Pontifícia Universidade Católica
do Rio de Janeiro

Fernanda da Cunha Duarte

Classificação de Gliomas Utilizando Índices de Biodiversidade e de Diversidade Filogenética em Imagens por Ressonância Magnética Através de uma Abordagem *Radiomics*

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós–graduação em Informática da PUC-Rio .

Orientador : Prof. Marcelo Gattass
Co-orientador: Prof. Aristófanes Corrêa Silva

Rio de Janeiro
Maio de 2019

Pontifícia Universidade Católica
do Rio de Janeiro

Fernanda da Cunha Duarte

Classificação de Gliomas Utilizando Índices de Biodiversidade e de Diversidade Filogenética em Imagens por Ressonância Magnética Através de uma Abordagem *Radiomics*

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós–graduação em Informática da PUC-Rio . Aprovada pela Comissão Examinadora abaixo.

Prof. Marcelo Gattass
Orientador
Departamento de Informática – PUC-Rio

Prof. Aristófanes Corrêa Silva
Co-orientador
Universidade Federal do Maranhão – UFMA

Prof. Waldemar Celes Filho
Departamento de Informática – PUC-Rio

Prof. Helio Côrtes Vieira Lopes
Departamento de Informática – PUC-Rio

Rio de Janeiro, 21 de Maio de 2019

Fernanda da Cunha Duarte

Graduou-se em Matemática Aplicada pela Escola de Matemática Aplicada (EMAp) da Fundação Getúlio Vargas. Iniciou o curso de mestrado do Departamento de Informática da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) em 2017, na área de computação gráfica, e tornou-se bolsista pelo Instituo Tecgraf de Desenvolvimento de Software Técnico-Cientíco da PUC-Rio e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Ficha Catalográfica

Duarte, Fernanda

Classificação de Gliomas Utilizando Índices de Biodiversidade e de Diversidade Filogenética em Imagens por Ressonância Magnética Através de uma Abordagem *Radiomics* / Fernanda da Cunha Duarte; orientador: Marcelo Gattass; coorientador: Aristófanes Corrêa Silva. – Rio de janeiro: PUC-Rio , Departamento de Informática, 2019.

v., 84 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui bibliografia

1. Informática – Teses. 2. Classificação de Gliomas;. 3. Radiomics;. 4. Machine Learning;. 5. Imagens Médicas.. I. Gattass, Marcelo. II. C. Silva, Aristófanes. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

Agradecimentos

À minha família, que esteve sempre ao meu lado, me apoiando e ajudando, principalmente nos momentos mais difíceis. Obrigada por todo o amor e carinho.

Ao meu professor e orientador Marcelo Gattass, por compartilhar comigo parte de seus infinitos conhecimentos e ensinamentos, que levarei comigo para sempre. Obrigada pela paciência, pelo apoio, pela confiança e pelo carinho.

Ao professor Aristofanes Correa Silva, por ter me acompanhado desde o início deste trabalho, fornecendo tantos conhecimentos e conselhos essenciais para a sua conclusão. Muito obrigada.

Ao meu amigo Jeferson Coelho, por todos os conselhos pessoais e acadêmicos, por toda a ajuda e por todo o conhecimento compartilhado. Muito obrigada.

Aos professores Paulo Cezar Carvalho e Asla Sá, que me apresentaram ao mundo da computação gráfica e me ajudaram a embarcar nesta grande aventura, com tantos ensinamentos e tanto carinho. Obrigada por tudo.

A todos os meus amigos, que permaneceram sempre ao meu lado, acreditando em mim e me proporcionando tantos momentos de felicidade. Obrigada por todo o amor e carinho.

Ao Instituto Tecgraf, por todas as oportunidades e por todo o apoio e suporte.

Ao CNPq, pelo suporte financeiro.

Muito obrigada a todos.

Resumo

Duarte, Fernanda; Gattass, Marcelo; C. Silva, Aristófanes. **Classificação de Gliomas Utilizando Índices de Biodiversidade e de Diversidade Filogenética em Imagens por Ressonância Magnética Através de uma Abordagem *Radiomics***. Rio de Janeiro, 2019. 84p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Gliomas estão entre os tumores cerebrais malignos mais comuns. Eles podem ser classificados entre gliomas de baixo e alto grau e sua identificação precoce é fundamental para o direcionamento do tratamento aplicado. Utilizando uma abordagem *radiomics*, o presente trabalho propõe o uso de índices de biodiversidade e de diversidade filogenética, definidos no campo da biologia, no problema de classificação de gliomas. O método proposto apresentou resultados promissores, com AUC-ROC (*area under the* ROC *curve*), acurácia, sensibilidade e especificidade de 0,951, 0,930, 0,967 e 0,827, respectivamente.

Palavras-chave

Classificação de Gliomas; Radiomics; Machine Learning; Imagens Médicas.

Abstract

Duarte, Fernanda; Gattass, Marcelo (Advisor); C. Silva, Aristófanes (Co-Advisor). **Radimocs Analysis for Glioma Grading Using Biodiversity and Phylogenetic Diversity Indices on Multi-Modal Magnetic Resonance Imaging**. Rio de Janeiro, 2019. 84p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Gliomas are among the most common malignant brain tumors. They can be classified into low-grade and high-grade gliomas and their early identification is crucial for treatment direction. Using a radiomics approach, the present work proposes the use of biodiversity and phylogenetic diversity biology indices to handle the glioma classification problem. The proposed method presented promising results, with AUC-ROC (*area under the* ROC *curve*), accuracy, sensitivity and specificity of 0,951, 0,930, 0,967 and 0,827, respectively.

Keywords

Glioma Classification; Glioma Grading; Radiomics; Machine Learning; Medical Images.

Sumário

1 Introdução **14**
1.1 Objetivos 15
1.2 Estrutura do Trabalho 15

2 Trabalhos Relacionados **16**

3 Fundamentação Teórica **19**
3.1 Gliomas 19
3.2 Radiomics 20
3.3 Imagens por Ressonância Magnética 22
3.4 Processamento Digital de Imagens 23
3.4.1 Histogram Matching 24
3.4.2 Extração de Regiões de Interesse 26
3.5 Índices de Biodiversidade 27
3.6 Índices de Diversidade Filogenética 30
3.7 Machine Learning e Algoritmos de Classificação 32
3.7.1 Support Vector Machine 32
3.7.1.1 Problemas linearmente separáveis 34
3.7.1.2 Problemas não linearmente separáveis 36
3.7.2 Random Forest 38
3.7.3 Multilayer Perceptron 40
3.7.4 Validação Cruzada 45
3.7.5 Seleção de atributos 46
3.7.5.1 ANOVA 47
3.7.5.2 Extremely Randomized Trees 47
3.7.5.3 Recursive Feature Elimination 48
3.7.6 Métricas de Desempenho 48

4 Metodologia Proposta **52**
4.1 Base de Dados 52
4.2 Pré-processamento 54
4.3 Extração de Atributos 55
4.3.1 Primeira Abordagem 56
4.3.2 Segunda Abordagem 56
4.3.3 Terceira Abordagem 56
4.3.4 Quarta Abordagem 57
4.4 Classificação 57
4.5 Validação dos Resultados 58

5 Resultados **60**
5.1 Resultados da Primeira Abordagem 60
5.2 Resultados da Segunda Abordagem 62
5.3 Resultados da Terceira Abordagem 64
5.4 Resultados da Quarta Abordagem 68

5.5 Discussão 71

6 Conclusão e Trabalhos Futuros **74**

Referências **76**

Lista de figuras

Figura 3.1 Exame por ressonância magnética de um paciente com glioma de alto grau, obtido através da base de dados do desafio BraTS'18. a) Fatia do exame mostrando o cérebro com a lesão; b) Imagem mostrando apenas a estrutura da lesão. 20
Figura 3.2 Imagem extraída de Kumar et al. (2012) [63], ilustrando as etapas associadas à abordagem *radiomics* e seus desafios. 21
Figura 3.3 Exemplos de fatias bidimensionais extraídas de imagens volumétricas obtidas através das quatro modalidades de MRI utilizadas no presente estudo [58, 59]. a) T1; b) T1ce; c) T2; d) FLAIR. 23
Figura 3.4 Gráfico ilustrativo mostrando a correspondência entre duas funções de distribuição acumulada. 25
Figura 3.5 Exemplos de imagens e histogramas extraídos de nossa base de dados antes e depois da aplicação do método de especificação de histograma. (a) Histograma da imagem tomada como referência e centésima fatia da imagem; (b) Histograma de uma imagem I_1 e exemplo de fatia, antes e depois da aplicação do método, respectivamente; (c) Histograma de uma imagem I_2 e exemplo de fatia, antes e depois da aplicação do método, respectivamente. 26
Figura 3.6 a) Exemplo de fatia de um exame de MRI; b) Exemplo de máscara utilizada para extração do tumor da fatia analisada; c) Resultado da extração. 27
Figura 3.7 Exemplo de relações filogenéticas representadas por um cladograma. Neste exemplo, o valor máximo de voxel encontrado no volume de interesse é 1252. 31
Figura 3.8 Etapas de um algoritmo de classificação. 32
Figura 3.9 a) Exemplo de diferentes hiperplanos separando um conjunto de dados com duas classes linearmente separáveis. Neste caso, cada instância possui dois atributos (representados por cada eixo) e sua cor (azul ou vermelho) representa a classe correspondente. b) Hiperplano ótimo h_2 com distância m (margem) para as instâncias de cada classe mais próximas a ele. Tais instâncias são chamadas de vetores de suporte (*support vectors*). 33
Figura 3.10 Exemplo de hiperplanos paralelos traçados levando em conta a margem m 34
Figura 3.11 Exemplo de classificação utilizando SVM de margem suave. 36
Figura 3.12 Demonstração de como funciona uma árvore de decisão. Os atributos x_{ij} não necessariamente são escolhidos na ordem em que aparecem nos vetores de instâncias x_i, neste caso foram colocados em ordem apenas para simplificar a demonstração. 38
Figura 3.13 Esquema ilustrando a classificação realizada pelo algoritmo *random forest*. 40

Figura 3.14 a) Representação simplificada de um neurônio biológico (extraída e adaptada de [37]). As sinapses são as conexões entre os terminais do axônio de um neurônio com os dendritos de outros neurônios; b) Representação de um neurônio artificial em uma camada l recebendo sinais de k neurônios pertencentes à camada $(l-1)$ anterior e enviando um sinal de valor a_j^l para os neurônios da camada seguinte. 41
Figura 3.15 Representação de uma rede neural MLP através de um grafo direcionado. 42
Figura 3.16 Ilustração demonstrando o funcionamento de uma validação cruzada 5-*fold*. M é o valor resultante da métrica utilizada para medir o desempenho do modelo. 46
Figura 3.17 Matriz de confusão para um problema de classificação binária. 49
Figura 3.18 Ilustração indicando a curva ROC e a área AUC associada a ela. 51

Figura 4.1 Metodologia proposta. 52
Figura 4.2 Exemplos de fatias de cada modalidade de MRI relativas a um indivíduo presente da base de dados fornecida pelo desafio BraTS 2018. Da esquerda para a direita: modalidades T1, FLAIR, T2, T1ce e anotação das regiões intratumorais. De cima para baixo: fatias 45, 55, 65 e 75, respectivamente. 53
Figura 4.3 Imagem mostrando as diferentes subregiões que podem aparecer em um glioma. (a) A região total do tumor está indicada em amarelo e sua segmentação é feita a partir das modalidades T2 e FLAIR. (b) O centro (núcleo) tumoral está mostrado em vermelho e é visualizado através da modalidade T2. (c) As subregiões ET e NCR internas ao centro tumoral são visíveis em T1ce e estão representadas em azul e verde, respectivamente. (d) Combinação de todas as informações anteriores para formar as subregiões fornecidas pela base de dados: a região total menos o centro e suas estruturas internas formam o edema peritumoral (amarelo); o centro tumoral menos as estruturas NCR e ET, formam a subregião NET (vermelho); as partes em azul indicam apenas a área ET; e em verde, a estrutura NCR. Vale ressaltar que nos dados fornecidos, NCR e NET são representadas como uma classe só. Imagem obtida em [69]. 54
Figura 4.4 Exemplo de anotações para cada subregião intratumoral: NCR/NET (cinza escuro), ED (cinza claro) e ET (branco). 55
Figura 4.5 Exemplo de fatia 2D da imagem de anotação após a divisão em camadas (terceira abordagem), na qual cada cor representa uma camada diferente. 57

Figura 5.1 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na primeira abordagem, sem o uso do algoritmo de *histogram matching*. 63

Figura 5.2 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na primeira abordagem, com o uso do algoritmo de *histogram matching*. 63
Figura 5.3 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na segunda abordagem, com o uso do algoritmo de *histogram matching*. 65
Figura 5.4 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na segunda abordagem, sem o uso do algoritmo de *histogram matching*. 65
Figura 5.5 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na terceira abordagem, sem o uso do algoritmo de *histogram matching*. 67
Figura 5.6 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na terceira abordagem, com o uso do algoritmo de *histogram matching*. 67
Figura 5.7 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na quarta abordagem, sem o uso do algoritmo de *histogram matching*. 70
Figura 5.8 Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na quarta abordagem, com o uso do algoritmo de *histogram matching*. 70
Figura 5.9 a) Tumor HGG classificado corretamente; b) Tumor HGG classificado como LGG; c) Tumor LGG classificado corretamente; d) Tumor LGG classificado como HGG. 72

Lista de tabelas

Tabela 3.1 Correspondência entre os diferentes contextos 28
Tabela 3.2 Exemplos de funções de *kernel* 37

Tabela 4.1 Definições dos parâmetros de cada algoritmo variados entre os diferentes treinamentos. O nome de cada parâmetro corresponde aos nomes apresentados nas funções da biblioteca Scikit-learn. 58

Tabela 5.1 Tabela mostrando os melhores resultados obtidos na primeira abordagem utilizando o SVM. As métricas de desempenho são apresentadas em conjunto com os desvios-padrão correspondentes. 61
Tabela 5.2 Tabela mostrando os melhores resultados obtidos na primeira abordagem utilizando o RF. 61
Tabela 5.3 Tabela mostrando os melhores resultados obtidos na primeira abordagem utilizando o MLP. 62
Tabela 5.4 Tabela mostrando os melhores resultados obtidos na segunda abordagem utilizando o SVM. 64
Tabela 5.5 Tabela mostrando os melhores resultados obtidos na segunda abordagem utilizando o RF. 64
Tabela 5.6 Tabela mostrando os melhores resultados obtidos na segunda abordagem utilizando o MLP. 66
Tabela 5.7 Tabela mostrando os melhores resultados obtidos na terceira abordagem utilizando o SVM. 66
Tabela 5.8 Tabela mostrando os melhores resultados obtidos na terceira abordagem utilizando o RF. 68
Tabela 5.9 Tabela mostrando os melhores resultados obtidos na terceira abordagem utilizando o MLP 68
Tabela 5.10 Tabela mostrando os melhores resultados obtidos na quarta abordagem utilizando o SVM 69
Tabela 5.11 Tabela mostrando os melhores resultados obtidos na quarta abordagem utilizando o RF 69
Tabela 5.12 Tabela mostrando os melhores resultados obtidos na quarta abordagem utilizando o MLP 71
Tabela 5.13 Comparação de resultados para o problema de classificação de gliomas utilizando atributos *radiomics* em dados de MRI. 73

Lista de Abreviaturas

MRI – *Magnetic Resonance Imaging*
HGG – *High Grade Gliomas*
LGG – *Low Grade Gliomas*
WHO – *World Health Organization*
CT – *Computed Tomography*
PET – *Positron Emission Tomography)*
LOOCV – *Leave-one-out Cross Validation*
ROC – *Receiver Operating Characteristic*
AUC – *Area Under the Curve*
CNN – *Convolutional Neural Network*
XGBoost – *gradient boosting tree*
mRMR – Máxima Relevância e Mínima Redundância
RF – *Random Forest*
SVM – *Support Vector Machines*
MLP – *Multilayer Perceptrons*
SNC – Sistema Nervoso Central
GBM – Glioblastoma Multiforme
ROI – *Region of Interest*
PCT – *Perfusion Computed Tomography*
IVIM – *Intravoxel Incoherent Motion*

1
Introdução

Gliomas são tumores do sistema nervoso central que se originam nas células gliais, células do tecido nervoso que auxiliam na manutenção e proteção dos neurônios. Estão entre os tipos de tumores cerebrais primários - ou seja, tumores que se originam no cérebro - malignos mais comuns e apresentam alta taxa de mortalidade entre homens e mulheres [1]. Eles são comumente classificados entre gliomas de baixo grau (LGG, *low grade gliomas*) e de alto grau (HGG, *high grade gliomas*). Gliomas de baixo grau compreendem os graus I e II definidos pela Organização Mundial de Saúde (WHO, *World Health Organization*), enquanto que os de alto grau incluem os graus III e IV [2, 3].

Gliomas de baixo grau tendem a apresentar um melhor prognóstico para o paciente, sendo os de grau I geralmente benignos, podendo em alguns casos ser retirados através de cirurgia, sem a necessidade de tratamento por quimioterapia [4]. Gliomas de grau II, apesar de normalmente também apresentarem um melhor prognóstico, eventualmente se desenvolvem para os estágios mais avançados da doença (graus III e IV), considerados extremamente agressivos e apresentando baixa taxa de sobrevivência [5].

A identificação e caracterização precoce do tumor por um especialista é fundamental para o direcionamento do tratamento, tanto para tentar impedir que o tumor avance para graus superiores, quanto para buscar oferecer um melhor prognóstico para pacientes que já se encontram nos estágios mais avançados da doença. Os principais meios utilizados para a realização do diagnóstico de gliomas envolvem a realização de exames de imagem, como imagens por ressonância magnética (MRI, *magnetic resonance imaging*), tomografia computadorizada (CT, *computed tomography*) e tomografia por emissão de pósitrons (PET, *positron emission tomography*), e exames de biópsia, que consistem na retirada de uma porção do tumor para posterior análise em microscópio [6].

Com o avanço cada vez mais notável de algoritmos de *machine learning* e a crescente disponibilidade de dados médicos, muito estudos que combinam métodos computacionais e conhecimentos da medicina têm surgido em uma tentativa de criar técnicas que auxiliem no diagnóstico e tratamento de doenças. Um dos tópicos que mais tem ganhado destaque nesse contexto está relacionado à análise e interpretação de imagens médicas e deu origem a um

procedimento chamado *radiomics*, cuja premissa é a utilização de atributos (*features*) quantitativos extraídos de imagens médicas digitalizadas como eficientes indicadores de características genéticas e fisiopatológicas, capazes de auxiliar nas decisões tomadas por especialistas [7]. Além de oferecer informações que muitas vezes não são percebidas pelo olho humano, tais atributos extraídos por métodos computacionais podem oferecer uma excelente alternativa aos procedimentos invasivos atualmente utilizados para a confirmação de determinados diagnósticos. Mais ainda, o uso do *radiomics* pode otimizar o tempo que especialistas gastam para analisar o grande volume de exames que recebem.

1.1 Objetivos

Este trabalho se propõe a trazer conceitos do campo da biologia para a área de imagens médicas, introduzindo atributos *radiomics* ainda não utilizados no problema de classificação de gliomas entre LGG e HGG. Através das imagens por ressonância magnética fornecidas pelo desafio BraTS 2018 (*Multimodal Brain Tumor Segmentation Challenge* 2018) [9, 10], são calculados índices de biodiversidade e de diversidade filogenética, que serão utilizados como atributos em um algoritmo de classificação.

1.2 Estrutura do Trabalho

O presente trabalho está dividido em 5 capítulos. No primeiro capítulo, introduzimos o problema estudado, discutindo sua importância e apresentando os objetivos aqui pretendidos. No segundo discutimos sobre trabalhos relacionados, expondo as metodologias propostas por cada um. O terceiro capítulo foi dedicado inteiramente à exposição de toda a fundamentação teórica necessária para a compreensão do problema estudado e da metodologia proposta para abordá-lo. No quarto capítulo falamos sobre os materiais e métodos utilizados. No quinto capítulo apresentamos os resultados obtidos, discutindo-os e comparando-os com aqueles já publicados por outros estudos. Por fim, no sexto e último capítulo, apresentamos a conclusão e uma discussão sobre os possíveis trabalhos futuros.

2
Trabalhos Relacionados

Existem diversos trabalhos que se dedicaram ao estudo de gliomas. Claus et al. (2015) [5] defenderam o uso de características histológicas, marcadores tumorais, informações genotípicas e resultados de tratamentos cirúrgicos para aprimorar os tratamentos direcionados para indivíduos com LGG, atentando para o fato de que gliomas de baixo grau, apesar de geralmente apresentarem melhores prognósticos, costumam se desenvolver para graus superiores. Charlotte et al (2017) [65] conduziu um estudo para estimar o tempo de sobrevida livre de progressão de pacientes com glioblastoma multiforme (GBM) - tipo de glioma de alto grau com maior incidência - utilizando dados demográficos e clínicos coletados rotineiramente de 50 pacientes submetidos a tratamentos de radioterapia e quimioterapia.

Em relação à classificação de gliomas, diversas abordagens já foram propostas, utilizando diferentes tipos de dados. Ellika et al. (2007) [67] propuseram o uso de imagens de perfusão por tomografia computadorizada (PCT, *Perfusion Computed Tomography*) de 19 pacientes (14 HGG e 5 LGG) para o problema de classificação. Tais dados fornecem informações de vascularidade e angiogênese do tumor e foram capazes de obter valores de sensibilidade e especificidade de 0,929 e 0,100, respectivamente. Togao et al. (2015) [68] analisaram o desempenho no uso de dados do movimento incoerente intravoxel (IVIM, *Intravoxel Incoherent Motion*) extraídos de imagens por ressonância magnética para classificar o grau de gliomas, obtendo valores de sensibilidade, especificidade e área abaixo da curva ROC (AUC) de 0,966, 0,812 e 0,950, respectivamente. Embora tais estudos tenham apresentado resultados promissores, os tipos de dados utilizados são mais complexos e dificilmente encontrados para serem utilizados em pesquisas independentes.

Uma das principais ideias defendidas por abordagens *radiomics* é a criação de bases de dados de fácil acesso que facilitem pesquisas nas mais diversas áreas utilizando imagens médicas. Muitos trabalhos já ofereceram metodologias *radiomics* para tratar eficientemente problemas da medicina. Hui et al. (2016) [21] utilizaram uma abordagem *radiomics* para estimar o risco de recorrência de câncer de mama utilizando imagens por ressonância magnética, mostrando que atributos como o tamanho e a textura da lesão apresentaram

alto poder discriminativo na previsão do prognóstico do paciente, com AUC de 0,880. Jiangwei et al. (2017) [23] investigaram o uso de redes neurais para criar atributos *radiomics* para a predição da sobrevida geral entre pacientes com glioblastoma multiforme. Combinando os atributos gerados com dados clínicos, o modelo proposto obteve um índice de concordância (C-*index*) de 0,729. Outro trabalho que propôs novos atributos *radiomics* foi o publicado por Neto et al. (2018) [12], que utilizou atributos baseados em índices de diversidade filogenética, calculados a partir de imagens por ressonância magnética, para classificar células não pequenas de câncer de pulmão. Os resultados obtidos foram extremamente promissores, apresentando uma acurácia de 0,985 e um valor de AUC de 0,999.

Muitos estudos recentes utilizaram atributos *radiomics* baseados em imagens por ressonância magnética para lidar com o problema de classificação de gliomas, fundamentando-se na ideia de que gliomas de graus mais elevados tendem a apresentar uma estrutura mais heterogênea [8] e utilizando a extração de características de textura e forma para mensurar tal heterogeneidade [13, 14, 15, 16, 17, 18].

Um estudo conduzido por Zhang et al. (2017) [15] comparou o desempenho de diferentes técnicas de *machine learning* na classificação de gliomas entre LGG e HGG e na diferenciação entre os graus II, III e IV, utilizando dados de ressonância magnética (T1ce, FLAIR, ASL, DWI e DCE) de 120 indivíduos. Através de 25 algoritmos de aprendizado e 8 métodos de seleção de atributos, eles analisaram a eficácia de cada combinação de algoritmos utilizando uma validação cruzada *leave-one-out* (LOOCV, *Leave-one-out cross validation*). Ao final do estudo, verificou-se que o classificador que apresentou melhor desempenho foi o *support vector machine* (SVM), que combinado à técnica de eliminação recursiva de atributos (*recursive feature elimination*, RFE) obteve acurácia de 0,960 e AUC de 0,960 na classificação entre LGG e HGG.

Outro estudo publicado por Cho et al. (2017) [16] criou um modelo de classificação de gliomas utilizando atributos *radiomics* baseados em histograma, forma e matriz de coocorrência de níveis de cinza, utilizando a base de dados de MRI (T1, T1ce, T2 e FLAIR) fornecida pelo desafio BraTS 2015. Através de uma validação cruzada 10-*folds* foi avaliada a eficácia de um modelo de classificação empregando uma regressão logística com seleção de atributos, atingindo valores de acurácia e AUC de 0,898 e 0,887, respectivamente. Utilizando a mesma base de dados, um trabalho proposto por Chen et al (2018) [17] ofereceu um sistema automático para classificação de gliomas, aplicando técnicas de *deep learning* para realizar uma segmentação automática das lesões anterior à extração de atributos *radiomics*. Deste modo, foi treinada uma

rede neural convolucional 3D (CNN, *convolutional neural network*) através das anotações das lesões fornecidas pela base de dados e, em seguida, a partir das regiões segmentadas, foram extraídos os atributos *radiomics* para servirem como entrada do modelo de classificação. Utilizando o algoritmo de aprendizado *gradient boosting tree* (XGBoost) em conjunto com uma etapa de seleção que resultou na escolha de 25 atributos, o modelo de classificação gerado apresentou uma acurácia de 0,913 e AUC de 0,960.

Um artigo recentemente publicado por Cho et al. (2018) [18] apresentou um estudo de classificação de gliomas utilizando a base de dados de MRI fornecida pelo desafio BraTS 2017, mesma base oferecida pelo desafio em 2018 e utilizada em nosso estudo. Considerando as segmentações de três regiões intratumorais oferecidas pelos dados anotados, foram extraídos 468 atributos *radiomics* e, através da técnica de seleção de Máxima Relevância e Mínima Redundância (mRMR), foram selecionados 5 atributos. Utilizando uma validação cruzada 5-*fold* para avaliar os algoritmos de classificação testados, o melhor resultado foi obtido através do algoritmo *random forest* (RF), com valores de acurácia e AUC de 0,888 e 0, 921, respectivamente.

3
Fundamentação Teórica

Neste capítulo são apresentados os conceitos necessários para a compreensão do presente trabalho.

3.1
Gliomas

O sistema nervoso central (SNC) do corpo humano é formado pelo encéfalo e pela medula espinhal e é responsável pela troca de informações necessárias para o correto funcionamento do organismo [19]. Nele estão presentes as chamadas células gliais, cuja principal função é auxiliar no desenvolvimento, manutenção e reparo do sistema nervoso através de suporte fornecido aos neurônios [20]. Glioma é um tipo de tumor do SNC que se origina nas células gliais e está entre os tipos de tumores cerebrais mais comuns. Conforme já foi dito anteriormente, seguindo as definições propostas pela WHO, os gliomas podem ser classificados entre LGG (graus I e II) e HGG (graus III e IV).

Diferentes tipos de glioma exigem diferentes estratégias de tratamento. Em alguns casos, indivíduos com gliomas de baixo grau, ao receberem o diagnóstico, ficam apenas em observação, sendo monitorados frequentemente através de exames de MRI para que o desenvolvimento da lesão seja acompanhado. Estima-se que em torno de 50% dos pacientes que apresentam gliomas de baixo grau precisam de cirurgia dentro de 2 a 3 anos após o início do acompanhamento médico [29]. Quando a estratégia de monitoramento não é uma opção, as alternativas de tratamento geralmente envolvem cirurgia, para a ressecção parcial ou total do tumor, radioterapia e quimioterapia.

Tendo em vista a importância em se definir um tratamento específico para cada caso, a identificação precoce do tipo de glioma com que se está lidando é de extrema importância. Certos tipos mais agressivos, como o glioblastoma multiforme, possuem baixa expectativa de vida, apresentando uma taxa de sobrevida média abaixo de 2 anos após o início do tratamento [30]. Sendo a ressonância magnética o principal exame utilizado atualmente para a identificação de gliomas, estabelecer um mecanismo que auxilie especialistas no processo de diagnóstico através de dados de MRI, de forma que tanto o tempo, quanto a acurácia da análise de exames sejam otimizados, pode ajudar

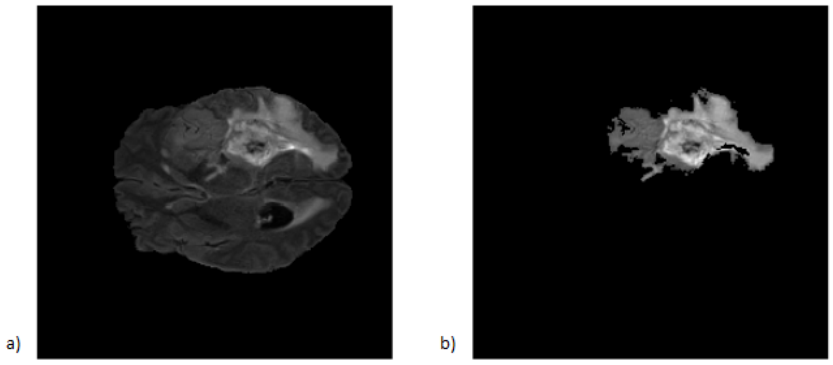

Figura 3.1: Exame por ressonância magnética de um paciente com glioma de alto grau, obtido através da base de dados do desafio BraTS'18. a) Fatia do exame mostrando o cérebro com a lesão; b) Imagem mostrando apenas a estrutura da lesão.

no direcionamento do tratamento e possivelmente oferecer um prognóstico mais favorável para o paciente. A Figura 3.1 mostra um exemplo de glioma de alto grau.

3.2 Radiomics

De modo geral, o termo *radiomics* refere-se ao processo de extração de atributos quantitativos de imagens médicas digitalizadas e o uso de tais atributos para auxiliar na tomada de decisão por especialistas. Em conjunto com outros dados relacionados aos pacientes, as características obtidas com o uso de uma abordagem *radiomics* podem ser de grande utilidade para a formulação e verificação de hipóteses relacionadas à patologia estudada [7]. Uma das principais vantagens associadas ao processo é a possibilidade de se gerar atributos mais complexos que sejam capazes de capturar propriedades importantes presentes nas imagens analisadas, sobretudo imagens volumétricas obtidas através de exames de tomografia e ressonância magnética.

Segundo Kumar et al. (2012) [63], a abordagem *radiomics* pode ser dividida em cinco diferentes processos principais: (1) Aquisição e reconstrução de imagens; (2) Segmentação e renderização de imagens; (3) Extração e qualificação de atributos; (4) Construção e compartilhamento de bases de dados; e (5) Posteriores análises em cima dos dados adquiridos. A Figura 3.2 ilustra tais etapas e os desafios associados a cada uma. Um dos objetivos da introdução desta metodologia é unificar e centralizar procedimentos inerentes às pesquisas relacionadas a imagens médicas, de modo que facilite a extração e o uso de informações relevantes dos dados utilizados.

Em relação ao processo de aquisição e reconstrução de imagens, um dos principais desafios apontados por Kumar et al. é a variação existente entre um

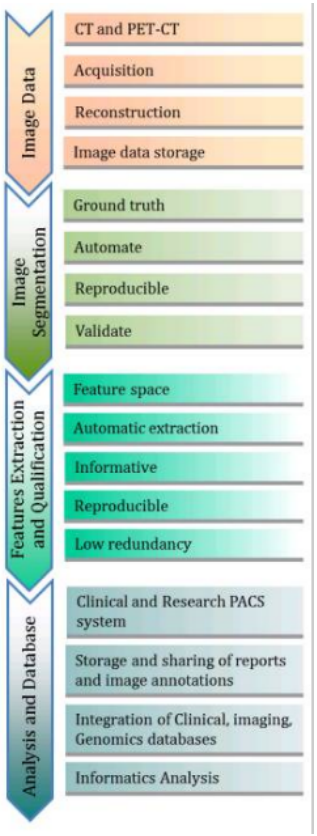

Figura 3.2: Imagem extraída de Kumar et al. (2012) [63], ilustrando as etapas associadas à abordagem *radiomics* e seus desafios.

mesmo tipo de exame realizado em diferentes instituições. Características como tipos de aparelhos, resolução das imagens, a posição do paciente ao realizar o exame e os algoritmos de reconstrução utilizados por especialistas para formar a imagem do exame final, podem variar muito de uma instituição para outra, o que dificulta a comparação de exames que envolvem parâmetros clínicos semelhantes, como o estágio de determinada doença em diferentes pacientes, por exemplo.

Para a extração dos atributos, a ideia é conseguir oferecer atributos que possam ser extraídos de diferentes tipos de imagens e utilizados eficientemente em diversos contextos, de tal forma que as análises resultantes destes atributos sejam robustas o suficiente para serem aplicadas por especialistas. Por exemplo, modelos que sejam capazes de realizar predições automáticas apenas através de imagens e atributos extraídos delas estão atualmente entre os principais objetivos do *radiomics*.

Outra importante questão é o incentivo à criação e centralização de grandes bases de dados cujo acesso seja facilitado para qualquer estudioso e especialista que queira conduzir pesquisas relacionadas a imagens médicas. Enquanto não houver um processo padronizado para a obtenção das imagens, Kumar et al. defenderam a ideia de que uma grande base de dados pode

compensar a heterogeneidade resultante dos diferentes processos de aquisição dos dados.

Atualmente, a área da medicina que mais se beneficia com o uso do *radiomics* é a oncologia. Tendo sido a propulsora de abordagens *radiomics*, ela utiliza os volumes de interesse definidos pelos tumores, ou subregiões deles (*habitats*), para extrair atributos relacionados a intensidade de histograma, tamanho e formato, por exemplo. Muitos estudos, como os publicados por Kuo et al. (2007) [19], Wibmer et al. (2015) [20], Hui et al. (2016) [21], Jiangwei et al. (2017) [23] e Neto et al. (2018) [12], comprovaram que tais características são capazes de fornecer informações relevantes para a detecção, classificação, diagnóstico, prognóstico, monitoramento ou até mesmo direcionamento de tratamento de diversos tipos de tumores.

O uso do *radiomics* por especialistas pode significar não apenas a otimização do tempo gasto na análise de exames de imagens, mas também o aumento da precisão com que tais análises são feitas. Com a crescente disponibilidade de dados médicos digitalizados, já existem muitos estudos fornecendo soluções computacionais satisfatórias para diversos problemas existentes na medicina e o presente trabalho se propõe a contribuir ainda mais para esta área em ascensão.

3.3 Imagens por Ressonância Magnética

O exame de imagem por ressonância magnética (*Magnetic Resonance Imaging*, MRI) é atualmente um dos principais métodos utilizados para a identificação, diagnóstico e monitoramento de gliomas [22]. A principal vantagem em se utilizar este tipo de exame está associada à capacidade de gerar imagens com alta resolução de tecidos do corpo, sem que haja emissão de radiação ionizante (raios-X) [24]. Através de um campo magnético e de ondas de rádio, o aparelho de ressonância magnética capta sinais emitidos pelo corpo que são posteriormente utilizados por métodos computacionais para formar imagens em escalas de cinza das regiões analisadas [27]. As diferentes propriedades inerentes aos tecidos anatômicos apresentam contrastes distintos nas imagens resultantes do exame, permitindo uma análise detalhada de suas estruturas internas. O escaneamento do corpo é feito através de "fatias" sequenciais, que podem ser utilizadas para formar imagens bidimensionais e tridimensionais [28].

Através de parâmetros previamente definidos, é possível controlar a forma como é realizada a emissão de sinais durante um exame de MRI e, consequentemente, modificar as imagens finais geradas [25]. Cada combinação

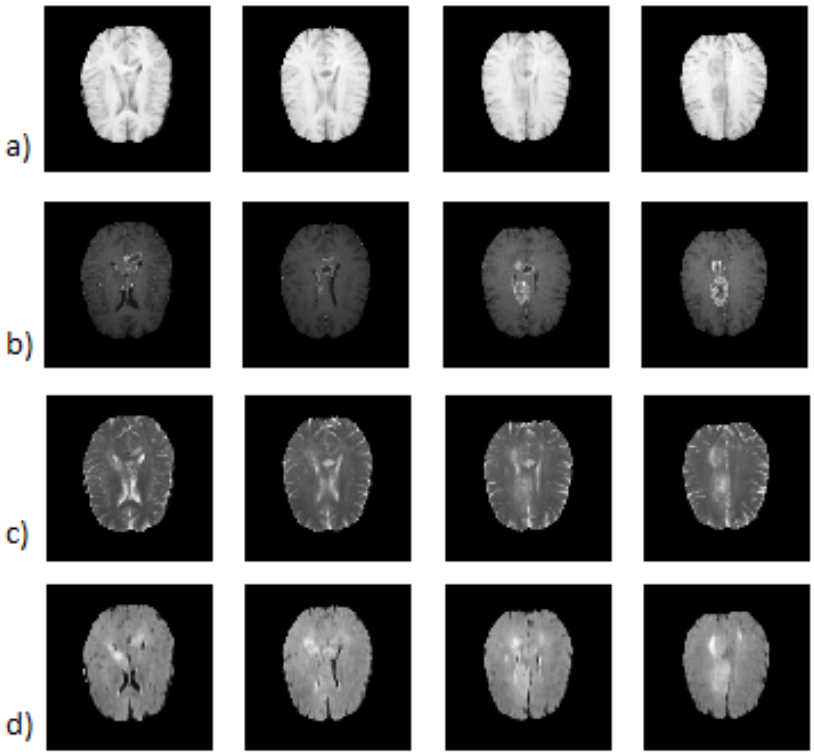

Figura 3.3: Exemplos de fatias bidimensionais extraídas de imagens volumétricas obtidas através das quatro modalidades de MRI utilizadas no presente estudo [58, 59]. a) T1; b) T1ce; c) T2; d) FLAIR.

de parâmetros define uma modalidade de exame distinta, sendo as modalidades T1, T1ce, T2 e FLAIR as mais comumente utilizadas para a análise inicial de gliomas, cada uma fornecendo informações biológicas distintas [9, 26]. A modalidade T1ce corresponde à modalidade T1 após a injeção de Gadolíneo no paciente, substância de contraste que permite uma melhor visualização de estruturas internas ao núcleo do tumor. Os dados utilizados no presente trabalho correspondem a imagens volumétricas obtidas através dessas quatro modalidades e a Figura 3.3 ilustra alguns exemplos.

3.4 Processamento Digital de Imagens

Uma imagem tridimensional em tons de cinza pode ser definida como uma função $f : U \subset \mathbb{R}^3 \rightarrow C \subset \mathbb{R}$ que a cada ponto (x, y, z) do subconjunto U contido no espaço $\mathbb{R}^3$ associa um valor $c = f(x, y, z)$ pertencente ao subconjunto C contido em $\mathbb{R}$, tal que c indica a intensidade de cinza da imagem naquele ponto. Quando trabalhamos com imagens digitais, utilizamos quantidades discretas para representar as coordenadas x, y, z e a intensidade c correspondente [56, 57]. Como se trata do caso tridimensional, cada elemento

que compõe a imagem digital é denominado voxel.

Em imagens volumétricas obtidas por aparelhos de ressonância magnética, cada voxel corresponde à intensidade dos sinais emitidos pelos tecidos do corpo em resposta às ondas eletromagnéticas de rádio impulsionadas pelo aparelho. Os dados de MRI utilizados em nosso estudo correspondem a imagens digitais 3D em níveis de cinza e, portanto, estaremos lidando aqui com o conceito de voxel.

Podemos chamar de processamento digital de imagens qualquer método computacional que possui imagens como entrada e saída, ou que, a partir de uma imagem, é capaz de extrair informações presentes nela [56]. Em nosso trabalho, lidamos com ambos os tipos de processamento. No primeiro caso, utilizando uma técnica de pré-processamento chamada *histogram matching*, ou especificação de histograma, para tentar contornar desvantagens associadas à aquisição dos dados, que serão discutidas com maiores detalhes no Capítulo 4. Já no segundo caso, aplicando métodos computacionais convencionais para calcular os índices de biodiversidade (Seção 3.5) e diversidade filogenética (Seção 3.6) sobre regiões de interesse exraídas da imagem.

3.4.1 Histogram Matching

A técnica de especificação de histograma (*histogram matching*) possui como objetivo modificar uma imagem para que o histograma associado a ela passe a apresentar um formato previamente especificado. Podemos definir o histograma de uma imagem em tons de cinza, cujos valores de pixels estão inseridos em um intervalo $[0, L-1]$, como uma função discreta $h(r_k) = n_k$, em que r_k é o k-ésimo tom de cinza do intervalo $[0, L-1]$ e n_k é a quantidade de pixels que apresentam o valor r_k. Para obter uma estimativa da probabilidade $p(r_k)$ de ocorrência do valor r_k, podemos construir o histograma normalizado, definido por $p(r_k) = n_k/N$, com N sendo a quantidade total de pixels na imagem [56].

A especificação de histograma é muito utilizada no contexto de imagens médicas, quando se tem exames obtidos através de diferentes aparelhos e instituições, o que pode gerar grandes variações nas intensidades dos pixels das imagens resultantes. No presente estudo, utilizamos o método como uma etapa de pré-processamento, em que foram selecionados os exames de MRI de um determinado indivíduo para servirem como referência. Assim, para cada modalidade de exame, calculamos o histograma da imagem volumétrica do indivíduo escolhido e em seguida aplicamos o método de especificação de histograma sobre as imagens dos demais pacientes.

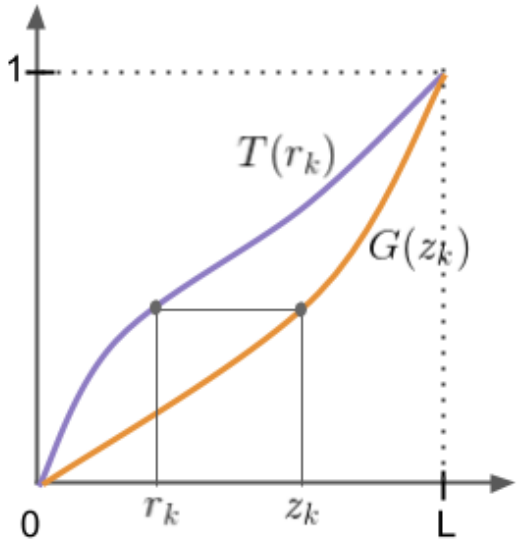

Figura 3.4: Gráfico ilustrativo mostrando a correspondência entre duas funções de distribuição acumulada.

Diante do exposto, considere uma imagem com valores de cinza r_k, tal que $k = 0, 1, 2, ..., L-1$. Suponha que queremos modificar essa imagem de tal modo que ela passe a apresentar um histograma $p_z(z)$ específico, calculado sobre outra imagem de referência. A ideia é utilizar a função de distribuição acumulada (fda) para encontrar as intensidades de cinza com distribuições equivalentes entre as duas imagens e, a partir dessa informação, mapear cada intensidade r_k para o valor z_k correspondente (Figura 3.4).

Primeiramente, vamos definir a função $T(r_k)$ que calcula a fda associada à imagem de entrada:

$$T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{N} \ , \quad k = 0, 1, 2, ..., L-1 \tag{3-1}$$

Analogamente, tendo conhecimento do histograma $p_z(z_k)$ que desejamos alcançar, podemos definir a fda $G(z_k)$ da imagem de referência como

$$G(z_k) = \sum_{i=0}^{k} p_z(z_i) = \sum_{i=0}^{k} \frac{n_i}{N} \ , \quad k = 0, 1, 2, ..., L-1 \tag{3-2}$$

Queremos encontrar as intensidades z_k da imagem de referência que satisfazem a igualdade

$$G(z_k) = T(r_k) \tag{3-3}$$

Logo, assumindo que a inversa de $G^{-1}(z_k)$ existe, podemos calcular os valores z_k através de

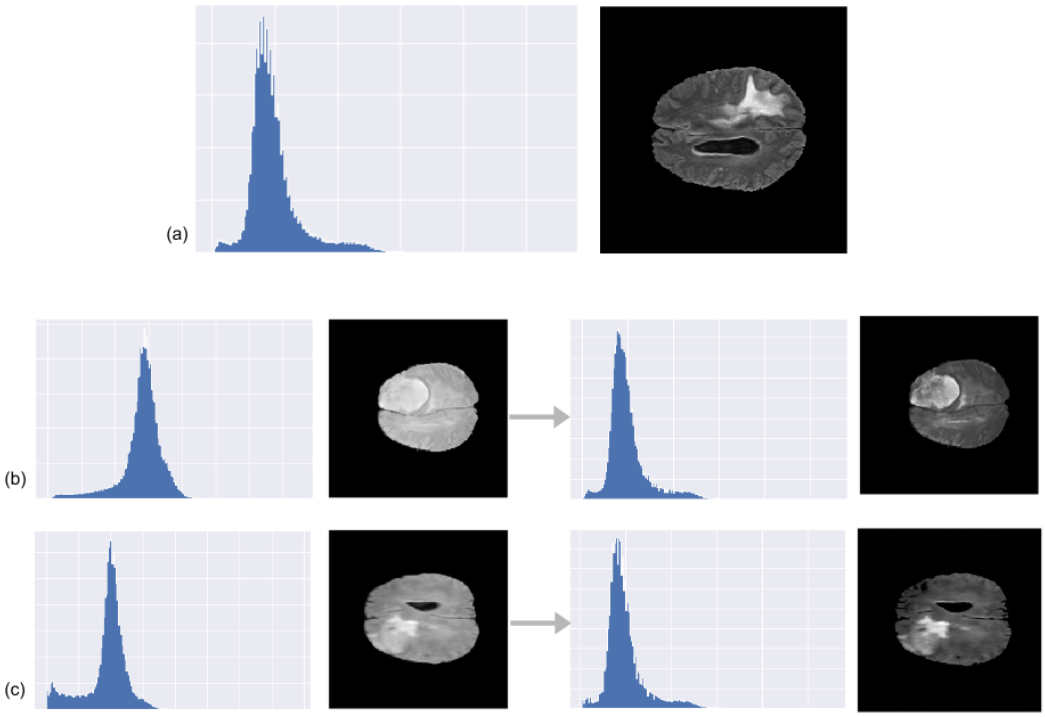

Figura 3.5: Exemplos de imagens e histogramas extraídos de nossa base de dados antes e depois da aplicação do método de especificação de histograma. (a) Histograma da imagem tomada como referência e centésima fatia da imagem; (b) Histograma de uma imagem I_1 e exemplo de fatia, antes e depois da aplicação do método, respectivamente; (c) Histograma de uma imagem I_2 e exemplo de fatia, antes e depois da aplicação do método, respectivamente.

$$z_k = G^{-1}(T(r_k)) \tag{3-4}$$

Vale ressaltar que, na prática, os valores obtidos com a Equação 3-4 são aproximações das intensidades apresentadas pela imagem que define o histograma de referência. A Figura 3.5 ilustra alguns exemplos de histogramas construídos através dos nossos dados, antes e depois da aplicação do método.

3.4.2 Extração de Regiões de Interesse

Uma região de interesse (*region of interest*, ROI) de uma imagem corresponde a um grupo de pixels sobre o qual se realiza uma análise ou procedimento específico. Quando se tem informação sobre a região exata que se deseja analisar, podemos utilizar máscaras para extrair uma ROI, que são imagens com mesma dimensão da imagem original em que os valores dos pixels servem como filtros para se captar a região que se deseja. Por exemplo, no caso de uma máscara binária, podemos extrair a ROI atribuindo o valor 1 a todos os pixels que

Figura 3.6: a) Exemplo de fatia de um exame de MRI; b) Exemplo de máscara utilizada para extração do tumor da fatia analisada; c) Resultado da extração.

desejamos analisar e 0 àqueles que queremos descartar.

A Figura 3.6 mostra um exemplo de máscara binária utilizada em nosso trabalho para extrair os voxels correspondentes aos tumores em cada exame de ressonância magnética. Os valores dos voxels extraídos foram submetidos a distintas análises e utilizados como entrada de funções para calcular os índices utilizados na metodologia proposta.

3.5 Índices de Biodiversidade

Índices de biodiversidade são muito utilizados para se medir a diversidade de espécies presentes em determinada região. De acordo com Morris et al. (2014) [64], a biodiversidade representa a variedade e heterogeneidade de organismos em qualquer nível de organização de seres vivos, desde moléculas até ecossistemas. Levando em conta a noção de que uma maior heterogeneidade intratumoral geralmente está associada a um pior prognóstico [7] e que, portanto, gliomas de mais alto grau tendem a apresentar uma estrutura mais heterogênea, fez-se uma adaptação de índices da biologia para o contexto de imagens [47, 11, 12] de modo a tentar utilizar a ideia de biodiversidade para capturar e mensurar as diferentes estruturas características de cada tumor.

Utilizando diferentes índices de biodiversidade, o objetivo é mostrar que lesões com estruturas mais heterogêneas apresentam uma maior diversidade se comparadas a estruturas mais homogêneas. A adaptação dos índices para o contexto de imagens médicas foi baseada no trabalho publicado por Silva et al. (2016) [47], em que cada voxel corresponde a um indivíduo e o valor assumido por ele, à sua espécie. Sendo assim, o conjunto de voxels que compõem a lesão define a comunidade estudada. O raciocínio utilizado está esquematizado na Tabela 3.1.

A seguir são descritos os 12 índices de biodiversidade [48, 49, 50] utilizados como atributos em nosso estudo.

1. Quantidade de espécies (S) : Corresponde à quantidade de diferentes

Tabela 3.1: Correspondência entre os diferentes contextos

Biologia	**Metodologia proposta**
Comunidade	Volume de interesse nos dados de MRI
Indivíduo	Voxel
Espécie	Valor do voxel

espécies presentes na comunidade estudada.

2. Quantidade de indivíduos (N) : Número total de indivíduos na comunidade.

3. Índice de Margalef (D_{Mg}) : É uma medida de riqueza de espécies baseada na hipótese de que existe uma relação linear entre a quantidade total de espécies e o logaritmo da quantidade de indivíduos. É dado por

$$D_{Mg} = \frac{S-1}{lnN} \tag{3-5}$$

4. Índice de Menhinck (D_{Mn}) : Também é uma medida de riqueza de espécies, mas apresenta menor variação entre amostras de diferentes comunidades, se comparado ao índice anterior.

$$D_{Mn} = \frac{S}{\sqrt{N}} \tag{3-6}$$

5. Índice de Odum (D_O) : Corresponde à quantidade de espécies por mil indivíduos multiplicada pela quantidade total de indivíduos.

$$D_O = \frac{S\ N}{1000} \tag{3-7}$$

6. Índice de Hulbert (PIE) : É a probabilidade de que dois indivíduos escolhidos ao acaso, em uma amostra aleatória de determinada comunidade, pertençam a espécies diferentes.

$$PIE = (\frac{N}{N-1})(1 - \sum_{i=1}^{S} p_i^2) \tag{3-8}$$

em que p_i é a fração de indivíduos dentro da amostra pertencentes à espécie i, ou a probabilidade de escolher um indivíduo ao acaso e ele pertencer à espécie i.

7. Índice de McNaughton (I) : Corresponde à porcentagem das duas espécies com maior quantidade de indivíduos na amostra da comunidade estudada.

$$I = \frac{n_1 + n_2}{N} 100 \tag{3-9}$$

Sendo n_1 e n_2 a quantidade de indivíduos pertencentes à primeira e à segunda espécies mais populosas, respectivamente.

8. Índice de Simpson (D_S) : Indica a probabilidade de que dois indivíduos escolhidos aleatoriamente em uma amostra pertençam à mesma espécie.

$$D_S = \frac{\sum_{i=1} n_i(n_i - 1)}{N(N-1)} \tag{3-10}$$

Em que n_i é a quantidade de indivíduos pertencentes à espécie i.

9. Índice de Shannon (H') : Mede o grau de incerteza médio ao se tentar prever a qual espécie pertence determinado indivíduo escolhido aleatoriamente em uma amostra.

$$H' = -\sum_{i=1}^{S} \frac{n_i}{N}\ ln(\frac{n_i}{N}) \tag{3-11}$$

10. Índice de Uniformidade (Evenness index, E) : Mede o quão uniforme é a distribuição de indivíduos entre as espécies existentes na amostra. Seu valor máximo é atingido quando todas as espécies possuem a mesma quantidade de indivíduos.

$$E = \frac{H'}{H'_{max}} \tag{3-12}$$

Em que H'_{max} corresponde ao valor máximo do índice de Shannon.

11. Índice de Redundância (Redundancy Index, R) : A redundância é inversamente proporcional à quantidade de espécies, podendo ser pensada como uma medida do quanto a abundância da amostra pode ser explicada por uma ou mais espécies.

$$R = \frac{H'_{max} - H'}{H'_{max} - H'min} \tag{3-13}$$

Sendo H'_{min}o valor mínimo do índice de Shannon.

12. Índice geométrico (B) : O índice compara um sistema de S espécies com um sistema de $S+k$ espécies, usando para tal comparação a relação entre o volume de duas esferas de raio r existentes em espaços de dimensão S e $S + k$, respectivamente.

$$B_k(S, r) = \frac{V_s(r)}{V_{s+k}(r)} = \alpha_k(S)\beta_k(r), \quad r \neq 0 \tag{3-14}$$

$$V_s(r) = \frac{\pi^{S/2}}{\Gamma((S/2)+1)} r^S \tag{3-15}$$

Em que $V_s(r)$ é o volume de uma esfera com raio r em um espaço S-dimensional e $\Gamma(x)$ é a função Gamma com as seguintes propriedade: $\Gamma(u+1) = u\Gamma(u)$, $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(3/2) = (1/2)\sqrt{\pi}$, $\Gamma(5/2) = (3/4)\sqrt{\pi}$ etc. $\alpha_k(S)$ e $\beta_k(r)$ são definidos por

$$\alpha_k(S) := \frac{\Gamma((S+k+2)/2)}{\pi^{k/2}\Gamma((S+2)/2)}, \qquad \beta_k(r) := \frac{1}{r^k}, \quad r = \sqrt{\sum_{i=1}^{S} p_i^2}$$

3.6 Índices de Diversidade Filogenética

Ao analisar a diversidade de espécies em determinada comunidade, é importante notar que, além de existirem diferentes espécies, há também uma noção de distância evolutiva entre elas, cujo principal objetivo é mensurar o quão diferentes duas espécies são entre si. Uma das formas de se calcular essa distância evolutiva é através das relações filogenéticas que descrevem a evolução das espécies ao longo do tempo, tentando estabelecer conexões entre cada espécie e seus ancestrais. A representação dessas relações pode ser feita através das chamadas árvores filogenéticas, em que as ramificações representam a forma como espécies, ou grupo de espécies, evoluíram a partir de seus ancestrais comuns. Assim, pode-se dizer que duas espécies possuem uma curta distância filogenética se elas possuem um ancestral comum mais recente, ou uma longa distância caso contrário.

Inspirado por resultados publicados por estudos recentes [11, 12] que utilizaram índices filogenéticos em problemas envolvendo a classificação de nódulos pulmonares, nosso estudo combina índices de diversidade filogenética aos índices de biodiversidade anteriormente descritos para serem usados como atributos no problema de classificação de gliomas entre HGG e LGG. O objetivo é utilizar tais índices para tentar medir o nível de heterogeneidade intratumoral.

Para capturar a noção de distância entre espécies em nosso contexto de imagens médicas, foram reproduzidos os mesmos procedimentos utilizados nos trabalhos citados anteriormente [47, 11, 12]. Utilizando as adaptações da Tabela 3.1 e um cladograma, como o indicado na Figura 3.7, representa-se através de nós terminais (ou folhas) cada espécie existente na região de interesse analisada e utiliza-se arestas para indicar as distâncias entre elas. Os nós internos representam os ancestrais comuns entre as espécies. Deste modo, a distância d_{ij} entre duas espécies i e j é dada pela quantidade de arestas existentes no menor caminho que vai do nó i ao nó j.

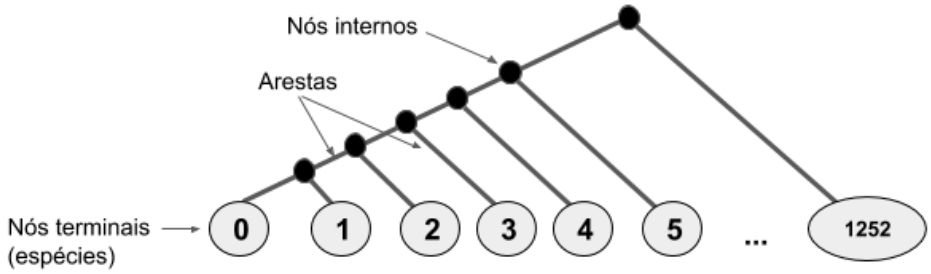

Figura 3.7: Exemplo de relações filogenéticas representadas por um cladograma. Neste exemplo, o valor máximo de voxel encontrado no volume de interesse é 1252.

A seguir são descritos os 5 índices de diversidade filogenética utilizados em nosso trabalho [51, 52, 53, 54, 55]:

1. Entropia quadrática intensiva (J) : Fornece a distância taxonômica média entre duas espécies selecionadas ao acaso.

$$J = \frac{\sum_{i=0}^{N} \sum_{j=0}^{N} d_{ij}}{S^2} \tag{3-16}$$

2. Entropia quadrática extensiva (F): É dada pela soma total das distâncias entre as espécies existentes.

$$F = \sum_{i=0}^{N} \sum_{j=0}^{N} d_{ij} \tag{3-17}$$

3. Distinção taxonômica média ($AvTD$): Representa a distância esperada entre dois indivíduos de diferentes espécies escolhidos ao acaso.

$$AvTD = \frac{\sum_{i=0}^{N} \sum_{j=0}^{N} d_{ij}}{\frac{S(S-1)}{2}} \tag{3-18}$$

4. Distinção taxonômica total (TTD): Corresponde à distinção taxonômica média somada sobre todas as espécies.

$$TTD = \frac{\sum_{i=0}^{N} \sum_{j=0}^{N} d_{ij}}{(S-1)} \tag{3-19}$$

5. Medida de pura diversidade (Dd): Equivale à distância entre uma espécie i e a espécie j_{min} mas próxima a ela, somada sobre todas as espécies.

$$Dd = \sum_{i=0}^{N} \sum_{j=0}^{N} d_{ij_{min}} \tag{3-20}$$

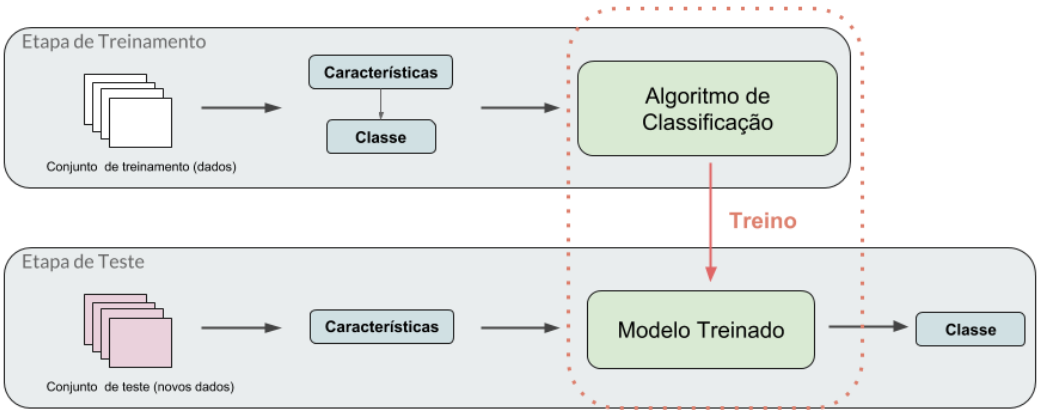

Figura 3.8: Etapas de um algoritmo de classificação.

3.7 Machine Learning e Algoritmos de Classificação

Machine learning, ou aprendizagem de máquina, é um processo de aprendizado automatizado em que um computador é programado para aprender determinada tarefa através de dados que são passados para ele. Sendo assim, um algoritmo de aprendizado recebe como entrada um conjunto de dados de treinamento e, através da experiência obtida com esses dados, apresenta como saída algum conhecimento capaz de desempenhar um procedimento desejado [31].

Entre os diversos tipos de problemas resolvidos por algoritmos de *machine learning*, temos os chamados problemas de classificação (Figura 3.8). Através de um conjunto de dados separados por classes previamente conhecidas - sendo, portanto, um método de aprendizado supervisionado - e contendo certa quantidade de atributos (*features*), utiliza-se um algoritmo de classificação para ser treinado sobre os dados recebidos, de modo que ele produza um modelo capaz de aprender as características inerentes a cada classe. Em seguida, com o modelo devidamente treinado, novos dados são fornecidos para testar a sua capacidade de classificar dados ainda não vistos.

Em nosso trabalho, foram testados três algoritmos de classificação distintos. Como estamos lidando com apenas duas classes (HGG e LGG), trata-se de um problema de classificação binária. Nas seções a seguir são descritos os algoritmos utilizados e outros métodos importantes para a implementação do processo de aprendizagem.

3.7.1 Support Vector Machine

Support vector machine (SVM), ou máquina de vetores de suporte, é um método de aprendizado muito utilizado em problemas de classificação. Uma

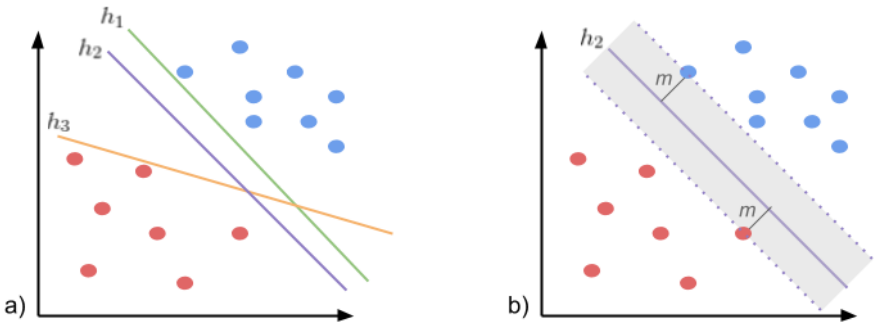

Figura 3.9: a) Exemplo de diferentes hiperplanos separando um conjunto de dados com duas classes linearmente separáveis. Neste caso, cada instância possui dois atributos (representados por cada eixo) e sua cor (azul ou vermelho) representa a classe correspondente. b) Hiperplano ótimo h_2 com distância m (margem) para as instâncias de cada classe mais próximas a ele. Tais instâncias são chamadas de vetores de suporte (*support vectors*).

de suas principais vantagens está associada ao seu alto poder de generalização [32], ou seja, à sua capacidade de obter bons resultados quando aplicado a novos dados. Dado um conjunto de dados de entrada associados a duas classes previamente conhecidas, o SVM tem como objetivo encontrar o hiperplano que separa da forma mais eficiente possível os vetores de instâncias entre as duas classes existentes. No caso linearmente separável, um hiperplano ótimo é aquele capaz de realizar tal partição encontrando a máxima distância entre ambas as classes.

A Figura 3.9-a) ilustra um exemplo de três hiperplanos (retas, no caso bidimensional) capazes de separar as instâncias de acordo com suas respectivas classes. Existem infinitos hiperplanos capazes de particionar o conjunto de dados, no entanto, pode-se notar que os hiperplanos h_1 e h_3, apesar de realizarem uma partição bem-sucedida, se fossem submetidos a pequenas mudanças, passariam a classificar algumas instâncias incorretamente. Seria natural então determinar uma quantidade finita que garantisse uma distância mínima de segurança entre o hiperplano escolhido e os dados de cada classe. Definindo margem como sendo a distância entre o hiperplano e a instância mais próxima a ele (Figura 3.9-b)), o SVM corresponde a um problema de otimização cujo objetivo é maximizar a menor distância entre o hiperplano e o conjunto de dados de treinamento, ou seja, encontrar aquele de margem máxima.

Um hiperplano capaz de particionar o conjunto de dados, de modo que não existam classes distintas em um mesmo lado da partição, só existe se os dados em questão forem linearmente separáveis. No entanto, muitos problemas de classificação existentes não possuem tal propriedade e, para contornar esse

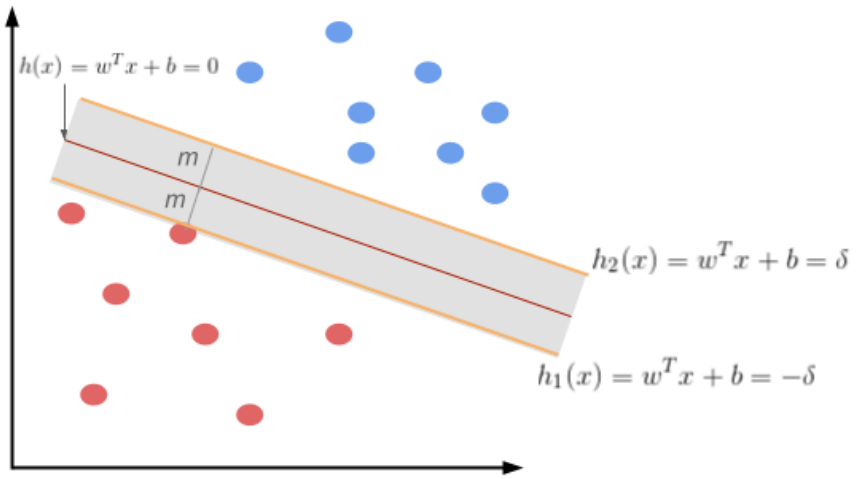

Figura 3.10: Exemplo de hiperplanos paralelos traçados levando em conta a margem m

tipo de situação, os dados de treinamento podem ser mapeados para um espaço de dimensão superior em que eles sejam linearmente separáveis. As seções a seguir são baseadas em [32, 33, 34, 35] e descrevem detalhadamente esses diferentes casos tratados pelo treinamento com SVM, apresentando a formulação de cada problema de otimização a ser resolvido.

3.7.1.1 Problemas linearmente separáveis

Suponha um problema de classificação binária em que se tem um conjunto de dados de treinamento com n instâncias, cada uma definida por um par $(x_i, y_i)_{i=1}^N$, em que $x_i \in \mathbb{R}^p$ é a i-ésima instância de treinamento inserida em um espaço de atributos p-dimensional e $y_i \in \{-1, 1\}$ é a classe associada a ela. Suponha ainda que se trata de um conjunto de dados linearmente separáveis, ou seja, existe um hiperplano definido por $h(x) = w^T x + b = 0$ tal que

$$y_i = \begin{cases} -1, & \text{se } w^T x_i + b < 0 \\ 1, & \text{se } w^T x_i + b > 0 \end{cases} \tag{3-21}$$

De modo equivalente, podemos escrever

$$y_i(w^T x_i + b) > 0 \tag{3-22}$$

Como já foi dito anteriormente, o hiperplano ótimo é aquele que apresenta maior margem em relação aos dados de treinamento. Considere o hiperplano $h(x)$ da Figura 3.10. A partir da margem m definida pelo ponto mais próximo a ele, podemos traçar dois hiperplanos paralelos $h_1(x)$ e $h_2(x)$ com distância m em relação a $h(x)$

$$h_1(x) = w^T x + b = -\delta \tag{3-23}$$

$$h_2(x) = w^T x + b = \delta \tag{3-24}$$

com $\delta > 0$. É fácil perceber que é possível aumentar m se encontrarmos um $h(x)$ que aproxima $h_2(x)$ do ponto mais próximo pertencente à classe azul, de modo que m assuma seu valor máximo e, consequentemente, tanto $h1(x)$ quanto $h_2(x)$ estejam o mais afastados possível de $h(x)$. Nosso objetivo é então resolver um problema de otimização capaz de encontrar os parâmetros w e b do hiperplano que maximiza a distância total $2m$.

Seja a distância d de um ponto x_i do conjunto de dados para o hiperplano $h(x)$ definida por

$$d = \frac{|w^T x_i + b|}{\|w\|} \tag{3-25}$$

Para fins de simplificação e sem perda de generalidade podemos modificar a escala dos vetores w e b multiplicando cada um de seus componentes por $1/\delta$, resultando nas seguintes equações para $h_1(x)$ e $h_2(x)$

$$h_1(x) = w^T x + b = -1 \tag{3-26}$$

$$h_2(x) = w^T x + b = 1 \tag{3-27}$$

Estamos interessados apenas nas distâncias dos chamados vetores de suporte, ou seja, dos pontos mais próximos a $h(x)$ e que estão sobre $h_1(x)$ e $h_2(x)$, representados aqui por $x_i^{(s)}$. Pelas equações 3-25, 3-26 e 3-27 temos que a distância m de cada um desses pontos ao hiperplano $h(x)$ é dada por

$$m = \frac{|w^T x_i^{(s)} + b|}{\|w\|} = \frac{1}{\|w\|} \tag{3-28}$$

Assim, maximizar $2m$ implica em minimizar a Equação 3-28, o que corresponde a resolver o seguinte problema de otimização convexa

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}\|w\|^2 \\ \text{s.a.} \quad & y_i(w^T x_i + b) \geq 1, \ i = 1, \ldots, N. \end{aligned} \tag{3-29}$$

em que a restrição $y_i(w^T x_i + b) \geq 1$ garante que cada lado de $h(x)$ irá conter instâncias pertencentes a uma única classe e impede que haja pontos entre os hiperplanos $h_1(x)$ e $h_2(x)$, o que caracteriza esse problema como um SVM de margem rígida (*hard margin SVM*).

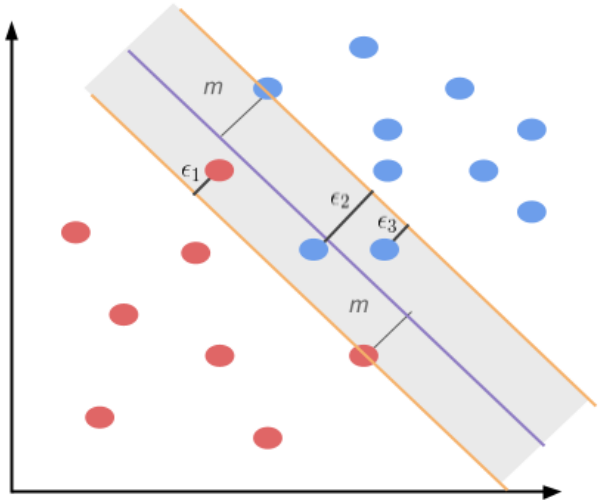

Figura 3.11: Exemplo de classificação utilizando SVM de margem suave.

Em muitas situações, o problema a ser resolvido possui alguns casos de sobreposição de classes, que pode ocorrer quando há a presença de *outliers* nos dados utilizados, ou seja, instâncias que fogem à tendência demonstrada por sua classe. Uma das formas de contornar isso é utilizar uma variável de folga ϵ para permitir que alguns pontos permaneçam a uma distância de $h(x)$ menor que a margem, ou até mesmo que permaneçam do lado de $h(x)$ que não corresponde à sua classe (Figura 3.11). Neste caso, utilizamos o um SVM de margem suave (*soft margin SVM*) e o problema de otimização a ser resolvido passa a ser

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\epsilon_i \\ \text{s.a.} \quad & y_i(w^T x_i + b) \geq 1 - \epsilon_i,\ i = 1, \ldots, N. \\ & \epsilon_i \geq 0 \end{aligned} \tag{3-30}$$

em que C é um parâmetro de "custo"que controla o quanto se deseja maximizar a margem em detrimento de um menor erro de treinamento.

3.7.1.2 Problemas não linearmente separáveis

Suponha agora que o conjunto de dados $\{(x_i, y_i) \mid x_i \in \mathbb{R}^p,\ y_i \in \{-1, 1\}\}$ não seja linearmente separável e mesmo que se utilize o SVM de margem suave, não há uma fronteira linear que irá classificá-los de forma eficiente. Em situações como essa, podemos tentar mapear os dados x_i para um espaço de dimensão superior em que seja possível encontrar uma solução linear para o problema de classificação.

Considere a função de mapeamento $\phi : x \in \mathbb{R}^p \rightarrow \phi(x) \in \mathbb{R}^P$ que mapeia

os dados x_i para um espaço em que sejam linearmente separáveis de acordo com suas respectivas classes. Queremos então encontrar o hiperplano ótimo $\hat{h}(x) = \hat{w}^T\phi(x) + \hat{b}$ capaz de particioná-los no novo espaço P-dimensional. Um modo extremamente eficiente de realizar tal mapeamento é através do chamado "truque do *kernel*" (*kernel trick*) que, utilizando uma função de *kernel*, exclui a necessidade de se definir a função ϕ explicitamente.

Primeiramente, vamos reescrever o problema de otimização 3-30 em sua forma dual

$$\begin{aligned} \max_{\alpha\in\mathbb{R}^N} \quad & \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j x_i^T x_j \\ \text{s.a.} \quad & 0 \le \alpha_i \le C, \quad i = 1, ..., N \\ & \sum_{i=1}^{N}\alpha_i y_i = 0 \end{aligned} \tag{3-31}$$

Com o mapeamento $\phi(x)$, nossa função objetivo passa a ser

$$\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \phi(x_i)^T\phi(x_j) \tag{3-32}$$

É neste ponto que entra o uso do *kernel*. Em alguns casos, calcular o produto interno $\left\langle \phi(x_i)^T, \phi(x_j) \right\rangle$ explicitamente pode ser muito custoso computacionalmente [31]. Ao invés disso, utilizamos uma função de *kernel* K previamente conhecida,

$$K(x, x') = \left\langle \phi(x_i)^T, \phi(x_j) \right\rangle \tag{3-33}$$

capaz de calcular os produtos internos no espaço de dimensão superior com pouco esforço computacional e sem a necessidade de se especificar a transformação $\phi(x)$.

A tabela 3.2 mostra alguns exemplos de funções de *kernel* comumente utilizadas. Todas as funções apresentadas nesta tabela foram testadas no presente trabalho.

Tabela 3.2: Exemplos de funções de *kernel*

Tipo de kernel	$\boldsymbol{K(x_i, x_j)}$
Linear	$x_i x_j$
Polinomial	$(\gamma(x_i x_j) + c)^d$
RBF	$exp(-\gamma\|x_i - x_j\|^2), \quad \gamma > 0$
Sigmoide	$tanh(\gamma x_i x_j + c)$

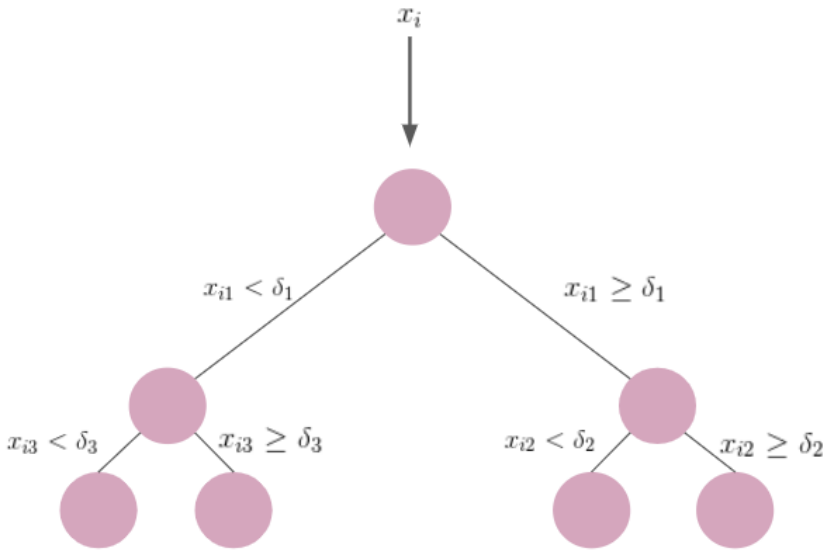

Figura 3.12: Demonstração de como funciona uma árvore de decisão. Os atributos x_{ij} não necessariamente são escolhidos na ordem em que aparecem nos vetores de instâncias x_i, neste caso foram colocados em ordem apenas para simplificar a demonstração.

3.7.2 Random Forest

O algoritmo *random forest* (RF), ou floresta aleatória, consiste na combinação de árvores de decisão para resolver problemas de regressão ou classificação. Em problemas de classificação, uma árvore de decisão $T = \{V, E\}$ é um método de aprendizado que, a partir do conjunto de dados de entrada, realiza a etapa de treinamento através da construção de uma estrutura de regras de decisão que define a forma como novas instâncias serão classificadas [36]. Assim, uma vez concluído o treinamento, quando uma nova instância é apresentada, ela percorre um caminho que vai do nó raiz a algum nó terminal (ou folha), no qual cada nó interno percorrido possui uma regra associada que indica a próxima aresta do caminho a ser percorrida, e cada nó terminal define a classe final prevista.

Em uma árvore binária de decisão, cada nó $v \in V$ representa uma regra associada a um atributo j, com $j \in \{1, \ldots, p\}$, e as duas arestas que saem de v indicam as duas respostas possíveis a tal regra, sendo essa geralmente definida por um valor de referência, de tal modo que os dados de entrada que apresentarem o valor de x_{ij} inferior àquele de referência, seguem pela aresta da esquerda, e os que apresentarem valor superior, vão pela aresta da direita (Figura 3.12).

A cada iteração, analisando o espaço de atributos definido pelos dados de entrada, o objetivo é encontrar o atributo cujo resultado se aproxima de

uma "divisão pura" dos dados, ou seja, que seja capaz de definir uma regra que envia todos os dados pertencentes a uma classe por uma aresta, e aqueles pertencentes à outra classe, pela outra aresta. Um das funções mais utilizadas para medir essa "pureza" associada à divisão é a chamada função Gini, definida por

$$\sum_{k \neq k'} p_{mk} p_{mk'} = \sum_{k=1}^{K} p_{mk}(1 - p_{mk}) \tag{3-34}$$

em que p_{mk} corresponde à proporção de instâncias pertencentes à classe k no nó m analisado e K é a quantidade de classes distintas existentes no problema em questão.

Quando estamos lidando com um problema de classificação binária ($K = 2$), chamando de p a proporção de uma das duas classes existentes, a Equação 3-34 é dada então por $2p(1-p)$. Interpretando $(1-p)$ como sendo a probabilidade da classe com proporção p ser classificada incorretamente, a função Gini pode ser interpretada como a taxa de erro de treinamento decorrente da regra associada ao nó em questão [33]. A escolha do atributo que definirá cada regra de decisão é feita levando em conta o "ganho" resultante da estratégia de divisão, em que ganho pode ser visto simplesmente como um decréscimo no erro [36]. Assim, utilizando a função Gini, é selecionado o atributo que apresentar o menor valor da (Equação 3-34).

O algoritmo RF combina então várias árvores de decisão, fornecendo como conjunto de treinamento para cada árvore um subconjunto dos dados recebidos, amostrado de acordo com alguma função de distribuição previamente definida. Além disso, o RF é caracterizado por uma modificação importante na forma como as árvores de decisão realizam a estratégia de divisão na definição de cada nó. Ao invés de analisar todo o espaço de *features*, cada iteração de uma árvore considera apenas uma amostra do espaço, também de acordo com uma distribuição previamente definida. Essas etapas de subamostragem do *Random Forest* estão entre as principais vantagens oferecidas pelo algoritmo, pois reduzem consideravelmente as chances de *overfitting* sobre os dados de entrada [36].

A construção de árvores de decisão é interrompida quando algum critério de parada é atingido, normalmente associado ao número de árvores existentes e ao ganho de informação associado à construção de uma nova árvore. Quando apresentado a uma nova instância, o modelo de classificação gerado retorna a classe prevista baseando-se em um voto majoritário sobre as previsões das árvores construídas durante a fase de treinamento (Figura 3.13).

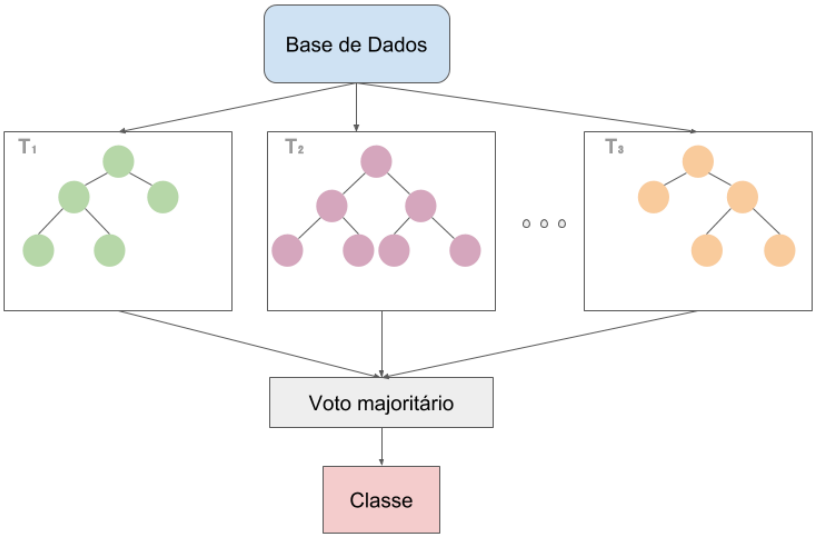

Figura 3.13: Esquema ilustrando a classificação realizada pelo algoritmo *random forest*.

3.7.3 Multilayer Perceptron

Multilayer Perceptron (MLP), ou perceptron de múltiplas camadas, corresponde a uma classe de modelos de redes neurais artificiais, que por sua vez são métodos de aprendizado inspirados pelo funcionamento dos neurônios. Em um problema de classificação, o objetivo de um modelo MLP é encontrar uma função $y = f^*(x)$ capaz de mapear cada dado de entrada x_i para sua respectiva classe y_i. Deste modo, definindo um mapeamento inicial $y = f(x; w)$, o algoritmo tenta aprender o conjunto de parâmetros w que melhor aproxima a função $f^*(x)$. Todas as explicações presentes nesta seção foram baseadas em [31], [37] e [38].

A intuição biológica por trás de um modelo MLP está relacionada à forma como os neurônios se comunicam entre si. Essa comunicação se dá através das chamadas sinapses, conexões entre neurônios que permitem a propagação do impulso nervoso. Cada neurônio possui um canal de entrada chamado dendrito, que recebe os sinais propagados por outros neurônios através da sinapse, e um canal de saída chamado axônio, que de acordo com a intensidade do sinal recebido, envia ou não sinais para serem propagados para outros neurônios (Figura 3.14-(a)).

Analogamente, uma rede neural MLP é composta por camadas de neurônios artificiais, definidos através de modelos computacionais, que se comunicam entre si para aprender uma determinada tarefa. Deste modo, dado o j-ésimo neurônio n_j^l pertencente à camada l, ele recebe uma certa quantidade de sinal a_i^{l-1} vindo de cada neurônio n_i^{l-1} da camada $(l-1)$ anterior, e pondera

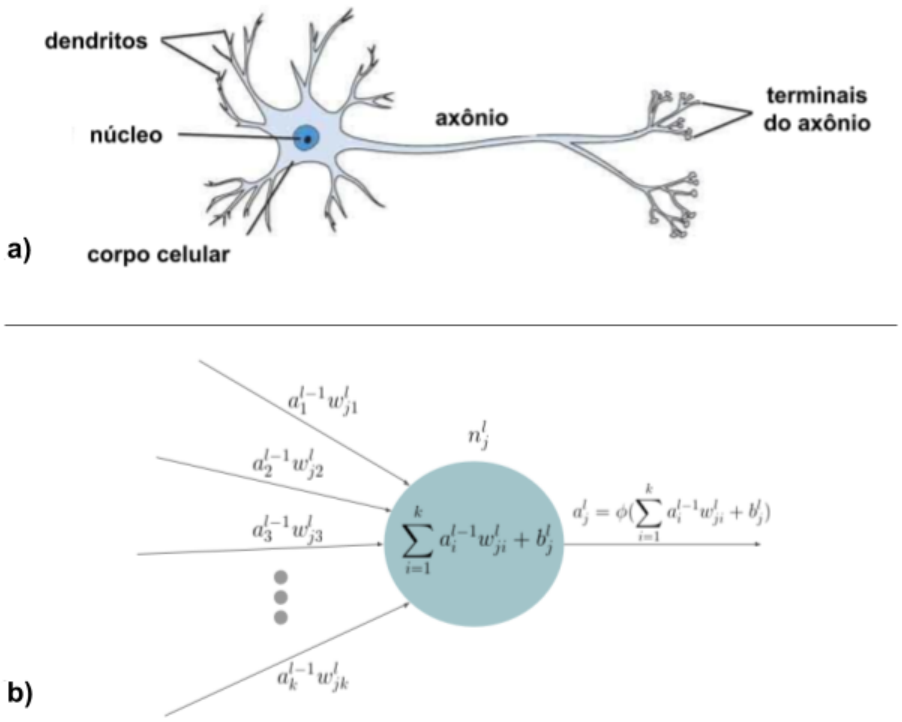

Figura 3.14: a) Representação simplificada de um neurônio biológico (extraída e adaptada de [37]). As sinapses são as conexões entre os terminais do axônio de um neurônio com os dendritos de outros neurônios; b) Representação de um neurônio artificial em uma camada l recebendo sinais de k neurônios pertencentes à camada $(l-1)$ anterior e enviando um sinal de valor a_j^l para os neurônios da camada seguinte.

este sinal de acordo com um peso w_{ji}^l associado à sinapse com n_i^{l-1}. Todos os sinais recebidos por n_j^l são então somados e o valor $\sum_i w_{ji}^l a_i^{l-1}$ resultante é somado a um termo de viés b_j^l e em seguida submetido a uma função de ativação $\phi(a; w, b)$, que definirá a intensidade do sinal de saída que será enviado de n_j^l para os neurônios da camada seguinte (Figura 3.14). A ideia por trás deste procedimento é que o peso w_{ji}^l associado à sinapse entre cada par de neurônios n_j^l e n_i^{l-1} controla a influência que um exerce sobre o outro e, portanto, tais pesos podem ser manipulados (ou aprendidos) para que a rede resultante seja capaz de retornar a resposta desejada.

O uso de uma função de ativação para definir o sinal de saída de cada neurônio é o que confere às redes neurais a capacidade de lidar bem com problemas não lineares. Um exemplo muito comum de função de ativação é a chamada função ReLU (*rectified linear unit*), mesma função de ativação utilizada em nosso trabalho e definida por

$$\phi(x) = max(0, x) \tag{3-35}$$

Uma rede MLP é usualmente representada por um grafo direcionado

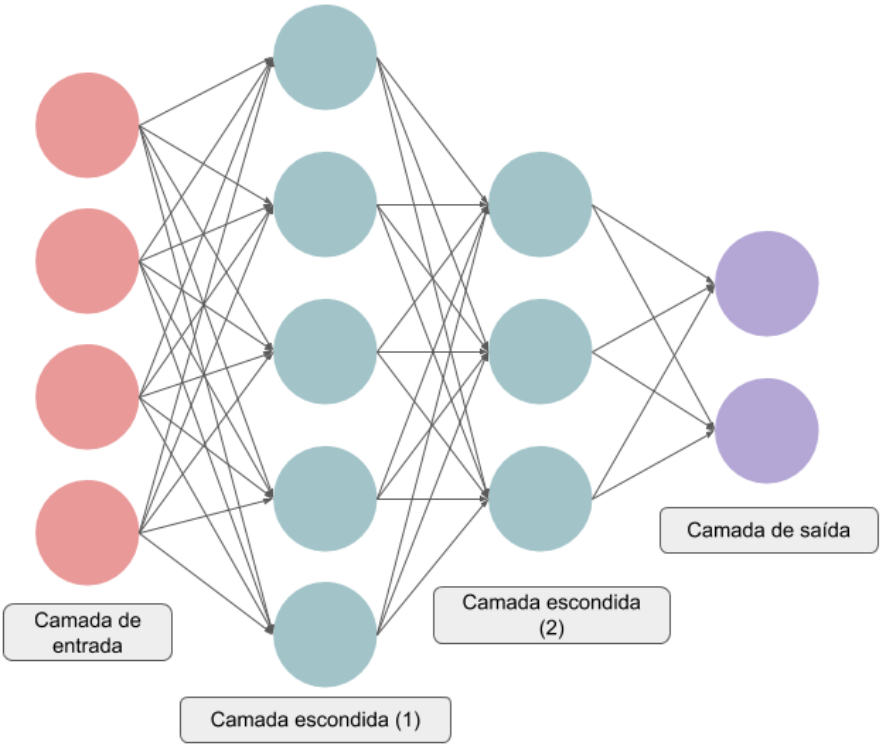

Figura 3.15: Representação de uma rede neural MLP através de um grafo direcionado.

acíclico, como o mostrado na Figura 3.15, no qual a informação é passada em um único sentido através das arestas que ligam cada par de nós. Vale ressaltar que neste tipo de rede neural, apesar dos nós de uma camada estarem ligados a todos os nós da camada seguinte (*fully connected layers*), não existe ligação entre nós de uma mesma camada. A primeira camada é chamada de "camada de entrada" (*input layer*) e é utilizada para representar os atributos de entrada inseridos na rede, de modo que cada neurônio representa um atributo diferente. Logo, se cada vetor de instância de treinamento x_i for composto por p atributos, a camada de entrada da rede irá apresentar p neurônios.

A última camada é chamada de "camada de saída" (*output layer*) e é responsável por retornar a resposta prevista pela rede. Em problemas de classificação, a quantidade de neurônios presentes na camada de saída geralmente equivale à quantidade de classes existentes na base de dados de treinamento, e o neurônio que retorna o maior valor indica a classe prevista pela rede. As demais camadas, por não apresentarem valores de entrada e saída desejados e previamente definidos, recebem o nome de "camadas escondidas" (*hidden layers*). A presença destas camadas escondidas conferem aos algoritmos de redes neurais um alto nível de complexidade e abstração. O modo como tais camadas devem processar os dados não é diretamente especificado, o que significa que elas são capazes de extrair automaticamente (e

"por conta própria") os atributos de cada dado de entrada para serem utilizados durante a etapa de treinamento.

Para modelar o treinamento de uma rede neural, é conveniente utilizar uma representação matricial. Conforme definido anteriormente, utilizando uma função de ativação ϕ previamente definida, podemos calcular o valor a_j^l que sai do j-ésimo neurônio da l-ésima camada através da seguinte expressão

$$a_j^l = \phi(\sum_i w_{ji}^l a_i^{l-1} + b_j^l) \tag{3-36}$$

Reescrevendo a Equação 3-36 utilizando uma notação matricial, temos

$$a^l = \phi(W^l a^{l-1} + b^l) \tag{3-37}$$

em que a^l e a^{l-1} são vetores cujos elementos correspondem aos valores de saída dos neurônios das camadas l e $l-1$, respectivamente; W^l é uma matriz que contém todos os pesos associados às arestas entre os neurônios das camadas $l-1$ e l; b^l é um vetor contendo os termos de viés de cada neurônio da camada l; e ϕ é a função de ativação aplicada elemento a elemento no vetor resultante da operação $w^l a^{l-1} + b^l$.

Tendo em vista que o objetivo do processo de treinamento é encontrar os parâmetros que resultam na melhor classificação dos dados de entrada, é natural analisarmos o desempenho do algoritmo medindo o quanto a classificação gerada se distancia daquela desejada. Assim, para mensurar o erro associado a cada iteração da etapa de treinamento, faremos uso de uma função de custo $C = \frac{1}{N}\sum_{x_i} C_{x_i}$, calculada através da média das funções de custo C_{x_i} obtidas com cada um dos N dados de entrada x_i. Como exemplo, temos a função de custo quadrática, que é frequentemente utilizada neste contexto e é definida por

$$C = \frac{1}{2N}\sum_{x_i} \|y(x_i) - a^L(x_i)\|^2 \tag{3-38}$$

sendo $y(x_i)$ o vetor contendo a resposta correta para a instância x_i e $a^L(x_i)$ o vetor de saída da rede para a mesma instância. Neste caso, podemos pensar em $y(x_i)$ como um vetor de zeros, apresentando o valor 1 na posição correspondente à classe de x_i, e em $a^L(x_i)$ como um vetor de probabilidades associadas a cada classe existente.

Uma vez definida a função de custo, o passo seguinte é especificar como utilizá-la para realizar a atualização dos parâmetros w e b de forma eficiente. Pela Equação 3-38 é fácil perceber que quanto maior for a diferença entre $a^L(x_i)$ e $y(x_i)$, maior será o valor de C. Logo, queremos encontrar os valores de w e

b que minimizam a função de custo. Para isso, iremos introduzir dois métodos importantes: o gradiente descendente (*gradient descent*) e a retropropagação de erro (*backpropagation*).

O método do gradiente descendente é comumente utilizado para resolver o problema de minimização da função de custo em redes neurais. Sendo o gradiente o vetor que fornece a direção de maior crescimento de uma função, a ideia é utilizar o negativo do gradiente da função de custo para atualizar os parâmetros w e b da forma mais eficiente possível, ou seja, escolhendo a direção que faz com que C decresça mais rapidamente. Assim, sendo C uma função de parâmetros $(w^1_{11}, w^1_{12}, ..., w^L_{j_n i_m}, b^1_1, b^1_2, ..., b^L_{j_n})$, seu gradiente é dado por

$$\nabla C = \{\frac{\partial C}{\partial w^1_{11}}, \frac{\partial C}{\partial w^1_{12}}, ..., \frac{\partial C}{\partial w^L_{j_n i_m}}, \frac{\partial C}{\partial b^1_1}, \frac{\partial C}{\partial b^1_2}, ..., \frac{\partial C}{\partial b^L_{j_n}}\} \tag{3-39}$$

com o uso das notações j_n e i_m para indicar os parâmetros do último neurônio da camada de saída (L).

Considerando a iteração correspondente à instância de treinamento x_i, o método consiste então em atualizar os parâmetros w^l_{ji} e b^l_j da camada l da seguinte forma

$$w^l_{ji} = w^l_{ji} - \alpha \frac{\partial C}{\partial w^l_{ji}} \tag{3-40}$$

$$b^l_j = b^l_j - \alpha \frac{\partial C}{\partial b^l_j} \tag{3-41}$$

em que α é uma taxa de aprendizado que controla o tamanho do "passo" que será dado na direção do gradiente.

Para calcular as derivadas parciais que definem o gradiente utiliza-se o método de retropropagação do erro, capaz de realizar este cálculo de forma extremamente eficiente. Considerando que os valores retornados pela rede são afetados por todos os parâmetros w e b presentes nela através de uma "reação em cadeia", a ideia é retropropagar o erro final associado à camada de saída, de forma que seja possível calcular eficientemente o erro δ associado à cada camada escondida. Assim, o erro δ^l da camada l é obtido utilizando o erro δ^{l+1} calculado na camada $(l+1)$. Os valores δ^l são então relacionados às quantidades de interesse $\frac{\partial C}{\partial w^l_{ji}}$ e $\frac{\partial C}{\partial b^l_j}$.

Diante do exposto, seja o erro δ^l_j do j-ésimo neurônio da camada l definido por

$$\delta^l_j = \frac{\partial C}{\partial z^l_j} \tag{3-42}$$

com $z_j^l = \sum_i w_{ji}^l a_i^{l-1} + b_j^l$. Podemos calcular o vetor de erros δ^L associado à camada de saída através da seguinte expressão

$$\delta^L = \nabla_a C \odot \phi'(z^L) \tag{3-43}$$

em que $\nabla_a C$ é o vetor de derivadas parciais $\frac{\partial C}{\partial a_j^L}$, o símbolo $\odot$ representa a multiplicação elemento a elemento (*Hadamard product*) e $\phi'(z^L)$ é a derivada da função de ativação em relação a z^L.

Para as demais camadas, o vetor δ^l pode ser obtido por

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \phi'(z^l) \tag{3-44}$$

Utilizando as Equações 3-43 e 3-44 podemos então obter cada elemento do gradiente

$$\frac{\partial C}{\partial w_{ji}^l} = a_i^{l-1} \delta_j^l \tag{3-45}$$

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l \tag{3-46}$$

As Equações 3-43, 3-44, 3-45 e 3-46 podem ser entendidas com maiores detalhes em [38], tal explicação foge do escopo do presente trabalho. Substituindo os valores 3-45 e 3-46 nas Equações 3-40 e 3-41, respectivamente, podemos então realizar a atualização dos parâmetros. Geralmente, esta atualização só é feita após a rede ter processado um determinado grupo (*batch*) de instâncias previamente definido. Quando todas os grupos terminam de ser processados, dizemos que o treinamento completou uma época (*epoch*) e as instâncias são novamente processadas para dar início a uma nova etapa de treinamento. Uma das condições de parada do aprendizado é limitar o número de épocas concluídas.

3.7.4 Validação Cruzada

A validação cruzada (*cross validation*) é um procedimento muito utilizado para avaliar o desempenho de algoritmos de aprendizado quando se tem uma base de dados de tamanho limitado. Supondo uma base de tamanho N, o algoritmo divide os dados em k grupos, ou *folds*, de tamanho N/k e realiza k iterações independentes de treinamento, de modo que a cada iteração um dos k grupos seja utilizado como conjunto de teste e os outros $k-1$ grupos restantes sirvam como conjunto de treinamento. Uma vez obtidas as métricas

Figura 3.16: Ilustração demonstrando o funcionamento de uma validação cruzada 5-*fold*. M é o valor resultante da métrica utilizada para medir o desempenho do modelo.

de desempenho dos k modelos gerados iterando sobre os k *folds*, utiliza-se a média dessas métricas como uma estimativa para a avaliação do algoritmo de aprendizado utilizado [36] (Figura 3.16).

No presente trabalho, foi utilizada uma base com poucos dados e com uma estrutura desbalanceada, questão que será abordada com maiores detalhes na Seção 4.4. Assim, o desempenho de cada algoritmo de classificação foi avaliado através de uma validação cruzada 5-*fold*.

3.7.5 Seleção de atributos

Suponha um conjunto de instâncias representadas por p atributos. A seleção de atributos consiste em escolher uma quantidade k de atributos para representar cada instância, tal que $k << p$. Existem diversas vantagens em se representar um conjunto de dados com um número reduzido de atributos. Entre elas, podemos citar uma diminuição de custo computacional, uma vez que uma quantidade menor de atributos requer um menor consumo de memória, além de tornar o processo de aprendizado mais rápido. Outra vantagem, e possivelmente a mais relevante, é a diminuição das chances de ocorrer um *overfitting* durante o treinamento, podendo melhorar de forma considerável a capacidade de generalização do modelo final [36].

Outro ponto relevante a ser destacado nesse contexto é que, em muitos conjuntos de dados, é comum a presença de atributos redundantes e pouco explicativos. As técnicas de seleção servem também para eliminar tais atributos

e tentar melhorar a eficácia do modelo. Vale ressaltar que a etapa de seleção é realizada antes de iniciar o processo de treinamento. No presente trabalho, foram estudadas três abordagens distintas para realizar a seleção de atributos. A escolha de tais abordagens foi feita levando em conta as três categorias de métodos de seleção de atributos existentes: *filter*, *embedded* e *wrapper*. Assim, com o objetivo de testar o desempenho de cada categoria em nosso estudo e baseando-se nos métodos de seleção utilizados por [13, 14], foram escolhidos os métodos ANOVA, *extremely randomized trees* e *recursive feature elimination*.

As seções a seguir apresentam as abordagens utilizadas, discorrendo brevemente sobre as principais características de cada uma.

3.7.5.1 ANOVA

A análise de variância (*analysis of variance*, ANOVA) é muito utilizada como uma medida de qualidade para o método de seleção de atributos por filtragem (*filter method*). Este tipo de método consiste em analisar cada atributo individualmente e independentemente dos demais, atribuindo a ele uma certa medida de qualidade. Escolhe-se então os k atributos que apresentaram as medidas mais elevadas, descartando o restante [36].

Em problemas de classificação binária, a análise de variância utiliza o teste F para avaliar a significância estatística de cada atributo na diferenciação entre as duas classes existentes, através do valor-F calculado para cada um. Uma explicação detalhada sobre como utilizar a ANOVA para a seleção de atributos pode ser encontrada em [39].

3.7.5.2 Extremely Randomized Trees

A seleção de atributos utilizando o algoritmo de *extremely randomized trees* (ERT), ou árvores extremamente aleatórias, faz parte da classe de métodos "embutidos" (*embedded*), métodos que se baseiam em procedimentos de seleção de atributos internos a certos tipos de algoritmos de aprendizado [40]. O ERT é muito semelhante ao algoritmo de *random forest* exposto na seção 3.9.2, se baseando em conjuntos de árvores de decisão para lidar com problemas de regressão ou classificação. A maior diferença entre ambos está no critério de escolha para a partição do conjunto de dados para a criação de um novo nó da árvore: ao invés de tentar encontrar o valor que resulta na partição ótima baseada no conjunto de atributos escolhido para aquele nó, como é feito no *random forest*, o algoritmo escolhe o valor para partição de forma aleatória [41].

Em métodos baseados na construção de árvores de decisão, existe uma noção de importância associada a cada atributo, medida através do nível de impureza resultante da partição associada ao nó que escolheu tal atributo. Deste modo, atributos escolhidos por nós cuja partição dos dados resultou em uma maior queda na impureza possuem uma maior importância. No caso de métodos que constroem diversas árvores de decisão, como é o caso do ERT, a importância final de um atributo pode ser calculada tomando a média das importâncias calculadas em todas as árvores de decisão criadas.

Assim, utilizando os valores de importância associados a cada atributo, o método de seleção de atributos simplesmente escolhe aqueles que apresentaram os valores mais elevados, baseando-se em algum critério de escolha, como uma determinada quantidade k de atributos, ou atributos que apresentaram um valor acima de um limite pré-definido, por exemplo.

3.7.5.3 Recursive Feature Elimination

O processo *recursive feature elimination* (RFE), ou eliminação recursiva de atributos, pertence à classe de métodos *wrapper*, métodos que escolhem um subconjunto de atributos baseando-se em um algoritmo de aprendizado específico [42]. Assim, através de um algoritmo de classificação escolhido, a primeira iteração do RFE consiste em treinar um modelo utilizando todo o conjunto de atributos. Ao final do treinamento, um valor de importância é atribuído para cada atributo e aqueles que apresentarem um valor abaixo de determinado limite são eliminados. Uma nova iteração é iniciada com os atributos que sobraram da iteração anterior, realizando um novo treinamento apenas sobre este novo subconjunto de atributos e, ao final, eliminando aqueles de menor importância. Este procedimento se repete até que o subconjunto resultante apresente o número desejado de atributos [43].

O critério que mede a importância de cada atributo dependerá do algoritmo de aprendizado utilizado. No caso do SVM, por exemplo, esta medida pode ser calculada baseando-se nos pesos obtidos após o treinamento, escolhendo aqueles que apresentarem maior valor de w_i^2 [44].

3.7.6 Métricas de Desempenho

Métricas de desempenho são utilizadas com o objetivo de mensurar a eficácia de um modelo de aprendizado para lidar com dados ainda não vistos. A escolha das métricas varia de acordo com os dados, o tipo de problema e o contexto em que se está trabalhando. Nesta seção são apresentadas as métricas

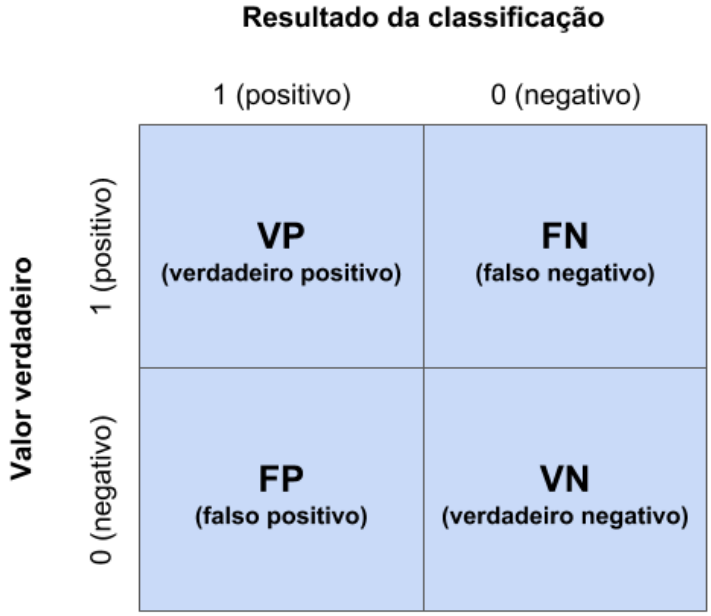

Figura 3.17: Matriz de confusão para um problema de classificação binária.

utilizadas para avaliar o método proposto. Tratando-se de um problema de classificação, todas elas são baseadas na chamada matriz de confusão, cujo objetivo é combinar informações sobre os valores reais de cada classe e aqueles preditos pelo modelo. Dito isto, suponha um problema de classificação binária, em que $y_i \in \{0, 1\}$ é a classe pertencente a cada instância x_i. Iremos chamar a classe 0 de classe negativa e a classe 1 de classe positiva. Considere um modelo M criado para lidar com esse problema. Se aplicarmos M em um conjunto de dados de teste, ou seja, dados ainda não vistos pelo modelo, podemos construir a matriz de confusão ilustrada na Figura 3.17 para analisar a classificação resultante.

A matriz de confusão apresenta quatro situações possíveis decorrentes do processo de classificação: dados da classe 1 classificados corretamente (VP); dados da classe 1 classificados como pertencentes à classe 0 (FN); dados da classe 0 classificados como pertencentes à classe 1 (FP); e dados da classe 0 classificados de forma correta (VN). Através de tais relações, podemos então calcular as seguintes métricas de desempenho:

1. Acurácia: Corresponde à razão entre a quantidade de dados classificados corretamente e a quantidade total de dados, dada por

$$\frac{VP + VN}{VP + FP + FN + VN} \tag{3-47}$$

Em alguns casos, utilizar a acurácia como única medida de desempenho pode resultar em uma avaliação enviesada. Isso porque uma alta acurácia não significa necessariamente que o classificador possui um alto poder

discriminativo. Em bases de dados desbalanceadas, ou seja, quando se tem muito mais dados de uma classe do que de outra, se o classificador for capaz de classificar corretamente todos os dados da classe mais numerosa, mas falhar em classificar os dados pertencentes à outra classe, a acurácia resultante terá um valor alto. No entanto, evidentemente não se trata de um bom classificador.

2. Sensibilidade: Mede a proporção de dados da classe positiva corretamente classificados entre todos os dados pertencentes à classe positiva:

$$\frac{VP}{VP + FN} \tag{3-48}$$

3. Especificidade: Analogamente à sensibilidade, mede a proporção de dados da classe negativa classificados como tal, entre todos os dados pertencentes à classe negativa:

$$\frac{VN}{VN + FP} \tag{3-49}$$

4. Área abaixo da curva ROC (AUC): Muitas vezes é desejável se ter uma probabilidade associada à predição do modelo de classificação utilizado, informando o seu nível de confiança no resultado retornado. Assim, em problemas de classificação binária, podemos estabelecer um limiar de probabilidade acima do qual a classe retornada é a classe positiva e abaixo do qual retornamos a classe negativa. A curva ROC (*Receiver Operating Characteristic*) nos diz a capacidade do modelo de distinguir entre as duas classes, mostrando a sua performance com diversos limiares diferentes de probabilidade [45, 46]. Ela pode ser obtida relacionando a sensibilidade (eixo-y) com o complemento da especificidade (ou seja, 1 - especificidade), também chamado de taxa de falsos positivos (eixo-x) (Figura 3.18). A área abaixo da curva (*area under the curve*, AUC) nos fornece então uma boa medida de performance do modelo final. Quanto mais próxima de 1, melhor é a capacidade preditiva do modelo.

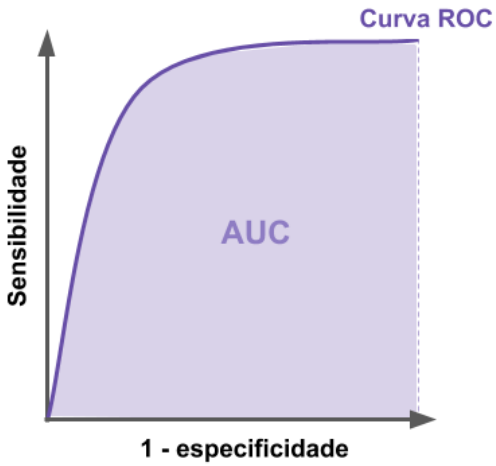

Figura 3.18: Ilustração indicando a curva ROC e a área AUC associada a ela.

4
Metodologia Proposta

Para avaliar o uso dos índices de biodiversidade e diversidade filogenética como atributos no problema de classificação de gliomas, é proposta uma metodologia realizada em cinco etapas principais (Figura 4.1). A primeira etapa consiste na aquisição da base de dados. A partir dos dados obtidos, realiza-se um pré-processamento, seguido pela etapa de extração de atributos, em que são realizadas as adaptações necessárias para o cálculo dos índices propostos. Em seguida, inicia-se a etapa de classificação, em que são testados diferentes algoritmos de aprendizado. Por fim, o resultado de cada método de classificação é avaliado através de diferentes técnicas de desempenho. As seções a seguir explicam cada etapa detalhadamente. Todos os procedimentos computacionais foram implementados na linguagem Python 3.5.5 [62], com o auxílio de bibliotecas já existentes.

4.1
Base de Dados

Foram utilizados os dados de treinamento oferecidos pelo desafio BraTS 2018 [58, 59]. Os dados foram fornecidos por 19 instituições e correspondem a exames de MRI de 285 indivíduos, com presença confirmada de gliomas, sendo 210 indivíduos com glioblastoma (GBM/HGG) e 75 com gliomas de baixo grau (LGG). Cada exame apresenta quatro sequências distintas de ressonância: ponderada em T1, ponderada em T1 pós-contraste (T1ce), ponderada em T2 e inversão-recuperação atenuante de fluido ponderada em T2 (*Fluid Attenuated*

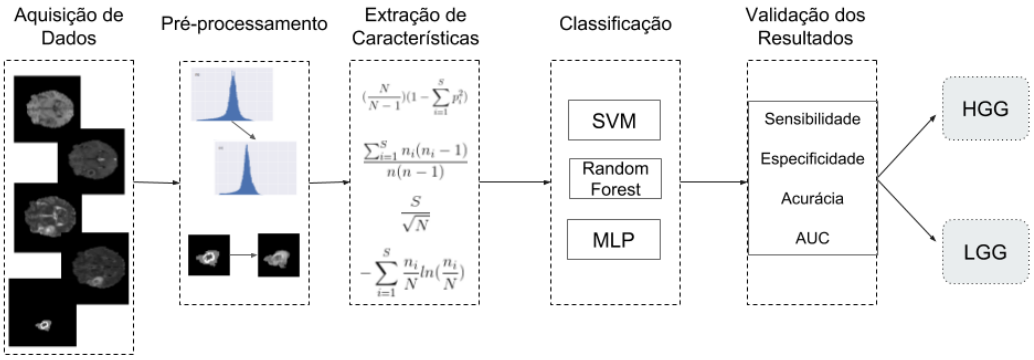

Figura 4.1: Metodologia proposta.

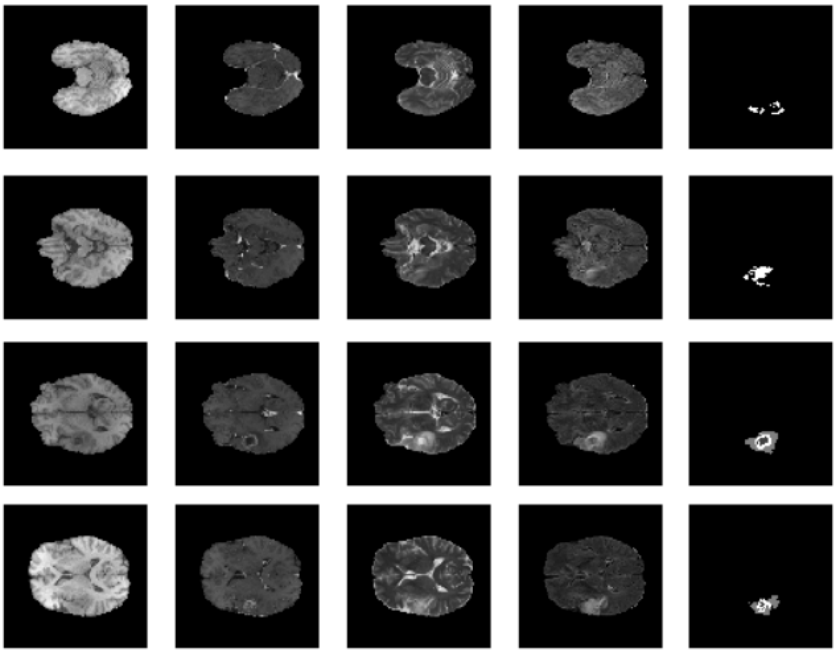

Figura 4.2: Exemplos de fatias de cada modalidade de MRI relativas a um indivíduo presente da base de dados fornecida pelo desafio BraTS 2018. Da esquerda para a direita: modalidades T1, FLAIR, T2, T1ce e anotação das regiões intratumorais. De cima para baixo: fatias 45, 55, 65 e 75, respectivamente.

Inversion Recovery, FLAIR). Cada modalidade de MRI corresponde a uma imagem volumétrica de dimensões 240x240x155 e a Figura 4.2 ilustra alguns exemplos.

As lesões de cada indivíduo foram anotadas manualmente e disponibilizadas em um arquivo à parte, fornecendo as posições dos voxels correspondentes ao volume de interesse a ser extraído de cada exame. As anotações especificam também as seguintes subregiões intratumorais: centro tumoral captante de contraste (*enhancing tumor core*, ET); edema peritumoral (ED); e centro tumoral necrótico e não captante de contraste (*necrotic and non-enhancing tumor core*, NCR/NET) (Figura 4.3). A Figura 4.4 ilustra um exemplo de imagem de anotação das subregiões fornecida pela base de dados. Antes de serem fornecidas para uso, todas as imagens foram submetidas a etapas de pré-processamento que consistiram em um co-registro para um modelo anatômico de referência, em uma interpolação para a mesma resolução (1 mm^3) e na extração da caixa craniana.

A análise de cada subregião intratumoral de forma independente pode fornecer importantes indicadores para a classificação de um glioma. Alguns estudos se propuseram a estudar a relação de regiões necróticas com a taxa de sobrevida de pacientes com glioma [70, 71, 72, 73], concluindo que a sua presença pode indicar um pior prognóstico. Tais regiões são comuns em

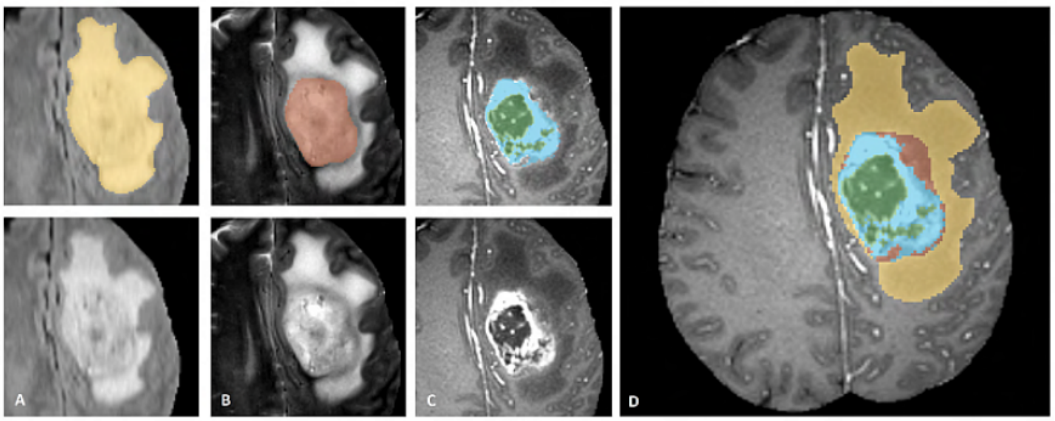

Figura 4.3: Imagem mostrando as diferentes subregiões que podem aparecer em um glioma. (a) A região total do tumor está indicada em amarelo e sua segmentação é feita a partir das modalidades T2 e FLAIR. (b) O centro (núcleo) tumoral está mostrado em vermelho e é visualizado através da modalidade T2. (c) As subregiões ET e NCR internas ao centro tumoral são visíveis em T1ce e estão representadas em azul e verde, respectivamente. (d) Combinação de todas as informações anteriores para formar as subregiões fornecidas pela base de dados: a região total menos o centro e suas estruturas internas formam o edema peritumoral (amarelo); o centro tumoral menos as estruturas NCR e ET, formam a subregião NET (vermelho); as partes em azul indicam apenas a área ET; e em verde, a estrutura NCR. Vale ressaltar que nos dados fornecidos, NCR e NET são representadas como uma classe só. Imagem obtida em [69].

glioblastomas e, portanto, constituem um importante fator para o diagnóstico de gliomas de alto grau. Elas costumam apresentar bordas irregulares e estruturas internas complexas [70], geralmente visíveis em exames de MRI, e a mensuração destas características pode fornecer atributos com alto poder discriminativo para a diferenciação de graus de gliomas.

Além disso, a extensão do edema peritumoral também já foi associada ao prognóstico dos pacientes e diferentes estudos [73, 74, 75] já forneceram evidências de que uma maior extensão está relacionada a um pior prognóstico, sugerindo que características específicas desta região também podem ser relevantes para a classificação de gliomas. Deste modo, o presente trabalho, além de estudar o volume de cada lesão como um todo, utiliza também informações provenientes de cada subregião intratumoral para complementar os atributos utilizados nos modelos de classificação.

4.2 Pré-processamento

A etapa de pré-processamento foi dividida em dois passos. Os dados de MRI utilizados foram obtidos através de instituições distintas e, portanto, derivam de diferentes protocolos clínicos e aparelhos. Assim, o primeiro passo do pré-processamento consistiu na aplicação do algoritmo de especificação de his-

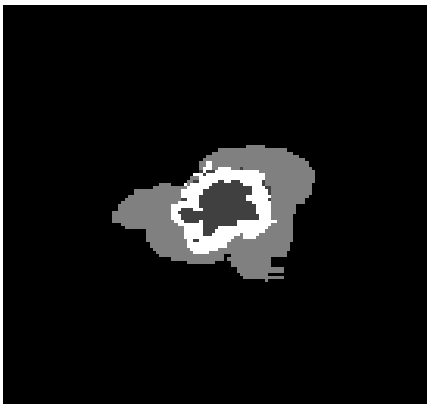

Figura 4.4: Exemplo de anotações para cada subregião intratumoral: NCR/NET (cinza escuro), ED (cinza claro) e ET (branco).

tograma (Seção 3.4.1), com o objetivo de diminuir os efeitos provocados pelos diferentes procedimentos de obtenção dos exames. A imagem de referência foi escolhida utilizando como critério a quantidade de fatias 2D do volume que continham todos os valores de voxel iguais a 0 (ou seja, fatias pretas), escolhendo o indivíduo cujos exames apresentaram a menor quantidade. O método foi aplicado em cada modalidade de MRI separadamente. Os algoritmos de classificação foram testados com e sem a aplicação da especificação de histograma, para analisar se o seu uso traz diferenças significativas para o modelo final. Em seguida, a imagem de anotação (Figura 4.4) de cada indivíduo foi utilizada como máscara para extrair o volume de interesse (Seção 3.4.2) de cada modalidade de MRI, ou seja, o conjunto de voxels correspondente à lesão em cada exame.

4.3 Extração de Atributos

Nesta etapa, foram testadas quatro abordagens distintas. A diferença entre cada uma está nas regiões utilizadas para o cálculo dos atributos, explicadas com maiores detalhes nas seções a seguir. Os atributos são baseados nos 12 índices de biodiversidade e nos 5 de diversidade filogenética apresentados nas seções 3.5 e 3.6, respectivamente, e calculados através das adaptações apresentadas nas mesmas seções.

O objetivo ao utilizar estas diferentes abordagens foi tentar verificar e quantificar diferenças existentes em determinadas regiões intratumorais quando comparadas entre as duas classes de gliomas. Cada abordagem gerou diferentes modelos e os resultados finais são apresentados no Capítulo 5.

4.3.1 Primeira Abordagem

Na primeira abordagem, utilizamos as três subregiões intratumorais fornecidas pela base de dados: NCR/NET, ED e ET. Utilizando a segmentação presente nos arquivos de anotação, extraímos os voxels pertencentes a cada subregião, separando-os em três grupos distintos. Para cada grupo separadamente, calculamos os 17 índices propostos. Este procedimento foi realizado para as 4 modalidades de MRI (T1, T1ce, T2, FLAIR), resultando em um total de 204 atributos para cada indivíduo.

4.3.2 Segunda Abordagem

Nesta abordagem, além dos grupos baseados nas subregiões fornecidas pelos dados, foram criados mais dois grupos concatenando informações dos três já existentes. O primeiro grupo concatenado corresponde ao núcleo do tumor e é formado pela junção das subregiões NCR/NET e ET. Já o segundo grupo corresponde ao tumor como um todo e é formado pela união das três subregiões NCR/NET, ET e ED. Logo, com um total de 5 grupos analisados separadamente, nas 4 modalidades de MRI, esta abordagem resultou no uso de 340 atributos para cada indivíduo.

4.3.3 Terceira Abordagem

Na terceira abordagem, foram definidas novas regiões intratumorais, sem levar em conta as estruturas internas conhecidas da lesão. Foi realizada uma divisão em k camadas, de fora para dentro, de forma que cada camada apresentasse dimensões (altura, largura e profundiade) com $1/k$ dos valores das dimensões do volume original. A Figura 4.5 ilustra exemplos de camadas obtidas após o procedimento de divisão adotado. O objetivo com essa nova abordagem foi verificar se regiões internas da lesão arbitrariamente escolhidas apresentam distinção em relação a regiões mais externas.

Após testes preliminares realizados com diferentes quantidades de camadas, a quantidade $k = 4$ foi a que apresentou os melhores resultados. Portanto, aplicando o procedimento nas 4 modalidades de MRI, esta abordagem resultou na extração de 272 atributos por paciente.

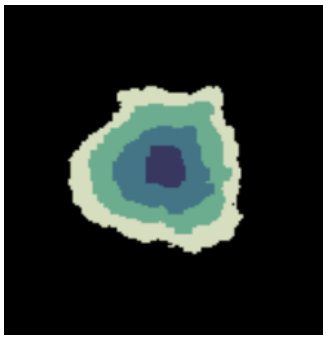

Figura 4.5: Exemplo de fatia 2D da imagem de anotação após a divisão em camadas (terceira abordagem), na qual cada cor representa uma camada diferente.

4.3.4 Quarta Abordagem

A quarta e última abordagem de extração de atributos não levou em conta nenhuma subregião da lesão. Cada índice foi calculado considerando o volume inteiro do tumor como um único grupo. Assim, com um grupo por modalidade de MRI, esta análise deu origem a 68 atributos.

4.4 Classificação

Utilizando uma validação cruzada 5-$fold$ (Seção 3.7.4), foram avaliados três algoritmos de classificação: SVM (Seção 3.7.1), RF (Seção 3.7.2) e MLP (Seção 3.7.3). Para cada um foram testados separadamente os três métodos de seleção de atributos apresentados na Seção 3.7.5, com o intuito de analisar suas diferentes vantagens oferecidas. A seleção foi feita em cada iteração da validação cruzada, de forma independente [60]. Assim, ao final do treinamento de cada modelo, ao analisarmos o seu desempenho, foram reportados apenas o atributos selecionados em 3 ou mais *folds*.

Para encontrar os melhores modelos de classificação produzidos por cada algoritmo, foram testadas diversas combinações de parâmetros. Por exemplo, ao testarmos o SVM, utilizamos quatro tipos de *kernels* diferentes: linear, polinomial, RBF e sigmoide. Com o RF, um dos parâmetros variados foi a profundiade máxima permitida para cada árvore de decisão criada. Já para a MLP, um importante parâmetro foi a arquitetura da rede utilizada, variando a quantidade de camadas escondidas e de neurônios em cada camada.

Para a implementação dos algoritmos de classificação, foram utilizadas as funções presentes na biblioteca Scikit-learn [61]. A Tabela 4.1 mostra os parâmetros variados para cada método de classificação, utilizando as funções disponíveis na biblioteca, e uma breve explicação sobre o seu significado.

Tabela 4.1: Definições dos parâmetros de cada algoritmo variados entre os diferentes treinamentos. O nome de cada parâmetro corresponde aos nomes apresentados nas funções da biblioteca Scikit-learn.

	Parâmetro	**Significado**
SVM	kernel	Função de kernel utilizada.
	C	Parâmetro de penalidade que controla o quanto se deseja priorizar a maximização da margem ao invés de minimizar o erro de treinamento.
	gamma	Coeficiente γ dos kernels RBF, polinomial e sigmoide (Tabela 3.2).
RF	n_estimators	A quantidade de árvores na floresta.
	max_depth	A profundidade máxima de cada árvore.
MLP	hidden_layer_sizes	Quantidade de camadas e de neurônios em cada camada escondida.
	learning_rate_init	A taxa de aprendizado utilizada para controlar o passo de atualização dos pesos.
	momentum	O termo de momento. É utilizado para ajudar a acelerar o processo de aprendizado dos pesos no gradiente descendente.

Vale ressaltar que a escolha da validação cruzada para a avaliação dos modelos foi influenciada pelo tamanho da base de dados e a quantidade de instâncias em cada classe. Conforme já foi dito anteriormente, a classe LGG contém 75 instâncias e a classe HGG 210. Tais valores demonstram uma base extremamente desbalanceada e dificultam abordagens de treinamento que utilizam apenas um conjunto de treino, validação e teste, já que a divisão pode acabar resultando em grupos mal distribuídos e o modelo final acaba se tornando enviesado. O uso da validação cruzada contorna tal desvantagem realizando várias divisões distintas e garantindo que todos os dados passem pelo conjunto de teste. Em geral, isso permite que a capacidade do modelo de aprendizado de classificar dados ainda não vistos seja avaliada com melhor precisão.

4.5 Validação dos Resultados

Para a validação dos resultados, utilizamos as métricas apresentadas na Seção 3.7.6: acurácia, sensibilidade, especificidade e AUC. HGG foi considerada a classe positiva e LGG, a negativa. Assim, instâncias verdadeiro-positivas, por exemplo, indicam indivíduos com gliomas da classe HGG que foram classificados como tal.

Cada métrica de desempenho apresentada nos resultados corresponde à

média e ao desvio-padrão dos valores obtidos para tal métrica nas 5 iterações da validação cruzada 5-*fold* (Seção 3.7.4).

5
Resultados

Neste capítulo, apresentamos os resultados obtidos com a metodologia proposta. Ela está dividida entre as quatro abordagens apresentadas no Capítulo 4. Para cada uma, apresentamos os resultados obtidos e os parâmetros utilizados. Em seguida, o melhor resultado é comparado com aqueles presentes na literatura.

5.1
Resultados da Primeira Abordagem

As Tabelas 5.1, 5.2 e 5.3 mostram os resultados obtidos com a primeira abordagem da metodologia. Para cada etapa modificada de cada algoritmo de classificação, mostramos o melhor resultado. Assim, a primeira linha da Tabela 5.1 indica o melhor resultado alcançado através de treinamentos com o SVM, sem pré-processamento e sem seleção de atributos. A segunda, treinamentos sem pré-processamento, mas utilizando seleção de atributos. E assim por diante. Para as etapas de seleção de atributos, mostramos apenas o melhor resultados entre os três algoritmos testados (Seção 3.7.5). A coluna de parâmetros indica os valores dos parâmetros utilizados em cada um dos resultados, seguindo a mesma ordem apresentada na Tabela 4.1.

Através da Tabela 5.1 pode-se observar que o uso do SVM com o kernel RBF forneceu bons resultados. O uso do algoritmo *histogram matching* não apresentou diferenças significativas neste caso, mas a etapa de seleção de atributos contribuiu para uma leve melhora dos resultados. Observando os resultados obtidos com o RF apresentados na Tabela 5.2, podemos perceber que o desempenho é equivalente ao oferecido pelo SVM, com a principal diferença sendo apenas os algoritmos de seleção de atributos que apresentaram melhores resultados. Em relação aos resultados obtidos pelo algoritmo de classificação MLP mostrados na Tabela 5.3, uma importante observação a ser mencionada é que houve uma melhora considerável se compararmos a métrica de especificidade (taxa de verdadeiros negativos) obtida, o que demonstra que o algoritmo foi capaz de classificar de forma mais eficaz a classe LGG, mesmo essa estando em menor quantidade. O algoritmo de seleção RFE foi o que apresentou melhor desempenho com a utilização de redes neurais para esta

Tabela 5.1: Tabela mostrando os melhores resultados obtidos na primeira abordagem utilizando o SVM. As métricas de desempenho são apresentadas em conjunto com os desvios-padrão correspondentes.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (kernel, C, gamma)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	RBF ; 2^4; 2^{-7}	0,936 ± 0,030	0,902 ± 0,045	0,933 ± 0,077	0,813 ± 0,122
-	Trees	RBF; 2^6; 2^{-6}	0,951 ± 0,040	0,930 ± 0,029	0,971 ± 0,028	0,813 ± 0,107
Sim	-	RBF; 2^5; 2^{-7}	0,961 ± 0,017	0,909 ± 0,013	0,952 ± 0,026	0,787 ± 0,098
Sim	ANOVA	RBF; 2^5; 2^{-7}	0,951 ± 0,035	0,930 ± 0,016	0,967 ± 0,012	0,827 ± 0,068

primeira abordagem.

Tabela 5.2: Tabela mostrando os melhores resultados obtidos na primeira abordagem utilizando o RF.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (n_estimators, max_depth)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	10; 16	0,961 ± 0,023	0,923 ± 0,028	0,957 ± 0,031	0,827 ± 0,116
-	RFE	500; None	0,957 ± 0,026	0,919 ± 0,045	0,967 ± 0,032	0,787 ± 0,115
Sim	-	500; None	0,957 ± 0,034	0,916 ± 0,050	0,967 ± 0,032	0,773 ± 0,131
Sim	Trees	20; 16	0,952 ± 0,027	0,916 ± 0,036	0,957 ± 0,020	0,800 ± 0,111

As Figuras 5.1 e 5.2 apresentam gráficos que ilustram a relação dos valores dos atributos *radiomics* extraídos na primeira abordagem com as classes HGG e LGG. É comum utilizar este tipo de gráfico em contextos *radiomics* para ilustrar o poder discriminativo dos atributos utilizados. O eixo x corresponde às instâncias presentes na base de dados, ou seja, às lesões estudadas, cada uma referente a um paciente diferente. O eixo y representa os 204 atributos utilizados na primeira abordagem, de modo que a cor do ponto (x_i, y_i) indica o valor do atributo y_i calculado sobre a lesão x_i.

Tabela 5.3: Tabela mostrando os melhores resultados obtidos na primeira abordagem utilizando o MLP.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (hid_layer_sizes, learn_rate_init, momentum)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	(143,) ; 0,001; 0,4	0,942 ± 0,023	0,900 ± 0,030	0,938 ± 0,024	0,787 ± 0,050
-	RFE	(285,); 0,1; 0,2	0,941 ± 0,013	0,902 ± 0,032	0,924 ± 0,038	0,840 ± 0,080
Sim	-	(285,); 0,001; 0,9	0,937 ± 0,026	0,920 ± 0,049	0,957 ± 0,044	0,813 ± 0,115
Sim	RFE	(2,); 0,06; 0,9	0,946 ± 0,018	0,902 ± 0,018	0,914 ± 0,012	0,867 ± 0,073

Para fins de comparação, os valores dos atributos foram normalizados (*z-scores*). A linha vertical em vermelho separa as instâncias de cada classe, de modo que as instâncias à esquerda dela pertencem à classe HGG e as instâncias à direita pertencem à classe LGG. Observando os gráficos, podemos constatar que há uma distinção perceptível entre os valores apresentados pelos atributos de instâncias pertencentes a classes distintas. Tal distinção mostra que os atributos escolhidos possuem um bom poder discriminativo no problema de classificação de gliomas.

5.2 Resultados da Segunda Abordagem

Seguindo o mesmo raciocínio da seção anterior, a seguir são apresentados os resultados obtidos com a segunda abordagem. Através das Tabelas 5.4, 5.5 e 5.6 podemos perceber que não houve diferença significativa em relação aos resultados fornecidos pela primeira abordagem. Novamente, o algoritmo MLP foi capaz de classificar melhor os elementos da classe LGG. Além disso, é interessante notar que os três algoritmos de classificação testados na segunda abordagem apresentaram melhores resultados com o uso dos métodos de seleção de atributos RFE e ANOVA, o que possivelmente mostra uma maior capacidade destes algoritmos em escolher os atributos com maior poder discriminativo na abordagem em questão.

Assim como na primeira abordagem, as Figuras 5.3 e 5.4 mostram uma distinção entre os valores dos atributos apresentados pelas instâncias de cada classe. Como era de se esperar, tendo em vista os resultados fornecidos através da classificação, os gráficos de ambas as abordagens apresentam grande

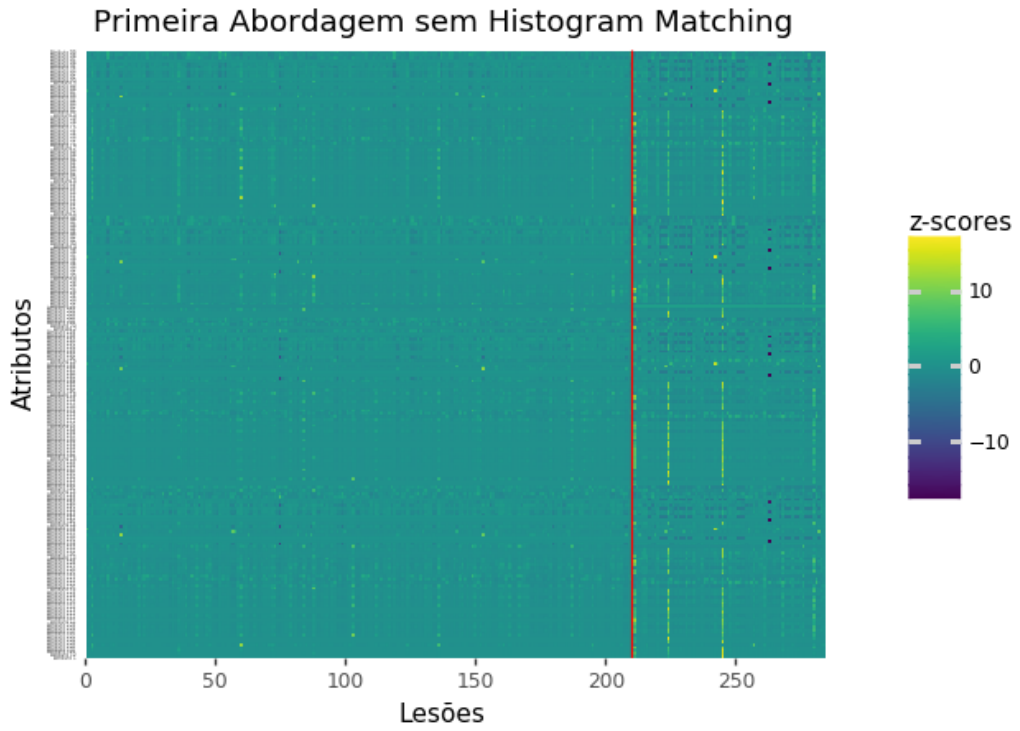

Figura 5.1: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na primeira abordagem, sem o uso do algoritmo de *histogram matching*.

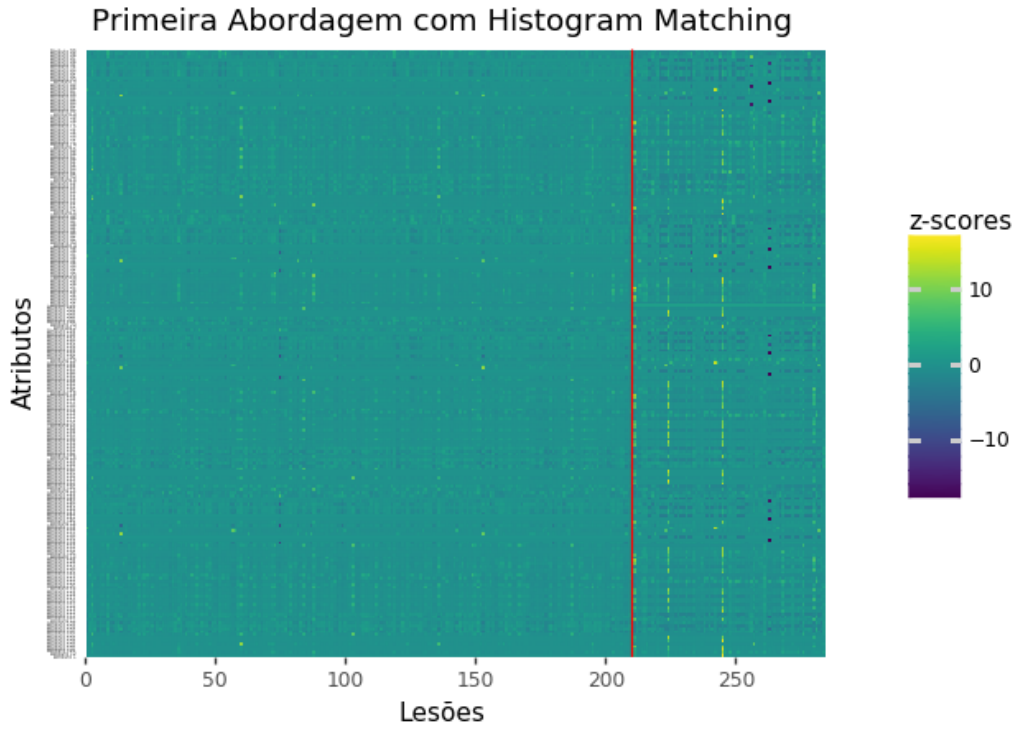

Figura 5.2: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na primeira abordagem, com o uso do algoritmo de *histogram matching*.

Tabela 5.4: Tabela mostrando os melhores resultados obtidos na segunda abordagem utilizando o SVM.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (kernel, C, gamma)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	RBF ; 2^3; 2^{-7}	0,949 ± 0,022	0,916 ± 0,034	0,967 ± 0,028	0,773 ± 0,068
-	RFE	RBF; 2^3; 2^{-7}	0,955 ± 0,021	0,919 ± 0,018	0,962 ± 0,019	0,800 ± 0,084
Sim	-	RBF; 2^3; 2^{-7}	0,958 ± 0,025	0,926 ± 0,047	0,971 ± 0,023	0,800 ± 0,146
Sim	ANOVA	RBF; 2^4; 2^{-7}	0,966 ± 0,009	0,916 ± 0,028	0,957 ± 0,031	0,800 ± 0,073

Tabela 5.5: Tabela mostrando os melhores resultados obtidos na segunda abordagem utilizando o RF.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (n_estimators, max_depth)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	200; 8	0,963 ± 0,013	0,905 ± 0,030	0,957 ± 0,031	0,760 ± 0,068
-	RFE	500; None	0,954 ± 0,025	0,919 ± 0,018	0,967 ± 0,024	0,787 ± 0,027
Sim	-	200; 8	0,952 ± 0,018	0,923 ± 0,028	0,971 ± 0,009	0,787 ± 0,107
Sim	ANOVA	20; 16	0,958 ± 0,023	0,909 ± 0,034	0,952 ± 0,034	0,787 ± 0,13

semelhança ao representar essa distinção entre as duas classes existentes.

5.3 Resultados da Terceira Abordagem

A terceira abordagem apresentou resultados consideravelmente inferiores em relação às duas abordagens anteriores, o que mostra que as regiões escolhidas não conseguem captar as particularidades de cada classe de glioma com a mesma eficácia que as regiões utilizadas nas abordagens anteriores. Isso indica que segmentar o volume do tumor de modo uniforme, sem levar em conta estruturas internas intratumorais específicas, neste caso, não foi suficiente para

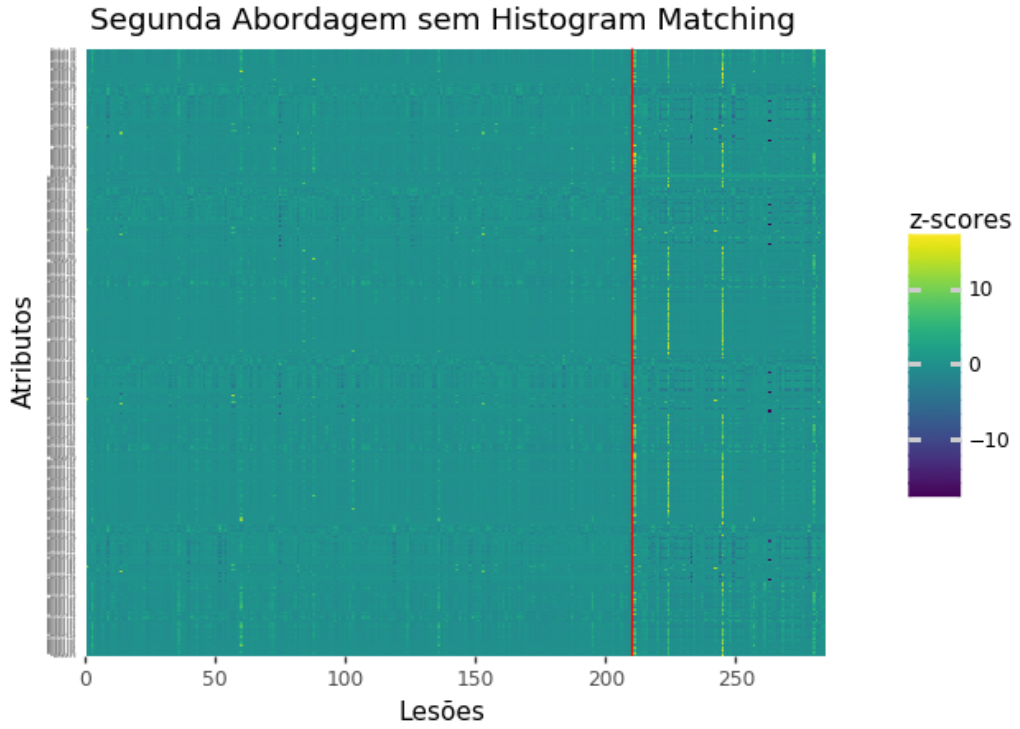

Figura 5.3: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na segunda abordagem, com o uso do algoritmo de *histogram matching*.

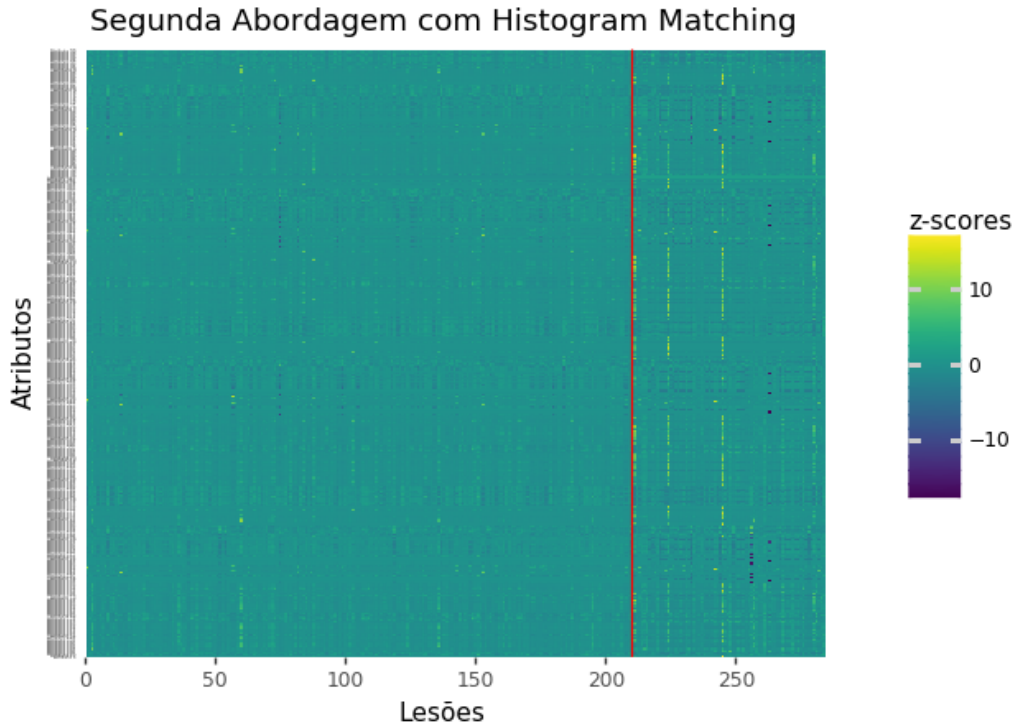

Figura 5.4: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na segunda abordagem, sem o uso do algoritmo de *histogram matching*.

Tabela 5.6: Tabela mostrando os melhores resultados obtidos na segunda abordagem utilizando o MLP.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (hid_layer_sizes, learn_rate_init, momentum)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	(285,) ; 0,02; 0,3	0,935 ± 0,057	0,912 ± 0,035	0,943 ± 0,039	0,827 ± 0,053
-	ANOVA	(285,) ; 0,06; 0,3	0,951 ± 0,023	0,926 ± 0,026	0,948 ± 0,035	0,867 ± 0,042
Sim	-	(100, 50); 0,001; 0,2	0,951 ± 0,041	0,923 ± 0,028	0,967 ± 0,024	0,800 ± 0,042
Sim	RFE	(285,); 0,001; 0,2	0,956 ± 0,017	0,923 ± 0,018	0,962 ± 0,019	0,813 ± 0,027

captar características essenciais que distinguem ambas as classes.

Outra distinção marcante em relação às abordagens anteriores é que o uso do *histogram matching* nesta abordagem mostrou diferenças mais significativas nos resultados, tornando-os melhores. O algoritmo MLP apresentou melhores valores de especificidade do que o SVM e o RF.

Tabela 5.7: Tabela mostrando os melhores resultados obtidos na terceira abordagem utilizando o SVM.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (kernel, C, gamma)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	RBF ; 2; 2^{-7}	0,888 ± 0,043	0,842 ± 0,046	0,933 ± 0,041	0,587 ± 0,088
-	Trees	RBF; 2^3; 2^{-7}	0,905 ± 0,043	0,877 ± 0,046	0,943 ± 0,024	0,693 ± 0,155
Sim	-	RBF; 2; 2^{-7}	0,916 ± 0,047	0,870 ± 0,038	0,962 ± 0,032	0,613 ± 0,122
Sim	RFE	RBF; 2^3; 2^{-6}	0,926 ± 0,025	0,902 ± 0,041	0,962 ± 0,028	0,733 ± 0,094

Através dos gráficos mostrados nas Figuras 5.5 e 5.6, podemos ver claramente que não existe uma distinção tão perceptível entre as instâncias das duas classes, como há nos gráficos mostrados anteriormente, o que ilustra o menor poder discriminativo dos atributos utilizados nesta abordagem.

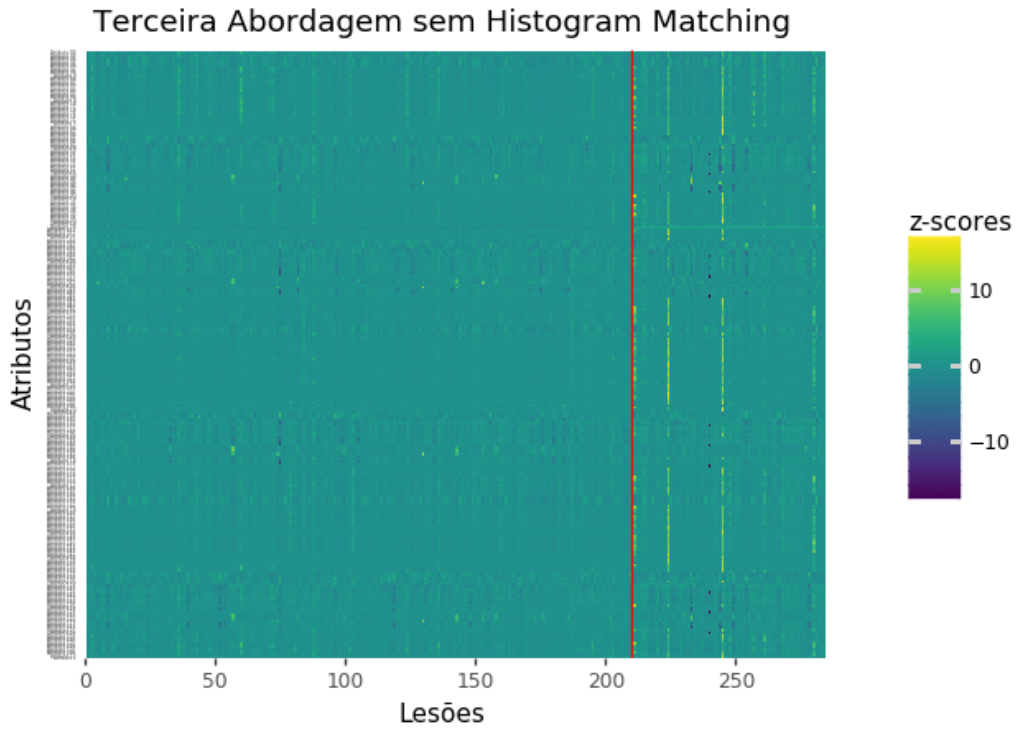

Figura 5.5: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na terceira abordagem, sem o uso do algoritmo de *histogram matching*.

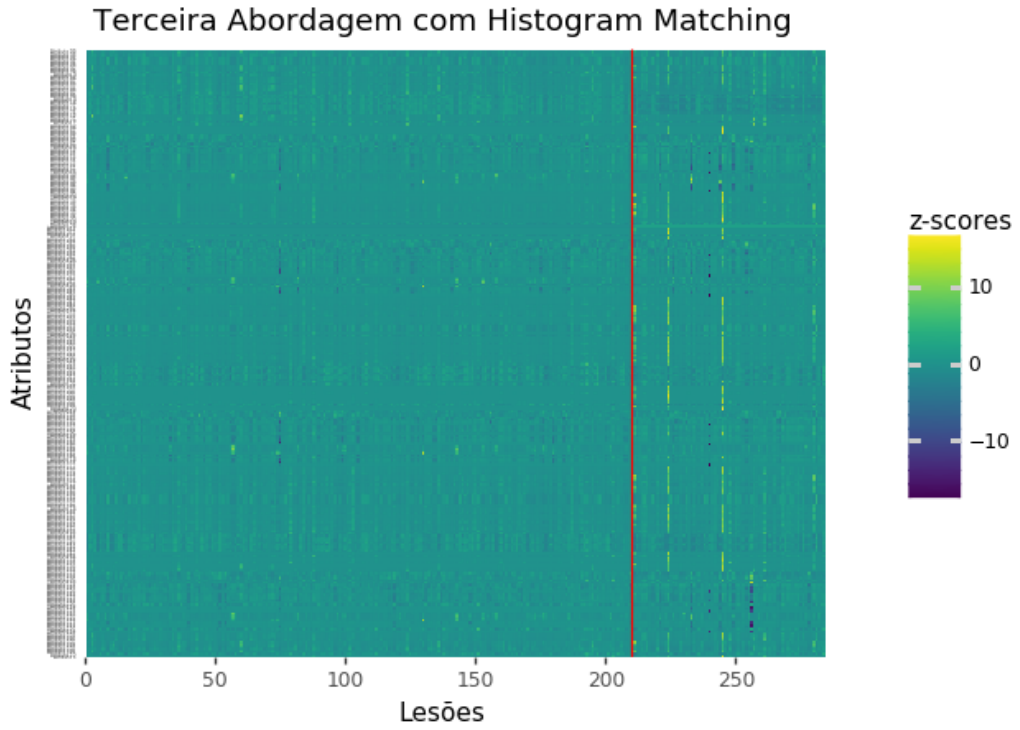

Figura 5.6: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na terceira abordagem, com o uso do algoritmo de *histogram matching*.

Tabela 5.8: Tabela mostrando os melhores resultados obtidos na terceira abordagem utilizando o RF.

Histogram Matching	**Seleção de Atributos**	**Parâmetros (n_estimators, max_depth)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	5; None	0,832 ± 0,032	0,824 ± 0,029	0,881 ± 0,058	0,667 ± 0,060
-	Trees	200; 8	0,898 ± 0,037	0,831 ± 0,042	0,909 ± 0,066	0,613 ± 0,129
Sim	-	100; 8	0,899 ± 0,059	0,884 ± 0,032	0,957 ± 0,035	0,680 ± 0,142
Sim	RFE	500; 8	0,895 ± 0,028	0,881 ± 0,020	0,948 ± 0,018	0,693 ± 0,053

Tabela 5.9: Tabela mostrando os melhores resultados obtidos na terceira abordagem utilizando o MLP

Histogram Matching	**Seleção de Atributos**	**Parâmetros (hid_layer_sizes, learn_rate_init, momentum)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	(143,) ; 0,001; 0,5	0,888 ± 0,052	0,867 ± 0,038	0,924 ± 0,023	0,707 ± 0,196
-	RFE	(287,) ; 0,001; 0,9	0,872 ± 0,054	0,867 ± 0,026	0,919 ± 0,044	0,720 ± 0,050
Sim	-	(287,); 0,001; 0,5	0,904 ± 0,054	0,895 ± 0,025	0,948 ± 0,009	0,747 ± 0,115
Sim	RFE	(285,); 0,001; 0,2	0,885 ± 0,062	0,881 ± 0,056	0,924 ± 0,059	0,760 ± 0,130

5.4 Resultados da Quarta Abordagem

A quarta e última abordagem foi a que apresentou os piores resultados. Os algoritmos SVM e RF apresentaram valores de especificidade inferiores a 60% em praticamente todas as etapas, sendo a única exceção a etapa do RF com o uso do *histogram matching*, o que mostra uma extrema ineficiência desses algoritmos ao utilizarem os atributos extraídos para distinguir a classe LGG.

Os resultados desta abordagem confirmam que as estruturas intratumorais, quando analisadas separadamente, possuem características distintas marcantes entre as duas classes de gliomas e análises que não levam em conta essas

regiões específicas acabam perdendo informações cruciais para a realização de uma classificação mais eficaz.

Tabela 5.10: Tabela mostrando os melhores resultados obtidos na quarta abordagem utilizando o SVM

Histogram Matching	**Seleção de Atributos**	**Parâmetros (kernel, C, gamma)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	RBF ; 2^{5}; 2^{-7}	0,836 ± 0,040	0,807 ± 0,024	0,909 ± 0,031	0,520 ± 0,050
-	Trees	RBF; 2^{3}; 2^{-4}	0,853 ± 0,053	0,803 ± 0,068	0,905 ± 0,095	0,520 ± 0,136
Sim	-	RBF; 2^{2}; 2^{-5}	0,871 ± 0,031	0,860 ± 0,022	0,962 ± 0,019	0,573 ± 0,053
Sim	Trees	RBF; 2^{3}; 2^{-5}	0,877 ± 0,028	0,863 ± 0,034	0,962 ± 0,036	0,587 ± 0,078

Tabela 5.11: Tabela mostrando os melhores resultados obtidos na quarta abordagem utilizando o RF

Histogram Matching	**Seleção de Atributos**	**Parâmetros (n_estimators, max_depth)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	20; 16	0,836 ± 0,016	0,796 ± 0,038	0,890 ± 0,053	0,533 ± 0,073
-	RFE	100; 16	0,820 ± 0,031	0,800 ± 0,008	0,895 ± 0,032	0,533 ± 0,073
Sim	-	10; 16	0,848 ± 0,049	0,831 ± 0,048	0,895 ± 0,053	0,653 ± 0,098
Sim	RFE	20; None	0,861 ± 0,047	0,839 ± 0,030	0,938 ± 0,028	0,560 ± 0,144

Os gráficos mostrados nas Figuras 5.7 e 5.8, de forma ainda mais evidente que na terceira abordagem, ilustram a dificuldade em distinguir as duas classes através dos atributos utilizados nesta abordagem.

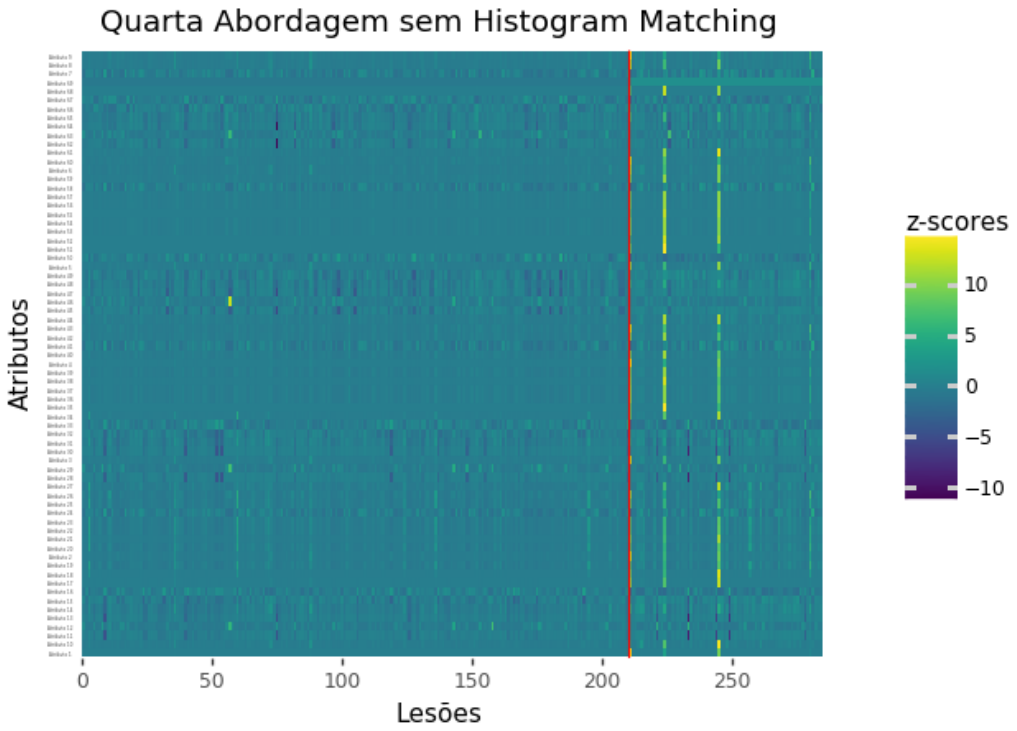

Figura 5.7: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na quarta abordagem, sem o uso do algoritmo de *histogram matching*.

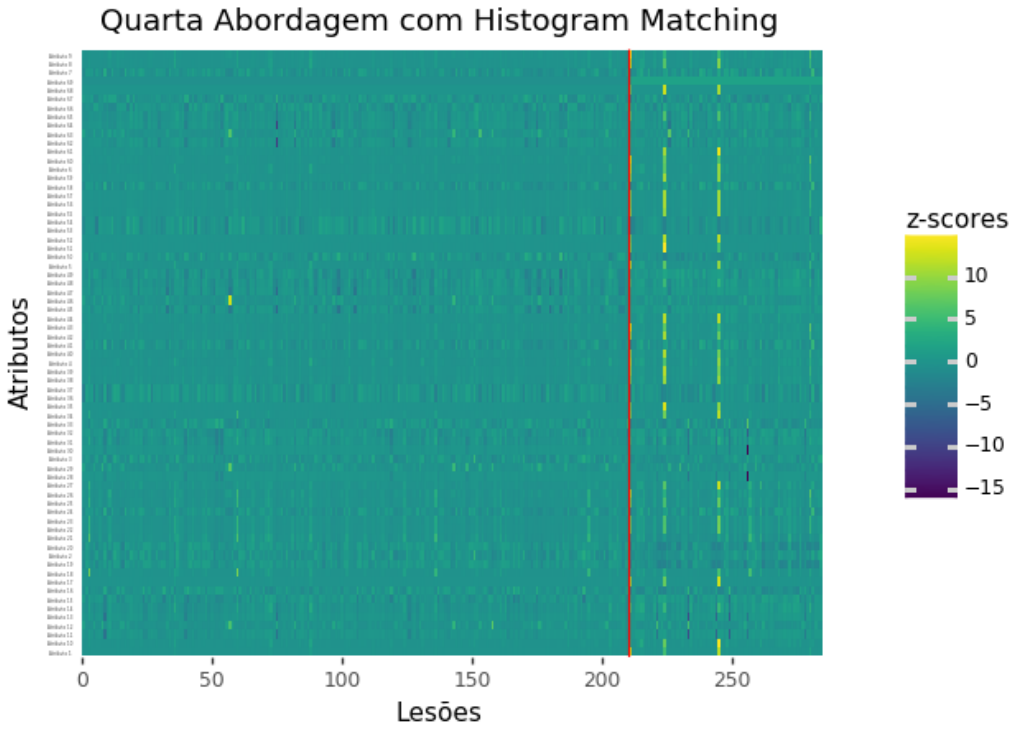

Figura 5.8: Gráfico ilustrando as diferenças de valores entre os atributos de cada instância utilizados na quarta abordagem, com o uso do algoritmo de *histogram matching*.

Tabela 5.12: Tabela mostrando os melhores resultados obtidos na quarta abordagem utilizando o MLP

Histogram Matching	**Seleção de Atributos**	**Parâmetros (hid_layer_sizes, learn_rate_init, momentum)**	**AUC**	**Acurácia**	**Sensibilidade**	**Especificidade**
-	-	(100, 50, 10); 0.02; 0.3	0,820 ± 0,034	0,758 ± 0,042	0,809 ± 0,067	0,613 ± 0,088
-	RFE	(287,) ; 0,001; 0,5	0,841 ± 0,064	0,838 ± 0,042	0,914 ± 0,044	0,627 ± 0,150
Sim	-	(143,); 0,001; 0,2	0,835 ± 0,051	0,860 ± 0,027	0,928 ± 0,030	0,667 ± 0,042
Sim	Trees	(287,); 0,001; 0,9	0,837 ± 0,026	0,831 ± 0,018	0,895 ± 0,032	0,653 ± 0,065

5.5 Discussão

Para realizar uma análise mais minuciosa dos resultados, os conjuntos de atributos selecionados em cada abordagem foram estudados. Nas duas primeiras abordagens, observou-se uma predominância de atributos derivados da modalidade T2 de MRI. Já na terceira e quarta abordagens, a modalidade que predominou foi a T1ce. Isso nos mostra que imagens obtidas com a modalidade T2 talvez tenham uma maior capacidade de destacar características importantes das estruturas intratumorais ET, NCR+NET e ED.

Outra importante análise a ser destacada está relacionada à predominância de atributos extraídos de regiões mais internas do volume tumoral. As duas primeiras abordagens obtiveram uma maior seleção de atributos provenientes da região ET, correspondente ao núcleo da lesão, com a primeira tendo aproximadamente 50% dos atributos selecionados associados a tal região. No caso da terceira abordagem, a camada mais interna foi a que predominou entre os atributos escolhidos. Estes resultados indicam que as características que mais influenciam na distinção entre as classes HGG e LGG estão concentradas nas regiões da lesão mais próximas ao núcleo. Mais ainda, ajudam a explicar os resultados inferiores obtidos pela quarta abordagem, tendo em vista que ela fornece atributos que combinam informações de todo o volume tumoral, fazendo com que informações de pouca relevância se combinem com informações de maior importância em um único atributo, o que dificulta o bom desempenho dos algoritmos de classificação.

Ainda sobre a etapa de seleção de atributos, analisamos também os índices mais escolhidos para determinar aqueles com maior poder discriminativo

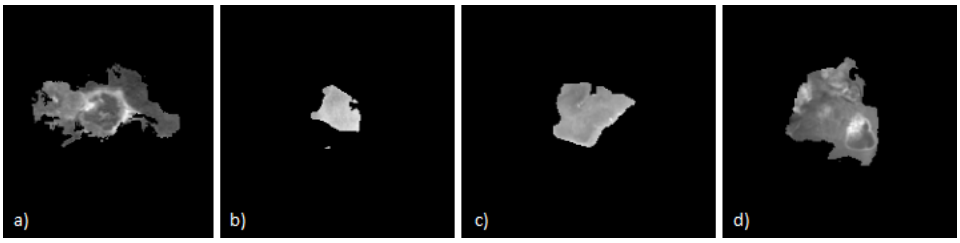

Figura 5.9: a) Tumor HGG classificado corretamente; b) Tumor HGG classificado como LGG; c) Tumor LGG classificado corretamente; d) Tumor LGG classificado como HGG.

no problema de classificação de gliomas. Destacaram-se os de biodiversidade Odum(O), Hulbert (PIE), McNaughton (I), Simpson (D_S) e Shannon (H'), e os de diversidade filogenética Entropia Quadrática Média (J), Entropia Quadrática Extensiva (F), Distinção Taxonômica Média ($AvTD$) e a Medida de Pura Diversidade (D_d).

A Figura 5.9 exibe alguns exemplos de classificação produzidos por um modelo de classificação utilizando o SVM, sendo todas as imagens mostradas correspondentes à modalidade T1ce, por possuir diferenças visualmente mais perceptíveis entre as classes. Observando as duas primeiras imagens correspondentes a gliomas de alto grau, pode-se notar que a imagem 5.9-b apresenta uma estrutura bem distinta da 5.9-a, com bordas mais suaves e uma textura mais homogênea, se assemelhando mais à imagem 5.9-c pertencente à classe LGG. Por outro lado, a imagem 5.9-d corresponde a um glioma de baixo grau, mas apresenta uma estrutura heterogênea com grande variação de escalas de cinza, estando mais próxima à estrutura apresentada por 5.9-a.

Tomando o índice de Simpson como exemplo, que mede a probabilidade de dois indivíduos escolhidos ao acaso pertencerem à mesma espécie - assumindo assim valores baixos em comunidades com maior diversidade de espécie (mais heterogêneas) e valores altos no caso contrário -, mesmo tendo se mostrado um índice com alto poder discriminativo, apresentou valores próximos entre 5.9-b e 5.9-c. Isso pode ser facilmente justificado pela presença de tons de cinza com menor variação e espalhados mais uniformemente pela imagem da lesão, originando comunidades com espécies distribuídas de forma mais homogênea.

Analisando todos os resultados expostos na seção anterior, escolhemos como referência aquele obtido com o uso do SVM na primeira abordagem, com atributos extraídos após a aplicação do *histogram matching* e da seleção de atributos com ANOVA. Tal escolha foi influenciada pelas suas altas taxas de verdadeiros positivos e verdadeiros negativos, além do elevado valor de AUC. Selecionando 97 atributos, o modelo atingiu valores de AUC, acurácia,

Tabela 5.13: Comparação de resultados para o problema de classificação de gliomas utilizando atributos *radiomics* em dados de MRI.

Trabalho	AUC	Acurácia	Sensibilidade	Especificidade	N^o de atributos
[8]	0,896	0,878	0,846	0,955	24
[15]	0,960	0,960	-	-	50
[16]	0,887	0,898	0,889	0,907	16-34
[17]	0,960	0,913	0,913	-	25
[18]*	0,921	0,888	0,943	0,733	5
Método proposto	**0,951**	**0,930**	**0,967**	**0,827**	**97**

* Trabalhos com a mesma base de dados aqui utilizada.

sensibilidade e especificidade de 0,951, 0,930, 0,967 e 0,827, respectivamente.

A Tabela 5.13 mostra uma comparação do nosso melhor resultado com outros estudos que também utilizaram uma abordagem *radiomics* para lidar com o problema de classificação de gliomas. Podemos observar que o nosso método, além de ter apresentado desempenho comparável a todos os resultados mostrados, forneceu resultados superiores aos obtidos por Cho et al. (2018) [18], que utilizou a mesma base de dados empregada em nosso estudo.

6
Conclusão e Trabalhos Futuros

Através da combinação de índices de biodiversidade e diversidade filogenética, o presente trabalho apresentou uma eficiente abordagem utilizando atributos *radiomics* para a classificação de gliomas entre HGG e LGG. A nossa metodologia conseguiu obter resultados satisfatórios utilizando a base de dados fornecida pelo desafio BraTS'18, mesmo com a presença de poucos dados e classes desbalanceadas, o que demonstra que os atributos propostos possuem grande poder discriminativo no problema abordado.

O melhor resultado obtido conseguiu alcançar valores de AUC, acurácia, sensibilidade e especificidade de 0,951, 0,930, 0,967 e 0,827, respectivamente, utilizando o algoritmo de classificação SVM. Analisando todas as abordagens testadas, apesar dos resultados promissores, notamos que, de forma geral, a capacidade do modelo em classificar instâncias da classe LGG foi inferior se comparada à classe HGG. Isso se deve principalmente ao número consideravelmente inferior de instâncias pertencentes à classe LGG. O uso da validação cruzada serviu como uma tentativa de mitigar o efeito desse desbalanceamento dos dados sobre o desempenho dos algoritmos testados.

A principal contribuição do nosso trabalho está na demonstração da eficácia no uso de índices do campo da biologia, associados à noção de diversidade de espécies, para a extração de informações relevantes presentes em imagens de gliomas, obtidas através de exames por ressonância magnética. Mais ainda, nosso trabalho corrobora a noção de que a comunicação entre diferentes áreas de conhecimento pode ser extremamente enriquecedora para o avanço da pesquisa científica.

Ao falar de trabalhos futuros, é natural pensar em uma busca por dados mais balanceados. Apesar da crescente disponibilidade de imagens médicas fornecidas para estudos acadêmicos, ainda há uma carência por uma quantidade maior de dados para o estudo de classificação de gliomas que sejam facilmente acessíveis. Além disso, podemos pensar também em uma busca por outros tipos de índices que formem atributos *radiomics* que melhorem ainda mais os modelos de classificação aqui gerados.

O uso do *radiomics* por especialistas pode significar não apenas a otimização do tempo gasto na análise de exames de imagens, mas também o aumento

da precisão com que tais análises são feitas. Deste modo, nosso trabalho reforça a eficácia de abordagens não invasivas utilizando atributos *radiomics* e técnicas de *machine learning* para lidar com a classificação de gliomas através de imagens por ressonância magnética. Mais do que tentar reduzir a necessidade de submeter pacientes a exames invasivos, modelos *radiomics* podem ser utilizados para auxiliar a tomada de decisão por especialistas, fornecendo o melhor prognóstico possível para o paciente.

Referências

[1] **NBTS 2017. Tumor Type: Understanding Brain Tumors**. http://braintumor.org/brain-tumor-information/understanding-brain-tumors/tumor-types/. Acessado em: Jan 2019.

[2] LOUIS, D. N.; OHGAKI, H.; WIESTLER, O. D.; CAVENEE, W. K.; BURGER, P. C.; JOUVET, A.; SCHEITHAUER, B. W. ; KLEIHUES, P.. **The 2007 WHO Classification of Tumours of the Central Nervous System**. Acta Neuropathologica, 114(2):97–109, Aug 2007.

[3] LOUIS, D. N.; PERRY, A.; REIFENBERGER, G.; VON DEIMLING, A.; FIGARELLA-BRANGER, D.; CAVENEE, W. K.; OHGAKI, H.; WIESTLER, O. D.; KLEIHUES, P. ; ELLISON, D. W.. **The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary**. Acta Neuropathologica, 131(6):803–820, Jun 2016.

[4] WHITTLE, I. R.. **The Dilemma of Low Grade Glioma**. Journal of Neurology, Neurosurgery & Psychiatry, 75(suppl 2):ii31–ii36, 2004.

[5] CLAUS, E. B.; WALSH, K. M.; WIENCKE, J. K.; MOLINARO, A. M.; WIEMELS, J. L.; SCHILDKRAUT, J. M.; BONDY, M. L.; BERGER, M.; JENKINS, R. ; WRENSCH, M.. **Survival and Low-Grade Glioma: The Emergence of Genetic Information**. Neurosurgical Focus FOC, 38(1), 2015.

[6] **Glioma Diagnosis and Treatment**. https://www.mayoclinic.org/diseases-conditions/glioma/diagnosis-treatment/drc-20350255. Acessado em: Jan 2019.

[7] GILLIES, R. J.; KINAHAN, P. E. ; HRICAK, H.. **Radiomics: Images Are More Than Pictures, They Are Data**. Radiology, 278(2), 2016.

[8] MCGRANAHAN, N.; SWANTON, C.. **Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution**. Cancer Cell, 28, Issue 1:141, July 2015.

[9] MENZE, B. H. ET AL. . **The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)**. IEEE Transactions on Medical Imaging, 34(10):1993–2024, Oct 2015. 10.1109/TMI.2014.2377694, 0278-0062.

[10] BAKAS, S.; AKBARI, H.; SOTIRAS, A.; BILELLO, M.; ROZYCKI, M.; KIRBY, J. S.; FREYMANN, J. B.; FARAHANI, K. ; DAVATZIKOS, C.. **Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features**. Scientific data, 4(170117), 2017. doi:10.1038/sdata.2017.117.

[11] DE CARVALHO FILHO, A. O.; SILVA, A. C.; CARDOSO DE PAIVA, A.; NUNES, R. A. ; GATTASS, M.. **Computer-Aided Diagnosis of Lung Nodules in Computed Tomography by Using Phylogenetic Diversity, Genetic Algorithm, and SVM**. Journal of Digital Imaging, 30(6):812–822, Dec 2017. ISSN 1618-727X. https://doi.org/10.1007/s10278-017-9973-6.

[12] NETO, A. C. D. S.; DINIZ, P. H. B.; DINIZ, J. O. B.; CAVALCANTE, A. B.; SILVA, A. C.; DE PAIVA, A. C. ; DE ALMEIDA, J. D. S.. **Diagnosis of Non-Small Cell Lung Cancer Using Phylogenetic Diversity in Radiomics Context**. In: Campilho, A.; Karray, F. ; ter Haar Romeny, B., editors, IMAGE ANALYSIS AND RECOGNITION, p. 598–604, Cham, 2018. Springer International Publishing. 978-3-319-93000-8.

[13] ZACHARAKI, E. I.; WANG, S.; CHAWLA, S.; YOO, D. S.; WOLF, R.; MELHEM, E. R. ; DAVATZIKOS, C.. **Classification of Brain Tumor Type and Grade Using MRI**. Magnetic Resonance in Medicine, 2009.

[14] BO QIN, J.; LIU, Z.; ZHANG, H.; SHEN, C.; CHUN WANG, X.; TAN, Y.; WANG, S.; FENG WU, X. ; TIAN, J.. **Grading of Gliomas by Using Radiomic Features on Multiple Magnetic Resonance Imaging (MRI) Sequences**. Med Sci Monit, 2017.

[15] ZHANG, X.; YAN, L.-F.; YU-CHUAN HU, G. L.; YANG, Y.; HAN, Y.; SUN, Y.; LIU, Z.-C.; TIAN, Q.; HAN, Z.-Y.; LIU, L.-D.; HU, B.-Q.; QIU, Z.-Y.; WANG, W. ; CUI, G.-B.. **Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features**. Oncotarget, 2017.

[16] CHO, H.-H.; PARK, H.. **Classification of low-grade and high-grade glioma using multi-modal image radiomics features**. volumen 2017, p. 3081–3084. 39th Annual International Conference of the

IEEE Engineering in Medicine and Biology Society (EMBC), 07 2017. 10.1109/EMBC.2017.8037508.

[17] CHEN, W.; LIU, B.; PENG, S.; SUN, J. ; QIAO, X.. **Computer-Aided Grading of Gliomas Combining Automatic Segmentation and Radiomics**. International Journal of Biomedical Imaging, 2018(2512037):11, 2018.

[18] CHO, H.; LEE, S.; KIM, J. ; PARK, H.. **Classification of the glioma grading using radiomics analysis**. PeerJ 6:e5982, 2018. https://doi.org/10.7717/peerj.5982.

[19] KUO, M. D.; GOLLUB, J.; SIRLIN, C. B.; OOI, C. ; CHEN, X.. **Radiogenomic Analysis to Identify Imaging Phenotypes Associated with Drug Response Gene Expression Programs in Hepatocellular Carcinoma**. Journal of Vascular and Interventional Radiology, 18(7):821 – 830, 2007.

[20] WIBMER, A.; HRICAK, H.; GONDO, T.; MATSUMOTO, K.; VEERARAGHAVAN, H.; FEHR, D.; ZHENG, J.; GOLDMAN, D.; MOSKOWITZ, C.; W FINE, S.; E REUTER, V.; EASTHAM, J.; SALA, E. ; ALBERTO VARGAS, H.. **Haralick Texture Analysis of Prostate MRI: Utility for Differentiating Non-cancerous Prostate From Prostate Cancer and Differentiating Prostate Cancers with Different Gleason Scores**. European radiology, 25, 05 2015.

[21] LI, H.; ZHU, Y.; BURNSIDE, E.; DRUKKER, K.; A HOADLEY, K.; FAN, C.; D CONZEN, S.; J WHITMAN, G.; J SUTTON, E.; M NET, J.; GANOTT, M.; HUANG, E.; MORRIS, E.; M PEROU, C.; JI, Y. ; GIGER, M.. **MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays**. Radiology, 281:152110, 05 2016.

[22] UPADHYAY, N.; WALDMAN, A. D.. **Conventional MRI Evaluation of Gliomas**. The British Journal of Radiology, 84(special_issue_2):S107–S111, 2011. PMID: 22433821.

[23] LAO, J.; CHEN, Y.-S.; LI, Z.; LI, Q.; ZHANG, J.; LIU, J. ; ZHAI, G.. **A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme**. Scientific Reports, 7, 12 2017.

[24] **Imagiologia de Ressonância Magnética**. http://www.radiacao-medica.com.br/tipos-de-imagens-medicas/outros-tipos-de-imagiologia-medica/imagiologia-de-ressonancia-magnetica/. Acessado em: Jan 2019.

[25] **Ponderações em RM**. https://ampoladigital.wordpress.com/2017/01/26/ponderacoes-em-rm/. Acessado em: Jan 2019.

[26] VILLANUEVA-MEYER, J. E.; MABRAY, M. C. ; CHA, S.. **Current Clinical Brain Tumor Imaging**. Neurosurgery, 81(3):397–415, 05 2017.

[27] **Ressonância Nuclear Magnética**. http://www.oncoguia.org.br/conteudo/ressonancia-nuclear-magnetica/6795/842/. Acessado em: Jan 2019.

[28] **3D Magnetic Resonance Imaging (3D MRI)**. https://www.myvmc.com/investigations/3d-magnetic-resonance-imaging-3d-mri/. Acessado em: Jan 2019.

[29] **Glioma in adults**. https://www.cancerresearchuk.org/about-cancer/brain-tumours/types/glioma-adults. Acessado em: Jan 2019.

[30] **Neoplasias malignas**. http://genoma.ib.usp.br/en/node/611. Acessado em: Jan 2019.

[31] GOODFELLOW, I.; BENGIO, Y. ; COURVILLE, A.. **Deep Learning**. MIT Press, 2016. http://www.deeplearningbook.org.

[32] CORTES, C.; VAPNIK, V.. **Support-vector networks**. Machine Learning, 20(3):273–297, Sep 1995.

[33] HASTIE, T.; TIBSHIRANI, R. ; FRIEDMAN, J.. **The Elements of Statistical Learning**. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

[34] JAMES, G.; WITTEN, D.; HASTIE, T. ; TIBSHIRANI, R.. **An Introduction to Statistical Learning: With Applications in R**. Springer Publishing Company, Incorporated, 2014.

[35] POCHET, A. D. J.. **Modeling of Geobodies: AI for seismic fault detection and all-quadrilateral mesh generation**. Pontifical Catholic University of Rio de Janeiro, 2018.

[36] SHALEV-SHWARTZ, S.; BEN-DAVID, S.. **Understanding Machine Learning: From Theory to Algorithms**. Cambridge University Press, New York, NY, USA, 2014.

[37] STANFORD UNIVERSITY. **CS231n: Convolutional Neural Networks for Visual Recognition**. `http://cs231n.github.io/`. Acessado em: Fev 2019.

[38] NIELSEN, M. A.. **Neural Networks and Deep Learning**. Determination Press, 2015. `http://http://neuralnetworksanddeeplearning.com/`.

[39] DING, H.; FENG, P.-M.; CHEN, W. ; LIN, H.. **Identification of bacteriophage virion proteins by the ANOVA feature selection and analysis**. Molecular bioSystems, 10, 06 2014.

[40] AKASH DUBEY. **Feature Selection Using Random forest**. `https://towardsdatascience.com/feature-selection-using-random-forest-26d7b747597f`. Acessado em: Mar 2019.

[41] GEURTS, P.; ERNST, D. ; WEHENKEL, L.. **Extremely randomized trees**. Machine Learning, 63(1):3–42, Apr 2006.

[42] USMAN MALIK. **Applying Wrapper Methods in Python for Feature Selection**. `https://stackabuse.com/applying-wrapper-methods-in-python-for-feature-selection/`. Acessado em: Mar 2019.

[43] GUYON, I.; WESTON, J.; BARNHILL, S. ; VAPNIK, V.. **Gene selection for cancer classification using support vector machines**. Machine Learning, 46(1):389–422, Jan 2002.

[44] SANZ, H.; VALIM, C.; VEGAS, E.; OLLER, J. M. ; REVERTER, F.. **Svm-rfe: selection and visualization of the most relevant features through non-linear kernels**. BMC Bioinformatics, 19(1):432, Nov 2018.

[45] SARANG NARKHEDE. **Understanding AUC - ROC Curve**. `https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5`. Acessado em: Mar 2019.

[46] JOCELYN D'SOUZA. **Let's learn about AUC ROC Curve!** `https://medium.com/greyatom/`

lets-learn-about-auc-roc-curve-4a94b4d88152. Acessado em: Mar 2019.

[47] SILVA, G. L. F. A. D.; CARVALHO FILHO, A. O. D.; SILVA, A. A. C. A.; PAIVA, A. C. D. ; GATTASS, M.. **Taxonomic indexes for differentiating malignancy of lung nodules on CT images**. Research on Biomedical Engineering, 32:263 – 272, 09 2016. ISSN 2446-4740. http://dx.doi.org/10.1590/2446-4740.04615.

[48] KARYDIS, M.; TSIRTSIS, G.. **Ecological indices: A biometric approach for assessing eutrophication levels in the marine environment**. Science of The Total Environment, 186:209–219, 07 1996. 10.1016/0048-9697(96)05114-5.

[49] CAMPOS, D.; ISAZA, J. F.. **A geometrical index for measuring species diversity**. Ecological Indicators, 9(4):651 – 658, 2009. ISSN: 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2008.07.007.

[50] LAMB, E.; BAYNE, E.; HOLLOWAY, G.; SCHIECK, J.; BOUTIN, S.; HERBERS, J. ; HAUGHLAND, D.. **Indices for monitoring biodiversity change: Are some more effective than others?** Ecological Indicators, 9:432–444, 05 2009. doi: 10.1016/j.ecolind.2008.06.001.

[51] IZSÁK, J.; PAPP, L.. **A link between ecological diversity indices and measures of biodiversity**. Ecological Modelling, 130(1):151 – 156, 2000.

[52] CLARKE, K. R.; WARWICK, R. M.. **A taxonomic distinctness index and its statistical properties**. Journal of Applied Ecology, 35(4):523–531, 1998. doi: 10.1046/j.1365-2664.1998.3540523.x.

[53] CLARKE, K.; WARWICK, R.. **Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. Primer-E Ltd: Plymouth, UK**. 01 2001.

[54] WEITZMAN, M. L.. **On diversity***. Quarterly Journal of Economics, 107(2):363–405, 1992.

[55] FAITH, D.. **Phylogenetic pattern and the quantification of organismal biodiversity**. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 345:45–58, 08 1994.

[56] GONZALEZ, R. C.; WOODS, R. E.. **Digital Image Processing**. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 1992.

[57] SA, A.; CARVALHO, P. C. ; VELHO, L.. **High Dynamic Range Image Reconstruction**. Morgan and Claypool Publishers, 2007.

[58] MENZE, B. H.; JAKAB, A.; BAUER, S.; KALPATHY-CRAMER, J.; FARAHANI, K.; KIRBY, J.; BURREN, Y.; PORZ, N.; SLOTBOOM, J.; WIEST, R.; LANCZI, L.; GERSTNER, E.; WEBER, M.; ARBEL, T.; AVANTS, B. B.; AYACHE, N.; BUENDIA, P.; COLLINS, D. L.; CORDIER, N.; CORSO, J. J.; CRIMINISI, A.; DAS, T.; DELINGETTE, H.; DEMIRALP, ; DURST, C. R.; DOJAT, M.; DOYLE, S.; FESTA, J.; FORBES, F.; GEREMIA, E.; GLOCKER, B.; GOLLAND, P.; GUO, X.; HAMAMCI, A.; IFTEKHARUDDIN, K. M.; JENA, R.; JOHN, N. M.; KONUKOGLU, E.; LASHKARI, D.; MARIZ, J. A.; MEIER, R.; PEREIRA, S.; PRECUP, D.; PRICE, S. J.; RAVIV, T. R.; REZA, S. M. S.; RYAN, M.; SARIKAYA, D.; SCHWARTZ, L.; SHIN, H.; SHOTTON, J.; SILVA, C. A.; SOUSA, N.; SUBBANNA, N. K.; SZEKELY, G.; TAYLOR, T. J.; THOMAS, O. M.; TUSTISON, N. J.; UNAL, G.; VASSEUR, F.; WINTERMARK, M.; YE, D. H.; ZHAO, L.; ZHAO, B.; ZIKIC, D.; PRASTAWA, M.; REYES, M. ; LEEMPUT, K. V.. **The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)**. IEEE Transactions on Medical Imaging, 34(10):1993–2024, Oct 2015. 10.1109/TMI.2014.2377694, 0278-0062.

[59] BAKAS, S.; AKBARI, H.; SOTIRAS, A.; BILELLO, M.; ROZYCKI, M.; KIRBY, J. S.; FREYMANN, J. B.; FARAHANI, K. ; DAVATZIKOS, C.. **Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features**. Scientific data, 4(170117), 2017. doi:10.1038/sdata.2017.117.

[60] CHRISTOS-IRAKLIS TSATSOULIS. **How NOT to perform feature selection!** `https://www.nodalpoint.com/not-perform-feature-selection/`. Acessado em: Mar 2019.

[61] **Scikit-learn v0.20.3 API Reference**. `https://scikit-learn.org/stable/modules/classes.html#module-sklearn.ensemble`. Acessado em: Mar 2019.

[62] **Python**. `www.python.org`. Acessado em: Mar 2019.

[63] KUMAR, V.; GU, Y.; BASU, S.; BERGLUND, A.; ESCHRICH, S. A.; SCHABATH, M. B.; FORSTER, K.; AERTS, H. J.; DEKKER, A.; FENSTERMACHER, D.; GOLDGOF, D. B.; HALL, L. O.; LAMBIN, P.; BALAGURUNATHAN, Y.; GATENBY, R. A. ; GILLIES, R. J.. **Radiomics: the**

process and the challenges. Magnetic Resonance Imaging, 30(9):1234 – 1248, 2012. Quantitative Imaging in Cancer.

[64] MORRIS, E. K.; CARUSO, T.; BUSCOT, F.; FISCHER, M.; HANCOCK, C.; MAIER, T. S.; MEINERS, T.; MÜLLER, C.; OBERMAIER, E.; PRATI, D.; SOCHER, S. A.; SONNEMANN, I.; WÄSCHKE, N.; WUBET, T.; WURST, S. ; RILLIG, M. C.. **Choosing and using diversity indices: insights for ecological applications from the german biodiversity exploratories**. Ecology and Evolution, 4(18):3514–3524, 2014.

[65] KELLY, C.; MAJEWSKA, P.; IOANNIDIS, S.; RAZA, M. H. ; WILLIAMS, M.. **Estimating progression-free survival in patients with glioblastoma using routinely collected data**. Journal of Neuro-Oncology, 135(3):621–627, Dec 2017.

[67] ELLIKA, S.; JAIN, R.; PATEL, S.; SCARPACE, L.; SCHULTZ, L.; ROCK, J. ; MIKKELSEN, T.. **Role of perfusion ct in glioma grading and comparison with conventional mr imaging features**. American Journal of Neuroradiology, 28(10):1981–1987, 2007.

[68] HIWATASHI, A.; HONDA, H.; KIKUCHI, K.; YAMASHITA, K.; YOSHIMOTO, K.; OBARA, M.; VAN CAUTEREN, M.; MIZOGUCHI, M.; SUZUKI, S. O.; IWAKI, T. ; TOGAO, O.. **Differentiation of high-grade and low-grade diffuse gliomas by intravoxel incoherent motion MR imaging**. Neuro-Oncology, 18(1):132–141, 08 2015.

[69] **Multimodal Brain Tumor Segmentation Challenge 2018**. `https://www.med.upenn.edu/sbia/brats2018/tasks.html`. Acessado em: Mar 2019.

[70] LIU, S.; WANG, Y.; XU, K.; WANG, Z.; FAN, X.; ZHANG, C.; LI, S.; QIU, X. ; JIANG, T.. **Relationship between necrotic patterns in glioblastoma and patient survival: Fractal dimension and lacunarity analyses using magnetic resonance imaging**. Scientific Reports, 7, 12 2017.

[71] NOCH, E.; KHALILI, K.. **Molecular mechanisms of necrosis in glioblastoma: The role of glutamate excitotoxicity**. Cancer biology therapy, 8:1791–7, 10 2009.

[72] RAZA, S.; LANG, F.; AGGARWAL, B.; N FULLER, G.; WILDRICK, D. ; SAWAYA, R.. **Necrosis and glioblastoma: A friend or a foe? a review and a hypothesis**. Neurosurgery, 51:2–12; discussion 12, 08 2002.

[73] WU, C.-X.; LIN, G.-S.; LIN, Z.-X.; ZHANG, J.-D.; LIU, S.-Y. ; ZHOU, C.-F.. **Peritumoral edema shown by mri predicts poor clinical outcome in glioblastoma**. World journal of surgical oncology, 13:496, 12 2015.

[74] POPE, W. B.; SAYRE, J.; PERLINA, A.; VILLABLANCA, J. P.; MISCHEL, P. S. ; CLOUGHESY, T. F.. **Mr imaging correlates of survival in patients with high-grade gliomas**. American Journal of Neuroradiology, 26(10):2466–2474, 2005.

[75] A. HAMMOUD, M.; SAWAYA, R.; SHI, W.; THALL, P. ; E. LEEDS, N.. **Prognostic significance of preoperative mri scans in glioblastoma multiforme**. Journal of neuro-oncology, 27:65–73, 02 1996.