

Universidade do Minho
Escola de Engenharia

José Adriano Azevedo da Silva Ribeiro Pinto

**Segmentação do Tecido Tumoral Cerebral
em Imagens de Ressonância Magnética**



Universidade do Minho
Escola de Engenharia

José Adriano Azevedo da Silva Ribeiro Pinto

Segmentação do Tecido Tumoral Cerebral em Imagens de Ressonância Magnética

Dissertação de Mestrado
Mestrado Integrado em Engenharia Biomédica
Ramo de Eletrónica Médica

Trabalho efetuado sob a orientação do
Professor Doutor Carlos Alberto Batista Silva
e da
Professora Doutora Deolinda M. L. D. Rasteiro

Declaração

Nome: José Adriano Azevedo da Silva Ribeiro Pinto

Endereço eletrónico: jadriano.pinto@gmail.com

Número de Cartão de Cidadão: 14173491

Título de Dissertação: Segmentação do Tecido Tumoral Cerebral em Imagens de Ressonância Magnética

Orientadores:

Professor Doutor Carlos Alberto Batista Silva

Professora Doutora Deolinda Maria Lopes Dias Rasteiro

Ano de conclusão: 2015

Designação do Mestrado: Mestrado Integrado em Engenharia Biomédica

Ramo: Electrónica Médica

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA DISSERTAÇÃO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, ___ / ___ / ___

Assinatura: _____

Agradecimentos

Em primeiro lugar quero referir que esta etapa me permitiu contactar com novos problemas, fazendo-me crescer como pessoa. Assim, quero desde já agradecer a todos aqueles que me acompanharam ao longo desta aventura.

Gostaria de expressar a minha profunda gratidão para com o meu orientador Carlos Silva, pelo seu apoio excepcional e confiança depositada em mim. O sucesso alcançado ao longo deste trabalho é resultado do seu aconselhamento e motivação. Espero nunca desapontar.

Quero também agradecer ao Sérgio com o qual tive a oportunidade de partilhar conhecimentos e discutir ideias, ao longo deste trabalho. Imensamente grato pela sua disponibilidade e sugestões apresentadas.

Aos meus colegas e amigos do ramo de Eletrónica Médica, Hugo, Maria João, Meireles e Zamith, um muito obrigado pelos bons momentos passados juntos, pelas horas de trabalho, pela motivação e apoio. Assumir e enfrentar desafios ao vosso lado foi uma honra.

Aos meus amigos Marcos e Nuno, por me fazerem ver que há amizades que nunca se perdem e por me fazerem lembrar que há mais vida para além do trabalho. Muito obrigado por levarem a amizade a sério.

À malta do laboratório, por me acolherem e fomentarem o espírito de amizade, um obrigado. Conseguiram fazer do laboratório um ótimo local de investigação e trabalho.

À família de Biomédica 2010/11, com os quais passei grandes vivências e momentos inesquecíveis.

Por último, mas não menos importante, quero agradecer ao meus pais e ao meu irmão, pelo apoio incondicional e pela capacidade de compreensão, por me ajudarem a alcançar os meus objetivos e me incentivarem a perseguir as minhas ambições. À minha avó, tias Lina e Teresa, bem como o meu tio Filipe e primo Nuno, aos quais agradeço o carinho e conselhos dados.

Adriano Pinto

Resumo

Atualmente, a facilidade de acesso aos cuidados de saúde, aliada a uma melhoria nas técnicas de diagnóstico, permite uma maior e melhor identificação de problemas de saúde de difícil tratamento. Dentro destes problemas destaca-se o cancro, cuja expressão tem vindo a aumentar gradualmente. O cancro é das principais causas de morte em todo o mundo, onde os tumores cerebrais se evidenciam por apresentarem uma das taxas de sobrevivência mais baixa. A identificação e classificação precoce dos diferentes tecidos tumorais existentes é fundamental, para que a intervenção cirúrgica de ressecção tumoral seja possível. Mesmo em tumores num estado avançado, a segmentação tem também relevo para um correto planeamento médico do tratamento e respetiva monitorização.

O procedimento de segmentação manual é um procedimento demorado, que requer bastantes recursos humanos, e sujeito a uma variabilidade inter- e intra-especialista. Por outro lado, métodos de segmentação automática ainda não são uma prática comum em ambiente clínico. Assim, na presente dissertação tem-se como finalidade desenvolver uma metodologia automática de segmentação de tumores cerebrais nos seus diferentes tipos de tecidos: necrose, edema, região ativa contrastante e região ativa não contrastante, com recurso à informação providenciada pelas imagens de ressonância magnética.

Vários autores, ao longo dos últimos 20 anos, projetaram diferentes abordagens na área de segmentação de tumores cerebrais. Ao longo deste tempo, com o surgir de novas ideias de pensar o problema, a evolução dos algoritmos permitiu que algumas das metodologias mais recentes apresentassem resultados robustos e com capacidade de generalização. Recentemente, a tendência das metodologias de segmentação assenta na utilização de classificadores supervisionados, capazes de assegurar uma classificação de múltiplas classes em simultâneo. O método proposto nesta dissertação é também baseado numa classificação supervisionada. Neste tipo de classificação um conjunto de características, extraídas através das imagens de ressonância magnética, procura discriminar os diferentes tipos de tecido a segmentar. Dada a variabilidade associada à técnica de ressonância magnética, etapas prévias de pré-processamento dos dados são implementadas, com o intuito de assegurar uma uniformização da informação presente nas imagens. Através da computação de características locais, de vizinhança e de contexto é treinada uma floresta de decisão (classificador), que posteriormente, na fase de teste, permite a obtenção da segmentação de um paciente, ou até um conjunto de pacientes.

A realização dos testes à metodologia proposta recorreu-se da base de dados pública (*NCI-MICCAI Grand Challenge BraTS 2013*), que permitiu também a validação dos resultados realizada pela sua plataforma *online Virtual Skeleton Database*. Desde a extração das características até à fase de pós-processamento foram testadas várias configurações, procurando sempre otimizar os resultados prévios.

A melhor configuração, testada no conjunto de *Challenge*, do método de segmentação projetado

obteve um valor de Dice de 0,85 para o Tumor Completo, 0,79 para o Tumor Central e 0,75 para a região ativa contrastante. Pela mesma ordem, a métrica *Positive Predictive Value* obteve valores de 0,87, 0,85 e 0,71. Complementarmente, a sensibilidade obtida foi de 0,84, 0,75 e 0,72 para o Tumor Total, Tumor Central e região ativa contrastante respectivamente. Estes resultados permitiram alcançar a 5ª posição, em 35 métodos, no conjunto de dados *Challenge*.

Em suma, o método proposto é capaz de obter resultados competitivos, comparando com métodos presentes na literatura e na plataforma de avaliação providenciada pela base de dados utilizada.

Abstract

Nowadays, the easy access to medical care, as well as the improvements in medical imaging techniques, ensure a better diagnosis capability and also a better identification of health problems of difficult treatment, being cancer among them. Cancer is one of the major causes of death all over the world, where brain tumours stand out by having one of the lowest rates of survival. In this kind of tumour, an early detection and evaluation of the different types of tumorous tissues are key factors in order to guarantee the appropriate treatment planning or the correct surgical intervention (for cancer resection). Therefore, this kind of information is only possible with access to brain tumour segmentation. Even for brain tumours in an advanced stage, tissue segmentation plays an important role in treatment and motorization processes.

Over the past years, automatic brain tumour segmentation, despite not being used as a common practice in clinical routines, has proven to be a viable and robust method to overcome the large costs of human resources and the intra- and inter-rater variabilities associated with manual segmentation. Therefore, this masters thesis' goal is to develop an automatic brain tumour segmentation method capable of distinguishing the different brain tumour tissues: necrosis, oedema, enhanced region and non-enhanced region, based upon the information provided by magnetic resonance images.

Several authors, over the past 20 years, have proposed different approaches in the brain tumour segmentation field. Along this period, the evolution of algorithms and the arise of new ways of thinking allowed the appearance of methods capable of assuring robustness and generalization. Recently, the proposed methods tend to apply supervised classifiers with the capability of multi-label classification. The proposed method is also based on a supervised classifier. The chosen classifier (Extra-Trees) was trained and tested with a set of features (local features, context features and neighbouring features), extracted from magnetic resonance images in order to differentiate the different types of brain tissue present. Due to the high variability of magnetic resonance images, previous to feature extraction, pre-processing stages are implemented in order to assure an uniformity in the analysed information.

The method was tested with the publicly available database from *NCI-MICCAI Grand Challenge BraTS 2013* allowing also an evaluation performed by its online platform Virtual Skeleton Database. From the feature extraction to the post-processing stage, several approaches were tested, with the main goal of improving the previous results.

For the Challenge dataset, the best configuration of the proposed method obtained a Dice score of 0.85 for the complete tumour, 0.79 for the core and 0.75 for the enhanced tumor. The positive predictive value achieved a score of 0.87, 0.85 and 0.71 for the complete tumor, core tumor and the enhanced tumor, respectively. In the same order, the sensitivity was 0.84, 0.75 and 0.82. The results previously presented achieved the 5th among 35 for the Challenge dataset.

To sum up, the method proposed achieved competitive results, being among the top 5 methods in the three different evaluation sets of BraTS 2013, concluding that the method is capable of providing good generalization.

Índice

Agradecimentos	i
Resumo	iii
Abstract	v
Lista de Figuras	xi
Lista de Tabelas	xv
Lista de Acrónimos	xix
1 Introdução	1
1.1 Motivação	1
1.2 Objetivo	2
1.3 Contribuições	2
1.4 Estrutura da Dissertação	3
2 Fundamentos Clínicos	5
2.1 Cancro - Contextualização Mundial	5
2.2 Tumores do Sistema Nervoso Central	7
2.3 Anatomofisiologia do Glioma	8
2.4 Detecção e Diagnóstico	10
2.4.1 Ressonância Magnética	10
2.4.2 Imagem	13
2.4.2.1 Segmentação Manual	13
2.5 Tratamento	15
2.5.1 Histologia e Ressonância Magnética	16
2.6 Sumário	17

3	Estado da Arte	19
3.1	Componentes Típicos na Segmentação de Tumores Cerebrais	19
3.2	Métodos de Segmentação de Tumores Cerebrais	21
3.2.1	Modelos Generativos	22
3.2.2	Modelos Discriminativos	25
3.3	Sumário	29
4	Fundamentos Teóricos	31
4.1	Algoritmo Típico de Segmentação Tumoral	31
4.2	Pré-processamento	32
4.2.1	Correção da Falta de Homogeneidade das Intensidades	32
4.2.1.1	Métodos Prospetivos	34
4.2.1.2	Métodos Retrospectivos	35
4.2.2	Normalização das Intensidades	36
4.3	Aprendizagem Computorizada - Classificação	40
4.3.1	Fundamentos da Floresta de Decisão	40
4.3.1.1	Árvores de Decisão	41
4.3.1.2	Formulação Matemática	42
4.3.1.3	Treino e Teste	43
4.3.2	Floresta de Decisão	45
4.3.2.1	Influência dos Hiper-parâmetros de uma Floresta	46
4.3.3	<i>Extra-Tree</i>	48
4.4	Sumário	49
5	Implementação Prática	51
5.1	Desenvolvimento do Método de Segmentação	51
5.2	Pré-processamento	52
5.2.1	Correção da Falta de Homogeneidade	52
5.2.2	Normalização das Intensidades	52
5.3	Características	53
5.3.1	Características Locais	55
5.3.2	Características de Contexto	56
5.3.3	Características de Textura	57
5.3.3.1	Extração das Características de Haralick	60
5.3.4	Características de Segunda Ordem	61
5.4	Classificação	63
5.4.1	Hiper-parâmetros da Floresta	63
5.4.2	Balanceamento da Floresta	63
5.4.3	Classificação em Dois Estágios	66
5.5	Pós-processamento	67

5.6	Sumário	68
6	Resultados	69
6.1	Base de Dados	69
6.2	Avaliação do Desempenho	70
6.3	Metodologia de Testes	71
6.4	Pré-processamento	71
6.4.1	Estudo das Características de Haralick	72
6.4.2	Efeito das Características	78
6.5	Classificação	80
6.5.1	Efeito do Classificador	80
6.5.2	Estudo dos Híper-parâmetros da <i>Extra-Trees</i>	83
6.5.3	Efeito da Amostragem	86
6.5.4	Estudo da Relevância de T1	88
6.6	Nova Proposta de Normalização	89
6.7	Classificação em Dois Estágios	91
6.8	Pós-processamento	96
6.9	Tempos de computação	99
6.10	NCI-MICCAI Grand Challenge BraTS 2013	100
6.11	Sumário	105
7	Conclusões/Perspetivas Futuras	107
7.1	Conclusão	107
7.2	Perspetivas Futuras	109
	Bibliografia	111
8	Anexos	123
8.1	Características de Haralick	123
8.2	Características de Segunda ordem	128
8.3	Estudo do Classificador	129
8.4	Configuração dos Híper-parâmetros do <i>Extra-Trees</i>	130
8.5	Efeito da Amostragem	132
8.6	Análise da Sequência T1	133
8.7	Novos Parâmetros de Normalização	134
8.8	Classificação em Dois Estágios	135
8.9	Pós-processamento	137

Lista de Figuras

2.1	Estimativas da: (a) Taxa de incidência de cancro no sexo masculino e (b) Taxa de mortalidade do cancro do mesmo sexo por 100 000 pessoas, no ano de 2012. Resultados com a idade estandardizada. O cancro não melanoma da pele é excluído destes registos. Adaptado de [11].	6
2.2	Estimativas da: (a) Taxa de incidência de cancro no sexo feminino e (b) Taxa de mortalidade do cancro do mesmo sexo por 100 000 pessoas, no ano de 2012. Resultados com a idade estandardizada. O cancro não melanoma da pele é excluído destes registos. Adaptado de [11].	6
2.3	Referencial adotado para contextualização dos campos magnéticos aplicados, na realização de uma ressonância magnética.	11
2.4	Átomo com spin: (a) Na ausência de campo magnético e (b) Na presença do campo magnético B_0 , em que possui um movimento de rotação e de precessão	11
2.5	Decaimento do sinal de ressonância magnética após aplicação do pulso de RF de 90° . O decaimento de indução livre é extremamente rápido devido a efeitos da não homogeneidade do campo magnético.	12
2.6	Fatia de ressonância magnética, nas diferentes sequências disponibilizadas pela base de dados, para a realização de uma segmentação cerebral tumoral manual e o respetivo resultado (marcação manual).	15
4.1	Exemplo de falta de homogeneidades: (a) Imagem de ressonância magnética original; (b) Estimação do campo de falta de homogeneidades e (c) Imagem original corrigida. Adaptado de [79].	33
4.2	Análise do histograma de uma imagem com uma distribuição: (a) uni-modal e (b) bimodal. Adaptado de [108]	38
4.3	Exemplo de duas imagens de treino $j = 1, 2$ utilizadas para a procura dos parâmetros do histograma de referência. Adaptado de [110].	39
4.4	Função de mapeamento das intensidades para o caso geral de l percentis. Adaptado de [110].	40
4.5	Estrutura de um grafo com a configuração em árvore. Adaptado de [37].	41

4.6	Resultado produzido por diferentes árvores treinadas em condições aleatórias e portanto distintas. Adaptado de [37].	45
4.7	Mapa de certeza de uma floresta treinada com diferentes funções de divisão dos dados: o modelo linear (a) de eixos alinhados e modelos curvilíneos de linhas orientadas (b) e secções cónicas (c). Adaptado de [37].	47
5.1	Estruturação do algoritmo de segmentação tumoral de gliomas proposto.	51
5.2	Abordagem implementada para a extração do contexto das imagens. Por questões de perceção apenas se encontram 2 dos 6 planos implementados.	57
5.3	Vizinhança bidimensional em torno de um pixel de análise x com as possíveis direções utilizadas para o cálculo das matrizes de coocorrência.	58
5.4	Exemplo de características de um paciente de HGG numa fatia de T1c. Na primeira coluna encontram-se as características base sobre as quais são computadas as características da segunda e terceira coluna de figuras. (a) Intensidades da sequência normalizada; (b) Exponencial; (c) Logaritmo; (d) Extração da média local de 3 sobre as intensidades (a); (e) Extração da média local com $w = 3$ sobre a exponencial (b); (f) Extração do Gradiente Gaussiano, com $\sigma = 3$ sobre as intensidades (c); (g) Extração do Gradiente Gaussiano, com $\sigma = 3$ sobre a exponencial (b); (h) Extração do Gradiente Gaussiano, com $\sigma = 3$ sobre o logaritmo (c).	62
5.5	Quartis das intensidades de cada uma das quatro classes de tecido tumoral (1- necrose; 2 - edema; 3 - tecido ativo não contrastante; 4 - tecido ativo contrastante) de HGG, nas sequências mais relevantes T1c (a) e T2 (b), após normalização cruzada. A linha a vermelho corresponde à mediana das intensidades da classe, ao par que os limites superior e inferior (de cada caixa) designam o quartil 75 % e 25% respetivamente. Os <i>outliers</i> são marcados com pontos em formato '+'.	65
5.6	Quartis das intensidades de cada uma das quatro classes de tecido tumoral (1- necrose; 2 - edema; 3 - tecido ativo não contrastante; 4 - tecido ativo contrastante) de LGG, nas sequências mais relevantes T1c (a) e T2 (b), após normalização. A linha a vermelho corresponde à mediana das intensidades da classe, ao par que os limites superior e inferior (de cada caixa) designam o quartil 75 % e 25% respetivamente. Os <i>outliers</i> são marcados com pontos em formato '+'.	66
5.7	Estruturação do algoritmo de segmentação tumoral em dois estágios.	67
5.8	Aplicação do pós-processamento com remoção de regiões achatadas, em <i>HG_310</i> . (a) Sequência T1c normalizada; (b) Segmentação sem pós-processamento; (c) Segmentação com pós-processamento.	68

6.1	Gráficos de barras da distância de Bhattacharyya das 13 características de textura de Haralick, em fatias das quatro sequências de RM de um único paciente de HGG, nos quatro tipos de tecido tumoral: necrose (azul), edema (laranja), região ativa não contrastante (cinzento) e região ativa contrastante (amarelo). Em (a) encontram-se as características f1, f2, f3, f4, f5 e f6 de Haralick; (b) compreende as características f7, f8, f9, f11, f12 e f13. Por questões de escala e visualização (c) apenas compreende a característica f10.	74
6.2	Gráfico de barras da distância de Bhattacharyya das três características de Haralick: f5, f6 e f11, considerando todo o volume cerebral de todos os pacientes do conjunto <i>Training</i> e todas as classes de tecido existentes: necrose (azul claro), edema (laranja), região ativa não contrastante (cinzento), região ativa contrastante (amarelo) e tecido cerebral saudável (azul escuro).	75
6.3	Gráfico de barras das médias das métricas obtidas relativas as segmentações provenientes de um classificador sem características de textura (cinzento) e um classificador que considera as características de textura associadas às características de primeira ordem, em HGG (a) e em LGG do conjunto <i>Training</i>	77
6.4	Gráficos de barras das médias das métricas obtidas relativas as segmentações provenientes de um classificador com as características de primeira ordem (cinzento) e um classificador com as características de segunda ordem (azul) em HGG (a) e em LGG, no conjunto <i>Training</i>	79
6.5	Gráficos de barras das médias das métricas avaliadas, relativas as segmentações resultantes de uma <i>Extra-Trees</i> (azul) e de uma Floresta de Decisão (cinzento), em HGG (a) e em LGG do conjunto <i>Training</i>	81
6.6	Gráfico de barras das médias das métricas avaliadas, das segmentações provenientes de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica <i>Grid Search</i> (azul) e um classificador com os parâmetros do terceiro classificado (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto <i>Training</i>	84
6.7	Gráfico de barras das médias das métricas avaliadas, de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica <i>Grid Search</i> (azul) e o classificador com os parâmetros aplicados até ao momento, na secção 6.3 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto <i>Training</i>	86
6.8	Gráfico de barras das médias das métricas avaliadas, de um classificador sem amostragem (azul) e o classificador balanceado através do procedimento de amostragem (cinzento), em LGG no conjunto <i>Training</i>	87
6.9	Gráfico de barras das médias das métricas avaliadas, de um classificador que contabiliza todas as sequências da RM disponíveis (azul) e um classificador sem a sequência T1 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto <i>Training</i>	89

6.10	Gráficos de barras das médias das métricas avaliadas com etapas de normalização das intensidades distintas: normalização com a base de dados BraTS2013 (azul) e normalização com a base de dados BraTS2015 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG, no conjunto <i>Training</i>	91
6.11	Gráficos de barras das médias das métricas da região tumoral total, relativas ao classificador multi-classe (azul) e ao primeiro estágio de HEGS (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto <i>Training</i>	93
6.12	Gráficos de barras das médias das métricas de todas as regiões, relativas ao classificador multi-classe (azul) e ao segundo estágio de HEGS considerando a informação da sequência T1 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto <i>Training</i>	94
6.13	Gráficos de barras das médias das métricas de todas as regiões, relativas ao classificador multi-classe (azul) e ao segundo estágio de HEGS excluindo a sequência T1 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto <i>Training</i>	95
6.14	Gráficos de barras das médias das métricas de todas as regiões, comparando as metodologias de segmentação sem a realização da etapa de pós-processamento (azul) e com a aplicação do pós-processamento (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto <i>Training</i>	97
6.15	Resultados relativos ao paciente de <i>Challenge HG_310</i> , à saída do classificador, bem como os resultados provenientes da aplicação das duas versões de pós-processamento distintas propostas: (a) corresponde ao resultado sem aplicação de qualquer pós-processamento; (b) pós-processamento em que apenas é mantido o componente ligado de maiores dimensões; (c) pós-processamento com remoção de regiões achatadas e pequenas marcações. Em (a) a região a vermelho corresponde à região achatada a remover, ao passo que as regiões a cinzento correspondem a tumor cerebral.	98
6.16	Gráfico de barras das médias das métricas de todas as regiões, relativas ao paciente 310 do conjunto <i>Challenge</i> , com a aplicação do pós-processamento sem remoção de regiões achatadas (azul) vs. pós-processamento com remoção de regiões achatadas (cinzento).	99
6.17	Classificação obtida pela metodologia proposta no conjunto <i>Training</i>	103
6.18	Classificação obtida pela metodologia proposta no conjunto <i>Challenge</i>	104

Lista de Tabelas

2.1	Tabela ilustrativa do comportamento imagiológico dos diferentes tecidos presentes no cérebro, quando sujeitos aos dois tipos base de protocolos de Ressonância Magnética.	13
5.1	Exemplo de visualização das intensidades dos diferentes tecidos tumorais, marcados manualmente, nas quatro sequências providenciadas pela base de dados, num tumor de elevado grau.	54
5.2	Gama de valores definidos para a transformação linear das intensidades de FLAIR, T1, T1c e T2 para posterior extração das características exponencial e logaritmo.	56
5.3	Limites de valores das intensidades, em T1c e T2, para a filtragem de amostras de necrose e tecido ativo contrastante, em pacientes HGG com a normalização de LGG.	65
6.1	Definição dos hiper-parâmetros fulcrais do classificador aquando da realização dos diferentes testes.	71
6.2	Tempo médio de computação, para cada paciente, das 13 características, ao longo das 4 direções, por cada sequência.	72
6.3	Tempos de computação totais relativos ao conjunto <i>Training</i> após a realização de todos os <i> folds</i>	81
6.4	Três melhores combinações obtidas pelo <i>Grid Search</i> , num conjunto de 100 candidatos.	83
6.5	Limites de valores das intensidades, em T1c e T2, para a filtragem de amostras de necrose e tecido ativo contrastante, em pacientes HGG com a normalização de LGG, em conformidade com os novos parâmetros de normalização.	90
6.6	Tempo médio de computação, em segundos, para efetuar a segmentação automática de um paciente.	99
6.7	Média de Dice, PPV e sensibilidade no conjunto <i>Challenge</i> da base de dados BraTS 2013.	102
8.1	Sobreposição e Distância de Bhattacharyya, para um único paciente, das características f1 até f7 de Haralick.	124
8.2	Sobreposição e Distância de Bhattacharyya, para um único paciente, das características f8 até f13 de Haralick.	125

8.3	Distância de Bhattacharyya, de f5, f6 e f11, computada tendo em consideração todos os pacientes do conjunto de <i>Training</i> , em ambos os níveis de glioma.	126
8.4	Média das métricas obtidas relativas as segmentações provenientes de um classificador com textura e um classificador sem textura.	127
8.5	Média das métricas obtidas relativas as segmentações provenientes de um classificador com as características de 1ª ordem e um classificador com as características de 2ª ordem (no conjunto de <i>Training</i>).	128
8.6	Média das métricas obtidas relativas as segmentações provenientes de uma Floresta de Decisão e de uma <i>Extra-Trees</i>	129
8.7	Média das métricas obtidas relativas as segmentações provenientes de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica <i>Random-Search</i> e um classificador com os parâmetros do terceiro classificado (no conjunto de <i>Training</i>). . . .	130
8.8	Média das métricas obtidas relativas as segmentações provenientes de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica <i>Random-Search</i> e o classificador com os parâmetros aplicados até ao momento, na secção 6.3 (no conjunto de <i>Training</i>).	131
8.9	Média das métricas, avaliadas pelo concurso, na ausência/presença da amostragem. A diferença entre estas duas abordagens permite quantificar o impacto da realização da amostragem.	132
8.10	Média das métricas obtidas relativas as segmentações de um <i>Extra-Trees</i> que contabiliza características da sequência T1 e um <i>Extra-Trees</i> que não considera o protocolo em estudo (no conjunto de <i>Training</i>).	133
8.11	Média das métricas obtidas das segmentações resultantes de duas etapas de normalização das intensidades distintas. Uma primeira onde a fase de treino da normalização considera o conjunto de treino BraTS2013 e a restante engloba um número mais alargado de pacientes (BraTS2015). Denote-se que ambos os resultados dizem respeito ao conjunto <i>Training</i>	134
8.12	Média das métricas da região de tumor total resultantes da aplicação de um único estágio de classificação (utilizando todas as sequências) vs. classificação binária (na ausência da T1). Resultados relativos ao conjunto <i>Training</i>	135
8.13	Média das métricas, em todas as classes, relativas as segmentações de um <i>Extra-Trees Multi-Label</i> vs. média das métricas da metodologia de classificação em dois estágios. Na metodologia em dois estágios é considerada a sequência T1 no segundo estágio. Denote-se que ambos os resultados dizem respeito ao conjunto <i>Training</i>	135
8.14	Média das métricas, em todas as classes, relativas as segmentações de um <i>Extra-Trees Multi-Label</i> vs. média das métricas da metodologia de classificação em dois estágios. Na metodologia em dois estágios, em ambas as fases de classificação não é considerada a sequência T1. Denote-se que ambos os resultados dizem respeito ao conjunto <i>Training</i>	136
8.15	Média das métricas obtidas relativas as segmentações sem pós-processamento e com pós-processamento.	137

8.16 Média das métricas, avaliadas pelo concurso, na ausência/presença da amostragem. A diferença entre estas duas abordagens permite quantificar o impacto da realização da amostragem.	137
---	-----

Lista de Acrónimos

CA *Celular Automata*. 24

CIC *Cavidade Intra-Craniana*. 22

CRF *Conditional Random Fields*. 25

DoG *Difference of Gaussians*. 54, 59, 75

DP *Densidade de Protões*. 12, 22

DRF *Discriminative Random Fields*. 25

EM *Expectation Maximization*. 23

Extra-Trees *Extremely Randomized Trees*. xiv, 61, 78–80, 86, 88, 91, 98, 122, 125, 127, 128

FCM *Fuzzy C-Means*. 28

FID *free induction decay*. 12

FLAIR *Fluid Attenuated Inversion Recovery*. 2, 23, 26–28, 65

HEGS *Hierarchical Extra-Trees for Glioma Segmentation*. 64, 89

HGG *High-Grade Glioma*. x–xiii, 9, 15, 17, 60–63, 65, 67, 69, 70, 72, 74–77, 79–84, 86–91, 93–95

i.i.d. *Independente e Identicamente Distribuído*. 21, 26

KB *Knowledge-Based*. 21, 22, 28

LGG *Low-Grade Glioma*. x–xiii, 9, 61–65, 67, 69, 74–77, 79, 82–95

LS *Level Set*. 24

MCD *Minimum Covariance Determinant*. 23

MRF *Markov Random Field*. 23, 25

N3 *Nonparametric Nonuniform intensity Normalization*. 34

N4ITK *Improved N3 Bias Correction*. 34

OMS *Organização Mundial de Saúde*. 5–10, 16, 17

PPV *Positive Predicted Value*. xiii, 67, 68, 77, 80, 84, 88, 90–92, 96, 98–100

RF *Radiofrequência*. ix, 12, 13

RM *Ressonância Magnética*. xi, 10, 12, 13, 15, 20–32, 34, 35, 37, 49–51, 53, 54, 58, 63, 67, 70, 72, 86, 87, 94, 95, 115

SVM *Support Vector Machine*. 24–27

T1c *Ressonância Magnética T1 com agente contrastante*. 2, 23, 26–28, 65

TE *Tempo de Eco*. 12, 13

TI *Tempo de Inversão*. 13

TNM *Tumour Nodes and Metastasis*. 8

TR *Tempo de Repetição*. 12, 13

VSD *Virtual Skeleton Database*. 96, 97

Introdução

O presente capítulo apresenta a motivação sobre a qual assenta o desenvolvimento do trabalho exposto e objetivos associados. Complementarmente, encontram-se também as contribuições realizadas e a estruturação da dissertação.

1.1 Motivação

De acordo com o último estudo realizado por Ferlay *et al.* [1] em 2012, os tumores cerebrais, juntamente com os tumores associados ao sistema nervoso central são a décima sétima causa de morte, mais comum, por cancro, no mundo. No estudo em questão, os tumores cerebrais e do sistema nervoso central apresentam um aumento de cerca de 2% de novos casos (256 000) e 2.3% do total de mortes tumorais são associadas a estas classes de tumores. Em termos de taxa de incidência e mortalidade, esta é superior nos países desenvolvidos, em parte explicada pela facilidade de acesso ao diagnóstico [1].

Gliomas originários das células astrócitas, oligodendrogiais e ependimiais correspondem aproximadamente a 70% dos tumores cerebrais [2]. A classe mais frequente e mais agressiva do glioma, o glioblastoma, compreende 65% dos casos identificados [3]. Embora se desconheça a causa, este tipo de cancro tem maior expressão no sexo masculino do que no sexo feminino [4]. O prognóstico de pacientes portadores de glioma é bastante reservado, sendo que menos de 3% dos pacientes com glioblastomas resistem aos primeiros 5 anos após a deteção do tumor [3]. Com a introdução da ressonância magnética e das técnicas computadorizadas de tomografia, a identificação e diagnóstico de tumores cerebrais aumentou consideravelmente. Com a introdução das referidas técnicas de imagem médica ocorreu um aumento abrupto da taxa de incidência deste tipo de tumores, que foi posteriormente explicado pelas melhorias na qualidade do diagnóstico. Atualmente as taxas de incidência encontram-se estáveis [1–3].

Assim sendo, atualmente, o foco relativo à identificação/tratamento dos tumores cerebrais é uma rápida identificação e distinção entre os diferentes tipos de tecido tumoral, ou seja, a segmentação. A segmentação manual é um processo temporalmente exigente, que requer grande precisão. O procedimento de monitorização tumoral assenta fundamentalmente na medição do respetivo volume, sendo com base nesta informação planeados o tratamento e possível intervenção cirúrgica [5]. Os gliomas, de baixo grau e de elevado grau, são estruturas passíveis de grande variabilidade no seu formato e aparência, pro-

vocando por vezes deformações de estruturas adjacentes. Deste modo, a segmentação manual é uma tarefa ainda mais árdua de se realizar, e com variabilidade intra- e inter-especialista consideráveis [5, 6].

1.2 Objetivo

Um glioma apresenta na sua constituição, geralmente, quatro tipos de tecidos tumorais de comportamento celular distinto. Uma região ativa, desdobrada em contrastante e não contrastante, responsável pela expansão do tumor. A região ativa circunda a porção de tecido tumoral morto, ao qual se atribui a designação de região necrótica ou cística do tumor. Na grande maioria dos gliomas dada a atividade celular da região ativa, esta é rodeada por um agregado de células tumorais infiltrativas, designado de edema. Nem todos os tipos de tumores apresentam os quatro tipos de tecidos descritos [7]. A título de exemplo, nos gliomas de baixo grau é comum identificar-se apenas a presença de região ativa não contrastante e edema [7, 8].

Neste trabalho procura-se o desenvolvimento de uma metodologia de segmentação automática dos diferentes tecidos que constituem o glioma, tendo por base a informação providenciada por quatro sequências de ressonância magnética (FLAIR, T1, T1c e T2). A utilização de diversas sequências prende-se com o facto de diferentes estruturas tumorais possuírem comportamentos próprios e distintos dos restantes tipos de tecidos presentes, daí que a conjugação de toda a informação seja benéfica. É com base na informação presente nas imagens de ressonância magnética que são extraídos conjuntos de características, o mais discriminativas e adequadas ao problema da segmentação. As imagens utilizadas para avaliar o desempenho da metodologia proposta encontram-se publicamente disponíveis, sendo parte integrante do *Grand Challenge on Multimodal Brain Tumor Segmentation (BraTS 2013)*, organizado por K. Farahani, M. Reyes, B. Menze, E. Gerstner, J. Kirby e J. Kalpathy-Cramer, em parceria entre a conferência MICCAI (*Medical Imaging Computing and Computer Assisted Intervention*) e NCI (*National Cancer Institute*).

O desenvolvimento desta metodologia requer o domínio de diversas áreas do conhecimento, nomeadamente a área do processamento de imagem, técnicas avançadas de programação, *machine learning* e a área das ciências da saúde.

1.3 Contribuições

Ao longo desta dissertação foram desenvolvidos alguns métodos que, no conhecimento do autor, podem ser considerados contribuições originais, nomeadamente:

- Um método de segmentação automática de tumores cerebrais, baseado num classificador supervisionado (Floresta de Decisão), onde se destaca uma nova abordagem para a extração de características de contexto.
 - Método presente no artigo da quarta conferência portuguesa IEEE em Bioengenharia [9].
- Desenvolvimento de características de segunda ordem capazes de assegurar uma melhor delimitação da região tumoral, bem como a distinção entre as diferentes classes de tumor. A extração de

características tendo por base a informação fornecida pela aplicação de não-linearidades (transformações exponencial e logaritmo) sobre as intensidades, demonstrou ter resultados positivos. Assim a utilização de pseudo-sequências comprovou ser um fator chave para posterior computação de características com maior capacidade de discriminação.

- O método em causa encontra-se descrito no artigo da conferência EMBC 2015 [10].
- Aplicação de uma metodologia de classificação com duas *Extra-Trees* (tipo de Florestas de Decisão), que se recorrem apenas de 3 das 4 sequências disponibilizadas pela base de dados. Classificação em dois estágios, que para além dos recursos computacionais serem consideravelmente menores, a segmentação de tecidos tumorais apresentou-se mais robusta. Destacando-se uma distinção entre tecido saudável e tecido tumoral mais eficaz.
- Metodologia submetida ao *NCI-MICCAI – Grand Challenge BraTS 2013* obtendo-se o 5/35 lugar no conjunto *Challenge* e 4/38 lugar no conjunto *Training*¹.

1.4 Estrutura da Dissertação

A presente dissertação encontra-se dividida em 7 capítulos. Primeiramente na introdução encontram-se a motivação, contribuições e objetivos do trabalho. O segundo capítulo apresenta uma contextualização clínica sobre o cancro, contendo dados demográficos relativos à taxa de incidência e taxa de morte. Complementarmente, encontram-se também descritas a anatomofisiologia dos gliomas, a sua identificação e tratamento, sendo ainda expostas as etapas que definem o procedimento de segmentação manual. De seguida, no capítulo 3 descrevem-se e analisam-se os diferentes métodos de segmentação automática de tumores (estado da arte). Posteriormente, o capítulo 4 contém a fundamentação teórica necessária para compreender e desenvolver uma metodologia de segmentação automática, assente numa classificação supervisionada. Conceitos desde o pré-processamento, extração de características, segmentação através de uma floresta de decisão, até ao pós-processamento são introduzidos. O capítulo 5 diz respeito à implementação do algoritmo nos seus diferentes estágios. De seguida, no capítulo 6 é descrita a modalidade de testes, sendo também expostos os diversos estudos efetuados. Os estudos em causa cobrem as diferentes etapas que constituem o classificador, com o intuito de aumentar continuamente os resultados produzidos. Por último, no capítulo 7 encontram-se as principais conclusões e possíveis sugestões de trabalho futuro.

¹<https://www.virtualskeleton.ch/BRA/TS/Start2013> - consultado em 20/05/2015

Fundamentos Clínicos

Nesta secção procura-se dar uma noção geral do impacto do cancro na sociedade, bem como o trabalho que tem vindo a ser desenvolvido e implementado pelas organizações responsáveis. Seguidamente, focam-se as atenções para a compreensão dos princípios fisiológicos e anatómicos que dizem respeito a um tipo de cancro cerebral, o glioma. Adicionalmente, é efetuada uma referência à modalidade médica mais adequada para o diagnóstico tumoral cerebral, com o intuito de conhecer o comportamento deste tipo de cancro.

2.1 Cancro - Contextualização Mundial

A designação cancro é um termo que compreende um vasto grupo de doenças que podem afetar qualquer parte corporal. Designações mais restritas como neoplasmos e tumores são correntemente utilizadas. O termo neoplasmo compreende uma fase inicial do desenvolvimento de um cancro, isto é, o surgir de uma região de células com um comportamento anormal. Anormal, porque a referida região de células cresce para além da sua fronteira natural. O neoplasmo adquire a designação de tumor, aquando da existência de diferentes tipos de tecido na estrutura cancerígena, sinónimo de uma diferenciação do neoplasmo, que tem a capacidade de invadir estruturas próximas e alastrar a órgãos próximos, sendo que este último evento é classificado de metastização. A capacidade de metastizar é uma propriedade de um tipo de tumor, o tumor maligno, afetando geralmente órgãos vizinhos e o sistema linfático. A presença de metástases é a maior causa de morte por cancro [4].

Em conformidade com os dados da Organização Mundial de Saúde¹ (OMS) em 2008, 7,6 milhões de pessoas morreram de cancro, correspondendo a uma percentagem de 13% de todas as mortes a nível global. O último estudo realizado, pela mesma organização, em 2012, indica que 14,1 milhões de novos casos de cancro foram detetados, 8,2 milhões de mortes se devem ao cancro e um total de 32,6 milhões de pessoas combatem esta patologia (indivíduos com menos de 5 anos após diagnóstico) [1, 2]. Tendo em conta estas estimativas, convém destacar que os países menos desenvolvidos englobam 57% dos novos casos (correspondendo a cerca de 8 milhões), 65% das mortes (5,3 milhões) e 48% (15,6 milhões) de casos de cancro com 5 anos de prevalência. Por outro lado, ao analisar estas percentagens e a sua

¹Termo inglês: *World Health Organization*

evolução ao longo dos sucessivos estudos realizados, verifica-se que o aumento da facilidade em disponibilizar meios de diagnóstico e recursos humanos nos países desenvolvidos são fatores preponderantes para uma diminuição contínua da expressão deste tipo de patologias [1].

No que diz respeito à taxa de incidência, agrupada por género, esta é 25% superior nos homens do que nas mulheres. O que em grandezas numéricas corresponde a uma taxa de 205/100 000 no caso do sexo masculino e 165/100 000 para o sexo oposto. A taxa masculina apresenta grande variabilidade entre regiões, ao contrário do que se regista para o sexo feminino [1].

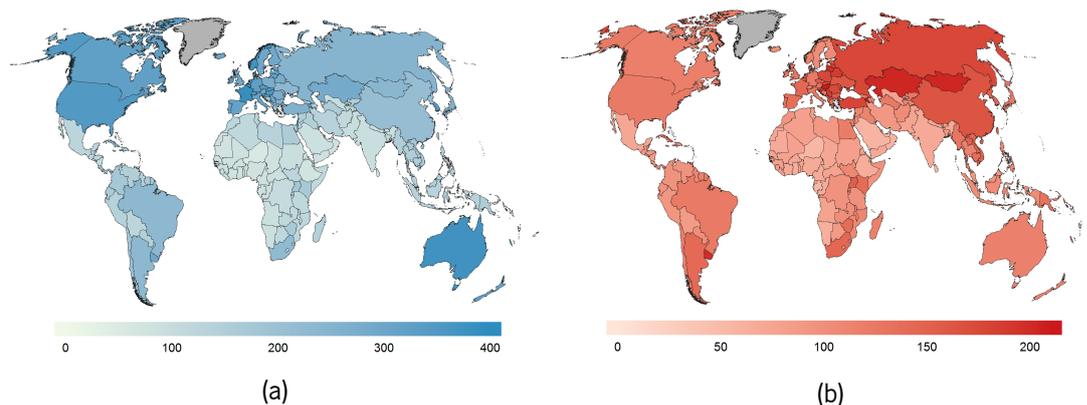


Figura 2.1: Estimativas da: (a) Taxa de incidência de cancro no sexo masculino e (b) Taxa de mortalidade do cancro do mesmo sexo por 100 000 pessoas, no ano de 2012. Resultados com a idade estandardizada. O cancro não melanoma da pele é excluído destes registos. Adaptado de [11].

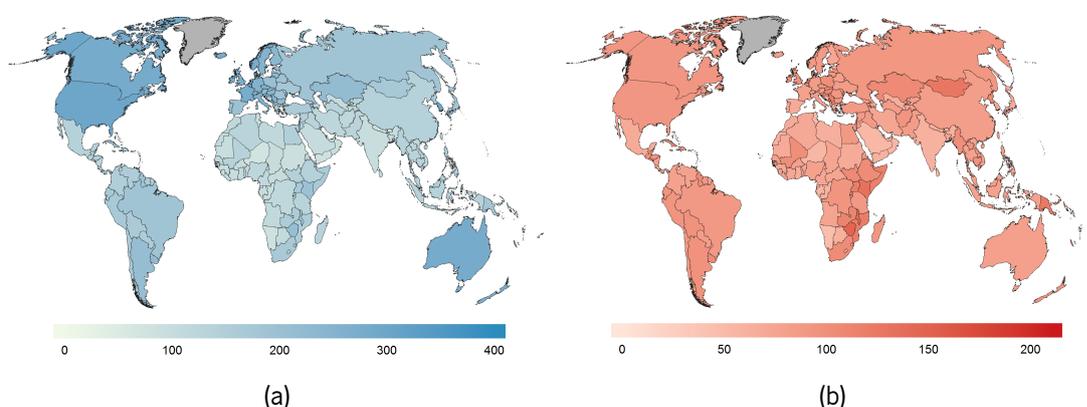


Figura 2.2: Estimativas da: (a) Taxa de incidência de cancro no sexo feminino e (b) Taxa de mortalidade do cancro do mesmo sexo por 100 000 pessoas, no ano de 2012. Resultados com a idade estandardizada. O cancro não melanoma da pele é excluído destes registos. Adaptado de [11].

Dentro da classe de tumores, os tumores cerebrais e do sistema nervoso central são responsáveis pelo aparecimento de 256 000 novos casos e 189 000 mortes, no ano de 2012, contabilizando cerca de 1,8% dos novos casos de cancro e 2,3% das mortes por cancro. A incidência e mortalidade são taxas mais elevadas nos países desenvolvidos, em parte explicado pela maior facilidade de acesso ao diagnóstico [1, 3].

Um dos principais objetivos da OMS é a diminuição do impacto do cancro na sociedade. A organização enuncia que este problema pode ser reduzido e controlado através da implementação de estratégias de

controlo e prevenção de cancro. A OMS defende que grande parte dos cancros apresentam grande probabilidade de cura se detetados e tratados precocemente. Assim sendo, propôs um conjunto de medidas à escala mundial, sendo elas [3, 4]:

- **Prevenção** — evitar um conjunto de fatores de risco, que demonstram estar relacionados com o aparecimento do cancro, como por exemplo o tabaco, radiação ionizante e não ionizante e a contaminação por hepatite B. A OMS estima que pelo menos um terço dos cancros podem ser prevenidos se os fatores de risco forem tidos em conta.
- **Deteção Precoce** — O registo de uma região neoplásica no seu estado inicial apresenta grandes possibilidades de cura. Assim, a consciencialização da população para os primeiros sintomas, bem como a disponibilização dos meios de diagnóstico apropriados são filosofias chave.
- **Tratamento** — Após o correto diagnóstico, o tratamento selecionado deve ser o adequado ao tipo de cancro a tratar, de modo a atuar fundamentalmente sobre a região de interesse e danificar o mínimo de tecido saudável. O regime de tratamento engloba uma ou mais modalidades, como a cirurgia, radioterapia e/ou quimioterapia. O tratamento procura curar o cancro ou prolongar a vida do paciente, evitando uma degradação da sua qualidade de vida. Os cuidados paliativos desempenham um papel importante neste último tópico.
- **Cuidados Paliativos** — Direcionados para indivíduos já numa fase avançada do cancro, e por isso com menores probabilidades de cura. Os cuidados procuram fundamentalmente aliviar o desconforto e evitar o surgir de outras complicações, ao invés da cura. Assim sendo, o tratamento paliativo procura evitar a deterioração da qualidade de vida do paciente.

A perspetiva futura da OMS aponta que, no ano de 2030, o número de novos casos de cancro será de cerca 20,3 milhões, contabilizando ambos os sexos e 13,2 milhões de mortes. As previsões indicam para um aumento de 5 milhões de mortes relacionadas com cancro (quando comparados com os dados de 2012). O número de novos casos detetados aumenta cerca de 6,2 milhões (contabilizando ambos os sexos). A previsão realizada tem em consideração o efeito demográfico. Deste modo, dado o impacto crescente que o cancro tem vindo a ter na sociedade é imperativo apostar em metodologias mais imediatas de deteção de cancro, que por sua vez possibilitará um diagnóstico mais eficaz, aumentando as probabilidades de erradicação da doença [1].

2.2 Tumores do Sistema Nervoso Central

O sistema nervoso central, até ao momento, engloba tumores do tecido neuroepitelial, tumores dos nervos cranianos e para-espinais, tumores das meninges, linfomas e neoplasmos hematopoiéticos, tumores das células germinativas, tumores da sela turca e tumores metastáticos [12].

A realização do estadiamento tumoral, isto é, conhecer o estágio em que o tumor se encontra, é um procedimento fulcral com o intuito de determinar a existência de metástases e auxiliar no planeamento e intervenção médica, bem como o estabelecimento de uma base de referência para posteriores análises comparativas. O tumor é classificado em 4 fases distintas, de acordo com os critérios da OMS, tendo-se as seguintes fases [12]:

- Fase I — Etapa inicial, pelo que o tumor está confinado a uma dada porção. Lesão onde as células apresentam uma divisão celular descontrolada, mas com baixo potencial de proliferação. Região que exige mais do sistema sanguíneo e linfático do que as estruturas que a rodeiam, pois a divisão celular é um processo com gastos energéticos. O tratamento deste tipo de tumor é alcançável através da intervenção cirúrgica (remoção do tecido).
- Fase II — Os neoplasmos deste tipo são definidos pela sua capacidade de infiltração e, apesar do seu baixo nível de proliferação é comum esta existir. Nesta fase os neoplasmos tendem a evoluir para níveis malignos, sendo-lhes atribuída uma das duas fases restantes.
- Fase III — Tumor com evidência histológica de malignidade e com grande capacidade de difusão e expansão do seu volume, afetando mais do que uma estrutura. Apresentam um aumento da atividade celular, uma grande atividade mitótica e atipia nuclear (detecção anormal de núcleos celulares).
- Fase IV — Tecido de citologia maligna, caracterizado pelas suas regiões de proliferação microvascular e/ou necrose. O tumor apresenta metástases, ou seja, a proliferação engloba mais do que um órgão. Maior parte das vezes são associados a uma evolução rápida pré e pós-operação, tendo consequências fatais.

A classificação do tumor numa das quatro fases enunciadas assenta na informação proveniente da realização do estadiamento. O estadiamento é um procedimento que fornece um panorama provável do comportamento do organismo a determinadas terapias. Um estadiamento preliminar, quando realizado, é fundamentalmente direcionado para a contabilização da porção maligna existente num neoplasmo. Apenas uma classificação mais cuidada, realizada posteriormente, é que permite a atribuição de uma fase à porção tumoral em análise. A este último estadiamento designa-se de método TNM. A sigla TNM designa 3 componentes: T - a extensão do tumor primário, N - presença de metástases no sistema linfático local, M - detecção de metástases em locais distantes. O desenvolvimento de critérios na atribuição das fases do tecido tumoral é um bom indicador para a estimativa do prognóstico [13].

2.3 Anatomofisiologia do Glioma

O glioma, tal como a própria designação o sugere, é um cancro originário numa família de células de suporte cerebral muito específicas, as células da glia. As células da glia compreendem vários tipos de células, destacando-se os astrócitos, os oligodendrócitos, células de Schwann, células da microglia e as células endimócitas. As células da glia são uma constante em todo o sistema nervoso, quer no sistema nervoso periférico quer no sistema nervoso central. A propriedade comum a todos os diferentes tipos de células supracitadas é desempenhar a função de suporte e proteção às células nervosas, quer através da capacidade de fornecimento de nutrientes e oxigénio, quer através da proteção contra agentes patogénicos [14].

O comportamento imagiológico dos gliomas é muito variado em termos de forma, tamanho, aparência e localização, sendo um tipo de tumores que apresenta multifocalidade e multi-centralidade. A multi-centralidade ocorre quando se detetam várias lesões tumorais em diferentes locais do cérebro, todas elas

com origem independente. Já a multifocalidade consiste na difusão de uma lesão original para restantes áreas cerebrais [15].

Em conformidade com o que foi descrito na secção anterior os gliomas são divididos em 4 classes distintas, de acordo com as suas características histológicas. Assim sendo, tem-se alguns exemplos de gliomas separados nas quatro fases, na [12, 16]:

- Fase I — Astrocitoma pilocítico;
- Fase II — Astrocitoma difuso e Oligodendroglioma;
- Fase III — Astrocitoma e Oligodendrogliomas anaplásticos; e
- Fase IV — Glioblastoma. O glioblastoma compreende outras modalidades pouco vulgares como os gliossarcomas, glioblastoma de células gigantes, glioblastomas de células pequenas e glioblastomas com uma componente oligodendroglial.

O último tipo de glioblastoma, referido na fase IV, gera alguma controvérsia na comunidade médica. Este tipo de glioblastoma é caracterizado por uma região necrótica, que lhe confere uma agressividade tumoral elevada. Patologistas defendem que este tipo de glioblastoma merece uma classe própria (reconhecimento como uma nova entidade por parte da OMS), em vez de estar embebida nos glioblastomas. Porém, atualmente, e fundamentalmente devido à ausência de um volume de casos estatisticamente considerável, o glioblastoma com componente oligodendroglial permanece na classe glioblastoma. A subcategorização das fases do estadiamento não devem ser menosprezadas dado que o planeamento e intervenção médica são mais robustos [16].

O glioma de elevado grau² (HGG) é dos tumores cerebrais malignos mais comuns, e quando comparado com os restantes tumores, despoletados em diferentes órgãos, é considerado dos tumores mais agressivos, visto que o tempo médio de vida dos pacientes com este tipo de cancro é reduzido (12 a 18 meses). Nos Estados Unidos, dos 22 500 casos anuais de tumores cerebrais malignos primários, cerca de 70% correspondem a HGG. O HGG é o termo utilizado para englobar as fases III e IV, do estadiamento tumoral. Por sua vez, as fases I e II definem os gliomas de baixo grau³ (LGG), cuja esperança média de vida é bastante superior. Por exemplo, indivíduos com gliomas anaplásticos (fase II) têm uma esperança média de vida de 2 a 5 anos [17].

Os tumores são constituídos por uma zona ativa e um edema e por vezes apresentam ainda uma zona necrótica. A região necrótica é mais comum nos gliomas de elevado grau, como o caso do oligodendroglioma. Quanto à zona ativa, esta pode ainda ser dividida em duas classes distintas de tecido: a zona contrastante do tumor e a zona não contrastante do tumor. Em termos fisiológicos a zona não contrastante do tumor representa a progressão da zona ativa, invadindo o edema, enquanto a zona contrastante corresponde ao núcleo da zona ativa. A definição da zona ativa e do edema, independentemente do grau do tumor em causa, são fundamentais para o diagnóstico e posterior tratamento [18].

Estudos recentes apontam que transformações malignas nas células da glia se devem em grande parte a acumulações sucessivas de aberrações genéticas e ainda à desregulação de sequências específicas relacionadas com o crescimento celular. Os fatores responsáveis pelo despoletar desta neoplasia, na grande maioria dos casos são desconhecidos. Até ao momento, o único fator de risco comprovado é a

²Termo inglês: *High Grade Glioma*

³Termo inglês: *Low Grade Glioma*

exposição a radiação ionizante, vulgar em tratamentos cancerígenos. Vários estudos têm sido levados a cabo para compreender o que de facto tem influência no desenvolvimento dos gliomas, como por exemplo o contributo da utilização dos telemóveis, exposição a campos magnéticos e ainda o tabagismo. Nenhum dos casos demonstrou ser um fator de risco no desenvolvimento deste tipo de tumor cerebral [19–21]. Em conformidade com um estudo realizado por Wen e Kesari [22] apenas 5% dos gliomas têm histórico familiar.

No mundo ocidental o glioma tem despertado cada vez mais a atenção médica devido ao aumento da incidência deste tipo de tumor. O referido aumento pode ser justificado pelos progressos na área da medicina, que permitem a acessibilidade a técnicas imagiológicas com maior capacidade de deteção deste tipo de estruturas. Complementarmente o aumento da esperança média de vida, devido à erradicação e tratamento de determinadas doenças torna possível o desenvolvimento de novas doenças [1, 2].

2.4 Deteção e Diagnóstico

De acordo com a OMS, o diagnóstico e prognóstico de tumores cerebrais é mais robusto quando é realizada a biopsia à zona com indícios de região tumoral [12]. Esta fase de diagnóstico é precedida por uma fase preliminar de deteção, onde intervêm as técnicas de imagiologia, de entre as quais se destacam a Tomografia Axial Computadorizada (TAC), a Tomografia de Emissão de Positrão (TEP) e a Ressonância Magnética (RM). Seguindo a linha de pensamento sobre a qual a presente dissertação foi desenvolvida, apenas será abordado como se processa a deteção e diagnóstico para a RM, após uma descrição prévia do seu modo de funcionamento e o comportamento.

2.4.1 Ressonância Magnética

A Ressonância Magnética é um método de obtenção de imagem médica minimamente invasivo e não nocivo, que garante uma elevada resolução e bom contraste. Tendo em conta a designação desta técnica imagiológica, facilmente se retira que opera com campos magnéticos que não têm, até ao momento, qualquer grau de nocividade associados. Por vezes, a RM é auxiliada com a aplicação de um agente contrastante daí que seja uma técnica minimamente invasiva [23].

O princípio base da RM é o conceito de *spin* nuclear, que é uma capacidade característica de determinados núcleos atómicos. O *spin* consiste num movimento de rotação de origem magnética, presente em átomos cujo núcleo contém um número ímpar de prótons/neutrões. Devido a esta desproporcionalidade o próton desencadeia um movimento giratório em torno do seu próprio eixo, o que permite a criação de um campo magnético de pequeno valor, que se designa de *spin* magnético. Na ausência de qualquer perturbação magnética externa o somatório dos *spins* magnéticos de um átomo é nulo ($M_{total} = 0$). Átomos como o carbono, flúor e o hidrogénio têm *spin* nuclear. Os três átomos referidos são constituintes fundamentais do organismo humano, principalmente o hidrogénio (H). Na ressonância magnética, o hidrogénio é então o átomo de eleição a ser excitado por um campo magnético externo, dada à abundância de água no organismo (H_2O) [23].

Com a aplicação de um campo magnético estático e de valor elevado, os *spins* magnéticos dos átomos de hidrogénio são alterados para que $M_{total} \neq 0$. A este campo magnético externo designa-se por B_0 . Em conformidade com a notação adotada, B_0 corresponde ao eixo z , tal como se pode constatar no referencial da figura 2.3. A finalidade da aplicação deste campo magnético é a polarização das estruturas atômicas, para que estas iniciem um movimento de precessão. O movimento de precessão é caracterizado pela sua frequência, a frequência de precessão que não é mais do que a frequência na qual o átomo apresenta um movimento sobre si mesmo e simultaneamente uma rotação em torno do eixo B_0 , como se pode visualizar na figura 2.4 [23].

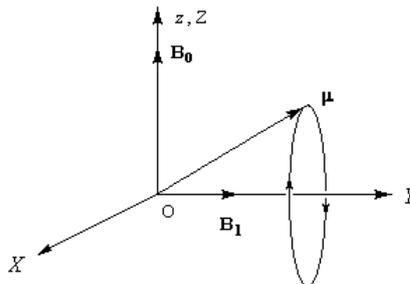


Figura 2.3: Referencial adotado para contextualização dos campos magnéticos aplicados, na realização de uma ressonância magnética.

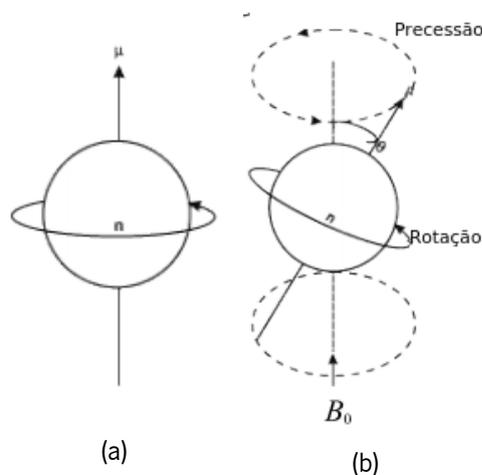


Figura 2.4: Átomo com spin: (a) Na ausência de campo magnético e (b) Na presença do campo magnético B_0 , em que possui um movimento de rotação e de precessão

A frequência de precessão é alternativamente designada de frequência de Larmor. A orientação dos *spins* magnéticos na presença de B_0 pode ser de dois tipos, paralela e antiparalela. O M_{total} é neste momento paralelo a B_0 , o que impossibilita o registo de qualquer tipo de campo magnético. Esta incapacidade é explicada pelo facto de M_{total} ser de baixo valor energético e portanto não mensurável. Deste modo, um pulso de radiofrequência (RF), de 90° relativamente ao campo B_0 , que se encontra em ressonância com a frequência de Larmor é aplicado, originando-se um campo magnético B_1 . A onda de RF de 90° impõe que todos os prótons fiquem num estado de elevada energia e com uma magnetização puramente transversal (orientação antiparalela) [23]. Complementarmente, um pulso de RF de 180° é

aplicado para que o vetor de magnetização transversal demore mais tempo a regressar ao seu estado fundamental.

Após a aplicação dos pulsos de RF, estes são retirados para que se registre a fase de decaimento. A fase de decaimento é que permite a construção de uma imagem de RM. Os prótons ao regressar à sua posição de repouso (estado desorganizado) são radiotransmissores, sendo este o tipo de sinal captado pela instrumentação do equipamento de RM. O sinal registado corresponde ao sinal de ressonância magnética conhecido como decaimento de indução livre (FID – *free induction decay*). Este fenómeno pode ser visualizado na figura 2.5 [23].

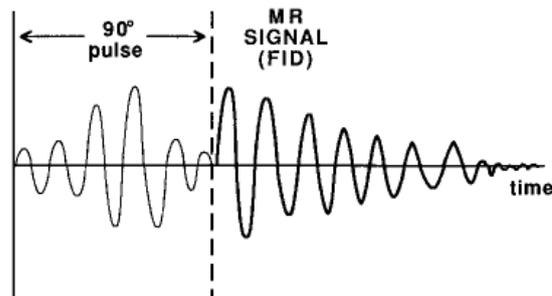


Figura 2.5: Decaimento do sinal de ressonância magnética após aplicação do pulso de RF de 90° . O decaimento de indução livre é extremamente rápido devido a efeitos da não homogeneidade do campo magnético.

A forma como os tecidos se distinguem depende do tempo de relaxação, após ausência do pulso de RF. A relaxação dos átomos pode ser dividida em dois tipos: [23]

- T2 – primeiro evento a registar-se, onde os prótons que se encontram em fase passam a um estado desorganizado, levando a um decréscimo da magnetização transversal;
- T1 – prótons no estado de elevada energia regressam ao estado de menor energia, aumentando a magnetização longitudinal. A energia libertada sob a forma de calor é absorvida pelos tecidos circundantes.

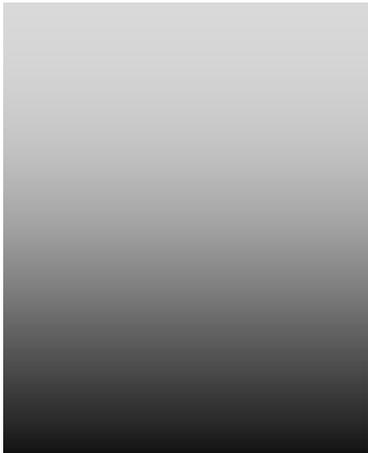
O modo de excitação do spin é controlado por dois parâmetros que auxiliam na distinção dos tecidos nos diferentes sinais. Os dois parâmetros são o tempo de repetição e o tempo de eco (TR e TE respetivamente). TR corresponde ao tempo entre dois pulsos de radiofrequência. TE diz respeito ao intervalo de tempo entre a aplicação dos primeiros dois pulsos, de 90° (responsável pelo FID) e o pulso de 180° , até ao momento em que a bobine regista um sinal de RF. Através da manipulação de TR e de TE surgem várias modalidades de ressonância magnética, as quais se designam de sequências. A sequência T1 surge com um TE e um TR curtos, ao par que a sequência T2 surge com TE e TR longos. A sequência DP (densidade de prótons) é alcançada com TR longo e um TE curto. Uma sequencial especial da RM é o caso da FLAIR⁴ que se serve do conceito de tempo de inversão (TI). O tempo de inversão consiste em aplicar de forma inversa os pulsos de RF, ou seja, primeiramente é aplicado o pulso de 180° e só depois o de 90° . O vetor de magnetização inicia-se com valores negativos, passa pelo valor nulo de campo magnético, e regressa aos seus valores positivos de equilíbrio na presença de B_0 [23]. A aquisição da FLAIR resulta da aplicação de um TI longo.

⁴Termo inglês: *Fluid Attenuated Inversion Recovery*

2.4.2 Imagem

Nesta subsecção procura-se compreender o comportamento imagiológico da RM nos diferentes tipos de tecido presentes num indivíduo com portador de um glioma. Com base nos conceitos teóricos presentes na secção 2.4.1, o sinal de ressonância magnética captado pela instrumentação de receção de radiofrequência é convertido em níveis de cinza. Independentemente da sequência aplicada, tempos de recuperação elevados correspondem a regiões com níveis de cinza mais elevados (zonas mais claras). De modo contrário, tempos de relaxação mais curtos são caracterizados com níveis de cinza mais baixos (zonas escuras). O que distingue um protocolo dos restantes reside no facto de os tecidos apresentarem diferentes tempos de relaxamento mediante o TR e TE aplicados. Na tabela 2.1 seguinte encontra-se o comportamento dos três principais tipos de tecido cerebral, mediante a aplicação do protocolo T1 e T2.

Tabela 2.1: Tabela ilustrativa do comportamento imagiológico dos diferentes tecidos presentes no cérebro, quando sujeitos aos dois tipos base de protocolos de Ressonância Magnética.

Intensidade	T1	T2
	Tecido adiposo	Líquido Cefalorraquidiano
	Matéria branca e cinzenta	Matéria branca e cinzenta
	Líquido Cefalorraquidiano	Tecido adiposo

As propriedades presentes na tabela anterior são as bases fundamentais para a segmentação manual. O modo como a segmentação manual é realizada não deve ser um procedimento de todo menosprezado, visto que, permite conhecer como um radiologista procede para detetar a região tumoral. Conhecimento este que assegura alguma sensibilidade aquando do desenvolvimento de metodologias automáticas de segmentação tumoral.

2.4.2.1 Segmentação Manual

Na segmentação manual a dificuldade reside em discriminar as fronteiras de dois ou mais tecidos tumorais presentes (variação mediante a classificação do tumor). Algumas peculiaridades em termos de comportamento imagiológico dos tecidos tornam a tarefa minuciosa. A título de exemplo, na sequência T1c é possível delinear a zona ativa como sendo a região com um sinal hiper-intenso (níveis de cinza muito claros). Porém, nos gliomas de elevado grau existe sempre uma porção não cística e não necrótica,

circundante desta região ativa, mas que se destaca do edema. A região em causa, já enunciada na secção 2.3, corresponde à zona não contrastante tumoral. Conclui-se então que, para a deteção correta de um tumor cerebral deste tipo, é importante a necessidade de recorrer a diferentes sequências para alcançar uma segmentação robusta [8].

Vários autores [8] definiram um conjunto de classes que devem ser identificados, quando presentes, aquando da realização da segmentação tumoral manual. Cada uma das classes tem associada um dos quatro tipos de tecido tumoral possível, sendo elas:

- Classe 1 – Região necrótica, cística e hemorrágica;
- Classe 2 – Edema;
- Classe 3 – Zona tumoral não contrastante;
- Classe 4 – Zona contrastante.

Em relação ao procedimento de segmentação manual pode ser estruturado em três fases distintas. Numa primeira fase procura-se localizar espacialmente o edema tumoral, onde as sequências vantajosas para a realização desta tarefa são a T2 e a FLAIR. O edema, tecido ativo e necrose têm um comportamento homogéneo hiper-intenso em T2 garantindo a sua distinção das regiões saudáveis. A FLAIR é útil quando a região tumoral se aproxima dos ventrículos pois, contrariamente ao que se regista para a T2, na FLAIR o seu comportamento é hipo-intenso. Adicionalmente, em gliomas de baixo grau, a FLAIR permite compensar a incapacidade de delimitação da T2 em regiões fronteiriças entre a zona ativa e o edema [8].

A segunda etapa passa pela definição da zona ativa, ou seja, a zona não contrastante e contrastante do tumor (classe 3 e 4 respetivamente). A segmentação da região não contrastante exige alguma experiência, visto que se serve em simultâneo de duas sequências que nem sempre apresentam as mesmas características. Através do protocolo T1c e T2 constata-se que a zona não contrastante na T2 pertence a uma região vasta do núcleo tumoral ativo, destacando-se do edema por apresentarem um sinal de baixa intensidade e heterogéneo. Porém, essa região na T1c pode ser dividida nas duas regiões da zona ativa, pois a zona não contrastante tem um comportamento semelhante ao do edema, de baixa intensidade [8].

Por último em T1c é identificada a região necrótica que se encontra sempre próxima (geralmente contida) da região contrastante da zona ativa, de comportamento hiper-intenso, quando comparada com uma região com um sinal muito baixo (hipo-intenso homogéneo). Contudo, embora que raramente, o fluido necrótico não se encontra contido na região ativa e portanto a FLAIR e a T2 devem ser utilizadas em simultâneo para averiguar a segmentação efetuada na T1c. Por último é definido o edema, que corresponde à região hiper-intensa da massa branca do cérebro em T2 e FLAIR, que frequentemente apresenta digitações (em formato de luva) para o sub-córtex giral, não identificada nas etapas anteriores [8].

No que diz respeito à sequência T1, comparativamente com as restantes sequências, apesar de não existir um consenso entre autores, a referida sequência na grande maioria das vezes não acarreta informação de relevo para a realização de uma segmentação cerebral tumoral [8, 24].

Na figura 2.6 encontram-se as quatro sequências de RM de um paciente com HGG. Levando a cabo as fases descritas obtêm-se a segmentação presente na figura 2.6 (e).

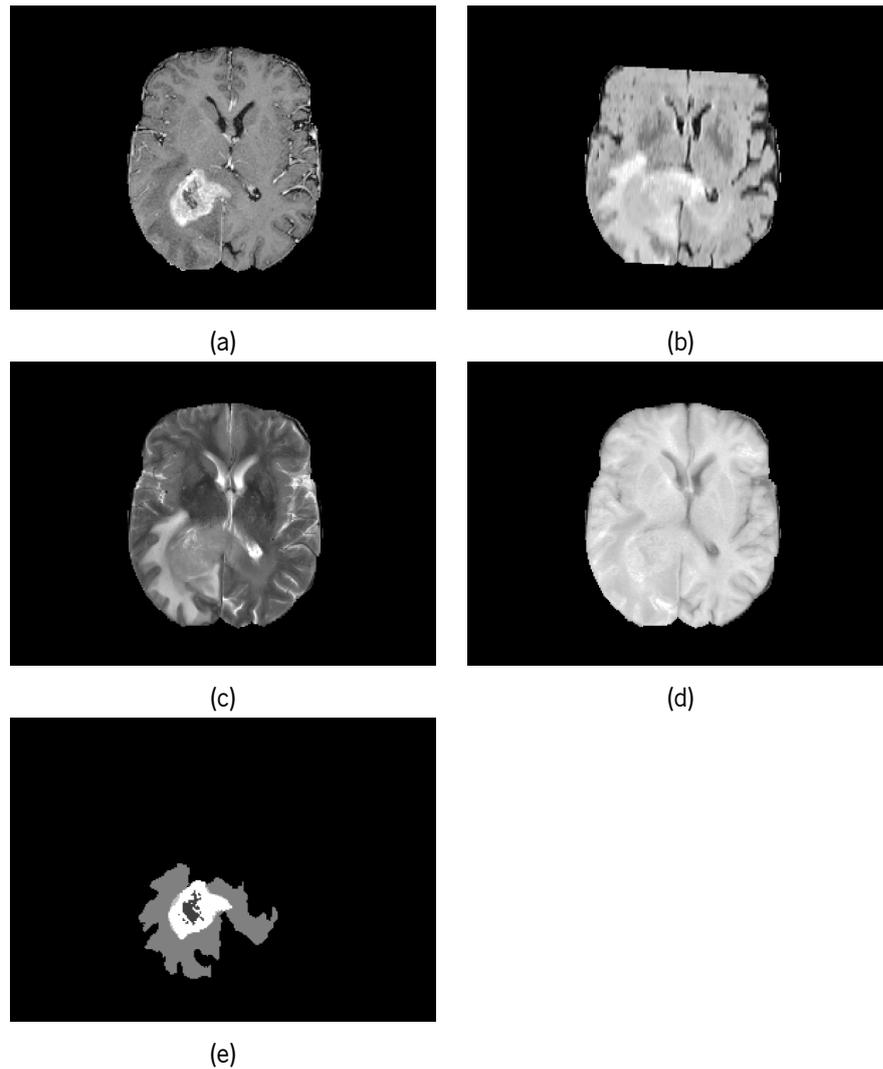


Figura 2.6: Fatia de ressonância magnética, nas diferentes sequências disponibilizadas pela base de dados, para a realização de uma segmentação cerebral tumoral manual e o respetivo resultado (marcação manual).

2.5 Tratamento

A metodologia de tratamento do glioma é ajustada principalmente consoante a capacidade de difusão do tumor, ou seja, se o tumor é difuso ou não difuso. No caso particular do glioma a ressecção tumoral, complementada com a radioterapia focal, ainda permanece como o tratamento mais comumente aplicado, existindo métodos alternativos como a quimioterapia e outros agentes biológicos, cujo papel destes últimos ainda não é bem claro. Aquando da deteção de um neoplasmo glial procura-se sempre a realização de uma ressecção total, porém as ferramentas imagiológicas e técnicas de neurocirurgia até ao momento disponíveis não têm a capacidade de alcançar, na totalidade, tal objetivo daí que terapias adjuvantes, pós-operativas, são aplicadas, como é o caso da radioterapia [5, 7].

A cirurgia, para gliomas difusos, cada vez menos é aplicada, pois a infiltração do tumor nos tecidos da matéria branca cerebral, induzindo também a sua remoção, não garantem uma melhoria do estado de saúde e por vez podem acarretar complicações futuras. Por outro lado, no caso dos gliomas não difusos

(tumores focais), estes são facilmente acessíveis com o auxílio de técnicas imagiológicas e cirúrgicas. Apenas para este tipo de tumores a remoção cerebral, independentemente de ser uma abordagem total ou parcial, geralmente não requer fases de posterior tratamento (radioterapia) [7].

O método radioterapêutico é também distinto, mediante a capacidade infiltrativa do glioma. No caso de indivíduos com gliomas infiltrativos o tratamento passa pela aplicação de um feixe externo, local, de dose controlável e fixa. Na ausência deste tratamento estima-se que a taxa de sobrevivência seja aproximadamente 20 semanas. Na situação de um neoplasmo glial não infiltrativo a radioterapia é usada seguida da cirurgia de remoção tumoral [5].

Relembrando a classificação da OMS, existem pequenas variações no tratamento consoante o nível de glioma. No caso dos gliomas de baixo grau, é aplicado todo o espectro de possíveis terapêuticas em voga, ou seja, radioterapia, cirurgia de remoção e quimioterapia, que por vezes é ignorada quando os sintomas são recentes. Na fase III do glioma, que já diz respeito a um tumor de elevado grau, a cirurgia é o primeiro procedimento em mente, adjuvada da radioterapia. Por último, na presença de um glioma de nível IV apenas os tratamentos de quimioterapia e radioterapia são utilizados, conjuntamente. A combinação destas duas abordagens, especialmente em casos recorrentes, demonstrou que a quimioterapia tem impacto positivo, embora que não seja estatisticamente significativo [5].

Contudo, estudos clínicos comprovam que a radioterapia atualmente permanece ainda como terapia adjuvante mais eficaz comparativamente com a quimioterapia. Não obstante, esta última terapia ainda se encontra em contínua expansão e estudo. Porém, o impacto da quimioterapia no tratamento de gliomas de elevado grau, ainda permanece uma incógnita. A radioterapia é um procedimento mais seletivo e local, cujo planeamento é mais eficaz [5, 7].

2.5.1 Histologia e Ressonância Magnética

Os gliomas, na ressonância magnética cerebral, apresentam características que têm uma elevada correlação com as suas propriedades histológicas, reforçando o lugar de destaque que esta técnica imagiológica tem na área dos tumores cerebrais [6].

No caso de um tumor de baixo grau difuso, este é detetado quando, no protocolo FLAIR, existem porções cerebrais que apresentam edema. O edema em termos histológicos corresponde a uma região de crescimento celular sem as estruturas sanguíneas circundantes (ausência de angiogénese) [6].

Gliomas de grau III, conforme a OMS, na sequência T1c contêm regiões nodulares indicativas da formação de uma nova região de vasos sanguíneos. Regiões nodulares são responsáveis pelo acelerar do processo de expansão tumoral, estando por isso vulgarmente associadas a uma porção de edema consideravelmente grande [6].

Para os gliomas de IV grau é possível visualizar, na sequência T1c, uma massa central rodeada por uma região nodular. Em termos histológicos, o núcleo corresponde a regiões de necrose, simbolizando um tumor com grande capacidade infiltrativa. A sequência T2 é também capaz de corroborar a existência de tumores deste tipo visto que nesta sequência o edema e a porção infiltrante (região nodular) têm a mesma resposta em termos de radiofrequência, produzindo em termos imagiológicos uma região hiper-intensa [6].

O planeamento do posterior tratamento carece da identificação de uma região anormal com um comportamento hiper-intenso na sequência T1c, o que leva a um diagnóstico de um tumor de elevado grau (HGG). Tal dedução não é geral pois posteriores biopsias podem indicar que se trate de um tumor difuso fibrilar de grau inferior, mas a este tipo de tumor, devido à sua crescente expansão, é aconselhável o mesmo tipo de tratamento de um tumor de elevado grau. Já a ausência de uma área hiper-intensa não tem uma correlação direta com o grau do tumor presente, nem o seu tratamento. Esta peculiaridade deve-se a possível existência de um tecido tumoral nodular com taxas de crescimento consideráveis e para tal classificado como HGG por parte da OMS [6].

2.6 Sumário

O cancro é uma doença cujo impacto na sociedade tem aumentado num passado recente. Em conformidade com os dados apresentados pela OMS, relativos à demografia e etiologia desta doença, compreende-se o relevo que uma deteção e diagnóstico precoce do cancro têm, bem como o seu controlo, em a aumentar a probabilidade de sucesso do planeamento e intervenção médicos.

No caso concreto dos cancros do Sistema Nervoso constatou-se que a sua taxa de incidência e mortalidade mundiais, consideráveis, é explicada, em grande parte, pela expressão crescente de tumores cerebrais denominados gliomas. Estes encontram-se divididos em duas classes, gliomas de baixo grau e de elevado grau. A deteção dos tumores cerebrais é efetuada com recurso às técnicas de imagiologia: TAC, RM e PET. Contudo, dada a sua capacidade multiespectral a RM destaca-se por permitir discriminar as diversas estruturas internas tumorais. Dada esta última propriedade, a RM desempenha um papel auxiliar importante no planeamento e acompanhamento médico, visto que permite realizar uma caracterização prévia do tumor, com um grau de detalhe considerável.

Posteriormente, à deteção da região tumoral, o diagnóstico da estrutura cerebral é feito através da biopsia, com a finalidade de caracterizar a estrutura e determinar o tratamento mais adequado. A primeira hipótese para o tratamento dos tumores cerebrais passa pela remoção da porção de cérebro afetada pelo tumor. Porém, o tratamento em estágios avançados, dado que a remoção da porção tumoral é complexa, compreende terapias adjuvantes, como a quimioterapia e a radioterapia. Não obstante, por vezes mesmo após a ressecção tumoral, é também aplicada terapia adjuvante.

Estado da Arte

Num passado recente, o desenvolvimento de sistemas de segmentação de tecidos cerebrais tumorais tem despertado o interesse de vários grupos de investigação. Apesar de não existirem sistemas de segmentação automática de tumores cerebrais que possam ser usados atualmente na prática clínica, vários métodos de segmentação têm demonstrado resultados robustos, mostrando assim, o potencial para reduzir os recursos humanos necessários e a variabilidade associada à segmentação manual. Complementarmente, a presença de métodos de segmentação automáticos, cada vez mais rápidos e precisos, são uma necessidade para o diagnóstico e planeamento cirúrgico e/ou terapêutico.

Neste capítulo, primeiramente, procura-se compreender os dois modelos base nos quais assentam as metodologias de segmentação cerebral de tumores. Seguidamente, são apresentadas as contribuições de alguns autores na área. Métodos baseados em *machine learning* (metodologias de aprendizagem computadorizada), em lógica difusa (algoritmos *fuzzy*) e em formulações estatísticas serão descritos.

3.1 Componentes Típicos na Segmentação de Tumores Cerebrais

Os algoritmos de segmentação de tumores, geralmente, encontram-se estruturados em três fases: pré-processamento, classificação e pós-processamento. Em conformidade com o modo de classificação implementado, os métodos de segmentação de tumores cerebrais podem ser divididos em dois grandes grupos: generativos e discriminativos [25].

Num modelo generativo a etapa de classificação é realizada tendo em consideração a probabilidade conjunta, $p(x, y)$, dos diferentes pontos da imagem, x , e as respetivas classes, y . A seleção da classe y mais provável, em cada ponto x é obtida, geralmente, pela regra de Bayes que calcula a probabilidade posterior e condicional $p(y|x)$ [26]. Neste tipo de modelo, a definição da probabilidade conjunta incorpora informação prévia relativa à aparência e distribuição dos diferentes tipos de tecidos a identificar [25], como por exemplo o crescimento ou atrofia da região. Vulgarmente, o crescimento ou atrofia de uma região são caracterizados através de metodologias de registo de atlas [27–29]. O registo de atlas é uma técnica baseada na informação espacial, cuja principal finalidade é modelar a anatomia do paciente alvo. Por

defeito, esta técnica de registo não contabiliza qualquer informação patológica. Assim sendo, a localização do tumor é extraída com base no pressuposto de que regiões de tecido tumoral causam um desvio considerável na distribuição estatística dos tecidos. Em termos imagiológicos este evento corresponde a regiões de deformação. Posteriormente, a demarcação das fronteiras do tumor é efetuada através da distribuição da probabilidade conjunta das intensidades, observadas na RM, e o tecido a classificar (alvo) [30].

As principais desvantagens dos modelos generativos são a sua vulnerabilidade ao ruído presente e ainda à variabilidade da informação entre sujeitos. Apesar destes problemas serem transversais ao tipo de modelo aplicado, é nos modelos generativos em que estes problemas têm mais impacto [25, 30, 31]. O modo como o modelo é inicializado afeta o restante processo de identificação dos diferentes tipos de tecido em análise [32]. Complementarmente, um único modelo estatístico apresenta dificuldades na captura de informação relevante, quando a definição dos parâmetros destes englobam um conjunto de informação vasto [32, 33]. Por forma a superar estas dificuldades alguns métodos servem-se da interação humana (modelos semiautomáticos) para a colocação de marcadores capazes de uma inicialização mais robusta do método [25].

Por outro lado, no modelo discriminativo, a probabilidade condicional, $p(y|x)$, é diretamente modelada [26]. Para tal, um modelo paramétrico de aprendizagem é definido e ajustado ao contexto do problema através de uma fase de treino. Durante a fase de treino o classificador procura capturar padrões capazes de discriminar os diferentes tipos de tecido, condicionados pela distribuição de um conjunto de características¹ extraídas através das intensidades de um ou mais protocolos de RM. Posteriormente, na fase de teste, é efetuada a classificação de cada ponto da imagem (*voxel*), sendo-lhe atribuído uma classe por parte do classificador. O desempenho do classificador é extremamente influenciado pelo procedimento de amostragem e a respetiva fase de treino, pois são fatores chave em assegurar um modelo com capacidade de generalização [25, 31].

Tal como foi descrito, um modelo discriminativo requer um conjunto de características extraídas para cada *voxel* do volume da RM. A extração de características tem como ponto de partida a informação fornecida pelas intensidades da RM. Daí que uma vantagem dos modelos discriminativos seja a capacidade em gerar características de elevada dimensão, isto é, que não dependem diretamente do valor das intensidades da RM, permitindo a computação de características com uma maior capacidade discriminativa das diferentes classes a segmentar [25, 26].

Os modelos discriminativos podem ainda ser desdobrados em duas subclasses: supervisionados e não supervisionados. Em métodos não supervisionados, tal como a própria designação o indica, os dados de treino não se encontram classificados, sendo o modelo responsável por criar uma estrutura capaz de agrupar os dados num determinado número de classes. Metodologias como *k-means* [34] (técnicas de *clustering*) e metodologias *fuzzy* [35, 36] são incluídas nos métodos discriminativos não supervisionados. No caso dos métodos supervisionados, na fase de treino, existe a informação sobre as classes a atribuir a cada *voxel*, o que no contexto do problema corresponde à presença de uma segmentação manual de todo o volume cerebral. A vantagem na aplicação de modelos supervisionados é em parte devido ao facto de se desconhecer o modelo de formação dos dados, da RM. Num método

¹Do inglês: *features*

supervisionado existe um modelo previamente definido, cujos parâmetros são ajustados ao conjunto de dados de treino. Uma vez determinados os parâmetros do modelo supervisionado estes permanecem fixos. Contudo, a aplicação de métodos supervisionados é condicionada pela existência de uma base de dados com a etiquetagem dos dados. Por outro lado, quando o modelo de formação dos dados é conhecido as metodologias não supervisionadas são as ideais, se as inicializações forem as corretas. Nas metodologias não supervisionadas, o modelo é constantemente adaptado a cada novo conjunto de dados, daí que o tempo de computação desta classe de metodologias é superior [25].

Na área da segmentação tumoral cerebral, os métodos supervisionados representam a metodologia mais recente [25]. Nestes destacam-se as *Random Forests* [37] e as *Support Vector Machines* [38].

Não obstante, um modelo generativo e um modelo discriminativo apresentam algumas semelhanças. Num modelo generativo, bem como num modelo discriminativo, as intensidades são o ponto de partida das metodologias de segmentação automática. Adicionalmente, ambos os modelos contêm funções probabilísticas que assentam no princípio independente e identicamente distribuído (i.i.d.), o que pode resultar em pontos isolados, ou agrupamentos de pontos, mal classificados [25]. Independentemente do modelo é fulcral assegurar uma coerência na informação das intensidades em diferentes regiões dos diversos pacientes em análise. A fase responsável por garantir tal objetivo é a etapa de pré-processamento, que engloba passos como a correção da falta de homogeneidade² das intensidades da RM, normalização das intensidades [39, 40] e técnicas de registo linear ou não-linear [41, 42].

3.2 Métodos de Segmentação de Tumores Cerebrais

Os primeiros passos dados na área da segmentação de tumores cerebrais ocorreram na década de 90. Em 1990, Cline *et al.* [45] propôs um modelo generativo capaz de realizar a segmentação tridimensional de um volume de RM. Através da interação humana, a primeira fase passa pela marcação de pontos de modo a caracterizar os diferentes tipos de tecidos a classificar. Seguidamente, partindo do pressuposto de que cada tecido segue uma distribuição normal bivariável, é computada a distribuição das probabilidades associada a cada tecido. Com base nas probabilidades é construído um mapa. Posteriormente, em cada ponto é atribuída uma classe de tecido através da determinação do valor máximo do conjunto de probabilidades associadas às diferentes classes de tecido definidas. Por último, como etapa de pós-processamento, intervém uma filtragem passa-baixo de modo a suavizar a imagem (fronteiras menos rígidas) e ainda morfologia matemática (conectividade) para eliminar pontos mal classificados. Contudo, o trabalho mais reconhecido na área de segmentação de gliomas, da década de 90, é proposto por Clark *et al.* [46], assente na filosofia *Knowledge Based* (KB).

Contrariamente ao método de Cline *et al.* [45], esta metodologia de segmentação, totalmente automática, é apenas capaz de realizar a análise de imagens bidimensionais, ou seja, a fatias³ do cérebro. A obtenção do volume final de tecido cerebral segmentado é alcançada através da junção das fatias. Na fase de pré-processamento é efetuada uma classificação prévia, com recurso a um modelo de *fuzzy c-means clustering* para a identificação das fatias detentoras de regiões tumorais, descartando-se as restantes.

²Do inglês: *Bias-Field correction*

³Do inglês: *slices*

Nesta mesma etapa é efetuada a extração de regiões anatómicas não relevantes, como por exemplo o crânio, ar e gordura, surgindo uma máscara, que define a região de operação. Posteriormente, o sistema KB intervém de modo a fornecer um conhecimento prévio da distribuição de cada classe, para posteriores etapas de classificação. A vantagem de um método KB é a sua capacidade de refinamento, das propriedades de cada classe, ao longo das sucessivas iterações. Juntamente com um *threshold* adaptativo aos histogramas de duas sequências de RM: T1 e Densidade de Protões (DP); o tecido tumoral é discriminado das restantes classes de tecidos presentes, rejeitando os pontos que pertencem à classe de tecido cerebral saudável. Para assegurar uma maior robustez na segmentação é aplicado ainda um procedimento no qual é construído um histograma tridimensional de densidades, que permite descartar ainda mais pontos de cérebro saudáveis mal classificados. Finalmente, através da análise da média e desvio-padrão, realiza-se uma filtragem aos pontos pertencentes à classe tumoral. Os componentes que não se incluem na distribuição do tecido tumoral são rejeitados. Como se pode averiguar, o método proposto por Clark *et al.* [46] procura sucessivamente descartar pixels que não pertencem à região tumoral.

Na subsecção 3.2.1 e 3.2.2 são analisados, respetivamente, os modelos generativos e discriminativos, mais relevantes, na área da segmentação tumoral cerebral de gliomas.

3.2.1 Modelos Generativos

A abordagem inovadora proposta por Clark permitiu que os modelos generativos adquirissem popularidade.

Dentro dos modelos generativos, Kaus *et al.* [47] impulsiona o desenvolvimento de métodos de segmentação tumoral cerebral baseados em atlas [27–29]. Na metodologia proposta, é implementado um processo iterativo de classificação estatística, sendo que no término de cada uma das iterações a técnica de registo intervém, por forma a atribuir uma de 4 classes. As 4 classes definidas são: fundo preto (*background*), pele, cérebro e ventrículos. Como o atlas não tem qualquer informação sobre informação patológica, os pontos tumorais são incluídos numa superclasse designada Cavidade Intracraniana (CIC) que contém as classes tecido cerebral, ventrículos e tecido tumoral. Terminada a fase de registo a classe CIC é alvo de duas iterações, através da informação prévia contida no registo, de modo a que contenha apenas *voxels* de tumor. Apenas nesta fase é possível a distinção da região de tecido saudável da região tumoral. O presente algoritmo parte do pressuposto que o tumor nunca se encontra numa região mais externa do cérebro, sendo por isso uma desvantagem visto que nem sempre se constata tal fenómeno. Adicionalmente, o método parte do pressuposto que a aparência do tumor na RM é extremamente homogénea.

Zou *et al.* [48] propõe um modelo de segmentação iterativo puramente probabilístico que procura estimar uma mistura de duas distribuições, com diferentes parâmetros de forma, para as duas classes a classificar: tumor e tecido não tumoral. A inicialização do método em causa serve-se do método de registo semiautomático, proposto por Kaus *et al.* [47]. A segmentação é obtida através da estimação dos parâmetros de cada distribuição e dos *thresholds* a aplicar às probabilidades. Este último procedimento é assegurado por um conjunto de métricas, que permite a definição dos *thresholds* ótimos, através de uma filosofia de maximização da certeza da segmentação, sendo por isso extremamente dependentes do

objetivo de segmentação.

Prastawa *et al.* [49], tal como Clark *et al.* [46], propôs uma metodologia de segmentação bidimensional. O método de segmentação multispectral, serve-se da informação contida nas sequências FLAIR, T1, T1c e T2. Primeiramente, pela técnica de atlas a região tumoral é delimitada. De seguida, o algoritmo generativo de *Expectation Maximization* (EM) é aplicado para distinguir duas subclasses da região tumoral: a região central do tumor e o edema. Na primeira etapa do algoritmo EM, a *Expectation*, é descrita uma função de probabilidade para a modelação dos dados das diferentes classes. Na segunda etapa, a classificação, os dados são atribuídos a cada uma das distribuições, previamente definidas, de modo a maximizar a verosimilhança entre os dados. Como pós-processamento Prastawa *et al.* [49] aplicou restrições volumétricas e anatómicas, para lidar com falsas deteções. Posteriormente, Prastawa *et al.* [50] propôs uma filosofia diferente, baseada unicamente na sequência T2 para a deteção de *outliers*, que correspondem às porções de tumor. O primeiro estágio de deteção tumoral manteve-se idêntico, através da técnica de registo num atlas de cérebro saudável. Um tratamento estatístico, responsável por agrupar os pontos tumorais (*Minimum Covariance Determinant* - MCD), é seguido por um modelo deformável implementado com o intuito de separar o núcleo do tumor do edema. A técnica proposta destaca-se pela implementação de uma estimação não paramétrica das distribuições, ao invés da comum mistura Gaussiana. Contudo, os autores detetaram que apesar de a sua base de dados ser reduzida, que as técnicas de amostragem baseadas em atlas, não são robustas a deformações tumorais exageradas.

Bauer e Reyes [51] propuseram uma nova filosofia para a aplicação da técnica de atlas na realização da segmentação tumoral de gliomas. Na sua metodologia, a técnica de registo não é efetuada de forma direta, quando aplicada à imagem de RM cerebral com presença de regiões tumorais. A deformação causada pela região tumoral é modelada matematicamente através das diferenças finitas de uma dada vizinhança local de cada ponto da imagem, com recurso à técnica de *Markov Random Field* (MRF). Posteriormente, Bauer *et al.* [52] propõe alterações ao MRF implementado. Neste último método, no MRF, os cliques de uma vizinhança procuram minimizar a função de energia que modela as propriedades biomecânicas celulares cerebrais, que no método proposto corresponde ao módulo de *Young*. Após um primeiro estágio de registo é definido com alguma certeza a localização do centro de massa do tumor. Partindo desta observação, o método efetua um processo iterativo de crescimento volumétrico que assenta na minimização da referida função de energia.

Menze *et al.* [53] propôs uma filosofia semelhante à de Bauer e Reyes [51], procurando modelar matematicamente o comportamento do tecido tumoral. Este modelo probabilístico, totalmente automático, é suscetível de ser aplicado a imagens de RM multidimensionais, encontrando-se dividido em duas fases. Numa primeira fase é aplicada a técnica de atlas para obtenção de um *prior* probabilístico que modela cada uma das classes de tecido cerebral a segmentar. Para cada ponto do volume da RM o atlas estima uma distribuição de cada classe. Um segundo atlas latente, de variação espacial, é utilizado para a estimação do tumor. Assim, na segunda fase, a segmentação é realizada através da aplicação de uma técnica de Máxima Verosimilhança de dois parâmetros: *prior* do atlas com k classes e um parâmetro de modelação da distribuição das intensidades. O último parâmetro apontado resulta da modelação das intensidades com base no pressuposto que estas seguem uma distribuição Gaussiana. Tratando-se de um processo iterativo, o método procura assim dar um relevo constante às intensidades em cada uma

das etapas. Após a estimação de ambos os parâmetros do modelo, é avaliada a probabilidade de cada *voxel* apresentar o comportamento característico de um tecido tumoral. A proposta de Menze *et al.* [53] foi dos primeiros modelos generativos a realizar uma inicialização dos parâmetros supracitados através da análise de uma segmentação prévia, em três classes de tecido cerebral saudável (matéria branca, cinzenta e fluido cefalorraquidiano). e uma classe de tecido tumoral.

Um outro método de segmentação de gliomas, através da técnica de registo, é apresentado por Kwon *et al.* [54]. A proposta apresentada pelos autores, à qual designaram de *GLioma Image SegmenTation and Registration* (GLISTR) compreende uma técnica de registo realizada através de atlas de pacientes saudáveis. Neste é embebido um modelo de crescimento tumoral inicializado pela colocação de sementes. O modelo de registo é útil para posterior computação de um mapa das probabilidades que caracteriza uma entidade única, o tumor. Por sua vez este mapa é responsável por gerar um novo atlas que permite a separação de edema da restante porção tumoral. Assim sendo, resta apenas a distinção da região ativa da região necrótica, pelo que um algoritmo EM é aplicado combinando a probabilidade *a posteriori* de cada classe com os parâmetros do modelo de crescimento descrito, o segundo atlas e as intensidades da imagem de RM.

Dentro dos modelos generativos, por último, Hamamci *et al.* [55] apresentou um método designado de *Tumor Cut*. O algoritmo apenas se recorre da informação das intensidades presentes na T1c, para a aplicação de uma metodologia de segmentação baseada na colocação de sementes - *Celular Automata* (CA). O utilizador é responsável pela marcação de sementes relativas às duas grandes classes: tumor e *background*. O método é ainda capaz de discriminar, dentro da classe tumor, dois tipos de tecidos, a necrose e a porção tumoral ativa. As sementes permitem a construção de mapas de probabilidade que indicam a proximidade de cada *voxel* às classes definidas, designando-se aos presentes mapas de mapas de força. Os mapas de força são posteriormente fundidos com recurso ao algoritmo CA. Irregularidades características do problema são posteriormente corrigidas por um algoritmo de *Level Set* (LS), inicializado através de uma superfície fechada definida por probabilidades superiores ou iguais a 0,5. O método de LS é posteriormente propagado até convergência, assegurando uma continuidade espacial na segmentação. De seguida a metodologia agrupa as amostras identificadas como tecido tumoral em duas classes: necrose e tecido ativo. O agrupamento das amostras é realizado pela aplicação de um *threshold* sobre as intensidades dos pontos da RM, que posteriormente servem de sementes para um novo andar de CA com a finalidade de assegurar continuidade à segmentação final. Finalmente é atribuído a cada *voxel* a classe que apresenta maior probabilidade. O algoritmo é mais robusto quanto maior for a certeza na marcação dos *voxels* responsáveis pela inicialização e caracterização das classes. Deste modo, o algoritmo proposto depende diretamente do conhecimento por parte do utilizador, sendo um fator limitante.

A grande maioria dos métodos generativos expostos requerem uma intervenção humana ou, no melhor dos casos, a retenção de algum conhecimento humano prévio, para posterior caracterização do comportamento do tumor em termos das suas intensidades ou até uma inicialização da sua localização [30, 56–60]. É com base na aplicação deste conhecimento que os modelos generativos têm a capacidade de distinguir o tumor das restantes classes de tecido cerebral saudável [56]. A criação de um modelo probabilístico robusto, tendo por base o conhecimento *a priori* é difícil. Estes argumentos foram suficientes para uma mudança na filosofia de abordagem para a segmentação de tecidos tumorais cerebrais.

3.2.2 Modelos Discriminativos

Nos modelos discriminativos, o classificador mais utilizado nos primeiros métodos de segmentação tumoral, cerebral foi o *Support Vector Machine* (SVM) [25, 31]. Na fase de treino, característica deste tipo de classificadores, define-se um hiperplano no espaço n-dimensional de características que se encontre o mais distante possível de ambas as classes e ao mesmo tempo as permita separar. As características permitem a criação de um espaço de dimensões elevadas de modo a permitirem a separação das amostras, sendo efetuada com recurso a um *kernel*. Geralmente, o *kernel* aplicado é linear, podendo ser sigmoide, tangencial ou até exponencial, exigindo, estes últimos, maior capacidade computacional [38]. Neste tipo de modelos, na classificação, a probabilidade de um determinado *voxel* não tem em consideração a respetiva vizinhança. Deste modo, não é contabilizada qualquer noção de contexto [61].

Por forma a superar o problema supracitado, Lee *et al.* [62] aliou aos SVMs a metodologia de *Random Fields*, atribuindo-lhe a designação de *Support Vector Random Fields*. Os *Random Fields* permitem compensar a falta de contexto presente no SVM, tendo sido o foco em estudo na metodologia proposta. De modo a sustentar sua proposta, esta foi comparada com várias modalidades de classificação, tendo ao todo 6 classificadores distintos: Classificador de Máxima Verosimilhança, Classificador de Regressão Logística, SVM, MRF, *Conditional Random Fields* (CRF) e *Discriminative Random Fields* (DRF). Os diferentes métodos foram aplicados a diferentes tarefas de segmentação, por forma a caracterizar a robustez e falhas associadas. Em primeiro lugar, é testada a capacidade de deteção da região tumoral ativa. Seguidamente foi avaliada a robustez em definir a região do edema e num teste final, mais complexo, foi analisado o desempenho do método à distinção entre tecido saudável e edema. Após a obtenção das probabilidades posteriores, os autores concluíram que o MRF apresenta algumas limitações na sua inicialização. Por outro lado, no CRF, como não necessita de modelar a distribuição de cada uma das classes, a computação da probabilidade posterior é mais imediata e com uma menor variância. A definição de um fator de penalização entre classes, por parte do CRF, é uma propriedade apontada como vantajosa para o tratamento destes problemas. Em relação ao DRF, que não é mais do que uma abordagem multi-dimensional do CRF, este apresenta resultados ainda mais positivos. Contudo os custos computacionais são exigentes, o que conduziu ao desenvolvimento de um *Pseudo-Conditional Random Fields* de modo a não comprometer a segmentação. Ao aliar o SVM com o método de *Pseudo-Conditional Random Fields* os autores verificaram um aumento do desempenho na atribuição de classes quando este classificador é utilizado isoladamente.

Bauer *et al.* [63] propõe também uma metodologia de classificação com SVM, aliada a uma fase de regularização espacial. Neste método é apresentado um conjunto de características multiespectrais que permitem modelar a textura presente nos volumes da RM. A metodologia permite classificar o tecido saudável nas suas três classes mais importantes: matéria branca, matéria cinzenta e fluido cefalorraquidiano. No que respeita a identificação do tecido tumoral, este algoritmo é capaz de dividir o tumor em duas classes: região necrótica e região ativa. Para aumentar a robustez e rapidez do método, este foi estruturado de um modo hierárquico. Em fases distintas procura-se classificar diferentes tecidos, em que no final de cada fase é aplicada a regularização espacial, providenciada pelo CRF. O método é extremamente dependente do valor das intensidades e para tal a implementação contém um pré-processamento

severo: correção da falta de homogeneidade das intensidades, reescalamento das intensidades associado a uma manipulação do histograma e ainda uma filtragem para redução do ruído. Apesar do desempenho da metodologia apresentar resultados promissores, os autores apontaram algumas limitações do SVM, nomeadamente a dificuldade de modelar problemas multi-classe. Posteriormente, Bauer *et al.* [64] propôs uma mudança do seu classificador para uma Floresta de Decisão assegurando uma metodologia mais robusta.

Apesar de os SVMs apresentarem algumas limitações para o contexto do problema, vários métodos foram propostos com estes classificadores [65, 66]. Li e Fan [65] apresentaram uma metodologia multi-escala, semi-supervisionada, que concilia métodos de agregação de classes como SVM. O presente método pode ser dividido em duas fases. Numa primeira fase o SVM intervém para obtenção de uma distribuição probabilística de cada uma das classes nos pontos do volume da RM. O classificador binário (tumor ou tecido saudável) recebe como parâmetros um vetor de intensidades do ponto em análise e as intensidades da sua vizinhança. Após obtenção do mapa de probabilidades segue-se uma agregação multi-escala. Esta última fase é responsável por tratar regiões salientes e isoladas no mapa de probabilidades, que surgem devido à presença de ruído não eliminado ou uma discrepância das intensidades. Para contornar este fator limitativo, integrou-se uma metodologia de segmentação de agregação pesada⁴ capaz de produzir vários níveis de decomposição da imagem. Este algoritmo recursivo tem com o intuito obter níveis formados por um único conjunto de tecido. Posteriormente, sobre o mapa das probabilidades é efetuada a agregação dos diferentes níveis para permitir o realce das regiões que efetivamente correspondem ao tumor.

A segmentação de gliomas através de uma Floresta de Decisão surge por Zikic *et al.* [67]. A introdução da Floresta de Decisão, pelo presente grupo de trabalho, é argumentada pela sua capacidade intrínseca em realizar uma classificação multi-classe, contrariamente ao que se constata nos métodos assentes em SVMs. O método proposto aplica o conceito das florestas de decisão para projeção de uma metodologia de segmentação automática multicanal. O algoritmo proposto focou-se apenas na análise de gliomas de elevado grau, sendo capaz de identificar 4 tipos de tecidos distintos: tecido cerebral saudável, necrose, edema e região ativa; recorrendo à informação das intensidades das sequências: FLAIR, T1, T1c e T2. O método tira partido das vantagens dos modelos discriminativos e generativos, através da inclusão de um modelo de misturas gaussianas (porção generativa) como um dos parâmetros de entrada das florestas de decisão (fase discriminativa). As probabilidades extraídas do modelo multidimensional de misturas gaussianas resultam de um processo iterativo de estimação da distribuição das intensidades que caracteriza cada uma das classes presentes, pelo que na verdade esta metodologia apresenta duas fases de treino. Como características de entrada da floresta de decisão o modelo implementa, para além das misturas gaussianas, características de contexto sobre o qual o classificador assenta, com o intuito de evitar posteriores etapas de regularização espacial. As características de contexto projetadas foram: a diferença de intensidades de um *voxel* numa dada sequência e um *voxel* a uma determinada distância d (numa sequência distinta), diferença entre as médias das intensidades de um cuboide definido em torno do *voxel* em análise e a média de um cuboide distanciado (numa sequência distinta) e, por último, a diferença entre os valores máximo e mínimo das intensidades de uma linha 3D (na mesma

⁴Do inglês: *Segmentation Weighted Agregation*

sequência). Os parâmetros de seleção das características de contexto, como o tamanho da vizinhança ou as sequências a comparar, são aleatórios, o que resulta num grande conjunto de características que garante a redundância. O método proposto apresenta resultados bastante promissores passíveis de serem melhorados, daí o grande reconhecimento que metodologias assentes em Floresta de Decisão tiveram nos últimos tempos.

Geremia *et al.* [68] procura combater a falta de regularidade nas segmentações realizadas por modelos discriminativos. Para tal, apresenta um método totalmente automático assente em florestas de decisão espaciais (*Spatial Decision Forests*). As características implementadas têm por base um trabalho prévio realizado com florestas de decisão, em lesões de Esclerose Múltipla [69]. O método procura a segmentação da região central tumoral e edema, servindo-se dos 4 protocolos da RM: FLAIR, T1, T1c e T2. A noção espacial é fornecida pela computação de características de contexto e ainda características de simetria. O contexto enunciado corresponde à captura da diferença das intensidades entre *voxel* em análise e a médias de dois cuboides aleatórios, definidos numa sequência distinta. As características de simetria extraídas, partem do pressuposto de que os tumores são estruturas assimétricas e que provocam a deformação de tecidos saudáveis circundantes, resultam da diferença entre dois *voxels* simétricos entre si em relação ao plano sagital médio.

Islam *et al.* [70] desenvolveu uma metodologia cuja ideia principal passa pelo desenvolvimento de um conjunto de características robusto e discriminativo. Deste modo, desenvolveu um conjunto de características de textura assentes na análise fractal (análise de geometria), que são úteis em caracterizar objetos com uma geometria irregular, como é o caso dos gliomas. Através da aplicação da teoria de movimento multifractal Browniano [71] procura-se a extração de características que permitam modelar a variação espacial do tumor cerebral. Complementarmente, propõe alterações a um tipo de classificador específico, que tem como base os SVMs, designado de *Adaboost* [72]. A proposta de alteração do classificador permite garantir a capacidade de o *Adaboost* lidar com os padrões de difícil classificação provenientes das características de textura, intervindo para tal na etapa de cálculo dos pesos do classificador atribuídos a cada característica. Para uma definição robusta da região tumoral, o método proposto serve-se da informação das intensidades presentes em FLAIR, T1, T1c e T2. Os resultados obtidos pela metodologia proposta apenas permitiam a distinção de tecido saudável de tecido tumoral. Não obstante, os resultados alcançados permitem assegurar uma delimitação robusta das fronteiras tumorais, bem como a marcação do seu interior. Com o impacto da introdução das Florestas de Decisão no problema da segmentação de gliomas, mais tarde Reza e Iftekhharuddin [73] propuseram uma alteração do classificador para uma Floresta de Decisão, alcançando resultados ainda mais positivos. Complementarmente, este último método tem a capacidade de distinguir o edema da porção central do tumor.

Tustison *et al.* [74] implementou um método de segmentação em duas fases, cuja classificação é realizada pela Floresta de Decisão. Numa primeira fase é aplicado um modelo de mistura gaussiana para permitir a definição das diferentes classes de tumor. Só após a atribuição das classes a cada um dos pontos do volume de RM é possível a extração de características. As características extraídas correspondem às intensidades, após normalização, à geometria de componentes ligados, modelada pelas intensidades, um conjunto vasto de estatísticas de primeira ordem determinadas para uma vizinhança local e ainda características que quantificam a assimetria bilateral provocada pela presença do tumor.

O método proposto, reconhece, a grande dependência do classificador face à presença das características de assimetria resultantes da técnica de registo, nas 4 sequências utilizadas (FLAIR, T1, T1c, T2). Posteriormente, a Floresta de Decisão permite a obtenção de mapas de probabilidades para cada uma das classes de tecido tumoral: necrose, edema e tecido ativo. Os mapas de probabilidades resultantes servem de base para nova extração de um novo conjunto de características e posterior realização de uma classificação final.

Goetz *et al.* [75], desenvolveu uma metodologia de segmentação assente em Extra-Trees. Os autores defendem que é fulcral a etapa de pré-processamento, especialmente a normalização das intensidades. Para tal, com o intuito de assegurar uma coerência na forma dos histogramas de diferentes pacientes, Goetz *et al.* [75] aplica uma função de normalização definida por um conjunto de múltiplas transformações lineares. O conjunto de características utilizado por Goetz *et al.* [75] compreende características estatísticas de primeira e segunda ordem, intensidades e características associadas ao histograma das imagens.

De modo a acrescentar maior capacidade de generalização aos modelos discriminativos Kleesiek *et al.* [76] propôs uma ferramenta de aprendizagem para segmentação, denominada *ilastik*. A fase de classificação é realizada em paralelo por 10 classificadores supervisionados, pelo que o resultado final advém da combinação da saída de cada um dos classificadores. As características extraídas, para além das intensidades normalizadas, compreendem um conjunto numeroso de diferentes modalidades de filtragem Gaussiana. Como pós-processamento Kleesiek *et al.* [76] recorre a duas filtragens consecutivas, para remoção de possíveis marcações isoladas, seguidas de uma regularização espacial. Meier *et al.* [77] por outro lado aplica apenas um único classificador do mesmo tipo, *Decision Random Forest*, que é responsável por gerar um mapa de probabilidades de cada classe, onde posteriormente a técnica de CRF é aplicada, para regularização espacial. O conjunto de características neste último método é também distinto dado que são utilizadas características de aparência: intensidades, textura de primeira ordem, características assentes no cálculo de gradientes, e características de contexto.

Por último, na área da segmentação tumoral cerebral surgiram metodologias baseadas em modelos não supervisionados, nomeadamente algoritmos de lógica difusa. Os algoritmos de lógica difusa começam por atribuir a cada ponto da imagem um valor de correspondência (entre 0 e 1). O valor de correspondência pode ser entendido como uma distância a uma determinada classe, ou seja, uma quantificação do grau de semelhança entre o ponto e as várias classes definidas. Nesta classe de métodos destaca-se o trabalho de Fletcher-Heath *et al.* [35] que combinou uma classificação não supervisionada de FCM (*Fuzzy C-Means*) com um processamento KB para segmentação de tumores de baixo grau. Após agrupamento dos dados com recurso ao FCM, é aplicada uma filtragem a cada um dos pontos pertencentes a cada uma das classes definidas. Recorrendo apenas a três sequências da RM (T1, T2 e PD), o FCM é implementado para definir um agrupamento grosseiro dos dados em 10 classes de dimensões. Posteriormente, o algoritmo KB tem a função de refinar as 10 classes iniciais em 7, em conformidade com as especificações da fase de treino. Este método procura apenas segmentar a necrose e pequenas porções do edema tumoral, de modo a permitir uma quantificação final do volume tumoral.

Moonis *et al.* [78], tendo noção da variabilidade presente nas intensidades de uma mesma classe de tecido, propôs uma normalização das intensidades através de uma transformação não linear do histo-

grama. Este grupo de trabalho defende que a variabilidade das intensidades resulta do comportamento heterogéneo dos tecidos, ruído e alterações do comportamento das intensidades do *background*. Deste modo, a normalização aplicada procura associar a cada tipo de tecido presente uma gama de intensidades característica, mapeando as intensidades de cada um dos protocolos de RM (FLAIR, T1 e T1c) para uma escala de intensidades *standard* definida heurísticamente. Esta fase de tratamento dos dados é considerada fulcral para o método, visto que a classificação é realizada através de um algoritmo de Conexão Difusa (*Fuzzy Connectedness*) que agrupa os *voxels* em classes de elevada conectividade mediante a informação das intensidades. Complementarmente, o método proposto é sensível ao facto de as diferentes sequências utilizadas terem a capacidade de fornecer diferente informação.

3.3 Sumário

Vários algoritmos de segmentação de tumores cerebrais, mais concretamente gliomas, foram apresentados neste capítulo. Dada a variedade e o número de metodologias propostas, por diversos autores, é possível constatar que a segmentação de tumores cerebrais é uma área de relevo na comunidade científica. O interesse da comunidade científica na área da segmentação tumoral cerebral é, grande parte, explicado pela dificuldade do problema bem como a capacidade em desenvolver uma metodologia capaz de obter resultados precisos e com capacidade de generalização.

Diversos métodos propostos procuram definir um modelo que se adapte melhor ao conjunto de dados do problema, aos quais se designam de métodos generativos. Dentro deste tipo de modelo, existem propostas que recorrem à construção de atlas [27–29, 47, 49, 51, 53, 54], resultantes da informação de diversos pacientes, com a finalidade de posteriormente determinar a região de maior probabilidade de localização do tumor. Por último, destaque para metodologias assentes no algoritmo de EM, implementado com o intuito de obtenção de um mapa de probabilidades resultante, na maioria dos casos, de Modelos de Misturas Gaussianas [49, 50, 54].

Por outro lado, tem-se os modelos discriminativos que assentam em classificadores supervisionados. Algoritmos de aprendizagem como o SVM [62, 63, 65] e a Floresta de Decisão [67, 68, 73, 74, 76] viram o seu desempenho testado na segmentação de tumores, providenciando resultados competitivos. A grande diferença entre os métodos deste tipo reside numa etapa prévia à classificação, que compreende o desenvolvimento de características capazes de discriminar positivamente cada um dos tipos de tecido a classificar e por conseguinte uma melhor adequação do método ao contexto do problema. Uma outra abordagem, designada de KB, passa pela intervenção humana, de modo a que o seu conhecimento permita dar uma noção inicial das propriedades e estrutura tumorais. Após a fase de inicialização intervém um procedimento iterativo de segmentação [35, 78].

Neste capítulo é ainda possível verificar uma mudança do objetivo de segmentação tumoral cerebral. Nas primeiras abordagens o desafio consistia na delimitação do tumor como uma estrutura única, ao par que, as abordagens mais recentes procuram identificar as diferentes classes de tecido tumoral.

Fundamentos Teóricos

No presente capítulo explicam-se os conceitos teóricos necessários para o desenvolvimento de uma metodologia de segmentação automática. Na primeira secção, encontra-se uma descrição das linhas gerais que caracterizam os algoritmos propostos na área da segmentação cerebral tumoral. De seguida são abordados com mais detalhe os métodos mais relevantes em cada uma das etapas, sobre os quais assenta a abordagem desenvolvida no âmbito da presente dissertação.

4.1 Algoritmo Típico de Segmentação Tumoral

As metodologias de segmentação cerebral tumoral seguem uma ordem típica constituída por três etapas: pré-processamento, classificação e pós-processamento [43, 79].

A etapa de pré-processamento, tanto nos modelos generativos como nos modelos discriminativos, tem bastante relevo. A sua finalidade principal é assegurar uma homogeneidade e coerência das propriedades das imagens de RM a nível intra e inter-paciente [43]. O pré-processamento procura ainda remover artefactos e elementos estruturais que afetam o desempenho do classificador [79]. Como etapas de pré-processamento comumente utilizadas, destacam-se a correção da falta de homogeneidade do campo magnético, registo espacial, remoção do crânio, normalização das intensidades e filtragem para remoção de ruído.

A presença da falta de homogeneidade nas intensidades não é um fenómeno exclusivo da ressonância magnética encontrando-se também presente na tomografia computadorizada, microscopia e ultrassons [79]. A falta de homogeneidade na RM surge devido a um processo de aquisição incorreto, que se traduz numa variação suave das intensidades ao longo da imagem [79]. Deste modo, um determinado tecido é incapaz de apresentar uma gama de valores de intensidades regular em toda a sua extensão, visto que estas variam de acordo com a localização espacial.

Os algoritmos assentes na informação providenciada pela RM cada vez mais se servem da capacidade de sua análise multiespectral para uma classificação mais robusta [64, 67, 68, 70, 74]. Assim sendo, as técnicas de registo são fundamentais para que haja uma coerência espacial e resolução idêntica nos diferentes protocolos utilizados. O processamento dos dados, de uma forma geral, requer duas fases em que, primeiramente, se estima o vetor de transformação, determinado sobre uma imagem de

referência (pré-selecionada), que posteriormente é aplicado às restantes imagens de modo a garantir um alinhamento espacial da informação [42, 80].

Na remoção do crânio, tal como a própria designação o indica, procura-se eliminar as regiões que não trazem qualquer relevo para a análise do cérebro. O crânio, no contexto do problema, não é uma região anatómica alvo de análise. Além disso, a presença do crânio é prejudicial para algumas metodologias. Na RM o crânio corresponde a valores de intensidade elevados, pelo que esta estrutura pode ser confundida com classes de tecido cerebral tumoral, nomeadamente tecido tumoral ativo, afetando assim o desempenho do classificador [81].

Na RM o valor das intensidades não tem um significado associado, devido à grande variabilidade característica desta técnica de imagem médica. Como consequência direta deste fenómeno, as intensidades dos diferentes tecidos cerebrais e tumorais variam mediante o protocolo e a máquina de aquisição selecionados. Através da aplicação de várias etapas de pré-processamento procura-se combater este entrave e alcançar uma consistência na informação das intensidades. Em suma, a presença do pré-processamento tem como objetivo principal a atribuição de um significado às intensidades, o que consequentemente melhora o desempenho da classificação [79, 81].

A fase de pós-processamento tem como finalidade a deteção da presença de pequenos agrupamentos de pontos isolados, que são mal classificados por parte do classificador. Comummente, o pós-processamento assenta em métodos de morfológica matemática.

4.2 Pré-processamento

Na presente secção é descrita a fundamentação teórica sobre a qual assentaram as etapas de pré-processamento realizadas nesta dissertação, a correção da falta de homogeneidade das intensidades e a normalização das intensidades.

4.2.1 Correção da Falta de Homogeneidade das Intensidades

A falta de homogeneidade das intensidades na ressonância magnética, apesar de ter um efeito por vezes dificilmente detetável por inspeção visual, influencia o desempenho dos métodos de análise de imagem médica, tais como metodologias de classificação e técnicas de registo, dada a sua sensibilidade às variações das intensidades. Por forma a reduzir este problema, foram propostos vários tipos de métodos de correção da falta de homogeneidade, que podem ser agrupados em duas grandes classes: prospetivos e retrospectivos. Um método prospetivo foca-se em desenvolver melhorias no processo de aquisição da imagem de RM e ainda em procedimentos de calibração dos equipamentos. Por outro lado, os métodos retrospectivos assentam no tratamento da informação presente na imagem de RM e por vezes em informação prévia acerca do modo de aquisição, por exemplo, informação relativa às bobinas de receção do equipamento [79].

A falta de homogeneidade das intensidades é entendida como uma função de variação suave que varia com o espaço tridimensional, o que se traduz em alterações diferentes em intensidades semelhantes que pertencem ao mesmo tipo de tecido cerebral. Os primeiros passos dados na correção da falta

de homogeneidade das intensidades surgiram em meados da década de 80 [82, 83]. De imediato se constatou que a melhor forma, para eliminar esta distorção, assenta na formulação de modelos adequados que podem ser agrupados em duas classes. Uma das classes engloba os modelos que consideram a falta de homogeneidade como um efeito multiplicativo, permitindo modelar de forma bastante similar a falta de homogeneidade associada à sensibilidade da bobine de recepção do equipamento de RM. Por outro lado, a restante classe compreende modelos que descrevem a falta de homogeneidade como um fator aditivo devido às correntes induzidas e da excitação não uniforme [79].

Independentemente da abordagem aplicada para a remoção da falta de homogeneidade, torna-se necessário primeiramente compreender o modelo matemático de formação de uma imagem de RM. Na formação de uma imagem de RM, Prima *et al.* [85], Styner *et al.* [86] e Brechbühler *et al.* [87] defendem que este procedimento apresenta duas fontes de ruído: ruído biológico e o ruído do *scanner*. Denote-se que o ruído biológico abrange uma distorção das intensidades intrínseca dos tecidos, independente da variação espacial, que não deve ser confundida com o artefacto introduzido pela falta de homogeneidade do campo magnético. Por outro lado, o ruído do *scanner* corresponde às imperfeições do dispositivo de aquisição. Geralmente apenas uma das fontes de ruído é tida em consideração, sendo que o modelo mais comum se preocupa em modelar o ruído do *scanner*. Assim, de acordo com o último modelo descrito, a imagem resultante da aquisição da ressonância magnética é dada por [88, 89]:

$$v(x) = u(x)b(x) + n(x) \quad (4.1)$$

No modelo descrito na equação 4.1, v designa a imagem corrompida resultante da aquisição, u a imagem não corrompida (original), b a falta de homogeneidade do campo e n contabiliza todos os tipos ruído do *scanner*. Como se pode averiguar a falta de homogeneidade é um artefacto tido como multiplicativo no presente modelo. Em termos visuais na figura 4.1 encontra-se um exemplo da presença da falta de homogeneidade das intensidades.

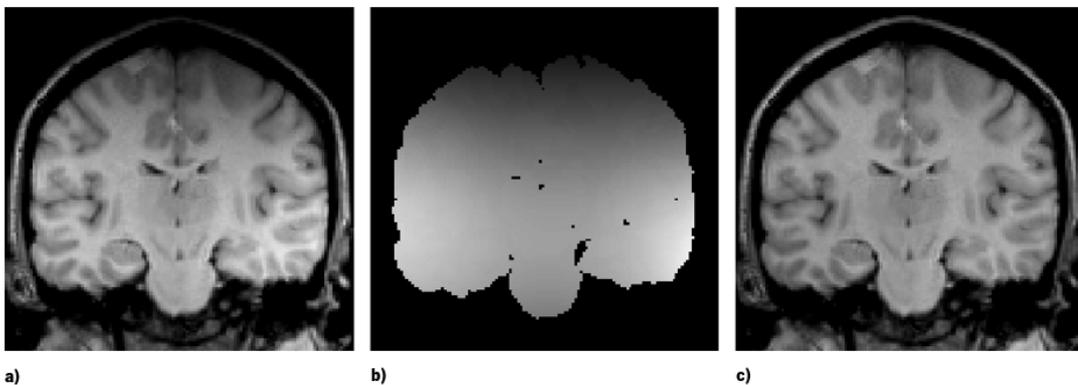


Figura 4.1: Exemplo de falta de homogeneidades: (a) Imagem de ressonância magnética original; (b) Estimção do campo de falta de homogeneidades e (c) Imagem original corrigida. Adaptado de [79].

4.2.1.1 Métodos Prospetivos

Num método prospetivo, a falta homogeneidade das intensidades da ressonância magnética é vista como um erro sistemático, resultante do processo de aquisição. De forma a minimizar o impacto da falta homogeneidade, os modelos prospetivos sugerem a aquisição de imagens adicionais utilizando materiais de referência (fantomas), alteração das bobines de aquisição ou a utilização de protocolos de RM específicos [79].

A estimação da falta de homogeneidade do campo com recurso aos fantomas serve-se de um material uniforme, com propriedades físicas conhecidas. Após aquisição da imagem são aplicadas operações de suavização e redimensionamento das imagens, permitindo assim modelar a falta de homogeneidade presente. Os fantomas geralmente utilizados são a água e o óleo, visto serem materiais com propriedades muito comuns às dos constituintes do organismo humano [90, 91]. Porém, aquando da inexistência de falta de homogeneidades, o presente método, não é capaz de operar. Uma outra desvantagem deste método é a elevada dependência da bobine de aquisição. A bobine de aquisição apresenta variações temporais e espaciais consideráveis, alterando o modelo do campo de falta de homogeneidades, o que requer a aquisição regular de imagens com fantomas. De forma a amenizar este último problema, Wicks *et al.* [92] sugere que aquando da aquisição da imagem com o fantoma seja também registada a orientação e a posição deste, bem como das bobines, para que posteriormente seja possível a aplicação de modelos de transformação geométrica do campo das falta de homogeneidades.

Na RM são utilizados dois tipos de bobines: de superfície e de volume. As bobines de superfície, apesar da boa relação sinal ruído, despoletam uma falta de homogeneidade do campo elevada. Nas bobines de volume, o comportamento é exatamente oposto, isto é, fraca relação sinal ruído, mas baixa falta de homogeneidade. Narayana *et al.* [93] propõe assim a utilização de ambos os tipos de bobines para conciliar as vantagens das bobines de volume e de superfície. O campo de falta de homogeneidade é estimado pela suavização da imagem resultante da divisão da imagem adquirida com as bobines de superfície pela imagem resultante da imagem adquirida com as bobines de volume [93, 94]. As principais desvantagens dos métodos de multi-bobines são os elevados tempos de aquisição, bem como a necessidade de bobines com configurações especiais. Para além destas desvantagens, a falta de homogeneidade da bobine de superfície permanece sempre na imagem final sendo uma correção incompleta [79].

A correção da falta de homogeneidade das intensidades com recurso a sequências de ressonância específicas é dependente da configuração do equipamento utilizado. Com base na distribuição espacial do ângulo de inversão, para determinados pulsos de sequência, é estimada a falta de homogeneidade. Para que seja possível a estimação do campo de falta de homogeneidades é necessária a aquisição de duas imagens, sendo que a segunda contém o dobro do ângulo de inversão da anterior [79]. Uma outra técnica deste tipo de correção é a codificação da sensibilidade através da utilização de bobines de receção múltiplas. Este último procedimento, apesar de ter sido pensado para adicionar rapidez ao processo de aquisição, permite também a redução da falta de homogeneidade [95].

4.2.1.2 Métodos Retrospectivos

Os métodos retrospectivos existentes não são específicos da técnica imagiológica de ressonância magnética, visto que apenas algumas considerações gerais são assumidas aquando da aquisição da imagem, sendo por isso métodos gerais para correção de imagens [79]. Métodos pertencentes a esta classe recorrem à informação presente na imagem, de modo a retirar informação anatômica e ainda a falta de homogeneidade das intensidades. A informação anatômica é descrita por um conhecimento espacial e/ou pela distribuição das probabilidades das intensidades. Esta informação *a priori* é útil para tornar mais célere o processo de tratamento da falta de homogeneidade das intensidades [79]. A vantagem destes métodos em relação aos métodos prospectivos é que, para além da remoção da falta de homogeneidade associada ao equipamento de aquisição da imagem, é ainda capaz de tratar a presença de falta de homogeneidade intrínseca do paciente. Os métodos retrospectivos podem ainda ser desdobrados em quatro subclasses, descritas de seguida.

Vários métodos consideram que a falta de homogeneidade pode ser representada por uma componente de baixa frequência [96–98]. Portanto, alguns métodos retrospectivos designados de métodos de filtragem, filtram as componentes de baixa frequência de forma a remover a falta de homogeneidade das estruturas anatômicas de interesse. Métodos baseados em filtragem são consistentes quando as estruturas anatômicas são de pequenas dimensões, não contendo estruturas passíveis de serem removidas pela filtragem passa-baixo [79]. De imediato se percebe que a implementação de métodos deste tipo na ressonância não acarreta resultados benéficos, visto que existem estruturas de dimensões vastas como é o caso da matéria cinzenta e a matéria branca. Como os espectros destes tecidos anatômicos se sobrepõe ao espectro da falta de homogeneidade, a aplicabilidade dos métodos de filtragem é limitada [99].

Uma outra abordagem para tratamento da falta de homogeneidade reside na extração de um conjunto de características que permitam descrever a sua presença.[79]. As características extraídas são sujeitas à aplicação de uma superfície paramétrica, geralmente numa formulação polinomial ou numa *spline*, que modela a falta de homogeneidade presente [100–102]. Posteriormente, com base no modelo que considera a falta de homogeneidade como multiplicativa, a imagem de entrada é corrigida. O tratamento da falta de homogeneidade, por parte dos métodos deste tipo, varia consoante o conjunto de características. Por exemplo, características extraídas através das intensidades e características computadas sobre o gradiente da imagem são sujeitas a uma superfície paramétrica distinta [79].

Como já foi referido, os métodos baseados em classificadores têm o seu desempenho diminuído devido à presença da falta de homogeneidade [79]. Porém, a aplicação de métodos de correção, após a realização da segmentação, permitem reduzir o impacto causado pela presença da falta de homogeneidade das intensidades. Assim sendo, existem métodos que efetuam a correção da falta de homogeneidade e em simultâneo uma segmentação [88, 103, 104]. A combinação destas duas etapas acarreta benefícios para cada uma delas, ou seja, existe uma melhoria simultânea na segmentação final bem como no método de correção da falta de homogeneidade. Contudo, este tipo de métodos retrospectivos de segmentação apresentam tempos computacionais elevados e uma implementação complexa.

Por último, dentro dos métodos retrospectivos, tem-se os métodos baseados na análise de histogramas, operando diretamente sobre as intensidades. A necessidade de uma inicialização e/ou conheci-

mento *a priori* da distribuição das probabilidades das intensidades das estruturas em causa, é mínima ou inexistente. Assim, este tipo de abordagens são totalmente automáticas e com grande capacidade de generalização, sendo independentes do problema ao qual vão ser aplicadas. Os métodos retrospectivos de segmentação podem operar do mesmo modo, contudo, nestes últimos, é produzida uma segmentação enquanto nos métodos retrospectivos de histogramas o resultado é a própria imagem inicial corrigida.

Os métodos retrospectivos baseados em histogramas contêm três abordagens distintas: minimização da informação, correspondência de histogramas e maximização da alta frequência. Na minimização da informação, a determinação da falta de homogeneidade do campo é obtida pela minimização da entropia da imagem. Nesta abordagem parte-se do pressuposto que a presença da falta de homogeneidade aumenta o valor da entropia da imagem. Por outro lado, na correspondência de histogramas, a imagem é dividida em várias escalas, sub-volumes, de modo a conter falta de homogeneidades constantes. A presença da falta de homogeneidade é detetada através de um histograma modelo, extraído através da imagem completa, que é ajustado para cada um dos histogramas dos sub-volumes. A maximização da alta frequência é a classe de métodos retrospectivos mais utilizada no momento, sendo o método mais comum o método N4, que surgiu de uma melhoria do método N3 [40, 79, 84].

De acordo com o modelo de formação de uma imagem de RM, descrito na equação 4.1, o método N3 (*Nonparametric Nonuniform intensity Normalization*) procura atenuar, iterativamente, a presença da falta de homogeneidade multiplicativa (b). Primeiramente é aplicado o logaritmo à equação 4.1, seguida de uma remoção do ruído n resultando em [84]:

$$\log(v(x)) = \log(u(x) \times b(x)) \equiv \log(u(x) + \log(b(x))) \quad (4.2)$$

Assim, a imagem resultante pode ser descrita pela soma de duas porções, a imagem original (não corrompida) e o campo de falta de homogeneidades. Através de um processo iterativo o campo de falta de homogeneidades é estimado e subtraído à imagem corrompida $v(x)$. A estimação do campo, no método N3, é dada pela aplicação de um operador suave definido por uma *B-spline* uniforme. A *B-spline* não é mais de que uma curva definida por um conjunto de pontos de controlo, igualmente espaçados entre si. Verificou-se que o método de N3 é incapaz de implementar distâncias curtas, na definição da sua *B-spline*, que levam a resultados mais robustos. Assim, surgiu o método *Improved N3 Bias Correction* (N4ITK), proposto por Tustison *et al.* [40], que substitui o modelo de *B-spline* de forma a permitir menores distâncias entre os pontos de controlo. Complementarmente, em cada nova iteração realizada, invés de se utilizar a imagem corrompida inicial é utilizada a imagem corrigida na iteração anterior, sendo o processo iterativo entendido como um refinamento gradual do campo das homogeneidades [40]. O método N4ITK foi o método selecionado para o processamento das falta de homogeneidades neste trabalho.

4.2.2 Normalização das Intensidades

A ressonância magnética foi uma das técnicas de imagiologia radiológica que revolucionou o estudo das estruturas internas do corpo humano. O facto de ser uma abordagem não invasiva, sem fatores de risco associados à saúde, conhecidos, torna esta técnica vantajosa [23]. Na RM é possível a aplicação de diferentes protocolos, o que leva a que diferentes tecidos tenham diferentes contrastes dentro da mesma

estrutura em análise, por exemplo o cérebro [8]. Porém, o valor das intensidades na RM não tem um significado qualitativo fixo associado, o que acarreta problemas para os métodos de segmentação de imagens [104, 105]. Para um mesmo protocolo, adquirido no mesmo equipamento, e para a mesma estrutura anatômica em análise do mesmo paciente, as intensidades tomam valores diferentes em cada uma das aquisições. A implicação imediata deste problema é que a visualização da imagem requer sempre um ajuste da gama das intensidades. Assim sendo, o desafio passa por garantir uma coerência nas intensidades das imagens resultantes de protocolos semelhantes. O objetivo pode ser alcançado através de uma calibração do equipamento da ressonância magnética, através do uso dos fantomas [106, 107], ou através de técnicas de processamento dos dados após a aquisição da imagem [39, 108–110]. A atribuição de uma mesma escala de intensidades com o método de calibração do sinal de RM, para além de ser um procedimento que consome recursos, não é capaz de tratar imagens que já foram adquiridas sem calibração com os fantomas. Já as técnicas de processamento que são aplicadas após aquisição das imagens, não necessitam de parâmetros especiais de aquisição, surgem como metodologias mais atrativas sendo estes os métodos abordados nesta secção.

À partida parece trivial o procedimento de normalização através de uma simples transformação do mínimo e máximo das intensidades de uma imagem para uma gama de intensidades fixa. Porém, esta abordagem não é capaz de assegurar valores de intensidades similares para diferentes pacientes. Uma outra abordagem reside no ajuste automático do contraste e brilho das imagens de RM para visualização. Apesar de, na visualização, ser possível detetar uma regularização do valor das intensidades, estas ainda não têm um significado característico de cada tecido. Após a aplicação da transformação, as intensidades não têm uma coerência a nível inter e intra-paciente. Nyúl e Udupa [108] desenvolveram um método de normalização que procura contornar os problemas dos restantes métodos. O método procura, com recurso a transformações não-lineares, manipular o histograma de uma dada imagem, para um histograma de referência (*standard*).

Na normalização de [108] esta é realizada em duas fases: uma primeira fase de treino, executada uma única vez, para um dado protocolo de RM de uma determinada estrutura, e uma fase de transformação das imagens. Na fase de treino são estimados os parâmetros do histograma de referência tendo por base a informação fornecida pelas imagens de treino (população de treino). Os parâmetros estimados são posteriormente utilizados na fase de transformação, onde os parâmetros da imagem a transformar são mapeados para os parâmetros do histograma de referência. Após o mapeamento dos parâmetros, o histograma da imagem a normalizar sofre uma deformação, para que este seja compatível com o histograma *standard*. A deformação do histograma traduz-se numa transformação das intensidades, sendo que este processo não-linear, é o responsável por atribuir uma nova escala de intensidades à imagem. Contudo, apesar das deformações aplicadas, a relação inicialmente existente entre as intensidades dos diferentes tecidos é mantida [108].

No modelo de normalização supracitado, considera-se que numa imagem de RM, a cauda superior e inferior do seu histograma correspondem a artefactos e a *outliers*. Tendo isto em mente, é definida uma gama de intensidades de interesse, estável, capaz de reduzir a gama das intensidades real da imagem. Assim, a estimação dos parâmetros em vez de considerar a gama das intensidades mínima e máxima reais, m_1 e m_2 (respetivamente), é realizada sobre uma gama de intensidades de interesse $[p_1, p_2]$

definida pelos percentis mínimo e máximo $[pc_1, pc_2]$. A fase de treino procura determinar os parâmetros relativos ao mínimo e ao máximo das intensidades de interesse da escala padrão, s_1 e s_2 respetivamente, e à moda da região a transformar, μ_s . As intensidades de interesse, $[pj_1, pj_2]$, e a moda μ_j de cada uma das j imagens de treino, são mapeadas para a escala das intensidades normalizadas, s_1 e s_2 , de modo linear. Após a aplicação do mapeamento é obtida a moda μ_s através da média das modas de cada uma das imagens de treino mapeadas μ_{js} . O mapeamento é descrito da seguinte forma [108]:

$$x' = s_1 + \frac{x - p_{1j}}{p_{2j} - p_{1j}}(s_2 - s_1) \quad (4.3)$$

Na fase de transformação, as imagens a normalizar sofrem primeiramente um mapeamento para a gama de intensidades definida pelo histograma de referência, em dois estágios distintos. Primeiramente, na imagem i a normalizar são mapeadas as intensidades do intervalo $[p_{1i}, \mu_i]$ para $[1, \mu_s]$. A segunda fase mapeia o intervalo $[\mu_i, p_{2i}]$ para $[\mu_s, s_2]$. O procedimento realizado na fase de transformação encontra-se presente na equação 4.4 [108].

$$\tau_i = \begin{cases} \mu_s + (x - \mu_i) \frac{s_1 - \mu_s}{p_{1i} - \mu_i}, & m_{1i} < x < \mu_i \\ \mu_s + (x - \mu_i) \frac{s_2 - \mu_s}{p_{2i} - \mu_i}, & \mu_i < x < m_{2i} \end{cases} \quad (4.4)$$

A definição da moda é dependente do tipo de histograma a normalizar [110]. Num histograma bimodal, isto é, com a presença de dois cumes (tal como se pode observar na figura 4.2a), μ representa a moda secundária obtida após eliminação da moda principal, que corresponde ao *background*, com recurso a um *threshold* aplicado ao valor médio da imagem. Num histograma uni-modal (figura 4.2b), não é possível definir uma moda secundária sendo que μ passa a ser caracterizado pela cauda da bossa principal (do *background*) [39, 110].

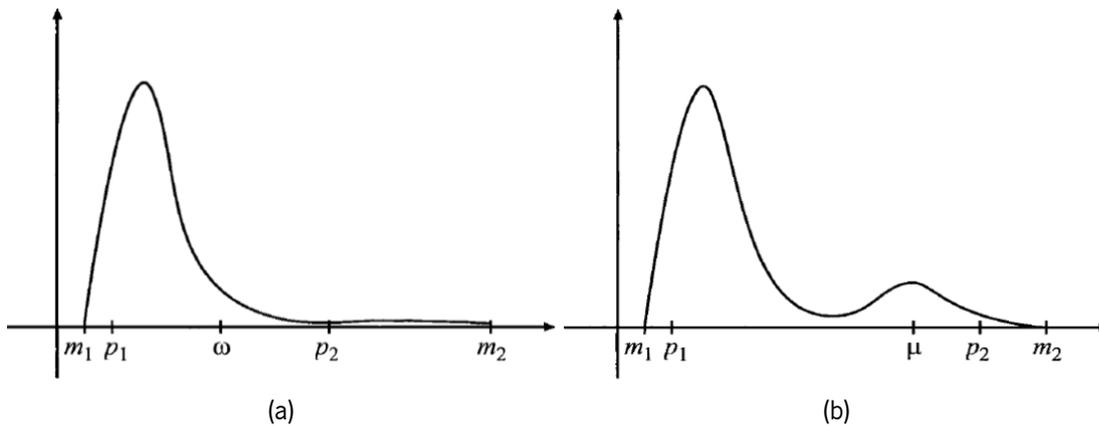


Figura 4.2: Análise do histograma de uma imagem com uma distribuição: (a) uni-modal e (b) bimodal. Adaptado de [108]

Posteriormente, [109] apresentaram algumas alterações ao método de normalização. Constatou-se que a metodologia baseada na moda não se encontra adequada ao tópico de segmentação de imagens, onde as intensidades devem apresentar um significado e coerência mais robustos, mesmo para gamas de intensidades pequenas. A alteração passa por adequar a transformação das intensidades de acordo com o problema em causa. Por exemplo, a definição de s_1 não deve conter o valor nulo caso intensidades

abaixo de p_1 necessitem de ser distinguidas de 0. Complementarmente, $s_2 - s_1$ devem ter uma gama que não provoque a sobreposição das intensidades vizinhas. Contudo, o maior contributo resulta da definição de uma escala designada de "escala aberta", que ocorre quando $pc_1 > 0$ e $pc_2 < 100$ onde ambos os valores $[m_{1i}, p_{1i}]$ e $[p_{2i}, m_{2i}]$ são mapeados para $[s_{1i}, s_1]$ e $[s_2, s_{2i}]$, respetivamente. Com esta abordagem torna-se possível definir gamas de intensidades de interesse melhores, sem qualquer tipo de perda de informação discriminativa nas caudas do histograma.

Como a moda μ é um parâmetro variável mediante o tipo de histograma, em Nyúl e Udupa [110] o valor das intensidades dos percentis intermédios resultam das médias das intensidades do conjunto de histogramas de treino. Nesta nova abordagem, invés de se estimar a moda, procura-se estimar a intensidade associada ao percentil. Após a determinação dos valores de intensidades do conjunto de treino, de um dado percentil do histograma de referência, obtém-se o respetivo valor da intensidade que não é mais do que a média do conjunto. Na figura 4.3 encontra-se um exemplo de treino, com duas imagens, para determinação dos parâmetros do histograma de referência.

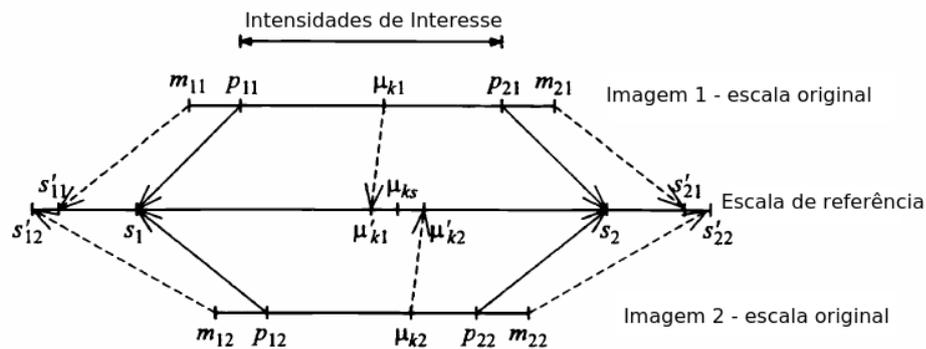


Figura 4.3: Exemplo de duas imagens de treino $j = 1, 2$ utilizadas para a procura dos parâmetros do histograma de referência. Adaptado de [110].

Nyúl e Udupa [110] apresentou ainda uma melhoria à sua metodologia, que passa pela utilização de um conjunto mais vasto de percentis que asseguraram imagens normalizadas com significados mais consistentes e gamas de intensidades mais apropriadas para os diferentes tecidos. Nesta última proposta, o método de Nyúl-Udupa apresenta ainda uma modalidade para normalizar os valores extremos da escala de intensidades de referência. Os valores extremos s'_{1i} e s'_{2i} , da escala de referência, resultam da aplicação do mesmo mapeamento linear aplicado aos segmentos mínimos e máximos das intensidades de interesse. Assim sendo, as caudas do histograma normalizado resultam da aplicação das transformações: $[p_{1i}, \mu_{1i}]$ para $[s_1, \mu_{1s}]$ e $[\mu_{li}, p_{2i}]$ para $[\mu_{ls}, s_2]$. A gama de intensidades $[m_{1i}, p_{1i}]$ e $[p_{2i}, m_{2i}]$, do histograma a normalizar, geralmente, encontram-se associadas a artefactos e a pontos *outliers* responsáveis pelas variações inter- e intra-paciente das imagens de RM. Daí que a normalização destas intensidades segue o mapeamento idêntico dos intervalos extremos da gama de intensidades de interesse, pois são regiões de estáveis das intensidades. Na figura 4.4 encontra-se a função de mapeamento das intensidades, aplicada na fase de transformação, geral para l percentis.

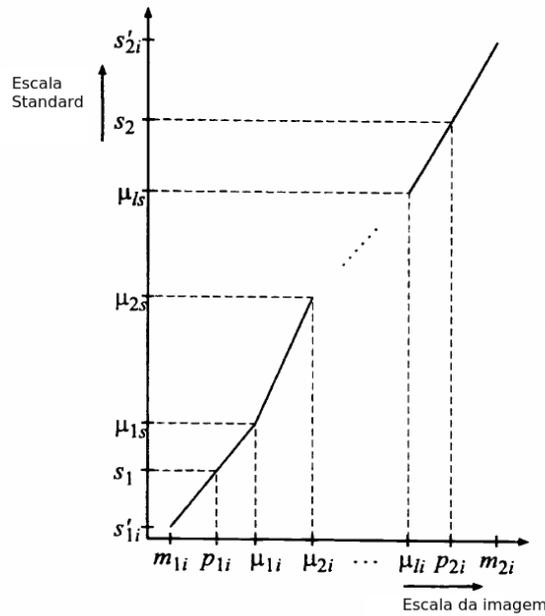


Figura 4.4: Função de mapeamento das intensidades para o caso geral de l percentis. Adaptado de [110].

4.3 Aprendizagem Computorizada - Classificação

A aprendizagem computorizada é um ramo da área do reconhecimento de padrões, cuja finalidade é a atribuição automática de uma categoria, ou classe, à amostra em análise. A amostra, ou objeto, compreende diversas entidades, como por exemplo uma imagem médica, um sinal ou outro tipo de dados passíveis de categorização. Mediante o número de classes a discriminar a modalidade da classificação é distinta. No caso em que apenas se procura a distinção de duas classes tem-se um problema de classificação binária. Em problemas de classificação que definem várias classes tem-se uma classificação multi-classe [61]. A atribuição de uma classe a uma amostra, procedimento designado de classificação, requer uma primeira etapa de aprendizagem. O sucesso da classificação é assim dependente da capacidade em identificar padrões no conjunto de dados em análise, representativos de cada uma das classes definidas na fase de aprendizagem. Mediante os fundamentos utilizados para a realização desta última fase, os algoritmos de aprendizagem computorizada podem ser divididos em três tipos: supervisionados, não-supervisionados e semi-supervisionados [61, 111].

No presente trabalho são implementados dois algoritmos de aprendizagem supervisionados: a Floresta de Decisão e a *Extra-trees*, com o intuito de classificar os *voxels* de uma imagem tridimensional, de ressonância magnética, em quatro classes de tecido tumoral (patológico) e uma classe correspondente ao tecido saudável (*background*).

4.3.1 Fundamentos da Floresta de Decisão

Durante a década de 90, a Floresta de Decisão teve forte concorrência por parte do método de classificação SVM. Neste último método procura-se definir um hiper-plano, no espaço de características

n-dimensional, por forma a distinguir as diferentes classes. A colocação do hiper-plano é realizada de modo a maximizar a distância entre o ponto mais próximo e o hiper-plano. A fácil compreensão e a possibilidade de desenvolver uma estrutura de otimização perceptível foram propriedades que forneceram popularidade ao método de classificação SVM. Porém, as Florestas de Decisão destacaram-se dada a sua capacidade intrínseca em lidar com problemas de segmentação multi-classe. A sua capacidade de generalização e eficiência foram também propriedades que permitiram dar ainda mais destaque à Floresta de Decisão [112–114].

A Floresta de Decisão resulta da conjugação de várias árvores de decisão. As árvores de decisão foram introduzidas por Breiman *et al.* [115] em meados da década de 80. A árvore de decisão é entendida como um classificador fraco, pois a sua capacidade de classificação é ligeiramente superior à de uma classificação aleatória. Porém, o treino conjunto destes classificadores resulta num classificador forte capaz de assegurar precisão e capacidade de generalização. Através da combinação linear dos vários classificadores fracos obtém-se um classificador final capaz de se correlacionar fortemente com a verdadeira classificação. Para a obtenção do classificador forte deve-se na fase de treino, de cada classificador fraco, atribuir pesos distintos às amostras. Deste modo, em cada estágio de treino o classificador fraco seguinte procura dar mais peso aos dados cujo desempenho do classificador fraco anterior foi menor. O procedimento descrito é designado de *boosting* [37, 111].

4.3.1.1 Árvores de Decisão

Uma árvore de decisão é um grafo com uma configuração própria, correspondendo a uma coleção de nodos e vértices dispostos de forma hierárquica. Os nodos internos são designados de nodos de divisão ao par que os nodos terminais correspondem às folhas. Todos os nodos exceto a raiz (nodo inicial) têm um e um só vértice de chegada, associado e dois vértices de partida, pelo que, contrariamente ao que se constata na maioria dos grafos, uma árvore de decisão não é um grafo cíclico. A estrutura de um grafo encontra-se presente na figura 4.5.

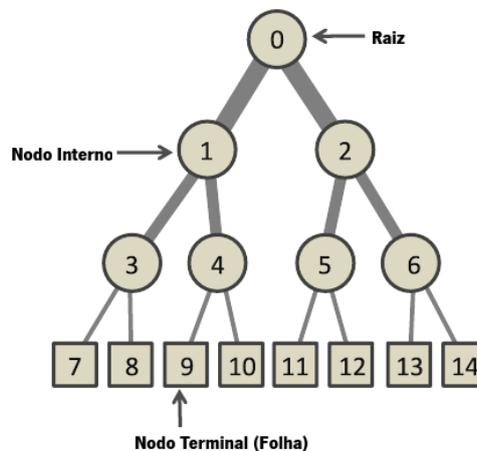


Figura 4.5: Estrutura de um grafo com a configuração em árvore. Adaptado de [37].

A árvore de decisão, dada a sua configuração hierárquica, é capaz de resolver problemas de elevada

complexidade através da aplicação de uma série de testes de baixa complexidade. Cada nodo interno é responsável pela aplicação de um dos testes de modo a permitir a sucessiva divisão dos dados até se alcançar uma folha. Deste modo, em cada folha encontra-se a resposta mais provável com base na série de divisões anteriormente efetuadas. Perante um conjunto de pontos, estes são sujeitos a um grupo de testes realizados sobre as propriedades que caracterizam cada ponto, gerando um resultado binário (verdadeiro ou falso). Com base no resultado é realizada a separação dos dados pelos nodos filhos até se alcançar o nodo terminal. Com o avançar dos testes realizados, o grau de confiança aumenta, explicado pelo aumento da separabilidade de dados não relacionados [37, 114].

O desempenho de uma árvore de decisão é superior quando os testes associados a cada nodo interno se encontram bem definidos e ainda os métodos de divisão dos dados são os corretos para o problema. A árvore de decisão, baseando-se num processo supervisionado, na fase de treino, otimiza os parâmetros de cada nodo interno, responsáveis pela divisão dos dados que minimizem a função de energia, selecionada previamente. Com a fase de treino pretende-se que na fase de teste cada nodo interno seja capaz de distinguir novos dados pelos seus nodos filhos, de forma apropriada [37, 114].

4.3.1.2 Formulação Matemática

Um objeto, quando alvo de análise por parte de uma árvore de decisão, é representado matematicamente por um vetor $v = (x_1, x_2, \dots, x_d) \in \mathfrak{R}_d$, onde cada um dos componentes de v corresponde à uma propriedade que o caracteriza. Aos elementos do conjunto, x_i , designam-se de características. As características computadas variam mediante o problema em causa. A título de exemplo, na área das ciências da visão v corresponde a um *voxel* na imagem, ao par que x_i é o resultado de um conjunto de filtragens efetuadas sobre a intensidade do respetivo *voxel*. O tipo e o número de características variam de acordo com o tipo de dados presentes e a aplicação. Em termos, teóricos o espaço dimensional definido pelo vetor de características pode ser extremamente vasto ou mesmo infinito. Contudo, em termos práticos, não existe a necessidade de um número vasto de características mas sim um subconjunto $\mathfrak{R}_{d'}$ capaz de assegurar e providenciar informação precisa e discriminativa, ou seja, $d' \ll d$. [37]

As características correspondem à informação sobre a qual cada nodo da árvore de decisão efetua o teste associado. Assim sendo, as características são sujeitas à função de teste de cada nodo. A função de teste de um nodo divisão j é dada por: [37]

$$h(v, \theta_j) : \mathfrak{R}_d \times T \rightarrow 0, 1 \quad (4.5)$$

O resultado da função de teste é binário, variando de acordo com condição aplicada. Na equação 4.5 $\theta_j \in T$ corresponde aos parâmetros de divisão dos dados associados ao nodo j , enquanto T representa o espaço de todos os parâmetros da árvore. Em conformidade com o resultado da função de teste, à chegada do ponto v ao nodo j este é enviado para o nodo filho esquerdo ou direito. Os modelos que definem as funções de teste podem ser de dois tipos: lineares e não-lineares. Nos modelos lineares a separação dos dados pela função de teste é realizada com recurso a uma linha ou um hiper-plano com coordenadas homogêneas. Por outro lado, no caso dos modelos não-lineares a separação dos pontos é realizada através da definição de hiper-planos com superfícies de graus de liberdade superior, como é o

caso de uma superfície quadrática [37].

4.3.1.3 Treino e Teste

Na fase de treino, alternativamente designada de fase *offline*, é utilizado um conjunto de dados de treino de modo a, automaticamente, definir a melhor combinação de parâmetros das funções de divisão $h(v, \theta)$ para cada nodo interno j . Em cada j nodo, de acordo com o subconjunto de treino de S_j é aprendida a função objetivo que melhor separa os dados de S_j pelos seus nodos filhos S_j^L e S_j^R . Com a aplicação da fase de treino procura-se então a maximização da função objetivo, descrita matematicamente pela expressão 4.6 [37].

$$\theta_j = \arg_{\theta \in T} \max I(S_j, \theta) \quad (4.6)$$

A maximização da função objetivo é alcançada através da procura da melhor combinação de variáveis, para a divisão de um conjunto de amostras S_j definidos por um conjunto de parâmetros discretos θ . Os subconjuntos resultantes da divisão S_L e S_R , que correspondem aos nodos filhos esquerdo e direito respetivamente, são definidos por: [37]

$$\begin{aligned} S_j^L(S_j, \theta) &\{(v, y) \in S_j | h(v, \theta) = 0\} \\ S_j^R(S_j, \theta) &\{(v, y) \in S_j | h(v, \theta) = 1\} \end{aligned} \quad (4.7)$$

Como se pode averiguar pelas expressões 4.7, na fase de treino existe uma marcação y associada a cada objeto que não é mais do que a classificação final que a árvore procura apresentar num dos seus nodos filho. Na divisão dos dados não ocorre sobreposição de informação, isto é, o subconjunto presente no nodo S_j^L e no nodo S_j^R não contêm objetos em comum. É esta última característica que permite que, à medida que se percorre a árvore, o subconjunto dos dados formados seja mais puro e intrínseco de uma classe c^* [37].

A função objetivo implementada é importante para assegurar precisão na classificação, visto que esta tem influência no percurso realizado pelas amostras, ao longo da árvore de decisão. A divisão dos dados, realizada pela função objetivo, tem por base modelos de energia. O modelo mais comum aplica o conceito de ganho de informação, I , para assegurar uma separação dos dados gradual e mais precisa. A separação dos dados em subconjuntos com características mais semelhantes resultam numa redução da medida de incerteza associada a uma variável aleatória, o que desencadeia um aumento de I . A redução de incerteza é monitorizada pela entropia, H [37].

$$I = H(S) - \sum_{i \in L, R} \frac{|S^i|}{|S|} H(S^i) \quad (4.8)$$

Na equação 4.8, H corresponde à entropia, que para o caso de distribuições discretas (problemas de classificação) é dada pela expressão 4.9 [37].

$$H(S) = - \sum_{c \in C} p(c) \log(p(c)) \quad (4.9)$$

A entropia contabiliza o conjunto de todas as classes C definidas para um dado problema. Na equação 4.9, $p(c)$ representa a distribuição empírica resultante das amostras de treino pertencentes ao conjunto S , ou seja, a expressão que os dados pertencentes a uma classe c têm no conjunto S . A distribuição é obtida através do histograma normalizado de cada uma das classes. O ganho de informação I , definido em 4.8 resulta então da diferença da entropia calculada para o conjunto original e a soma das entropias para os subconjuntos S, L .

Para além da entropia existem outros critérios de separação dos dados, destacando-se a razão de erros de classificação e a impureza de Gini [116].

A razão de erros de classificação, E , consiste na proporção de dados classificados como classe c^* mas que na realidade, na classificação manual correspondem a uma classe distinta c_i , conforme o que se encontra na expressão 4.10 [116, 117].

$$E(S) = \frac{1}{S} \sum_{i \in S} 1_{se}(c_i \neq c^*) = 1 - p(c^*) \quad (4.10)$$

Por outro lado, a impureza de Gini [118] é uma função objetivo que quantifica a probabilidade de um determinado ponto ser mal classificado caso este fosse aleatoriamente classificado em conformidade com a distribuição das classes num determinado nodo. Assim a combinação de parâmetros de um nodo é melhor quanto maior for a impureza de Gini, descrito pela expressão 4.11.

$$G(S) = \sum_{c \in C} p(c)(1 - p(c)) = 1 - \sum_{c \in C} p(c)^2 \quad (4.11)$$

O presente critério de separação dos dados é ainda uma ferramenta útil para computar a importância relativa de um dado conjunto de características. Esta última propriedade faz da impureza de Gini uma função objetivo interessante de ser utilizada na área da imagem médica, pois permite avaliar se as características que têm mais impacto no classificador permitem discriminar melhor as regiões ou estruturas associadas de relevo, por exemplo, regiões patológicas. A importância de uma dada característica é obtida através do somatório dos decréscimos da impureza de Gini, que ocorrem em todos os nodos da árvore. [37, 118]

Após a estimação dos parâmetros para cada um dos nodos internos da árvore, os nodos terminais (folhas) compreendem um subconjunto de pontos cujo comportamento é similar. Assim sendo, é possível associar a cada uma das folhas, uma das classes definidas para o problema, visto que os dados são similares. Através da distribuição condicionada definida pelo conjunto de dados presentes em cada folha, é possível definir as probabilidades posteriores de cada classe, $p(c|v)$. A classe, c^* é atribuída a um dado ponto através da probabilidade de maior valor¹ presente na folha [37], ou seja:

$$c^* = \arg_{c \max} p(c|v) \quad (4.12)$$

Mediante um ponto ainda não observado v' , são aplicados os vários testes previamente selecionados de modo a atribuir uma classificação ao objeto em questão. Com base nas características que definem

¹Do inglês: *Maximum A-Posteriori*

v são aplicadas as funções de teste, selecionadas durante a fase de treino, de modo a que o ponto alcance um dos nodos terminais. Iniciando-se o teste na raiz da árvore, de acordo com o resultado binário providenciado pela função de teste, o objeto é enviado para o nodo interno esquerdo ou direito. O procedimento é repetido até que v' alcance um nodo folha. No nodo folha, através do estimador é associado uma classe, c^* ao valor de entrada v' . Todo este processo é designado de teste de uma árvore de decisão que exige que as funções de divisão nos nodos internos e os estimadores nos nodos folha se encontrem já com os seus parâmetros otimizados, numa fase prévia de treino [37].

4.3.2 Floresta de Decisão

Uma floresta de decisão aleatória resulta de um agregado de árvores de decisão. Cada uma das árvores é aleatoriamente diferente das restantes, como se pode constatar na figura 4.6, onde 3 árvores diferentes, para o mesmo conjunto de dados v , produzem resultados distintos. Uma floresta de decisão aleatória assenta nos mesmos princípios enunciados para a árvore de decisão. O facto de cada uma das árvores ser aleatoriamente diferente permite diminuir a correlação dos dados entre as previsões realizadas por cada uma das árvores, o que resulta numa grande capacidade de generalização e robustez na fase de classificação. Através do parâmetro de aleatoriedade ρ , é controlada a quantidade de aleatoriedade intrínseca de cada árvore e ainda o grau de correlação entre as diferentes árvores de uma floresta [37, 114].

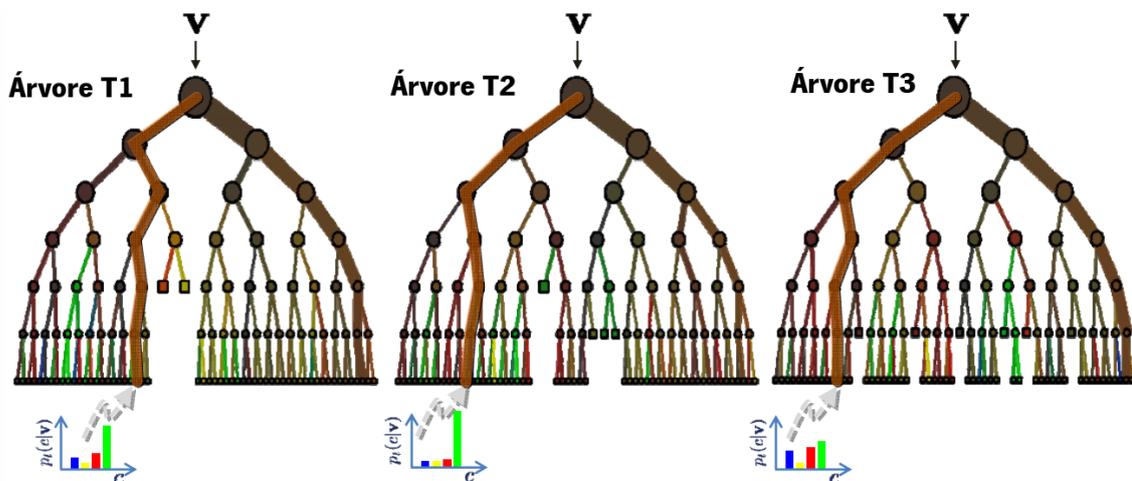


Figura 4.6: Resultado produzido por diferentes árvores treinadas em condições aleatórias e portanto distintas. Adaptado de [37].

A aleatoriedade é assegurada na fase de treino e pode ser assegurada de dois modos distintos. O primeiro modo baseia-se na realização de uma amostragem aleatória do conjunto de treino. A grande motivação por detrás desta filosofia é a redução do *overfitting*. A amostragem aleatória do conjunto de treino não se serve da totalidade dos dados o que leva a que o treino seja mais imediato. Contudo, a existência de desperdício de informação de treino para ajustar os parâmetros de cada árvore não é vantajoso visto que pode limitar a capacidade de generalização da floresta. Daí que surgiu uma segunda abordagem, designada de otimização aleatória dos nodos. Recordando a expressão 4.6, na otimização

aleatória do nodo, ao invés de se otimizar todo o conjunto de parâmetros T (que acarreta grande custos de eficiência) recorre-se de um pequeno conjunto $T_j \subset T$ quando se encontra a treinar o nodo j . Vulgarmente, $|T| = \infty$, isto é, o número de parâmetros a estimar é de elevada ordem, pelo que se torna necessária a definição do parâmetro extra $\rho = |T_j|$ que controla o grau de aleatoriedade, idêntico para todas as árvores. O parâmetro $\rho \in 1, \dots, T$ define então a quantidade de informação a utilizar em todos os nodos de todas as árvores, sendo que $\rho = 1$ significa que apenas um parâmetro aleatório é selecionado, enquanto que $\rho = |T|$ utiliza todos os parâmetros. Numa floresta, em que $\rho = |T|$ todas as árvores são idênticas sendo o seu grau de correlação máximo [37].

Após a fase de treino, durante o teste, cada um dos pontos v é simultaneamente colocado em cada uma das árvores da floresta. Mediante o resultado definido pelo vetor de características associado a v , este alcança os nodos terminais de cada uma das árvores presentes na floresta com uma classificação associada. Através da combinação de todas as classificações providenciadas pelas árvores da floresta, com recurso à técnica de MAP aplicada à probabilidade posterior total, obtém-se a classificação final. A classificação final pode ser alcançada através da média das probabilidades das árvores da florestas (equação 4.13) ou através da multiplicação dos resultados de cada uma das árvores seguida da aplicação de uma função de partição (Z). Visto que as árvores de decisão não são estatisticamente independentes, a função de partição tem como finalidade assegurar a normalização probabilística (4.14) [37, 114].

$$p(c|v) = \frac{1}{T} \sum_{t=1}^T p_t(c|v) \quad (4.13)$$

$$p(c|v) = \frac{1}{Z} \prod_{t=1}^T p_t(c|v) \quad (4.14)$$

4.3.2.1 Influência dos Híper-parâmetros de uma Floresta

No treino da Floresta de Decisão, para além de se otimizar a função objetivo de cada nodo das árvores de decisão procura-se também otimizar a estrutura de cada árvore. Este processo de otimização deve ser adequado ao problema de classificação. Em termos hipotéticos poder-se-ia pensar que uma Floresta de Decisão teria uma melhor desempenho quanto maior for a profundidade das árvores de decisão que a constituem. Porém, profundidades elevadas podem conduzir a que o número de pontos que chegam às folhas ser reduzido, conduzindo a um *overfitting* aos dados de treino. Deste modo, o desempenho e robustez de uma Floresta de Decisão é influenciado por um conjunto de híper-parâmetros [119]:

- Profundidade das árvores – com o aumento da profundidade as árvores adquirem uma maior sensibilidade na separação dos dados e por conseguinte classificações com um maior grau de confiança. Porém, em florestas com um valor de profundidade elevado, a divisão dos dados passa a ser de tal forma elevada que *outliers* podem passar a ser classificados como uma classe isolada, o que se traduz na presença de *overfitting*. O classificador fica adaptado ao conjunto de treino de modo a que precisão e generalização da floresta é limitada aquando da classificação do conjunto de teste. Adicionalmente, uma profundidade elevada pode resultar em árvores de decisão não

balanceadas. A regra de ouro para a definição da profundidade das árvores de uma floresta passa pela avaliação da dimensão do espaço de características e do tipo de problema em estudo [119, 120].

- Número de árvores — O tamanho de uma floresta deve ser tanto maior quanto os recursos computacionais o permitirem. Em cada árvore de decisão a função de separação define um limite rígido de separação entre os pontos das classes. Este fenómeno advém do facto de cada árvore ser treinada de forma diferente das restantes, pois o subconjunto de pontos de treino é distinto. Consequentemente, cada árvore apresenta um grau de confiança elevado (em regiões de incerteza) aquando da classificação de objetos de teste, o que é indesejado. Assim ao combinar o resultado das diferentes árvores pertencentes a uma mesma floresta, a classificação de objetos com características próximas dos objetos de treino é mais eficaz e ao mesmo tempo, para objetos fora da zona de confiança a classificação é mais insegura (difusa) [119].
- Função de divisão — os modelos selecionados para definir a função de divisão dos dados em cada um dos nodos internos de uma árvore desempenha um papel principal na qualidade da classificação e ainda na eficiência do resultado. O modo como varia a confiança do estimador à medida que os pontos de teste se distanciam dos dados de treino deve ser tido em consideração. A figura 4.7 contém um exemplo prático do desempenho de diferentes modelos [119].

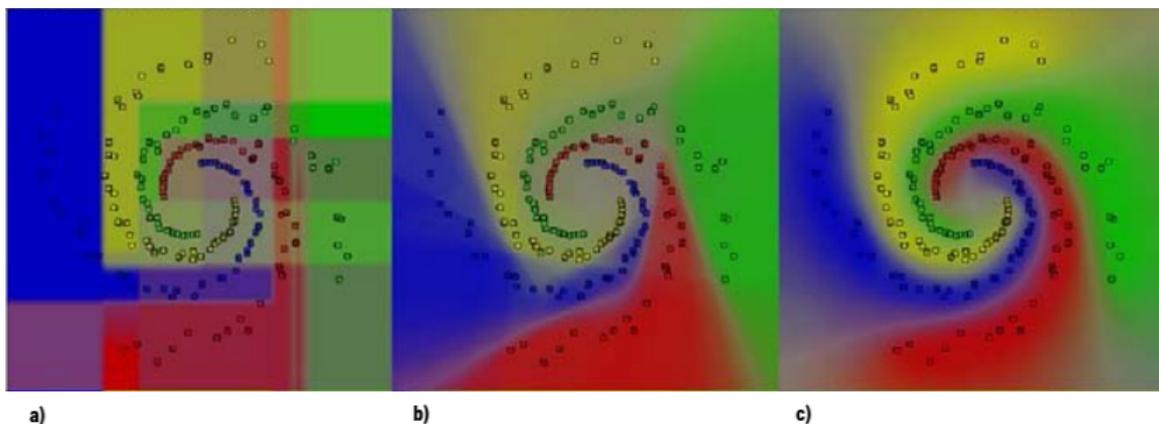


Figura 4.7: Mapa de certeza de uma floresta treinada com diferentes funções de divisão dos dados: o modelo linear (a) de eixos alinhados e modelos curvilíneos de linhas orientadas (b) e secções cónicas (c). Adaptado de [37].

- Aleatoriedade ρ — um maior valor de ρ conduz a uma menor correlação entre as árvores da mesma floresta, aumentando assim a capacidade de generalização, visto que as árvores se tornam únicas e diferentes entre si. A aleatoriedade é útil em reduzir o número de parâmetros a otimizar, especialmente em funções de divisão não lineares. Contudo, a confiança geral é diminuída, podendo este problema ser compensado por um aumento da profundidade das árvores. Quando ρ procura apenas otimizar um único parâmetro do espaço total de parâmetros de θ , a classificação torna-se computacionalmente demorada [119].
- Características — a fase de desenvolvimento das características a introduzir no classificador é dependente do problema em estudo, bem como das estruturas a identificar. Assim sendo, o espaço dimensional de características deve ser tal que permita a discriminação entre as diferentes classes, sem grande dispersão dos dados e sobreposição mínima entre as classes [119].
- Proporção das amostras — árvores não balanceadas, ou seja, com uma desproporção na expressão de cada uma das classes devem ser evitadas. O porquê desta afirmação reside no facto na presença de uma árvore não balanceada, esta corresponder a uma cadeia de nodos com pouca capacidade discriminativa entre as diferentes características, conduzindo a uma capacidade de generalização de baixo grau. A amostragem é então um passo fundamental para que exista equilíbrio na distribuição entre as diferentes classes [121].

4.3.3 Extra-Tree

A designação *Extra-Tree* corresponde a árvores de decisão extremamente aleatórias, que se destacam devido à sua capacidade em lidar com espaços de previsão mais complexos, ou seja, múltiplas classes. O algoritmo por detrás da *Extra-Tree* é bastante semelhante ao de uma Floresta de Decisão, introduzido com o intuito de fornecer ainda mais aleatoriedade às funções de divisão dos nodos internos. Com esta alteração procura-se assegurar maior precisão na classificação, através da diminuição da variância dos dados (o que conduz a uma redução no tempo de treino). Relembrando a formulação matemática sobre a qual foi enunciado o comportamento de uma árvore de decisão, numa árvore de decisão extremamente aleatória é possível identificar 4 propriedades diferentes [122, 123]:

- Por defeito cada árvore é construída de forma independente das restantes, ou seja, a profundidade de cada árvore não se encontra limitada.
- Não existe qualquer fase de amostragem dos dados de treino, sendo que cada árvore recebe a totalidade dos dados de treino disponíveis.
- As funções de divisão presentes nos nodos internos são modelos assentes em divisões de eixos alinhados, o que em termos matemáticos resulta numa divisão binária descrita pela formula 4.15.

$$h(v, (i, \tau))[x_i < \tau] \tag{4.15}$$

Cada um das características x_i , $1 \leq i \leq d$ apresenta assim um *threshold* associado definido em conformidade com a gama de valores da característica.

- A otimização da função de divisão é realizada, em cada nodo interno, através de uma seleção prévia de um conjunto aleatório de características, T_j , do espaço T . Dentro o conjunto definido são identificadas as que asseguram um melhor desempenho da função objetivo. Contrariamente ao que se realiza na árvore de decisão, os *thresholds* discretos não são valores a otimizar, mas sim constantes, definidas em conformidade com a sua gama.
- O crescimento de uma árvore de decisão extremamente aleatória, para além do parâmetro de profundidade máxima definido para a fase de treino (tal como numa árvores de decisão) tem ainda dois critérios que controlam a sua profundidade. Um primeiro que condiciona o crescimento da árvore quando as amostras presentes num nodo são puras e um segundo parâmetro que impõe um número mínimo de amostras para se realizar a divisão do nodo em nodos filhos.

4.4 Sumário

Com este capítulo descreveram-se os conceitos teóricos que permitiram o desenvolvimento do algoritmo de segmentação de tumores cerebrais da presente dissertação. Tratando-se de um modelo que se recorre de classificadores supervisionados, a fase de pré-processamento é importante. O alinhamento das sequências de RM, a eliminação de artefactos como a falta de homogeneidade do campo magnético, bem como assegurar um significado às intensidades, através da normalização, são procedimentos chave. A primeira etapa de pré-processamento descrita encontra-se já aplicada pela base de dados utilizada. Na correção da falta de homogeneidade foi selecionado o método N4 [40], dado apresentar-se como a técnica atual mais robusta para a remoção deste artefacto. No caso da normalização das intensidades, o método de Nyúl *et al.* [39] foi escolhido, dado que foi desenvolvido especificamente para a RM. Neste algoritmo são aplicadas transformações lineares aos histogramas de forma a coincidir com um histograma de referência. Os classificadores utilizados neste trabalho assentam nos conceitos da Floresta de Decisão, que resultam da conjugação de classificadores fracos treinados de forma independente — as árvores de decisão. A fase de treino é responsável pela otimização de uma função de divisão capaz de dividir os dados pelos nodos de cada árvore de decisão, tendo em conta a informação presente num conjunto de características. Assim, a otimização da função de divisão procura encontrar o melhor conjunto de *thresholds* a aplicar ao respetivo conjunto de características. A classificação final resulta da combinação das contribuições de cada uma das árvores, pelo cálculo da média ou multiplicação das probabilidades posteriores dos nodos folha.

Implementação Prática

Neste capítulo encontra-se a descrição da metodologia implementada e as diversas abordagens estudadas no desenvolvimento do algoritmo de segmentação automática de tumores. O funcionamento dos algoritmos utilizados será também explicado, independentemente de resultarem da aplicação de propostas desenvolvidas por terceiros ou algoritmos criados de raiz.

O método de segmentação tumoral segue a linha de pensamento de um algoritmo assente numa classificação supervisionada. Por forma a ter uma visão geral dos diferentes estágios que definem o algoritmo de segmentação, na figura 5.1 apresenta-se a linha de processamento da informação das diversas imagens de RM.

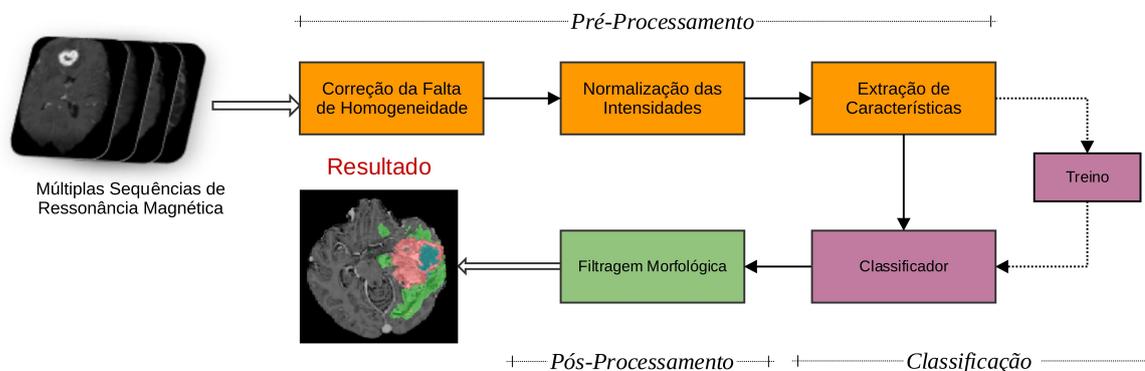


Figura 5.1: Estruturação do algoritmo de segmentação tumoral de gliomas proposto.

Em conformidade com a estrutura da metodologia proposta, exposta na figura 5.1, será descrita a implementação prática de cada uma das fases constituintes.

Previamente serão descritas as linguagens de programação utilizadas na projeção da metodologia de segmentação.

5.1 Desenvolvimento do Método de Segmentação

As funções desenvolvidas para cumprir as diversas etapas da metodologia de segmentação assentaram na linguagem de programação *Python* [124]. O projeto desenvolvido tem a capacidade de realizar a

leitura das imagens da base de dados, criação de máscaras, normalização, extração de características, classificação, armazenamento dos resultados e pós-processamento.

Na linguagem de programação utilizada encontram-se disponíveis vários pacotes, como por exemplo o *mahotas* [125], *numpy* [126], *scipy* [127], *pytables* e *sklearn* [124]. O acesso a estes pacotes foi vantajoso na medida em que permitiram uma implementação mais otimizada e imediata da metodologia de segmentação.

Complementarmente, algumas funções foram desenvolvidas em *C++* e *Cython* dada a existência de etapas chave que necessitavam de uma otimização da sua eficiência. Daí que, os algoritmos como a etapa de normalização e extração de características tenham sido desenvolvidos em *C++*.

5.2 Pré-processamento

Em conformidade com os conceitos teóricos descritos em 4.2 são, de seguida, apresentadas as etapas de pré-processamento efetuadas ao longo da presente dissertação.

5.2.1 Correção da Falta de Homogeneidade

A correção da falta de homogeneidade é um procedimento usualmente aplicado a imagens de RM, especialmente em situações em que são utilizados métodos cuja dependência do valor das intensidades é elevado, tal como no presente caso. A correção da falta de homogeneidade das intensidades assenta no método N4 [40], que se encontra disponível no *Advanced Normalization Tools* (ANTs) [128]. Os parâmetros foram os definidos por defeito.

5.2.2 Normalização das Intensidades

A normalização das intensidades, baseada na proposta de normalização de Nyúl *et al.* [39], foi um procedimento desenvolvido de raiz.

Tal como descrito por Nyúl *et al.* [39], para que seja possível a transformação do histograma é necessária uma fase prévia de treino, cuja finalidade é definir um histograma de referência e as marcas das intensidades associadas. Primeiramente são determinados os valores de intensidade mínima e máxima do histograma de referência, s_1 e s_2 . Os valores obtidos estão associados aos valores mínimos e máximo do intervalo de intensidades de interesse do histograma de referência. Os parâmetros s_1 e s_2 provêm de uma análise à variabilidade associada a cada um dos percentis p_{c_j} que contabilizam os valores das intensidades do conjunto de treino. Para tal é utilizada a diferença do desvio padrão, $\Delta\sigma_X$, caracterizando a variabilidade das intensidades entre dois percentis consecutivos. Perante a aplicação de um *threshold* sobre $\Delta\sigma_X$ são obtidos s_1 e s_2 .

De seguida, define-se um conjunto de percentis intermédios, do histograma de referência. O valor de intensidade de cada percentil resulta do cálculo da média das intensidades, do respetivo percentil, do conjunto de histogramas de treino. Vários conjuntos de percentis intermédios podem ser implementados

[39]. Na presente dissertação utilizou-se um conjunto de 9 percentis intermédios, definido na expressão 5.1:

$$L = \{s_1, 10, 20, 30, 40, 50, 60, 70, 80, 90, s_2\} \quad (5.1)$$

O procedimento de definição dos percentis é crítico. Uma má definição deste conjunto pode despoletar uma compactação ou alongamento de gamas de intensidades, deteriorando-se a qualidade da informação.

Após a fase de treino, tem-se a normalização das intensidades. A fase de normalização é realizada em duas etapas. Na primeira etapa, numa nova imagem são determinados os percentis intermédios do seu histograma e linearmente mapeadas para a escala de referência L . Na segunda etapa é realizado um novo mapeamento linear entre as marcas. A realização dos mapeamentos lineares resulta da aplicação da função 4.3.

Por último, é realizado o tratamento das caudas inferior e superior do histograma. Geralmente, estas regiões encontram-se associadas a *outliers*, onde a definição de um modelo próprio de mapeamento linear é extremamente difícil dada a variabilidade presente. Assim sendo, a gama inferior das intensidades resulta da aplicação do mesmo mapeamento linear, entre marcas, nas porções mais extremas das intensidades de interesse. O que leva a que a transformação da gama de intensidades $[p_{1i}, \mu_{1i}]$ seja idêntica à transformação aplicada para a gama $[s_1, \mu_{1s}]$, para a normalização da cauda inferior. A cauda superior compreendida pelo intervalo $[\mu_{1i}, p_{2i}]$ é normalizada pelo mapeamento linear de $[\mu_{1s}, s_2]$.

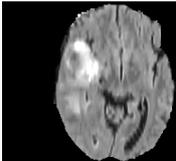
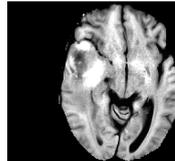
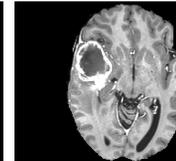
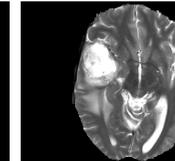
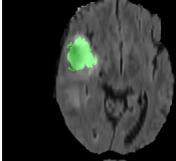
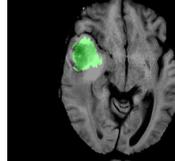
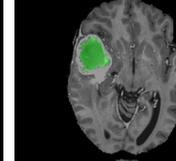
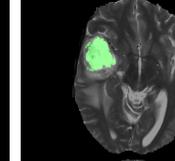
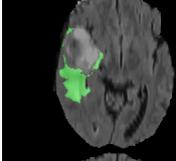
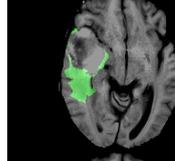
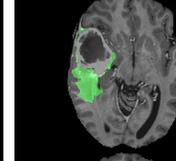
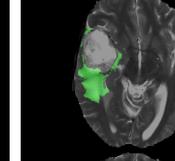
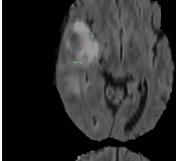
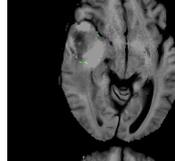
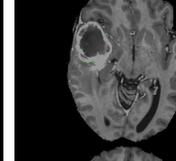
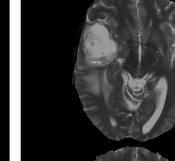
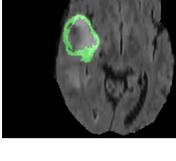
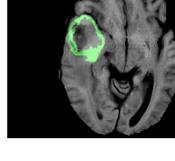
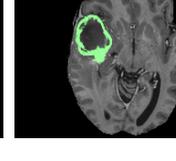
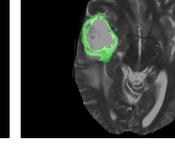
O método de normalização proposto por Nyúl *et al.* [39] apesar de ser direcionado para a normalização de imagens de RM cerebrais, não é um algoritmo que intrinsecamente seja capaz de contabilizar a presença de regiões tumorais. A grande maioria dos tecidos cerebrais tumorais, na técnica de RM, tem uma aparência hipo-intensa ou hiper-intensa [8]. Este fenómeno leva a que uma grande quantidade de amostras se localizem nos extremos do histograma, que correspondem a zonas características de grande variabilidade. Complementarmente, apesar de o mapeamento linear aplicado ser idêntico ao dos intervalos extremos da gama de intensidades de interesse, por defeito, não existe qualquer limite para as intensidades resultantes da transformação das caudas do histograma o que pode desencadear uma compactação ou alongamento do histograma indesejados. Intuitivamente pode-se pensar que um intervalo de intensidades de interesse maior pode contornar estes entraves. Contudo, um aumento do número de amostras vistas como *outliers* pode também afetar a qualidade da informação presente na região estável do histograma. Assim sendo, constata-se que modo como a normalização das caudas do histograma é realizado, bem como a definição do intervalo de intensidades de interesse, no contexto de segmentação cerebral tumoral é um procedimento de relevo.

5.3 Características

A marcação e delineação manual de um tumor, com base na informação providenciada pelas diferentes sequências da ressonância magnética, assenta fundamentalmente na perceção do contraste visual assegurada pelas intensidades [8]. Na presença de diferentes protocolos de ressonância magnética, é possível discriminar a estrutura tumoral do restante tecido saudável circundante e ainda identificar as

diferentes classes de tecido tumoral, através da comparação das propriedades dos tecidos. Na tabela 5.1 encontra-se a segmentação manual de cada uma das classes de tecido tumoral, sobreposta aos diferentes protocolos da RM, por forma a compreender o comportamento de cada um dos tecidos tumorais do glioma.

Tabela 5.1: Exemplo de visualização das intensidades dos diferentes tecidos tumorais, marcados manualmente, nas quatro sequências providenciadas pela base de dados, num tumor de elevado grau.

	FLAIR	T1	T1c	T2
Sequência Normalizada				
Região 1				
Região 2				
Região 3				
Região 4				

Diferentes protocolos da RM permitem fornecer percepções distintas da região tumoral. As sequências T2 e FLAIR permitem dar uma noção da localização espacial do tumor, ao par que a T1c discrimina com alguma robustez o tecido necrótico e tecido ativo. Por outro lado a T1 procura fundamentalmente dar uma noção de contextualização anatómica cerebral [8].

Várias metodologias de segmentação automática procuram modelar a percepção de um especialista aquando da segmentação manual, recorrendo-se da computação de um conjunto de características extraídas através das intensidades das imagens. A extração de características tem como finalidade caracterizar uma imagem através de informação representativa de cada uma das classes de tecido a segmentar [129].

No presente trabalho, as características extraídas podem ser agrupadas em três grandes classes: locais, contexto e textura. Nas secções que se seguem são descritos os diferentes conjuntos de características utilizadas, procurando evidenciar a sua capacidade de distinção e finalidade.

5.3.1 Características Locais

As características locais englobam um conjunto de propriedades vasto, extraído em relação ao ponto em análise ou a uma vizinhança definida em torno do mesmo. Características de intensidade, médias locais, Gradientes Gaussianos¹ e Diferença de Gradientes Gaussianos são parte integrante do conjunto de características locais.

As intensidades da ressonância magnética são a característica base do sistema de segmentação, que permitem a computação de características com um maior grau de complexidade. Aquando do processo de aquisição da RM o comportamento da região em análise é caracterizado pelos valores das intensidades, definidos numa escala de níveis de cinzento. O valor das intensidades de uma região varia consoante a sequência de RM, tirando-se partido desta capacidade multispectral para a classificação de tecidos cerebrais tumorais, tal como foi enunciado em 2.4.2.1. Níveis de cinzento mais baixos correspondem a zonas mais escuras, hipo-intensas, ao par que regiões com níveis de cinzento mais elevados compreendem regiões hiper-intensas e portanto zonas mais claras [8].

De acordo com a tabela 5.1, geralmente em T1c, a necrose apresenta um comportamento hipo-intenso e a zona tumoral ativa um comportamento hiper-intenso, pois são estes os tecidos que captam o agente contrastante. Esta informação é quase que exclusiva de T1c. Porém, a definição das restantes classes de tecido tumoral, edema e região ativa não contrastante, não é realizada do mesmo modo imediato que o descrito para a necrose e região ativa. O edema e a região ativa não contrastante requerem uma análise da informação de forma exclusiva. Para tal, numa primeira fase, pelas sequências T2 e FLAIR é delineado todo o tumor do restante tecido circundante saudável, pois o tumor tem um comportamento hiper-intenso. Assim, através da diferença das regiões identificadas pela T1c e as regiões T2 e FLAIR é possível obter o edema e região ativa não contrastante. Com recurso a uma segunda observação da T1c é identificada região ativa não contrastante, que é um tecido que não capta o agente contrastante e por isso tem baixa intensidade na T1c, mas superior à do tecido cerebral saudável. Por exclusão de partes tem-se a região restante, o edema [8].

Em concordância com conceitos supracitados, as características de intensidade utilizadas são 7, onde 4 correspondem às intensidades das próprias sequências (FLAIR, T1, T1c e T2) e as restantes às diferenças entre determinadas sequências, nomeadamente FLAIR-T2, T1c-FLAIR e T1c-T2. As intensidades resultantes das diferenças não são utilizadas para posterior computação de características, onde a sua finalidade é garantir uma discriminação mais robusta das diversas classes de tecido a classificar, entre as duas sequências, modelando o comportamento do especialista aquando da realização da segmentação manual.

Para além destas características foram ainda computados dois tipos de características, para cada uma das sequências, assentes na transformação não-linear das intensidades: a transformação exponencial e a logarítmica. Através da aplicação de uma exponencial às intensidades de uma sequência, procura-se saturar os níveis inferiores e superiores das intensidades, ou seja, dar mais relevo a regiões com comportamentos hipo-intensos e hiper-intensos. No caso concreto da T1c, toda a região do tecido

¹ Designação comumente utilizada para descrever o cálculo da magnitude do Gradiente efetuado sobre imagens após a aplicação de uma filtragem com um *kernel* Gaussiano

ativo, apresenta na característica exponencial, um valor máximo e regular, ao par que a região necrótica apresenta um valor mínimo e regular. Outra vantagem da extração da característica exponencial é permitir uma compressão das intensidades intermédias que correspondem maioritariamente às regiões de tecido saudável, tidas como estáveis aquando do treino da normalização de Nyúl-Udupa. Com a aplicação do logaritmo, devido à sua saturação imediata intrínseca, procura-se enfatizar transições nas intensidades de menor valor, o que é útil, no caso da T2, para que seja possível destacar o edema tumoral. A extração de ambas as características em análise é apenas efetuada após uma transformação linear das intensidades de cada uma das sequências para intervalos de interesse pré-definidos, presentes na tabela 5.2.

Tabela 5.2: Gama de valores definidos para a transformação linear das intensidades de FLAIR, T1, T1c e T2 para posterior extração das características exponencial e logaritmo.

	Exponencial	Logaritmo
FLAIR	[3, 7]	[1, 3]
T1	[2, 7]	[1, 2]
T1c	[2, 7]	[1, 6]
T2	[3, 7]	[1, 7]

Um outro tipo de características locais extraídas das imagens RM foram as médias locais. Com a introdução destas características procura-se obter uma perceção geral e grosseira da localização espacial do tumor. As médias locais são filtros passa-baixo que se servem de um *kernel* (máscara que define a região de operação), sobre o qual é calculada a média das intensidades e atribuída ao ponto central do *kernel* tridimensional. A definição do *kernel* sobre o qual são computadas as médias locais é um procedimento que não deve ser menosprezado. O *kernel* é parametrizado pela variável w , tendo sido selecionados diferentes valores de w para se obter noções distintas acerca da caracterização, de um modo grosseiro, das fronteiras tumorais. Assim, $w \in \{3 \text{ mm}, 9 \text{ mm}, 19 \text{ mm}\}$.

Por último, em termos de características locais foi ainda computado o Gradiente Gaussiano e a Diferença de Gaussianas (DoG), cuja intuição é assegurar detalhe no contorno entre as fronteiras entre tumor e tecido saudável e ainda entre edema e a região central. A filtragem gaussiana é uma filtragem passa-alto, o que intrinsecamente enfatiza as transições abruptas nas intensidades da imagem.

5.3.2 Características de Contexto

As características de contexto têm grande relevo na definição de regiões fronteiriças, em que o valor das intensidades se encontra numa região intermédia entre duas ou mais classes, tendo sido inicialmente introduzido na área da segmentação cerebral tumoral por Zikic *et al.* [67] e Geremia *et al.* [68].

As características de contexto compreendem características computadas tendo em conta a informação presente numa vizinhança. Neste tipo de características o ponto em análise não é tido em consideração.

O contexto procura quantificar o grau de similaridade de um determinado ponto com as regiões que o rodeiam. As características de contexto procuram modelar o procedimento de segmentação manual de uma imagem de ressonância magnética. Por natureza o especialista tem em consideração o contexto em que uma dada região da imagem se insere e com base na diferença entre as intensidades envolventes,

é atribuído ao *voxel* uma das classes possíveis. Assim sendo, neste tipo de características observam-se regiões de dimensões e localização espacial distintas da região que se encontra a segmentar.

Com o intuito de modelar a percepção de um radiologista, as características de contexto são controladas por dois parâmetros: a distância d e a dimensão do plano bidimensional l , extraídos ao longo de 6 planos distintos. Após a definição das regiões bidimensionais, nestas são determinados os valores máximo e mínimo das intensidades, armazenando-se as respectivas coordenadas, para cada um dos 6 planos (topo, fundo, lado esquerdo, lado direito, frente e trás). Posteriormente, sobre uma outra característica já extraída é atribuído a cada *voxel* o valor da característica dado pelas coordenadas previamente determinadas. Em termos visuais na figura 5.2 encontra-se um exemplo do contexto.

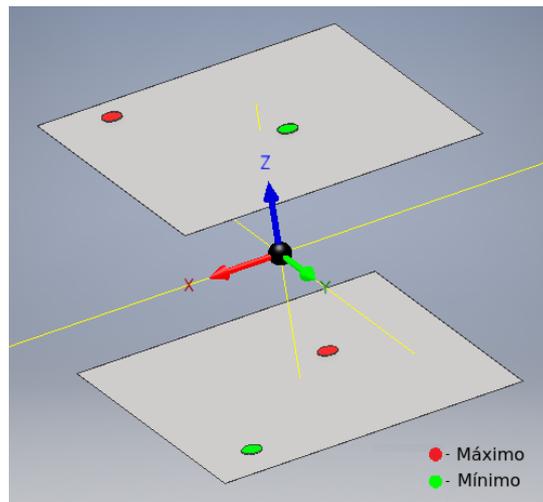


Figura 5.2: Abordagem implementada para a extração do contexto das imagens. Por questões de percepção apenas se encontram 2 dos 6 planos implementados.

Uma combinação de parâmetros $\{l, d\}$ resulta em 12 características computadas, visto que 6 resultam da obtenção das coordenadas relativas ao máximo da região bidimensional e as restantes ao mínimo da região. Na presente implementação $l \in \{3\}$ e $d \in \{4\}$.

5.3.3 Características de Textura

A textura de uma imagem não tem uma definição globalmente aceite, apresentando diferentes definições mediante a aplicação em particular. Contudo, a textura pode ser entendida como uma variação espacial das intensidades, o que se traduz numa contabilização espacial da intensidade dos pixels. A extração deste tipo de características é extremamente sensível ao ruído, pelo que fases de pré-processamento são importantes, para que não se introduzam pseudo-texturas dominantes. A análise da textura é um procedimento estatístico que pode ser dividido em texturas de primeira ordem, segunda ordem e elevada ordem [130].

Nas características de textura de primeira ordem, através da distribuição das intensidades (histograma) de uma imagem ou uma região, obtém-se o número de ocorrências em função do respetivo nível de intensidades, que posteriormente são normalizadas para que correspondam a probabilidades. Através das probabilidades é definido um conjunto de sete características que descrevem as propriedades do his-

tograma: média, variância, granularidade, coeficiente de assimetria, curtose, energia e entropia. Apesar da grande facilidade de computação, a capacidade de discriminação das presentes características, no contexto do problema, não é eficaz, dado que não existe nenhuma noção de correlação espacial [130].

As texturas de segunda ordem assentam no cálculo de matrizes de coocorrência. A matriz de coocorrência, corresponde a um histograma bidimensional de níveis de cinzento, caracterizada por uma dada distância d e direção θ . Os valores presentes na matriz de coocorrência representam a probabilidade de um pixel com intensidade i ser encontrado a uma distância, d , e direção θ do pixel com intensidade j , numa região de cálculo pré-definida. Este procedimento encontra-se descrito na figura 5.3, retirando-se que numa imagem com N_c níveis de cinzento, a matriz de coocorrência resultante terá uma dimensão de $N_c \times N_c$. Numa imagem bidimensional, um dado ponto em análise apresenta 8 vizinhos segundo as 4 direções principais (direções colineares têm a mesma adjacência) definindo-se por isso 4 matrizes de coocorrência para descrever a textura de segunda ordem de uma imagem (uma para cada direção). As matrizes originadas são matrizes simétricas, sendo que os valores únicos se encontram ao longo da sua diagonal. Os valores acima da diagonal da matriz são idênticos ao valores que se encontram abaixo desta [130]. A informação produzida por cada uma das matrizes é então uma probabilidade, onde os índices das linhas e colunas são codificadas pelas intensidades (ver expressão matricial 5.2).

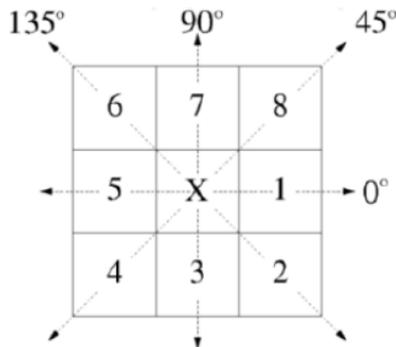


Figura 5.3: Vizinhança bidimensional em torno de um pixel de análise x com as possíveis direções utilizadas para o cálculo das matrizes de coocorrência.

$$G = \begin{bmatrix} p(1, 1) & p(1, 2) & \cdots & p(1, N_c) \\ p(2, 1) & p(2, 2) & \cdots & p(2, N_c) \\ \vdots & \vdots & \ddots & \vdots \\ p(N_c, 1) & p(N_c, 2) & \cdots & p(N_c, N_c) \end{bmatrix} \quad (5.2)$$

Com base na matriz de coocorrência, Haralick [131], em 1973, projetou 14 métricas com a finalidade de capturar a informação de textura através das matrizes de coocorrência. Estas podem ser agrupadas em duas classes. Na primeira classe caracterizam-se as propriedades intrínsecas das imagens, como é o caso da homogeneidade e contraste. Já na segunda classe define a complexidade e a natureza de transições de níveis de cinzento na imagem. Uma das características desenvolvidas por Haralick (f_{14}) é computacionalmente instável pelo que é globalmente aceite apenas o cálculo de 13 características, descritas do seguinte modo:

$$\begin{aligned}
 f_1 &= \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{d,\theta}(i, j)^2 && \text{Momento Angular Secundário} \\
 f_2 &= \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{d,\theta}(i, j)^2 \right\}, \text{ onde } n = |i - j| && \text{Contraste} \\
 f_3 &= \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{d,\theta}(i, j) \log(p_{d,\theta}(i, j)) && \text{Correlação} \\
 f_4 &= \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - \mu)^2 p_{d,\theta}(i, j) \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} && \text{Variância} \\
 f_5 &= \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{d,\theta}(i, j)^2 && \text{Diferença do Momento Inverso} \\
 f_6 &= \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{1}{1 + (i - j)^2} p_{d,\theta}(i, j) && \text{Soma das Médias} \\
 f_7 &= \sum_{i=0}^{2(N_g-1)} i \cdot p_{x+y}(i) && \text{Soma das Variâncias} \\
 f_8 &= \sum_{i=0}^{2(N_g-1)} (i - f_7)^2 p_{x+y}(i) && \text{Soma das Entropias} \\
 f_9 &= - \sum_{i=0}^{2(N_g-1)} p_{x+y}(i) \log(p_{x+y}(i)) && \text{Entropia} \\
 f_{10} &= \sum_{i=0}^{N_g-1} (i - f'_{10})^2 p_{x-y}(i) && \text{Diferença das variâncias} \\
 f_{11} &= - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i)) && \text{Diferença das entropias} \\
 f_{12} &= \frac{HXY - HXY1}{\max(HX, HY)} && \text{Medida de Correlação 1} \\
 f_{13} &= (1 - e^{[-2(HXY2 - HXY)])^{1/2}} && \text{Medida de Correlação 2}
 \end{aligned} \tag{5.3}$$

A textura extraída a partir das matrizes de coocorrência, é uma propriedade muito dependente da noção angular, pelo que o vetor de deslocamento d e o ângulo θ são parâmetros que devem ser alvo de um estudo experimental.

Um outro tipo de textura comumente utilizado é a textura de Gabor. A textura de Gabor é uma característica que advém da aplicação de um conjunto de filtros assentes nos conceitos da Transformada *Wavelet*. Deste modo, a textura de Gabor é capaz de modelar com confiança as várias estruturas dada a sua propriedade de multi-resolução e multi-direção. A filtragem é geralmente definida através de um *kernel* Gaussiano. Várias escalas e orientações podem ser implementadas para análise da textura de Gabor. Por vezes, dado o seu custo computacional são aproximadas a filtros bidimensionais discretos ou filtros ortogonais. A inclusão de características de Gabor no âmbito da segmentação cerebral tumoral encontra-se presente nos métodos propostos por Parisot *et al.* [132] e Subbanna *et al.* [133]

Alternativamente a textura pode ser computada pela aplicação direta da Transformada *Wavelet*, designando-se de textura fractal. A textura fractal tem como finalidade a caracterização da textura não homogénea aleatória e variável no espaço ao longo de diferentes escalas. Islam *et al.* [70] realiza a extração da textura fractal no seu algoritmo de segmentação tumoral.

Por último, a textura pode ser quantificada através de uma metodologia designada de banco de filtros de *Leung Malik*, testada por Zhang *et al.* [134] no âmbito da segmentação cerebral tumoral. As características correspondem à resposta a um conjunto de filtros. São aplicados 48 filtros, com diferentes resoluções e direções, sendo eles a 1ª e 2ª derivada em 6 orientações diferentes em 3 escalas diferentes, juntamente com 4 filtros Gaussianos e 8 filtros LoG (Laplaciano de Gaussianas).

5.3.3.1 Extração das Características de Haralick

A extração de características de textura apenas foi efetuada para a textura de Haralick. Para a extração deste tipo de características foi utilizado um algoritmo já desenvolvido, presente na livraria *mahotas* [125]. Este algoritmo permite o cálculo nas 4 direções definidas por Haralick de 13 características de textura.

Tal como foi referido, uma matriz de coocorrência tem as suas dimensões definidas pelo nível de cinza máximo, N_c presente numa imagem, resultando matrizes de dimensões $N_c \times N_c$. No caso das imagens de RM as matrizes de coocorrência geradas serão de elevadas dimensões, pois níveis máximos das intensidades normalizadas encontram-se no intervalo [1300, 3200]. Assim sendo, a obtenção das matrizes de coocorrência têm tempos de computação associados extremamente elevados. Por forma a ultrapassar este problema, numa primeira fase do estudo das características de Haralick implementou-se uma redução da gama de cinzentos nas imagens. Um outro parâmetro responsável por reduzir o tempo computacional foi a definição de uma região de cálculo de dimensões reduzidas, passando a região de interesse a ser definida por um plano bidimensional de 5×5 , contendo o ponto em análise.

A nova escala de intensidades de 256 níveis de cinzento, recorre-se de uma transformação não-linear das intensidades: a exponencial. Nesta transformação optou-se por definir uma transformação exponencial de duas caudas, semelhante a um polinómio de terceiro grau, com o intuito de minimizar a perda da informação presente nas caudas de um histograma, isto porque, dizem respeito às regiões hipo-intensas e hiper-intensas onde se concentram os tecidos tumorais, nas quatro sequências avaliadas.

Na equação 5.4 encontra-se definida a exponencial implementada, onde é possível confirmar a partição da escala de 256 níveis de cinza em três diferentes transformações não lineares.

$$\begin{aligned}
 y(x) &= (255 - 73) + e^{\frac{x-72}{offset_1-x}}, & \text{se } x > \lim_{superior} \\
 y(x) &= (128 - 73) + e^{\frac{x-128}{offset_2-x}}, & \text{se } \lim_{inferior} < x < \lim_{superior} \\
 y(x) &= (55 + e) \times \left[\frac{1 - e^{-\frac{x-min_{sequência}}{\sigma}}}{1 + e^{-\frac{x-min_{sequência}}{\sigma}}} \right], & \text{se } x < \lim_{inferior}
 \end{aligned} \tag{5.4}$$

Na equação anterior $\lim_{superior}$ e $\lim_{inferior}$ são *thresholds* definidos especificamente para cada uma das sequências, tendo em consideração a escala das intensidades do tecido saudável e tecido tumoral. A definição destes limites resultou de uma análise dos histogramas dos diversos pacientes, em ambos os tipos de glioma, por forma a obter um maior grau de discriminação das regiões hiper-intensas e hipointensas do tumor, localizadas nas caudas do histograma, em relação do tecido saudável. Para cada um dos tipos de transformações é ainda possível observar a existência de diferentes parâmetros (*offsets*) cujo intuito é adequar o comportamento da exponencial à nova escala de intensidades sobre a qual posteriormente serão extraídas as características de Haralick. A definição de cada *offsets* está definido na expressão 5.5.

$$\begin{aligned}
 offset_1 &= max_{sequência} + \frac{max_{sequência} - 72}{\log(72)} \\
 offset_2 &= \lim_{superior} + \frac{\lim_{superior} - 128}{\log(128)} \\
 \sigma &= \frac{\lim_{inferior}}{5}
 \end{aligned} \tag{5.5}$$

5.3.4 Características de Segunda Ordem

As características resultantes das transformações não-lineares às intensidades, a exponencial e o logaritmo, permitem delinear as zonas fronteiriças das diversas estruturas presentes no cérebro. Tal observação suscitou o interesse de utilizar estas características como base de computação para posterior extração de características, como o Gradiente Gaussiano e ainda as médias locais. As características DoG foram também extraídas sobre as não linearidades, porém a sua capacidade discriminativa em pouco se alterou. Daí que se optou por abandonar a computação deste tipo de características. O cálculo das médias locais, que ocorre apenas sobre a exponencial, foi também alvo de alterações, tendo-se redefinido o modo como estas são computadas. Na sua configuração inicial estas características tinham como região de cálculo uma vizinhança tridimensional. Colocou-se em questão o facto da região de cálculo contabilizar um número elevado de pontos, para o cálculo da média do *voxel* em análise, e portanto a variabilidade destes conjunto de pontos ser assinalável, limitando o desempenho do classificador. Para contornar esta suspeita as médias locais passaram a ser características computadas sobre regiões bidimensionais. A extração das características médias locais foi efetuada com o mesmo conjunto de parâmetros definidos na subsecção 5.3.1.

Com a definição desta nova modalidade de características, as características de contexto foram tam-

bém alvo de análise. Para tal, esta última classe de característica foi reforçada, tendo sido extraída em novos planos definidos em novas distâncias e tamanhos das vizinhanças, $l \in \{3, 5, 7\}$ e $d \in \{3, 5, 7\}$.

Na figura 5.4 é possível observar a diferença das características mediante a sua base de computação. Ao novo conjunto de características, dada a sua dimensionalidade, optou-se por designar de características de segunda ordem.

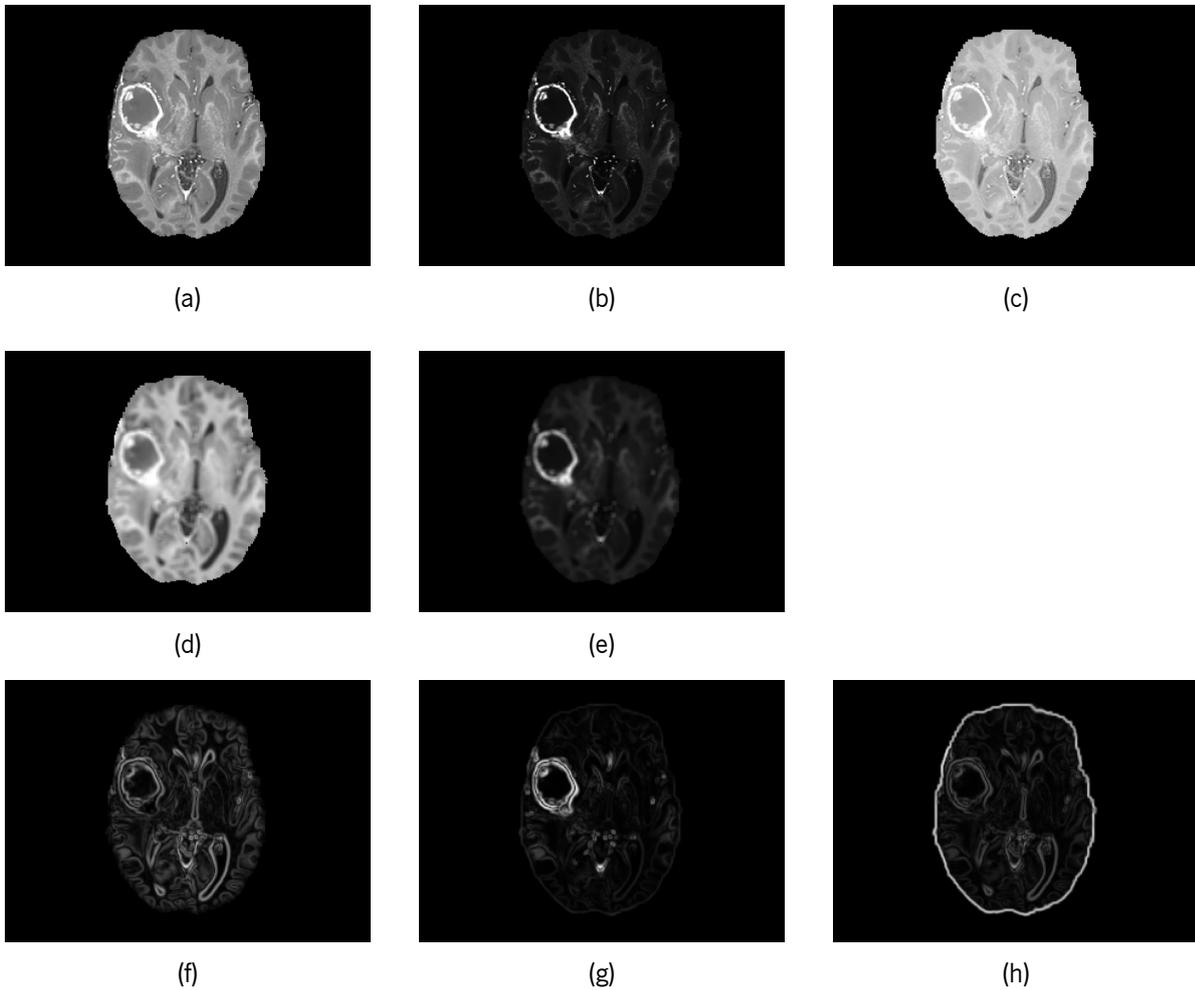


Figura 5.4: Exemplo de características de um paciente de HGG numa fatia de T1c. Na primeira coluna encontram-se as características base sobre as quais são computadas as características da segunda e terceira coluna de figuras. (a) Intensidades da sequência normalizada; (b) Exponencial; (c) Logaritmo; (d) Extração da média local de 3 sobre as intensidades (a); (e) Extração da média local com $w = 3$ sobre a exponencial (b); (f) Extração do Gradiente Gaussiano, com $\sigma = 3$ sobre as intensidades (c); (g) Extração do Gradiente Gaussiano, com $\sigma = 3$ sobre a exponencial (b); (h) Extração do Gradiente Gaussiano, com $\sigma = 3$ sobre o logaritmo (c).

5.4 Classificação

Na realização do presente trabalho foram abordados dois tipos de classificadores distintos, a Floresta de Decisão e Extra-Trees. Ambos os classificadores encontram-se já implementados no *scikit-learn* [127].

5.4.1 Híper-parâmetros da Floresta

A vantagem da utilização do pacote *scikit-learn* reside no facto de permitir a configuração dos hiper-parâmetros de ambos os tipos do classificadores aplicados. Em conformidade com os conceitos teóricos descritos em 4.3.2, os parâmetros mais relevantes, passíveis de configuração são:

- Número de estimadores (n) – número de árvores que constituem a floresta;
- Profundidade das árvores ($depth$) – sendo que na Extra-Trees, por defeito, cada árvore tem a capacidade de ter uma profundidade distinta das outras, ao par que na Floresta de Decisão tal não se verifica;
- Critério de Divisão ($split$) – Gini ou Entropia;
- N° máximo de características a avaliar em cada divisão (M_{feat})
- N° mínimo de amostras em cada nova folha (m_{leaf})
- N° mínimo de amostras num nodo interno (m_{node})

5.4.2 Balanceamento da Floresta

De acordo com a literatura, a fase de treino de uma Floresta não deve ser indiferente à falta de proporcionalidade das amostras de cada uma das classes a classificar. A falta de proporcionalidade pode implicar uma diminuição no desempenho do classificador supervisionado. Dada uma classe, com uma quantidade de dados reduzida, a precisão do classificador é diminuída em detrimento da classe com maior quantidade de dados [37].

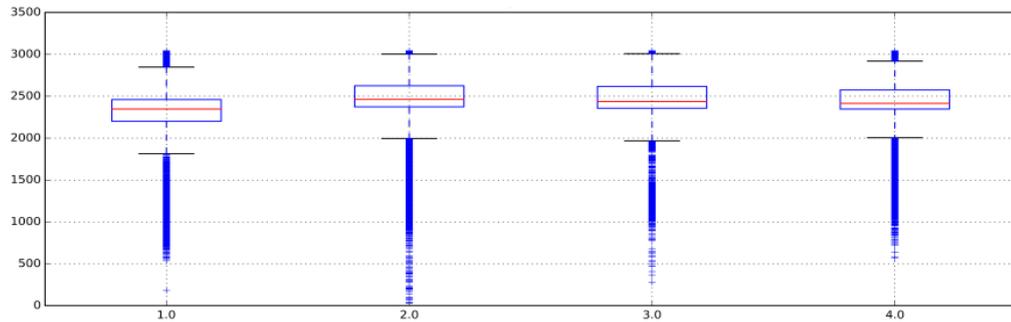
No contexto da segmentação cerebral tumoral, os dados utilizados para o treino da metodologia proposta apresentam a particularidade previamente descrita. Em HGG constata-se que a região tumoral corresponde a cerca de 9,57% de todas as amostras de tecido cerebral saudável. No caso de LGG dado que as regiões tumorais são menores estas correspondem apenas a 5,6% de todas as amostras de tecido cerebral saudável. Deste modo, a fase de treino do classificador utiliza todas as amostras de tecido tumoral, de todo o conjunto de pacientes de treino definido, e ainda um igual número de amostras de *background* (tecido cerebral saudável). Na amostragem do tecido pertencente à classe *background* foi ainda definido um critério para assegurar que amostras fronteira, isto é, próximas da borda do tumor seriam também analisadas pelo classificador. Com este procedimento procura-se assegurar um balanceamento entre as amostras das duas grandes classes de tecido cerebral e ainda assegurar alguma imunidade relativa ao tamanho variável que os tumores podem apresentar. Na fase de teste, o classificador recebe todas as amostras das quatro sequências de ressonância magnética, para cada um dos pacientes presentes na fase de teste.

Em HGG 58,8% das amostras, de tecido tumoral, dizem respeito a edema, 11,3% a necrose, 7,4% a região ativa não contrastante e 22,5% a região ativa contrastante. Retira-se que o edema é uma classe

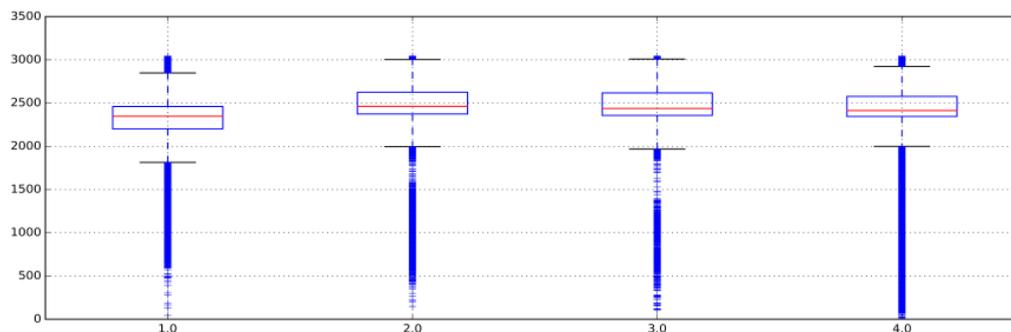
de expressão considerável, ao par que a a região ativa não contrastante apresenta-se com uma classe com poucas amostras. Contudo, em LGG a proporção das amostras de tecido tumoral é ainda mais discrepante, onde 2,7% das amostras de tumor dizem respeito a amostras de necrose, 51,5% a edema, 40,9% a região ativa não contrastante e 4,9% a região ativa contrastante. Neste último tipo de glioma são contabilizados tumores em estágios iniciais pelo que, contrariamente aos gliomas de elevado grau, existem tecidos característicos de uma fase tumoral mais avançada, que não estão presentes em alguns gliomas de baixo grau, sendo eles: região ativa contrastante e não contrastante e a necrose. Uma agravante a este problema de desproporcionalidade entre amostras, em LGG, é que se tratam de tecidos tumorais com relevo médico para quantificação da estrutura tumoral, avaliação e planeamento médicos, tal como descrito em 2.3. Daí que apenas se procurou combater este efeito neste tipo de glioma.

Várias hipóteses surgiram para combater a desproporcionalidade entre as amostras de cada classe. Porém, a proposta implementada resulta de uma normalização cruzada, isto é, normalizar HGG mas com os percentis e as marcas das intensidades definidos para LGG. Após a normalização cruzada e posterior extração de características, são então recolhidas amostras de necrose e tecido ativo contrastante, com recurso a um processo de amostragem.

A amostragem assenta na definição de filtros interquartil sobre as intensidades das sequências de ressonância magnética que têm maior grau de discriminação — a T1c e a T2. A designação deste processo é fundamentada pelo facto de os limites das intensidades, de cada um dos filtros, serem definidos com base no estudo realizado sobre os quartis das intensidades de cada uma das classes, ao longo dos pacientes de conjunto *Training*. Assim com a aplicação simultânea dos filtros a T1c e T2 procura-se seleccionar amostras com um comportamento transversal. Nas figuras 5.5 e 5.6 encontram-se os diagramas de caixas de cada uma das classes de tecido tumoral em HGG (após a aplicação da normalização cruzada) e LGG. As amostras resultantes do processo de filtragem são seguidamente introduzidas conjuntamente com as amostras de LGG para o treino do classificador.



(a)



(b)

Figura 5.5: Quartis das intensidades de cada uma das quatro classes de tecido tumoral (1- necrose; 2 - edema; 3 - tecido ativo não contrastante; 4 - tecido ativo contrastante) de HGG, nas sequências mais relevantes T1c (a) e T2 (b), após normalização cruzada. A linha a vermelho corresponde à mediana das intensidades da classe, ao par que os limites superior e inferior (de cada caixa) designam o quartil 75 % e 25% respetivamente. Os *outliers* são marcados com pontos em formato '+

Na tabela 5.3 encontram-se definidos os filtros interquartil aplicados sobre as intensidades de T1c e T2, heurísticamente. Convém clarificar que após a aplicação deste processo, as amostras resultantes contêm as características de todas as quatro sequências da RM.

Tabela 5.3: Limites de valores das intensidades, em T1c e T2, para a filtragem de amostras de necrose e tecido ativo contrastante, em pacientes HGG com a normalização de LGG.

	Necrose		Região ativa contrastante	
	$I_{\text{mínimo}}$	$I_{\text{máximo}}$	$I_{\text{mínimo}}$	$I_{\text{máximo}}$
T1c	450	1150	1650	2750
T2	1500	1700	1100	1650

A normalização cruzada seguida de amostragem procura assim evitar a utilização de procedimentos mais complexos para a criação de amostras sintéticas e ainda a realização de um treino conjunto com HGG, que exige mais recursos computacionais e tempos de computação mais elevados.

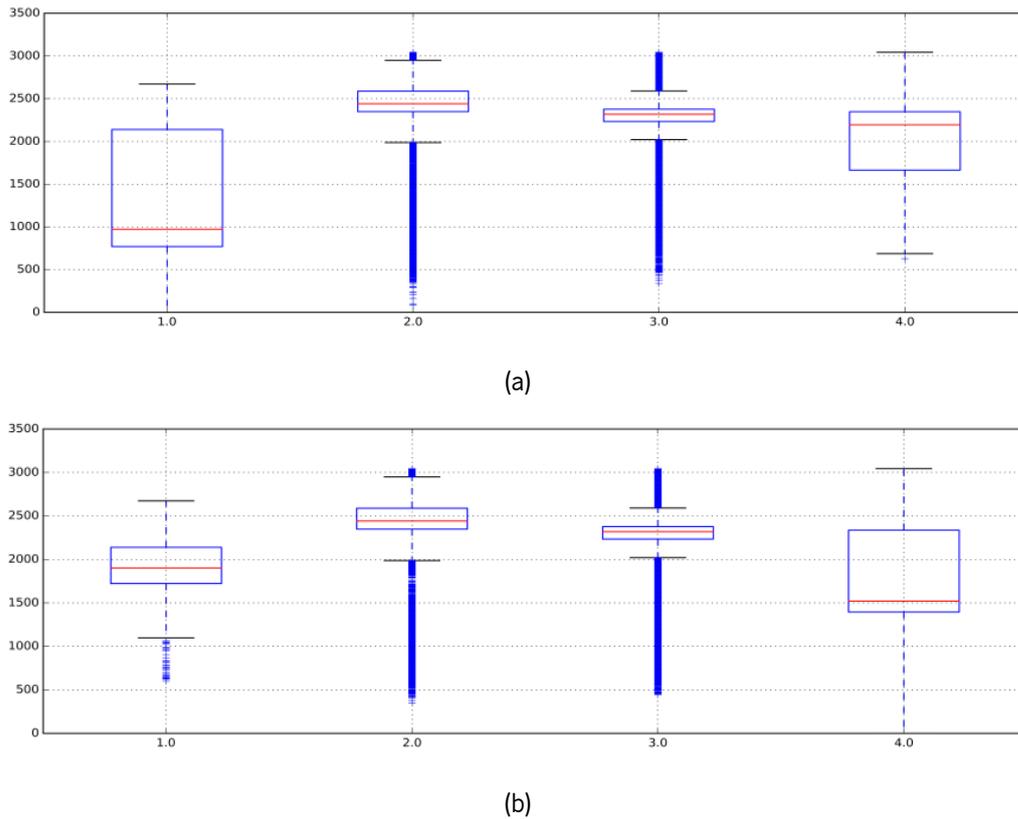


Figura 5.6: Quartis das intensidades de cada uma das quatro classes de tecido tumoral (1- necrose; 2 - edema; 3 - tecido ativo não contrastante; 4 - tecido ativo contrastante) de LGG, nas sequências mais relevantes T1c (a) e T2 (b), após normalização. A linha a vermelho corresponde à mediana das intensidades da classe, ao par que os limites superior e inferior (de cada caixa) designam o quartil 75 % e 25% respetivamente. Os *outliers* são marcados com pontos em formato '+'.
+'

5.4.3 Classificação em Dois Estágios

A capacidade de identificação e delimitação da região tumoral, como um todo, é um fator preponderante para posterior classificação das classes de tecido tumoral que o constituem. Assim pensou-se no desenvolvimento de uma abordagem distinta para a segmentação dos tumores cerebrais. Numa primeira fase procura-se definir as regiões tumorais presentes e, posteriormente, com base na informação do primeiro estágio, classificar as regiões identificadas nas 4 classes de tecido tumoral existentes. Assim sendo, tem-se uma linha hierárquica de classificadores.

Com a aplicação de uma classificação em dois estágios procura-se também modelar a percepção de um especialista aquando da realização da segmentação manual. Numa primeira fase, tem-se como objetivo delimitar a região tumoral e posteriormente, com base na informação do primeiro estágio, classificar a região identificada como tumor nas 4 classes de tecido tumoral. A esta abordagem atribui-se a designação de HEGS - *Hierarchical Extra-Trees for Glioma Segmentation*. Na figura 5.7 encontra-se a estrutura do algoritmo implementado.

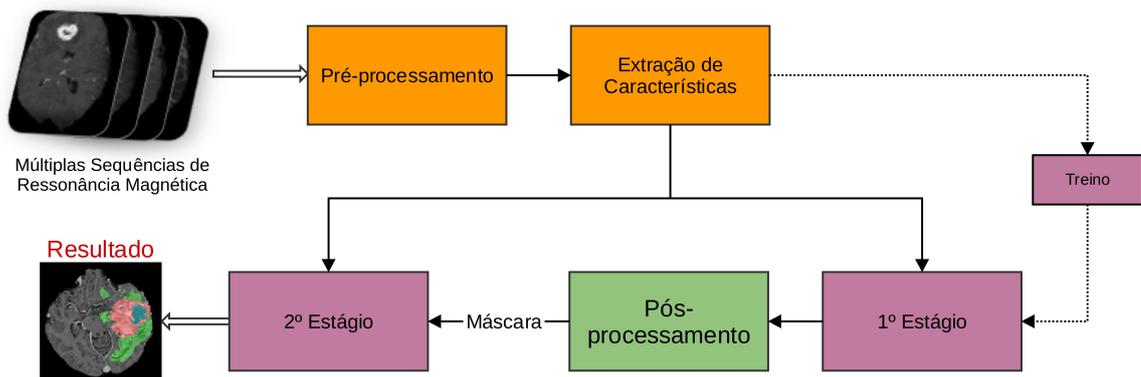


Figura 5.7: Estruturação do algoritmo de segmentação tumoral em dois estágios.

Em conformidade com a figura 5.7, o primeiro classificador realiza uma classificação binária, distinguindo o tecido tumoral do tecido cerebral saudável. Neste estágio apenas foram utilizadas três das quatro sequências, para a extração de características FLAIR, T1c e T2, visto que nestas existe uma melhor delimitação tumoral. Lembrando que todos os tipos de tecido tumoral, em FLAIR e T2, têm um valor de intensidades hiper-intenso, o agrupamento de todas as classes de tecido tumoral, numa única, tem como intuito aumentar o grau de confiança do classificador em regiões de fronteira e assim adquirir maior precisão na delimitação do tumor como uma entidade única. Com a introdução da sequência T1c tem-se como objetivo enfatizar as regiões de transição de tecido tumoral e tecido cerebral saudável que são definidas pelo tecido ativo contrastante.

Após a aplicação do pós-processamento, intervém um segundo classificador responsável por discriminar as quatro classes de tecido tumoral. Para tal, as segmentações resultantes do primeiro estágio definem a região sobre a qual é realizada a nova etapa de classificação.

5.5 Pós-processamento

Independentemente do nível do tumor a segmentar, HGG ou LGG, a presença de um tumor cerebral consiste numa estrutura de volume considerável, isto é, um grande componente ligado.

Deste modo, o pós-processamento aplica uma filtragem morfológica de modo a remover pequenas marcações isoladas e ainda pequenos componentes ligados. Ainda que raro, é possível que pacientes apresentem mais do que uma região de tumor cerebral. Por forma a não descartar esta hipótese, o pós-processamento foi alvo de alterações. Na segunda versão da etapa de pós-processamento é avaliado, após a eliminação de pequenos aglomerados, o formato das regiões restantes, classificadas como entidades tumorais. Nesta avaliação é caracterizado o formato de cada uma das regiões permanecendo apenas as que não apresentam um formato achatado, ou seja, uma variação reduzida do espaço de coordenadas tridimensional que caracteriza cada uma das estruturas. Na figura 5.8 encontra-se um exemplo prático da aplicação da segunda versão do pós-processamento.

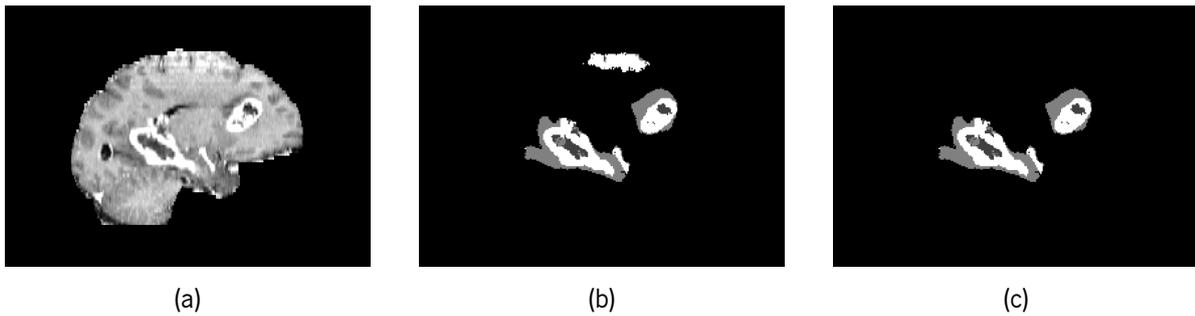


Figura 5.8: Aplicação do pós-processamento com remoção de regiões achatadas, em *HG_310*. (a) Sequência T1c normalizada; (b) Segmentação sem pós-processamento; (c) Segmentação com pós-processamento.

5.6 Sumário

No presente capítulo encontra-se exposta a estrutura do algoritmo proposto para a segmentação de tumores cerebrais. No pré-processamento, para a correção da falta da homogeneidade utilizou-se um pacote de ferramentas, *ITK-Advanced Normalization Tools (ANTs)* [128] que contém o método N4. Já a fase de normalização foi desenvolvida de raiz, em conformidade com os pressupostos de Nyúl *et al.* [39]. Encontram-se também os hiper-parâmetros utilizados por defeito pelos classificadores estudados. Em último lugar, foram apresentadas as características desenvolvidas para o problema de segmentação em estudo. Este conjunto procura uma delimitação precisa das regiões tumorais, mas também a discriminação do diversos tecidos tumorais e ao mesmo tempo assegurar uma regularização espacial. O conjunto de características utilizadas pode ser agrupado em três tipos: características locais, de contexto e de textura. Complementarmente, é também descrito o procedimento para a extração do conjunto de características utilizado.

Resultados

No presente capítulo, inicialmente, será descrito o método de avaliação dos resultados bem como o procedimento implementado para a realização dos testes. É ainda descrita a base de dados utilizada para a realização dos testes. De seguida, apresentar-se-ão os resultados obtidos dos diferentes componentes da abordagem proposta para a segmentação.

6.1 Base de Dados

As imagens de ressonância magnética públicas utilizadas foram preparadas no âmbito do desafio *NCI-MICCAI 2013 – Grand Challenge BraTS* sob a supervisão por K. Farahani, M. Reyes, B. Menze, E. Gerstner, J. Kirby e J. Kalpathy-Cramer. As imagens, totalmente anónimas, resultam de aquisições de RM de diferentes instituições (*ETH Zurich, University of Bern, University of Debrecen e University of Utah*), com diferentes equipamentos. Cada paciente tem todas as sequências corregistadas em relação à sequência T1c, realizando-se uma reamostragem para assegurar uma resolução de 1 mm isotrópico, para garantir um alinhamento espacial da informação.

Na base de dados existem três conjuntos distintos de imagens: *Training, Leaderboard e Challenge*, que correspondem às diferentes fases do desafio *NCI-MICCAI 2013 – Grand Challenge BraTS*. Na presente dissertação apenas se utilizou o conjunto *Training e Challenge*. O conjunto de *Training*, tem como intuito permitir, numa fase de desenvolvimento treinar o algoritmo de segmentação tumoral. Para tal, neste conjunto de dados, são disponibilizadas as segmentações manuais, realizadas por especialistas, de 20 pacientes com tumores de elevado grau (HGG) e 10 pacientes com tumores de baixo grau (LGG). Por outro lado, o conjunto de *Challenge* é constituído apenas por imagens de HGG (10) disponibilizadas especificamente com o intuito de assegurar uma avaliação imparcial do desempenho de cada um dos métodos, visto que as segmentações manuais não são públicas.

6.2 Avaliação do Desempenho

O desempenho de um algoritmo é avaliado de acordo com três métricas, *Dice*, Sensibilidade e PPV (*Positive Predicted Value*). A sensibilidade corresponde à probabilidade de alcançar um resultado positivo para as instâncias positivas, o que resulta numa medida que descreve a expressão dos falsos negativos na segmentação (expressão 6.1). Contrariamente, o PPV é uma medida que quantifica a expressão de falsos positivos numa segmentação (expressão 6.2). Em termos estatísticos, esta métrica, corresponde à probabilidade de um teste negativo para as instâncias negativas, ou seja, as amostras que não pertencem à classe ou classes em análise. [135].

$$\text{Sensibilidade} = \frac{TP}{TP + FN} \quad (6.1)$$

$$PPV = \frac{TP}{TP + FP} \quad (6.2)$$

Nas expressões anteriores, 6.1 e 6.2, TP, FP e FN designam a contagem dos verdadeiros positivos, falsos positivos e falsos negativos, respetivamente.

No âmbito da segmentação de tumores cerebrais, para uma dada classe, a métrica sensibilidade quantifica a expressão das amostras presentes no *ground truth* (segmentação manual) que foram também identificadas pela segmentação em análise. Assim sendo, na sensibilidade é quantificado o desempenho do algoritmo em identificar as regiões tumorais. Por outro lado, a métrica PPV permite quantificar a expressão de amostras presentes na segmentação manual que não foram identificadas pela segmentação em análise. Deste modo, nesta última métrica é quantificado o desempenho em identificar corretamente as amostras de uma classe [135].

Para além das duas métricas descritas, foi ainda utilizada a métrica de *Dice* (expressão 6.3). O *Dice* permite quantificar a sobreposição existente entre a segmentação final obtida e a segmentação manual. Por defeito, o *Dice* não é capaz de contabilizar a informação acerca do tipo de erro cometido na segmentação em análise, isto é, se a segmentação em causa é resultado de sub-segmentação (apenas uma pequena região é contabilizada pela classificação final) ou sobre-segmentação (a classificação final cobre toda a marcação manual). Deste modo, o *Dice* é comumente conjugado com outras métricas [135].

$$\text{Dice} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (6.3)$$

Este conjunto de métricas é utilizado na avaliação de algoritmos de segmentação por parte do *NCI-MICCAI 2013 – Grand Challenge BraTS*. Daí que, de modo a que seja possível a comparação com outras metodologias o valor de *Dice*, PPV e sensibilidade foram as métricas utilizadas, nesta dissertação. O cálculo das métricas foi realizado sobre as seguintes classes e regiões do tumor:

- Classe 1 – Necrose
- Classe 2 – Edema
- Classe 3 – Tecido ativo não contrastante
- Classe 4 – Tecido ativo contrastante

- Região Total - Conjunto de todas as classes de tecido tumoral: $(1 + 2 + 3 + 4)$
- Região Central – Porção central tumoral que engloba apenas as classes ativas e necrose, ou seja: $(1 + 3 + 4)$

Contudo, apenas as métricas relativas à região ativa contrastante, região total e região central serão avaliadas para discussão do desempenho das diferentes variantes do presente trabalho de dissertação. Deste modo, é possível realizar uma análise semelhante à realizada na plataforma online de *NCI-MICCAI 2013 – Grand Challenge BraTS*.

6.3 Metodologia de Testes

O modo de treino do classificador supervisionado variou mediante o tipo de conjunto de dados sobre o qual o método foi aplicado. No conjunto *Training* é utilizada uma estratégia de validação cruzada, para cada grau de tumor, dividindo-se a base de dados em grupos¹ de treino e teste. No caso do conjunto *Challenge* todas as imagens do respetivo nível tumoral a segmentar (HGG) foram utilizadas. No caso do conjunto *Training* foram realizados 10 grupos para 20 pacientes, em HGG, e 10 grupos para 10 pacientes de LGG. Assim sendo, para este conjunto, em HGG, cada grupo é constituído por 18 pacientes de treino e 2 pacientes de teste. No caso de LGG o grupo de treino contabiliza 9 pacientes permanecendo o restante para teste (metodologia *leave-one-out*).

O classificador supervisionado, independentemente dos diferentes tipos de classificador testados, partiu de uma configuração base cujos hiper-parâmetros são descritos na tabela 6.1. Os restantes hiper-parâmetros, não referidos na tabela, foram mantidos no seu valor por defeito.

Tabela 6.1: Definição dos hiper-parâmetros fulcrais do classificador aquando da realização dos diferentes testes.

Parâmetro	Atribuição
Número de estimadores	100
Profundidade da Floresta	40
Critério de divisão	Gini
Nº mínimo de amostras num nodo interno	2
Nº máximo de características a avaliar em cada divisão	$\sqrt{1}$
Nº mínimo de amostras em cada nova folha	1

Os testes foram todos realizados num processador Intel® Core™ i7-3930k 3.2 GHz (x12), com 47 GB de RAM, com o sistema operativo Ubuntu 14.04.

6.4 Pré-processamento

Estudos relativos às etapas prévias à classificação, nomeadamente os estudos do efeito de diferentes conjuntos de características e o impacto da normalização na segmentação de tumores cerebrais são apresentados nesta secção.

¹Do inglês: *folds*

6.4.1 Estudo das Características de Haralick

O primeiro estudo realizado para a avaliação do desempenho do algoritmo tem como objetivo compreender qual o impacto que as características de Haralick têm, na fase de classificação, quando associadas às características de primeira ordem. A extração deste conjunto de características não foi um processo imediato dada a dimensão inicial elevada das matrizes de coocorrência, necessárias para o cálculo das características de Haralick. Dado o elevado tempo de computação foi aplicado um reescalonamento das intensidades, descrito em 5.3.3, permitindo o cálculo das 13 características, em 4 direções. Em média, para cada paciente, o tempo de computação associado encontra-se na tabela 6.2.

Tabela 6.2: Tempo médio de computação, para cada paciente, das 13 características, ao longo das 4 direções, por cada sequência.

	Tempo (médio)
FLAIR	3 h 35 min
T1	4 h 05 min
T1c	4 h 20 min
T2	4 h 49 min

Com a redução dos tempos de computação, um estudo posterior foi realizado com o intuito de quantificar a capacidade discriminativa das características em questão, nas diferentes sequências da RM, dividido em 3 etapas:

- 1^a** - Avaliação da capacidade discriminativa das 13 características de Haralick num número reduzido de fatias do cérebro;
- 2^a** - Avaliação da capacidade discriminativa das características de Haralick mais discriminativas em todo o volume de cérebro; e
- 3^a** - Integração no classificador das características de Haralick.

A primeira e segunda fases procuram determinar se as 13 características de Haralick serão benéficas para o contexto do problema e definir, se possível, um conjunto menor de características capaz de discriminar cada um dos tipos de tecido tumoral de forma única.

No caso concreto da primeira fase de estudo apenas é considerada a informação de um único paciente de HGG, escolhido de forma a conter todas as classes de tecido tumoral. Para diminuir o tempo de extração das características, apenas são analisadas regiões que englobam o tumor cerebral e não todo o cérebro. Na primeira fase, a distinção entre tecido saudável e tecido tumoral não é alvo de análise, visto que não correspondem à motivação que despoletou a extração das características de Haralick.

Na extração das características de textura, a vizinhança utilizada para cálculo das matrizes de coocorrência foi uma vizinhança de 5×5 . Este procedimento apenas ocorreu na direção $d = 0^\circ$ indo de encontro com Bauer *et al.* [63], Islam *et al.* [70] e Zhang *et al.* [134] que defendem que a textura tem um maior grau de expressão ao longo do plano de simetria longitudinal.

Por forma a avaliar a capacidade discriminativa das características foi utilizada a distância de Bhattacharyya. A distância de Bhattacharyya é uma medida estatística que permite quantificar o grau de

similaridade entre duas distribuições probabilísticas (contínuas ou discretas). Assim sendo, é possível medir o nível de sobreposição entre duas populações e determinar o seu grau de proximidade. Na sua formulação mais simples, para amostras discretas, a distância de Bhattacharyya entre duas classes que seguem uma distribuição normal é obtida através da equação 6.4 [136].

$$D_B(p, q) = \frac{1}{4} \ln \left(\frac{1}{4} \left(\frac{\sigma_p^2}{\sigma_q^2} + \frac{\sigma_q^2}{\sigma_p^2} + 2 \right) \right) + \frac{1}{4} \left(\frac{(\mu_p - \mu_q)^2}{\sigma_p^2 + \sigma_q^2} \right) \quad (6.4)$$

$D_B(p, q)$ - corresponde à distância de Bhattacharyya entre duas distribuições p e q,

σ_p - variância da distribuição de p,

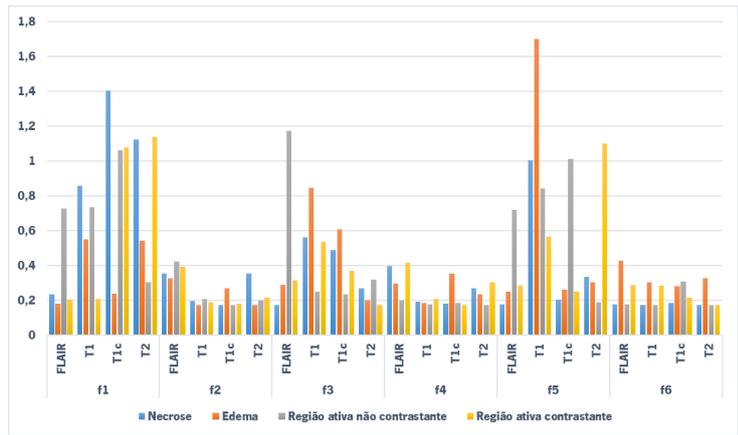
μ_p - média da distribuição de p,

σ_q - variância da distribuição de q,

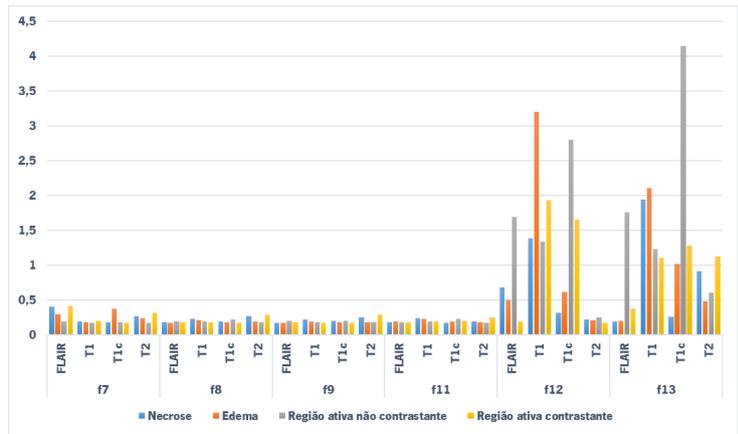
μ_q - média da distribuição de q, e

p, q - duas distribuições diferentes.

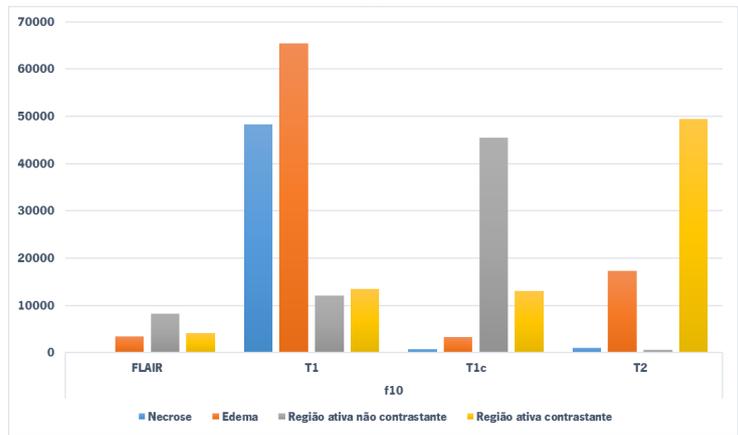
De seguida é então realizada a primeira fase do estudo das características de textura, onde na figura 6.1 se encontram as distâncias de Bhattacharyya das 13 características de Haralick.



(a)



(b)



(c)

Figura 6.1: Gráficos de barras da distância de Bhattacharyya das 13 características de textura de Haralick, em fatias das quatro sequências de RM de um único paciente de HGG, nos quatro tipos de tecido tumoral: necrose (azul), edema (laranja), região ativa não contrastante (cinzento) e região ativa contrastante (amarelo). Em (a) encontram-se as características f1, f2, f3, f4, f5 e f6 de Haralick; (b) compreende as características f7, f8, f9, f11, f12 e f13. Por questões de escala e visualização (c) apenas compreende a característica f10.

Pela análise da figura 6.1 constata-se que nem todas as características de Haralick permitem caracterizar de forma única cada classe de tecido tumoral. Tendo em análise os níveis de sobreposição quantificados pela distância de Bhattacharyya, as características f10, f12 e f13 apresentam níveis de

sobreposição comparativamente elevados (especialmente a primeira característica referida). Porém, é ainda possível averiguar, simultaneamente nas três sequências de maior relevo (FLAIR, T1c e T2), que na região ativa (classe 4) as características f5, f6 e f11 tem uma distância de Bhattacharyya de baixo valor. Apesar de a grande maioria das características de Haralick, na região ativa contrastante, apresentarem distâncias de Bhattacharyya de baixo valor, é em f5, f6 e f11 onde a sobreposição das três sequências combinadas tem menor expressão. As características f6 e f11 permitem ainda observar que em relação à necrose, esta apresenta baixos valores de sobreposição ao longo das 3 sequências supracitadas. Deste modo, apenas se optou pela computação de f5, f6 e f11, que correspondem às características de Haralick mais significativas, em concordância com o seu nível de sobreposição avaliado para as quatro classes de tecido tumoral.

No segundo estudo, após identificação das características de Haralick com poder mais discriminativo, f5, f6 e f11, estas foram extraídas para toda a região do cérebro em todos os pacientes do conjunto *Training*. Nesta fase, a sequência T1 não foi alvo de extração de características de textura. A justificação para tal procedimento reside no facto da distância de Bhattacharyya apresentar valores médios superiores aos das restantes sequências e portanto se considerar que é uma sequência de pouco relevo, no que diz respeito à extração de características de textura. Para além da sobreposição entre as diferentes classes de tumor, a capacidade do mesmo conjunto de caraterísticas discriminar o tecido tumoral do tecido saudável foi alvo de análise. Tratando-se de um estudo com um conjunto de dados mais vasto, a inclusão da região cerebral não pode de todo ser descartada pois, a região de tecido saudável que circunda a região tumoral é rica em textura. Na figura 6.2 tem-se a distância de Bhattacharyya para as três características de Haralick seleccionadas.

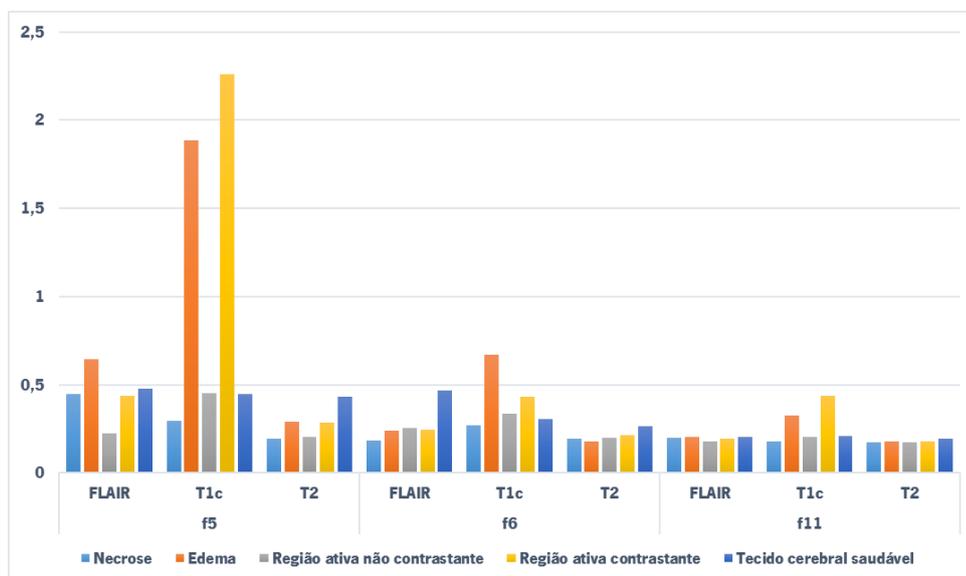


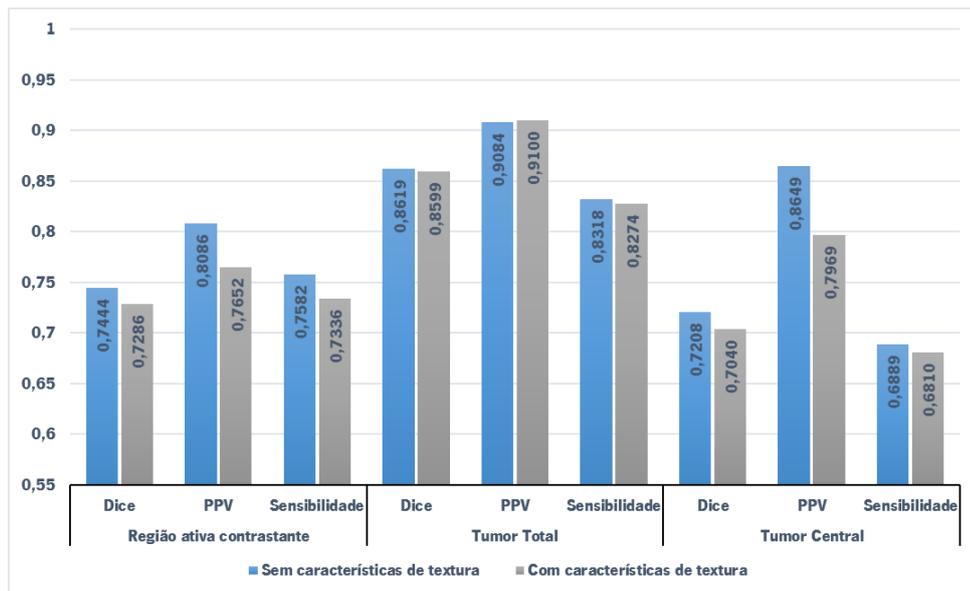
Figura 6.2: Gráfico de barras da distância de Bhattacharyya das três características de Haralick: f5, f6 e f11, considerando todo o volume cerebral de todos os pacientes do conjunto *Training* e todas as classes de tecido existentes: necrose (azul claro), edema (laranja), região ativa não contrastante (cinzento), região ativa contrastante (amarelo) e tecido cerebral saudável (azul escuro).

Comparando as medidas estatísticas da figura 6.2 com as obtidas para um único paciente de teste na figura 6.1 verifica-se que, os valores das distâncias de Bhattacharyya não apresentam discrepâncias assinaláveis. Esta última observação permite dar consistência ao teste realizado na primeira fase, dado que o paciente selecionado foi representativo da amostra do conjunto de dados disponível. Assim, as mesmas conclusões retiradas para f5, f6 e f11 na figura 6.1 são transversais. Focando as atenções na classe de tecido tumoral saudável, retira-se que nas sequências em FLAIR, T1c e T2, em f11, constata-se que o grau de sobreposição entre tecido saudável e tecido tumoral tem valores comparativamente menores. Não obstante, nas restantes duas características de Haralick o grau de sobreposição entre tecido saudável e tecido tumoral é de baixo valor. Posteriormente procedeu-se à integração, no conjunto de características de primeira ordem, de f5, f6 e f11 na fase de treino e teste do classificador.

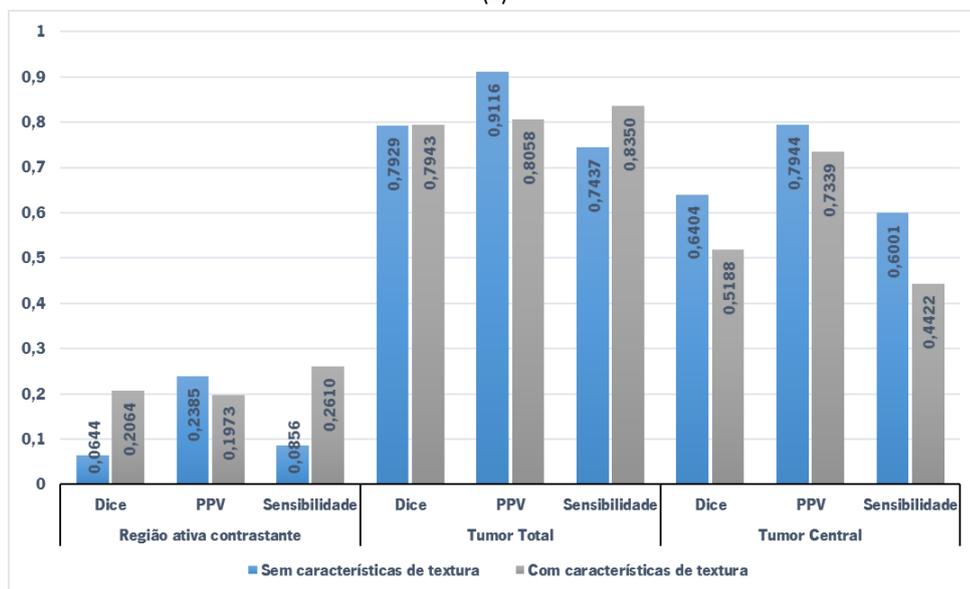
Na última fase de validação, em concordância com as fases anteriores, as características de Haralick f5, f6 e f11 foram incluídas no conjunto de características de primeira ordem descritas em 5.3.1 e 5.3.2. Por forma a quantificar a influência das características de Haralick, duas classificações distintas foram estudadas. Na primeira classificação as características de textura não se encontram presentes, ao passo que no segundo caso em estudo, as características de textura são incluídas na classificação. Em ambos os procedimentos utilizaram-se florestas de decisão, em conformidade com os parâmetros especificados na secção 6.3. Apesar de não se ter considerado a sequência T1 na segunda fase do estudo, nesta última fase de estudo todas as quatro sequências disponíveis foram utilizadas, para que o poder discriminativo do restante conjunto de características não seja afetado. Na figura 6.3 encontram-se os resultados obtidos pelos dois tipos de classificador.

Com base nos resultados presentes na figura 6.3 é possível averiguar que as características de textura influenciaram a divisão dos dados ao longo dos sucessivos nodos, resultando num treino do classificador e posterior fase de segmentação (teste) distintos. Esta afirmação é fundamentada pela diferença das métricas obtidas. Contudo, as características de textura introduzidas não são capazes de influenciar positivamente o desempenho do classificador em ambos os tipos de glioma. Analisando com mais detalhe as métricas presentes, em HGG, na ausência das características de Haralick o desempenho do classificador é superior. Através da métrica de Dice observa-se que a diferenciação das diferentes classes de tecido tumoral é superior, nas três métricas avaliadas. Destaca-se ainda o facto de na ausência de características de textura todas as métricas relativas à região ativa contrastante apresentarem um valor superior.

Em LGG, no geral, as características de Haralick têm um impacto positivo, especialmente numa classe de tecido tumoral difícil de segmentar neste tipo de gliomas, a região ativa contrastante. Verifica-se ainda que a delimitação do tumor completo tem melhor performance, aquando da inclusão das características de Haralick onde o valor de Dice é ligeiramente maior. Não obstante, a segmentação da região central do tumor apresenta menor qualidade, visto que a sensibilidade e o Dice são inferiores.



(a)



(b)

Figura 6.3: Gráfico de barras das médias das métricas obtidas relativas às segmentações provenientes de um classificador sem características de textura (cinzento) e um classificador que considera as características de textura associadas às características de primeira ordem, em HGG (a) e em LGG do conjunto *Training*.

No geral, conclui-se que a utilização de características de Haralick em metodologias de segmentação tumoral não acarretam a uma melhoria significativa dos resultados. Complementarmente, os tempos de computação destas características são ainda muito elevados o que levou a que as características de textura fossem excluídas do conjunto final de características. Neste momento, poder-se-ia questionar o porquê de não incluir também características que permitam caracterizar a forma de um tumor. Tal facto é contra-argumentado com a enorme variabilidade de forma que caracterizam as regiões tumorais, especialmente nas sub-regiões que constituem o glioma.

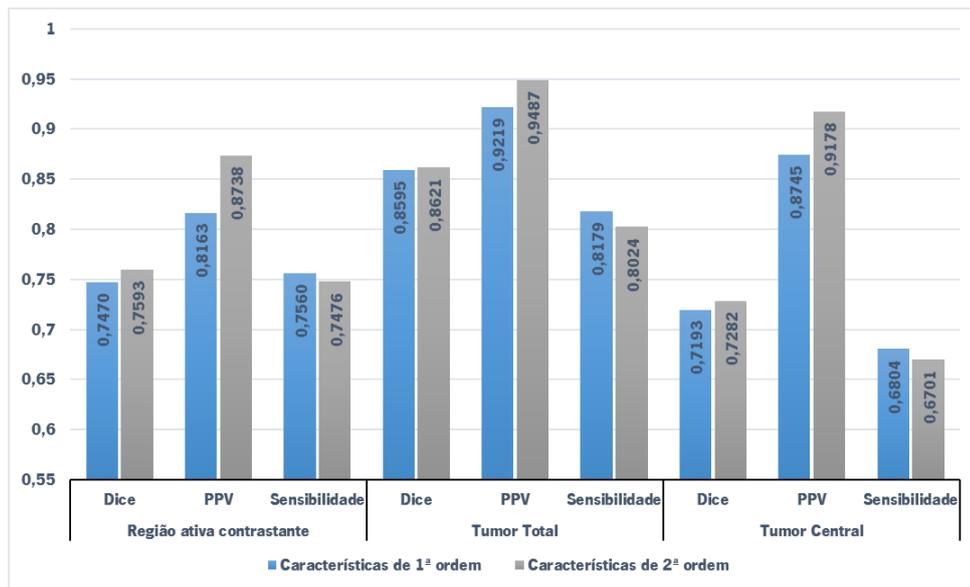
6.4.2 Efeito das Características

Com a realização de sucessivas avaliações ao algoritmo projetado, a capacidade discriminativa do conjunto de características de primeira ordem começou a ser posta em causa. A suspeita principal residia na baixa capacidade de discriminação introduzida pelo Gradiente Gaussiano, bem como as médias locais e respetivos contextos. No caso de características como a diferença de Gaussianas (DoG) para além de não introduzirem qualquer informação benéfica, suspeitou-se que este subconjunto de características afetava negativamente o desempenho do classificador. A identificação deste problema conduziu a uma nova fase de desenvolvimento de novas características. Com base nestas suposições surgiu um novo conjunto de características, as características de segunda ordem, descritas em 5.3.4. Na figura 6.4 encontram-se os resultados obtidos pela metodologia com o conjunto de características de primeira ordem e com o conjunto de característica de segunda ordem, em ambos os tipos de glioma.

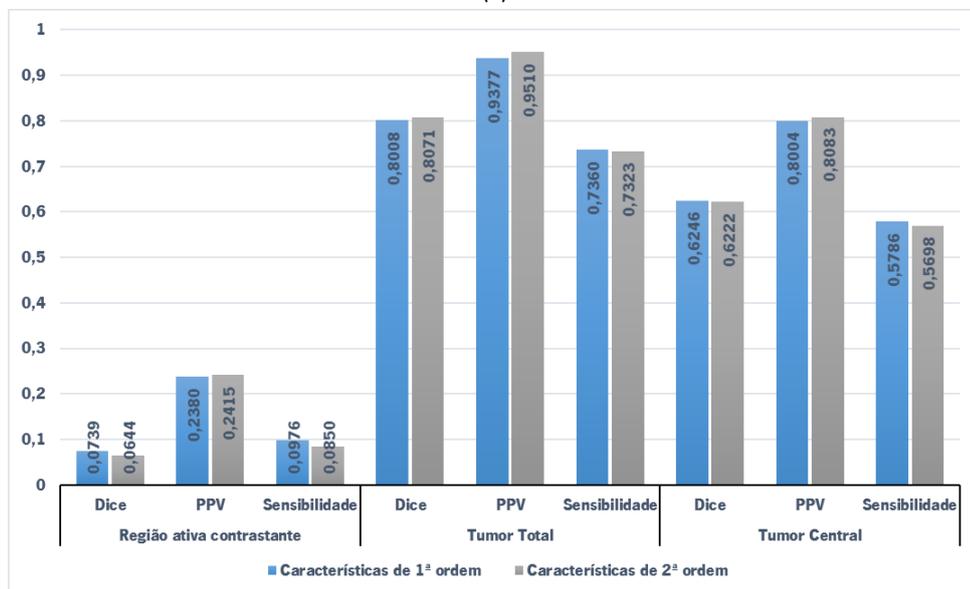
Pela análise da figura 6.4, observa-se que o as características de segunda ordem permitem obter um método mais robusto, visto que o desempenho é superior quando é aplicado este conjunto de características. A suspeita da baixa capacidade discriminativa por parte das características enunciadas anteriormente, pode ser explicada em grande parte devido à limitação que as intensidades da ressonância magnética impõe e não à incapacidade associada à própria característica, como se tinha em mente.

Através da análise das métricas obtidas é possível afirmar que a projeção do novo conjunto de características é mais discriminativo. A inclusão das características de segunda ordem resulta numa distinção superior entre as regiões fronteiriças do tumor e do tecido saudável resultando numa delineação do tumor mais precisa. Considerando apenas as três métricas da região Tumor Completo o seu desempenho é superior em 1,4%, em HGG, e 1,6%, em LGG, relativamente ao conjunto de características de primeira ordem. Em ambos os tipos de gliomas, nas características de segunda ordem, a métrica PPV na região Tumor Completo apresenta valores superiores, sinónimo de uma fraca expressão de falsos positivos. Ainda nesta classe, é possível averiguar que a sensibilidade do novo de conjunto de características projetado apresenta uma ligeira diminuição. Contudo, o referido aumento do PPV conduz a que no geral a métrica de Dice seja ligeiramente superior.

Focando a análise para a região Tumor Central é possível constatar um comportamento semelhante ao relatado para a região Tumor Completo. Nos gliomas de elevado grau esta região, apresenta um aumento considerável no PPV, produzindo um valor de Dice superior. Nesta região, as características de segunda ordem permitiram uma melhoria no desempenho de 4,2%, em HGG. Porém, no caso concreto de LGG, as características de segunda ordem, não foram capazes de assegurar uma melhoria na segmentação da região central tumoral. Através das três classes avaliadas observa-se que as características de segunda ordem têm um desempenho inferior em cerca de 0,6%. Em grande parte este fenómeno é explicado pela diminuição da sensibilidade da região ativa, em 1,3%, não acompanhada por um aumento no PPV, tal como se verificou em HGG.



(a)



(b)

Figura 6.4: Gráficos de barras das médias das métricas obtidas relativas às segmentações provenientes de um classificador com as características de primeira ordem (cinza) e um classificador com as características de segunda ordem (azul) em HGG (a) e em LGG, no conjunto *Training*.

Não obstante, considerando ambos os níveis de glioma, as características de segunda ordem asseguraram um incremento no desempenho do algoritmo em cerca de 10 %. Observando os resultados obtidos para cada classe e região de tecido avaliados, é ainda possível demarcar a existência de um compromisso. Verifica-se que, uma diminuição da capacidade em detetar *voxels* como pertencentes à classe tumor (diminuição da sensibilidade) acarreta, por vezes, um menor número de *voxels* que não pertencem a nenhuma das classes de tecido tumoral (falsos positivos), o que leva um aumento do PPV.

Em suma, esta nova proposta em que são aplicadas transformações não-lineares sobre as intensidades para posterior processo de extração de um segundo conjunto de características, provou ser uma abordagem positiva. Denote-se que, os tempos de computação das características não tiveram qualquer

peso na seleção do melhor conjunto. No conjunto de características de segunda ordem o espaço dimensional é superior ao inicial (de 37 para 95 características por sequência). Contudo, o incremento temporal medido não se considerou como relevante, sendo que em cada treino e teste do classificador se reportou um aumento médio de 6 minutos por cada grupo no conjunto *Training*.

6.5 Classificação

Dada a possibilidade de configuração de alguns dos hiper-parâmetros da classificação supervisionada, nesta secção é realizado um estudo sobre os hiper-parâmetros passíveis de configuração. Porém, previamente é comparado o desempenho dos dois tipos distintos de classificadores implementados, a Floresta de Decisão e a *Extra-Trees*. Por último, serão apresentadas as diferentes alterações à fase de treino e teste da classificação supervisionada, nomeadamente o efeito do balanceamento do classificador, estudo da relevância de T1 e, por último, a classificação em dois estágios.

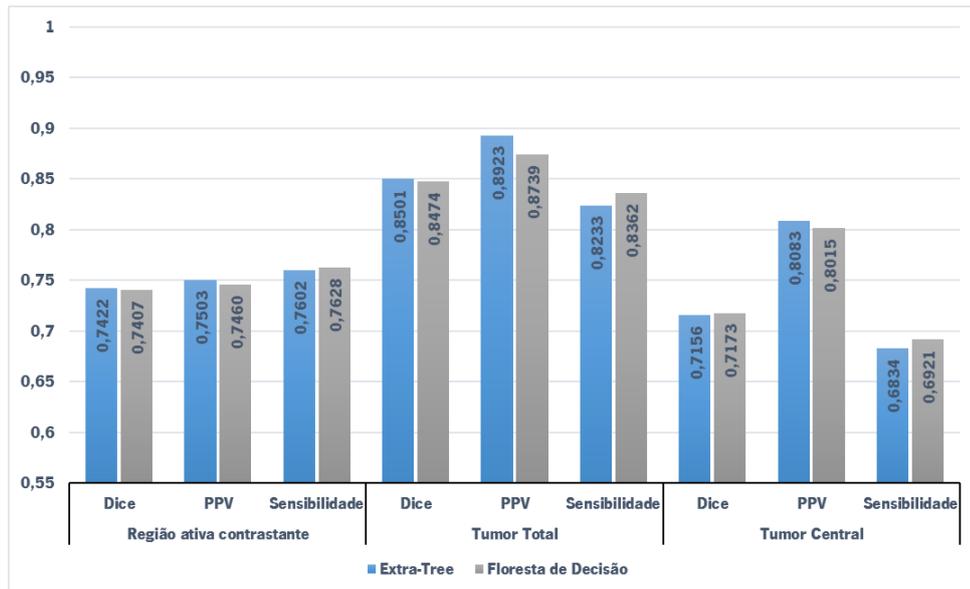
6.5.1 Efeito do Classificador

Relembrando os conceitos do classificador *Extra-Trees*, este permite assegurar um maior grau de aleatoriedade do que as Florestas de Decisão. Devido à sua aleatoriedade, no *Extra-Trees*, a seleção dos *thresholds* para a divisão das amostras pelos nodos filho é mais adequada, provocando uma redução da variância e consequentemente um aumento do desempenho.

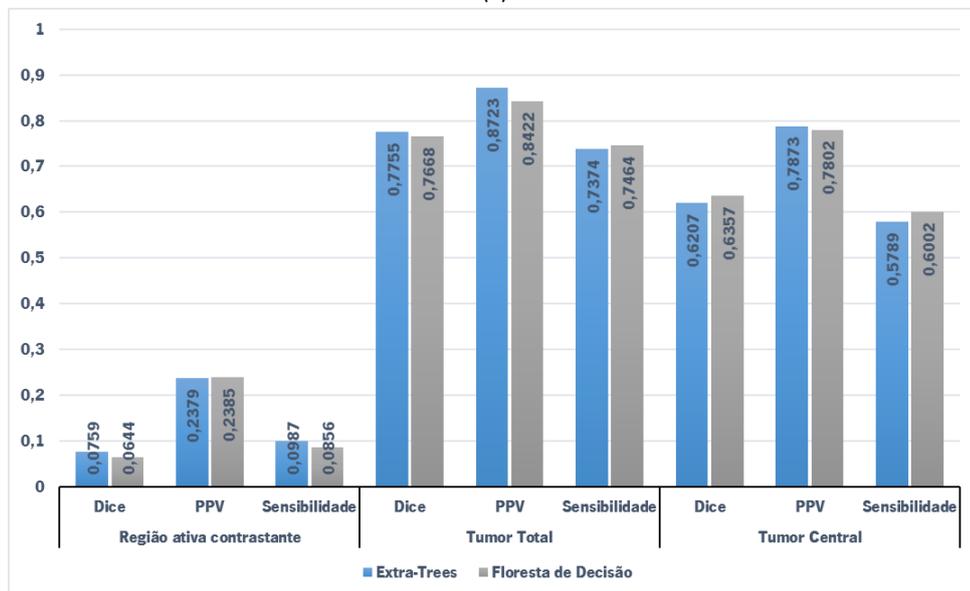
No *Extra-Trees*, em cada nodo, um subconjunto de características é escolhido aleatoriamente segundo uma distribuição uniforme e sem substituição, para avaliar a divisão dos dados pelos nodos filhos. Seguidamente, é determinado o *threshold* para cada uma das características, em conformidade com a sua gama, resultante de uma avaliação prévia de *thresholds* por parte do critério impuridade de Gini. De acordo com os seus autores, o classificador *Extra-Trees* é capaz de apresentar tempos de computação consideravelmente inferiores [123]. Assim sendo, procurou-se comparar o desempenho da Floresta de Decisão com o *Extra-Trees*. Os hiper-parâmetros de ambos os tipos de florestas foram idênticos, já descritos em 6.1, garantindo assim que apenas o desempenho da árvore seja a única variável de estudo. Para tal, o conjunto de características utilizadas, para este estudo, foi também idêntico para ambas as árvores, tendo sido utilizado o conjunto de características de primeira ordem. Na tabela 6.3 é possível observar os tempos de computação de ambas as modalidades em estudo. Por outro lado, na figura 6.5 encontram-se os resultados obtidos em ambos os tipos de glioma, sem aplicação do pós-processamento. Como se pretende comparar apenas a robustez dos dois classificadores, a realização da etapa de pós-processamento não é realizada. Ao excluir esta etapa procura-se também avaliar a expressão dos falsos negativos em cada uma das segmentações.

Tabela 6.3: Tempos de computação totais relativos ao conjunto *Training* após a realização de todos os *folders*.

	HG (Tempo total)	LG (Tempo Total)
Floresta de Decisão	32 h 08 min	6 h 08 min
<i>Extra-Trees</i>	10 h 41 min	1 h 44



(a)



(b)

Figura 6.5: Gráficos de barras das médias das métricas avaliadas, relativas as segmentações resultantes de uma *Extra-Trees* (azul) e de uma Floresta de Decisão (cinza), em HGG (a) e em LGG do conjunto *Training*.

Com base nos resultados obtidos constata-se que a utilização de Extra-Trees é a melhor abordagem, visto que a segmentação de tumores cerebrais é ligeiramente superior. O facto do Extra-Trees assegurar uma diminuição da variância permite que, na fase de testes, as probabilidades geradas, em cada *voxel* a segmentar, sejam mais confiantes e portanto produzam segmentações mais regulares. Sem a realização do pós-processamento a quantidade de marcações isoladas e pequenos aglomerados identificados como tumor é consideravelmente superior, no caso da Floresta de Decisão. Esta diferença detetada deve-se, em grande parte, à modalidade de seleção dos *thresholds* que diferencia os classificadores. No contexto do problema, definir um subconjunto de características escolhido aleatoriamente segundo uma distribuição uniforme e sem substituição, permite também aumentar a robustez da metodologia de segmentação.

O comportamento supracitado é explicado pelo valor das métricas obtidas. Considerando todas as métricas do Tumor Completo o Extra-Trees, nos gliomas de baixo grau, têm um desempenho cerca de 3% superior relativamente ao classificador Florestas de Decisão. O aumento do desempenho é devido em grande parte a um aumento do PPV. Neste mesmo tipo de glioma a região ativa contrastante também apresenta resultados superiores. Comparando todas as métricas associadas à região ativa contrastante o desempenho do Extra-Trees é 2,4% superior ao da Floresta de Decisão. A superioridade é explicada, em grande parte, devido ao aumento da sensibilidade para a referida classe. Em HGG, constata-se o mesmo comportamento, isto é, o classificador Extra-Trees aparenta ter uma maior capacidade em lidar com falsas classificações na demarcação do Tumor Total. Porém, neste tipo de glioma ocorre um decréscimo da sensibilidade do Tumor Total, ou seja, a capacidade de deteção de *voxels* de tumor é menor, diminuindo o impacto constatado para o PPV em LGG. Deste modo, o desempenho do Tumor Completo no Extra-Trees é apenas 0,8% superior à Floresta de Decisão. Nas restantes regiões avaliadas, considerando todas as métricas, para os gliomas de elevado grau observa-se o mesmo fenómeno pelo que o desempenho no Extra-Trees não é da mesma ordem do obtido para os gliomas de baixo grau, sendo apenas 0,8% superior. Contudo, verifica-se que o Extra-Trees apresenta melhores desempenhos quando os problemas de classificação possuem um menor balanceamento dos dados, como é o caso dos gliomas de baixo grau em conformidade com o que foi descrito em 5.4.2. Este fenómeno pode também ser explicado pela grande variabilidade associada ao Extra-Trees.

Um outro aspeto relevante que faz do Extra-Trees um classificador de eleição ao longo deste trabalho é o tempo de computação. Em termos percentuais, o Extra-Trees requer cerca de 30% do tempo de uma Floresta de Decisão para a realização do treino e teste, ou seja, é três vezes mais rápido. Esta última vantagem é fundamental em problemas de Processamento de Imagem, especialmente na área de Segmentação Automática de Tumores. Menores tempos de computação permitem a realização de uma maior quantidade de iterações, o que permite um maior grau de ajuste da metodologia ao problema. A redução do tempo de computação permite ainda o acesso a competições e concursos que exigem resultados numa curta janela temporal.

Conclui-se assim que, a utilização do Extra-Trees para problemas de segmentação de tumores é, dentro dos classificadores supervisionados estudados, o mais indicado. Todos os posteriores resultados apresentados resultam da aplicação deste classificador supervisionado.

6.5.2 Estudo dos Híper-parâmetros da Extra-Trees

Com a finalidade de otimizar a metodologia ao contexto da segmentação cerebral tumoral, foi levado a cabo um estudo dos híper-parâmetros de um classificador. Este estudo apenas se realizou sobre a Extra-Trees.

A procura dos melhores híper-parâmetros de uma floresta é um estudo vasto e demorado que, nesta dissertação, foi efetuado por um algoritmo *Grid Search*. Com este estudo procura-se obter um novo conjunto de parâmetros, mais adequado, capaz de substituir a configuração definida na secção 6.3.

O algoritmo *Grid Search*, com base num conjunto pré-definido de parâmetros é responsável por determinar a melhor combinação capaz de obter o melhor desempenho, quantificado por um critério específico, a média das métricas no conjunto de validação². Os parâmetros alvo de análise, bem como a respetiva gama de análise, foram:

- $n \in \{100, 125, 150, 175, 200\}$
- $depth \in \{30, 40, 50, None\}$
- $split \in \{Gini, Entropy\}$
- $M_{feat} \in \{ 'sqrt', 'log_2', 0.05, 0.10, 0.15, 0.20, 0.25 \}$
- $m_{leaf} = [1, 11] \cap m_{leaf} \in \mathbb{N}$
- $m_{node} = [1, 11] \cap m_{node} \in \mathbb{N}$

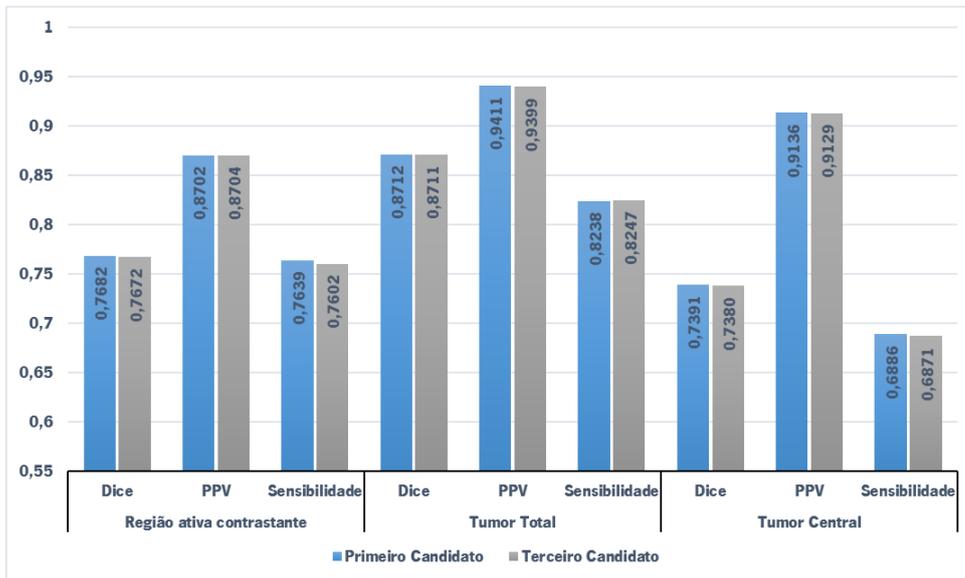
Como se pode constatar através da gama de valores permitidos para cada um dos parâmetros, o número de combinações possíveis é 33880. Dado este elevado número de combinações possíveis, apenas foram avaliados 100 candidatos, aleatoriamente escolhidos, de onde as três melhores configurações foram selecionadas, com base no critério supracitado. O procedimento demorou 4 dias 3 horas e 22 minutos, reforçando o porquê de este tipo de estudos ser temporalmente e computacionalmente limitativo. Os três melhores candidatos obtidos, de acordo com a média da validação, estão presentes na tabela 6.4.

Tabela 6.4: Três melhores combinações obtidas pelo *Grid Search*, num conjunto de 100 candidatos.

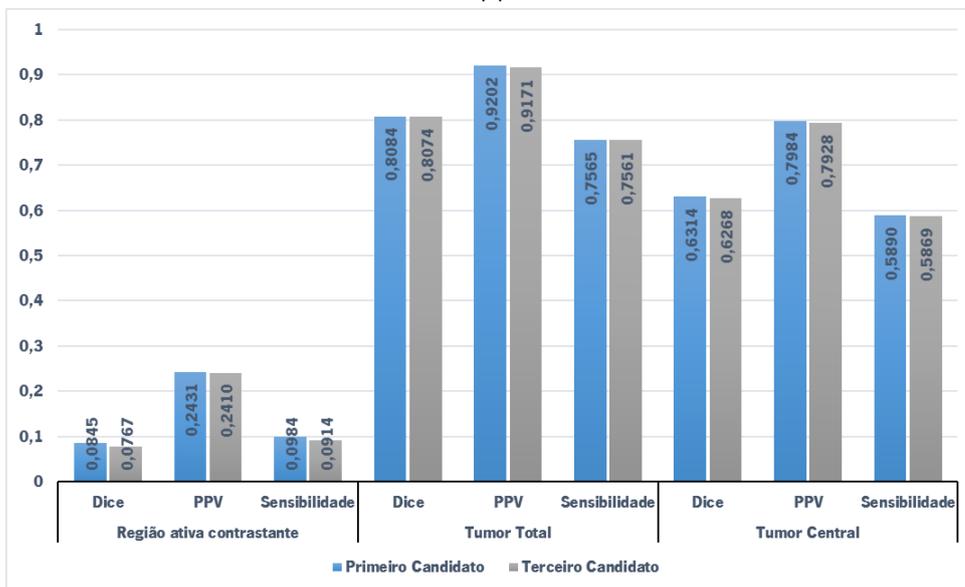
	Score	n	$depth$	$split$	M_{feat}	m_{leaf}	m_{node}
1º	0.982	125	30	Entropia	0.2	1	2
2º	0.982	100	None	Entropia	0.25	1	7
3º	0.981	100	50	Gini	0.25	2	6

²Do inglês: *mean validation score*

Após a obtenção dos três melhores conjuntos de parâmetros procedeu-se à implementação do primeiro e terceiro melhores candidatos, com o intuito de quantificar as alterações introduzidas por estes no desempenho da classificação. O terceiro candidato foi alvo de testes visto que, por vezes, o primeiro candidato pode apresentar *overfitting* ao conjunto de treino. Os resultados de ambas as abordagens encontram-se presentes na figura 6.6.



(a)



(b)

Figura 6.6: Gráfico de barras das médias das métricas avaliadas, das segmentações provenientes de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica *Grid Search* (azul) e um classificador com os parâmetros do terceiro classificado (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto *Training*.

Nos gliomas de elevado grau, numa primeira análise constata-se que ambas as configurações de parâmetros têm desempenhos bastante próximos. Neste tipo de glioma, nas três classes avaliadas, não existem diferenças assinaláveis no desempenho de ambas as configurações. Assim sendo, em HGG

não é possível afirmar qual das configurações de parâmetros da *Extra-Trees* é a mais adequada para a segmentação de tumores cerebrais.

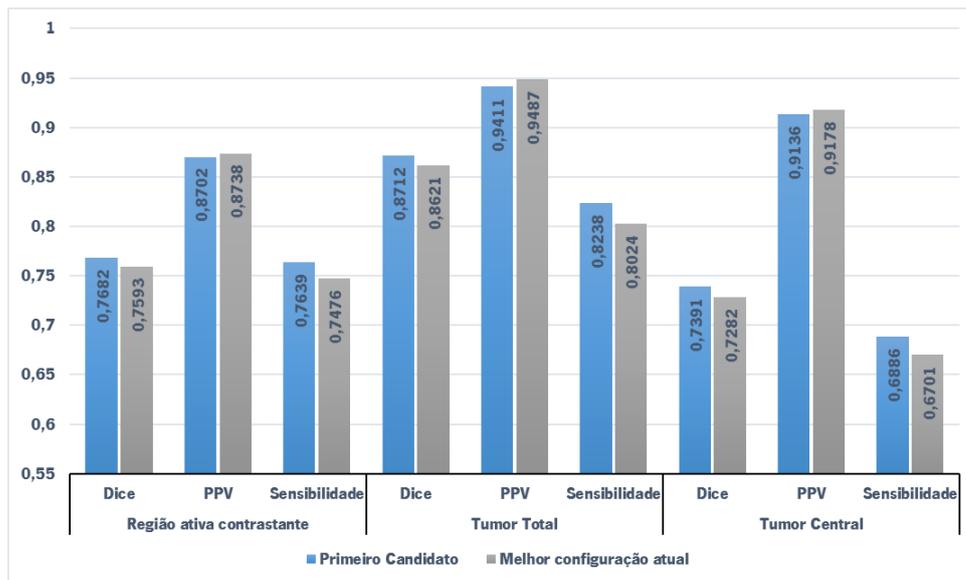
Focando a análise nos gliomas de baixo grau retira-se que o primeiro candidato é consideravelmente superior ao terceiro candidato. Avaliando o desempenho de todas as métricas, o primeiro candidato supera o terceiro por uma diferença de 3,4%. Esta diferença no desempenho é explicada essencialmente pela incapacidade de classificar corretamente a região ativa contrastante, onde o primeiro candidato tem um desempenho 1,6% superior ao terceiro candidato. Como este tipo de gliomas é um desafio para as metodologias de segmentação tumoral cerebral, o primeiro candidato é de facto a melhor das duas configurações analisadas.

Determinada a melhor das duas configurações de parâmetros avaliadas, torna-se necessária a comparação do primeiro candidato com os parâmetros implementados até ao presente momento, descritos na secção 6.3. Na figura 6.7 estão presentes os resultados obtidos nos dois tipos de gliomas a segmentar.

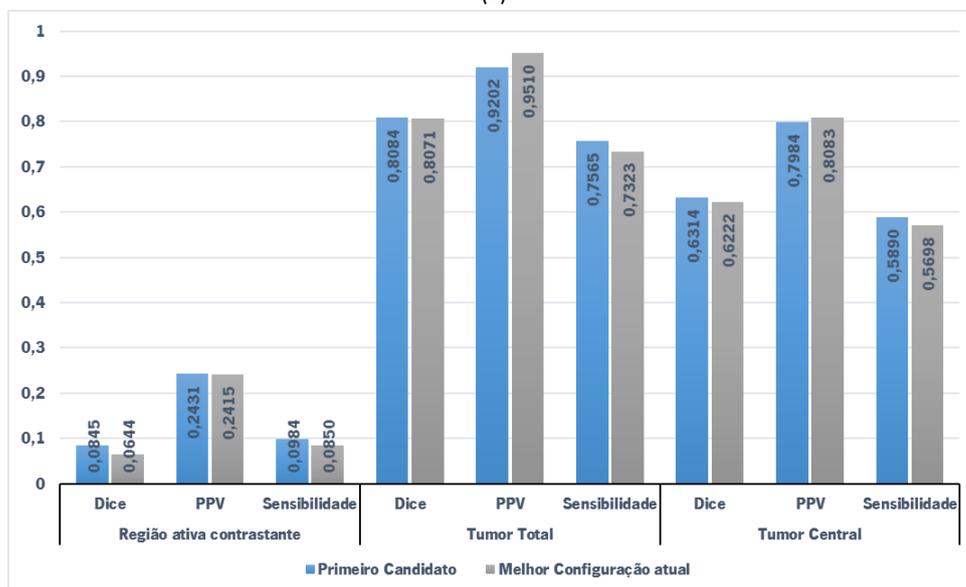
A nova configuração de parâmetros é capaz de assegurar um melhor desempenho da metodologia. No caso do HGG existe um aumento geral de 7%. Por outro lado, em LGG, o novo conjunto de parâmetros consegue apresentar uma melhoria geral no desempenho de 5%, considerando todo o conjunto de métricas. Uma particularidade detetada foi o facto de as métricas obtidas pela nova configuração apresentarem melhorias consistentes na sensibilidade e consequentemente em Dice. O valor do PPV em todas as classes, em ambos os tipos de glioma na nova configuração, não consegue superar a configuração de parâmetros, até ao presente momento utilizada. Este fenómeno pode ser explicado devido ao número de árvores e respetiva profundidade que definem o primeiro candidato. Comparando com os parâmetros até ao momento utilizados, o aumento do número de árvores conjuntamente com uma menor profundidade das árvores de decisão pode ser responsável por um maior ajuste do classificador aos dados e portanto uma maior capacidade em detetar pontos pertencentes à classe tumoral. Dado o aumento da confiança em identificar amostras como pertencentes aos diversos tecidos tumorais, traduzindo-se numa maior sensibilidade, existe associado um erro em considerar pontos de tecido cerebral saudável como tumor. Deste modo, pode-se assim justificar a pequena diminuição observada na métrica PPV.

Apesar desta melhoria significativa na qualidade das segmentações obtidas pelo novo conjunto de parâmetros, estes não foram aplicados. O porquê desta ação reside nos elevados tempos de computação associados a esta nova configuração do classificador. Utilizando o conjunto de hiper-parâmetros do primeiro candidato a metodologia requer 2 dias, 7 horas e 26 minutos (para todo o conjunto *Training*), ao passo que na configuração da secção 6.3 é de aproximadamente 13 horas. Os elevados tempos de computação associados à nova metodologia podem ser explicados pela dificuldade em otimizar a função de divisão dos nodos, em conformidade com o critério de otimização (entropia).

Sendo um dos principais objetivos da presente dissertação é o desenvolvimento de uma metodologia passível de ser implementada em contexto clínico, o elevado tempo de computação do primeiro candidato presente na tabela 6.4 é um entrave para assegurar um procedimento de segmentação rápido. Assim sendo, as configurações inicialmente propostas foram mantidas ao longo dos testes efetuados posteriormente, tendo sido mantidos os resultados da configuração inicial.



(a)



(b)

Figura 6.7: Gráfico de barras das médias das métricas avaliadas, de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica *Grid Search* (azul) e o classificador com os parâmetros aplicados até ao momento, na secção 6.3 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto *Training*.

6.5.3 Efeito da Amostragem

Após a aplicação das alterações anteriormente realizadas, o desempenho de LGG ainda não é competitivo e portanto não é o desejado. Através de uma análise das métricas avaliadas verifica-se que o classificador fica aquém do esperado para dois tipos de tecidos tumorais importantes: necrose e tecido ativo contrastante. Em LGG estes tecidos surgem em pacientes que apresentam um grau de diferenciação assinalável e/ou em fase de transição para um glioma de elevado grau. Dado o número reduzido de pacientes de LGG que contêm necrose e tecido ativo contrastante, estas classes tumorais são definidas por poucas amostras.

De forma a combater este problema foi estudado um procedimento de amostragem baseado na análise do histograma – filtros interquartil descritos em 5.4.2. Os novos resultados obtidos, bem como o melhor resultado prévio, encontram-se na figura 6.8.

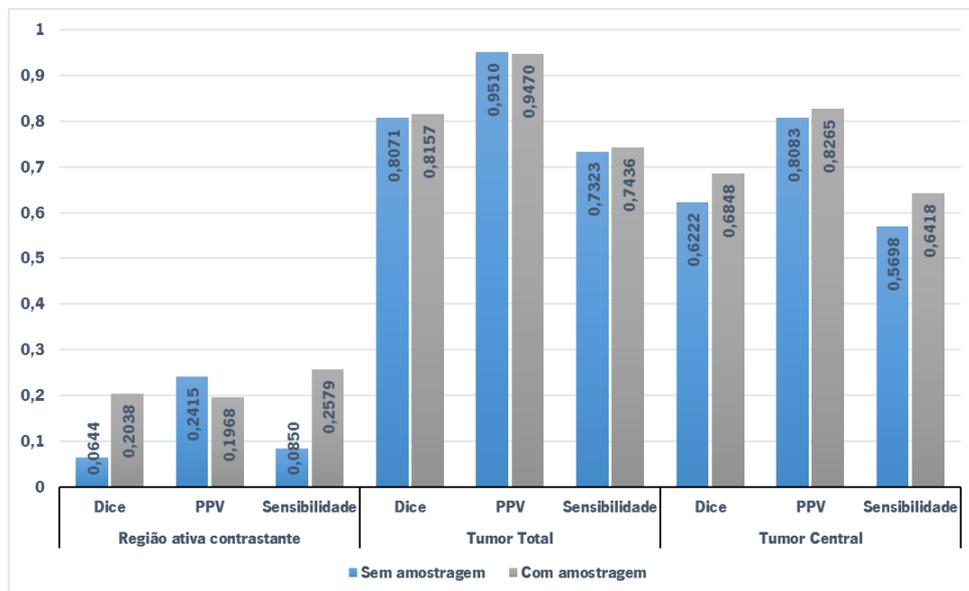


Figura 6.8: Gráfico de barras das médias das métricas avaliadas, de um classificador sem amostragem (azul) e o classificador balanceado através do procedimento de amostragem (cinzento), em LGG no conjunto *Training*.

O desempenho do classificador na segmentação de LGG melhorou consideravelmente com a inclusão de amostras específicas, com o intuito de balancear o treino do classificador. Realizando a diferença entre as três classes, combinando todas as métricas, é possível constatar um aumento no desempenho do classificador em cerca de 44%.

O balanceamento da floresta, aquando do seu treino, é um fator chave para alcançar segmentações mais robustas neste tipo de glioma. O procedimento de amostragem permite assim assegurar um classificador com uma maior capacidade em segmentar gliomas de baixo grau, que se encontram num estágio tumoral mais avançado e portanto sendo mais similar a um glioma de elevado grau. A seleção do tipo de amostras que são adicionadas, controlada pela aplicação do filtro interquartil, tem também impacto, não aumentando a variabilidade associada a cada classe de tecido a segmentar. A aplicação da amostragem, em LGG, é um procedimento importante, pelo que todos os testes posteriores realizados tiveram em consideração o procedimento em causa.

6.5.4 Estudo da Relevância de T1

No desenvolvimento de metodologias de segmentação automática cerebral, a presença da sequência T1 pode ser colocada em questão. Em termos de percepção visual a T1 não aparenta, por si só, ter boa distinção entre a região tumoral e a região cerebral saudável, bem como distinção entre os diferentes tecidos de tumor. Deste modo, foi colocado em causa o impacto desta sequência no treino/teste de um classificador supervisionado, isto é, se a T1 é capaz de fornecer características com capacidade discriminativa. Com base no melhor método alcançado até ao momento — Extra-Trees com características de segunda ordem e com um procedimento de amostragem em LGG — foi estudado o impacto da presença e ausência da sequência T1 e respetivas características na classificação. Na figura 6.9 encontram-se os resultados obtidos em ambos os tipos de glioma da melhor metodologia com e sem a sequência T1.

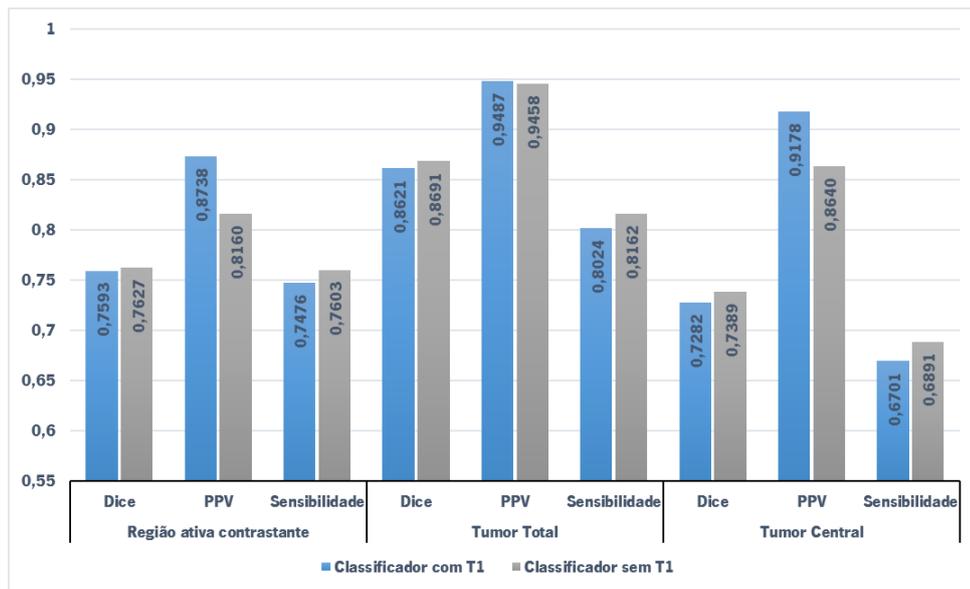
A ausência de uma sequência na fase de treino/teste do classificador permite uma redução dos recursos computacionais em cerca de 20%. Por si só esta propriedade é vantajosa, dado que permite uma redução dos requisitos mínimos necessários para a aplicação da metodologia.

Porém, ao avaliar as métricas expostas é possível detetar duas peculiaridades interessantes. Primeiramente, constata-se que em HGG na ausência da T1 o desempenho da metodologia é inferior em cerca de 5%. Em parte este resultado é explicado pela redução no desempenho da classificação da região ativa contrastante influenciando o desempenho do Tumor Central. No caso de LGG, assinala-se uma quebra considerável no desempenho do Tumor Central de 17%, contribuindo para que no geral a qualidade das segmentações sejam cerca de 15% inferiores. Torna-se possível concluir que as características extraídas a partir da sequência T1 têm impacto na discriminação do edema dos restantes tipos de tecido tumoral.

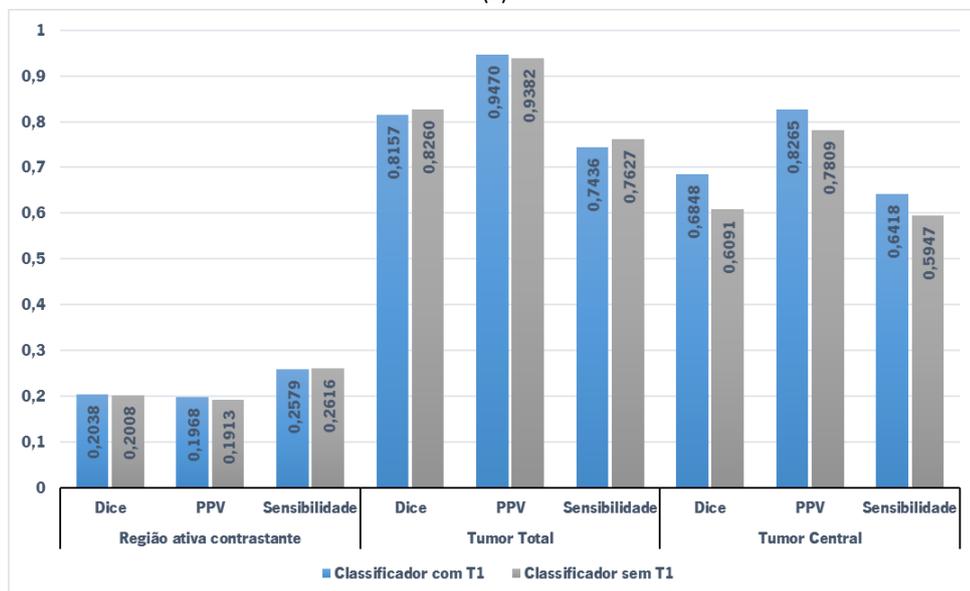
Contudo, em ambos os tipos de glioma verifica-se uma particularidade interessante. Na ausência da sequência T1 a metodologia é capaz apresentar uma melhor capacidade na delimitação do tumor total. Tanto em HGG como em LGG, a não utilização da T1 permite que a delimitação do tumor total seja melhor em 2%.

Este estudo permite fundamentar a partição do algoritmo em dois estágios. Com esta proposta procura-se assim tirar partido de uma melhor capacidade em delimitar o tumor como um todo, na ausência de T1, sem diminuir a capacidade em distinguir as diferentes classes de tecido tumoral, como se verificou na ausência da sequência em estudo.

Porém, previamente serão apresentados os novos resultados obtidos aquando da disponibilização de uma versão mais recente da base de dados utilizada, que permitiu a realização de uma nova etapa de normalização.



(a)



(b)

Figura 6.9: Gráfico de barras das médias das métricas avaliadas, de um classificador que contabiliza todas as sequências da RM disponíveis (azul) e um classificador sem a sequência T1 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto *Training*.

6.6 Nova Proposta de Normalização

O algoritmo de normalização é realizado tendo em conta o conjunto *Training* da base de dados BraTS 2013. Para cada tipo de glioma definiu-se uma normalização própria, ou seja, HGG e LGG são normalizados separadamente. Mais tarde, após a disponibilização da nova base de dados BraTS 2015, do conjunto de *Training*, foram definidos novos parâmetros de normalização. Esta última base de dados apresenta uma maior quantidade de pacientes portadores de glioma (220 pacientes HGG e 54 pacientes LGG) o que resulta numa definição dos parâmetros mais adequada ao contexto do problema.

Aquando da realização do estudo do impacto da sequência T1 na qualidade da segmentação, surge

a nova base de dados BraTS 2015 que contém uma maior quantidade de pacientes de treino portadores de glioma. Assim sendo, procedeu-se à realização de uma nova normalização dos pacientes da base de dados BraTS 2013, onde a fase de treino do algoritmo de normalização, para a definição dos percentis e posteriores marcas de intensidades, contabilizou a nova versão da base de dados. A alteração dos parâmetros de normalização despoletou nova extração de características e por consequência uma nova fase de classificação. Complementarmente, os novos parâmetros de normalização despoletaram a uma distribuição distinta dos valores das intensidades de cada classe de tecido cerebral a identificar. Deste modo, o procedimento de amostragem definido em 5.4.2 teve de ser reajustado. Na tabela 6.5 encontra-se a nova gama de valores, que definem o filtro interquartil. Posteriormente, é utilizada a melhor configuração atual em imagens com normalizações das intensidades distintas, por forma a estudar o impacto da nova proposta de normalização. Na figura 6.10 encontram-se as métricas obtidas em ambos os tipos de glioma.

Tabela 6.5: Limites de valores das intensidades, em T1c e T2, para a filtragem de amostras de necrose e tecido ativo contrastante, em pacientes HGG com a normalização de LGG, em conformidade com os novos parâmetros de normalização.

	Necrose		Região ativa contrastante	
	$I_{\text{mínimo}}$	$I_{\text{máximo}}$	$I_{\text{mínimo}}$	$I_{\text{máximo}}$
T1c	300	610	1100	1700
T2	1150	1800	1100	1650

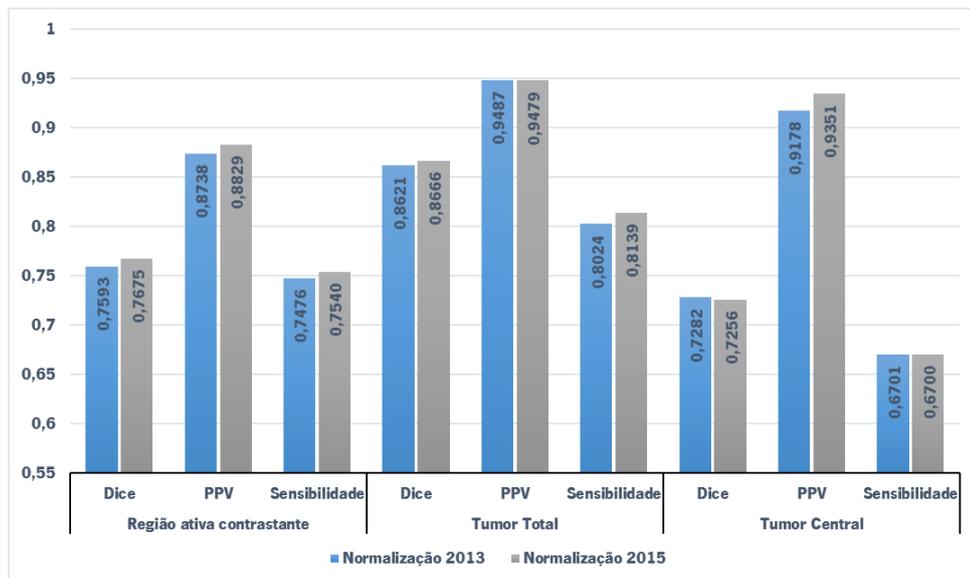
Com uma maior quantidade de dados de treino estima-se que, os novos parâmetros resultantes sejam capazes de assegurar uma melhor capacidade de generalização bem como uma delineação tumoral mais capaz.

Avaliando as métricas dispostas na figura 6.10 retira-se que a normalização assente num conjunto de dados mais vasto permite assegurar melhores resultados.

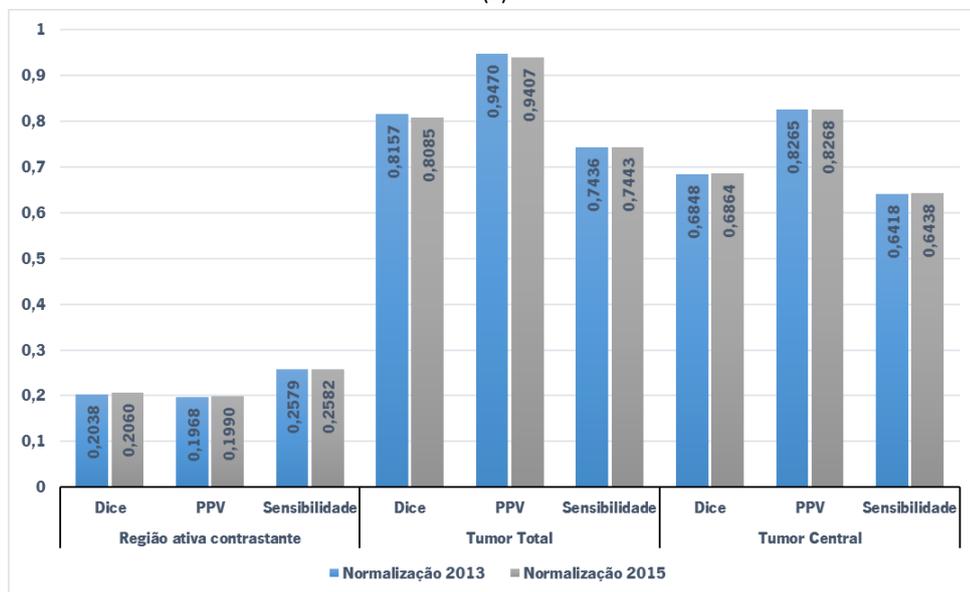
Em HGG verifica-se que, em todas as classes, a nova normalização é capaz de garantir melhorias no desempenho. Com a alteração da normalização obteve-se uma melhoria de 5,35%. A definição de percentis e marcas de intensidades, provenientes de um treino mais extenso, provocam melhorias significativas na métrica PPV. O Extra-Trees vê assim aumentada a sua capacidade em descartar uma maior quantidade de amostras de tecido cerebral saudável com propriedades semelhantes à do tecido tumoral, ou seja, amostras de tecido saudável semelhantes às de tecido tumoral da região tumoral.

No caso de LGG, com a nova proposta de normalização não se obteve uma melhoria idêntica à reportada para HGG. Neste tipo de glioma observou-se uma diminuição geral da classificação de 0,4%. Apesar de não ser uma diminuição no desempenho considerável, em grande parte esta diferença pode ser explicada pelo processo de amostragem associado a LGG. Tratando-se de um procedimento heurístico, a definição das gamas de intensidades, presentes na tabela 6.5, resultam de um estudo de iterativo de procura dos melhores parâmetros para um balanceamento correto do classificador.

Englobando ambos os tipos de glioma é possível averiguar que existe uma melhoria geral de 4,9% no desempenho da classificação, pelo que os novos parâmetros de normalização e o filtro interquartil definido, foram validados como a melhor abordagem atual.



(a)



(b)

Figura 6.10: Gráficos de barras das médias das métricas avaliadas com etapas de normalização das intensidades distintas: normalização com a base de dados BraTS2013 (azul) e normalização com a base de dados BraTS2015 (cinza); (a) compreende as métricas de HGG e (b) as métricas de LGG, no conjunto *Training*.

6.7 Classificação em Dois Estágios

Nesta nova proposta é definida uma hierarquia de classificadores (*Extra-Trees*). A filosofia proposta e estudada, primeiramente, tem como objetivo encontrar o tumor e seguidamente efetuar a distinção das diferentes classes de tumor. À abordagem em estudo atribui-se a designação de HEGS - *Hierarchical Extra-Trees for Glioma Segmentation*. Na figura 6.11 encontram-se os resultados obtidos, em HGG e LGG no conjunto *Training*, do classificador *Extra-Trees*, em que o treino é realizado no modo binário *versus* um *Extra-Trees* multi-classe. No classificador em modo binário apenas foram utilizadas três das

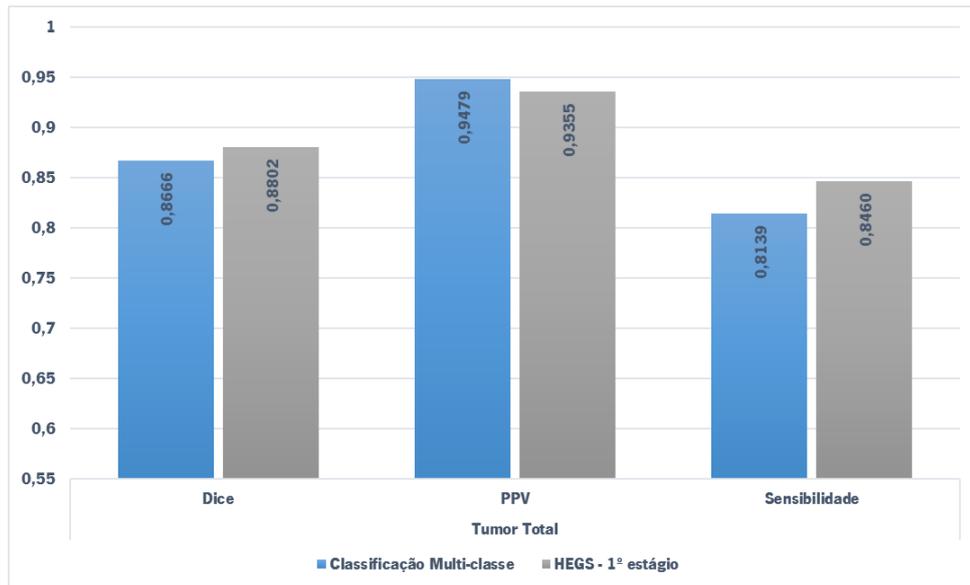
quatro sequências para a extração de características: FLAIR, T1c e T2, visto que na subsecção 6.5.4 se constatou que a exclusão de T1 permite uma melhor delimitação tumoral. Nas sequências T2 e FLAIR a gama das intensidades da grande maioria da região tumoral, já analisada na secção 2.4.2.1, tem um comportamento hiper-intenso, sendo o tumor a única estrutura capaz de originar intensidades de valor elevado em simultâneo na FLAIR e T2. Esta particularidade permite reforçar a presença de uma primeira fase onde o tumor é delimitado com maior confiança. Com a introdução da sequência T1c procura-se adicionar maior precisão na delimitação das regiões fronteiras quando estas são definidas pela região de ativo contrastante, que em T1c tem um comportamento hiper-intenso.

As segmentações obtidas foram avaliadas após a aplicação do pós-processamento, pois serão estes resultados que servem de máscara para o segundo estágio da segmentação. Os resultados encontram-se presentes na figura 6.11.

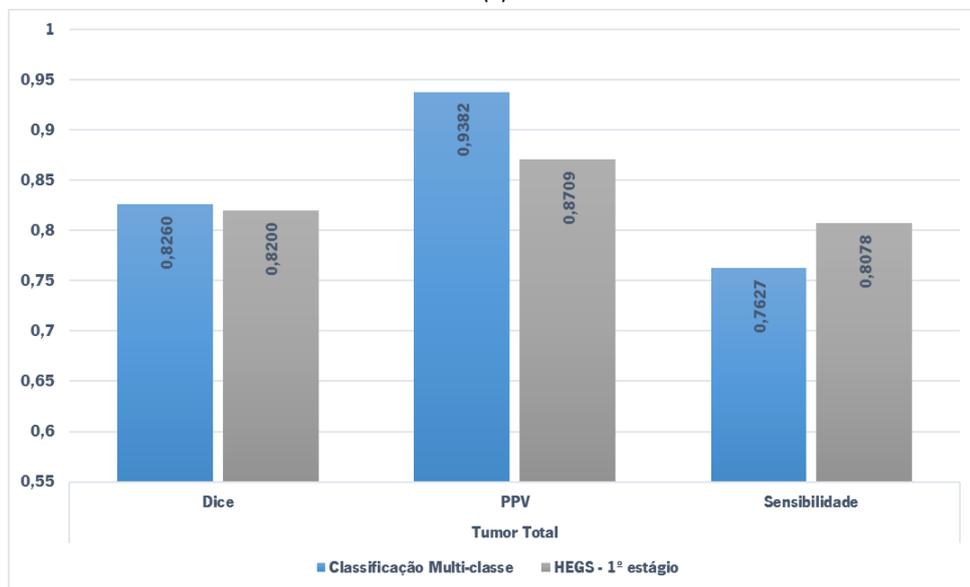
Denote-se que, nesta fase apenas se pretende averiguar a qualidade das metodologias com base na sua capacidade em definir a região tumoral total, pelo que apenas foi considerada esta região para o estudo. Pelas métricas obtidas em HGG, a classificação binária produz melhores resultados que a classificação multi-classe. No geral, no glioma de elevado grau constata-se uma melhoria no desempenho em cerca de 3,3%, em grande parte assegurada pelo aumento da sensibilidade da metodologia binária. Neste tipo de glioma, apesar de se averiguar uma descida no desempenho do PPV esta métrica tem ainda um desempenho elevado. No caso de LGG observa-se uma ligeira diminuição do desempenho de 0,7% na delimitação do tumor total. Contudo o desempenho das três métricas avaliadas é mais regular isto é a discrepância entre o valor de PPV e a sensibilidade não é tão elevado o que sugere que o método é mais equilibrado na classificação dos tecidos tumorais.

Na segunda fase do método proposto, a sequência T1 é introduzida. Nesta fase, partindo da região segmentada anteriormente, é treinado um Extra-Trees capaz de classificar unicamente as quatro classes de tumor cerebral, excluindo-se assim a classe de tecido cerebral saudável. Na figura 6.12 é comparado o desempenho, final, das métricas em todas as regiões com os resultados obtidos aquando da realização da nova normalização, num só estágio.

Em HGG, após a realização da segunda fase de classificação da metodologia HEGS, na presença da T1, é assinalável uma quebra no desempenho do algoritmo. Comparativamente com a classificação com um único estágio é possível verificar duas grandes desvantagens. A primeira é verificada para a região ativa contrastante, onde o desempenho do método desce em cerca de 5%. No tecido ativo contrastante, comparando as duas abordagens em estudo, existe um aumento na sensibilidade de 2%. Porém, a métrica PPV diminui em cerca de 7,4%, ou seja, a presença de falsos positivos na referida classe é maior. Deste modo, é possível constatar que o classificador apresenta uma sensibilidade aumentada para a atribuição de tecido ativo contrastante a pontos previamente identificados como tumor. Como consequência, o Tumor Central, em certa parte, apresenta uma diminuição do seu desempenho, pois o tecido ativo contrastante é integrante do Tumor Central. No Tumor Central, em termos comparativos, o desempenho desce cerca de 3%, em grande parte, uma vez mais, devido a uma diminuição acentuada no valor da métrica PPV (6,4%). Assim sendo, no geral, apesar de na primeira fase de HEGS se ter alcançado resultados positivos na delimitação do tumor total, os resultados provenientes da segunda fase demonstram uma diminuição do desempenho em cerca de 4%.



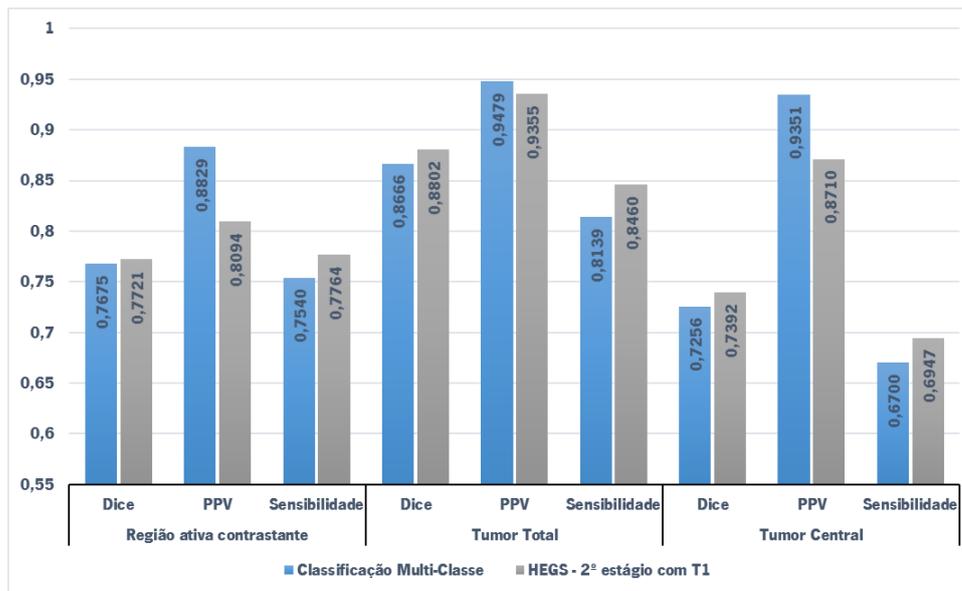
(a)



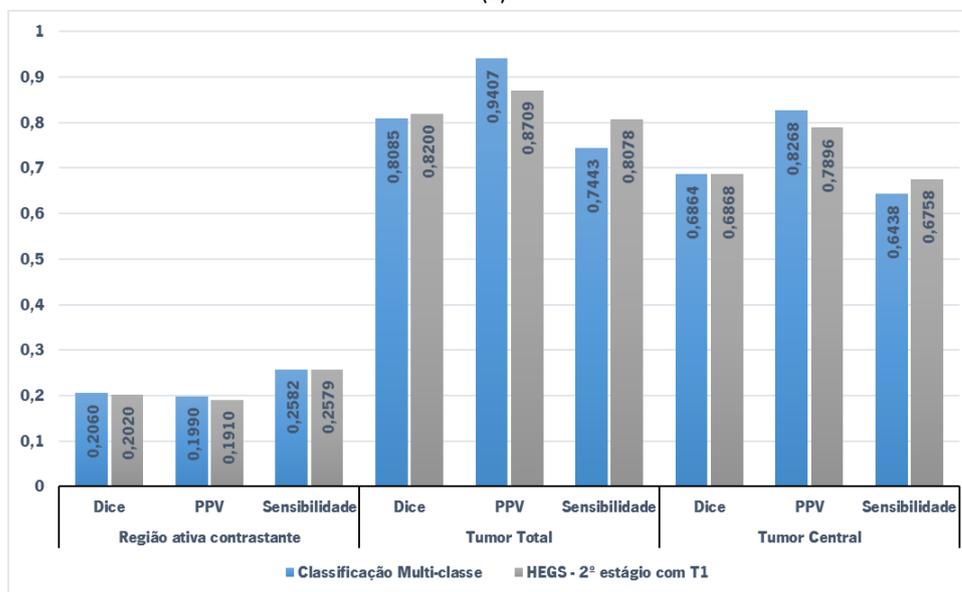
(b)

Figura 6.11: Gráficos de barras das médias das métricas da região tumoral total, relativas ao classificador multi-classe (azul) e ao primeiro estágio de HECS (cinza); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto *Training*.

No caso de LGG o desempenho da classificação no geral desce em 1,6%. Como anteriormente, o Tumor Total apresenta já um desempenho inferior seria de esperar que a região central e região ativa contrastante surgissem como classes afetadas por esse efeito, embora que mínimo. No caso do tecido ativo contrastante deteta-se uma queda no desempenho de 0,7% ao passo que na região tumoral central existe uma pequena quebra de 0,09%.



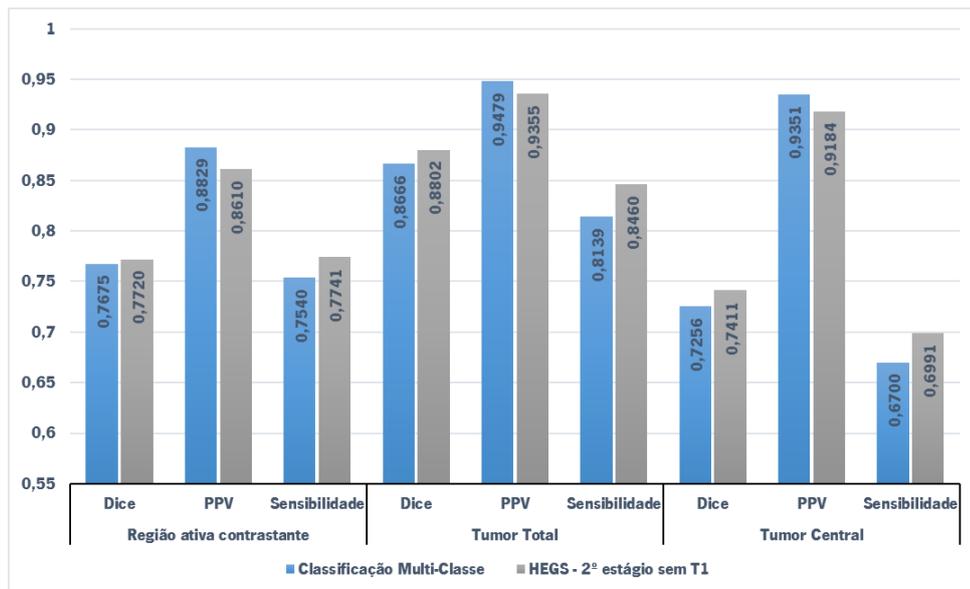
(a)



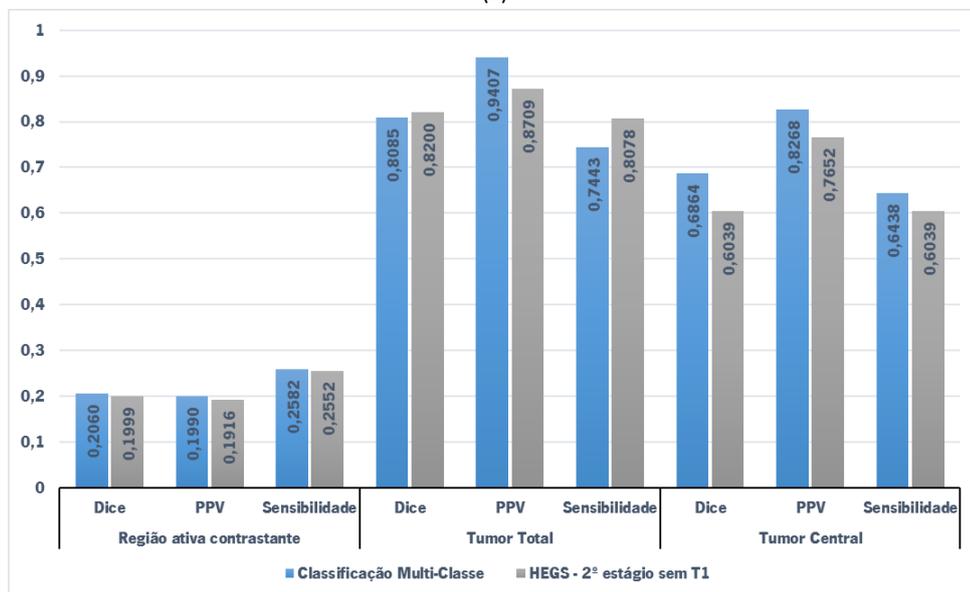
(b)

Figura 6.12: Gráficos de barras das médias das métricas de todas as regiões, relativas ao classificador multi-classe (azul) e ao segundo estágio de HEGS considerando a informação da sequência T1 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto *Training*.

Com base nos resultados obtidos na figura 6.12, procurou-se modificar a abordagem presente na segunda fase de classificação da metodologia HEGS. Apenas se colocou o foco na segunda fase pois a primeira foi capaz de apresentar melhorias, na definição de uma região de operação em fases mais avançadas. Assim sendo, o segundo estágio passou a contabilizar as três sequências idênticas às do primeiro estágio. Na figura 6.13 encontram-se os resultados obtidos da proposta de alteração da segunda etapa de classificação.



(a)



(b)

Figura 6.13: Gráficos de barras das médias das métricas de todas as regiões, relativas ao classificador multi-classe (azul) e ao segundo estágio de HEGS excluindo a sequência T1 (cinzento); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto *Training*.

Comparando as duas modalidades de segmentação aplicadas no segundo estágio, como se pode constatar pelas figuras 6.12 e 6.13 anteriores, existe uma discrepância nas conclusões a retirar para os dois tipos de glioma. No caso de HGG, a não inclusão da T1 permite um aumento no desempenho de 10,3%. Já no caso dos gliomas de baixo grau, a eliminação da informação contida nas características extraídas através da T1 provoca uma queda na qualidade da segmentação de 19,6%. Este fenómeno pode ser explicado pelo processo de normalização. As regiões tumorais de HGG correspondem a porções de volume considerável e consequentemente com um número de amostras elevado. Associada a esta característica tem-se o comportamento hipo-intenso e hiper-intenso que a necrose e tecido ativo contrastante devem apresentar num protocolo T1. Este comportamento leva a que as caudas do histograma

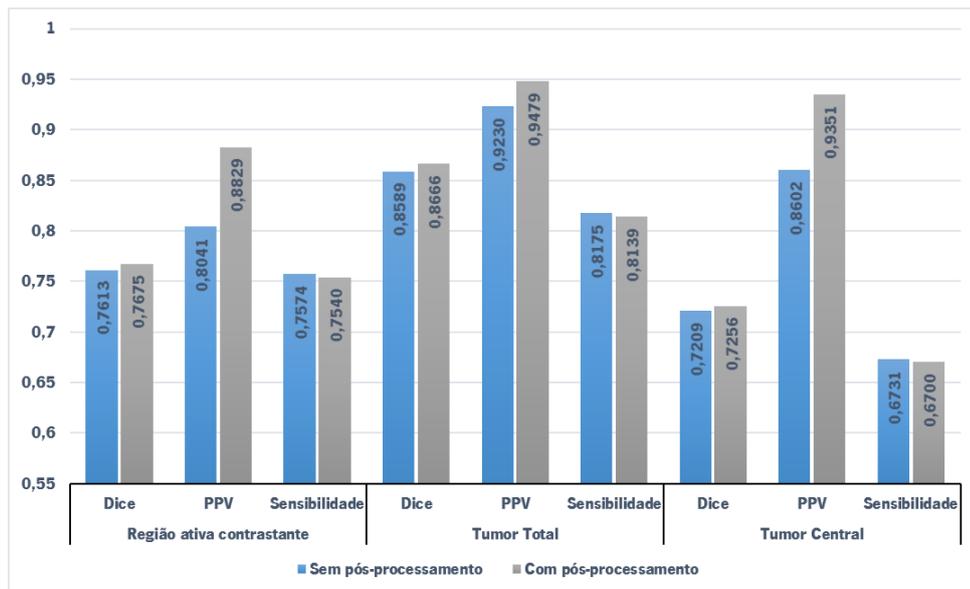
associado, eventualmente, contenham um número de amostras considerável de tecido tumoral de relevo para a segmentação. Com a aplicação da normalização, as amostras localizadas nas referidas regiões do histograma podem estar a ser alvo de um alargamento da gama da escala das intensidades nas caudas do histograma. O fenómeno de alargamento da gama de intensidades desencadeia uma gama de valores das respetivas características extraídas superior, dificultando a seleção do melhor *threshold* na separação dos dados na fase de treino do classificador e conseqüentemente uma diminuição da qualidade da segmentação. Como em LGG o número de amostras de tecido necrótico e tecido ativo contrastante é inferior, estima-se que a normalização deste tipo de glioma não sofre do fenómeno descrito e portanto a sequência T1 acarreta informação discriminativa relevante. Conclui-se assim que a metodologia HEGS passa a ser sensível ao tipo de glioma a segmentar, isto é, quando na presença de um glioma de elevado grau a sequência T1 é excluída, ao passo que em gliomas de baixo grau esta sequência é contabilizada para a classificação.

Em suma, a utilização de uma metodologia assente em dois estágios de classificação distintos é benéfica para a segmentação automática tumoral cerebral. Assim sendo, esta proposta representa a melhor metodologia desenvolvida no âmbito da presente dissertação.

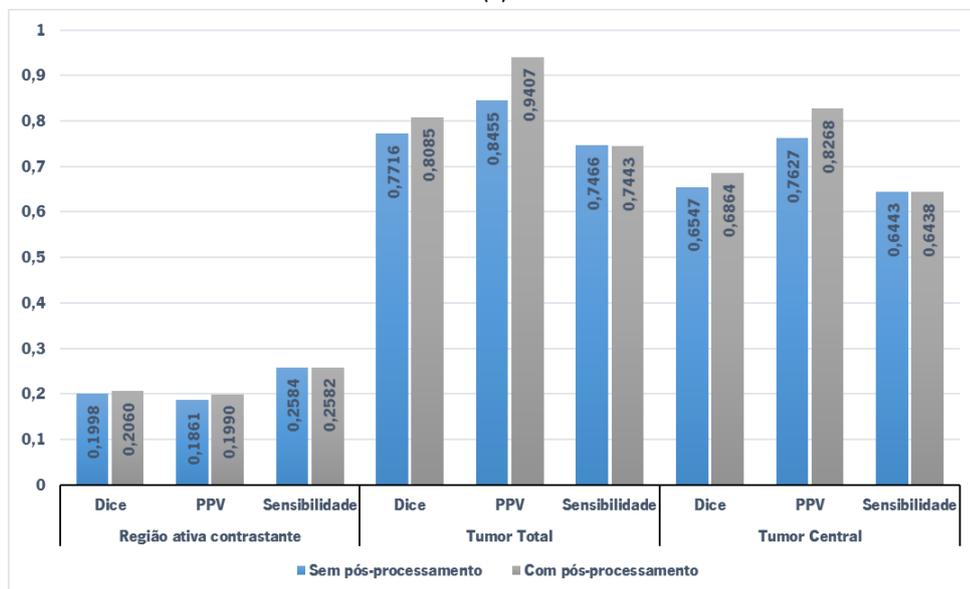
6.8 Pós-processamento

O pós-processamento utiliza-se com o intuito remover pequenos aglomerados, com recurso a morfologia matemática, partindo do pressuposto que os tumores cerebrais são definidos por regiões contínuas, ou seja, componentes ligados, de volume considerável. Assim, com a aplicação desta etapa procura-se eliminar pequenas regiões tridimensionais, diminuindo a expressão dos falsos positivos nas segmentações obtidas. O procedimento que permite eliminar pequenas marcações tridimensionais identificadas como tumor assenta na definição de um *threshold* sobre os componentes morfologicamente ligados que definem uma região. O *threshold* implementado é dinâmico, isto é, caso seja demasiado elevado para um determinado paciente em análise então o *threshold* passa a ser definido pela maior região presente. Tal procedimento é necessário, pois no caso de tumores de baixo grau as dimensões dos mesmos podem ser reduzidas, pelo que se optou pela definição de um *threshold* dinâmico. A figura 6.14 apresenta o valor médio das métricas em ambos os tumores de HGG e LGG, do conjunto de *Training*, no caso concreto da classificação com e sem pós-processamento da implementação apresentada em 6.6.

Os resultados permitem concluir que a presença de uma etapa posterior ao pós-processamento é útil para acrescentar robustez à metodologia de segmentação de tumores. Em HGG é assinalável um aumento no desempenho de 0,8% e em LGG um aumento de 7,2%.



(a)



(b)

Figura 6.14: Gráficos de barras das médias das métricas de todas as regiões, comparando as metodologias de segmentação sem a realização da etapa de pós-processamento (azul) e com a aplicação do pós-processamento (cinza); (a) compreende as métricas de HGG e (b) as métricas de LGG no conjunto *Training*.

Porém, quando da submissão de resultados para o conjunto *Challenge*, constatou-se que o pós-processamento assenta num pressuposto não válido, em que se assume que em cada paciente apenas existe um e um único volume de tumor cerebral. Através da análise das imagens de RM deste conjunto de testes, observou-se que um paciente de elevado grau (*HG_310*) apresenta duas lesões tumorais. Deste modo, a filosofia de pós-processamento teve de ser repensada. Para tal, desenvolveu-se uma técnica de pós-processamento em que são mantidos todos os volumes acima do mesmo *threshold* supracitado. Posteriormente é aplicado um procedimento de remoção de regiões achatadas. A remoção de regiões achatadas surge como uma filtragem auxiliar para eliminar regiões de volume considerável mas que

não apresentem variação espacial ao longo de um dos eixos tridimensionais. Geralmente, estas regiões surgem nos topos de uma imagem de ressonância do cérebro explicadas pela presença de ruído na aquisição da RM. Sendo que estas regiões são de comportamento hiper-intenso são entendidas pelo classificador como tecido tumoral.

A última versão do pós-processamento não acarreta qualquer impacto no conjunto de *Training*, pois não existem, neste conjunto de dados, pacientes com mais do que uma região tumoral. Para além deste pormenor, as regiões achatadas eram também removidas quando o pós-processamento procurava eliminar todas as regiões que não correspondessem ao maior volume segmentado. Apenas no conjunto de *Challenge* se constatou uma melhoria significativa na qualidade da segmentação do paciente *HG_310*. Na figura 6.15 é possível comparar o efeito da aplicação do novo método de pós-processamento, encontrando-se as duas regiões tumorais e uma terceira região achatada (de volume superior a uma das regiões tumorais) que é posteriormente removida com eficácia. Na figura 6.16 encontram-se os resultados, provenientes da plataforma *online VSD*.

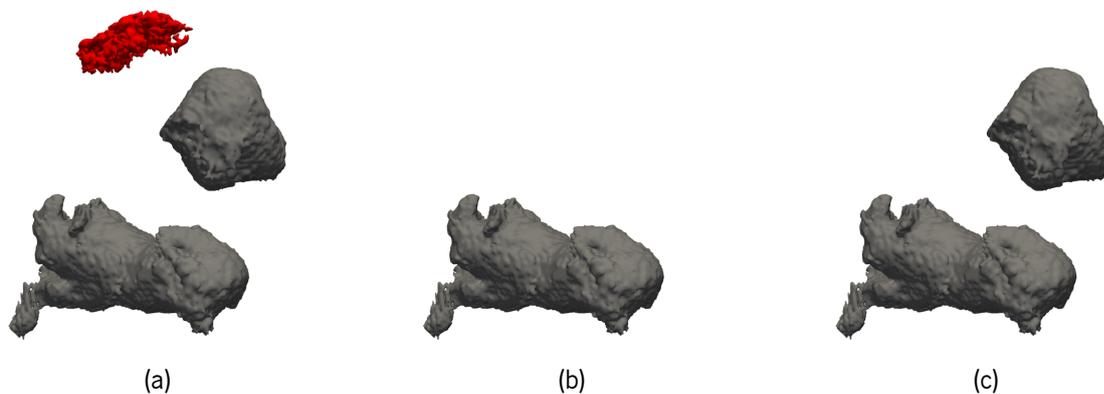


Figura 6.15: Resultados relativos ao paciente de *Challenge HG_310*, à saída do classificador, bem como os resultados provenientes da aplicação das duas versões de pós-processamento distintas propostas: (a) corresponde ao resultado sem aplicação de qualquer pós-processamento; (b) pós-processamento em que apenas é mantido o componente ligado de maiores dimensões; (c) pós-processamento com remoção de regiões achatadas e pequenas marcações. Em (a) a região a vermelho corresponde à região achatada a remover, ao passo que as regiões a cinzento correspondem a tumor cerebral.

Com a alteração do pós-processamento foi possível aumentar consideravelmente o desempenho da classificação em todas as métricas avaliadas. No paciente em análise, apenas se tem uma redução no PPV da região ativa contrastante, associada à introdução do segundo volume tumoral, que apresenta um volume tumoral ligeiramente superior à segmentação manual. Através dos resultados obtidos é possível concluir que a nova proposta de pós-processamento se encontra mais adequada ao contexto do problema, não descartando informação de relevo.

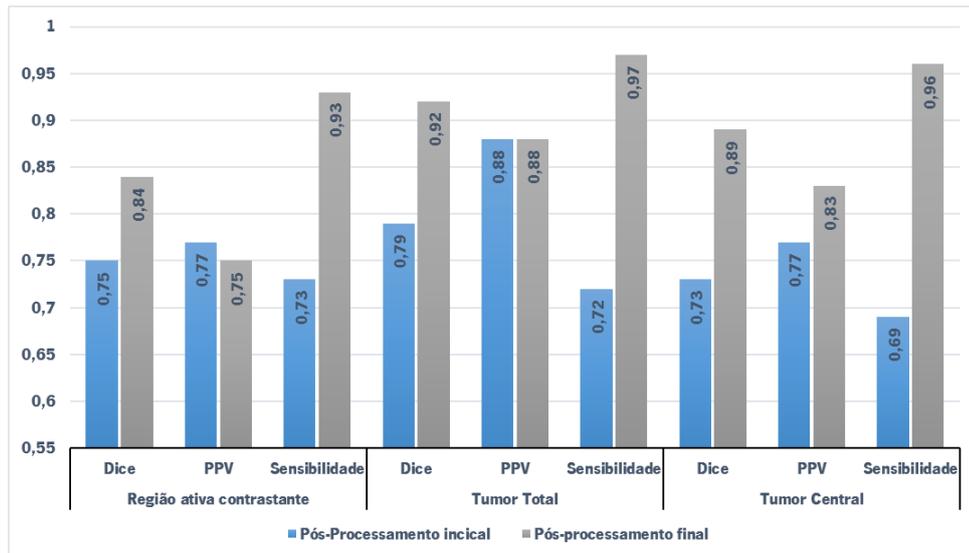


Figura 6.16: Gráfico de barras das médias das métricas de todas as regiões, relativas ao paciente 310 do conjunto *Challenge*, com a aplicação do pós-processamento sem remoção de regiões achatadas (azul) vs. pós-processamento com remoção de regiões achatadas (cinzento).

6.9 Tempos de computação

Os tempos de computação, relativos à melhor metodologia, encontram-se presentes na tabela 6.6 de acordo com as diferentes etapas chave que o constituem. Note-se que os resultados correspondem aos tempos de processamento médios por paciente, em segundos.

Tabela 6.6: Tempo médio de computação, em segundos, para efetuar a segmentação automática de um paciente.

Etapas de processamento	Tempo de computação médio (s/paciente)
Pré-Processamento	42
Extração das características	95,7
Classificação (Treino/Teste)	2200/27
Pós-Processamento	15

No presente momento, a extração de características e o treino/teste do classificador são as etapas mais críticas do trabalho. A extração das características é uma etapa que deve ser alvo de otimização, através da sua paralelização, com recurso aos vários *threads* que a máquina utilizada é capaz de fornecer.

6.10 NCI-MICCAI Grand Challenge BraTS 2013

Os resultados obtidos no conjunto *Training* motivaram a aplicação da metodologia aos restantes conjuntos existentes (*Challenge* e *Learderboard*), que constituem o concurso BraTS 2013, ainda aberto para submissões. Deste concurso, organizado durante a conferência MICCAI 2013, fazem parte equipas experientes e de renome no campo da Imagem Médica e *Machine Learning* (p.e. INRIA, França; Departamento de Radiologia, Universidade de Colúmbia, Nova Iorque; Instituto de Tecnologias Cirúrgicas e Biomecânica, Universidade de Berna; Laboratório de Visão Computorizada, Zurique, Suíça).

A avaliação do desempenho da metodologia proposta é efetuado por uma plataforma online VSD, através da análise das três métricas avaliadas ao longo desta dissertação.

Na figura 6.17 encontra-se a classificação obtida pelo conjunto *Training*, obtendo-se o quinto lugar. Denote-se que o desenvolvimento de um método capaz de apresentar resultados consistentes em ambos os tipos de glioma foi o que permitiu a classificação positiva. Em termos de métricas destaca-se o valor competitivo PPV obtido, significando que a quantidade de falsos positivos identificados como tumor tem pouca expressão. Contudo, analisando os dois primeiros classificados existem metodologias que já permitem resultados bastante positivos para este conjunto de dados. Porém pode-se suspeitar da existência de um *overfitting* com os dados de treino, pelo que estas foram testadas e apresentadas no conjunto *Challenge*.

Na figura 6.18 encontra-se a classificação obtida pelo conjunto *Challenge*, obtendo-se o quinto lugar. De modo a que seja possível realizar uma análise mais cuidada na tabela 6.7 encontram-se os resultados dos *Proceedings* do desafio [44, 137], que foram testados e apresentados na base de dados BraTS 2013. Adicionalmente as duas melhores metodologias, apresentadas ao longo deste capítulo, são incluídas.

A Melhor Configuração diz respeito a uma classificação HEGS, associada às características de segunda ordem, extraídas da base de dados BraTS 2013 normalizada com os parâmetros de treino de Nyúl-Udupa para BraTS 2015. O pós-processamento aplicado resulta do algoritmo de remoção de regiões achatadas. No caso da segunda Melhor Configuração, esta difere no modo de classificação, com um único estágio de classificação, e ainda na normalização das imagens. Neste último procedimento a base de dados é normalizada com os parâmetros provenientes do treino de BraTS 2013.

Pela análise da tabela anterior, constata-se que nenhum método consegue alcançar o melhor desempenho em todas as métricas em simultâneo, reforçando a dificuldade existente na área da segmentação tumoral cerebral. Adicionalmente, é ainda possível realizar uma comparação direta com o método proposto por Goetz *et al.* [75], já descrito. Comparando com a Melhor Configuração observa-se que esta é capaz de superar a variante proposta por Goetz *et al.* [75]. Assim sendo, é possível indicar que o conjunto de características projetado, na presente dissertação, contém uma maior capacidade para treinar um classificador com maior capacidade de generalização.

Focando a análise no PPV é possível averiguar que metodologia proposta apresenta resultados competitivos, o que leva a concluir que a Melhor Configuração tem a capacidade de rejeitar falsos positivos. A capacidade em descartar falsos positivos é uma mais-valia em posteriores estudos médicos, que procuram caracterizar as dimensões tumorais com robustez sem despoletar erros de análise.

No caso da região ativa contrastante (classe 4) observa-se que a Melhor Configuração tem também um

desempenho digno de relevo. Como vem sendo referido ao longo da presente dissertação, a caracterização da região ativa contrastante é de interesse para análise médica. Visto que esta classe é a responsável pela proliferação celular, a caracterização em termos volumétricos é fulcral para, por exemplo, a avaliação e planeamento do tratamento.

Comparando as duas melhores configurações retira-se que a melhor configuração, validada no conjunto *Training*, é capaz de assegurar uma melhor delimitação do tumor total. A melhor configuração supera ou iguala a segunda proposta em todas as métricas, expeto o PPV do tumor central. Deste modo, leva-se a concluir que o novo conjunto de parâmetros de normalização, aliado a uma classificação em dois estágios permitem um desempenho mais robusto e generalista da metodologia de segmentação tumoral cerebral.

Por outro lado, comparando a Melhor Configuração com métodos mais recentes, presentes na figura 6.18, é possível verificar que o desempenho é menor no caso da região tumoral total na métrica Dice e sensibilidade. Não obstante, o valor do PPV para a região tumoral total, dentro do conjunto de métodos expostos, é o terceiro classificado. Esta particularidade é indicativa de que a Melhor Configuração é capaz de classificar menos amostras de tecido cerebral saudável, em relação aos restantes métodos. Porém, devido a uma sensibilidade reduzida existem também falhas na deteção de *voxels* associados à classe tumoral. Na região ativa contrastante, uma vez mais, a Melhor Configuração é capaz de assegurar uma segmentação robusta. Tal facto é comprovado pelo valor de Dice e sensibilidade para a classe em questão.

Tabela 6.7: Média de Dice, PPV e sensibilidade no conjunto *Challenge* da base de dados BraTS 2013.

Métodos \ Métricas	Dice			PPV			Sensibilidade		
	Tumor Total	Tumor Central	Classe 4	Tumor Total	Tumor Central	Classe 4	Tumor Total	Tumor Central	Classe 4
Tustinson et al. [137]	0,87	0,78	0,74	0,85	0,74	0,69	0,89	0,88	0,83
Reza et al.[137]	0,83	0,72	0,72	0,82	0,81	0,70	0,86	0,69	0,76
Kwon et al.[54]	0,88	0,83	0,72	0,92	0,90	0,74	0,84	0,78	0,72
Kleesiek et al.[76]	0,87	0,76	0,74	-	-	-	-	-	-
Meier et al.[77]	0,84	0,73	0,68	0,80	0,80	0,72	0,89	0,70	0,70
Goetz et al.[44]	0,83	0,71	0,68	-	-	-	-	-	-
Melhor Configuração	0,85	0,79	0,75	0,87	0,85	0,71	0,84	0,75	0,82
2ª Melhor Configuração	0,83	0,78	0,73	0,86	0,86	0,71	0,82	0,75	0,79

Evaluation Results: Training

Patient

Position	User	Dice			Positive Predictive Value			Sensitivity			Kappa	Complete tumor Rank	Tumor core Rank	Enhancing tumor Rank
		complete	core	enhancing	complete	core	enhancing	complete	core	enhancing				
1	rezas1	0.92 (1)	0.91 (2)	0.68 (1)	0.97 (1)	0.95 (1)	0.70 (1)	0.89 (6)	0.88 (2)	0.66 (1)	1.00 (1)	2.25	1.75	5.75
2	ziked1	0.92 (3)	0.93 (1)	0.63 (2)	0.86 (12)	0.92 (2)	0.63 (6)	0.98 (1)	0.94 (1)	0.64 (2)	1.00 (3)	4.75	1.25	7.75
3	silvc1	0.87 (5)	0.73 (6)	0.58 (3)	0.93 (3)	0.86 (6)	0.62 (8)	0.84 (11)	0.69 (10)	0.61 (4)	1.00 (5)	6.00	7.25	4.00
4	tustn1	0.88 (4)	0.76 (4)	0.55 (8)	0.88 (8)	0.80 (9)	0.65 (4)	0.89 (5)	0.79 (5)	0.53 (8)	1.00 (4)	5.25	5.25	8.00
5	pinta1	0.86 (8)	0.72 (7)	0.58 (4)	0.91 (5)	0.88 (3)	0.64 (5)	0.83 (13)	0.69 (9)	0.60 (5)	1.00 (6)	8.00	6.75	4.00
6	zhaol1	0.86 (7)	0.75 (5)	0.55 (9)	0.88 (9)	0.86 (5)	0.65 (3)	0.86 (9)	0.72 (7)	0.50 (12)	1.00 (8)	8.25	5.75	9.75
7	havam2	0.87 (6)	0.79 (3)	0.53 (10)	0.92 (4)	0.82 (8)	0.56 (13)	0.83 (14)	0.80 (4)	0.54 (7)	1.00 (7)	7.75	4.75	13.50
8	kleej1	0.83 (12)	0.67 (10)	0.55 (7)	0.90 (7)	0.77 (11)	0.66 (2)	0.80 (20)	0.69 (12)	0.51 (9)	1.00 (10)	12.25	10.75	8.50

Figura 6.17: Classificação obtida pela metodologia proposta no conjunto *Training*.

Evaluation Results: Challenge

Patient

Position	User	Dice			Positive Predictive Value			Sensitivity			Kappa	Complete tumor Rank	Tumor core Rank	Enhancing tumor Rank
		complete	core	enhancing	complete	core	enhancing	complete	core	enhancing				
1	kwond1	0.88 (1)	0.81 (2)	0.71 (7)	0.92 (1)	0.85 (6)	0.71 (11)	0.85 (13)	0.79 (7)	0.72 (10)	0.99 (1)	5.00	5.00	6.44
2	peres1	0.85 (5)	0.81 (1)	0.74 (2)	0.83 (16)	0.81 (9)	0.69 (16)	0.89 (7)	0.84 (3)	0.83 (2)	0.99 (3)	9.33	4.33	6.78
3	urbag1	0.87 (2)	0.76 (5)	0.73 (5)	0.87 (7)	0.86 (4)	0.73 (7)	0.86 (12)	0.71 (15)	0.75 (7)	0.99 (4)	7.00	8.00	7.11
4	tustn1	0.87 (3)	0.78 (4)	0.74 (3)	0.85 (12)	0.74 (19)	0.69 (17)	0.89 (4)	0.88 (1)	0.83 (3)	0.99 (2)	6.33	8.00	7.33
5	pinta1	0.85 (6)	0.79 (3)	0.75 (1)	0.87 (8)	0.85 (7)	0.71 (13)	0.84 (16)	0.75 (10)	0.82 (4)	0.99 (5)	10.00	6.67	7.56
6	kleej1	0.85 (8)	0.74 (7)	0.73 (4)	0.83 (15)	0.76 (17)	0.77 (4)	0.87 (10)	0.75 (11)	0.72 (12)	0.99 (8)	11.00	11.67	9.78
7	silvc1	0.83 (14)	0.75 (6)	0.72 (6)	0.76 (23)	0.72 (22)	0.65 (25)	0.92 (2)	0.85 (2)	0.85 (1)	0.99 (17)	13.00	10.00	11.22
8	meier1	0.82 (15)	0.73 (9)	0.69 (10)	0.76 (24)	0.78 (11)	0.71 (12)	0.92 (1)	0.72 (14)	0.73 (9)	0.99 (19)	13.33	11.33	11.67

Figura 6.18: Classificação obtida pela metodologia proposta no conjunto *Challenge*.

6.11 Sumário

Este capítulo iniciou-se pela descrição da base de dados utilizada para testar as diferentes configurações da metodologia desenvolvida na presente dissertação. Composta por 20 pacientes de HGG e 10 de LGG, a base de dados foi obtida do NCI-MICCAI - BraTs 2013 contendo imagens resultantes da aquisição de diferentes instituições e equipamentos. De seguida foi descrito o mecanismo implementado para a avaliação e realização dos testes, nomeadamente as métricas utilizadas para a avaliação do desempenho e ainda os parâmetros de treino dos classificadores supervisionados.

No primeiro estudo é considerada a possibilidade de integração de características de textura num conjunto inicial de características já testado. Posteriormente, é ainda repensado o conjunto de características inicial, procurando-se, em ambos, assegurar uma maior capacidade de discriminação entre os diferentes tipos de tecido tumoral a classificar e por conseguinte melhor desempenho. De seguida, no estágio de classificação, foi avaliado e discutido qual o melhor dos dois classificadores estudados (Flores-tas de Decisão ou Extra-Trees) que permitem assegurar maior capacidade de generalização e robustez. Determinado o classificador mais adequado, foram posteriormente estudados vários hiper-parâmetros do classificador, de modo a garantir um maior grau de ajuste do classificador ao conjunto de dados. A ausência de proporcionalidade entre amostras, em LGG, foi também alvo de análise, tendo resultado num processo de amostragem interquartil para a seleção de amostras de comportamento similar às classes de menor expressividade (necrose e região ativa contrastante). Posteriormente, O impacto da sequência T1 foi também estudado, tendo despoletado uma abordagem de classificação em dois estágios que fora seguidamente testada. O pós-processamento da metodologia de segmentação foi também alvo de um estudo, tendo sido alterado o pressuposto inicial para a eliminação de aglomerados de menor volume. Uma última secção contabiliza a participação no NCI-MICCAI - Grand Challenge BraTS 2013 e uma comparação com os métodos mais recentes.

Conclusões/Perspetivas Futuras

7.1 Conclusão

A finalidade da presente dissertação consistiu no desenvolvimento de uma metodologia automática de segmentação de tumores cerebrais, identificando os diferentes tipos de tecido tumorais (necrose, edema e tecidos ativos contrastante e não contrastante). A informação providenciada por parte da segmentação é de extremo relevo para o planeamento de uma possível intervenção cirúrgica ou ainda para a monitorização e avaliação da radioterapia ou outro tipo de tratamento, tendo sido esta a motivação fundamental para a realização deste trabalho.

Diversas propostas de segmentação tumoral têm sido apresentadas nos últimos anos, procurando sempre substituir a segmentação manual e as suas desvantagens associadas, como por exemplo a variabilidade intra- e inter-especialista bem como os recursos humanos que a segmentação manual exige. Contudo, os métodos de segmentação tumoral automática ainda não se encontram implementados em larga escala, visto que a comunidade científica ainda considera que os resultados obtidos não são robustos o suficiente.

O trabalho realizado iniciou-se com uma análise à capacidade discriminativa das características de textura de Haralick, concluindo-se que este tipo de características tem influência no classificador e por conseguinte nos resultados obtidos. Numa primeira fase, através da distância de Bhattacharyya concluiu-se que nem todas as 13 características de Haralick têm poder discriminativo. Sendo o tempo de computação destas características demasiado elevado optou-se por, com base na medida estatística anterior, reduzir o conjunto de características a extrair para um subconjunto de três características de textura. Posteriormente, procurou-se comparar o desempenho do classificador quando este contém um conjunto predefinido de características com a adição das três novas características mais discriminativas de Haralick. Constatou-se que a textura, quando presente, é quantitativamente melhor para gliomas de baixo grau assegurando uma melhor delimitação do tumor total e da região ativa. Porém, no caso dos gliomas de elevado grau a textura influencia as funções de divisão dos nodos produzindo segmentações comparativamente mais fracas, quando as características de Haralick não são utilizadas. O desempenho comprometedor juntamente com o seu tempo de computação levaram ao abandono destas características.

O segundo estudo comparativo das características permitiu concluir que a utilização de não-linearidades, como volume base para a computação de características de maior dimensão, acarreta vantagens na capacidade de segmentação da floresta. Em ambos os graus de tumor cerebral foi possível constatar melhorias significativas nas médias das métricas avaliadas, visto que o classificador apresenta melhor capacidade de distinção entre as diversas classes de tecido tumoral e ainda a distinção entre tecido tumoral e tecido saudável. Porém, foi possível constatar que, ainda, nenhuma das configurações apresenta resultados robustos na segmentação da região não contrastante, em tumores HGG, e nas regiões ativa e necrótica de tumores LGG.

Posteriormente, efetuou-se a procura dos melhores parâmetros que definem uma floresta de decisão, verificou-se que este estudo é um procedimento demorado e temporalmente exigente. Não obstante, foi possível retirar que as configurações definidas inicialmente são as que melhor se coadunam com o contexto do problema, dado o baixo tempo de computação associado. Embora o conjunto de combinações estudadas, dentro de um vasto conjunto de combinações, seja consideravelmente pequeno (100 candidatos apenas) suspeita-se que a melhor combinação de parâmetros obtida, apesar de ter um desempenho superior à configuração inicial, seja passível de *overfitting* ao conjunto de *Training*. Por outro lado, o terceiro classificado também não provou ser a melhor abordagem, embora que muito semelhante às configurações inicialmente definidas para o treino e teste da floresta. A explicação para esta situação assenta fundamentalmente no número elevado de características utilizadas para a divisão de cada nodo, visto que se aumentou a variância dos dados associados em cada nodo de divisão.

Relativamente ao efeito da amostragem, o seu principal objetivo consistia na melhoria da segmentação dos pacientes de LGG. A adição de amostras de necrose e tecido ativo contrastante de HGG, normalizadas com os percentis de LGG, permitiu um melhor balanceamento do classificador, despoletando um aumento na qualidade das segmentações obtidas. Assim sendo, comprovou-se a suspeita de que anteriores resultados menos competitivos, em LGG, se deviam a uma fraca expressão das amostras de necrose e tecido ativo contrastante. Complementarmente, a aplicação de uma filtragem interquartil, às intensidades, dentro do conjunto de amostras existentes, foi fulcral. Com a definição de filtros interquartil assegurou-se que a amostragem incluiu amostras de comportamento próximo ou semelhante às amostras características de LGG, de modo a não aumentar a variabilidade de cada classe. A operação teve um impacto positivo, obtendo-se um aumento no desempenho em cerca de 44%.

Com a implementação de uma metodologia de segmentação em dois estágios, para além de diminuir os recursos computacionais exigidos, assegurou um melhor desempenho da segmentação tumoral. Esta metodologia proposta resulta de um estudo prévio relativo ao impacto da sequência T1 na classificação, onde se concluiu que para uma delimitação do tumor esta sequência não acarreta resultados benéficos. Assim numa primeira fase de identificação do tumor, a sequência T1 e respetivas características não foram tidas em consideração. Para além da diminuição dos recursos computacionais exigidos constatou-se um aumento na robustez da metodologia proposta pois, para uma delimitação inicial do tumor viu aumentado o seu desempenho em ambos os tipos de glioma. Porém, com a posterior introdução de uma segunda fase, de classificação do tumor nas diferentes classes de tecido tumoral verificou-se que o impacto da sequência T1 não pode ser generalizado para ambos os níveis de glioma. Em HGG o seu desempenho é superior ao na ausência da sequência T1, ao par que em LGG a presença da sequência

T1 teve um impacto considerável. Assim sendo, a metodologia proposta passou a ser sensível ao tipo de glioma a segmentar dado que o conjunto de sequências e respetivas características é distinto de HGG e em LGG. Não obstante conclui-se que com duas fases de classificação representa melhor metodologia de segmentação de tumores cerebrais desenvolvida no âmbito da presente dissertação.

Por último, em relação ao pós-processamento realizado, retira-se que o seu efeito, apesar de ser mínimo na grande maioria dos pacientes, é fundamental na remoção de falsas marcações. A segunda versão de pós-processamento proposta provou ser a mais adequada para o tratamento das segmentações à saída do classificador, visto que, esta última tem a capacidade de contabilizar a possível existência de dois volumes tumorais, num mesmo paciente.

Em conclusão, os resultados obtidos pelo método desenvolvido nesta dissertação asseguraram uma a segmentação de tumores cerebrais, consistente e competitiva. Quando comparada com restantes metodologias de investigação de equipas com bastante experiência na área da Imagem Médica, a metodologia proposta teve um desempenho bastante positivo.

7.2 Perspetivas Futuras

Métodos computacionais de segmentação de tumores cerebrais é uma área ainda em fase de plena investigação, constatando-se que vários grupos (de diferentes partes do globo) procuram dar o seu contributo e resolver os problemas associados a esta área. Num passado bastante recente novas metodologias assentes em *Deep-Learning* (redes neuronais) têm demonstrado ser propostas bastante capazes de assegurar segmentações de qualidade elevada.

Apesar dos resultados obtidos, o método desenvolvido apresenta ainda alguns pontos que podem ser alvo de estudo, e posterior melhoria. Como foi possível de retirar ao longo deste trabalho, os tempos de computação associados, às diferentes etapas, são consideráveis. Uma otimização do modo de computação de alguns estágios dever ser então implementada, especialmente na etapa de extração das características. O desenvolvimento de uma metodologia capaz de paralelizar o cálculo das características, tirando partido de toda a capacidade de processamento do equipamento, é essencial para a redução tempo de processamento.

Com a alteração dos percentis da normalização das intensidades, constatou-se que o pré-processamento é uma fase que condiciona o desempenho do classificador. Um estudo mais completo sobre as técnicas de normalização e o modo como estas devem ser aplicadas seria positivo para alcançar segmentações quantitativamente melhores.

Por último, a etapa de pós-processamento deve ser também alvo de um estudo. Vários autores, implementam após a floresta de decisão etapas de regularização espacial, que se serve das probabilidades das florestas para avaliar e qualificar a relação entre pontos vizinhos. Assim sendo, diferentes abordagens para processar os dados após a segmentação de um classificador devem ser analisadas.

Bibliografia

- [1] Jacques Ferlay, Isabelle Soerjomataram, Rajesh Dikshit, Sultan Eser, Colin Mathers, Marise Rebelo, Donald Maxwell Parkin, David Forman, e Freddie Bray. Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. *International Journal of Cancer*, 136(5): E359–E386, 2015.
- [2] Jacques Ferlay, Hai-Rim Shin, Freddie Bray, David Forman, Colin Mathers, e Donald Maxwell Parkin. Estimates of worldwide burden of cancer in 2008: Globocan 2008. *International journal of cancer*, 127(12):2893–2917, 2010.
- [3] Hiroko Ohgaki e Paul Kleihues. Epidemiology and etiology of gliomas. *Acta neuropathologica*, 109(1):93–108, 2005.
- [4] World Health Organization *et al.* Cancer fact sheet n°297. *Version February*, 2011.
- [5] Alba A Brandes. State-of-the-art treatment of high-grade brain tumors. In *Seminars in oncology*, volume 30, pages 4–9. Elsevier, 2003.
- [6] John W Henson, Paola Gaviani, e R Gilberto Gonzalez. Mri in treatment of adult gliomas. *The lancet oncology*, 6(3):167–175, 2005.
- [7] Sarah S Donaldson, Fred Laningham, e Paul Graham Fisher. Advances toward an understanding of brainstem gliomas. *Journal of Clinical Oncology*, 24(8):1266–1272, 2006.
- [8] PhD Dr. András Jakab, MD. *Segmenting Brain Tumors with the Slicer 3D Software*, 2012. Manual for providing expert segmentations for the BRATS-Ch Tumor Segmentation Challenge.
- [9] Adriano Pinto, Sergio Pereira, Hugo Dinis, Carlos A Silva, e Deolinda MLD Rasteiro. Random decision forests for automatic brain tumor segmentation on multi-modal mri images. In *Bioengineering (ENBENG), 2015 IEEE 4th Portuguese Meeting on*, pages 1–5. IEEE, 2015.
- [10] Adriano Pinto, Sergio Pereira, Higino Correia, Jorge Oliveira, Deolinda M. L. D. Rasteiro, e Carlos A Silva. Brain tumour segmentation based on extremely randomized forest with high-level features. *EMBC 2015*, 2015.

- [11] International Agency for Research on Cancer. The global cancer atlas, sep 2015. URL <http://globocan.iarc.fr/ia/World/atlas.html>.
- [12] David N Louis, Hiroko Ohgaki, Otmar D Wiestler, Webster K Cavenee, Peter C Burger, Anne Jouvett, Bernd W Scheithauer, e Paul Kleihues. The 2007 who classification of tumours of the central nervous system. *Acta neuropathologica*, 114(2):97–109, 2007.
- [13] Leslie H Sobin, Mary K Gospodarowicz, e Christian Wittekind. *TNM classification of malignant tumours*. John Wiley & Sons, 2011.
- [14] Vladimir Parpura, Michael T Heneka, Vedrana Montana, Stéphane HR Olié, Arne Schousboe, Philip G Haydon, Randy F Stout, David C Spray, Andreas Reichenbach, Thomas Pannicke, *et al*. Glial cells in (patho) physiology. *Journal of neurochemistry*, 121(1):4–27, 2012.
- [15] F Lafitte, S Morel-Precetti, N Martin-Duverneuil, A Guermazi, E Brunet, F Heran, e J Chiras. Multiple glioblastomas: Ct and mr features. *European radiology*, 11(1):131–136, 2001.
- [16] Hiroko Ohgaki e Paul Kleihues. Genetic pathways to primary and secondary glioblastoma. *The American journal of pathology*, 170(5):1445–1453, 2007.
- [17] Vairavan Narayanan. High grade gliomas: Pathogenesis, management. 12(4):24–29, 2012.
- [18] Freddie Bray, Jian-Song Ren, Eric Masuyer, e Jacques Ferlay. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *International Journal of Cancer*, 132(5):1133–1145, 2013.
- [19] Ruth A Kleinerman, Martha S Linet, Elizabeth E Hatch, Robert E Tarone, Peter M Black, Robert G Selker, William R Shapiro, Howard A Fine, e Peter D Inskip. Self-reported electrical appliance use and risk of adult brain tumors. *American journal of epidemiology*, 161(2):136–146, 2005.
- [20] Crystal N Holick, Edward L Giovannucci, Bernard Rosner, Meir J Stampfer, e Dominique S Michaud. Prospective study of cigarette smoking and adult glioma: dosage, duration, and latency. *Neuro-oncology*, 9(3):326–334, 2007.
- [21] Melissa Bondy, John Wiencke, Margaret Wrensch, e Athanassios P Kyritsis. Genetics of primary brain tumors: a review. *Journal of neuro-oncology*, 18(1):69–81, 1994.
- [22] Patrick Y Wen e Santosh Kesari. Malignant gliomas in adults. *New England Journal of Medicine*, 359(5):492–507, 2008.
- [23] William R Hendee e E Russell Ritenour. *Medical imaging physics*. John Wiley & Sons, 2003.
- [24] Stefan Bauer, Roland Wiest, Lutz-P Nolte, e Mauricio Reyes. A survey of mri-based medical image analysis for brain tumor studies. *Physics in medicine and biology*, 58(13):R97, 2013.
- [25] Bjoern Menze *et al*. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Imaging*, 2014. doi: 10.1109/TMI.2014.2377694.

-
- [26] A Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, 14:841, 2002.
- [27] Tammy Riklin Raviv, KV Leemput, e Bjoern H Menze. Multi-modal brain tumor segmentation via latent atlases. *Proc. MICCAI-BRATS*, pages 64–73, 2012.
- [28] Lior Weizman, L Ben Sira, Leo Joskowicz, Shlomi Constantini, Ronit Precel, Ben Shofty, e D Ben Bashat. Automatic segmentation, internal classification, and follow-up of optic pathway gliomas in mri. *Medical image analysis*, 16(1):177–188, 2012.
- [29] Martin Rajchl, JS Baxter, A Jonathan McLeod, Jing Yuan, Wu Qiu, Terry M Peters, e Ali R Khan. Asets: Map-based brain tissue segmentation using manifold learning and hierarchical max-flow regularization. In *MRBrainS 2013-MICCAI Grand Challenge on MR Brain Segmentation 2013*. 2014.
- [30] Jun Jiang, Yao Wu, Meiyang Huang, Wei Yang, Wufan Chen, e Qianjin Feng. 3d brain tumor segmentation in multimodal mr images based on learning population-and patient-specific feature sets. *Computerized Medical Imaging and Graphics*, 37(7):512–521, 2013.
- [31] Meiyang Huang, Wei Yang, Yao Wu, Jun Jiang, Wufan Chen, e Qianjin Feng. Brain tumor segmentation based on local independent projection-based classification. *Biomedical Engineering, IEEE Transactions on*, 61(10):2633–2645, 2014.
- [32] Ilkay Ulusoy e Christopher M Bishop. Generative versus discriminative methods for object recognition. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 258–265. IEEE, 2005.
- [33] Guillaume Bouchard e Bill Triggs. The tradeoff between generative and discriminative classifiers. In *16th IASC International Symposium on Computational Statistics (COMPSTAT'04)*, pages 721–728, 2004.
- [34] M Masroor Ahmed e Dzulkifli Bin Mohamad. Segmentation of brain mr images for tumor extraction by combining kmeans clustering and perona-malik anisotropic diffusion model. *International Journal of Image Processing*, 2(1):27–34, 2008.
- [35] Lynn M Fletcher-Heath, Lawrence O Hall, Dmitry B Goldgof, e F Reed Murtagh. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artificial intelligence in medicine*, 21(1):43–63, 2001.
- [36] Jianguo Liu, Jayaram K Udupa, Dewey Odhner, David Hackney, e Gul Moonis. A system for brain tumor volume estimation via mr imaging and fuzzy connectedness. *Computerized Medical Imaging and Graphics*, 29(1):21–34, 2005.
- [37] Antonio Criminisi e Jamie Shotton. *Decision forests for computer vision and medical image analysis*. Springer Science & Business Media, 2013.
-

- [38] Corinna Cortes e Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [39] László G Nyúl, Jayaram K Udupa, e Xuan Zhang. New variants of a method of mri scale standardization. *IEEE Trans. Med. Imaging*, 19(2):143–150, 2000.
- [40] Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, e James C Gee. N4itk: improved n3 bias correction. *IEEE Trans. Med. Imaging*, 29(6):1310–1320, 2010.
- [41] Arno Klein, Jesper Andersson, Babak A Ardekani, John Ashburner, Brian Avants, Ming-Chang Chiang, Gary E Christensen, D Louis Collins, James Gee, Pierre Hellier, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage*, 46(3):786–802, 2009.
- [42] Brian B Avants, Nicholas J Tustison, Michael Stauffer, Gang Song, Baohua Wu, e James C Gee. The insight toolkit image registration framework. *Frontiers in neuroinformatics*, 8, 2014.
- [43] Bjoern Menze, Andras Jakab, Stefan Bauer, Mauricio Reyes, Marcel Prastawa, e Koen Van Leemput. Proceedings of the miccai challenge on multimodal brain tumor image segmentation (brats) 2012. In *MICCAI Challenge on Multimodal Brain Tumor Image Segmentation (BRATS)*, page 77. MICCAI, 2012.
- [44] B Menze, A Jakab, S Bauer, J Kalpathy-Cramer, K Farahani, J Kirby, Y Burren, N Porz, J Slotboom, R Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats)(2014). URL <http://hal.inria.fr/hal-00935640>.
- [45] Harvey E Cline, William E Lorensen, Ron Kikinis, e Ferenc Jolesz. Three-dimensional segmentation of mr images of the head using probability and connectivity. *Journal of computer assisted tomography*, 14(6):1037–1045, 1990.
- [46] Matthew C Clark, Lawrence O Hall, Dmitry B Goldgof, Robert Velthuizen, F Reed Murtagh, e Martin S. Silbiger. Automatic tumor segmentation using knowledge-based techniques. *Medical Imaging, IEEE Transactions on*, 17(2):187–201, 1998.
- [47] Michael R Kaus, Simon K Warfield, Arya Nabavi, Peter M Black, Ferenc A Jolesz, e Ron Kikinis. Automated segmentation of mr images of brain tumors 1. *Radiology*, 218(2):586–591, 2001.
- [48] Kelly H Zou, William M Wells, Ron Kikinis, e Simon K Warfield. Three validation metrics for automated probabilistic image segmentation of brain tumours. *Statistics in medicine*, 23(8):1259–1282, 2004.
- [49] Marcel Prastawa, Elizabeth Bullitt, Nathan Moon, Koen Van Leemput, e Guido Gerig. Automatic brain tumor segmentation by subject specific modification of atlas priors 1. *Academic radiology*, 10(12):1341–1348, 2003.

-
- [50] Marcel Prastawa, Elizabeth Bullitt, Sean Ho, e Guido Gerig. A brain tumor segmentation framework based on outlier detection. *Medical image analysis*, 8(3):275–283, 2004.
- [51] Stefan Bauer e Mauricio Reyes. A markov random field-based biomechanical tumor growth model for atlas-based segmentation of brain tumor images. *This volume is dedicated to cancer patients*, page 76, 2010.
- [52] Stefan Bauer, Lutz-Peter Nolte, e Mauricio Reyes. Segmentation of brain tumor images based on atlas-registration combined with a markov-random-field lesion growth model. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pages 2018–2021. IEEE, 2011.
- [53] Bjoern H Menze, Koen Van Leemput, Danial Lashkari, Marc-André Weber, Nicholas Ayache, e Polina Golland. A generative model for brain tumor segmentation in multi-modal images. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 151–159. Springer, 2010.
- [54] Dongjin Kwon, Russell T Shinohara, Hamed Akbari, e Christos Davatzikos. Combining generative models for multifocal glioma segmentation and registration. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 763–770. Springer, 2014.
- [55] Andac Hamamci, Nadir Kucuk, Kutlay Karaman, Kayihan Engin, e Gozde Unal. Tumor-cut: segmentation of brain tumors on contrast enhanced mr images for radiosurgery applications. *Medical Imaging, IEEE Transactions on*, 31(3):790–804, 2012.
- [56] Bjoern H Menze, Erin Stretton, Ender Konukoglu, e Nicholas Ayache. Image-based modeling of tumor growth in patients with glioma. *Optimal control in image processing*, 2011.
- [57] L Görlitz, Bjoern H Menze, M-A Weber, B Michael Kelm, e Fred A Hamprecht. Semi-supervised tumor detection in magnetic resonance spectroscopic images using discriminative random fields. In *Pattern Recognition*, pages 224–233. Springer, 2007.
- [58] Elsa D Angelini, Julie Delon, Alpha Boubacar Bah, Laurent Capelle, e Emmanuel Mandonnet. Differential mri analysis for quantification of low grade glioma growth. *Medical image analysis*, 16(1): 114–126, 2012.
- [59] Ali Gooya, Kilian M Pohl, Michel Bilello, Luigi Cirillo, George Biros, Elias R Melhem, e Christos Davatzikos. Glistr: glioma image segmentation and registration. *Medical Imaging, IEEE Transactions on*, 31(10):1941–1954, 2012.
- [60] Lior Weizman, Liat Ben Sira, Leo Joskowicz, Daniel L Rubin, Kristen W Yeom, Shlomi Constantini, Ben Shofty, e Dafna Ben Bashat. Semiautomatic segmentation and follow-up of multicomponent low-grade tumors in longitudinal brain mri studies. *Medical physics*, 41(5):052303, 2014.
- [61] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.
-

- [62] Chi-Hoon Lee, Mark Schmidt, Albert Murtha, Aalo Bistritz, Jörg Sander, e Russell Greiner. Segmenting brain tumors with conditional random fields and support vector machines. In *Computer vision for biomedical image applications*, pages 469–478. Springer, 2005.
- [63] Stefan Bauer, Lutz-P Nolte, e Mauricio Reyes. Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 354–361. Springer, 2011.
- [64] Stefan Bauer, Thomas Fejes, Johannes Slotboom, Roland Wiest, Lutz-P Nolte, e Mauricio Reyes. Segmentation of brain tumor images based on integrated hierarchical classification and regularization. *Proc MICCAI-BRATS*, 2012.
- [65] Hongming Li e Yong Fan. Label propagation with robust initialization for brain tumor segmentation. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 1715–1718. IEEE, 2012.
- [66] Wei Wu, Albert YC Chen, Liang Zhao, e Jason J Corso. Brain tumor detection and segmentation in a crf (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features. *International journal of computer assisted radiology and surgery*, 9(2):241–253, 2014.
- [67] Darko Zikic, Ben Glocker, Ender Konukoglu, Antonio Criminisi, C Demiralp, Jamie Shotton, OM Thomas, T Das, R Jena, e SJ Price. Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel mr. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 369–376. Springer, 2012.
- [68] Ezequiel Geremia, Bjoern H Menze, e Nicholas Ayache. Spatial decision forests for glioma segmentation in multi-channel mr images. *MICCAI Challenge on Multimodal Brain Tumor Segmentation*, 34, 2012.
- [69] Ezequiel Geremia, Olivier Clatz, Bjoern H Menze, Ender Konukoglu, Antonio Criminisi, e Nicholas Ayache. Spatial decision forests for ms lesion segmentation in multi-channel magnetic resonance images. *NeuroImage*, 57(2):378–390, 2011.
- [70] Atiq Islam, Syed MS Reza, e Khan M Iftekharuddin. Multifractal texture estimation for detection and segmentation of brain tumors. *Biomedical Engineering, IEEE Transactions on*, 60(11):3204–3215, 2013.
- [71] Benoit B Mandelbrot e John W Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.
- [72] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, 28(2): 337–407, 2000.

-
- [73] S Reza e KM Iftekharuddin. Multi-fractal texture features for brain tumor and edema segmentation. In *SPIE Medical Imaging*, pages 903503–903503. International Society for Optics and Photonics, 2014.
- [74] Nicholas J Tustison, KL Shrinidhi, Max Wintermark, Christopher R Durst, Benjamin M Kandel, James C Gee, Murray C Grossman, e Brian B Avants. Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ants. *Neuroinformatics*, pages 1–17, 2014.
- [75] Michael Goetz, Christian Weber, Josiah Bloecher, Bram Stieltjes, Hans-Peter Meinzer, e Klaus Maier-Hein. Extremely randomized trees based brain tumor segmentation, 2014. URL <https://sites.google.com/site/miccaibrats2014/> [Accessed: March 23, 2014].
- [76] Jens Kleesiek, Armin Biller, Gregor Urban, Ullrich Köthe, Martin Bendszus, e Fred A Hamprecht. ilastik for multi-modal brain tumor segmentation.
- [77] Raphael Meier, Stefan Bauer, Johannes Slotboom, Roland Wiest, e Mauricio Reyes. Appearance- and context-sensitive features for brain tumor segmentation.
- [78] Gul Moonis, Jianguo Liu, Jayaram K Udupa, e David B Hackney. Estimation of tumor volume with fuzzy-connectedness segmentation of mr images. *American Journal of Neuroradiology*, 23(3): 356–363, 2002.
- [79] Uroš Vovk, Franjo Pernuš, e Boštjan Likar. A review of methods for correction of intensity inhomogeneity in mri. *Medical Imaging, IEEE Transactions on*, 26(3):405–421, 2007.
- [80] Pierre Hellier, Christian Barillot, Isabelle Corouge, Bernard Gibaud, Georges Le Goualher, D Louis Collins, Allan Evans, Grégoire Malandain, Nicholas Ayache, Gary E. Christensen, *et al.* Retrospective evaluation of intersubject brain registration. *Medical Imaging, IEEE Transactions on*, 22(9):1120–1130, 2003.
- [81] Christine Fennema-Notestine, I Burak Ozyurt, Camellia P Clark, Shaunna Morris, Amanda Bischoff-Grethe, Mark W Bondi, Terry L Jernigan, Bruce Fischl, Florent Segonne, David W Shattuck, *et al.* Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location. *Human brain mapping*, 27(2): 99–113, 2006.
- [82] John Haselgrove e Manfred Prammer. An algorithm for compensation of surface-coil images for sensitivity of the surface coil. *Magnetic Resonance Imaging*, 4(6):469–472, 1986.
- [83] ER McVeigh, MJ Bronskill, e RM Henkelman. Phase and sensitivity of receiver coils in magnetic resonance imaging. *Medical physics*, 13(6):806–814, 1986.
- [84] John G Sled, Alex P Zijdenbos, e Alan C Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *Medical Imaging, IEEE Transactions on*, 17(1):87–97, 1998.
-

- [85] Sylvain Prima, Nicholas Ayache, Tom Barrick, e Neil Roberts. Maximum likelihood estimation of the bias field in mr brain images: Investigating different modelings of the imaging process. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, pages 811–819. Springer, 2001.
- [86] Martin Styner, Christian Brechbühler, G Szckely, e Guido Gerig. Parametric estimate of intensity inhomogeneities applied to mri. *Medical Imaging, IEEE Transactions on*, 19(3):153–165, 2000.
- [87] Christian Brechbühler, Guido Gerig, e Gábor Székely. Compensation of spatial inhomogeneity in mri based on a parametric bias estimate. In *Visualization in Biomedical Computing*, pages 141–146. Springer, 1996.
- [88] Dzung L Pham e Jerry L Prince. An adaptive fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneities. *Pattern recognition letters*, 20(1):57–68, 1999.
- [89] David W Shattuck, Stephanie R Sandor-Leahy, Kirt A Schaper, David A Rottenberg, e Richard M Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876, 2001.
- [90] Andrew Simmons, Paul S Tofts, Gareth J Barker, e Simon R Arridge. Sources of intensity nonuniformity in spin echo images at 1.5 t. *Magnetic Resonance in Medicine*, 32(1):121–128, 1994.
- [91] Leon Axel, Jay Costantini, e John Listerud. Intensity correction in surface-coil mr imaging. *American Journal of Roentgenology*, 148(2):418–420, 1987.
- [92] David AG Wicks, Gareth J Barker, e Paul S Tofts. Correction of intensity nonuniformity in mr images of any orientation. *Magnetic Resonance Imaging*, 11(2):183–196, 1993.
- [93] PA Narayana, WW Brey, MV Kulkarni, e CL Sievenpiper. Compensation for surface coil sensitivity variation in magnetic resonance imaging. *Magnetic resonance imaging*, 6(3):271–274, 1988.
- [94] William W Brey e Ponnada A Narayana. Correction for intensity falloff in surface coil magnetic resonance imaging. *Medical Physics*, 15(2):241–245, 1988.
- [95] Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, Peter Boesiger, et al. Sense: sensitivity encoding for fast mri. *Magnetic resonance in medicine*, 42(5):952–962, 1999.
- [96] Benjamin H Brinkmann, Armando Manduca, e Richard A Robb. Optimized homomorphic unsharp masking for mr grayscale inhomogeneity correction. *Medical Imaging, IEEE Transactions on*, 17(2):161–171, 1998.
- [97] B Johnston, M Stella Atkins, B Mackiewicz, e M Anderson. Segmentation of multiple sclerosis lesions in intensity corrected multispectral mri. *Medical Imaging, IEEE Transactions on*, 15(2): 154–169, 1996.

- [98] LQ Zhou, YM Zhu, C Bergot, A-M Laval-Jeantet, V Bousson, J-D Laredo, e M Laval-Jeantet. A method of radio-frequency inhomogeneity correction for brain tissue segmentation in mri. *Computerized Medical Imaging and Graphics*, 25(5):379–389, 2001.
- [99] James B Arnold, Jeh-San Liow, Kirt A Schaper, Joshua J Stern, John G Sled, David W Shattuck, Andrew J Worth, Mark S Cohen, Richard M Leahy, John C Mazziotta, *et al.* Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects. *NeuroImage*, 13(5):931–943, 2001.
- [100] Benoit M Dawant, Alex P Zijdenbos, Richard Margolin, *et al.* Correction of intensity variations in mr images for computer-aided tissue classification. *Medical Imaging, IEEE Transactions on*, 12(4):770–781, 1993.
- [101] Julien Milles, Yue Min Zhu, Gérard Gimenez, Charles RG Guttmann, e Isabelle E Magnin. Mri intensity nonuniformity correction using simultaneously spatial and gray-level histogram information. *Computerized Medical Imaging and Graphics*, 31(2):81–90, 2007.
- [102] Charles R Meyer, Peyton H Bland, e James Pipe. Retrospective correction of intensity inhomogeneities in mri. *Medical Imaging, IEEE Transactions on*, 14(1):36–41, 1995.
- [103] Williams M Wells III, W Eric L Grimson, Ron Kikinis, e Ferenc A Jolesz. Adaptive segmentation of mri data. *Medical Imaging, IEEE Transactions on*, 15(4):429–442, 1996.
- [104] James C Bezdek, LO Hall, e L_P Clarke. Review of mr image segmentation techniques using pattern recognition. *Medical physics*, 20(4):1033–1048, 1992.
- [105] LP Clarke, RP Velthuizen, S Phuphanich, JD Schellenberg, JA Arrington, e MMRI Silbiger. Mri: stability of three supervised segmentation techniques. *Magnetic Resonance Imaging*, 11(1):95–106, 1993.
- [106] WA Edelstein, Paul A Bottomley, e Leah M Pfeifer. A signal-to-noise calibration procedure for nmr imaging systems. *Medical Physics*, 11(2):180–185, 1984.
- [107] T Yamamoto, T Nambu, K Miyasaka, e Y Morita. Accurate and practical calibration of mr signal intensities by the new transmission amplitude method: application of the numerical diagnosis to mri. *Radiology P*, 209:582, 1998.
- [108] László G Nyúl e Jayaram K Udupa. On standardizing the mr image intensity scale. *image*, 1081, 1999.
- [109] László G Nyúl e Jayaram K Udupa. New variants of a method of mri scale normalization. In *Information Processing in Medical Imaging*, pages 490–495. Springer, 1999.
- [110] Laszlo G Nyúl e Jayaram K Udupa. Standardizing the mr image intensity scales: making mr intensities have tissue-specific meaning. In *Medical Imaging 2000*, pages 496–504. International Society for Optics and Photonics, 2000.

- [111] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- [112] Yali Amit e Donald Geman. Randomized inquiries about shape: An application to handwritten digit recognition. Technical report, DTIC Document, 1994.
- [113] Yali Amit e Donald Geman. Shape quantization and recognition with randomized trees. *Neural computation*, 9(7):1545–1588, 1997.
- [114] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [115] Leo Breiman, JH Friedman, Richard A Olshen, e Charles J Stone. Classification and regression trees. wadsworth, 1984. *Intelligence*, pages 1002–1007, 1993.
- [116] Y-S Shih. Families of splitting criteria for classification trees. *Statistics and computing*, 9(4):309–315, 1999.
- [117] J. Ross Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.
- [118] Leo Breiman, Jerome H Friedman, Richard A Olshen, e Charles J Stone. Classification and regression trees. wadsworth. Belmont, CA, 1984.
- [119] Antonio Criminisi, Jamie Shotton, Duncan Robertson, e Ender Konukoglu. Regression forests for efficient anatomy detection and localization in ct studies. In *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*, pages 106–117. Springer, 2011.
- [120] Jamie Shotton, Matthew Johnson, e Roberto Cipolla. Semantic texton forests for image categorization and segmentation. In *Computer vision and pattern recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.
- [121] Antonio Torralba, Kevin P Murphy, e William T Freeman. Sharing features: efficient boosting procedures for multiclass object detection. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–762. IEEE, 2004.
- [122] Pierre Geurts. Contributions to decision tree induction: bias/variance tradeoff and time series classification. 2002.
- [123] Pierre Geurts, Damien Ernst, e Louis Wehenkel. Extremely randomized trees. *Machine learning*, 63(1):3–42, 2006.
- [124] Guido Van Rossum *et al*. Python programming language. In *USENIX Annual Technical Conference*, volume 41, 2007.
- [125] Luis Pedro Coelho. Mahotas: Open source software for scriptable computer vision. *Journal of Open Research Software*, 1, July 2013. doi: <http://dx.doi.org/10.5334/jors.ac>.

- [126] Stéfan van der Walt, S. Chris Colbert, e Gaël Varoquaux. The numpy array: A structure for efficient numerical computation. *Computing in Science and Engineering*, 13(2):22–30, 2011. doi: <http://dx.doi.org/10.1109/MCSE.2011.37>. URL <http://scitation.aip.org/content/aip/journal/cise/13/2/10.1109/MCSE.2011.37>.
- [127] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, e E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [128] Brian B Avants, Nick Tustison, e Gang Song. Advanced normalization tools (ants). *Insight J*, 2: 1–35, 2009.
- [129] Guneet Saini. *Texture analysis of CT scan images*. PhD thesis, THAPAR UNIVERSITY, PATIALA, 2008.
- [130] William Henry Nailon. *Texture analysis methods for medical image characterisation*. INTECH Open Access Publisher, 2010.
- [131] Robert M Haralick, Karthikeyan Shanmugam, e Its' Hak Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, pages 610–621, 1973.
- [132] Sarah Parisot, William Wells, Stéphane Chemouny, Hugues Duffau, e Nikos Paragios. Concurrent tumor segmentation and registration with uncertainty-based sparse non-uniform graphs. *Medical image analysis*, 18(4):647–659, 2014.
- [133] Nagesh Subbanna, Doina Precup, e Tal Arbel. Iterative multilevel mrf leveraging context and voxel information for brain tumour segmentation in mri. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 400–405. IEEE, 2014.
- [134] Jing Zhang, Daniel P Barboriak, Hasan Hobbs, e Maciej A Mazurowski. A fully automatic extraction of magnetic resonance image features in glioblastoma patients. *Medical physics*, 41(4):042301, 2014.
- [135] Aleksandra Popovic, Matias de la Fuente, Martin Engelhardt, e Klaus Radermacher. Statistical validation metric for accuracy assessment in medical image segmentation. *International Journal of Computer Assisted Radiology and Surgery*, 2(3-4):169–181, 2007.
- [136] A Bhattachayya. On a measure of divergence between two statistical population defined by their population distributions. *Bulletin Calcutta Mathematical Society*, 35:99–109, 1943.
- [137] Bjoern Menze, Mauricio Reyes, e Koen Van Leemput. The multimodal brain tumor image segmentation benchmark (brats). 2014.

Capítulo 8

Anexos

Nesta última secção encontram-se as tabelas com as médias das métricas de todas as classes de todo o conjunto de testes apresentados no capítulo 6 por ordem de apresentação.

8.1 Características de Haralick

Nas tabelas 8.1 e 8.2 tem-se os resultados relativos ao estudo da distância de Bhattacharyya para averiguar a capacidade discriminativa das características de Haralick. Neste estudo apenas foi utilizado um único paciente, representativo do restante conjunto de pacientes, em fatias de RM pré-definidas.

Tabela 8.1: Sobreposição e Distância de Bhattacharyya, para um único paciente, das características f1 até f7 de Haralick.

		Flair	T1	T1c	T2
	Classe Tumoral	Distância Bhattacharyya	Distância Bhattacharyya	Distância Bhattacharyya	Distância Bhattacharyya
f1	1	0.237	0.860	1.404	1.122
	2	0.183	0.555	0.239	0.542
	3	0.727	0.736	1.063	0.307
	4	0.207	0.208	1.076	1.138
f2	1	0.355	0.198	0.173	0.355
	2	0.330	0.176	0.271	0.176
	3	0.424	0.211	0.175	0.202
	4	0.394	0.190	0.183	0.218
f3	1	0.175	0.563	0.488	0.269
	2	0.288	0.846	0.609	0.202
	3	1.174	0.252	0.237	0.322
	4	0.316	0.537	0.371	0.175
f4	1	0.399	0.195	0.184	0.271
	2	0.296	0.188	0.357	0.237
	3	0.201	0.180	0.187	0.176
	4	0.418	0.208	0.174	0.307
f5	1	0.178	1.004	0.206	0.337
	2	0.253	1.699	0.263	0.305
	3	0.719	0.842	1.012	0.190
	4	0.287	0.566	0.2500	1.102
f6	1	0.179	0.173	0.187	0.175
	2	0.427	0.305	0.282	0.327
	3	0.180	0.173	0.3105	0.173
	4	0.289	0.287	0.215	0.173
f7	1	0.401	0.193	0.187	0.272
	2	0.295	0.188	0.375	0.246
	3	0.197	0.179	0.182	0.176
	4	0.419	0.208	0.175	0.320

Tabela 8.2: Sobreposição e Distância de Bhattacharyya, para um único paciente, das características f8 até f13 de Haralick.

	Classe Tumoral	Flair	T1	T1c	T2
		Distância Bhattacharyya	Distância Bhattacharyya	Distância Bhattacharyya	Distância Bhattacharyya
f8	1	0.181	0.233	0.193	0.269
	2	0.173	0.211	0.185	0.190
	3	0.197	0.194	0.221	0.183
	4	0.182	0.181	0.176	0.293
f9	1	0.179	0.219	0.203	0.252
	2	0.174	0.198	0.189	0.184
	3	0.200	0.188	0.203	0.186
	4	0.183	0.179	0.174	0.290
f10	1	185.818	48360.293	704.557	1076.368
	2	3449.55	65421.045	3265.933	17241.154
	3	8253.064	12132.258	45479.775	671.256
	4	4125.945	13546.939	13121.644	49368.798
f11	1	0.182	0.244	0.179	0.190
	2	0.197	0.229	0.191	0.184
	3	0.181	0.198	0.232	0.178
	4	0.188	0.190	0.203	0.251
f12	1	0.685	1.385	0.316	0.224
	2	0.499	3.199	0.616	0.210
	3	1.696	1.335	2.796	0.247
	4	0.193	1.937	1.654	0.179
f13	1	0.196	1.941	0.261	0.910
	2	0.204	2.109	1.021	0.484
	3	1.760	1.235	4.147	0.608
	4	0.373	1.105	1.281	1.122

Na tabela 8.3 é avaliada a distância de Bhattacharyya de três das treze características de Haralick, selecionadas aquando da realização do anterior estudo, em todos os pacientes do conjunto *Training* para toda a região cerebral.

Tabela 8.3: Distância de Bhattacharyya, de f5, f6 e f11, computada tendo em consideração todos os pacientes do conjunto de *Training*, em ambos os níveis de glioma.

		Flair	T1c	T2
Classe Tumoral		Distância Bhattacharyya	Distância Bhattacharyya	Distância Bhattacharyya
f5	1	0.449	0.295	0.196
	2	0.643	1.884	0.289
	3	0.226	0.451	0.203
	4	0.435	2.260	0.284
	5	0.479	0.450	0.434
f6	1	0.183	0.271	0.195
	2	0.238	0.670	0.179
	3	0.255	0.334	0.199
	4	0.246	0.431	0.213
	5	0.467	0.307	0.265
f11	1	0.197	0.180	0.174
	2	0.205	0.328	0.180
	3	0.180	0.203	0.176
	4	0.195	0.438	0.181
	5	0.205	0.207	0.192

Dentro do estudo das características de Haralick apresenta-se a tabela 8.4 onde se encontram as médias das métricas, de todas as classes e regiões de tecido tumoral, relativas à última fase de análise da capacidade discriminativa destas características de textura, tendo sido incluídas na classificação.

Tabela 8.4: Métricas obtidas relativas as segmentações provenientes de um classificador com textura e um classificador sem textura.

		Classificador sem textura		Classificador com textura	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.5008	0.1184	0.4693	0.1349
	PPV	0.6208	0.2137	0.6317	0.1178
	Sensibilidade	0.5142	0.1535	0.4382	0.2429
Região 2	Dice	0.7657	0.5584	0.7539	0.542
	PPV	0.7636	0.5922	0.7597	0.4572
	Sensibilidade	0.7872	0.5728	0.7694	0.7414
Região 3	Dice	0.0148	0.5543	0.045	0.2542
	PPV	0.1178	0.659	0.0898	0.757
	Sensibilidade	0.0118	0.5261	0.0414	0.1721
Região 4	Dice	0.7444	0.0644	0.7286	0.2064
	PPV	0.8086	0.2385	0.7652	0.1973
	Sensibilidade	0.7582	0.0856	0.7336	0.261
Tumor Completo (1+2+3+4)	Dice	0.8619	0.7929	0.8599	0.7943
	PPV	0.9084	0.9116	0.9101	0.8058
	Sensibilidade	0.8318	0.7437	0.8274	0.835
Tumor Central (1+3+4)	Dice	0.7208	0.6404	0.704	0.5188
	PPV	0.8649	0.7944	0.7969	0.7339
	Sensibilidade	0.6889	0.6001	0.681	0.4422

8.2 Características de Segunda ordem

Na tabela 8.5 tem-se a média das métricas de um classificador com o conjunto de características de 1ª ordem e um segundo classificador com as características de 2ª ordem.

Tabela 8.5: Média das métricas obtidas relativas as segmentações provenientes de um classificador com as características de 1ª ordem e um classificador com as características de 2ª ordem (no conjunto de *Training*).

		Características de 1ª ordem		Características de 2ª ordem	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.4931	0.1211	0.4662	0.0967
	PPV	0.6311	0.2138	0.639	0.2197
	Sensibilidade	0.4986	0.1565	0.4662	0.1110
Região 2	Dice	0.7622	0.5582	0.7585	0.5597
	PPV	0.7740	0.6013	0.7846	0.6135
	Sensibilidade	0.7717	0.5611	0.7546	0.5510
Região 3	Dice	0.0109	0.5342	0.0068	0.5419
	PPV	0.1096	0.6573	0.1330	0.6611
	Sensibilidade	0.0084	0.4977	0.0045	0.5079
Região 4	Dice	0.7470	0.0739	0.7593	0.0644
	PPV	0.8163	0.2380	0.8738	0.2415
	Sensibilidade	0.7560	0.0976	0.7476	0.0850
Tumor Completo (1+2+3+4)	Dice	0.8595	0.8008	0.8621	0.8071
	PPV	0.9219	0.9377	0.9487	0.9510
	Sensibilidade	0.8179	0.7360	0.8024	0.7323
Tumor Central (1+3+4)	Dice	0.7193	0.6246	0.7282	0.6222
	PPV	0.8745	0.8004	0.9178	0.8083
	Sensibilidade	0.6804	0.5786	0.6701	0.5698

8.3 Estudo do Classificador

Os resultados presentes na tabela 8.6 foram responsáveis por determinar qual o melhor classificador no âmbito da segmentação automática cerebral.

Tabela 8.6: Média das métricas obtidas relativas as segmentações provenientes de uma Floresta de Decisão e de uma *Extra-Trees*.

		Classificador RF		Classificador <i>Extra-Tree</i>	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.4962	0.1184	0.4894	0.1211
	PPV	0.616	0.2137	0.6272	0.2138
	Sensibilidade	0.5145	0.1535	0.4989	0.1565
Região 2	Dice	0.7489	0.5263	0.7525	0.5275
	PPV	0.727	0.5353	0.7447	0.5478
	Sensibilidade	0.791	0.5765	0.7782	0.5644
Região 3	Dice	0.0148	0.5509	0.0108	0.5315
	PPV	0.1157	0.6511	0.109	0.65
	Sensibilidade	0.0118	0.5261	0.0084	0.4978
Região 4	Dice	0.7407	0.0644	0.7422	0.0759
	PPV	0.746	0.2385	0.7503	0.2379
	Sensibilidade	0.7628	0.0856	0.7602	0.0987
Tumor Completo (1+2+3+4)	Dice	0.8474	0.7688	0.8501	0.7755
	PPV	0.8739	0.8422	0.8923	0.8723
	Sensibilidade	0.8362	0.7464	0.8233	0.7374
Tumor Central (1+3+4)	Dice	0.7173	0.6357	0.7156	0.6207
	PPV	0.8015	0.7802	0.8083	0.7873
	Sensibilidade	0.6921	0.6002	0.6834	0.5789

8.4 Configuração dos Híper-parâmetros do Extra-Trees

Na tabela 8.7 encontra-se a segunda fase do estudo da melhor configuração dos híper-parâmetros da Extra-Trees, que corresponde à avaliação do desempenho do primeiro e terceiro candidatos. Posteriormente na tabela 8.8 é comparado o desempenho do primeiro e candidato com o classificador com os híper-parâmetros em conformidade com o descrito em 6.3. Denote-se que em todas a variantes da metodologia de segmentação o conjunto de características é idêntico.

Tabela 8.7: Média das métricas obtidas relativas as segmentações provenientes de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica *Random-Search* e um classificador com os parâmetros do terceiro classificado (no conjunto de *Training*).

		Primeiro Candidato		Terceiro Candidato	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.5035	0.1152	0.4983	0.1104
	PPV	0.6603	0.2169	0.6489	0.2165
	Sensibilidade	0.5172	0.1453	0.5178	0.1362
Região 2	Dice	0.7655	0.5625	0.7651	0.5625
	PPV	0.7778	0.6009	0.7759	0.6024
	Sensibilidade	0.7733	0.5782	0.7749	0.5780
Região 3	Dice	0.0102	0.5420	0.011	0.5407
	PPV	0.1520	0.6553	0.1704	0.6512
	Sensibilidade	0.0077	0.5154	0.0082	0.5173
Região 4	Dice	0.7682	0.0845	0.7672	0.0767
	PPV	0.8702	0.2431	0.8704	0.2410
	Sensibilidade	0.7639	0.0984	0.7602	0.0914
Tumor Completo (1+2+3+4)	Dice	0.8712	0.8084	0.8711	0.8074
	PPV	0.9411	0.9202	0.9399	0.9171
	Sensibilidade	0.8238	0.7565	0.8247	0.7561
Tumor Central (1+3+4)	Dice	0.7391	0.6314	0.7380	0.6268
	PPV	0.9136	0.7984	0.9129	0.7928
	Sensibilidade	0.6886	0.5890	0.6871	0.5869

Tabela 8.8: Média das métricas obtidas relativas as segmentações provenientes de um classificador com os parâmetros obtidos pelo primeiro candidato da técnica *Random-Search* e o classificador com os parâmetros aplicados até ao momento, na secção 6.3 (no conjunto de *Training*).

		Primeiro Candidato		Classificador Atual	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.5035	0.1152	0.4662	0.0967
	PPV	0.6603	0.2169	0.639	0.2197
	Sensibilidade	0.5172	0.1453	0.4662	0.1110
Região 2	Dice	0.7655	0.5625	0.7585	0.5597
	PPV	0.7778	0.6009	0.7846	0.6135
	Sensibilidade	0.7733	0.5782	0.7546	0.5510
Região 3	Dice	0.0102	0.5420	0.0068	0.5419
	PPV	0.1520	0.6553	0.1330	0.6611
	Sensibilidade	0.0077	0.5154	0.0045	0.5079
Região 4	Dice	0.7682	0.0845	0.7593	0.0644
	PPV	0.8702	0.2431	0.8738	0.2415
	Sensibilidade	0.7639	0.0984	0.7476	0.0850
Tumor Completo (1+2+3+4)	Dice	0.8712	0.8084	0.8621	0.8071
	PPV	0.9411	0.9202	0.9487	0.9510
	Sensibilidade	0.8238	0.7565	0.8024	0.7323
Tumor Central (1+3+4)	Dice	0.7391	0.6314	0.7282	0.6222
	PPV	0.9136	0.7984	0.9178	0.8083
	Sensibilidade	0.6886	0.5890	0.6701	0.5698

8.5 Efeito da Amostragem

A realização da amostragem, com recurso a filtros interquartil, para gliomas de baixo grau é avaliada na tabela 8.9.

Tabela 8.9: Média das métricas, avaliadas pelo concurso, na ausência/presença da amostragem. A diferença entre estas duas abordagens permite quantificar o impacto da realização da amostragem.

		LGG (s/ amostragem)	LGG c/ amostragem	Diferença (c/ amostragem - s/ amostragem)
Região ativa (4)	Dice	0.0644	0.2038	0.1394
	PPV	0.2415	0.1968	-0.0447
	Sensibilidade	0.085	0.2579	0.1729
Região Central Tumoral (1+3+4)	Dice	0.8071	0.8157	0.0086
	PPV	0.951	0.947	-0.004
	Sensibilidade	0.7323	0.7436	0.0113
Região Tumoral Total (1+2+3+4)	Dice	0.6222	0.6848	0.0626
	PPV	0.8083	0.8265	0.0182
	Sensibilidade	0.5698	0.6418	0.072
				0.4363

8.6 Análise da Sequência T1

Na tabela 8.10 é avaliado o desempenho da metodologia de segmentação na presença e ausência da sequência T1, em ambos o tipos de glioma.

Tabela 8.10: Média das métricas obtidas relativas as segmentações de um Extra-Trees que contabiliza características da sequência T1 e um Extra-Trees que não considera o protocolo em estudo (no conjunto de *Training*).

		Extra-Trees com T1		Extra-Trees sem T1	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.4662	0.1531	0.4807	0.1721
	PPV	0.6390	0.1661	0.6718	0.1702
	Sensibilidade	0.4662	0.232	0.4814	0.2257
Região 2	Dice	0.7585	0.5638	0.7649	0.5616
	PPV	0.7846	0.6321	0.7854	0.5915
	Sensibilidade	0.7546	0.5443	0.7635	0.5721
Região 3	Dice	0.0068	0.4833	0.0155	0.4493
	PPV	0.133	0.6886	0.1867	0.6904
	Sensibilidade	0.0045	0.4229	0.0109	0.4003
Região 4	Dice	0.7593	0.2038	0.7627	0.2008
	PPV	0.8738	0.1968	0.8160	0.1913
	Sensibilidade	0.7476	0.2579	0.760	0.2616
Tumor Completo (1+2+3+4)	Dice	0.8621	0.8157	0.8691	0.8260
	PPV	0.9487	0.947	0.9458	0.9382
	Sensibilidade	0.8024	0.7436	0.8162	0.7627
Tumor Central (1+3+4)	Dice	0.7282	0.6848	0.7389	0.6091
	PPV	0.9178	0.8265	0.8640	0.7809
	Sensibilidade	0.6701	0.6418	0.6891	0.5947

8.7 Novos Parâmetros de Normalização

A definição de novos parâmetros de normalização, resultantes da libertação da base de dados BraTS2015, despoletou a necessidade de realização de um novo estudo, presente na tabela 8.11. Nesta são expostas as métricas relativas aos dois conjuntos de parâmetros de normalização de Nyúl-Udupa.

Tabela 8.11: Média das métricas obtidas das segmentações resultantes de duas etapas de normalização das intensidades distintas. Uma primeira onde a fase de treino da normalização considera o conjunto de treino BraTS2013 e a restante engloba um número mais alargado de pacientes (BraTS2015). Denote-se que ambos os resultados dizem respeito ao conjunto *Training*.

		Normalização BraTS2013		Normalização BraTS2015	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.4662	0.1531	0.4813	0.1507
	PPV	0.6390	0.1661	0.6706	0.1552
	Sensibilidade	0.4662	0.232	0.4716	0.2371
Região 2	Dice	0.7585	0.5638	0.7606	0.5555
	PPV	0.7846	0.6321	0.7816	0.6209
	Sensibilidade	0.7546	0.5443	0.7665	0.5460
Região 3	Dice	0.0068	0.4833	0.0131	0.4747
	PPV	0.1330	0.6886	0.2200	0.6907
	Sensibilidade	0.0045	0.4229	0.0099	0.4138
Região 4	Dice	0.7593	0.2038	0.7675	0.2060
	PPV	0.8738	0.1968	0.8829	0.1990
	Sensibilidade	0.7476	0.2579	0.7540	0.2582
Tumor Completo (1+2+3+4)	Dice	0.8621	0.8157	0.8666	0.8085
	PPV	0.9487	0.947	0.9479	0.9407
	Sensibilidade	0.8024	0.7436	0.8139	0.7443
Tumor Central (1+3+4)	Dice	0.7282	0.6848	0.7256	0.6864
	PPV	0.9178	0.8265	0.9351	0.8268
	Sensibilidade	0.6701	0.6418	0.6700	0.6438

8.8 Classificação em Dois Estágios

O desempenho da classificação em dois estágios, tal como a designação o indica, foi validado em duas fases. Uma primeira, presente na tabela 8.12, relativa à classificação binária. Posteriormente, nas tabelas 8.14 e 8.13 encontram-se as médias das três métricas avaliadas, em todas as classes e regiões tumorais na ausência e presença da T1, respetivamente.

Tabela 8.12: Média das métricas da região de tumor total resultantes da aplicação de um único estágio de classificação (utilizando todas as sequências) vs. classificação binária (na ausência da T1). Resultados relativos ao conjunto *Training*.

		Classificação Multi-Label		Classificação Binária	
		HGG	LGG	HGG	LGG
Tumor Completo (1+2+3+4)	Dice	0.8666	0.7699	0.8802	0.8162
	PPV	0.9479	0.8432	0.9355	0.8614
	Sensibilidade	0.8139	0.7458	0.8460	0.8090

Tabela 8.13: Média das métricas, em todas as classes, relativas as segmentações de um Extra-Trees *Multi-Label* vs. média das métricas da metodologia de classificação em dois estágios. Na metodologia em dois estágios é considerada a sequência T1 no segundo estágio. Denote-se que ambos os resultados dizem respeito ao conjunto *Training*.

		Classificação Multi-classe		Classificação Dois Estágios	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.4813	0.1528	0.5066	0.1455
	PPV	0.6706	0.1652	0.6646	0.1375
	Sensibilidade	0.4716	0.232	0.5067	0.2441
Região 2	Dice	0.7606	0.5257	0.7663	0.5578
	PPV	0.7816	0.5676	0.7660	0.5775
	Sensibilidade	0.7665	0.549	0.7946	0.6045
Região 3	Dice	0.0131	0.4813	0.064	0.4754
	PPV	0.2200	0.6809	0.1601	0.6777
	Sensibilidade	0.0099	0.4229	0.0122	0.4257
Região 4	Dice	0.7675	0.1976	0.7721	0.2020
	PPV	0.8829	0.1837	0.8094	0.1910
	Sensibilidade	0.7540	0.2581	0.7764	0.2579
Tumor Completo (1+2+3+4)	Dice	0.8666	0.7699	0.8802	0.8200
	PPV	0.9479	0.8432	0.9355	0.8709
	Sensibilidade	0.8139	0.7458	0.8460	0.8078
Tumor Central (1+3+4)	Dice	0.7256	0.6517	0.7392	0.6868
	PPV	0.9351	0.7595	0.8710	0.7896
	Sensibilidade	0.6700	0.6422	0.6947	0.6758

Tabela 8.14: Média das métricas, em todas as classes, relativas as segmentações de um Extra-Trees *Multi-Label* vs. média das métricas da metodologia de classificação em dois estágios. Na metodologia em dois estágios, em ambas as fases de classificação não é considerada a sequência T1. Denote-se que ambos os resultados dizem respeito ao conjunto *Training*.

		Classificação Multi-classe		Classificação Dois Estágios	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.4813	0.1528	0.5116	0.1636
	PPV	0.6706	0.1652	0.6815	0.1445
	Sensibilidade	0.4716	0.232	0.5139	0.2391
Região 2	Dice	0.7606	0.5257	0.7684	0.5399
	PPV	0.7816	0.5676	0.7697	0.5183
	Sensibilidade	0.7665	0.549	0.7931	0.6207
Região 3	Dice	0.0131	0.4813	0.0248	0.4448
	PPV	0.2200	0.6809	0.1719	0.6797
	Sensibilidade	0.0099	0.4229	0.0203	0.3973
Região 4	Dice	0.7675	0.1976	0.7720	0.1999
	PPV	0.8829	0.1837	0.8610	0.1916
	Sensibilidade	0.7540	0.2581	0.7741	0.2552
Tumor Completo (1+2+3+4)	Dice	0.8666	0.7699	0.8802	0.8200
	PPV	0.9479	0.8432	0.9355	0.8709
	Sensibilidade	0.8139	0.7458	0.8460	0.8078
Tumor Central (1+3+4)	Dice	0.7256	0.6517	0.7411	0.6039
	PPV	0.9351	0.7595	0.9184	0.7652
	Sensibilidade	0.6700	0.6422	0.6991	0.6039

8.9 Pós-processamento

Nesta última secção encontra-se, na tabela 8.15, a média das métricas perante a ausência e aplicação da primeira filosofia de pós-processamento. A segunda proposta do pós-processamento foi avaliada apenas para um paciente do conjunto *Challenge* cujos resultados estão presentes na tabela 8.16.

Tabela 8.15: Média das métricas obtidas relativas as segmentações sem pós-processamento e com pós-processamento.

		Sem pós-processamento		Com pós-processamento	
		HGG	LGG	HGG	LGG
Região 1	Dice	0.4820	0.1501	0.4813	0.1507
	PPV	0.6707	0.1537	0.6706	0.1552
	Sensibilidade	0.4723	0.2371	0.4716	0.2371
Região 2	Dice	0.7512	0.5276	0.7606	0.5555
	PPV	0.7512	0.5694	0.7606	0.6209
	Sensibilidade	0.7699	0.5509	0.7665	0.5460
Região 3	Dice	0.0131	0.4728	0.0131	0.4747
	PPV	0.2200	0.6828	0.2200	0.6927
	Sensibilidade	0.0099	0.4139	0.0099	0.4138
Região 4	Dice	0.7613	0.1998	0.7675	0.2060
	PPV	0.8041	0.1861	0.8829	0.1990
	Sensibilidade	0.7574	0.2584	0.7540	0.2582
Tumor Completo (1+2+3+4)	Dice	0.8589	0.7716	0.8666	0.8085
	PPV	0.9230	0.8455	0.9479	0.9407
	Sensibilidade	0.8175	0.7466	0.8139	0.7446
Tumor Central (1+3+4)	Dice	0.7209	0.6547	0.7256	0.6864
	PPV	0.8602	0.7627	0.9351	0.8268
	Sensibilidade	0.6731	0.6443	0.6700	0.6438

Tabela 8.16: Métricas, avaliadas pelo concurso, na ausência/presença da amostragem. A diferença entre estas duas abordagens permite quantificar o impacto da realização da amostragem.

		<i>HG_310</i> (pós-processamento inicial)	<i>HG_310</i> (pós-processamento final)	Diferença (inicial-final)
		Região ativa (4)	Dice	0.75
PPV	0.77		0.75	0.02
Sensibilidade	0.73		0.93	-0.23
Região Central Tumoral (1+3+4)	Dice	0.79	0.92	-0.13
	PPV	0.88	0.88	0.00
	Sensibilidade	0.72	0.97	-0.25
Região Tumoral Total (1+2+3+4)	Dice	0.73	0.89	-0.16
	PPV	0.77	0.83	-0.06
	Sensibilidade	0.69	0.96	-0.27
				-1.17

