

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**Segmentação de lesões causadas pela esclerose múltipla
utilizando machine learning em exames radiológicos**

Matheus de Lima Piovani

Trabalho de Conclusão de Curso - MBA em Ciência de Dados (CEMEAI)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Matheus de Lima Piovani

Segmentação de lesões causadas pela esclerose múltipla utilizando machine learning em exames radiológicos

Trabalho de Conclusão de Curso apresentado ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de MBA em Ciências de Dados. *VERSÃO REVISADA*

Área de Concentração: Ciências de Dados, Machine Learning e Radiologia

Orientador: Prof. Dr. Afonso Paiva Neto

USP – São Carlos
Fevereiro de 2022

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

d278s de Lima Piovani, Matheus
Segmentação de lesões causadas pela esclerose
múltiplautilizando machine learning em exames
radiológicos / Matheus de Lima Piovani; orientador
Afonso Paiva Neto. -- São Carlos, .
p.

Trabalho de conclusão de curso (MBA em Ciência
de Dados) -- Instituto de Ciências Matemáticas e de
Computação, Universidade de São Paulo, .

1. Ciência de Dados. 2. Machine Learning. 3.
Radiologia. I. Paiva Neto, Afonso, orient. II.
Título.

Matheus de Lima Piovani

**Segmentation of lesions caused by multiple sclerosis using
machine learning in radiological exams**

Completion of course work submitted to the Institute of
Mathematics and Computer Sciences – ICMC-USP, in
partial fulfillment of the requirements for the degree of
the MBA in Data Science. *FINAL VERSION*

Concentration Area: Data Science, Machine Learning
and Radiology

Advisor: Prof. Dr. Afonso Paiva Neto

USP – São Carlos
February 2022

Dedico esse trabalho a minha noiva Nicole que me deu muito suporte não só no desenvolvimento desse projeto como na vida. Te amo be!

AGRADECIMENTOS

Gostaria de agradecer a Nicole pelo ajuda psicológica e pelas inúmeras revisões no desenvolvimento, aos meus pais pelo suporte e amor que sempre me deram e ao pessoal do discord que esteve fazendo companhia enquanto estava na tensão aguardando resultados de modelos.

Keep It Simple, Stupid
(Kelly Johnson)

RESUMO

PIOVANI, M. L. **Segmentação de lesões causadas pela esclerose múltipla utilizando machine learning em exames radiológicos.** 2022. 58 p. Trabalho de Conclusão de Curso (MBA – Ciência de Dados (CEMEAI)) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

Com o crescimento do poder computacional, capacidade de processamento e quantidade de dados disponíveis a utilização de *machine learning* para identificação de padrões em imagens se mostra muito viável e eficiente. A esclerose múltipla é uma doença neurológica e crônica, que em determinados estágios é caracterizada por períodos de surto, esses podem ser detectados com identificação de atividade inflamatória na substância branca do cérebro através de uma ressonância magnética com contraste. O objetivo deste trabalho consiste em utilizar métodos de *machine learning* a fim de realizar a segmentação de áreas de lesão causadas pela doença de forma rápida e automatizada, os resultados serão utilizados como auxiliares para o aumento da eficácia e agilidade do trabalho de radiologistas no diagnóstico através do exame.

Palavras-chave: Ciências de Dados, *Machine Learning*, Radiologia.

ABSTRACT

PIOVANI, M. L. **Segmentation of lesions caused by multiple sclerosis using machine learning in radiological exams**. 2022. 58 p. Trabalho de Conclusão de Curso (MBA – Ciência de Dados (CEMEAI)) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

With the growth of computational power, processing capacity and amount of available data, the use of machine learning techniques to identify patterns in images proves to be very viable and efficient. Multiple sclerosis is a neurological and chronic disease, which at certain stages is characterized by periods of relapse, which can be detected with identification of inflammatory activity in the white matter of the brain through a contrast-enhanced magnetic resonance. The objective of this work is to use machine learning methods in order to realize the segmentation of areas of lesions caused by the disease in a fast and automated way, the results will be used as an aid to increase the effectiveness and agility of the work of the radiologist responsible for analyzing the exam.

Keywords: Data Science, Machine Learning, Radiology.

LISTA DE TABELAS

Tabela 1 – Lista de Estudos de Segmentação	27
Tabela 2 – Comparação de T1 × T2 × FLAIR (cérebro).	30
Tabela 3 – Comparação de resultados em relação ao BIANCA	50

LISTA DE ABREVIATURAS E SIGLAS

BIANCA	Brain Intensity AbNormality Classification Algorithm
CD	Ciências de Dados
DER	<i>Detection error rate</i>
DP	Desvio Padrão
DTL	Deep Transfer Learning
EM	Esclerose Múltipla
FC	Fluido Cerebrospinal
FLAIR	Recuperação de Inversão Atenuada por Fluido
IA	Inteligência Artificial
K-FOLD	K-Fold Cross Validation
KNN	<i>k - Nearest Neighbors</i>
ML	<i>Machine Learning</i>
OER	<i>Outline error rate</i>
PR	Pulso de Radiofrequência
RFN	Razão de Falso Negativos
RFP	Razão de Falso Positivos
RM	Ressonância Magnética
RNC	Redes Neurais Convolucionais
SI	<i>Dice Similarity Index</i>
SMM-SC	Distribuição de Modelo de Mistura de t-Student Espacialmente Restrito
SW	<i>Spatial Weighting</i>

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Estrutura do Trabalho	22
2	REVISÃO BIBLIOGRÁFICA	23
2.1	Considerações iniciais	23
2.2	Segmentação e Classificação de Lesões Cerebrais	24
3	METODOLOGIA	29
3.1	Imagens por Ressonância Magnética	29
3.2	<i>Features</i>	30
3.3	Algoritmo KNN	32
3.4	Método BIANCA	33
3.4.1	<i>Parâmetros e Configurações</i>	34
3.4.2	<i>Métricas</i>	36
4	DESENVOLVIMENTO	37
4.1	Considerações Iniciais	37
4.2	Base utilizada	38
4.3	Atividades Realizadas	38
4.3.1	<i>Pré-Processamento</i>	39
4.3.2	<i>Cálculo das Features</i>	40
4.3.3	<i>Aplicação do algoritmo</i>	44
5	RESULTADOS	49
5.1	Análise Quantitativa	49
5.2	Análise Qualitativa	50
5.3	Dificuldades e Limitações	53
6	CONCLUSÃO	55
	REFERÊNCIAS	57

INTRODUÇÃO

Os estudos em Ciências de Dados (CD) evoluíram muito nos últimos anos, com o crescimento do poder computacional e quantidade de dados disponíveis as aplicações de suas técnicas se tornaram cada vez mais viáveis. Tecnologias de CD vêm sendo aplicadas em diversas áreas, entre elas a medicina no setor de diagnóstico por imagens, atuando como auxiliares ao profissional de saúde, tornando os diagnósticos mais rápidos e eficientes.

A Esclerose Múltipla (EM) é uma doença desmielinizante, neurológica e crônica, que causa lesões no sistema nervoso central. Em determinados estágios ela é caracterizada por períodos de surto, o acompanhamento de pacientes diagnosticados com EM se dá pela realização periódica de exames de Ressonância Magnética (RM) com contraste, nelas um radiologista pode detectar lesões e possível atividade inflamatória na substância branca do cérebro.

Entre as tecnologias de CD o *Machine Learning* (ML) é um dos conhecimentos de Inteligência Artificial (IA) que possui grande potencial quando utilizado no reconhecimento e classificação de padrões em imagens. O ML consiste no estudo de algoritmos de computador capazes de se aprimorar automaticamente, através de um conjunto de dados de treinamento. Podemos utilizar um modelo treinado para realizar previsões ou tomar decisões em relação a uma nova amostragem.

Entre os métodos de ML temos o algoritmo *k - Nearest Neighbors* (KNN), que categoriza uma nova amostra de interesse buscando o quão similar o seu vetor de característica é com os dos elementos de seu conjunto de treinamento.

As aplicações e os impactos de ML na radiologia são apontados em (CHOY *et al.*, 2018). Já existem estudos para sua utilização especificamente para EM como em (JUNIOR, 2019), (MCKINLEY *et al.*, 2020) e (SALEM *et al.*, 2018), mostrando que a sua aplicação deve gerar grandes impactos na forma como as análises de RM são realizadas hoje.

Atualmente o diagnóstico gerado pelos radiologistas é realizado de forma manual, esse

processo consome bastante tempo e há um problema em padronizar os resultados, dado que diferentes radiologistas podem realizar diferentes diagnósticos para um mesmo exame de RM. O objetivo desse trabalho é auxiliar na precisão e velocidade do diagnóstico, utilizando o algoritmo KNN para realizar a segmentação em áreas de lesões, saudáveis e de incerteza. As áreas classificadas como incerteza representam as possíveis divergências em diagnósticos e serão os locais nos quais será necessário uma maior atenção dos radiologista no diagnóstico.

1.1 Estrutura do Trabalho

O presente trabalho está estruturado da seguinte forma: o [Capítulo 2](#) apresenta trabalhos relacionados ao tema. No [Capítulo 3](#) é apresentado os conceitos básicos e o referencial teórico do trabalho. O [Capítulo 4](#) a metodologia desenvolvida e os seus aspectos computacionais. O [Capítulo 5](#) contém os resultados obtidos. Finalmente, no [Capítulo 6](#) é apresentada a conclusão e sugestões para trabalhos futuros.

REVISÃO BIBLIOGRÁFICA

Essa capítulo contém uma revisão bibliográfica com referenciais de pesquisa sobre a identificação de lesões na substância branca do cérebro. Serão apresentadas as pesquisas recentes que utilizam métodos computacionais a fim de auxiliar no diagnóstico realizado por radiologistas.

2.1 Considerações iniciais

Atualmente as diretrizes para diagnóstico e acompanhamento da EM citam como o protocolo padrão a utilização de RM com contraste por gadolínio. O resultado da RM dessa forma apresenta os melhores resultados para revelar informações de atividade inflamatória, do surgimento de novas lesões ou do crescimento de lesões pré-existentes ([TRABOULSEE *et al.*, 2016](#); [WATTJES *et al.*, 2015](#)).

As diretrizes existentes hoje surgem para guiar a forma de diagnóstico e apontam como um dos problemas futuros a necessidade de implementar padrões de aquisição e interpretação da RM de pacientes com EM. Dessa forma podemos propor a utilização de técnicas computacionais para auxiliar na padronização da metodologia e interpretação dos resultados de RM.

As lesões causadas pela EM no encéfalo ocorrem na substância branca e podem se espalhar por toda a área. Por esse motivo e pela natureza da inflamação as suas formas e tamanhos são variadas, dificultando assim a sua identificação. Por causa do uso prolongado de gadolínio ao qual os pacientes com EM são expostos, também ocorre um acúmulo da substância no organismo, dificultando dessa forma a diferenciação de novas lesões ([FREIRE, 2019](#)).

A fim de realizar a identificação das lesões e atividades inflamatórias causadas pela EM será realizada a segmentação que tem como objetivo obter, através da imagem de RM, quais são as áreas de interesse, ou seja, locais com anormalidades, sendo assim identificadas como lesões.

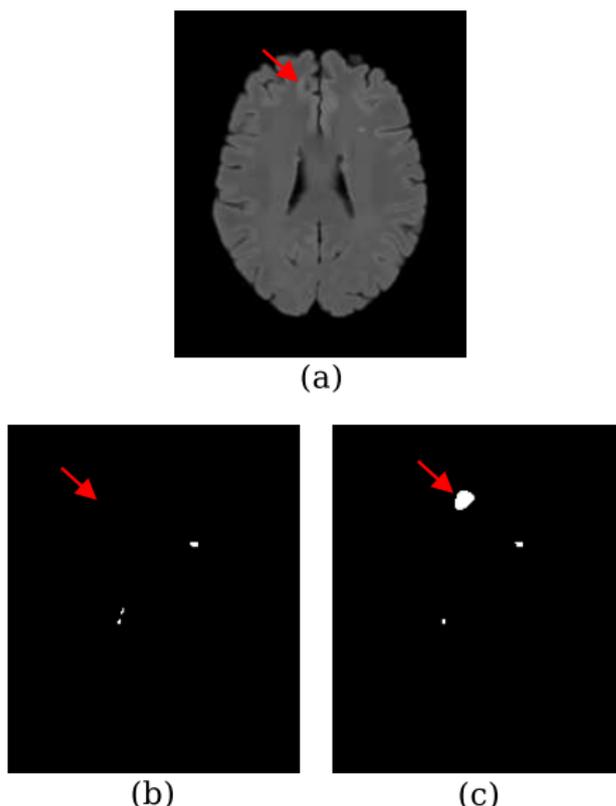
No estudo de Freire além da segmentação foi realizada uma classificação da atividade de EM, para isso utilizou um do classificador de texturas GLCM e RLM pela sua simplicidade

e extensibilidade para três dimensões. Ele utiliza as informações espaciais de sua vizinhança são utilizadas para realizar a classificação entre lesões ativas e não ativas. Os resultados obtidos foram bem promissores, de 157 lesões identificadas por especialistas 154 foram identificadas corretamente e sem a utilização da informação do gadolínio. As três lesões não identificadas foram todas do mesmo paciente e todas possuíam como característica tamanho e intensidade reduzidos (FREIRE, 2019, p. 57).

2.2 Segmentação e Classificação de Lesões Cerebrais

A identificação de regiões durante a segmentação possui uma alta variação quando comparado com as análises de especialistas, porém quando comparamos análises dos próprios especialistas identificamos grande variação entre eles como observado em [Figura 1](#). A [Figura 1.\(a\)](#) mostra a imagem da RM original, enquanto que na [Figura 1.\(b\)](#) e [Figura 1.\(c\)](#) podemos ver duas análises diferentes realizadas por dois especialistas distintos.

Figura 1 – Exemplo de discordância entre especialistas.



Fonte: [Freire \(2019\)](#).

[Freire \(2019\)](#) realizou um estudo a fim de realizar a segmentação utilizando Distribuição t-Student e a classificação de lesões de EM com o classificador XGBoost. Os resultados obtidos foram muito promissores e realizaram a classificação de lesões sem a necessidade da utilização do

contraste por gadolínio, resolvendo assim problemas como o de acúmulo da substância no corpo ao longo da utilização e também auxiliando no diagnóstico para pacientes que não podem fazer o uso da substância como contraste. No entanto os resultados em relação a fase de segmentação apresentaram uma alta taxa de falsos positivos. Segundo ele os estudos sobre classificação e segmentação de lesões causadas pela EM são considerados problemas desafiadores e estão abertos para novas abordagens no campo de processamento de imagens médicas.

A utilização do modelo de mistura t-Student para realizar a segmentação é justificada por (FREIRE, 2019, p. 57) pela distribuição possuir uma cauda pesada intrínseca, sendo assim menos suscetível ao efeito de amostras que estão muito longe dos centroides no espaço das características. Também cita que pela distribuição de t-Student tender a uma distribuição Gaussiana quando v tende a infinito, fica fácil de generalizá-la e comparar ambas as distribuições se necessário. Pesquisas utilizando t-Student também foram feitas em (NGUYEN; WU, 2012) para segmentação de tecidos do cérebro com imagens de MR utilizando t-Student, os autores também incorporaram informações espaciais usando distribuição de Dirichlet, essa abordagem foi chamada de Distribuição de Modelo de Mistura de t-Student Espacialmente Restrito (SMM-SC).

No estudo de Roura *et al.* (2015) a segmentação de lesões de EM foi realizada em duas etapas, primeiro segmentando a imagem em três clusters com o algoritmo SPM8/12 e utilizando o cluster de GMGrai Matter obtido para selecionar os voxels considerados outliers como candidatos a lesões de EM. Então na segunda etapa é realizado um pós processamento a fim de remover falsos positivos, melhorando a acurácia do método.

A ferramenta Brain Intensity AbNormality Classification Algorithm (BIANCA) (GRIF-FANTI *et al.*, 2016) foi criada como uma ferramenta para automatizar a segmentação de hiperintensidades na matéria branca do cérebro, a proposta é que ela seja flexível, robusta, computacionalmente enxuta e disponível de graça. O algoritmo utilizado pela ferramenta é o de KNN, eles combinam as informações das diversas seções da MR com pesos a fim de obter uma quantidade confiável de características.

Diniz *et al.* (2018) também realiza um estudo de métodos computacionais para automatizar a detecção de regiões lesionadas na matéria branca em RM, para obter os resultados são utilizadas técnicas de processamento de imagem e reconhecimento de padrões através de Redes Neurais Convolucionais (RNC), o método obteve acurácia de 98.73%.

Uma acurácia de 100% na classificação de abnormalidades no cérebro com uma base de 613 RM foi obtida em (TALO *et al.*, 2019) utilizando Deep Transfer Learning (DTL) que faz parte do conhecimento de RNC. O estudo também tem como proposta a classificação automática, sem a utilização de características obtidas de forma manual. Foi obtido um baixo custo de computação dos parâmetros, eliminando o longo processo de validação necessário.

Recentemente Junior (2019) e McKinley *et al.* (2020) utilizaram RNC para segmentação de lesões causadas pela EM, com objetivo de realizar o acompanhamento da doença através

da análise de pacientes com a mesmas. Os resultados foram muito positivos indicando que a utilização de RNC como método se mostra como um dos mais eficientes métodos atuais.

Em 2015 houve uma competição para segmentação de lesões de EM e um artigo foi publicado posteriormente (CARASS *et al.*, 2017), essa competição disponibilizou uma database com 21 data sets para treinamento e 61 data sets de teste e contava com um sistema automatizado para submissão e avaliação dos resultados dos concorrentes. O site da competição ainda está disponível em <https://smart-stats-tools.org/lesion-challenge> onde ainda é possível acessar as databases e enviar os resultados para avaliação no sistema, até o presente momento (2021) o site continua recebendo submissões.

A Tabela 1 sumariza os estudos de CD para análise de RM citados nessa revisão, nela mostramos os artigos, os algoritmos utilizados e se o estudo realizado foi específico para EM. O objetivo desse trabalho será realizar a segmentação de lesões utilizando KNN assim como realizado por (GRIFFANTI *et al.*, 2016) no desenvolvimento do BIANCA, a fim obter melhores resultados novas *features* serão utilizadas e também será considerada a área de incerteza que foi demonstrada em Figura 1.

Tabela 1 – Lista de Estudos de Segmentação

Referência	Algoritmo	EM
(MCKINLEY <i>et al.</i> , 2020)	RNC	✓
(FREIRE, 2019)	t-Student	✓
(JUNIOR, 2019)	RNC	✓
(TALO <i>et al.</i> , 2019)	DLT(RNC)	✗
(DINIZ <i>et al.</i> , 2018)	RNC	✗
(GRIFFANTI <i>et al.</i> , 2016)	KNN	✗
(ROURA <i>et al.</i> , 2015)	SPM8/12	✓
(NGUYEN; WU, 2012)	SMM-SC	✗

METODOLOGIA

Essa seção aborda com mais detalhes os principais pontos necessários para o desenvolvimento do estudo. Será explicado o funcionamento e o formato das imagens obtidas por RM, assim como as *features* a serem utilizadas nas análises, mostraremos a definição do algoritmo KNN e por fim serão apresentadas as definições da ferramenta BIANCA ([GRIFFANTI et al., 2016](#)).

3.1 Imagens por Ressonância Magnética

O tópico apresenta a base necessária para a interpretação da RM, o funcionamento da RM será apresentado sem se aprofundar muito na física por trás, então definiremos as sequências de imagens obtidas pelas RM e também a sua interpretação, focando em tecidos cerebrais.

A RM é uma das mais comuns técnicas de imagem médica utilizadas, ela gera imagens tridimensionais de tecidos nervosos e moles do corpo, permitindo assim aos médicos analisarem e realizarem o diagnóstico de má formações, áreas lesionadas ou com inflamações ([MAGNETIC... , 2021](#)), nosso foco aqui será na identificação dessas duas últimas.

O funcionamento da RM se dá explorando a interação dos campos magnéticos, íons de hidrogênio e Pulso de Radiofrequência (PR). Quando colocamos um paciente em um campo magnético forte, os íons de hidrogênio se alinham na direção do campo magnético. Aplicar um PR deve mudar a direção de alinhamento desses íons de hidrogênio. Quando o PR é desligado, os íons tentam se realinhar com o campo magnético novamente, liberando um sinal. A força desse sinal depende do tipo de tecido (gordura, músculo, água) que o íons está. Usando esses princípios, você pode ajustar a máquina para detectar sinais de faixas variadas e variando os planos de magnetização – disso vem os “imagens ponderadas em” ([HIRD, 2021](#)).

Em exames de RM do cérebro as imagens ponderadas podem ser classificadas em T1, T2 e Recuperação de Inversão Atenuada por Fluido (FLAIR) são as mais utilizadas. Imagens do tipo

T1 mostram melhor a anatomia do cérebro, elas detectam a estrutura de gordura, mostrando o Fluido Cerebrospinal (FC) e os ossos em preto, a matéria cinza como cinza e a matéria branca o tecido adiposo como brancos. Imagens T2 apresentam melhor as estruturas com maior quantidade de água, servindo melhor para detectar alterações patológicas no tecido neural, ela mostra o FC e o tecido adiposo como brancos, matéria cinza como cinza escuro, a matéria branca em branco e os ossos em preto.(VASKOVIĆ, 2021) A sequência FLAIR é semelhante a T2 exceto que os tempos de exposição e repetição dos PR são bem mais longos, com isso anormalidades ficam em branco mas o FC é atenuado e aparece em preto, ela é mais sensível a patologias, facilitando a identificação das anormalidades (MAGNETIC..., 2021).

O artigo (MAGNETIC..., 2021) apresenta uma Tabela 2 interessante que facilita a compreensão e interpretação das imagens de RM do cérebro, sendo o córtex constituído de matéria cinza. Por ela podemos verificar que as imagens FLAIR identificam somente as inflamações como brilhantes, facilitando a identificação, o cruzamento de informações dela com a T2 melhora a análise, o BIANCA (GRIFFANTI *et al.*, 2016) consegue realizar segmentações utilizando somente essas duas sequencias de imagens, sendo a T1 opcional.

Tabela 2 – Comparação de T1 × T2 × FLAIR (cérebro).

Tecido	T1	T2	FLAIR
FC	Escuro	Brilhante	Escuro
Matéria Branca	Claro	Cinza Escuro	Cinza Escuro
Córtex	Cinza	Cinza Claro	Cinza Claro
Gordura	Brilhante	Claro	Claro
Inflamação	Escuro	Brilhante	Brilhante

Fonte: Adaptada de Magnetic... (2021).

3.2 Features

As *features* são as características dos dados, utilizadas para treinar os modelos de IA, nesse estudo os nossos dados serão as imagens de RM. Cada sequência da RM gera uma imagem tridimensional, cada uma será pré processadas individualmente, a fim de gerar as suas próprias *features*. Como as RMs geram imagens tridimensionais iremos gerar as *features* para cada um de seus voxels, que são o equivalente ao pixels para imagens volumétricas.

Para as imagens serão geradas *features* de dois tipos, sendo elas espaciais e de dispersão. Como descritor espacial será utilizada a própria posição do voxel, a sua utilização se justifica, como citado em (GRIFFANTI *et al.*, 2016), pelo fato que algumas áreas do cérebro são mais propensas que as outras a apresentarem inflamações, enquanto em outras elas não devem ocorrer, para utilizar essa informação é necessário garantir que as imagens estão padronizadas de forma que as mesmas regiões do cérebro estejam apresentadas nas mesmas áreas para todas as imagens utilizadas.

Features de textura serão geradas para cada voxel, elas descrevem o comportamento de cada um deles buscando representar informações como a intensidade das suas cores ou de sua vizinhança e como é a sua morfologia. Para as características de intensidade utilizaremos o descritor de intensidade local média.

O descritor de textura local utilizado consiste em para cada voxel calcular a média de intensidade da sua vizinhança de tamanho N , para isso será gerado um cubo de aresta N centrado no próprio voxel, o valor da cor de todos os voxels do cubo será somado e então dividido por N^3 . Como nos voxels de borda não é possível inserir o cubo eles serão ignorados no processamento, esse fator não deve impactar nos resultados pois as bordas das imagens de RM devem consistir de uma área preta indicando que estão vazias. A [Figura 2](#) exemplifica o cálculo considerando uma imagem 2D, utilizando a vizinhança de tamanho 3×3 .

Figura 2 – Exemplo de cálculo da intensidade local 2D em uma vizinhança 3×3 .

0	0	4	2	2					
1	2	10	4	1		2,44	3,44	3,89	
3	1	1	7	4		2,89	3,44	3,78	
4	2	2	2	3		1,89	2,56	3,00	
1	2	1	5	2					

Fonte: Elaborada pelo autor.

Como descritor de dispersão iremos utilizar o Desvio Padrão (DP), que indica a dispersão dos dados em torno da média amostral. Valores baixos indicam que os valores estão próximos da média, enquanto valores altos indicam que os valores estão bem distantes da média, indicando dessa forma que temos um comportamento não muito uniforme na região. O cálculo do DP σ é dado pela seguinte fórmula:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2},$$

onde x_i representa os elementos, μ é a média amostral e N é o número total de elementos. As *features* de DP serão calculadas para blocos da mesma forma que para intensidade local, [Figura 3](#) mostra como é feito o cálculo do DP.

Figura 3 – Exemplo de cálculo de DP em uma imagem 2D em uma vizinhança 3×3 .

0	0	4	2	2					
1	2	10	4	1		2,95	3,06	2,81	
3	1	1	7	4		2,69	2,91	2,82	
4	2	2	2	3		0,99	1,95	1,89	
1	2	1	5	2					

Fonte: Elaborada pelo autor.

3.3 Algoritmo KNN

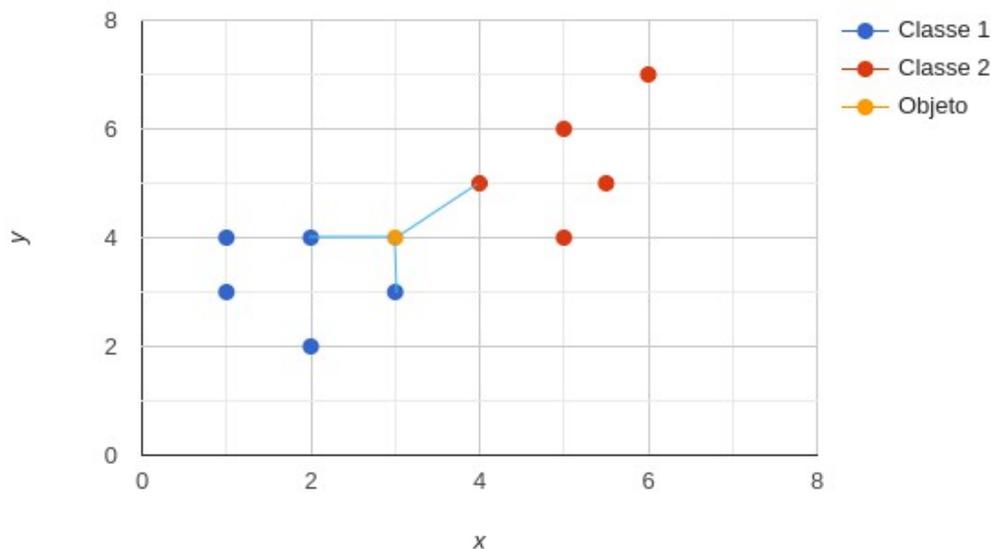
O algoritmo KNN consiste em calcular quais os objetos são mais semelhantes baseado na menor distância entre seus atributos, é um algoritmo de classificação não paramétrico que pode ser utilizado para classificação ou regressão. No estudo utilizaremos o algoritmo para classificação, dessa forma o saída do método é a uma classe a qual o objeto de estudo deve pertencer.

Para medir o quão próximo são os objetos podemos usar medidas de proximidade de dois tipos, sendo elas de similaridade e dissimilaridade. Nas medida de similaridade quanto maior o número melhor, os objetos possuem mais em comum, sendo 1 o valor máximo, indicando que os objetos são iguais. Já nas medidas de dissimilaridade objetos mais semelhantes devem possuir um valor menor, mais próximo de zero.

O funcionamento do KNN segue os seguintes passos:

1. Identificar os k-vizinhos mais próximos do vetor de *features* do objeto que queremos classificar.
2. Determinar o número de vizinhos de cada classe.
3. Classificar o objeto como pertencente à classe com maior número de vizinhos.

Figura 4 – Exemplo do algoritmo KNN.



Fonte: Elaborada pelo autor.

A Figura 4 mostra objetos de Classe 1 em azul, Classe 2 em vermelho e o objeto a ser classificado em amarelo. Utilizando o método KNN com $k = 3$ no primeiro passo identificamos os três vizinhos mais próximos através da distância euclidiana, na figura eles são indicados pela linha azul claro, o passo seguinte é contar o número de vizinhos de cada classe, no caso dois elementos da Classe 1 e um elemento da Classe 2. Como no final tivemos a classe um com maior número de vizinhos mais próximo então classificamos o objeto como Classe 1

Um ponto a ser destacado no KNN é que pelo fato de calcular as distâncias é interessante normalizar os valores já que determinadas *features* possuem escalas diferentes e com isso as de maior escala possuem um peso muito maior na classificação. Vale também destacar que o algoritmo em si apenas memoriza os objetos do conjunto de treinamento, o cálculo para classificação do objeto é realizado na própria fase de classificação.

3.4 Método BIANCA

A ferramenta BIANCA será utilizada como base e referência para o desenvolvimento, ela é capaz de detectar alterações na massa branca utilizando o algoritmo KNN. Nela é possível passar diferentes parâmetros que serão utilizados para decidir quais faixa das RMs serão utilizadas, como as *features* que serão calculadas e também o peso. A seguir serão descritos o funcionamento dos parâmetros da ferramenta e os cálculos das métricas que foram utilizadas para avaliação dos resultados obtidos pela ferramenta.

3.4.1 Parâmetros e Configurações

A ferramenta possui diversas configurações disponíveis, esses parâmetros são brevemente apresentados na [Figura 5](#) retirada do artigo original (GRIFFANTI *et al.*, 2016), eles serão traduzidos e melhor detalhados abaixo.

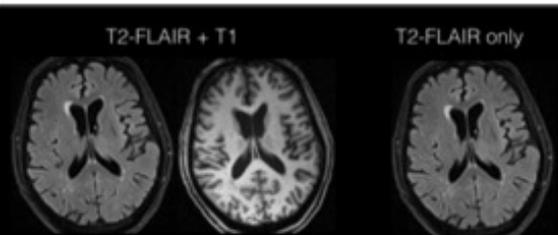
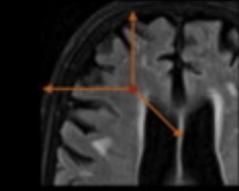
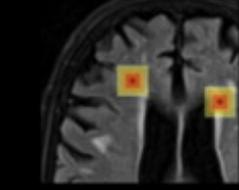
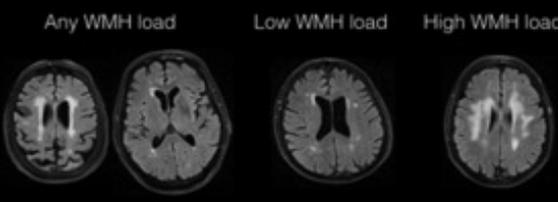
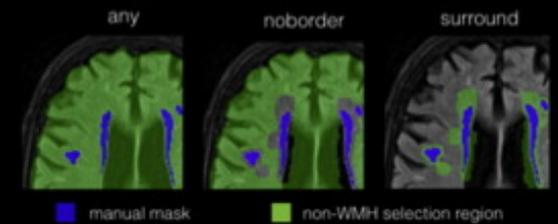
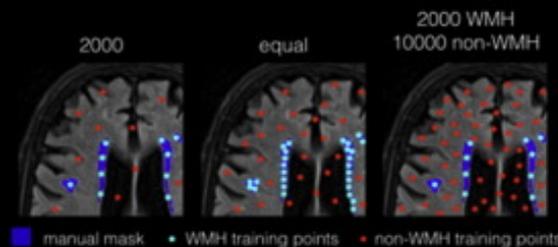
A primeira das opções da ferramenta são as RMs que a serem utilizadas para gerar as *features* do modelo, podendo ser qualquer combinação de T1, T2 e FLAIR, todas as RMs possuem a sua variância normalizada. A segunda configuração disponível é o peso espacial que consiste em utilizar as próprias coordenadas dos voxels como *feature* e atribuir um peso chamado de *Spatial Weighting* (SW), dessa forma quanto maior o SW passado maior o impacto da posição dos voxels nos cálculos do KNN. A opção *patch* é passada como parâmetro *D*, o valor de *D* é utilizado como tamanho da aresta do voxel a ser utilizado para cálculo da *feature* de intensidade local média do voxel.

Os pacientes a serem utilizados no treinamento podem ser definidos de acordo com a quantidade de alterações na matéria branca, podendo ser “qualquer quantidade de alterações”, “poucas alterações” e “muitas alterações”. A localização dos pontos de matéria branca sem alteração a serem utilizados no treinamento é passada para a ferramenta das seguintes formas:

- Todos os pontos da massa cerebral;
- Excluir bordas das regiões de lesão;
- Somente as bordas das lesões.

A última configuração passada é o número máximo de pontos de treinamento a serem utilizados. Como as RMs são de imagens tridimensionais, valores altos dessa configuração podem impactar muito no tempo de processamento necessário para a execução do algoritmo.

Figura 5 – Opções do método BIANCA.

BIANCA OPTION	Description	Values/options tested (schematic representation)
MRI modality	Intensity features	
Spatial weighting (sw)	Spatial features (x,y,z)	 <p>• training point of spatial coordinates x,y,z</p> <p>The higher the value of spatial weighting, the more spatially local is the neighbor selection. Values tested were: sw = 1, 5 and 10.</p>
Patch (D)	Additional intensity features	 <p>• training point</p> <p>Patch for local averaging, with kernel of size D (D=3, D=6, D=9)</p>
Subjects included in the training set (WMH load)	Which subjects to use for the training dataset in terms of the amount of WMHs	
Location of training points	Where to select the non-WMH voxels	 <p>■ manual mask ■ non-WMH selection region</p>
Number of training points	(Maximum) number of training points to use	 <p>■ manual mask ● WMH training points ● non-WMH training points</p>

Fonte: Griffanti *et al.* (2016).

3.4.2 Métricas

Os resultados finais serão comparados com o da ferramenta, por isso utilizaremos as mesmas métricas para a análise, a seguir traduzimos e apresentamos as definições de como elas são calculadas:

- *Dice Similarity Index* (SI): calculado como $2 * (\text{número de voxels de lesão verdadeiro positivos}) / (\text{número de voxels de lesão} + \text{voxels positivos})$.
- Razão de Falso Positivos (RFP): número de voxels classificados incorretamente como lesões dividido pelo número total de lesões realmente existentes.
- Razão de Falso Negativos (RFN): número de voxels classificados incorretamente como saudáveis dividido pelo número total de lesões realmente existentes.
- *Detection error rate* (DER): a detecção de erro é a soma de voxels de lesão pertencentes aos falsos negativos ou falsos positivos. Para obter o DER dividimos a detecção de erro pela média da área total, calculado como a média do volume das lesões preditas pelo modelo e da máscara original.
- *Outline error rate* (OER): o erro de contorno é obtido pela soma dos voxels pertencentes aos verdadeiros positivos. O OER é dado pelo erro de contorno dividido pela média do volume das lesões preditas pelo modelo e da máscara original, de forma semelhante ao DER.

DESENVOLVIMENTO

Nesse capítulo serão descritos como as definições e métodos apresentados nos capítulos anteriores serão utilizados no desenvolvimento do projeto a fim de realizar a classificação de seções do cérebro em áreas de lesão, saudáveis ou áreas de incerteza. O início do capítulo apresenta uma breve descrição da forma e objetivo do desenvolvimento e continua com a apresentação da base a ser utilizada. O capítulo segue com a descrição das atividades realizadas, mostrando como os dados serão pré processados, a fim realizar o cálculo das *features*. Ao final será apresentada a seleção das *features* com melhores estimativas para serem utilizadas na configuração do modelo para utilização no KNN.

4.1 Considerações Iniciais

A seção 3.4 descreve o funcionamento do BIANCA, mostrando as *features* utilizadas e os seus parâmetros de funcionamento, vide figura 5. Na versão do BIANCA a ser implementada no desenvolvimento não serão utilizados os três últimos parâmetros, dessa forma utilizaremos todas as RMs dos pacientes disponíveis na base de treino, sem limitação do número de pontos de lesão no treino, com isso o tempo de processamento deve crescer drasticamente, porém o objetivo com isso é utilizar o máximo de informação disponível nas imagens no modelo a fim de melhorar os resultados obtidos. Para utilizar os resultados do modelo implementado pelo BIANCA será mantida a proporção de 5:1 na quantidade de pontos saudáveis em relação ao número de lesões.

Na figura 1 mostramos que existem discrepâncias entre os diagnósticos realizados por diferentes radiologistas, por isso iremos cobrir áreas de incerteza na classificação, assim teremos ao final do desenvolvimento a classificação em pontos saudáveis, áreas de consenso de lesão e áreas de incerteza que devem ser analisadas com maior atenção.

Outra diferença em relação ao BIANCA será a utilização do DP como *feature* de

dispersão, ele será calculado para os mesmos *patches* utilizados para cálculo da *feature* de intensidade.

4.2 Base utilizada

Para o desenvolvimento será utilizada a base de treino disponibilizada em <https://smart-stats-tools.org/lesion-challenge-2015>, essa base é a utilizada na competição de (CARASS *et al.*, 2017), a base de teste da competição não será utilizada pois as máscaras com diagnóstico realizado pelos especialistas não são disponibilizadas, a avaliação do método em relação a base de teste só é possível ao submeter para análise no site da competição.

A base de treinamento possui RMs de cinco pacientes, sendo que cada um possui um certo número de RMs realizadas em intervalos de aproximadamente um ano, totalizando 21 RMs. Para todas as RMs a base possui os dados nas formas originais, pré processadas e resultados, para os estudos as RMs nos formatos originais serão desconsideradas. Dois resultados diferentes estão disponíveis para cada imagem, cada um deles é referente ao diagnóstico realizado por um radiologista, isso se deve ao fato observado na [Figura 1](#) de que existe discrepância nos resultados. As RMs utilizadas serão separadas em teste e treinamento seguindo uma proporção semelhante ao da competição, dessa forma 5 RMs irão compor a base de treinamento e as outras 16 a base de teste. Todas as RMs estão disponíveis ponderadas em T1, T2 e FLAIR, sempre do período pré aplicação do contraste, dificultando assim a identificação das lesões.

As imagens disponibilizadas estão no formato NIfTI com a extensão *.nii.gz*, para ler esses arquivos será utilizada a biblioteca de Python, `NiBabel`¹, que possui funções para a leitura, escrita e utilitários para auxiliar no desenvolvimento.

4.3 Atividades Realizadas

Para separação do processo de desenvolvimento as atividades serão divididas em 3 etapas:

1. Pré-Processamento;
2. Cálculo das *Features*;
3. Aplicação do algoritmo KNN.

A saída de cada uma das etapas será utilizada como entrada para a próxima, por isso durante o processo de desenvolvimento serão gerados arquivos com a saída de cada etapa.

¹ <https://nipy.org/nibabel/>

4.3.1 Pré-Processamento

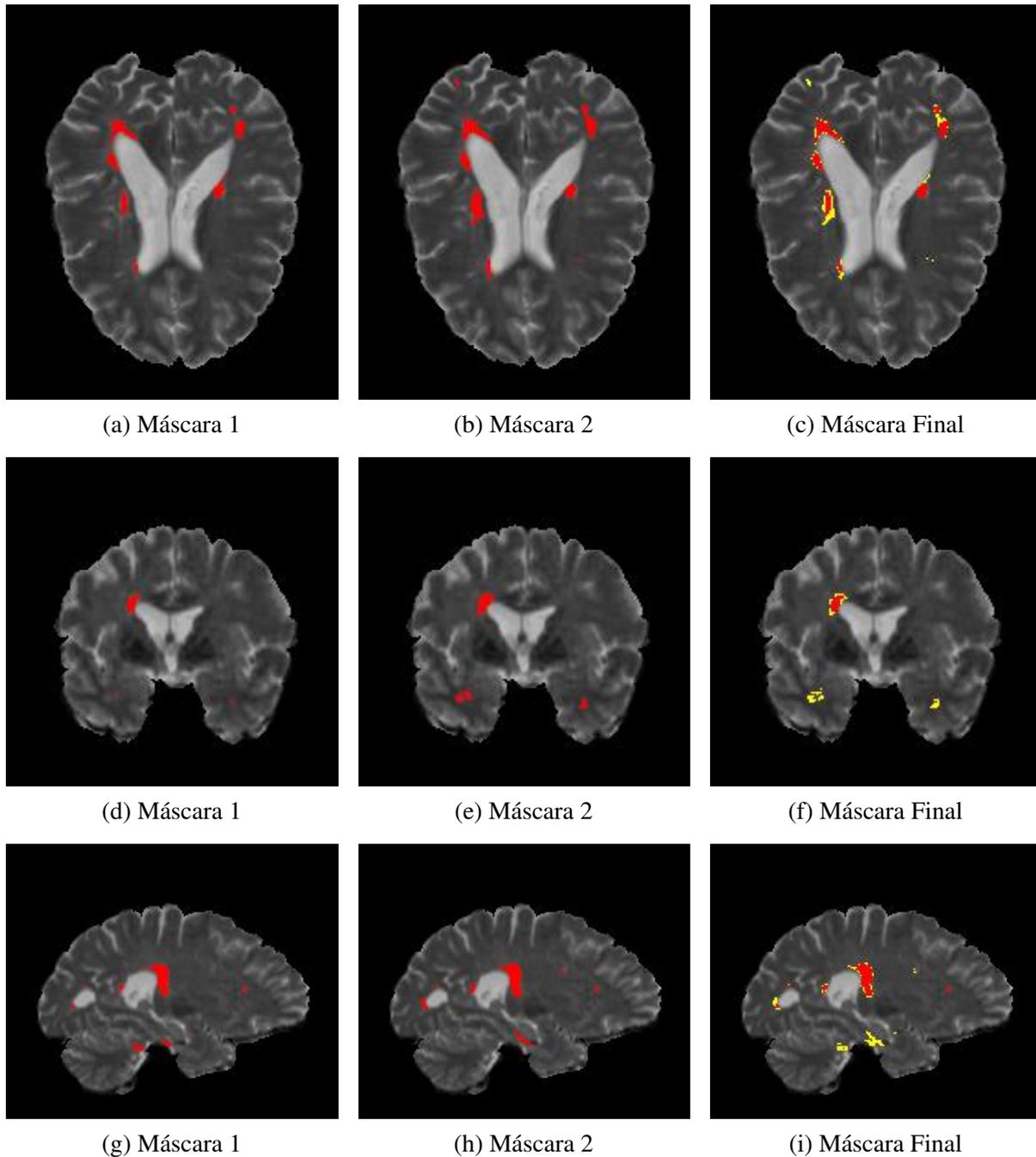
A base utilizada já possui as imagens de RM pré processadas, com remoção de ruídos e padronização espacial, como essa parte foge do escopo do estudo essas imagens serão utilizadas no desenvolvimento.

Como a base possui duas máscaras de resultado com o diagnóstico de dois especialistas para cada RM, ambas serão utilizados para definição da área de incerteza, com isso as novas classes para a nossa base de treinamento serão da seguinte forma:

- **Saudável:** pontos que não foram definidos como área de lesão por nenhum dos radiologistas;
- **Incerteza:** pontos que foram definidos como lesão por somente um dos radiologistas;
- **Lesão:** pontos que foram definidos como lesão por ambos os radiologistas.

Na [Figura 6](#) demonstramos as áreas de incerteza calculadas para algumas seções das RMs da base utilizada, cada linha da imagem representa uma seção, para cada linha a primeira e segunda imagem representam os diagnósticos realizados pelos especialistas 1 e 2, respectivamente, em ambas as imagens a área vermelha representa o que foi marcado como lesão. A terceira imagem de cada linha representa a máscara final, em amarelo mostramos a área de incerteza e em vermelho somente a área de consenso de lesão.

Figura 6 – Áreas de incerteza.



Fonte: Elaborada pelo autor.

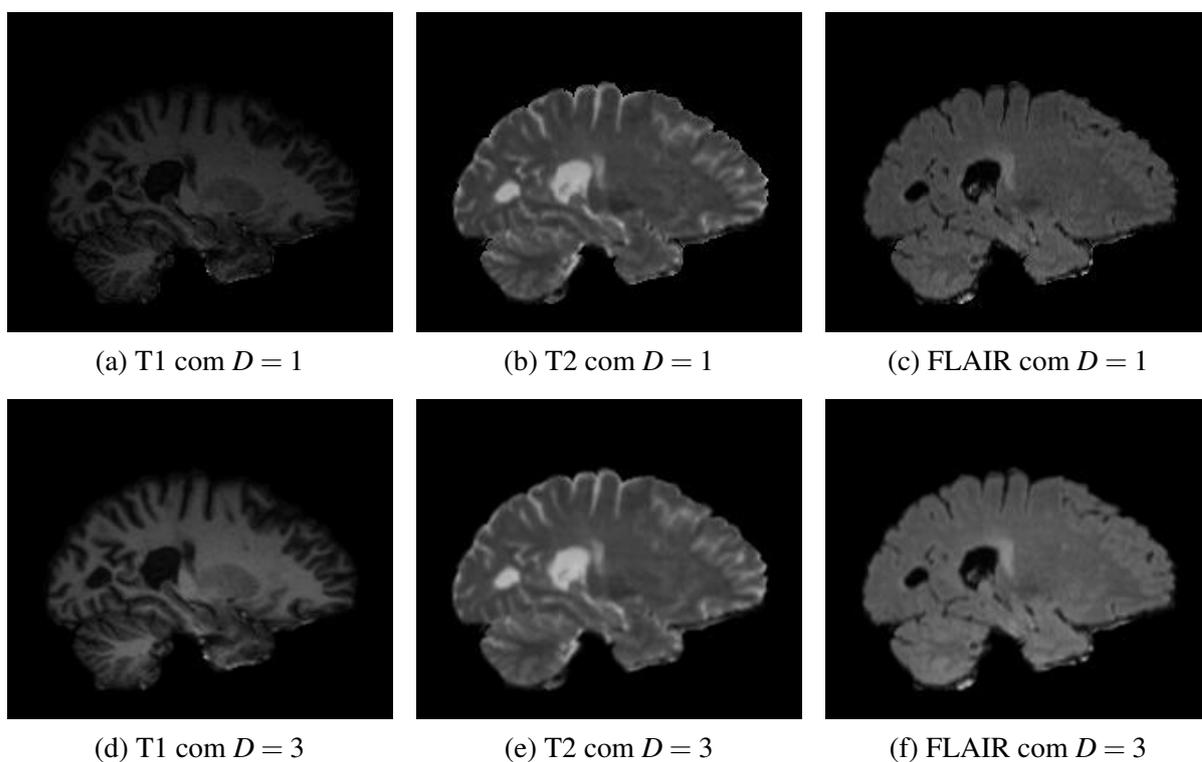
4.3.2 Cálculo das Features

O cálculo das *features* será realizado para as RMs dos tipos T1, T2 e FLAIR de toda a base. Para os *patches* D utilizados para cálculo das *features* de intensidade usamos os valores $D = 1$ e $D = 3$, os mesmo valores utilizados pelo BIANCA. Para o cálculo do DP de um *patch*, o tamanho do patch será denotado por σ_D , nesse caso os valores a serem calculados serão $\sigma_D = 3$, $\sigma_D = 5$ e $\sigma_D = 9$. As imagens utilizadas no modelo serão selecionadas em um momento posterior,

na definição dos parâmetros de inicialização, no entanto nessa etapa já deixaremos todas as *features* pré-calculadas para posterior uso no KNN. Uma das *features* utilizadas no modelo será a coordenada dos voxels, esses não precisam de nenhum cálculo pois ao utilizar as imagens pré-processadas as suas coordenadas já estão padronizadas.

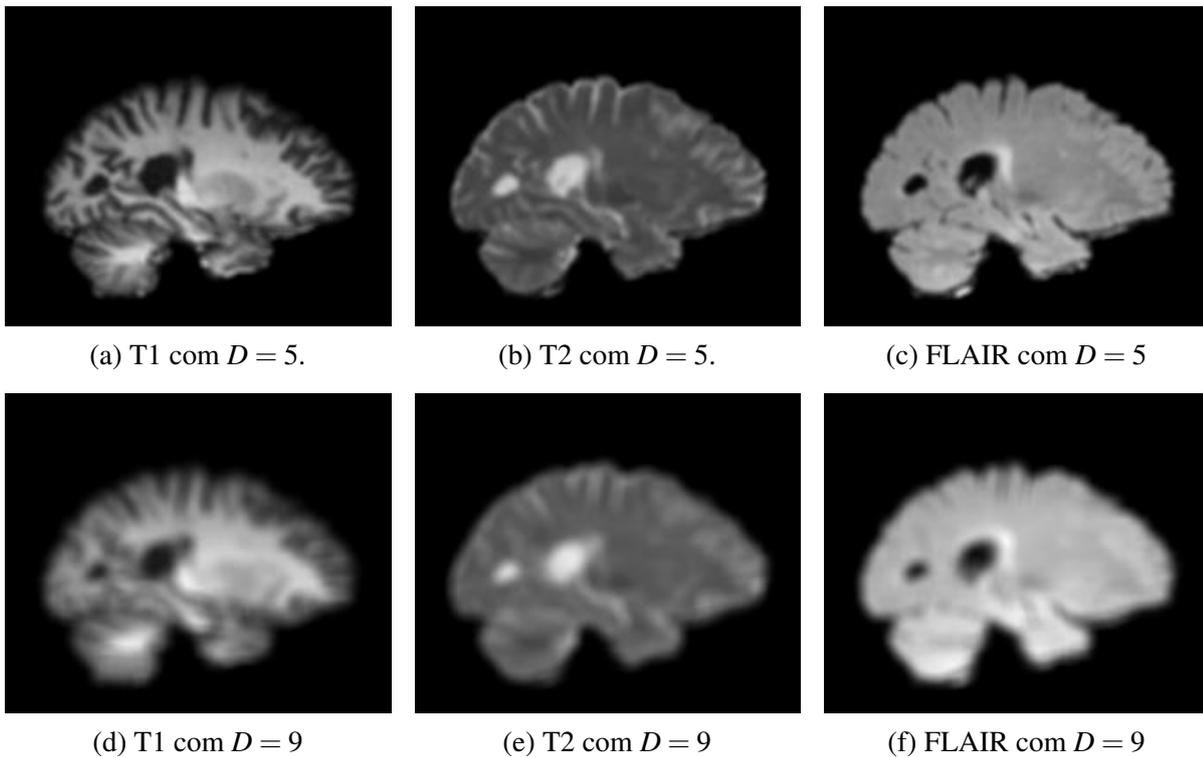
Na [Figura 7](#) mostramos as RMs em T1, T2 e FLAIR para os valores $D = 1$ e $D = 3$, podemos observar nas imagens que após aplicação da média de intensidade local 3 ([Figura 7d](#), [Figura 7e](#), [Figura 7f](#)) estão suavizadas e apresentam menor ruído, em [Figura 7f](#) podemos identificar visualmente uma lesão ao centro.

Figura 7 – RMs com $D = 1$ e $D = 3$.



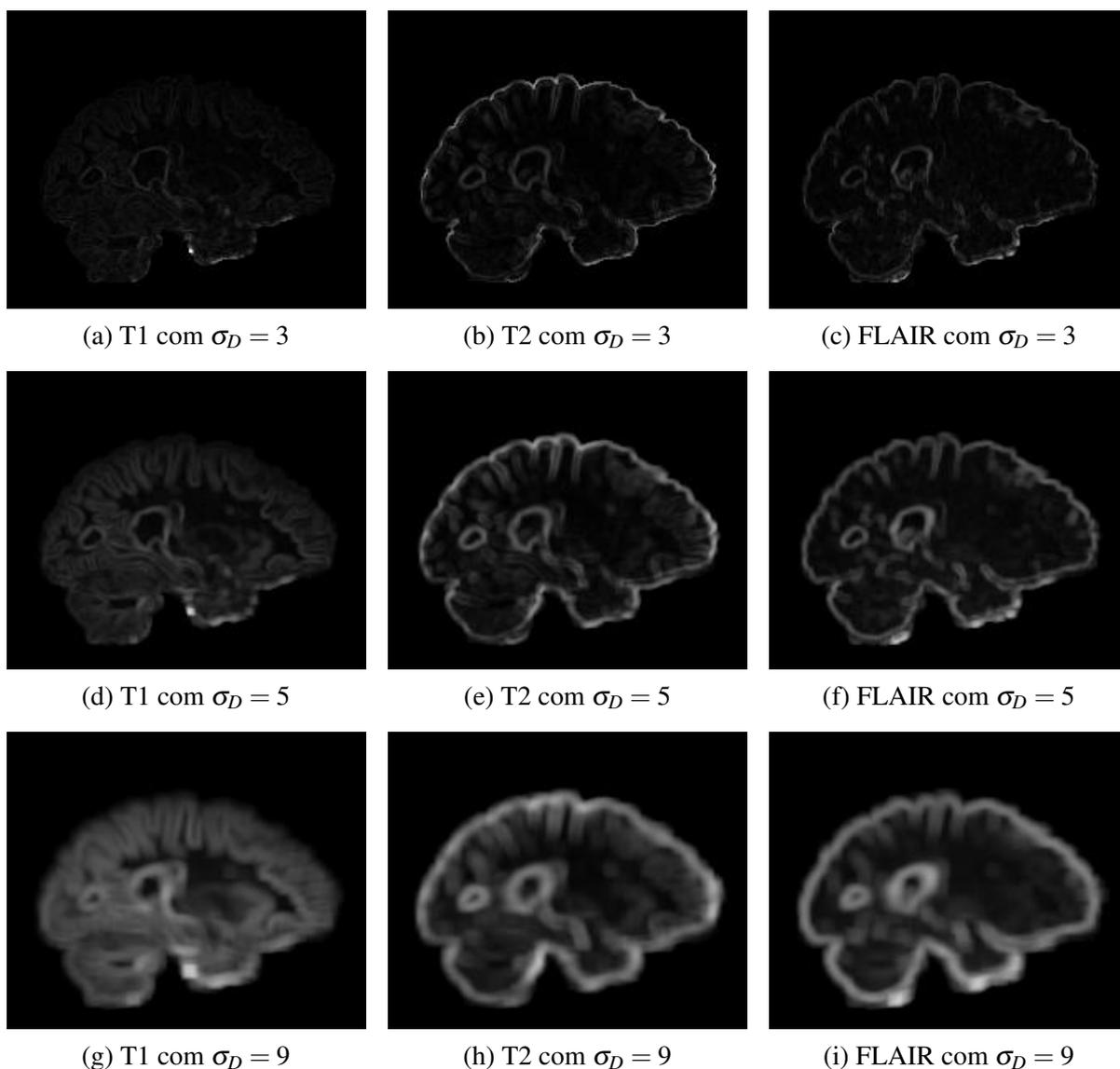
Fonte: Elaborada pelo autor.

O motivo do BIANCA ter demonstrado piores resultados piores com valores maiores de D pode ser explicado com a [Figura 8](#), podemos observar na imagem que após a aplicação ocorre uma perda de definição e a imagem se torna borrada. Em [Figura 8f](#) a lesão observada antes começa a se misturar com os pixels a sua volta.

Figura 8 – RMs com $D = 5$ e $D = 9$.

Fonte: Elaborada pelo autor.

Na [Figura 9](#) podemos observar as RMs após a utilização dos valores $\sigma_D = 3$, $\sigma_D = 5$ e $\sigma_D = 9$, com essa *feature* podemos observar melhor os contornos das áreas do cérebro. O objetivo da utilização do DP como *feature* é possibilitar ao algoritmo diferenciar regiões de borda para identificá-las na classificação, o BIANCA utiliza uma opção para executar o modelo sem as regiões de borda, a *feature* do DP é uma alternativa para informar ao modelo as informações das bordas. Assim como na aplicação do σ_D podemos visualizar uma perda de definição nas imagens ponderadas em T2 e FLAIR, tornando as bordas mais grossas conforme o DP aumenta ([Figura 9e](#)).

Figura 9 – RMs com $\sigma_D = 3$, $\sigma_D = 5$ e $\sigma_D = 9$.

Fonte: Elaborada pelo autor.

Para as *features* do *patch* D e σ_D foram calculados somente os voxels não completamente pretos em todas as faixas da RM, esses valores foram desconsiderados na aplicação do algoritmo a fim de otimizar tanto o tempo de processamento quanto o espaço de armazenamento. Ao gerar todas as *features* para todos os voxels de uma RM foi obtido um arquivo CSV com tamanho aproximado de 1.8GB enquanto o arquivo gerado após a remoção dos voxels completamente pretos possui cerca de 850MB. A redução no tempo de processamento observada foi semelhante, caindo de 5 horas para duas horas e meia para o processamento de todas as *features* dessa RM. Os resultados finais não devem ser impactado pois são removidas somente regiões não cranianas das RMs.

4.3.3 Aplicação do algoritmo

Para a aplicação do algoritmo KNN serão utilizados os arquivos CSV gerados como saída dos cálculos das *features*, a biblioteca Pandas² será utilizada para construir os *dataframes* a partir dos arquivos. O algoritmo de treinamento, as métricas e os métodos para seleção de modelos utilizados serão da biblioteca Sklearn³. Para gerar os gráficos dos resultados obtidos utilizaremos a biblioteca Matplotlib⁴.

Como o KNN é um algoritmo que calcula distância, as *features* com escalas maiores podem apresentar um maior impacto no resultado. A fim de evitar possíveis problemas de peso entre as variáveis, todas as *features*, para toda a base, serão normalizadas para valores entre 0 e 1 antes da utilização no algoritmo. O parâmetro SW passado que é o peso da *feature* espacial será aplicado multiplicando pelo valor das coordenadas x, y e z após a normalização. O valor de K do KNN será fixado em 40 que é o mesmo fixado pelo BIANCA e apresenta bons resultados.

Para a seleção dos pontos de treinamento utilizaremos valores semelhantes aos que obtiveram bons resultados demonstrados no BIANCA. A ferramenta utiliza 2000 pontos de lesão e 10000 pontos saudáveis no treinamento, no nosso modelo utilizaremos todos os pontos de lesão que são 49653 para a nossa base de treinamento e também todos os pontos de incerteza, que contabilizam mais 28127. A fim de manter a proporção o número total de pontos de saudáveis utilizados será 5 vezes o total de pontos de lesão e incerteza, assim para nosso modelo serão utilizados 383900 pontos saudáveis.

Conforme citado na Seção 4.2, a base será dividida da seguinte forma: 5 RMs para o treinamento e as outras 16 para o teste, assim para os treinamento utilizaremos os arquivos CSV referentes as RMs da base de teste. Após finalização do modelo KNN os CSVs referentes as RMs da base de teste serão fornecidos ao modelo para que seus voxels sejam classificados em uma das três classes: lesão, saudável e incerteza.

O desenvolvimento tem como objetivo a definição dos parâmetros do modelo para então a execução do treinamento. A fim de encontrar os parâmetros ideais utilizaremos a validação K-Fold Cross Validation (K-FOLD), uma breve explicação do seu funcionamento pode ser encontrada em <<https://machinelearningmastery.com/k-fold-cross-validation/>>. No K-FOLD, iremos validar diversas possibilidades para os parâmetros aceitos no nosso modelo, a fim de realizar o treinamento com os melhores resultados. Será utilizado o parâmetro $D = 3$ para o *patch* da *feature* de intensidade pois é o mesmo utilizado para o BIANCA. Em relação as outras *features* serão geradas e estimadas as configurações possíveis para os seguintes parâmetros:

- Modalidades de RM (T1), (T2), (FLAIR), (T1, T2), (T1, FLAIR), (T2, FLAIR) e (T1, T2, FLAIR);

² <<https://pandas.pydata.org/>>

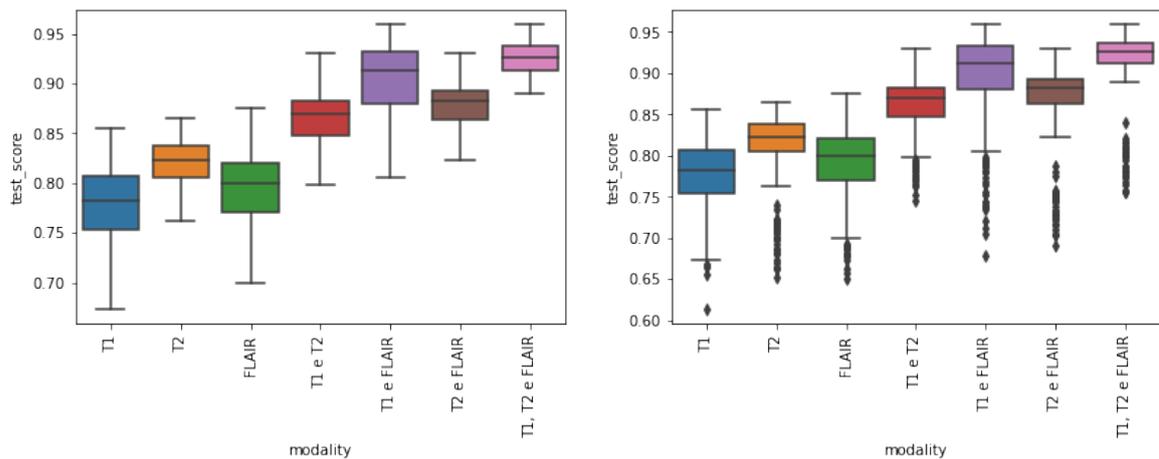
³ <<https://scikit-learn.org/stable/>>

⁴ <<https://matplotlib.org/>>

- 3, 5, 9 para σ_D . Também será gerado o caso sem a *feature*;
- 0, 0.5 e 1 como valores para SW de peso da *feature* espacial.

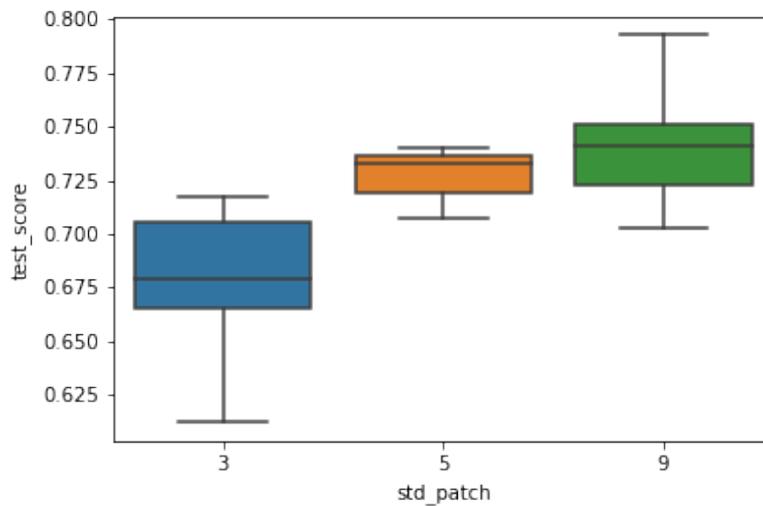
A primeira avaliação de parâmetros será em relação a quais modalidades de RM serão utilizadas, para isso foram realizadas simulações com o K-FOLD, na [Figura 10](#) podemos ver os resultados obtidos para as modalidades. Para gerar a tabela foram simulados todos os valores de σ_D , nesse momento o SW utilizado foi 0 pois ele é independente e será avaliado posteriormente nesse capítulo. Pelos gráficos apresentados podemos verificar que as melhores opções para a modalidade são (T1 e FLAIR) e (T1, T2 e FLAIR), no BIANCA as opções utilizadas foram (T1 e FLAIR) e (FLAIR) pois eram as únicas RMs disponíveis nas suas bases. Nesse modelo optaremos pela opção (T1, T2 e FLAIR) que apresentou os melhores resultados, tanto para a taxa de acerto quanto para a sua variância.

Figura 10 – Precisão por Modalidade



Fonte: Elaborada pelo autor.

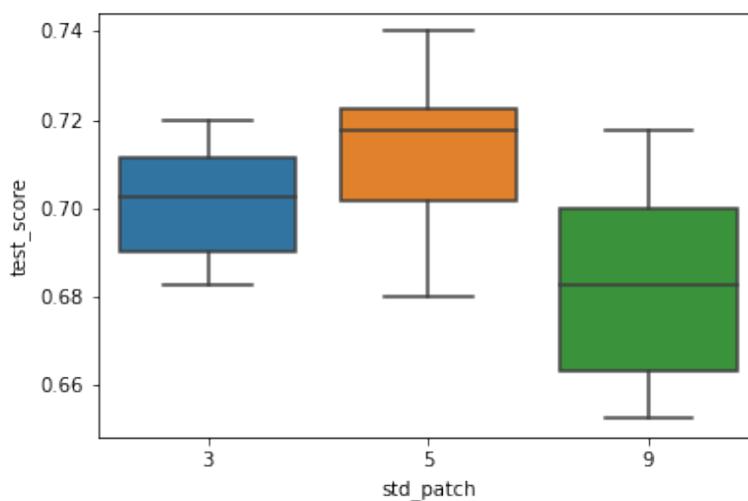
Para os valores de σ_D iremos realizar simulações através do K-FOLD para cada faixa de imagem, zerando o parâmetro SW e utilizando todas as faixas de RMs. Pelo observado em [Figura 11](#) iremos o valor de $\sigma_D = 9$ em T1 no nosso modelo, também podemos verificar pela [Figura 9](#) que as imagens em T1 começam a apresentar maior definição nas bordas somente para valores maiores de σ_D .

Figura 11 – Precisão T1 conforme o σ_D 

(a)

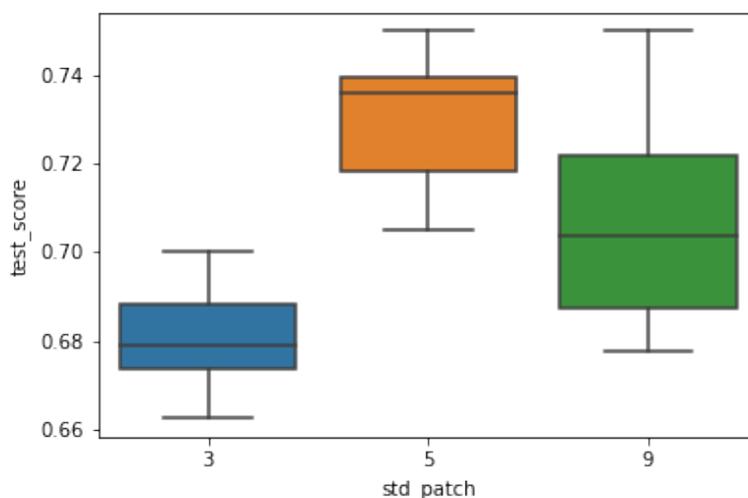
Fonte: Elaborada pelo autor.

Em [Figura 12](#) e [Figura 13](#), podemos verificar que os melhores resultados para o σ_D nas RMs ponderadas tanto em T2 como FLAIR são as com o valor 5 para a *feature*, por isso ele será utilizado no modelo. Pela [Figura 9](#) vemos que essas RMs, diferente da T1, começam a apresentar bordas grossas com perda de definição para valores maiores de forma mais rápida.

Figura 12 – Precisão T2 conforme o σ_D 

(a)

Fonte: Elaborada pelo autor.

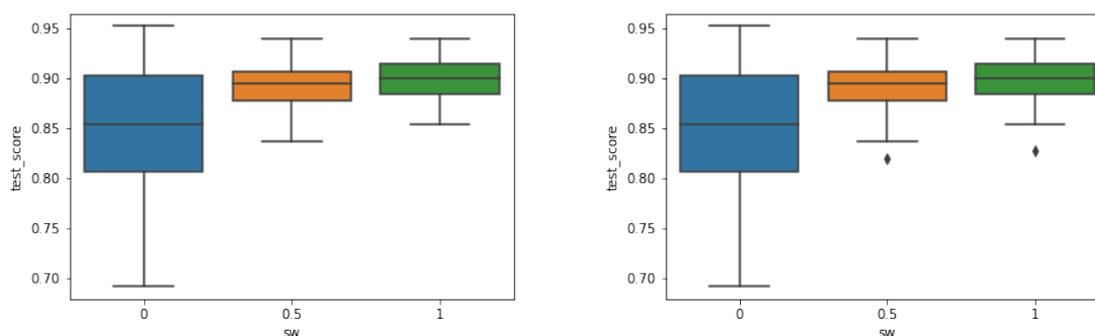
Figura 13 – Precisão FLAIR conforme o σ_D 

(a)

Fonte: Elaborada pelo autor.

Na [Figura 14](#) observamos os resultados ao utilizar o SW, as simulações foram realizadas utilizando todas as faixas de RMs com $D = 3$. A utilização desse parâmetro no BIANCA se da pois determinados pontos do FC possuem maior probabilidade de apresentarem inflamações. É importante ressaltar que valores muito altos para essa *feature* podem piorar os resultados, isso se deve ao fato que se considerássemos a *feature* de forma isolada o modelo do KNN decoraria que posições específicas do cérebro são lesões, assim as áreas de lesões definidas serão enviesadas pelas apresentadas na base de treinamento, inclusive ao apresentar um cérebro saudável no teste, o modelo tenderia a responder que determinado ponto é uma lesão. No próprio artigo ([GRIFFANTI *et al.*, 2016](#)) o autor fala sobre o possível problema ao atribuir pesos altos para essa *feature*.

Figura 14 – Precisão pelo peso espacial



Considerando os resultados apontados nesse capítulo definimos todos os parâmetros que serão passados para o modelo, eles serão:

- Modalidades: T1, T2 e FLAIR
- T1: $D = 3$ e $\sigma_D = 9$
- T2: $D = 3$ e $\sigma_D = 5$
- FLAIR: $D = 3$ e $\sigma_D = 5$
- SW: 1

RESULTADOS

Nesse capítulo iremos falar sobre os resultados obtidos após a aplicação do algoritmo KNN com os parâmetros definidos no capítulo anterior. Iniciamos apresentando os modelos utilizados pelo BIANCA, assim como as suas bases e então apresentamos os resultados obtidos pelo estudo em relação ao BIANCA. Em seguida mostramos imagens das RMs com as máscaras preditas pelo modelo para uma avaliação visual e comparação com as do diagnóstico real dos especialistas. No final do capítulo falaremos sobre as principais dificuldades e limitações encontradas no desenvolvimento.

5.1 Análise Quantitativa

Nessa seção avaliaremos os resultados obtidos pelo modelo desenvolvido nesse estudo em comparação aos 3 modelos divulgados no BIANCA, por simplicidade chamaremos eles de BIANCA 1, BIANCA 2 e BIANCA 3.

O BIANCA 1 utiliza como configuração as RMs ponderadas em FLAIR e T1, $SW = 1$, $D = 1$, seleciona pacientes com alta taxa de lesões, e faz uso de 2000 pontos de lesão e 10.000 pontos saudáveis para o treinamento. O método também apresenta um *threshold* para aceitar o ponto como lesão, o valor definido foi de 0,9. A base de dados utilizada foi de pessoas com idade de 57 a 91 anos, 25 com provável doença de Alzheimer e 24 com comprometimento cognitivo leve, 11 com comprometimento cognitivo subjetivo e 25 de controle com sistema cognitivo saudável. As RMs disponíveis são as ponderadas em FLAIR e T1, a única diferença do BIANCA 1 para o BIANCA 2 é a utilização do *patch* $D = 3$ no modelo.

Para o BIANCA 3 as configurações foram, $D = 1$, $SW = 1$, 20 pacientes com alta taxa de lesão inclusos no treino, *threshold* 0,9 e somente imagens ponderadas em FLAIR. A base utilizada para o teste desse modelo possui 474 pacientes que passaram por um pequeno derrame não debilitante ou por um ataque isquêmico transitório, idade de 20 até 102 anos. O modelo se

utiliza somente das RMs ponderadas em FLAIR por serem as únicas disponíveis na base.

Na tabela [Tabela 3](#) apresentamos os resultados obtidos pelo BIANCA e os valores observados pelo modelo desenvolvido nesse estudo. Como o modelo do BIANCA só apresenta duas classes, o resultado do estudo foi adaptado para realizar a comparação, para isso os valores de incerteza das máscaras originais e preditas foram considerados como lesão.

Tabela 3 – Comparação de resultados em relação ao BIANCA

MODELO	SI	RFP	RFN	DER	OER
ESTUDO	0.33	2.73	0.25	1.33	0.33
BIANCA 1	0.75	0.22	0.26	0.03	0.47
BIANCA 2	0.76	0.22	0.25	0.03	0.46
BIANCA 3	0.52	0.46	0.45	0.19	0.76

Fonte: Adaptada de [Magnetic... \(2021\)](#).

Pelos valores observados na ta tabela vemos uma grande RFP que é a razão entre os valores classificados erroneamente como positivos e o número real de positivos, isso indica que o modelo classificou uma grande quantidade de pontos saudáveis como positivo ou área de incerteza.

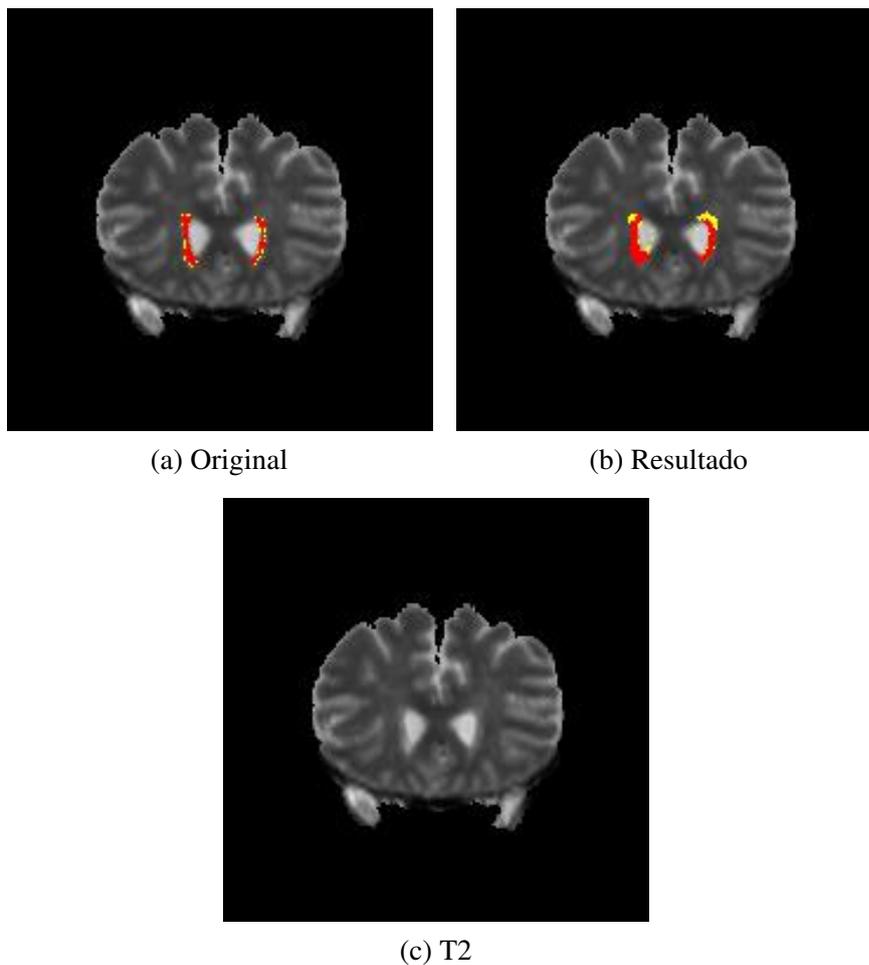
A RFN obtida foi semelhante a observada no BIANCA 1 e BIANCA 2, isso indica que o modelo errou pouco ao classificar pontos de lesão como saudáveis. Esse dado junto ao RFP indica que o modelo classificou grande quantidade de pontos como positivos.

5.2 Análise Qualitativa

Como o modelo realiza predições de RMs que são imagens volumétricas, a interpretação dos resultados de forma numérica não é tão intuitiva. Essa seção apresenta alguns exemplos de resultados, mostrando as máscaras originais, as preditas pelo modelo e a RM original ponderada em T2.

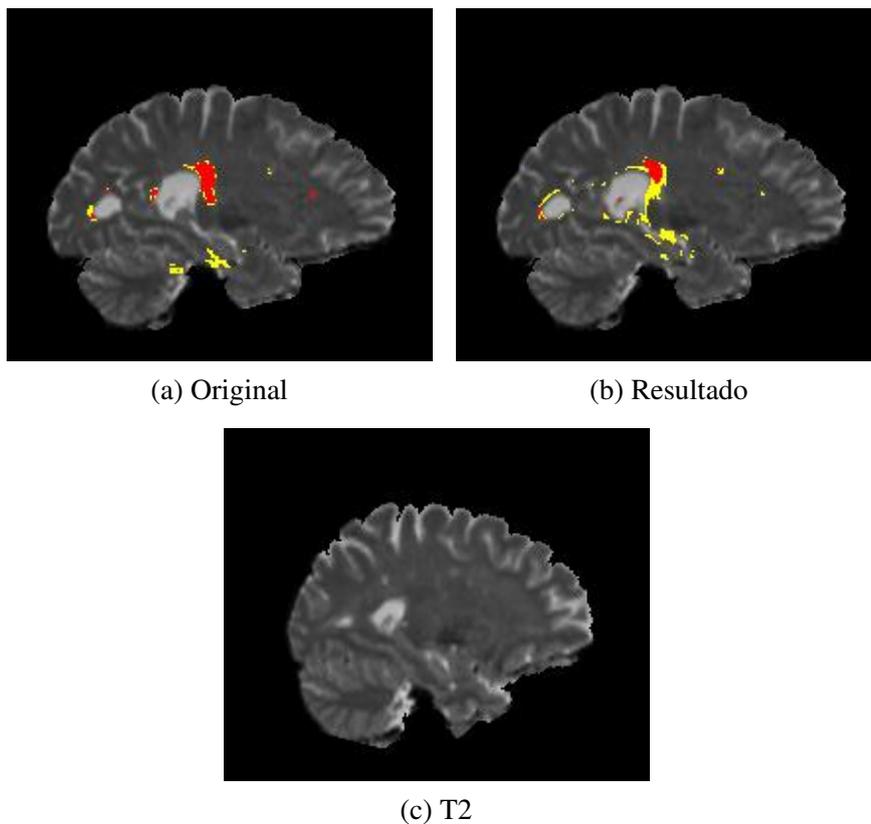
Em [Figura 15](#) podemos visualizar a comparação entre as máscaras uma seção da RM de um paciente, vemos nela que o local da predição está bem próximo do original porém existe uma discordância entre os voxels classificados como lesão e áreas de incerteza.

Figura 15 – Resultados em relação a máscara



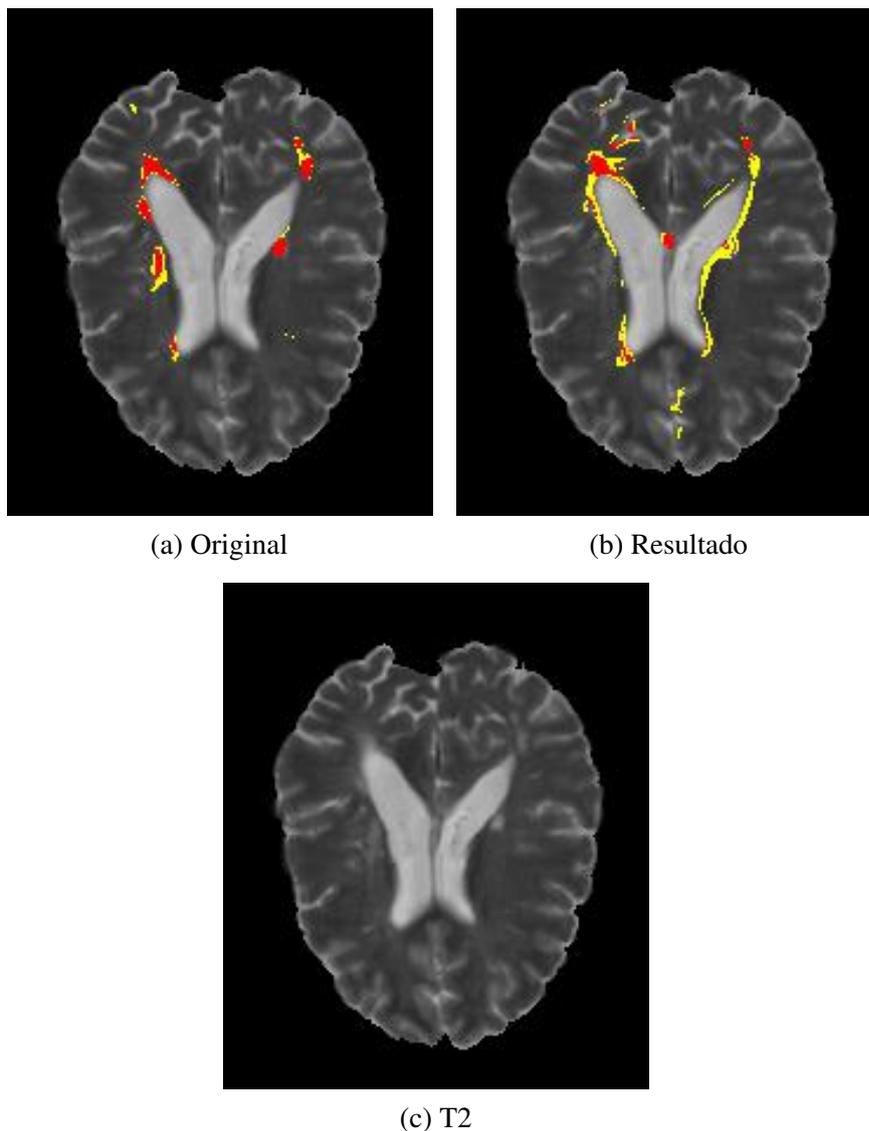
Ao observar a [Figura 16](#), vemos uma discordância maior entre as a máscara original e a predição. O modelo acertou regiões maiores e marcou a maioria das regiões da original como lesão ou incerteza, porém ainda deixa alguns pontos escaparem como a parte amarela na parte inferior esquerda da imagem original. O modelo também marcou uma área muito maior como incerteza, pela imagem vemos que a maior parte que foi confundida como área de incerteza é referente a regiões de borda entre seções do cérebro. Esse resultado confirma o que era previsto pela RFP e RFN obtida pelo modelo.

Figura 16 – Resultados em relação a máscara



Também podemos observar na [Figura 17](#) um comportamento semelhante ao observado na imagem anterior, a região central foi inclusive interpretada como lesão pelo modelo, isso se deve possivelmente ao fato de que é uma região de borda entre duas seções.

Figura 17 – Resultados em relação a máscara



5.3 Dificuldades e Limitações

O principal problema encontrado no desenvolvimento foi em relação ao tempo de processamento para cálculo das *features*. Considerando a base de treinamento e teste foram utilizados 21 exames de RM, cada um deles ponderado em T1, T2 e FLAIR, cada RM com o tamanho de 181x217x181, resultando em 21.327.411 voxels por exame e 447.875.631 para toda a base. Ao tentar executar todas as etapas de cálculo das *features* mais o algoritmo KNN padrão da biblioteca *scikit learn*, o programa não conseguiu chegar a um resultado após cinco dias em uma máquina com 32GB de RAM e processador com 4 núcleos e 4.5GHz de frequência. Ao observar os recursos do computador notou-se que toda a RAM do computador estava sendo utilizada, dessa forma o processamento foi dividido em etapas, gerando arquivos a cada uma das etapas de cálculo das *features* e também com os resultados finais. A pasta com todos os arquivos intermediários e finais gerados durante o desenvolvimento foi de pouco mais de 200GB.

Outro objetivo era submeter o modelo para análise na competição <<https://smart-stats-tools.org/lesion-challenge>>, no entanto por mais que o site possuía um botão para criar uma conta, não foi possível finalizar o cadastro pois ele é submetido para análise e em nenhuma das diversas tentativas realizadas no segundo semestre de 2021 e início de 2022 foi obtida alguma resposta.

CONCLUSÃO

A utilização do BIANCA para realizar identificação de lesões no cérebro em RMs já havia sido demonstrada, com a base utilizada nesse projeto foi possível realizar o diagnóstico sem a necessidade da utilização das imagens com o contraste. O modelo acrescenta o DP como *feature* e propõe nos resultados uma nova forma de classificação adicionando as áreas de incerteza nas classes finais, auxiliando assim o radiologista para que seja necessário maior atenção somente em determinadas áreas na sua análise, acelerando dessa forma o tempo para o diagnóstico que é crucial para pacientes com EM.

Como trabalho futuro e continuidade do projeto, o objetivo inicial é melhorar os resultados obtidos no estudo, encontrando novas *features* como o GLCM e realizando uma seleção de atributos a fim de evitar problemas da "Maldição da Alta Dimensionalidade" que pode afetar algoritmos como o KNN. Com a melhora dos resultados o objetivo é disponibilizar uma ferramenta online ou os códigos e técnicas para que outros possam utilizar como auxiliar no diagnóstico. Uma das grandes dificuldades encontradas no desenvolvimento foi em relação de processamento, por isso será desenvolvida uma nova implementação utilizando CUDA no cálculo das *features* a fim de reduzir o tempo necessário para a execução. O código também deve ser adaptado para ficar no formato de submissão da competição de 2015 para segmentação de EM, novas tentativas de cadastro serão realizadas e o objetivo é comparar os resultados com os outros métodos apresentados na competição e acrescentar aos seus estudos o método desenvolvido.

REFERÊNCIAS

CARASS, A.; ROY, S.; JOG, A.; CUZZOCREO, J. L.; MAGRATH, E.; GHERMAN, A.; BUTTON, J.; NGUYEN, J.; PRADOS, F.; SUDRE, C. H.; Jorge Cardoso, M.; CAWLEY, N.; CICCARELLI, O.; WHEELER-KINGSHOTT, C. A.; OURSELIN, S.; CATANESE, L.; DESHPANDE, H.; MAUREL, P.; COMMOWICK, O.; BARILLOT, C.; TOMAS-FERNANDEZ, X.; WARFIELD, S. K.; VAIDYA, S.; CHUNDURU, A.; MUTHUGANAPATHY, R.; KRISHNAMURTHI, G.; JESSON, A.; ARBEL, T.; MAIER, O.; HANDELS, H.; IHEME, L. O.; UNAY, D.; JAIN, S.; SIMA, D. M.; SMEETS, D.; GHAFOORIAN, M.; PLATEL, B.; BIRENBAUM, A.; GREENSPAN, H.; BAZIN, P.-L.; CALABRESI, P. A.; CRAINICEANU, C. M.; ELLINGSEN, L. M.; REICH, D. S.; PRINCE, J. L.; PHAM, D. L. Longitudinal multiple sclerosis lesion segmentation: Resource and challenge. **NeuroImage**, v. 148, p. 77–102, 2017. ISSN 1053-8119. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1053811916307819>>. Citado nas páginas 26 e 38.

CHOY, G.; KHALILZADEH, O.; MICHALSKI, M.; DO, S.; SAMIR, A. E.; PIANYKH, O. S.; GEIS, J. R.; PANDHARIPANDE, P. V.; BRINK, J. A.; DREYER, K. J. Current applications and future impact of machine learning in radiology. **Radiology**, v. 288, n. 2, p. 318–328, 2018. PMID: 29944078. Disponível em: <<https://doi.org/10.1148/radiol.2018171820>>. Citado na página 21.

DINIZ, P. H. B.; VALENTE, T. L. A.; DINIZ, J. O. B.; SILVA, A. C.; GATTASS, M.; VENTURA, N.; MUNIZ, B. C.; GASPARETTO, E. L. Detection of white matter lesion regions in mri using slic0 and convolutional neural network. **Computer Methods and Programs in Biomedicine**, v. 167, p. 49–63, 2018. ISSN 0169-2607. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169260717304455>>. Citado nas páginas 25 e 27.

FREIRE, P. G. de lima. **Automatic computational scheme for segmentation, volumetric assessment and analysis of multiple sclerosis lesions in magnetic resonance images of the human brain**. São Carlos: [s.n.], 2019. Citado nas páginas 23, 24, 25 e 27.

GRIFFANTI, L.; ZAMBONI, G.; KHAN, A.; LI, L.; BONIFACIO, G.; SUNDARESAN, V.; SCHULZ, U. G.; KUKER, W.; BATTAGLINI, M.; ROTHWELL, P. M.; JENKINSON, M. Bianca (brain intensity abnormality classification algorithm): A new tool for automated segmentation of white matter hyperintensities. **NeuroImage**, v. 141, p. 191–205, 2016. ISSN 1053-8119. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1053811916303251>>. Citado nas páginas 25, 26, 27, 29, 30, 34, 35 e 47.

HIRD, R. **The Basics of MRI Interpretation**. 2021. Disponível em: <<https://geekymedics.com/the-basics-of-mri-interpretation/>>. Acesso em: 16/10/2021. Citado na página 29.

JUNIOR, J. C. N. **Deteção automática de lesões novas em exames consecutivos de pacientes com esclerose múltipla**. Dissertação (Mestrado) — Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2019. Disponível em: <<https://doi.org/10.11606/D.17.2020.tde-24082020-094901>>. Citado nas páginas 21, 25 e 27.

MAGNETIC Resonance Imaging (MRI) of the Brain and Spine: Basics. 2021. Disponível em: <<https://case.edu/med/neurology/NR/MRI%20Basics.htm>>. Acesso em: 16/10/2021. Citado nas páginas 29, 30 e 50.

MCKINLEY, R.; WEPFER, R.; GRUNDER, L.; ASCHWANDEN, F.; FISCHER, T.; FRIEDLI, C.; MURI, R.; RUMMEL, C.; VERMA, R.; WEISSTANNER, C.; WIESTLER, B.; BERGER, C.; EICHINGER, P.; MUHLAU, M.; REYES, M.; SALMEN, A.; CHAN, A.; WIEST, R.; WAGNER, F. Automatic detection of lesion load change in multiple sclerosis using convolutional neural networks with segmentation confidence. **NeuroImage: Clinical**, v. 25, p. 102104, 2020. ISSN 2213-1582. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2213158219304516>>. Citado nas páginas 21, 25 e 27.

NGUYEN, T. M.; WU, Q. M. J. Robust student's-t mixture model with spatial constraints and its application in medical image segmentation. **IEEE Transactions on Medical Imaging**, v. 31, n. 1, p. 103–116, 2012. Citado nas páginas 25 e 27.

ROURA, E.; OLIVER, A.; CABEZAS, M.; VALVERDE, S.; PARETO, D.; VILANOVA, J. C.; RAMIÓ-TORRENTÀ, L.; ROVIRA, À.; LLADÓ, X. A toolbox for multiple sclerosis lesion segmentation. **Neuroradiology**, v. 57, n. 10, p. 1031–1043, Oct 2015. ISSN 1432-1920. Disponível em: <<https://doi.org/10.1007/s00234-015-1552-2>>. Citado nas páginas 25 e 27.

SALEM, M.; CABEZAS, M.; VALVERDE, S.; PARETO, D.; OLIVER, A.; SALVI, J.; ROVIRA Àlex; LLADÓ, X. A supervised framework with intensity subtraction and deformation field features for the detection of new t2-w lesions in multiple sclerosis. **NeuroImage: Clinical**, v. 17, p. 607–615, 2018. ISSN 2213-1582. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2213158217302954>>. Citado na página 21.

TALO, M.; BALOGLU, U. B.; YILDİRİM Özal; Rajendra Acharya, U. Application of deep transfer learning for automated brain abnormality classification using mr images. **Cognitive Systems Research**, v. 54, p. 176–188, 2019. ISSN 1389-0417. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1389041718310933>>. Citado nas páginas 25 e 27.

TRABOULSEE, A.; SIMON, J.; STONE, L.; FISHER, E.; JONES, D.; MALHOTRA, A.; NEWSOME, S.; OH, J.; REICH, D.; RICHERT, N.; RAMMOHAN, K.; KHAN, O.; RADUE, E.-W.; FORD, C.; HALPER, J.; LI, D. Revised recommendations of the consortium of ms centers task force for a standardized mri protocol and clinical guidelines for the diagnosis and follow-up of multiple sclerosis. **American Journal of Neuroradiology**, American Journal of Neuroradiology, v. 37, n. 3, p. 394–401, 2016. ISSN 0195-6108. Disponível em: <<http://www.ajnr.org/content/37/3/394>>. Citado na página 23.

VASKOVIĆ, J. **Normal brain MRI**. 2021. Disponível em: <<https://www.kenhub.com/en/library/anatomy/normal-brain-mri>>. Acesso em: 16/10/2021. Citado na página 30.

WATTJES, M. P.; ROVIRA, À.; MILLER, D.; YOUSRY, T. A.; SORMANI, M. P.; STEFANO, N. de; TINTORÉ, M.; AUGER, C.; TUR, C.; FILIPPI, M.; ROCCA, M. A.; FAZEKAS, F.; KAPPOS, L.; POLMAN, C.; BARKHOF, F.; MONTALBAN, X.; GROUP, o. b. o. t. M. s. Magnims consensus guidelines on the use of mri in multiple sclerosis—establishing disease prognosis and monitoring patients. **Nature Reviews Neurology**, v. 11, n. 10, p. 597–606, Oct 2015. ISSN 1759-4766. Disponível em: <<https://doi.org/10.1038/nrneurol.2015.157>>. Citado na página 23.

