicMSlesions (TL)	0,903
NicMSlesions (FT)	NicMSlesions (TL)	0,399

Por fim, ao analisar a Diferença Absoluta de Volume, o teste de Kruskal-Wallis verificou que houve diferença relevante entre os modelos, com $p=0,036$. A seguir, o teste de Dunn identificou diferenças entre dois pares de modelos: FLEXCONN e *New Light Weighted Unet* e FLEXCONN e NicMSlesions (FT-treinamento completo), conforme pode ser visto na Tabela 7.10.

Tabela 7.10: Teste de Dunn para análise de DV.

Modelo 1	Modelo 2	Valor-p
FLEXCONN	New Light Weighted Unet	0,042
FLEXCONN	NicMSlesions (FT)	0,001
FLEXCONN	NicMSlesions (TL)	0,075
New Light Weighted Unet	NicMSlesions (FT)	0,430
New Light Weighted Unet	NicMSlesions (TL)	0,801
NicMSlesions (FT)	NicMSlesions (TL)	0,298

Capítulo 8

Conclusões

O presente capítulo reporta brevemente os objetivos e atividades realizadas ao longo deste trabalho. Em seguida, as principais conclusões e limitações deste estudo são apresentadas. Por fim, sugestões para trabalhos futuros são relatados.

8.1 Conclusões Finais

Esta dissertação teve como objetivo explorar métodos de processamento e análise de imagens de ressonância magnética, especificamente técnicas de pré-processamento das imagens de MRI para posterior segmentação das lesões cerebrais de pacientes diagnosticados com Esclerose Múltipla, através de métodos de *Machine Learning*. A segmentação automática é muito importante como ferramenta para auxiliar o acompanhamento médico, pois reduz o tempo necessário de realizar tais atividades, em relação à segmentação manual, além de diminuir possíveis subjetividades dos observadores.

Diferentes trabalhos de segmentação de lesões escleróticas foram analisados nesta dissertação. Tais métodos envolveram arquiteturas com concatenação (FLEXCONN), em formato de U (*Light Weighted Unet*) e até modelos em cascata (NicMSlesions). O pré-processamento adequado das imagens foi fundamental para que ocorresse segmentações com boa acurácia. Através dos resultados alcançados, também percebeu-se que ao incluir no banco de treino imagens com mesmo domínio de intensidade das imagens de teste, houve melhora significativa na performance dos modelos.

Estatisticamente, ao analisar os coeficientes Dice alcançados, não houve diferença relevante entre os melhores modelos de cada arquitetura, embora o modelo NicMSlesions (FT) tenha obtido maior mediana. Em relação ao Valor Preditivo Positivo, o programa FLEXCONN alcançou os melhores valores, inclusive obteve desempenho superior ao *New Light Weighted Unet* e ao NicMSlesions (TL). Por último, a análise da Diferença Absoluta de Volume indicou que o modelo FLEXCONN apresentou performance inferior ao modelos *New Light Weighted Unet* e NicMSlesions (FT). Diante do exposto, o modelo NicMSlesions (FT) foi considerado o melhor programa para segmentar as lesões escleróticas. Este trabalho indicou a existência de grandes possibilidades

de utilização de técnicas de segmentação automática, a exemplo de *Deep Learning*, em especial *Convolutional Neural Networks*, com a finalidade de identificar lesões escleróticas em imagens médicas.

8.2 Limitações e Trabalhos Futuros

Modelos de *Deep Learning* normalmente requerem muitos dados durante o treinamento, para garantir que os resultados sejam satisfatórios. No caso de aplicações médicas, as imagens de ressonância magnética de pacientes não são simples de obter devido a questões éticas e dificuldade de segmentar manualmente grande quantidade de imagens. Sugere-se que mais estudos com transferência de aprendizado deva ser realizado, para averiguar a possibilidade de ultrapassar tal dificuldade.

Além disso, o volume das lesões é muito pequeno em relação ao volume total da imagem, resultados em casos de *data imbalance*. Como forma de amenizar esse efeito, seria interessante estudar a influência de diferentes funções de perdas, que considerem a probabilidade de localização específica das lesões.

Atualmente, o grupo de pesquisa está a trabalhar na criação de códigos em Python para realizar as etapas de pré-processamento das imagens de ressonância magnética sem a necessidade de softwares externos. Por fim, o grande objetivo de investigar segmentação automática de lesões é a aplicabilidade em clínicas médicas. Diante disso, recomenda-se a realização de novos estudos para avaliar o desempenho dos modelos ao segmentar imagens clínicas sujeitas a ruído e outras interferências.

Referências

- [1] Multiple Sclerosis International Federation, "Atlas of M.S.: Mapping multiple sclerosis around the world. Disponível em: <https://www.msif.org/wp-content/uploads/2020/10/Atlas-3rd-Edition-Epidemiology-report-EN-updated-30-9-20.pdf>. Acesso em 2/outubro/2020.
- [2] R. Pinto. "Segmentação automática de lesões de esclerose múltipla em imagens de ressonância magnética". Dissertação de mestrado. Faculdade de Engenharia, Universidade do Porto, Porto, PT, 2017.
- [3] M.A. McAteer and R. P. Choudhury. "Applications of nanotechnology in molecular imaging of the brain". *Progress in Brain Research*, 180:72–96, 2009.
- [4] K. Rejdak, S. Jackson, and G. Giovannoni. "Multiple sclerosis: A practical overview for clinicians". *British medical bulletin*, 95:79–104, 2010.
- [5] A. Compston and C. Alasdair. "Multiple sclerosis". *Lancet*, 359:1221–31, 2002.
- [6] R. Dobson and G. Giovannoni. "Multiple sclerosis - a review". *European journal of neurology*, 26,1:27–40, 2019.
- [7] S. Hunter. "Overview and diagnosis of Multiple Sclerosis". *The American journal of managed care*, 22:141–150, 2016.
- [8] S. Roy, J. Butman, D. Reich, P. Calabresi, and D. Pham. "Multiple sclerosis lesion segmentation from brain mri via fully convolutional neural networks". *ArXiv Preprint 1803.09172*, 2018.
- [9] A. Carass, J. Cuzzocreo, M.B. Wheeler, P.L. Bazin, S.M. Resnick, and J.L. Prince. "Simple paradigm for extra-cerebral tissue removal: Algorithm and analysis". *NeuroImage*, 56:1982–92, 2011.
- [10] I. Rizwan i Haque and J. Neubert. "Deep learning approaches to biomedical image segmentation". *Informatics in Medicine Unlocked*, 18:100297, 2020.
- [11] A. Danelakis, T. Theoharis, and D. A. Verganelakis. "Survey of automated multiple sclerosis lesion segmentation techniques on magnetic resonance imaging". *Computerized Medical Imaging and Graphics*, 70:83 – 100, 2018.
- [12] Regina Alvarenga, Marcos Alvarenga e Claudia Vasconcelos. Esclerose Múltipla Rio. Disponível em: <http://esclerosemultiplario.com.br/o-que-e-esclerose-multipla/o-que-e-esclerose-multipla/>. Acesso em 16/outubro/2020.

- [13] W. Brownlee, T. Hardy, F. Fazekas, and D. Miller. "Diagnosis of multiple sclerosis: progress and challenges". *The Lancet*, 389:1336–1346, 2016.
- [14] A. Ascherio. "Environmental factors in multiple sclerosis". *Expert Review of Neurotherapeutics*, 13:3–9, 2013.
- [15] B. M. Segal. "Multiple Sclerosis", in *Clinical Immunology*, 2nd ed, Elsevier, 2019, pp.891-902.
- [16] F.D. Lublin, S.C. Reingold, J.A. Cohen, G.R. Cutter, P.S. Sørensen, A.J. Thompson, J.S. Wolinsky, L.J. Balcer, B. Banwell, F. Barkhof, B. Bebo, P.A. Calabresi, M. Clanet, G. Comi, R.J. Fox, M.S. Freedman, A.D. Goodman, M. Inglese, L. Kappos, B.C. Kieseier, J.A. Lincoln, C. Lubetzki, A.E. Miller, X. Montalban, P.W. O'Connor, J. Petkau, C. Pozzilli, R.A. Rudick, M.P. Sormani, O. Stüve, E. Waubant, and C.H. Polman. "Defining the clinical course of Multiple Sclerosis: the 2013 revisions". *Neurology*, 83,3:278–286, 2014.
- [17] C.H. Polman, S.C. Reingold, B. Banwell, M. Clanet, J.A. Cohen, M. Filippi, K. Fujihara, E. Havrdova, M. Hutchinson, L. Kappos, F.D. Lublin, X. Montalban, P. O'Connor, M. Sandberg-Wollheim, A.J. Thompson, E. Waubant, B. Weinshenker, and J.S. Wolinsky. "Diagnostic criteria for multiple sclerosis: 2010 revisions to the mcdonald criteria". *Ann Neurol*, 69:292–302, 2011.
- [18] F. A. G. Cardoso. "Atuação fisioterapêutica na esclerose múltipla forma recorrente-remitente". *Revista Movimenta*, 3(2):69–75, 2010.
- [19] E. van Beek, C. Kuhl, Y. Anzai, P. Desmond, R.L. Ehman, Q. Gong, G. Gold, V. Gulani, M. Hall-Craggs, T. Leiner, C.C. Tschoyoson Lim, J. G. Pipe, S. Reeder, C. Reinhold, M. Smits, D. K. Sodickson, C. Tempny, A. Vargas, and M. Wang. "Value of mri in medicine: More than just another test?". *Journal of Magnetic Resonance Imaging*, 49(7):14–25, 2019.
- [20] I. J. Sanches. "Sobreposição de imagens de termografia e ressonância magnética: uma nova modalidade de imagem médica tridimensional". Dissertação de Ph.D., Universidade Tecnológica Federal do Paraná, Curitiba, PR, 2009.
- [21] A. Paiva and R. Seixas e M. Gattass. "Introdução à Visualização Volumétrica". Pontifícia Universidade Católica do Rio de Janeiro, 1999.
- [22] A. X. Falcão and N. J. Leite. "Fundamentos de processamento de imagem digital". Instituto de Computação, Universidade de Campinas, 2003.
- [23] C. Westbrook. *MRI at Glance*, Wiley-Blackwell, 3rd ed, Cambridge, 2015, pp.12-28.
- [24] A. A. Mazzola. "Ressonância magnética: princípios de formação da imagem e aplicações em imagem funcional". *Revista Brasileira de Física Médica*, 3:117–129., 2009.
- [25] B.M. Dale, M. A. Brown, and R. C. Semelka. *MRI basic principles and applications*, Wiley-Blackwell, 5th ed, West Sussex, 2015, pp. 17-26.
- [26] Andrew Murphy, Matthew Skalski, and Frank Gaillard. "MRI Sequence Parameters". Disponível em: <https://radiopaedia.org/articles/mri-sequence-parameters>. Acesso em 12/janeiro/2021.

- [27] David C. Preston. "MRI Basics". Disponível em: <https://case.edu/med/neurology/NR/MRI%20Basics.htm>. Acesso em 19/dezembro/2020.
- [28] C.Egger, R. Opfer, C. Wang, T. Kepp, M.P. Sormani, L. Spies, M. Barnett, and S. Schippling. "MRI flair lesion segmentation in multiple sclerosis: Does automated segmentation hold up with manual annotation?". *NeuroImage: Clinical*, 13:264 – 270, 2017.
- [29] M. K. Santos, J.R. Ferreira Júnior, D.T. Wada, A.P. Tenório, M.H. Barbosa, and P.M. de Azevedo Marques. "Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine". *Radiologia Brasileira*, 52:387 – 396, 2019.
- [30] P. C. Klein. "Detecção de lesões de esclerose múltipla em imagens de ressonância magnética do tipo fluid attenuated inversion recovery (flair)". Dissertação de mestrado, Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, 2016.
- [31] S. Ira Ktena Martin Rajchl and Nick Pawlowski. "An introduction to biomedical image analysis with tensorflow and dltk". Disponível em: <https://medium.com/tensorflow/an-introduction-to-biomedical-image-analysis-with-tensorflow-and-dltk-2c25304e7c13>. Acesso em 10/outubro/2020.
- [32] G. Gerig, O. Kubler, R. Kikinis, and F. Jolesz. "Nonlinear anisotropic filtering of MRI data". *IEEE Transactions on Medical Imaging*, 11:221–32, 1992.
- [33] M. M. Santos. "Classificação de padrões de imagens: Função objetivo para perceptron multicamada e máquina de aprendizado extremo convolucional". Dissertação de doutorado, Universidade Federal de Pernambuco, Recife, PE, 2019.
- [34] M. M. George, S. Kalaivani, and M.S. Sudhakar. "A non-iterative multi-scale approach for intensity inhomogeneity correction in mri". *Magnetic Resonance Imaging*, 42:43–59, 2017.
- [35] Uro Vovk, Franjo Pernus, and Botjan Likar. "A review of methods for correction of intensity inhomogeneity in mri". *IEEE Transactions on Medical Imaging*, 26(3):405–421, 2007.
- [36] L. Lefkovits, S. Lefkovits, and M. Vaida. "An atlas based performance evaluation of inhomogeneity correcting effects". *5th International Conference on Recent Achievements in Mechatronics, Automation, Computer Science and Robotics (MACRo)*. Tirgu-Mures, Romania, 2015, pp. 79-90.
- [37] M. Shah, Y. Xiao, N. Subbanna, S. Francis, D.L. Arnold, D.L. Collins, and T. Arbel. "Evaluating intensity normalization on mris of human brain with multiple sclerosis". *Medical Image Analysis*, 15(2):267 – 282, 2011.
- [38] M. F. Silva Pinto. "Evaluation and prediction of multiple sclerosis disease progression". Dissertação de mestrado, Universidade de Coimbra, Coimbra, PT, 2018.
- [39] S. Valverde, M. Salem, M. Cabezas, D. Pareto, J.C. Vilanova, L. Ramió-Torrentà, À. Rovira, J. Salvi, A. Oliver, and X. Lladó. "One-shot domain adaptation in Multiple Sclerosis lesion segmentation using convolutional neural networks". *NeuroImage: Clinical*, 21:101638, 2019.

- [40] J. C. Reinhold, B. E. Dewey, A. Carass, and J. L. Prince. "Evaluating the impact of intensity normalization on mr image synthesis". *Proc. SPIE 10949, Medical Imaging 2019: Image Processing*. 2019, pp. 890-898.
- [41] J. Bernal, K. Kushibar, D.S. Asfaw, S. Valverde, A. Oliver, R. Martí, and X. Lladó. "Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review". *Artificial Intelligence in Medicine*, 95:64–81, 2019.
- [42] P. Kalavathi and V. B. Surya Prasath. "Methods on skull stripping of mri head scan images: A review". *Journal of Digital Imaging*, 29:365–79, 2016.
- [43] S. Smith. "Fast robust automated brain extraction". *Human Brain Mapping*, 17:143–55, 2002.
- [44] A.X. Falcão, F.P.G. Bergo, F.O. Favretto, G.C.S. Ruppert, P.A.V. Miranda, and F.A.M. Capabianco. *Neurociências e Epilepsia*. Ed. Plêiade, São Paulo, 2008, pp. 97-124.
- [45] L. G. Brown. "A survey of image registration techniques". *ACM Computing Surveys*, 24(4):325–376, 1992.
- [46] J.V. Manjón. *Imaging Biomarkers*. Springer, Cham, 2017, pp. 53-63.
- [47] X. Zhang, Y. Feng, W. Chen, X. Li, A.V. Faria, Q. Feng, and S. Mori. "Linear registration of brain mri using knowledge-based multiple mediator libraries". *Frontiers in Neuroscience*, 13:909, 2019.
- [48] R. Millionini and E. Iori. E. Murphy P. Tessari S. Sbrignadello, A. Tura. "The inter and intraoperator variability in manual spot segmentation and its effect on spot quantitation in two-dimensional electrophoresis analysis". *Electrophoresis*, 31(10):1739–42, 2010.
- [49] J. Iglesias. "Globally optimal coupled surfaces for semi-automatic segmentation of medical images". *IPMI 2017 Information Processing in Medical Imaging*. Boone, NC, 2017, pp. 610-621.
- [50] A. S. Lundervold and A. Lundervold. "An overview of deep learning in medical imaging focusing on mri". *Zeitschrift für Medizinische Physik*, 29(2):102–127, 2019.
- [51] Data Science Academy. *Deep Learning Book*. Disponível em: <http://www.deeplearningbook.com.br/>. Acesso em 10/dezembro/2020.
- [52] D. García-Lorenzo, S. Francis, S. Narayanan, D.L. Arnold, and D. Louis Collins. "Review of automatic segmentation methods of Multiple Sclerosis white matter lesions on conventional magnetic resonance imaging". *Medical Image Analysis*, 17(1):1–18, 2013.
- [53] C. C. Aggarwal. *Neural Networks and Deep Learning: A Textbook*, Springer, 2018, pp. 315-371.
- [54] S. Cohen. *Artificial Intelligence and Deep Learning in Pathology*, Elsevier, 2021, pp. 13-40.
- [55] J. Ker, L. Wang, J. Rao, and T. Lim. "Deep learning applications in medical image analysis". *IEEE Access*, 6:9375–9389, 2018.
- [56] R. Yamashita, M. Nishio, R. Do, and K. Togashi. "Convolutional neural networks: an overview and application in radiology". *Insights into Imaging*, 9:611–629, 2018.

- [57] E. L. Skeika. "Utilização de redes neurais completamente convolucionais para identificação e medição de crânios fetais". Dissertação de mestrado, Universidade Tecnológica Federal do Paraná, Ponta Grossa, PR, 2019.
- [58] M. C. Cavalcante. "Segmentação de imagens de lesões de pele usando a rede neural convolucional u-net". Dissertação de mestrado, Universidade Federal do Espírito Santo, Vitória, PR, 2019.
- [59] Jones Granatyr. "Deep learning com python de a a z: o curso completo". Udemey course. Disponível em: <https://www.udemy.com/course/deep-learning-com-python-az-curso-completo/>. Acesso em 15/novembro/2020.
- [60] S. Jadon. "A survey of loss functions for semantic segmentation". *2020 IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology*. Fully virtual, 2020.
- [61] C. Q. Silva. "Redes neurais aplicadas no reconhecimento de símbolos matemáticos manuscritos online". Dissertação de mestrado, Universidade de São Paulo, São Paulo, SP, 2019.
- [62] Rafael Meier. "Treinamento de modelos". Grupo de Estudos em Aprendizagem de Máquina (GEAM). Universidade Federal de Santa Catarina. Disponível em: <https://geam.paginas.ufsc.br/files/2020/02/Treinando-modelos.pdf> Acessado em: 28/dezembro/2020.
- [63] Nick Winovich. Deep learning optimization. Disponível em: https://www.math.purdue.edu/~nwinovic/deep_learning_optimization.html. Acesso em 03/janeiro/2021.
- [64] W. S. Oliveira. "Consenso de segmentações de imagens usando classificação de padrões". Dissertação de doutorado, Universidade Federal de Pernambuco, Recife, PE, 2017.
- [65] A. Plaatt. *Learning to Play: Reinforcement Learning and Games*. Springer, 2020, pp 135-194.
- [66] O. Ronneberger, P. Fischer, and T. Brox. "U-net: Convolutional networks for biomedical image segmentation". In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, Cham, 2015, pp 234–241.
- [67] Towards Data Science. "Convolutional neural network". Disponível em: <https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529>. Acesso em 15/outubro/2020.
- [68] S. Valverde, M. Cabezas, E. Roura, S. González-Villà, D. Pareto, J.C. Vilanova, L. Ramió-Torrentà, À. Rovira, A. Oliver, and X. Lladó. "Improving automated multiple sclerosis lesion segmentation with a cascaded 3d convolutional neural network approach". *NeuroImage*, 155:159–168, 2017.
- [69] H. M. Rehan Afzal, S. Luo, S. Ramadan, J. Lechner-Scott, M. Ruhul Amin, and M. Kamran Afzal J. Li. "Automatic and robust segmentation of multiple sclerosis lesions with convolutional neural networks". *Cmc -Tech Science Press-*, 66:977–991, 2020.
- [70] M. Havaei, N. Guizard, N. Chapados, and Y. Bengio. "HeMIS: Hetero-modal image segmentation". In *19th International Conference on Medical Image Computing and Computer Assisted Intervention*. Athenes, 2016.

- [71] A. Kumar, , O.N. Murthy, Shrish, P. Ghosal, A. Mukherjee, and D. Nandi. "A dense u-net architecture for Multiple Sclerosis lesion segmentation". In *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, pages 662–667. Kerala, 2019.
- [72] B. Menze et al. "The multimodal brain tumor image segmentation benchmark (brats)". *IEEE TMI*, 34(10):1993–2024, 2015.
- [73] M. Styner, J. Lee, B. Chin, M.S. Chin, O. Commowick, H. Tran, S. Markovic-Plese, V. Jewells, and S. Warfield. "3D segmentation in the clinic: A grand challenge ii- mss lesion segmentation". *MIDAS Journal*, page 1–6, 2008.
- [74] A. Birenbaum and H. Greenspan. "Multi-view longitudinal cnn for multiple sclerosis lesion segmentation". *Engineering Applications of Artificial Intelligence*, 65:111 – 118, 2017.
- [75] A. Carass et al. Longitudinal multiple sclerosis lesion segmentation: Resource and challenge. *NeuroImage*, 148:77–102, 2017.
- [76] O. Commowick, F. Cervenansky, and R. Ameli. "MSSEG challenge proceedings: Multiple sclerosis lesions segmentation challenge using a data management and processing infrastructure". *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Athenes, Greece, 2016.
- [77] P. Ghosal, P. K. C. Prasad, and D. Nandi. "A light weighted deep learning framework for multiple sclerosis lesion segmentation". In *2019 5th International Conference on Image Information Processing (ICIIP)*, pages 526–531. Shimla, India, 2019.
- [78] H. Zhang, A. M. Valcarcel, R. Bakshi, R. Chu, F. Bagnato, T.R. ShinoharaKilian, and H. Oguz. "Multiple sclerosis lesion segmentation with tiramisu and 2.5d stacked slices". In *22nd International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 338–346. Shenzhen, China, 2019.
- [79] Z. Akkus, A. Galimzianova, A. Hoogi, D. Rubin, and B. Erickson. "Deep learning for brain mri segmentation: State of the art and future directions". *Journal of Digital Imaging*, 30:449–459, 2017.
- [80] S. Grahl, V. Pongratz, P. Schmidt, C. Engl, M. Bussas, A. Radetz, G. Gonzalez-Escamilla, S. Groppa, F. Zipp, C. Lukas, J. Kirschke, C. Zimmer, M. Hoshi, A. Berthele, B. Hemmer, and M. Mühlau. "Evidence for a white matter lesion size threshold to support the diagnosis of relapsing remitting multiple sclerosis". *Multiple Sclerosis and Related Disorders*, 29:124–129, 2019.
- [81] M. Alshayegi, M.A. Al-Rousan, H. Ellethy, and S. Abed. "An efficient Multiple Sclerosis segmentation and detection system using neural networks". *Computers Electrical Engineering*, 71:191–205, 2018.
- [82] Ž. Lesjak, A. Galimzianova, A. Koren, M. Lukin, F. Pernuš, B. Likar, and Ž. Špiclin. "A novel public mr image dataset of multiple sclerosis patients with lesion segmentations based on multi-rater consensus". *Neuroinformatics*, 16(1):51–63, 2018.
- [83] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, JC Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J.V. Miller, S. Pieper, and R. Kikinis. "3D slicer as an image computing platform for the quantitative imaging network". *Magnetic Resonance Imaging*, 30(9):1323–1341, 2012.

- [84] S. Bauer, T. Fejes, and M. Reyes. "A skull-stripping filter for itk". *Insight Journal*, 2013.
- [85] H. Johnson, G. Harris, and K. Williams. "BRAINSFIT: Mutual information registrations of whole-brain 3d images, using the insight toolkit". *Insight Journal*, 2007.
- [86] N. J. Tustison, B. B. Avants, P. A. Cook, A. Egan Y. Zheng, P. A. Yushkevich, and J. C. Gee. "N4ITK: improved n3 bias correction. ieee transactions on medical imaging". *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, 2010.