UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

CAIO FERNANDES ARAÚJO

SEGMENTAÇÃO DE IMAGENS 3D UTILIZANDO COMBINAÇÃO DE IMAGENS 2D

Recife
2016

CAIO FERNANDES ARAÚJO

SEGMENTAÇÃO DE IMAGENS 3D UTILIZANDO COMBINAÇÃO DE IMAGENS 2D

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Tsang Ing Ren

Recife

2016

Catalogação na fonte
Bibliotecária Joana D'Arc Leão Salvador CRB 4-572

A663s Araújo, Caio Fernandes.
Segmentação de imagens 3D utilizando combinação de imagens 2D / Caio Fernandes Araújo. – 2016.
128 f.: fig., tab.

Orientador: Tsang Ing Ren.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIN. Ciência da Computação, Recife, 2016.
Inclui referências.

1. Ressonância magnética. 2. Imagens médicas. 3. Segmentação – 3D 4. Cérebro. I.Tsang, Ing Ren (Orientador). II. Titulo.

538.36 CDD (22. ed.) UFPE-MEI 2017-12

Caio Fernandes Araújo

Segmentação de imagens 3D utilizando combinação de imagens 2D

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação.

Aprovado em: 12/08/2016

BANCA EXAMINADORA

_____

Prof. Dr. George Darmiton da Cunha Cavalcanti
Centro de Informática / UFPE

_____

Prof. Dr. Jorge de Jesus Gomes Leandro
Motorola Mobility LLC

_____

Prof. Dr. Tsang Ing Ren
Centro de Informática / UFPE
(**Orientador**)

Dedico este trabalho aos meus familiares, que me apoiaram e incentivaram ao longo de todo o desenvolvimento do curso e desse trabalho e aos meus amigos que me motivaram, ajudaram e estiveram sempre presentes quando foi preciso.

AGRADECIMENTOS

Ao meu orientador, Professor Dr. Tsang Ing Ren, pelos ensinamentos, conselhos, apoio e confiança depositada em mim no decorrer do mestrado.

Aos meus familiares, meus pais, irmãos, primos, avós e tios, que me deram apoio, incentivo e ajuda quando foi necessário e sem eles eu não estaria aqui.

Aos meus amigos, que me ajudaram tanto com apoio quanto com ajuda e dicas para a realização de algumas das tarefas da dissertação.

À Universidade Federal de Pernambuco e ao Centro de Informática, por disponibilizar a estrutura necessária para a pesquisa, dia e noite.

À Capes, pelo apoio financeiro durante o mestrado.

RESUMO

Segmentar imagens de maneira automática é um grande desafio. Apesar do ser humano conseguir fazer essa distinção, em muitos casos, para um computador essa divisão pode não ser tão trivial. Vários aspectos têm de ser levados em consideração, que podem incluir cor, posição, vizinhanças, textura, entre outros. Esse desafio aumenta quando se passa a utilizar imagens médicas, como as ressonâncias magnéticas, pois essas, além de possuírem diferentes formatos dos órgãos em diferentes pessoas, possuem áreas em que a variação da intensidade dos pixels se mostra bastante sutil entre os vizinhos, o que dificulta a segmentação automática. Além disso, a variação citada não permite que haja um formato pré-definido em vários casos, pois as diferenças internas nos corpos dos pacientes, especialmente os que possuem alguma patologia, podem ser grandes demais para que se haja uma generalização. Mas justamente por esse possuírem esses problemas, são os principais focos dos profissionais que analisam as imagens médicas. Este trabalho visa, portanto, contribuir para a melhoria da segmentação dessas imagens médicas. Para isso, utiliza a ideia do *Bagging* de gerar diferentes imagens 2D para segmentar a partir de uma única imagem 3D, e conceitos de combinação de classificadores para uni-las, para assim conseguir resultados estatisticamente melhores, se comparados aos métodos populares de segmentação. Para se verificar a eficácia do método proposto, a segmentação das imagens foi feita utilizando quatro técnicas de segmentação diferentes, e seus resultados combinados. As técnicas escolhidas foram: binarização pelo método de Otsu, ~~o~~ *K-Means*, rede neural SOM e o modelo estatístico GMM. As imagens utilizadas nos experimentos foram imagens reais, de ressonâncias magnéticas do cérebro, e o intuito do trabalho foi segmentar a matéria cinza do cérebro. As imagens foram todas em 3D, e as segmentações foram feitas em fatias 2D da imagem original, que antes passa por uma fase de pré-processamento, onde há a extração do cérebro do crânio. Os resultados obtidos mostram que o método proposto se mostrou bem sucedido, uma vez que, em todas as técnicas utilizadas, houve uma melhoria na taxa de acerto da segmentação, comprovada através do teste estatístico T-Teste. Assim, o trabalho mostra que utilizar os princípios de combinação de classificadores em segmentações de imagens médicas pode apresentar resultados melhores.

Palavras-chave: Segmentação. 3D. Imagens médicas. Combinação de classificadores. Ressonância magnética. Cérebro. Matéria cinza.

ABSTRACT

Automatic image segmentation is still a great challenge today. Despite the human being able to make this distinction, in most of the cases easily and quickly, to a computer this task may not be that trivial. Several characteristics have to be taken into account by the computer, which may include color, position, neighborhoods, texture, among others. This challenge increases greatly when it comes to using medical images, like the MRI, as these besides producing images of organs with different formats in different people, have regions where the intensity variation of pixels is subtle between neighboring pixels, which complicates even more the automatic segmentation. Furthermore, the above mentioned variation does not allow a pre-defined format in various cases, because the internal differences between patients bodies, especially those with a pathology, may be too large to make a generalization. But specially for having this kind of problem, those people are the main targets of the professionals that analyze medical images. This work, therefore, tries to contribute to the segmentation of medical images. For this, it uses the idea of Bagging to generate different 2D images from a single 3D image, and combination of classifiers to unite them, to achieve statistically significant better results, if compared to popular segmentation methods. To verify the effectiveness of the proposed method, the segmentation of the images is performed using four different segmentation techniques, and their combined results. The chosen techniques are the binarization by the Otsu method, K-Means, the neural network SOM and the statistical model GMM. The images used in the experiments were real MRI of the brain, and the dissertation objective is to segment the gray matter (GM) of the brain. The images are all in 3D, and the segmentations are made using 2D slices of the original image that pass through a preprocessing stage before, where the brain is extracted from the skull. The results show that the proposed method is successful, since, in all the applied techniques, there is an improvement in the accuracy rate, proved by the statistical test T-Test. Thus, the work shows that using the principles of combination of classifiers in medical image segmentation can obtain better results.

Keywords: Segmentation. 3D. Medical images. Classifiers combination. MRI. Brain. Gray matter.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

Bagging	Bootstrap Aggregating
BET	Brain Extraction Tool
CSF	Cerebrospinal Fluid
DSC	Sørensen-Dice Coefficient
EM	Expectation-Maximization
FMRIB	Functional MRI of the Brain
FSL	FMRIB Software Library v5.0
GM	Gray Matter
GMM	Gaussian Mixture Model
MRI	Magnetic Resonance Imaging
NKI-RS	Nathan Kline Institute/Rockland sample
OASIS-TRT	Open access series of imaging studies test–retest (“reliability”) sample
SOFM	Self-Organizing Feature Maps
SOM	Self-Organizing Maps
SPM	Statistical Parametric Mapping
STAPLE	Simultaneous Truth and Performance Level Estimation
SUIT	Spatially Unbiased Atlas Template of the Cerebellum and Brainstem
UCL	University College London
WM	White Matter

SUMÁRIO

1 INTRODUÇÃO

Durante todo o processo de dispersão da população ao redor do mundo ao longo do tempo, junto com as evoluções culturais e contatos e conflitos interpopulacionais, houve mudanças da relação entre os homens e a natureza, seja animada ou inanimada. E cada uma dessas mudanças resultou no surgimento de doenças contagiosas novas ou não familiares [A. J. McMichael, 2004]. Visto isso, com o objetivo de sempre melhorar a qualidade de vida, várias pessoas se dedicam a pesquisar para descobrir mais sobre essas doenças. Contudo, à medida que o tempo passa e novas doenças são diagnosticadas, apenas o exame clínico de um enfermo passa a não ser mais suficiente para saber de que mal ele sofre. Então, surge a necessidade de se desenvolver equipamentos que possam fazer uma análise interna do paciente, gerando imagens de como o corpo está pelo lado de dentro, sem que seja necessário um procedimento intrusivo. Alguns dos métodos de se obter essas imagens internas mais conhecidos são os Raios X, o Ultrassom e as Ressonâncias Magnéticas.

Ao longo dos anos, cada vez mais equipamentos vêm sendo desenvolvidos e aperfeiçoados, de modo que os métodos que fazem as análises fiquem cada vez mais complexos e precisos. Isso pode ser facilmente visto em equipamentos que geram imagens. Basta ver a Figura 1 para se perceber a evolução do que os equipamentos podem produzir, mesmo utilizando os mesmos tipos de técnicas. A imagem à esquerda corresponde a uma das primeiras ressonâncias magnéticas, tendo sido feita na altura do abdômen, onde é bastante complicado para uma pessoa sem extremo conhecimento identificar regiões específicas, como o cólon, o duodeno, o fígado ou o pâncreas (todos presentes na imagem). A da direita, mostra uma imagem atual da mesma região.

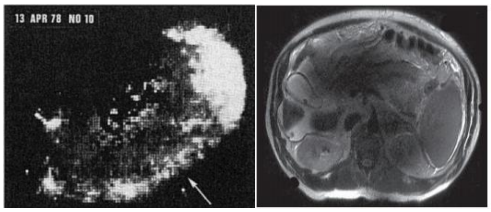

Figura 1 – Ressonâncias magnéticas do abdômen.
Fonte: [P. Mansfield and A. A. Maudsley, 1977], [D. Cornfeld, and J. Weinreb, 2008]

Os crescimentos do número de equipamentos e de exames a serem analisados [R. Smith-Bindman, D. L. Miglioretti, and E. B. Larson, 2016] não foram acompanhados proporcionalmente do crescimento do número de pessoas capacitadas a fazer diagnósticos através desses exames [HIS, 2016]. Desse modo, muitas vezes as pessoas não conseguem ter seus estados de saúde avaliados a tempo, pois a fila para que se consiga um atendimento médico é maior do que deveria [N. Y. Times, 2016], [FECOMERCIOSP, 2016], fora a grande quantidade de tempo que pode ser desperdiçada pelos profissionais analisando pessoas que não estão com problema algum de saúde, em detrimento de pessoas que realmente necessitam da análise. Desse modo, mais e mais técnicas para se automatizar ou pelo menos acelerar as análises dessas imagens têm sido desenvolvidas, de modo a não se perder tempo com quem esta saudável e, principalmente, levar a tratamento de maneira rápida quem necessita [A. Tanács, E. Máté, and A. Kuba, 2005].

As segmentações de imagens na área médica têm sido extremamente úteis, pois auxiliam nos diagnósticos dos profissionais, quando já não fazem o diagnóstico de maneira 100% automática. Elas permitem um processamento de um volume muito maior de imagens do mesmo espaço de tempo, se comparadas com um profissional, e ainda podem trazer à tona detalhes que poderiam passar despercebidos. Contudo, assim como os profissionais, as técnicas de segmentação e análise também possuem problemas, e não existe uma técnica que seja superior para qualquer tipo de análise [N. Sharma, and L. M. Aggarwal, 2016]. Desse modo, as pesquisas na área continuam, sempre visando aperfeiçoar ainda mais a segmentação, de modo que se possa ajudar as pessoas com cada vez mais eficiência.

1.1. Motivação

Os métodos de segmentação, para conseguirem bons resultados, devem ser tolerantes a vários tipos de variações nas imagens que serão analisadas. Caso contrário, seria bastante incomum encontrar algum objeto de interesse que estivesse exatamente com as mesmas intensidades de pixel (ou voxel) e posições em várias imagens diferentes, especialmente em imagens médicas. Estas variações vêm tanto da forma de aquisição das imagens, quanto pelas próprias diferenças intrínsecas aos objetos de interesse, em relação aos outros do mesmo tipo. Variações como altura, peso, sexo, entre várias outras, modificam de maneira intensa o

interior do corpo, levando a uma quantidade praticamente infinita de pequenas variações possíveis entre imagens.

Uma boa técnica de segmentação de imagens médicas deve, obrigatoriamente, conseguir levar em consideração as variações produzidas pelas imagens de diferentes pacientes, seja pela diferença natural entre as pessoas que estiverem sendo analisadas, seja pela diferença de equipamento para a captura das imagens (considerando que o tipo de imagem ainda continua sendo o mesmo). Apesar de, em alguns casos, ser uma tarefa razoavelmente simples para humanos, especialmente os bem treinados e com experiência na área, as técnicas de segmentação automáticas esbarram em problemas de difícil solução, como a análise de bordas. Além disso, muitas das técnicas desenvolvidas são financiadas por empresas, fazendo com que apenas os resultados sejam divulgados, sem que se possa replicar ou melhorar essas técnicas livremente.

Contudo, a pesquisa na área continua, e muitas técnicas têm sido apresentadas, cada uma com um tipo de abordagem diferente e um alvo específico a ser tratado, melhorando pontualmente cada segmentação [N. Sharma, and L. M. Aggarwal, 2016]. Assim, as técnicas têm se mostrado cada vez mais úteis para o auxílio de quem precisa.

Como o tempo dos profissionais que analisam esse tipo de imagem é muito valioso, é sempre interessante conseguir com que estes atendam o máximo de pessoas possíveis. Para isso, é preciso que se gaste menos tempo em pacientes saudáveis, e que esse tempo extra seja ocupado por pessoas que não estejam saudáveis. Desse modo, é extremamente interessante desenvolver modelos que gerem segmentações dessas imagens médicas, que farão com que pelo menos uma parte dos pacientes saudáveis não precise ser avaliado por um profissional, viabilizando uma celeridade maior, seja por conseguir melhores resultados e demonstrar esses resultados, seja por conseguir obter resultados semelhantes a outras técnicas, mas com um ganho de tempo para a geração desses resultados. Com bons desempenhos, mais pessoas serão diagnosticadas com algum problema de maneira antecipada e menos pessoas serão diagnosticadas incorretamente, com o verdadeiro negativo, onde se encontra algo quando não há nada ou o pior caso, o falso positivo, que corresponde a não encontrar nada de anormal quando realmente há algum problema.

1.2. Descrição do Problema

Por ser necessária a análise de inúmeras imagens médicas com alta precisão, a observação dessas imagens se torna, mais necessária. Desse modo, a cada dia mais e mais imagens chegam às mãos dos profissionais, que, por estarem em número limitado, não são capazes de analisar todas essas imagens, levando ao ponto de que é preciso ter um meio de diminuir essa quantidade de imagens a ser analisadas. Contudo, diferentemente dos seres humanos, as técnicas de segmentação de imagens não conseguem ter essa facilidade para diferenciar algumas áreas e tecidos destas imagens médicas, seja por limitação das técnicas, seja por causa das próprias imagens, que podem estar com baixa resolução ou com uma grande quantidade de ruídos. Com a segmentação automática é possível diminuir o tempo dispendido em analisar essas imagens, já que é possível analisar uma quantidade muito superior de imagens na mesma quantidade de tempo, seja por encontrar anomalias, seja por simplesmente evitar que um profissional perca tempo analisando uma imagem de uma pessoa saudável. Assim, a automatização da análise dessas imagens, via segmentação, visa então o auxílio dos profissionais da área médica, e não a sua substituição.

Para realizar a segmentação da imagem médica, é preciso primeiro que se obtenha a imagem, de preferência, de alta resolução, pois, quanto maior a qualidade da imagem, mais fácil de evitar alguns erros no momento da segmentação. Neste trabalho, foram escolhidas as imagens de ressonância magnética do cérebro para a realização dos testes, como mostrado na Figura 2, pois trata-se de uma das áreas mais importantes do corpo.

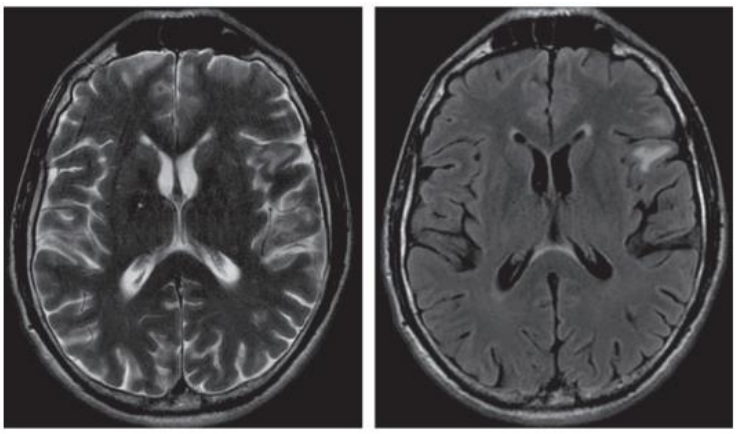

Figura 2– Ressonância magnética de um cérebro humano.

Fonte: [R. Bitar, 2006]

O cérebro é o principal órgão do sistema nervoso do corpo humano e possui uma estrutura incrivelmente complexa. Ele é uma massa de tecido esponjosa, delicada e não substituível, que fica em uma região protegida pelo crânio, que impede que este sofra alguns danos, mas também dificulta o estudo sobre seu funcionamento, tanto da saúde quanto na doença [S. K. Bandhyopadhyay and T. U. Paul, 2013]. Contudo, o cérebro pode ser acometido por problemas que afetam seu comportamento natural, como uma lesão formada pelo crescimento desenfreado de células, o tumor [W.-C. Lin, E. C.-K. Tsao, and C.-T. Chen, 1992]. Além desse crescimento, os tumores também podem destruir as células saudáveis do órgão, devido à pressão causada por esse crescimento, tanto contra o cérebro, quanto contra o crânio.

Ao longo dos últimos 20 anos, a incidência média de cânceres, incluindo o de cérebro, tem aumentado mais de 10% por ano, de acordo com o *National Cancer Institute* [N. C. Institute, 2015]. O *National Brain Tumor Foundation* levantou que, apenas nos Estados Unidos, cerca de 29.000 pessoas são diagnosticadas com um tumor no cérebro a cada ano, das quais aproximadamente 13.000 vêm a óbito [D. Bhattacharyya and T.-h. Kim, 2011]. Em crianças, tumores no cérebro correspondem a cerca de 25% das mortes provocadas por câncer.

O tumor cerebral causa um crescimento anormal das células do cérebro. As células arteriais que alimentam o cérebro são fortemente ligadas, o que torna testes laboratoriais inadequados para uma análise da química do cérebro. Desse modo, profissionais da área utilizam as Ressonâncias Magnéticas para o estudo do cérebro de maneira não invasiva [M. Sonka, S. K. Tadikonda, and S. M. Collins, 1996]. A intensidade da Ressonância Magnética depende da densidade de prótons de moléculas de àgua presentes no cérebro, que formam as imagens. Já a segmentação da imagem do cérebro é feita isolando a matéria cinza (*Gray Matter* - GM), isolando a matéria branca (*White Matter* - WM) e o Líquido Cefalorraquidiano (*Cerebrospinal Fluid* - CSF). Uma segmentação correta das imagens é muito importante, já que anomalias podem ser detectadas no processamento, o que poderia levar a uma economia muito grande de tempo dos profissionais que fazem essa análise, pois eles olhariam apenas as imagens que apresentassem alguma deformação inesperada. Contudo, essa segmentação automática pode ser bastante trabalhosa, devido ao baixo contraste entre as regiões, levando a segmentações com altos índices de erro. O trabalho, portanto, visa desenvolver um novo método para realizar essa segmentação, utilizando das características tridimensionais dessas

imagens de ressonâncias magnéticas, para atingir resultados expressivos, mas de maneira menos custosa computacionalmente que outras técnicas de segmentação.

1.3. Abordagem Proposta

O cérebro é um sistema que possui diversas áreas, tecidos, vasos diferentes, de modo que sua análise é bastante complexa. Várias técnicas diferentes de segmentação são utilizadas, onde cada uma obtém resultados variados dependendo do tipo de imagem escolhida. Para um modelo de segmentação realizar o que lhe é designado, cada algoritmo leva em consideração características diferentes, como a cor dos pixels, as cores dos pixels nas vizinhanças, formas geométricas ou formas pré-determinadas, entre outras. Cada característica é importante, de modo que cada técnica de segmentação pode ser melhor para diferentes casos, não havendo uma que seja melhor em todos os casos. Para que não haja um pensamento tendencioso em favor de alguma técnica, foi escolhida uma abordagem que visa utilizar conhecimentos empregados em diversas áreas, incluindo a segmentação de imagens, para que haja uma melhoria, independentemente da técnica escolhida. O objetivo é fazer com que a união dos resultados combinados aumente a quantidade de acertos, sem que haja uma necessidade de implementação de algoritmos mais elaborados ou custosos.

O método desenvolvido utiliza conceitos de reconhecimento de padrões e combinação de classificadores, com a combinação final dos resultados sendo feito através do voto majoritário. As imagens 3D são fatiadas de maneiras diferentes para gerar várias amostras para segmentação diferentes, que são imagens 2D, baseado no modo de funcionamento do *Bagging*, para que se tenha uma quantidade maior de resultados de segmentação, independentemente da técnica de segmentação escolhida. O *Bagging* é uma técnica que visa aumentar a diversidade das amostras de entrada de um classificador, para que se diversifique também suas saídas para posterior combinação ou seleção de resultados [L. Breiman, 1994]. As segmentações são feitas individualmente em cada fatia 2D e posteriormente unidas, para formar novamente uma imagem 3D. Por fim, as imagens 3D segmentadas são combinadas, através do voto majoritário, e o resultado final é analisado para verificar se há um ganho em relação à segmentação fatiando apenas uma vez.

A combinação de classificadores é uma técnica muito útil em diversos problemas, incluindo na segmentação de imagens. Ela pode ser usada em diversos problemas,

especialmente em casos onde apenas um único classificador não funciona bem, e tem tido um reconhecimento cada vez maior, além de ter sua utilização cada vez mais difundida atualmente [L. I. Kuncheva, 2004]. A combinação de classificadores, no trabalho, vai ser utilizada no início, utilizando o conceito de uma de suas técnicas, para conseguir uma maior diversidade das amostras, e também no final, onde se combinam os resultados, utilizando o voto majoritário, mas podendo se utilizar outras técnicas mais elaboradas posteriormente.

Já na segmentação, a ideia é utilizar técnicas com conceitos diferentes, para que se teste a robustez do método proposto. Assim, foram escolhidas técnicas que funcionam de maneiras distintas: a binarização, mais precisamente utilizando o método de Otsu, pois é um dos modos mais simples de segmentação. Depois foi utilizado o método *K-Means*, que se utiliza do valor das amostras de entrada e sua distância Euclidiana até os centroides dos grupos, onde cada grupo representa uma das classes desejadas, para que se defina os limiares dos valores para a segmentação. O terceiro método escolhido foi o SOM, uma rede neural artificial que se utiliza de cada amostra individualmente para que se atualize tanto o limiar atual quanto os limiares mais próximos, onde estes também servirão para definir as classes de cada pixel. Por fim, o GMM, que se utiliza de técnicas estatísticas para obter valores ótimos para os limiares citados anteriormente. Como as técnicas de segmentação são bastante distintas, acredita-se que, caso haja um aumento estatisticamente relevante na segmentação de todas estas, o método terá cumprido seu objetivo, podendo vir a ser testado posteriormente mesmo com técnicas mais elaboradas.

1.4. Objetivos

Partindo da hipótese que a segmentação é um dos passos mais importantes para a análise de uma imagem médica [A. Norouzi, M. Rahim, et al., 2014], essa dissertação tem como objetivos principais propor um método que melhore, de maneira geral, a segmentação de imagens médicas tridimensionais, e avaliar se realmente houve um ganho no desempenho, fazendo a verificação de maneira estatística, utilizando bases de dados reais, para que haja uma precisão maior dos resultados coletados. Pontualmente, os principais objetivos a serem alcançados nessa dissertação são:

- Gerar um novo modo de se analisar uma imagem 3D, utilizando uma imagem apenas para gerar diversas novas entradas, que serão fatias 2D da imagem original;

- Segmentar essas fatias utilizando técnicas já conhecidas e utilizadas na área;
- Utilizar técnicas de combinação de classificadores para juntar os resultados apresentados e obter um novo resultado para a segmentação.

1.5. Visão geral da solução proposta

De forma a limitar o escopo, foram avaliadas nesse trabalho apenas técnicas muito conhecidas e utilizadas para a segmentação de imagens, além de ter sido utilizado apenas um método para a combinação dos resultados gerados. As imagens inicialmente passaram por um pré-processamento, depois foram segmentadas de maneiras diferentes utilizando técnicas diferentes, uma por vez, os resultados foram combinados e suas respostas foram analisadas estatisticamente, para verificar a qualidade da segmentação. A Figura 3 mostra o processo descrito.

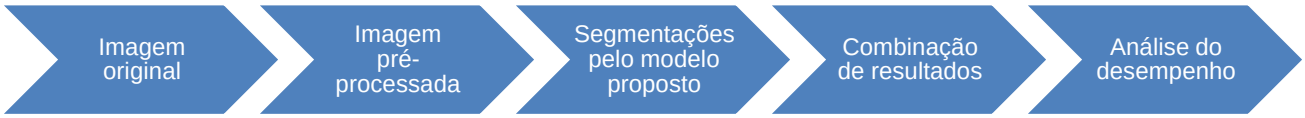

Figura 3– Processo simplificado da solução proposta da segmentação.

1.6. Escopo

Esta dissertação tem como escopo:

- Utilização de imagens reais, sem adição de ruídos ou criação de imagens que possuam características desejadas de antemão, como imagens que contenham apenas o cérebro já extraído do crânio. As imagens são pré-processadas de maneira igual, e podem possuir qualquer tamanho, contanto que sejam tridimensionais e na forma de um paralelepípedo, pois imagens com tamanhos variáveis exigem uma conversão para análise e posterior arredondamento de valores.
- Avaliar os métodos utilizados na dissertação: nos experimentos, somente foram usadas as técnicas de segmentação descritas no Capítulo 3 (binarização pelo método de Otsu, *K-Means*, a rede neural SOM e o GMM) e o método de combinação, que foi o voto majoritário.

1.7. Estrutura da dissertação

Esta dissertação está disposta em seis capítulos, que estão organizados da seguinte maneira:

- Capítulo 2: Levantamento de métodos muito utilizados atualmente para a segmentação de imagens médicas.
- Capítulo 3: Descreve os fundamentos necessários, as técnicas utilizadas na segmentação e uma descrição do modelo proposto.
- Capítulo 4: Descreve a metodologia utilizada nos testes, desde a validação dos algoritmos, a base de dados, o pré-processamento, os experimentos de como cada técnica se comportou na segmentação e uma análise estatística para verificação do desempenho do método proposto.
- Capítulo 5: Apresenta considerações em relação ao trabalho desenvolvido e sugestões para trabalhos futuros.

2. TRABALHOS RELACIONADOS

Neste capítulo, apresentamos uma visão geral do estado da arte das pesquisas que foram realizadas sobre a segmentação de imagens médicas, utilizando quaisquer técnicas e para imagens tanto 2D quanto 3D. Primeiramente, serão mostrados os tipos de imagens médicas utilizadas na segmentação mais comuns atualmente. Posteriormente, serão mostradas algumas das técnicas mais conhecidas e utilizadas para a segmentação de imagens.

2.1. Métodos de Obtenção de Imagens Médicas Utilizadas na Segmentação

Com o surgimento de doenças novas ao longo dos anos de existência da humanidade, o exame clínico de um enfermo passa a não ser mais suficiente para saber de que mal ele sofre. Deste problema, surge a necessidade de se desenvolver equipamentos que possam fazer uma análise interna do paciente, gerando imagens internas do corpo, sem que seja necessário um procedimento intrusivo. Algumas das imagens médicas mais conhecidas são obtidas por raios X, o ultrassom e as ressonâncias magnéticas.

2.1.1. Raios X

Uma das primeiras técnicas usadas para gerar imagens internas do corpo foram os raios X. Os raios X foram descobertos pelo físico alemão Wilhelm Conrad Röntgen, em torno do ano de 1895 [E. P. Bertin, 2012], tendo rendido a ele o primeiro Prêmio Nobel de Física, em 1901. Após fazer um experimento deixando os raios X passarem por 15 minutos através da mão de sua esposa com uma chapa fotográfica do outro lado, Röntgen verificou as sombras dos ossos de sua esposa nessa chapa, na primeira radiografia da história [A. Assmus, 1995], como mostrado na Figura 4. Como certa quantidade dos raios era absorvida pela mão, sendo diferente dependendo da densidade e composição da seção, foi provida uma representação 2D das estruturas internas da mão. Röntgen, então, apresentou seu achado em 1896 e, no mesmo ano, a novidade foi adotada na medicina. A partir de então, seria possível visualizar ossos quebrados e órgãos doentes dentro do corpo humano, fazendo com que os raios X rapidamente fossem utilizados no tratamento do câncer e para estudos anatômicos. Por serem radioativos, os raios X não podem ser utilizados a todo instante, pois podem causar efeitos colaterais nos usuários. Além disso, também existem limites, uma vez que radiografia comum

nunca foi eficiente para visualização de tecidos moles (como o fígado, os intestinos ou o cérebro), pois estes deixam a radiação passar quase completamente não criando, portanto, bons contrastes. A visualização destes órgãos só foi possível com o surgimento da tomografia computadorizada, em 1979, que é uma evolução dos raios X. Para obtenção destas imagens, o paciente fica no interior de um grande anel, que emite e capta a radiação de muitos ângulos diferentes, gerando um resultado equivalente a cerca de 130.000 raios X.

Figura 4 - Primeira radiografia já feita e uma atual.
Fonte: [NobelPrize, 2016]

2.1.2. Ultrassom

Outro tipo de imagem criado para o auxílio da análise interna do ser humano foi o Ultrassom. Apesar das ondas de *Ultra High Frequency* já terem sido descobertas desde meados do século XIX, apenas nos anos 1930 elas passaram a ser usadas para gerar imagens médicas. Dois irmãos austríacos, Karl Theo e Friedrich Dussik, neurologista e físico respectivamente, perceberam que, ao utilizar o ultrassom, era possível verificar diferenças na quantidade de energia transmitida através de algum órgão. Desse modo, seria possível criar um padrão que representasse a forma desse órgão [E. Yoxen, 1987], como mostrado na Figura 5. Apesar das dificuldades encontradas na época, tanto pela tecnologia rudimentar, quanto pela Segunda Guerra Mundial, Dussik pretendia encontrar anormalidades na forma dos ventrículos do cérebro sem a necessidade de usar raios X.

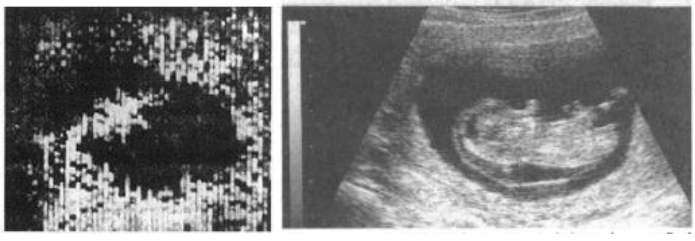

Figura 5 – Diferença entre uma Imagem de Ultrassom feita por Karl Dussik e uma atual.
Fonte: [K. H. Nicolaides et al.,1992]

2.1.3. Ressonância Magnética

Outra técnica de aquisição de imagens muito utilizada atualmente é a imagem por ressonância magnética (*Magnetic Resonance Imaging* – MRI). Criada em 1973 por um químico dos Estados Unidos chamado Paul Christian Lauterbur [P. C. Lauterbur, 1973], esta técnica visa investigar a anatomia e a fisiologia do corpo. Os scanners de MRI utilizam fortes campos magnéticos e ondas de rádio para gerar as imagens, mas sem a exposição causada pela radiação ionizante, presente nos raios X. Para isso, foi utilizado um campo magnético com comportamento espacial já conhecido para que se pudesse codificar o sinal recebido pela ressonância magnética. Desta forma foram obtidas as primeiras imagens, como mostrado na Figura 6, utilizando o método de *back projection*, um sistema de projeção de uma imagem por trás de uma placa translúcida para resultar num plano em combinação com outra imagem sobreposta, já utilizado na tomografia de raios X. As imagens de ressonâncias magnéticas foram selecionadas para a realização do trabalho, pela possibilidade de se analisar a cabeça, além de ser possível encontrar imagens em alta definição, o que ajuda na segmentação.

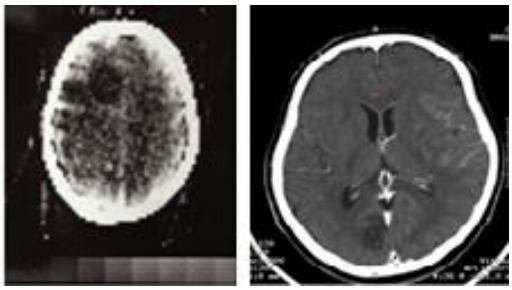

Figura 6 – Diferença entre uma ressonância magnética dos anos 70 e uma atual.
Fonte: [L. Lança, 2012]

2.2. Técnicas de Segmentação

Uma segmentação eficiente de imagens médicas depende da aparência visual e da forma das estruturas, para que se consiga discriminar melhor entre essas estruturas e o fundo da imagem, assim como conhecimento prévio sobre características e formas para um conjunto de imagens [A. Elnakib et al., 2011]. Os principais meios de segmentação atuais são listados a seguir.

2.2.1. Características de Aparência

Usualmente, aparências visuais de áreas desejadas a ser segmentadas estão associadas com a intensidade individual dos pixels (ou voxels), e com as interações espaciais entre as intensidades, geralmente com seus vizinhos.

- Intensidade: intensidades individuais podem guiar a segmentação caso seus valores, o do objeto de interesse e o do fundo, sejam diferentes por uma distância considerável. Desse modo, pelo menos a maior parte dos pixels pode ser separada facilmente apenas com a comparação das intensidades, como o uso de um ou mais limiares. No entanto, usualmente, características discriminantes mais poderosas têm de ser utilizadas.
- Interação Espacial: A aparência de certas texturas pode ser associada com padrões espaciais de variações de intensidade em pixels locais, ou probabilidades empíricas de intensidades, chamadas de matrizes de co-ocorrência. Modelos de interação probabilística consideram as imagens como amostras de um campo aleatório de intensidades independentes, especificadas pelas suas distribuições de probabilidades.

Tentativas iniciais foram focadas em intensidade das bordas, pelo fato de que bordas de texturas finas têm uma densidade espacial maior do que as bordas de texturas grosseiras [K. I. Laws, 1980]. Foram usados filtros lineares com inibição não linear por [J. Malik and P. Perona, 1990] para computar as diferenças entre dois *offsets* de gaussianas para modelar a percepção humana visual de textura. Filtros no domínio da frequência assumem que o sistema visual humano decompõe uma imagem para análise de textura em componentes de frequência orientadas [F. W. Campbell and J. Robson, 1968]. Filtros multibanda foram utilizados por [J. M. Coggins and A. K. Jain, 1985], para classificar texturas tanto naturais quanto sintéticas. Já modelos fractais foram introduzidas nos anos 1970, por [B. B. Mandelbrot, D. E. Passoja, and A. J. Paullay, 1984], e explorados posteriormente por [M. F. Barnsley, 2014], [M. Haindl, 1993]. Esses modelos são independentes de escalas, e são capazes de modelar texturas naturais, como nuvens, folhas ou rios, assim como imagens médicas [H. Li, K. R. Liu, and S.-C. Lo, 1997]. Uma imagem pode ser segmentada de acordo com um ou mais parâmetros fractais associados com a região de interesse. Há também os modelos probabilísticos gaussianos, como mostrado em [R. C. Dubes and A. K. Jain, 1989], onde é assumido que sinais contínuos de imagens estão em uma extensão infinita, mesmo imagens digitais tendo

uma quantidade limitada de intensidades. A distribuição de probabilidade depende de dois parâmetros, a imagem média de tamanho XY e a matriz de covariância simétrica de tamanho $(XY)^2$, para serem aprendidas das imagens de treinamento.

2.2.2. Características de Forma

As características espaciais (interações espaciais e de intensidade) podem falhar, se utilizadas de maneira isolada, caso uma imagem possua, por exemplo, ruído ou baixa resolução, além de fronteiras difusas. Estruturas médicas muitas vezes possuem formas bem restritas, dentro de uma família de formas [M. Rousson and N. Paragios, 2002]. Desse modo, suas segmentações podem ser realçadas ao se incorporar modelos probabilísticos de forma, especificando uma forma média e a variação possível para um certo objeto. A forma inicial é estimada através de uma base de dados de treinamento de imagens alinhadas de um objeto. A Figura 7 mostra os passos básicos para reconstrução de uma forma específica em quatro imagens de ressonância magnética da cabeça. São selecionadas quatro imagens para treinamento, onde é feita uma segmentação da área de interesse e posteriormente é calculada a imagem média dessas quatro áreas, formando a imagem inicial para as segmentações futuras.

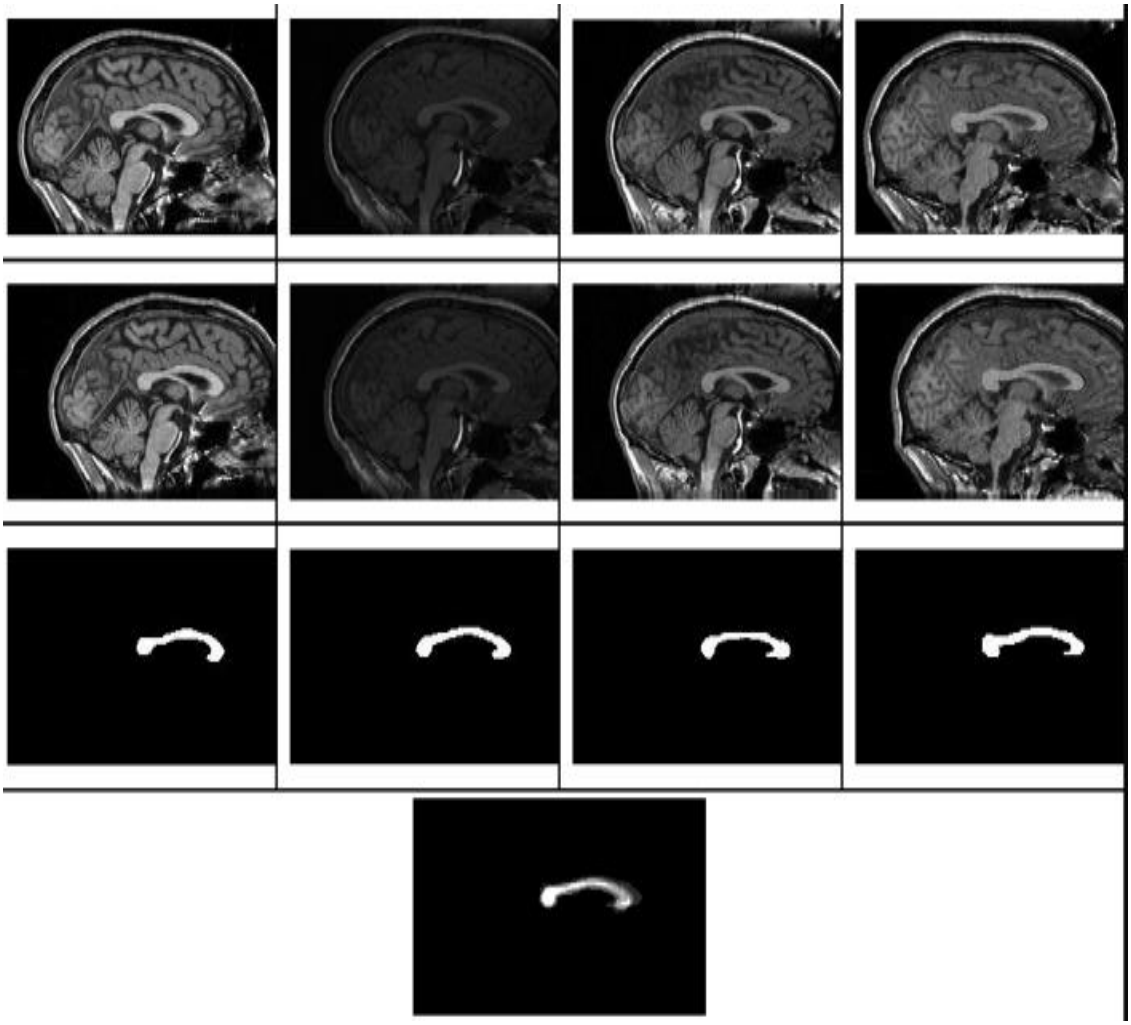

Figura 7 – Quatro imagens de cérebros com a reconstruçao de uma forma especifica e a sua imagem média.
Fonte: [A. Elnakib et al., 2011] p. 8

2.3. Técnicas de Segmentação

Altas complexidades e variabilidades de aparências e formas de estruturas anatômicas fazem com que a segmentação de imagens médicas seja uma das tarefas mais desafiadoras e essenciais em qualquer análise com assistência de um computador, pois cada alteração citada pode acarretar em novos erros. Devido à diversidade de objetos de interesse, modalidades de imagens, e limitações das técnicas computacionais, não existe uma técnica de segmentação universal, como mostrado no teorema *no free lunch* [D. H. Wolpert and W. G. Macready, 1995]. Algumas das técnicas mais populares para esse tipo de segmentação são as baseadas em regras, estatísticas, baseadas em atlas e modelos deformáveis. A seguir são descritas suas vantagens e desvantagens.

2.3.1. Segmentação Baseada em Regras

Nesse caso, características da imagem em uma região específica obedecem um conjunto de regras heurísticas. Limiarização simples é vastamente utilizada, devido à sua simplicidade e velocidade, para uma segmentação inicial ou em um estágio intermediário, mas muitas vezes não obtém um bom resultado quando usada de maneira individual. A limiarização mais simples divide a imagem em duas regiões, que correspondem ao objeto de interesse e ao fundo. O limiar pode ser fixo em toda a imagem (limiarização global) ou variar de acordo com a localização (limiarização local ou adaptativa). Normalmente, o valor do limiar não é escolhido ao acaso, mas sim após uma análise específica sobre a imagem a ser segmentada [C. Li and P. K.-S. Tam, 1998]. A Figura 8 mostra como ficam essas limiarizações em alguns exemplos.

Caso as distribuições de intensidades para o objeto e o fundo interceptem-se, mesmo que parcialmente, uma comparação com um limiar global não irá funcionar. Além disso, a limiarização não garante a conectividade dos objetos de interesse, que é um requerimento básico em várias imagens médicas [M. Sezgin et al., 2004], [O. Wirjadi, 2007], [R. M. Haralick and L. G. Shapiro, 1985]. Algumas das técnicas mais conhecidas são a binarização pelo método de Otsu, que utiliza o histograma para o cálculo do limiar, o *Region Growing* [L. G. Shapiro, 1985], que garante a conectividade de uma região para cada objeto segmentado de acordo com uma série de critérios, mostrado na Figura 9, ou o *Region Split-and-merge* [S. Hojjatoleslami and F. Kruggel, 2001], [S.-Y. Wan and W. E. Higgins, 2003], que particiona uma imagem em um número de regiões e depois iterativamente as une de acordo com uma regra de homogeneidade.

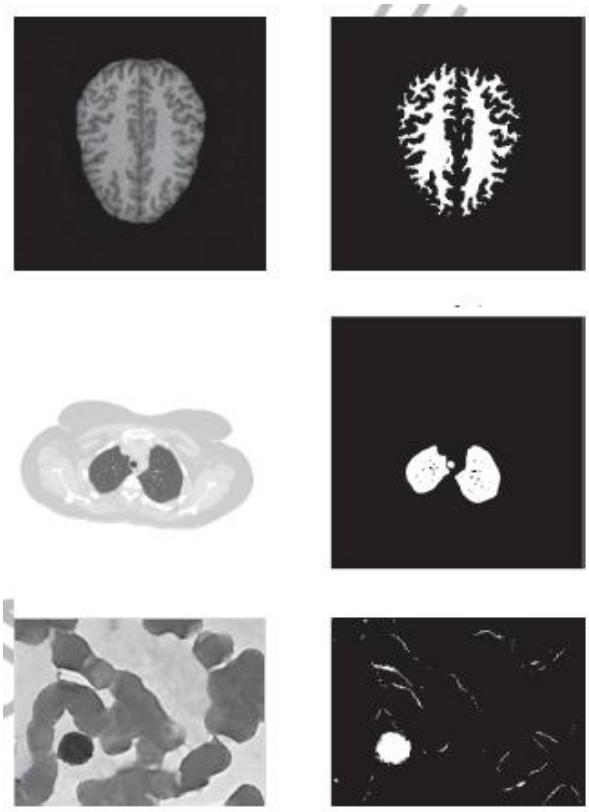

Figura 8 – Limiarizações de imagens médicas, do cérebro, pulmões e células.

Fonte: [S. K. Bandhyopadhyay and T. U. Paul, 2013] pp. 13

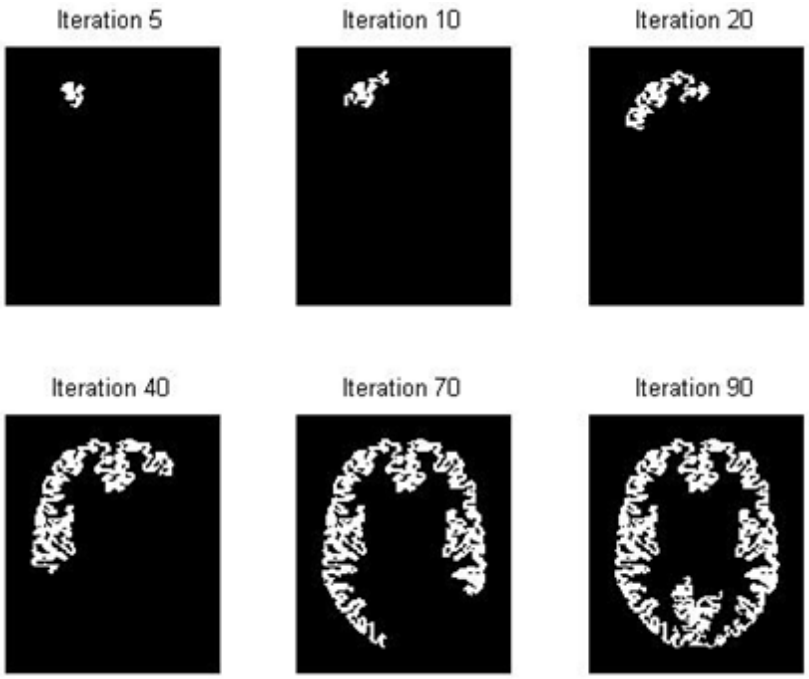

Figura 9 – Segmentação por *Region Growing*, com semente definida pelo usuário.

Fonte: [S. Franiatte, 2015]

2.3.2. Segmentação por Interferência Estatística Ótima

Segmentação de imagens usando técnicas estatísticas envolve modelos probabilísticos paramétricos e não paramétricos para aparência e forma dos objetos desejados, além de inferências ótimas, como, por exemplo, *expectation maximizaton* [J. C. Bezdek, L. Hall, and L. Clarke, 1992], [P. E. Hart, D. G. Stork, and R. O. Duda, 2001], [C. M. Bishop, 2006]. Modelos de densidade de probabilidade não paramétricos mais comuns utilizam o *k-nearest neighbor* (K-NN) e a janela de Parzen para suas estimações [D. W. Scott, 2015]. Já os modelos paramétricos mais comuns exploram representações analíticas que permitam aprendizagem de parâmetros numéricos razoáveis computacionalmente. Em particular, parâmetros de um modelo de misturas gaussianas (GMM) são aprendidas parte numericamente e parte analiticamente com o *expectation maximizaton* [P. E. Hart, D. G. Stork, and R. O. Duda, 2001], [C. M. Bishop, 2006], [J. Friedman, T. Hastie, and R. Tibshirani, 2001]. A Figura 10 mostra os resultados da segmentação de tecidos pulmonares de uma tomografia, usando distribuições de probabilidades unidimensionais como modelos de região. A distribuição das intensidades dos pixels na imagem são aproximadas e separadas em probabilidades do peito e dos pulmões, usando tanto o GMM quanto o *expectation maximizaton* para estimar seus parâmetros, e a janela de Parzen para estimar as distribuições.

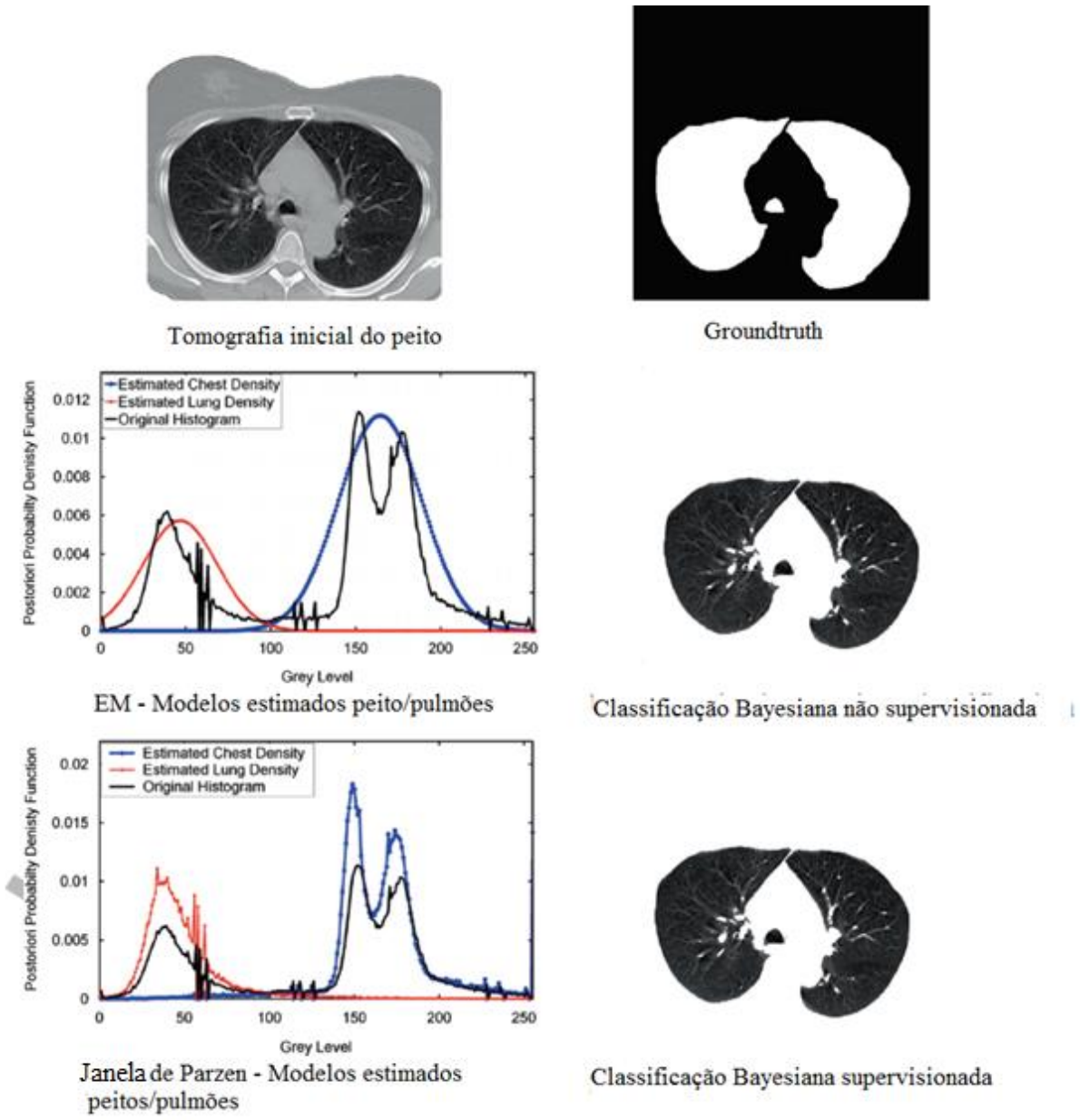

Figura 10 – Segmentação de pulmões usando modelos de densidade de probabilidade diferentes, o *expectation maximizaton* e a janela de Parzen.

Fonte: < [A. Elnakib et al., 2011] pp. 14>.

2.3.3. Segmentação Baseadas em Atlas

O uso de atlas anatômicos como imagens de referência para guiar segmentações de imagens novas é bastante popular em várias aplicações médicas, como para segmentar pulmões, corações ou órgãos abdominais [T. Rohlfing and C. R. Maurer, 2007], [S. L. Hartmann et al., 1999], [J. L. Marroquín et al., 2002], [B. Li et al., 2003], como mostrado na

Figura 11. Os atlas tipicamente retratam localizações e formas de estruturas anatômicas junto com suas relações espaciais [T. Rohlfing et al., 2005]. Os métodos conhecidos baseados em atlas podem ser classificados em segmentação baseada em atlas simples ou multi atlas. O método simples utiliza apenas um atlas, que é construído de uma ou mais imagens segmentadas rotuladas previamente. Uma vez que o atlas é criado, ele é registrado na imagem alvo, e a região de interesse é obtida pela chamada propagação de rótulos, que transfere os rótulos do atlas até a imagem, usando algum mapeamento geométrico. Obviamente, o desempenho da segmentação depende do mapeamento, além de ser a parte mais custosa computacionalmente. Uma imagem única para construir o atlas pode ser selecionada aleatoriamente [I. Isgum et al., 2009]. Caso o atlas seja construído por várias imagens, uma imagem pode ser selecionada como referência e todas as outras são registradas a ela. Com todas as imagens registradas, é calculada alguma média, e a imagem segmentada média é usada como o atlas [J. Stancanello et al., 2006].

A segmentação baseada em multi atlas registra vários atlas construídos independentemente, e então combina seus rótulos. A ideia inicial é que combinar vários classificadores independentes pode produzir classificações melhores [J. Kittler et al., 1998]. Em uma estratégia muito popular, o STAPLE (*simultaneous truth and performance level estimation*), o desempenho de cada classificador é avaliado iterativamente, e usa o EM para encontrar a melhor segmentação final [T. Rohlfing, D. B. Russakoff, and C. R. Maurer, 2004].

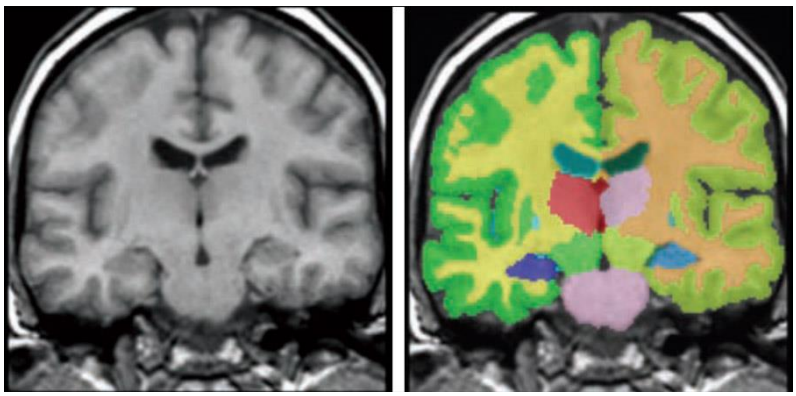

Figura 11 – Segmentação do cérebro usando Atlas.

Fonte: [K. R. Rao, 2015] p. 7

2.3.4. Segmentação com Modelos Deformáveis

Desde a publicação do artigo [M. Kass, A. Witkin, and D. Terzopoulos, 1988], modelos deformáveis se tornaram uma das técnicas dominantes e de maior sucesso na área de segmentação de imagens, detecção de contornos e rastreamento de estruturas em imagens médicas. Um modelo deformável, ou ativo, é uma curva em uma imagem 2D ou superfície 3D que evolui até contornar um objeto de interesse. Sua evolução é guiada por forças internas e externas. As forças internas servem para manter as bordas suaves, sendo calculadas por fatores interiores à curva. Já as forças externas são tiradas da imagem como um todo e servem para direcionar a curva rumo às bordas desejadas. Por representação e implementação, os modelos deformáveis são divididos em duas classes: paramétricas [N. Duta and M. Sonka, 1998], [D. Shen, E. H. Herskovits, and C. Davatzikos, 2001] e geométricas [A. Tsai et al., 2003], [J. Yang, L. H. Staib, and J. S. Duncan, 2004], [A. Tsai, 2004].

Os modelos deformáveis paramétricos são ferramentas muito robustas na segmentação de imagens biomédicas [C. Xu, D. L. Pham, and J. L. Prince, 2000], [D. Terzopoulos, 1987]. Essas imagens ordinariamente contêm estruturas complexas e irregulares, fazendo com que a representação e a segmentação dessas estruturas por descritores locais sejam bastante trabalhosas. Esses modelos são fisicamente planejados para bordas usando curvas paramétricas ou superfícies que deformam sob a influência de forças externas e internas. Para se delinear as bordas de um objeto em uma imagem, é preciso colocar uma curva fechada perto das bordas desejadas, de modo que essa curva passe por um processo iterativo de relaxação. A Figura 12 mostra como ocorre a deformação.

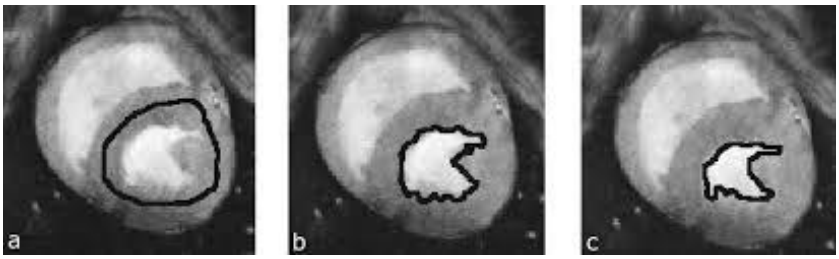

Figura 12 – Segmentação usando contornos ativos.
Fonte: [S. Angenent, E. Pichon, and A. Tannenbaum, 2006]

Modelos deformáveis não-paramétricos são baseados na teoria da convolução da curva. A evolução da curva independe da parametrização e, portanto, as alterações da topologia podem ser realizadas automaticamente. No entanto, elas são computacionalmente mais custosas [W. S. Salem, A. F. Seddik, and H. F. Ali, 2013].

2.3.5. Redes Neurais

As redes neurais são modelos computacionais que simulam o sistema nervoso central, e são muito úteis em diversos casos, incluindo a segmentação de imagens. Elas podem ser usadas no reconhecimento de padrões, que é um dos modelos de utilização de maior potencial das redes. As redes neurais farão a categorização dos padrões das imagens de entrada, seja supervisionada ou não-supervisionada, para identificar qual é a melhor resposta e assim ter a saída mais adequada. As redes neurais com treinamento utilizam uma base de dados, um conjunto de treinamento, e a cada iteração há um aprendizado ao utilizar os rótulos, que deve diferir do conjunto de testes. As redes neurais não supervisionadas utilizam apenas as características das amostras de entrada, sem que haja rótulos para se saber se o treinamento está sendo bem feito. O SOM faz isso através de treinamento e ajustes de pesos dos nodos, para que eles se pareçam o máximo com os padrões de entrada e possam representá-los corretamente, como mostrado na Figura 13.

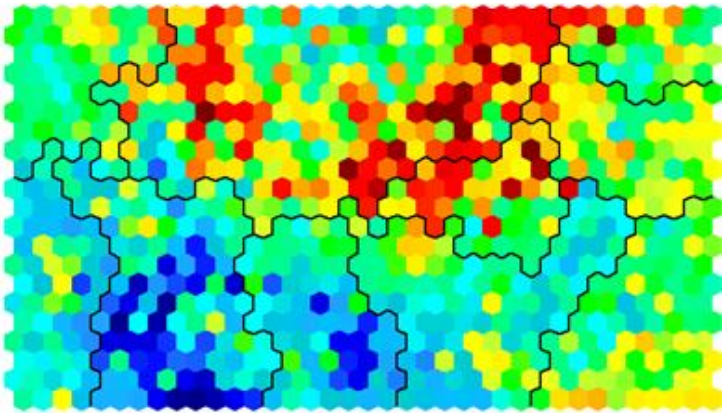

Figura 13 – Segmentação e visualização usando *SOM*.

Fonte: [S. Lynn, 2016]

3. TÉCNICAS UTILIZADAS

Neste capítulo, serão mostradas todas as técnicas utilizadas no trabalho apresentado, incluindo os algoritmos usados na segmentação e as técnicas de avaliação de desempenho, além de uma explicação sobre a combinação de classificadores e o porquê de seu uso e uma descrição sobre o que é reconhecimento de padrões. Também serão explicados o método de seleção de resultados e o *Bagging*, método que inspirou o modelo proposto.

3.1. Conhecimentos Básicos Sobre Classificação

Aqui é feita uma descrição dos conhecimentos básicos necessários para que se entenda o que é reconhecimento de padrões, o que são os sistemas de classificadores múltiplos e o método de seleção escolhido no trabalho.

3.1.1. Fundamentos do Reconhecimento de Padrões

De acordo com [L. I. Kuncheva, 2004], reconhecimento de padrões consiste em atribuir rótulos para os objetos. Estes, por sua vez, são descritos como um conjunto de medidas chamado também de atributos ou características. Seus estudos iniciais são datados das décadas de 1960 e 1970 e, devido às dificuldades para resolver problemas reais, apesar de décadas de pesquisas, teorias modernas na área ainda coexistem com ideias antigas. Isso é mostrado pela quantidade e variedade de técnicas e métodos desenvolvidos e disponíveis para quem deseja pesquisar sobre a área.

3.1.1.1. Ciclo do Reconhecimento de Padrões

A Figura 14 mostra os estágios para o reconhecimento de padrões. Um usuário apresenta um problema e um conjunto de dados. A tarefa é transformar esse problema em terminologia de reconhecimento de padrões, resolvê-lo e, por fim, comunicar a solução novamente ao usuário. Se o conjunto de dados não é fornecido, um experimento então é criado e um conjunto de dados é coletado. As características mais relevantes do conjunto de dados devem ser designadas e medidas. O conjunto de características deve ser o mais discriminante possível, contendo, inclusive, características que possam parecer menos relevantes inicialmente, mas que possam vir a ser relevantes com outras características. A

coleta dessas características pode ter alguns problemas, que vão desde a dificuldade em conseguir os dados, até a dificuldade de medir características subjetivas (como avaliar danos em um acidente, por exemplo).

Existem duas abordagens principais para tratamento de problemas de reconhecimento de padrões: os não supervisionados e os supervisionados. Na categoria dos não supervisionados, o problema consiste em descobrir a estrutura das características do conjunto de dados, caso haja alguma, ou seja, quer-se saber se há agrupamentos nos dados, e quais características fazem os objetos similares dentro do próprio grupo e diferentes dos outros grupos existentes. Já na categoria dos supervisionados, cada objeto do conjunto de dados possui com um rótulo já pré-assumido. A tarefa, portanto, é treinar um classificador, de modo que ele faça uma classificação que condiga com os rótulos do treinamento. Para isso, o classificador aprende um conjunto de técnicas de aprendizagem e tem o conjunto de dados classificados apresentado a ele. O que vai definir se essas técnicas de aprendizagem foram adequadas para o problema em questão será a precisão dos resultados obtidos.

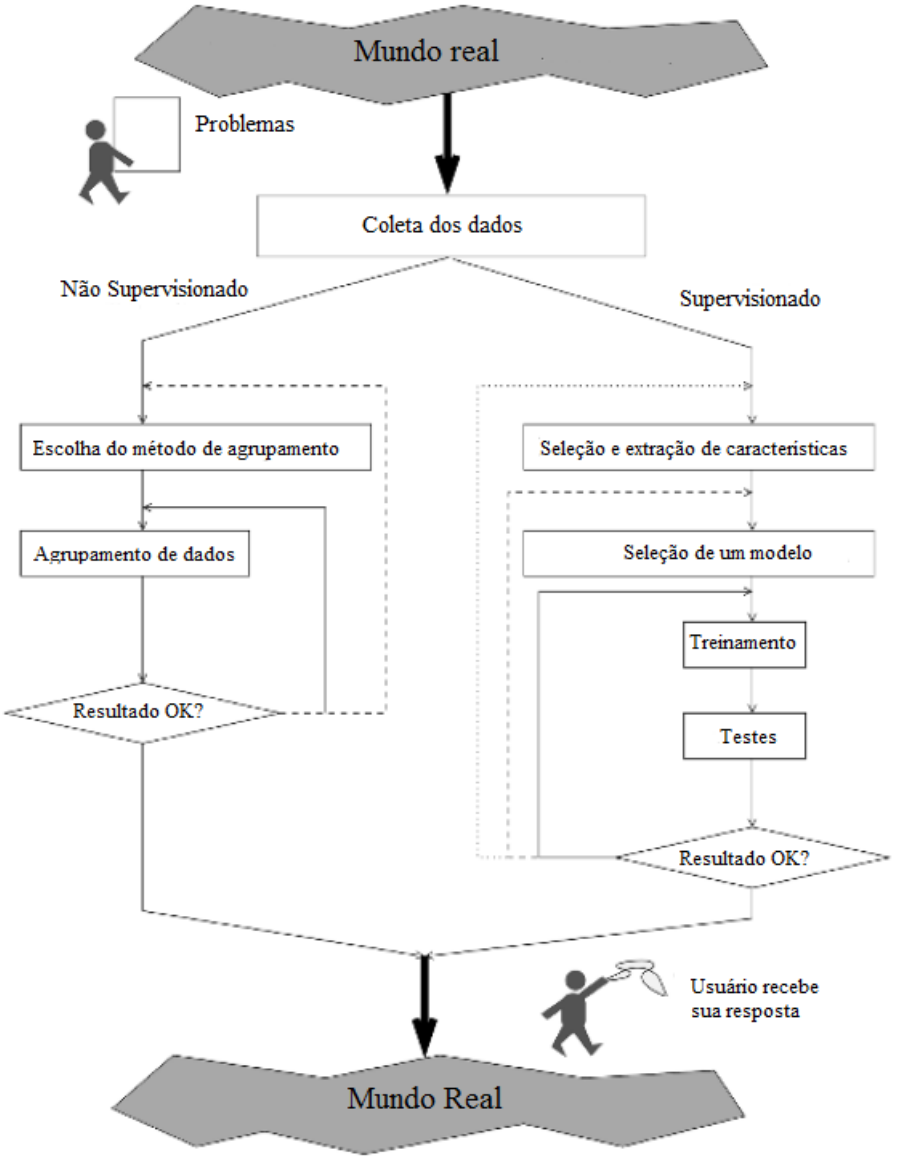

Figura 14- Ciclo do reconhecimento de padrões.

Fonte: [L. I. Kuncheva, 2004] p. 2

As características não são todas igualmente relevantes. Algumas delas são importantes apenas na relação com outras características, enquanto outras podem ser apenas um tipo de "ruído" no contexto. Seleção e extração de características são usadas para melhorar a qualidade da descrição. Seleção, treinamento e testes de um modelo de classificador formam a base de reconhecimento de padrões supervisionado. As linhas pontilhadas da Figura 14 mostram que o laço de refinamento do modelo pode ser interrompido em diferentes lugares. É possível refazer o treinamento usando o mesmo modelo, alterando parâmetros, ou mesmo

alterar o modelo. Por fim, quando uma solução aceitável é alcançada, então o usuário pode utilizá-la para testes e aplicações.

3.1.1.2. Classes e Características

Uma classe contém objetos similares, enquanto objetos dissimilares pertencem a classes diferentes. Algumas classes possuem uma divisão bastante clara, e, no caso mais simples, são mutuamente excludentes. Por outro lado, há também problemas em que classes possam ser difíceis de definir, como por exemplo, pesquisas médicas, como imagens médicas ou dados coletados, que possuem grande dificuldade de interpretação devido à variabilidade natural dos objetos de estudo. Assim, em um problema, haverá uma determinada quantidade de classes, em que cada classe terá um rótulo.

Característica é aquilo que descreve ou representa o objeto de estudo. As características podem ser qualitativas ou quantitativas. As quantitativas podem ser divididas em contínuas ou discretas. Características que possuem uma pequena quantidade de possibilidades são consideradas qualitativas, que podem ser divididas em ordinais ou nominais. Objetos podem ser representados por múltiplos subconjuntos de características. Esses subconjuntos específicos são medidos para cada análise de modalidade de detecção, e então o vetor de características é composto por subvetores. Essa subdivisão é chamada de Representação de Padrões Distintos (*distinct patern representation*) [J. Kittler et al., 1998].

A informação que é necessária para construir um classificador está usualmente na forma de um conjunto de dados rotulado. A Figura 15 mostra um exemplo de dígitos escritos à mão, que devem ser classificados em 10 classes diferentes. Para se construir o conjunto de dados, as imagens em preto e branco precisam ser transformadas em vetores de características, o que pode nem sempre ser tão simples. No exemplo, pode-se citar como características a quantidade de traços verticais ou horizontais ou mesmo o número de círculos na imagem do dígito. Escolher um conjunto de características com alto poder discriminante, portanto, pode determinar o sucesso do sistema reconhecedor de padrões.

Figura 15 – Dígitos manuscritos em letra cursiva.
Fonte: [L. Cun, B. Boser, et al., 1990] p. 3

3.1.2. Sistemas de Classificadores Múltiplos

Ao combinar classificadores diferentes, o objetivo é conseguir classificações mais precisas, ao custo de aumentar a complexidade do sistema. A dúvida é saber quando uma combinação de classificadores é justificável. De acordo com [T. K. Ho, 2002], o próximo passo para a combinação de classificadores múltiplos deve ser olhar pelo melhor conjunto de classificadores ou o melhor método de combinação, ao invés de buscar o melhor conjunto de características ou o melhor classificador especificamente. É possível imaginar que, em breve, a busca será pelo melhor conjunto de métodos de combinações e o melhor método de usar todos.

Combinar classificadores é um passo natural quando já se acumulou uma grande quantidade de conhecimento para modelos de classificadores únicos. Assim, a combinação de classificadores tem crescido rapidamente e ganho importância dentro dos ramos de reconhecimento de padrões e aprendizagem de máquina.

3.1.2.1. Conceitos Básicos

De acordo com [T. G. Dietterich, 2000], há três motivos para se usar uma combinação de classificadores ao invés de usar apenas um classificador:

- Estatístico:

Supondo que haja um número de diferentes classificadores com bom desempenho em uma base de dados específica, ao escolher apenas um desses classificadores como a solução,

corre-se o risco de escolher um classificador que não seja o melhor. Desse modo, uma maneira mais segura de obter um bom resultado é usar todos os classificadores e, através de alguma métrica, escolher a média dessas respostas. Desse modo, o novo classificador pode não ser melhor do que o melhor classificador unitário, mas irá diminuir ou eliminar o risco de se escolher um classificador unitário inadequado. A Figura 16 mostra uma ilustração do que foi descrito acima. A forma *H* representa o espaço de modelos com todos os classificadores. A região interna em azul contém todos os que, para aquela base de dados, possuem um bom desempenho. O melhor classificador para o problema é chamado de *f*. O esperado é que, ao se juntar os classificadores, tenha-se um classificador resultante mais próximo de *f* do que um classificador aleatoriamente escolhido do espaço de classificadores.

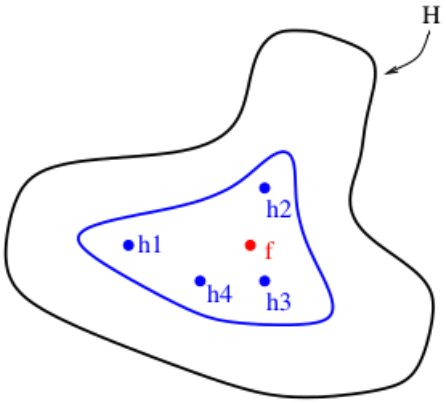

Figura 16 – Motivo estatístico para se combinar classificadores, onde *f* é o melhor classificador.
Fonte: [S. Shah, 2015]

- Computacional:

Alguns algoritmos de treinamento realizam *hill-climbing* ou fazem buscas aleatórias, o que pode levar a se encontrar valores diferentes de ótimos locais. A Figura 17 mostra exatamente essa situação. Assume-se que o processo de treinamento de cada classificador individual comece em algum lugar no espaço de parâmetros do modelo e termine mais próximo do melhor classificador *f*. Alguma forma de agregação pode levar a um classificador que seja uma aproximação melhor de *f* do que de algum classificador *h* qualquer.

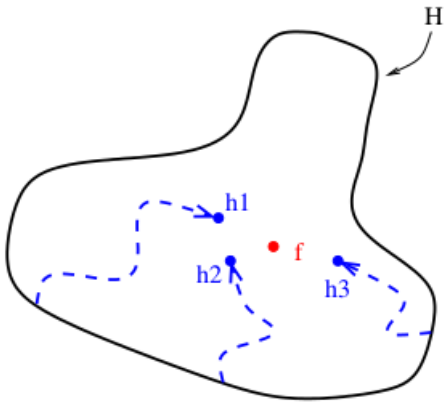

Figura 17 – Motivo computacional para se combinar classificadores, onde *f* é o melhor classificador.

Fonte: [S. Shah, 2015]

- Representacional:

  Em alguns casos, é possível que o espaço de classificadores considerado para o problema não contenha o classificador ótimo. Por exemplo, se o classificador ótimo de um problema for não linear e o espaço de classificadores for restrito para classificadores lineares, então o classificador ótimo estará fora do espaço de classificadores possíveis. No entanto, um conjunto de classificadores lineares pode aproximar qualquer fronteira de decisão com qualquer precisão predefinida. Para um problema como esse, pode-se treinar um novo classificador de maneira específica, contudo sua complexidade pode ser tão alta, que uma combinação de classificadores pode ser uma opção melhor para esse problema. A Figura 18 mostra o caso em que o classificador ótimo *f* está fora do espaço de classificadores.

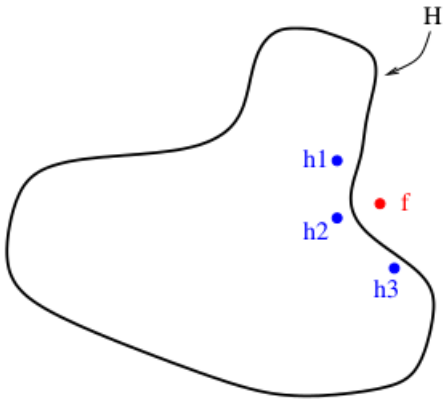

Figura 18 – Motivo representacional para se combinar classificadores, onde *f* é o melhor classificador.

Fonte: [S. Shah, 2015]

É preciso salientar que uma melhora no desempenho do melhor classificador ou na média do grupo não é garantida. O que foi descrito são motivos heurísticos que se mostram razoáveis. No entanto, trabalhos experimentais e as teorias desenvolvidas até agora demonstram o sucesso dos métodos de combinação de classificadores [L. I. Kuncheva, 2004].

3.1.2.2. Treinamento

Havendo um conjunto de dados a ser utilizado para treinar e testar um classificador, deseja-se utilizar o máximo de dados possível para fazer o treinamento do classificador, e também deixar o máximo possível de dados não ser visto para poder realizar testes melhores. Contudo, caso seja usado todo o conjunto de dados para o treinamento e o mesmo conjunto para os testes, é possível que haja um *overtraining*, ou seja, o classificador acabaria aprendendo apenas aqueles dados, e não generalizaria para outros dados de entrada que viessem a aparecer. Desse modo, é importante que haja uma separação entre o conjunto de dados a ser utilizado para o treinamento e o conjunto de dados que será utilizado na fase de testes.

Para treinar um conjunto de classificadores, algumas abordagens podem ser adotadas. Caso haja um conjunto de dados grande, as opções mais conhecidas são:

- Treinar um classificador (possivelmente muito complexo);
- Treinar os classificadores base usando conjuntos de treinamento que não se intersectam;
- Avaliar o conjunto e os classificadores individualmente de maneira muito precisa (usando um grande conjunto de testes) para que se se possa decidir qual é a melhor opção na prática.

Enquanto há muitas possibilidades quando se possui um grande conjunto de dados, conjuntos de dados pequenos podem ser um desafio grande. De acordo com [R. P. Duin, 2002], a estratégia de treinamento tem um ponto crucial no caso da utilização de combinação de classificadores, e sugere as seguintes recomendações:

- Se um conjunto de dados de treinamento for usado com um combinador não treinável, como o voto majoritário (combinador que não tem parâmetros extras que precisem ser treinados, ou seja, o combinador está pronto para operação assim que os

classificadores base estiverem treinados), então é preciso garantir que esses classificadores base não sofram *overtraining*. As precisões desses classificadores base são tidas como menos importantes para o sucesso do conjunto do que a confiabilidade das decisões e as estimativas de probabilidade;

- Se um conjunto de dados de treinamento for usado com um combinador treinável, então os classificadores base devem ficar subtreinados, e subsequentemente completar esse treinamento do combinador com o conjunto de treinamento. Aqui, é assumido que o conjunto de treinamento tem um potencial de treinamento. Então, para que seja possível treinar o combinador de maneira apropriada, os classificadores base não devem usar todo o potencial;
- Usar conjuntos de treinamento separados para os classificadores base e para os combinadores. Então, os classificadores base podem sofrer *overtraining* no conjunto de treinamento. A inclinação vai ser corrigida ao se treinar o combinador com o conjunto de treinamento separado.

3.1.2.3. Tipos de Saídas de Classificadores

Os modos possíveis de combinar as saídas dos classificadores em um combinador dependem de que tipo de informação é obtida de cada classificador individual. De acordo com [L. Xu, A. Krzyzak, and C. Y. Suen, 1992], há três tipos de saídas distintas:

- Tipo 1 (Nível Abstrato): Cada classificador produz um rótulos. No nível abstrato, não há informação sobre a precisão dos rótulos resultantes, nem alguma outro rótulo sugerida como opção. Por definição, qualquer classificador é capaz de produzir um rótulo, então o nível abstrato é o nível mais conhecido;
- Tipo 2 (Nível de Ranking): Nesse nível, cada classificador ordena as alternativas por ordem de plausibilidade de cada rótulo [T. K. Ho, J. J. Hull, and S. N. Srihari, 1994], [J. D. Tubbs and W. O. Alltop, 1991]. Esse tipo é especialmente adequado para problemas com um grande número de classes, como, por exemplo, caracteres, faces, reconhecimento de discurso ou voz, entre outros;
- Tipo 3 (Nível de Medição): Cada classificador produz um vetor, de tamanho igual ao número de classes, onde cada posição do vetor representa o percentual de probabilidade de que o aquela entrada pertença àquela classe.

3.1.3. Voto Majoritário

Neste trabalho, foi escolhida a abordagem do tipo 1. A técnica escolhida foi o voto majoritário, que resulta em apenas um rótulo por entrada para cada classificador. O voto majoritário é uma das mais antigas estratégias para tomada de decisões. Uma revisão da evolução do conceito foi feita por [W. H. E. Day, 1988], que mostra que esse tipo de decisão é usado desde tempos antigos, como as cidades da Grécia antiga ou mesmo o Senado Romano.

Três padrões consensuais, unanimidade, maioria simples e pluralidade são mostradas na Figura 19. Ao assumir que preto, cinza e branco correspondem a classes diferentes, e quem toma as decisões são os classificadores individuais do combinador, o rótulo final será preto para todos os três padrões. O voto majoritário será, portanto, a classe que mais vezes tiver sido escolhida pelos classificadores. Em caso de empate, haverá alguma solução arbitrária para resolução. Ela vai então coincidir com o caso de maioria simples para problemas com apenas duas classes.

Figura 19- Padrões de consenso em um grupo de 10 tomadores de decisão: unanimidade, maioria simples e pluralidade. Em todos os três casos, a decisão final do grupo foi por preto.
Fonte: [L. I. Kuncheva, 2004] p. 113

Vários estudos são dedicados ao uso do voto majoritário para a combinação de classificadores [D. Ruta and B. Gabrys, 2002], [L. Lam and C. Y. Suen, 1994], [X. Lin et al., 2003]. Para perceber o porquê de o voto majoritário ser um dos métodos de combinação de classificadores mais populares, basta analisar algumas de suas propriedades. Assumindo que:

- Exista um número de classificadores L que seja ímpar;
- A probabilidade de que cada classificador escolha a classe correta é p;
- As saídas dos classificadores são independentes.

Desse modo, o voto majoritário retornará uma classificação se pelo menos $\lfloor L/2 \rfloor$ (menor inteiro menor que L) + 1 classificadores fizerem a classificação de maneira correta. Então, a precisão da classificação do combinador segue a Equação 1:

$$P_{maj} = \sum_{m=\lfloor \frac{L}{2} \rfloor + 1}^{L} \binom{L}{m} p^m (1-p)^{L-m} \quad (1)$$

As probabilidades de classificação correta da combinação para p = 0.6, 0.7, 0.8 e 0.9, com L = 3, 5, 7 e 9 são mostradas na Tabela 1 [L. I. Kuncheva, 2004].

Tabela 1 – Valores de precisão do voto majoritário para L classificadores independentes com precisão individual p.

	$L = 3$	$L = 5$	$L = 7$	$L = 9$
$p = 0.6$	0.6480	0.6826	0.7102	0.7334
$p = 0.7$	0.7840	0.8369	0.8740	0.9012
$p = 0.8$	0.8960	0.9421	0.9667	0.9804
$p = 0.9$	0.9720	0.9914	0.9973	0.9991

Fonte: [L. I. Kuncheva, 2004] p. 114

O resultado a seguir também é conhecido como o Teorema do Júri de Condorcet [L. Shapley and B. Grofman, 1984]:

1. Se $p > 0.5$, então P_{maj} na Equação 1 cresce monotonicamente e
$$P_{maj} \to 1 \quad \text{para} \quad L \to \infty$$
2. Se $p < 0.5$, então P_{maj} na Equação 1 decresce monotonicamente e
$$P_{maj} \to 0 \quad \text{para} \quad L \to \infty$$
3. Se $p = 0.5$, então $P_{maj} = 0.5$ para qualquer L.

Esse resultado confirma a intuição de que seja esperada uma melhoria sobre a precisão p apenas quando o próprio p é maior que 0.5. Há estudos ainda, como o de [L. Lam and S. Suen, 1997], que analisam o caso de L ser um número par e o efeito da precisão da combinação de classificadores ao se adicionar ou remover novos classificadores, onde é provado que, caso o custo de um erro seja maior do que o custo de uma rejeição, o desempenho é melhor com L sendo par.

3.1.4. Bagging

O *Bagging* foi a técnica que inspirou o método proposto. Desse modo, apesar dele não ser efetivamente utilizado no trabalho, cabe uma explicação sobre seu funcionamento e quando utilizá-lo.

O termo *Bagging* foi introduzido por [L. Breiman, 1994] como um acrônimo para *Bootstrap AGGregatING*. A ideia do *Bagging* é que um combinador é feito de classificadores que foram treinados com diferentes conjuntos de treinamento. Ao final, as saídas dos classificadores são combinadas pelo voto de pluralidade [L. Breiman, 1996]. Idealmente, os conjuntos de treinamento deveriam ser gerados aleatoriamente da distribuição do problema. Contudo, na prática, geralmente tem-se apenas um conjunto de dados para treinamento, e é preciso imitar o processo de uma maior quantidade de conjuntos de dados de treinamento. Logo, é feita uma amostragem desses dados, com reposição (amostragem *bootstrap* [B. Efron and R. J. Tibshirani, 1994]), para criar um novo conjunto de treinamento. Para se fazer uso dessas diferenças dos dados de treinamento, deve-se usar um classificador base instável, ou seja, um classificador em que, caso haja pequenas diferenças no conjunto de treinamento, ele deve gerar saídas muito diferentes. Do contrário, o combinador resultante nada mais seria do que uma coleção de classificadores praticamente idênticos, o que não levaria a melhorias em relação ao desempenho do classificador de maneira individual. Exemplos de classificadores instáveis são redes neurais, ou árvores de decisão. A Figura 20 mostra o fluxo do *Bagging*, onde há a seleção, o treinamento e a combinação dos classificadores, para que seja feita a predição.

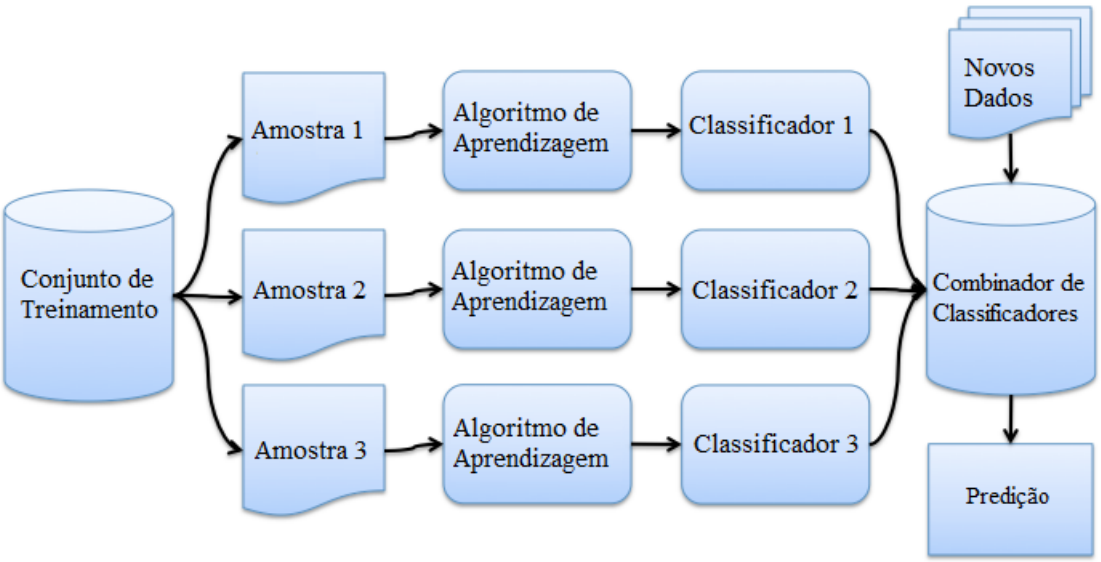

Figura 20 – O funcionamento do algoritmo *Bagging*.

Fonte: [Breiman, 1996]

Se as saídas dos classificadores forem independentes e os classificadores tiverem a mesma precisão, então o voto majoritário é garantido de melhorar o desempenho individual. O *Bagging* visa desenvolver classificadores independentes, pegando amostras aleatórias e diferentes entre si dos conjuntos de treinamento. Desse modo, as amostras serão pseudo-independentes, já que são retiradas do mesmo conjunto. Assim, com o aumento de classificadores treinados usando o *Bagging*, a tendência é que o número de acertos aumente e o erro diminua.

3.2. Técnicas Utilizadas

Aqui serão feitas as descrições de todas as técnicas utilizadas no trabalho proposto, descrevendo seu funcionamento, com suas equações e imagens, para demonstrar como elas funcionam na prática. As técnicas selecionadas foram a binarização de Otsu, o *K-Means*, a rede neural SOM (*Self-Organizing Maps*) e o GMM (*Gaussian Mixture Model*). Além disso, é descrito também o método de avaliação escolhido, o *Dice Coefficient* (*Sørensen-Dice Coefficient*).

3.2.1. Binarização de Otsu

Desenvolvido por Nobuyuki Otsu, é um método não paramétrico e não supervisionado para seleção automática de limiares para segmentação de imagens [N. Otsu, 1975]. Um limiar ótimo é selecionado utilizando o critério do discriminante, de modo a maximizar a separabilidade das classes resultantes em níveis de cinza. O procedimento utiliza apenas os momentos de ordem zero e primeira ordem dos histogramas da imagem, em tons de cinza. Ele também pode ser estendido, de modo a ser utilizado em problemas que exijam mais de um limiar.

O método propõe iterar por todos os valores de limiar possíveis na imagem a ser analisada, buscando o valor que minimize a soma das variâncias intraclasses da imagem. Logo, o valor selecionado será o melhor limiar, que vai separar frente e fundo (caso haja apenas um limiar), e vai atribuir uma cor para cada uma das classes.

Para avaliar se o limiar é o melhor, calcula-se a variância intraclasse utilizando a seguinte equação:

$$\sigma^2{}_W = w_0\sigma^2{}_0 + w_1\sigma^2{}_1 \tag{2}$$

onde w é o peso para cada uma das classes. Desse modo, a Equação 2 serve para calcular a probabilidade de um pixel selecionado pertencer a classe 0 (fundo) ou 1 (o primeiro plano), e esse cálculo deverá ser realizado para todos os valores de limiar possíveis para a imagem.

A escolha será o valor de limiar que minimizar o valor da variância intraclasse, para a binarização. Contudo, para se calcular a variância, é preciso que se saiba os pesos e os valores das médias de cada classe, sendo assim muito custoso computacionalmente. Otsu, na sua pesquisa, conseguiu mostrar que pode-se substituir a Equação 2 pelo cálculo da variância interclasse, o que diminui bastante o custo computacional do algoritmo. Desse modo, tem-se :

$$\sigma^2{}_B = \sigma^2 - \sigma^2{}_W \tag{3}$$

$$\sigma^2{}_B = w_0(\mu_0 - \mu)^2 + w_1(\mu_1 - \mu)^2 \tag{4}$$

onde:

$$\mu = w_0\mu_0 + w_1\mu_1 \tag{5}$$

Logo: $$\sigma^2{}_B = w_0 w_1 (\mu_1 - \mu_0)^2 \quad (6)$$

A diferença ao se aplicar esse caso, é que agora o algoritmo busca a variância interclasse. Portanto, o valor que atingir essa meta será o valor de limiar que vai minimizar a variância intraclasse e, assim, será o valor ideal para o limiar.

A Figura 21 mostra, em um fluxograma, o funcionamento do algoritmo de Otsu. Primeiro, é preciso que se calcule o histograma da imagem a ser binarizada, que deve estar em tons de cinza. Então, para cada valor possível de limiar, vão ser calculados as médias e pesos correspondentes, para cada uma das classes desejadas (o mais comum é utilizar apenas duas classes, mas pode-se aumentar esse valor). Após o término das iterações, o limiar selecionado vai ser aquele que apresentar o maior valor para a variância interclasses. Por fim, cada pixel da imagem será analisado e, dependendo do seu valor, ele passará a receber o valor pré-determinado para aquela classe onde ele se encontra.

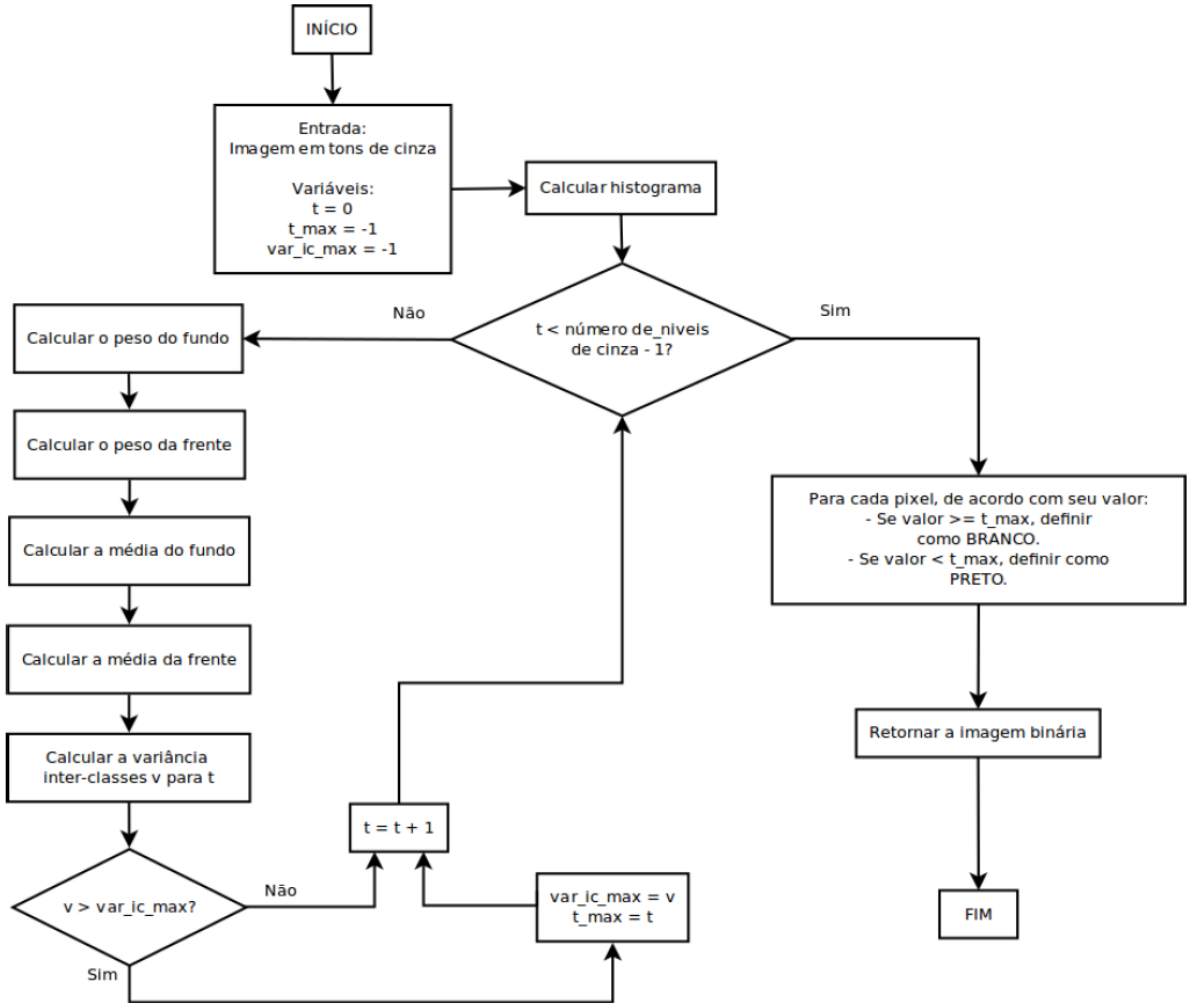

Figura 21 - O fluxograma do algoritmo de Otsu.

Fonte: [L. Torok, 2015], p. 2

A Figura 22 mostra a imagem original em tons de cinza, o histograma com o valor do limiar representado pela linha vermelha e a aplicação do algoritmo de Otsu para a segmentação da imagem. Vale notar que a segmentação não é perfeita para diferenciar o que é fundo do que não é.

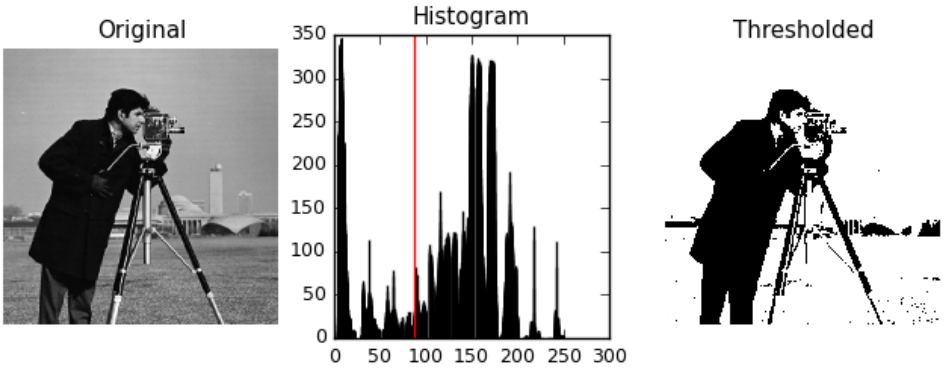

Figura 22 – Aplicação do algoritmo de binarização de Otsu.

Fonte: [Scikit-image, 2016]

3.2.2. *K-Means*

O algoritmo *K-Means* é um algoritmo de agrupamento não-supervisionado, que faz a classificação do dado de entrada baseado em suas distâncias entre si. O algoritmo assume que as características formam um espaço de vetores e tenta encontrar um agrupamento natural neles [S. Tatiraju and A. Mehta, 2008]. Ele geralmente é utilizado para realizar uma segmentação inicial da imagem [G. N. Abras and V. L. Ballarín, 2005], onde as áreas mais grosseiras são atenuadas. O *K-Means* é muito utilizado devido a sua simplicidade e ao seu baixo custo computacional. Além disso, é apto para segmentação de imagens médicas, pois usualmente o número de grupos (*K*) é conhecido de antemão [H. Ng et al., 2006].

Para se iniciar o *K-Means*, é primeiro preciso definir o número de grupos. Feito isso, os *K* centroides deverão ser inicializados em pontos aleatórios e, a partir daí, as entradas irão calcular qual dos centroides está mais próximo desta entrada, seguindo a Equação 7:

$$\arg\min \sum_{i=1}^{k} \sum_{x=1}^{n} \|\boldsymbol{x} - \boldsymbol{\mu}_i\|^2 \tag{7}$$

onde k é o número de grupos, n é o número de entradas, $\boldsymbol{x}$ é a entrada e $\boldsymbol{\mu}$ é o centroide a ser analisado.

Após definir quais entradas irão corresponder a cada centroide, o valor médio do centroide vai ser atualizado, seguindo a Equação 8:

$$\boldsymbol{\mu}_i^{(t+1)} = \frac{1}{\boldsymbol{S}_i^t} \sum_{\boldsymbol{x}_j \in \boldsymbol{S}_i^{(t)}} \boldsymbol{x}_j \tag{8}$$

onde $\boldsymbol{S_i}^t$ é a quantidade de entradas que estão mais próximas do centroide selecionado na iteração t.

Para visualizar melhor o funcionamento, observa-se a Figura 23, que mostra passo a passo como funciona o algoritmo. Primeiro, define-se o número de grupos como 2, para se facilitar o entendimento. Após isso, inicia-se o processo com os centroides em pontos iniciais aleatórios iniciais. Pode ser qualquer lugar do plano, para em seguida começar as iterações e encontrar os resultados.

Os dois pontos aleatórios são colocados no gráfico representando os centroides, e uma linha é calculada na metade da distância dos centroides vermelho e azul, delineando a qual centroide cada objeto se refere. Com este segmento, os objetos que estiverem acima da linha tracejada fazem parte do grupo vermelho e os de baixo da linha fazem parte do grupo azul. A primeira iteração do algoritmo calcula a distância de todos os pontos que estão no grupo até o centroide, e então atualiza a posição do centroide para o novo ponto, que é a posição média de todos os objetos que se ligaram àquele centroide. Essa alteração de posição do centroide pode alterar os objetos que fazem parte do grupo. Após o cálculo da média, alguns objetos mudaram de centroide, os objetos marcados em verde passaram para o grupo vermelho, e o marcado em azul passou para o grupo azul. Quando não ocorre mais nenhuma mudança de objetos com relação aos centroides, então o algoritmo é finalizado e os grupos ficam separados definitivamente.

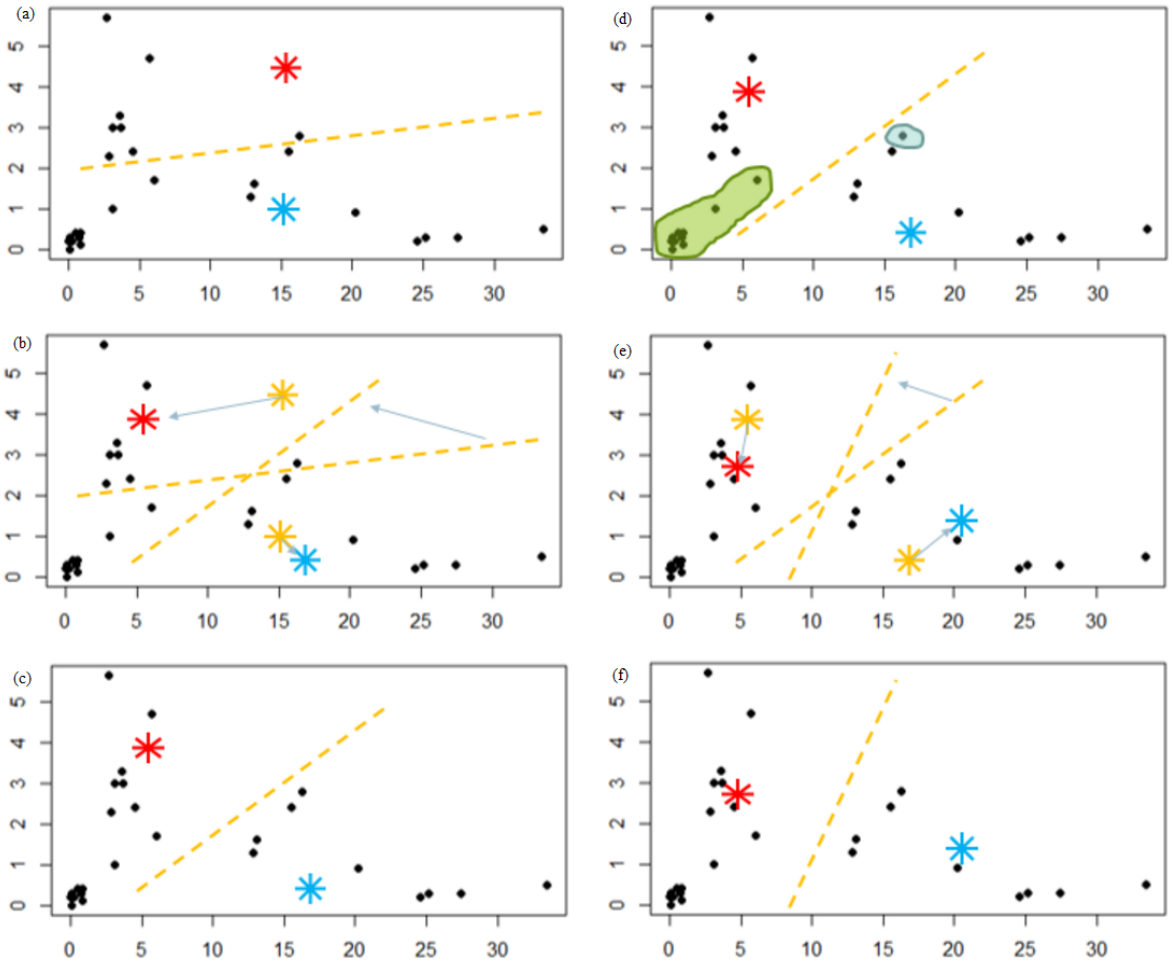

Figura 23 – Funcionamento do *K-Means*, onde (a) inicialização e separação dos objetos entre os centroides, (b) e (c) atualização da posição dos centroides, (d) objetos que mudaram de grupo, (e) nova atualização dos centroides e (f) posição final dos centroides, devido a nenhum ojeto ter mudado de grupo.

Fonte: [D. Nogare, 2016]

Na segmentação, é possível usar o *K-Means* para definir os valores que devem corresponder cada pixel da imagem. A Figura 24 mostra como o *K-Means* age na imagem.

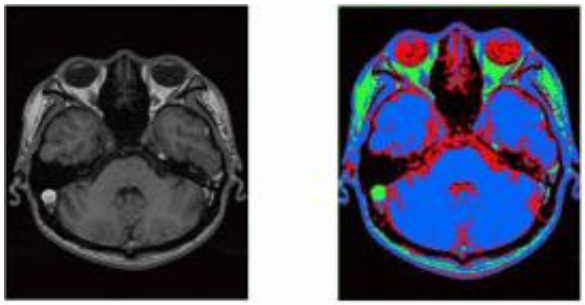

Figura 24 – (a) Imagem original (b) Segmentação com K-Means.

Fonte: [H. P. Ng et al., 2006]

3.2.3. SOM (*Self-Organizing Maps*)

Os Mapas Auto-Organizáveis (*Self-Organizing Maps* - SOM), também conhecidos como Mapas Auto-Organizáveis de Características (*Self-Organizing Feature Maps* - SOFM), foram propostos por [T. Kohonen, 1982]. O SOM é uma rede neural artificial que utiliza aprendizagem não-supervisionada para fazer uma redução de dimensões, mapeando a dimensão elevada de algum espaço para um espaço com dimensão menor, geralmente uma ou duas dimensões, como mostrado na Figura 25, havendo a menor perda de informações possível. Para que sejam mantidas as características topológicas da entrada com dimensão maior, é utilizada uma função de vizinhança. O mapa topológico é ordenado automaticamente, ao serem comparados, por várias vezes, os vetores de características de entrada e os vetores de peso de cada nodo existente na rede. Desse modo, cada região do mapa acaba representando um conjunto de características semelhantes, o que leva a uma categorização das áreas próximas.

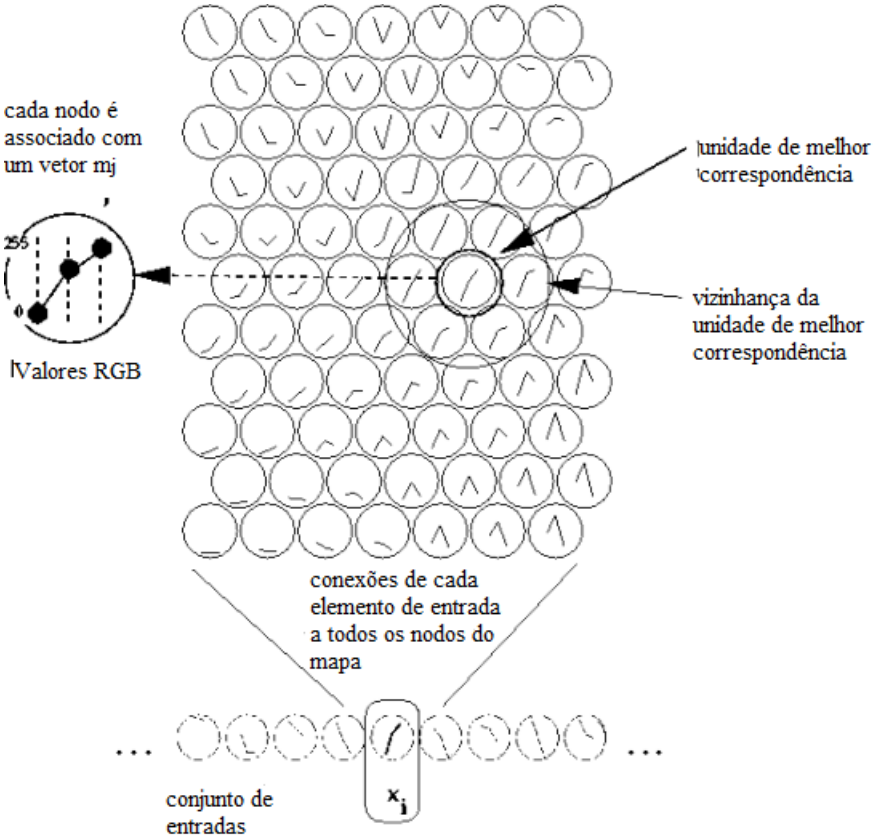

Figura 25 – Topologia do SOM. A entrada verifica o nodo com características mais próximas e atualiza seu valor.

Fonte: [T. Honkela, 1997] p. 16

Atualmente, o SOM é utilizado para várias funcionalidades [M. L. Gonçalves, M. L. de Andrade Netto, and J. A. F. Costa], como, por exemplo, reconhecimento automático de discurso, análise clínica de voz, análise dos sinais elétricos do cérebro, mineração e compressão de dados, entre outros.

O SOM funciona da seguinte maneira: os vetores de peso dos nodos da rede são inicializados com valores aleatórios e, de preferência, em posições bem próximas para evitar que vizinhos estejam muito afastados no início. Após a inicialização, um vetor de características de entrada será apresentado à rede, que irá escolher o vetor de peso que for mais semelhante ao vetor de características, ou seja, o nodo que estiver mais perto, já que a função para o cálculo da distância é a distância Euclidiana. O nodo escolhido é o nodo vencedor. À medida que o treinamento for se realizando, a rede atualiza os nodos para ficaram mais semelhantes aos vetores de características e, além disso, também atualiza seus vizinhos, de acordo com a Equação 9 abaixo:

$$\omega_i(k) = \omega_i(k-1) + \alpha(k-1).h_{i,j}(k-1).(s_k(k-1) - \omega_i(k-1)) \qquad (9)$$

onde $\omega_i(k)$ é o vetor de pesos da rede na unidade *i* no instante k, α é a função que estabelece o valor da taxa de aprendizagem da rede que decresce à medida que o valor de k aumenta, variando entre 0 e 1, $h_{i,j}$ é a função de vizinhança da rede, que depende do nodo vencedor e decresce à medida que a distância para o nodo vencedor aumenta, variando também entre 0 e 1 e $s_k(k)$ é o vetor de características de entrada no instante k.

Desse modo, ao fazer a atualização não só do nodo vencedor, mas também de seus vizinhos, a rede induz que o mapa gerado tenha características topológicas semelhantes aos dados de entrada, ou seja, eles responderão de maneira semelhante a vetores de características de entrada semelhantes.

O treinamento é feito com os vetores de características de entrada sendo utilizados em ordem aleatória, para que a rede não considere diferenças de momentos de apresentações no agrupamento espacial dos padrões e obtenha resultados incorretos caso um novo vetor seja apresentado. Assim, o treinamento é interrompido caso o número de iterações definido seja alcançado ou caso o erro médio entre o nodo antes e depois da atualização seja pequeno, mostrando que houve convergência na rede e não é mais preciso treinamento.

Na segmentação, também é possível usar o SOM para definir os valores que devem corresponder cada pixel da imagem. A Figura 26 mostra como o SOM age na imagem, mostrando a imagem original e após a segmentação.

Figura 26 – Imagem original e segmentada utilizando o SOM.
Fonte: [Jaffar, 2011] p. 3

3.2.4. GMM (*Gaussian Mixture Model*)

O GMM é uma função de densidade de probabilidade paramétrica, representada como uma soma ponderada de densidades de componentes gaussianos [D. Reynolds, 2015]. O GMM é mais comumente utilizado como um modelo paramétrico da distribuição de probabilidade de medidas contínuas, ou de características em sistemas biométricos. Seus parâmetros são estimados dos dados de treinamento, usando geralmente o algoritmo de *Expectation-Maximization* (EM).

O GMM é uma soma ponderada de M densidades gaussianas de componentes, dadas pela Equação 10:

$$\mathrm{p}(\mathbf{x}|\lambda) = \sum_{i=1}^{M} w_i \; \mathrm{g}(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma_i}) \tag{10}$$

onde $\mathbf{x}$ é um vetor de dimensão D, w_i, onde $i = 1,..., M$, são os vetores de mistura e $\mathrm{g}(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma_i})$, onde $i = 1,..., M$, são as M densidades gaussianas de componente. Cada densidade de componente é uma função gaussiana da forma da Equação 11:

$$\mathrm{g}(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma_i}) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)'{\Sigma_i}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)\right\} \tag{11}$$

com o vetor de média $\boldsymbol{\mu}_i$ e a matriz de covariância Σ_i. Os pesos da mistura satisfazem a restrição $\sum_{i=1}^{M} w_i = 1$.

O modelo completo é parametrizado pelos vetores de médias, matrizes de covariâncias e pesos das misturas de todas as densidades de componentes. Esses parâmetros são representados pela notação da Equação 12:

$$\lambda = \{w_i, \boldsymbol{\mu}_i, \Sigma_i\} \tag{12}$$

Existem diversas variantes sobre o GMM mostrado na Equação 11. As matrizes de covariância, Σ_i, podem ser completas ou restritas a ser diagonais. Além disso, os parâmetros podem ser partilhados, ou amarrados, entre os componentes da gaussiana, assim como ter uma matriz de covariância comum para todos os componentes. A escolha da configuração do modelo (número de componentes, as matrizes de covariância completas ou diagonais, e parâmetros) é muitas vezes determinada pela quantidade de dados disponíveis para estimar os parâmetros do GMM, e como o GMM é utilizado para uma aplicação particular. É também importante notar que, como os componentes da gaussiana atuam em conjunto para modelar a função de densidade global, as matrizes de covariância completas não são necessárias, mesmo se as características não forem estatisticamente independentes. A combinação linear de gaussianas de covariância diagonais é capaz de modelar as correlações entre os vetores elementos. O efeito do uso de um conjunto de *M* matrizes de covariância completas pode ser obtido por meio de um conjunto maior de gaussianas de covariância diagonais.

3.2.4.1. EM *(Expectation-Maximization)*

Na estatística, o algoritmo EM é um método iterativo para encontrar a máxima verossimilhança de parâmetros em modelos estatísticos, onde o modelo depende de variáveis latentes não observáveis. A iteração do EM alterna entre realizar um passo de Expectativa (E), que cria uma função para a expectativa da log-verossimilhança avaliada, usando a atual estimativa para os parâmetros, e um passo de Maximização (M), que computa os parâmetros maximizando a máxima log-verossimilhança encontrada no passo E. Esses parâmetros estimados então são usados para determinar a distribuição das variáveis latentes no próximo passo E.

Dado o modelo estatístico que gera um conjunto $\mathbf{X}$ de dados observados, um conjunto de dados latentes não observáveis $\mathbf{Z}$ e um vetor de parâmetros desconhecidos $\boldsymbol{\theta}$, junto com a função de verossimilhança $\boldsymbol{L}(\boldsymbol{\theta}; \mathbf{X}, \mathbf{Z}) = p\ (\mathbf{X}, \boldsymbol{Z}|\boldsymbol{\theta})$, a estimativa de máxima verossimilhança dos parâmetros desconhecidos é determinada pela verossimilhança marginal dos dados observáveis, mostrada na Equação 13:

$$\boldsymbol{L}(\boldsymbol{\theta}; \mathbf{X}) = p(\mathbf{X}|\boldsymbol{\theta}) = \sum_{\mathbf{Z}} p(\mathbf{X}, \mathbf{Z}|\boldsymbol{\theta}) \tag{13}$$

O algoritmo EM tenta encontrar a estimativa de máxima verossimilhança da verossimilhança marginal aplicando, iterativamente, os dois passos a seguir:

- Passo de Expectativa (E): Calcula o valor esperado da função de log-verossimilhança, com respeito à distribuição condicional de $\mathbf{Z}$ dado $\mathbf{X}$, sob a estimativa atual dos parâmetros $\boldsymbol{\theta}^{(t)}$, mostrado na Equação 14:

$$\boldsymbol{Q}\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(t)}\right) = \mathbf{E}_{\mathbf{Z}|\mathbf{X},\boldsymbol{\theta}^{(t)}}[\log \boldsymbol{L}(\boldsymbol{\theta}; \mathbf{X}, \mathbf{Z})] \tag{14}$$

- Passo de Maximização (M): Encontra o parâmetro que maximiza a Equação 15:

$$\boldsymbol{\theta}^{(t+1)} = \underset{\boldsymbol{\theta}}{\arg\max}\, Q\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(t)}\right) \tag{15}$$

Em modelos que são tipicamente aplicados o EM, os dados observados $\mathbf{X}$ podem ser discretos ou contínuos, os valores que faltam $\mathbf{Z}$ são discretos, e há uma variável latente por dados observado e os parâmetros são contínuos, e são de dois tipos: parâmetros que são associados com todos os dados e parâmetros associados com um valor particular de uma variável latente.

A Figura 27 mostra uma imagem sementada utilizando o GMM com EM, tendo o GMM apenas 3 componentes.

Figura 27 – Imagem original e segmentada utilizando o GMM com 3 componentes.

Fonte: [Kittipatkampa, 2015]

3.2.5. Dice Coefficient (Sørensen-Dice Coefficient - DSC)

O coeficiente Sørensen-Dice é um método estatístico utilizado para comparar a similaridade entre duas amostras. Foi desenvolvido, coincidentemente e independentemente, pelos botânicos Thorvald Sørensen e Lee Raymond Dice, em 1948 e 1945, respectivamente [T. Sørensen, 1948], [L. R. Dice , 1945].

A Equação 16 foi desenvolvida para ser aplicada na presença e ausência de dados:

$$QS = \frac{2|X \cap Y|}{|X| + |Y|} \tag{16}$$

onde $|X|$e$|Y|$ são as quantidades de valores onde há presença de dados, ou seja, onde existe relevância. Ou seja, o numerador da equação corresponde à interseção entre as duas amostras analisadas, onde são iguais, e o denominador corresponde às posições totais que não são nulas, ou seja, possuem dados. O coeficiente varia entre 0 e 1, onde 0 corresponde a totalmente diferente e 1 corresponde a totalmente igual.

O *Dice-Coefficient* tem sido bastante útil para dados na comunidade ecológica [J. Looman and J. Campbell, 1960], já que seu uso é primariamente empírico, ao invés de teórico (embora [D. W. Roberts, 1986] mostre que ele pode ser justificado como a interseção de dois conjuntos *fuzzy*). Se comparado com a distância Euclidiana, o coeficiente possui sensibilidade em conjuntos de dados mais heterogêneos, além de dar menos peso para *outliers* [B. McCune, J. B. Grace, and D. L. Urban, 2002]. Mais recentemente, tem sido muito comum seu uso na segmentação de imagens, em particular para comparação de resultados de algoritmos com as suas respectivas máscaras em aplicações médicas.

3.2.6. T-Teste

O T-Teste é um teste de diferença entre duas médias, tanto para grupos relacionados quanto para grupos independentes [M. M. Reis, 2016]. No caso deste trabalho, o teste será aplicado entre medias para dados pareados, ou seja, a mesma população.

De acordo com o trabalho, será verificado se a média de antes é menor do que a média de depois. O ponto de partida, que servirá para a definição da hipótese H_0, é que o trabalho não faz efeito, ou seja, as médias antes e após o tratamento são iguais (costuma-se colocar em

H_0 o contrário do que se quer provar), ou seja a diferença entre as médias deve ser supostamente zero. Desse modo, tem-se a Equação 17:

$$\begin{matrix} H_0: \mu_d = 0 \\ H_0: \mu_d < 0 \end{matrix} onde\ \mu_d = \mu_{antes} - \mu_{depois} \quad (17)$$

Após esse passo, é preciso definir o nível de significância, que no trabalho será de $\alpha = 5\%$. Então, será identificada a variável de teste. Como os exemplos terão menos de 30 elementos, a variável de teste que será utilizada será a variável t_{n-1} da distribuição t-*Student*. Após, será definida a região de aceitação de H_0, de acordo com o tipo de teste e variável. Neste caso, trata-se de um teste unilateral à esquerda (com 5% de significância), ou seja, para qualquer valor superior ao valor crítico, que é o valor mostrado pela Equação 18, H_0será aceito:

$$t_{n-1}, crítico = t_{n-1}, 0.05 \quad (18)$$

O próximo passo é através dos valores das amostras de antes e depois, calcular a diferença d_i entre cada par de valores, mostrado na Equação 19:

$$d_i = x_{antes} - x_{depois} \quad (19)$$

Então, serão calculados a diferença média e o desvio padrão da diferença média, como mostrado nas Equações 20 e 21:

$$\bar{d} = \frac{\sum d_i}{n} \quad (20)$$

$$s_d = \sqrt{\frac{\sum d_i^2 - [(\sum d_i)^2/n]}{n-1}} \quad (21)$$

Ainda haverá o cálculo da variável de teste, usando a Equação 22:

$$t_{n-1} = \frac{\bar{d}}{(s_d/\sqrt{n})} \quad (22)$$

Por fim, será feita a decisão pela aceitação ou rejeição de H_0. Conforme visto anteriormente, se o valor da variável de teste for menor do que $t_{n-1}, 0.05$ a hipótese H_0 será rejeitada. Portanto, rejeitando H_0, conclui-se que, com 95% de confiança, o modelo proposto

no trabalho contribuiu para o aumento do desempenho na segmentação das imagens do cérebro.

4. MÉTODO PROPOSTO

Nesse capítulo será descrito o modelo proposto. Como ele irá funcionar, seus conceitos, suas razões e uma demonstração de como ficará uma imagem 3D ao passar pelo método proposto para segmentação. Haverá também uma demonstração de como cada fatia 2D será selecionada através da imagem 3D original.

4.1. Método Proposto

Várias técnicas de segmentação de imagens foram desenvolvidas ao longo dos últimos anos [R. Dass and S. Devi, 2012]. Contudo, o que se tem observado é que cada técnica parece ter um desempenho melhor apenas para um determinado grupo de imagens selecionado pelo criador da nova técnica. Outra opção é o desenvolvimento de novas técnicas, que aparentemente possuem bons resultados, mas que, por serem desenvolvidas em pesquisas de empresas privadas, acabam não saindo para conhecimento público seu funcionamento, dificultando seu entendimento.

Baseado nesses problemas, foi pensado um modelo que servisse não como uma técnica de segmentação em si, mas sim um método para que seja possível melhorar a segmentação de uma imagem 3D, utilizando alguns dos conceitos de reconhecimento de padrões e de seleção e combinação de classificadores. A ideia do trabalho é, ao utilizar segmentações em fatias 2D de uma imagem 3D, independentemente da técnica de segmentação 2D utilizada, melhorar o resultado final, de modo que qualquer técnica que utilize segmentação 2D possa vir a ter um ganho de desempenho no resultado final da segmentação 3D.

Tradicionalmente, as segmentações de imagens 3D são feitas de duas maneiras: ou faz-se uma análise geral da imagem 3D e a segmentação é feita de uma vez, ou então o modo mais comum, onde a segmentação é feita através de fatias 2D da imagem 3D, ocorrendo a segmentação por fatia e tendo seu resultado combinado, gerando a resposta final. Desse modo, o que mais acontece, pela facilidade de se trabalhar desse modo, é a utilização das fatias em cortes feitos nos planos mais comuns, ou seja, nos planos do espaço cartesiano:

- Plano xy: onde se varia apenas z;

- Plano xz: onde se varia apenas y;
- Plano yz: onde se varia apenas x.

Esses planos podem ser facilmente visualizados na Figura 28, que mostram o espaço cartesiano para facilitar a visualização. As segmentações tradicionais são feitas escolhendo algum desses planos, e fatiando a imagem 3D em imagens 2D na direção escolhida e, por fim, juntando os resultados também no mesmo sentido que as imagens foram fatiadas, para que se recomponha a estrutura inicial.

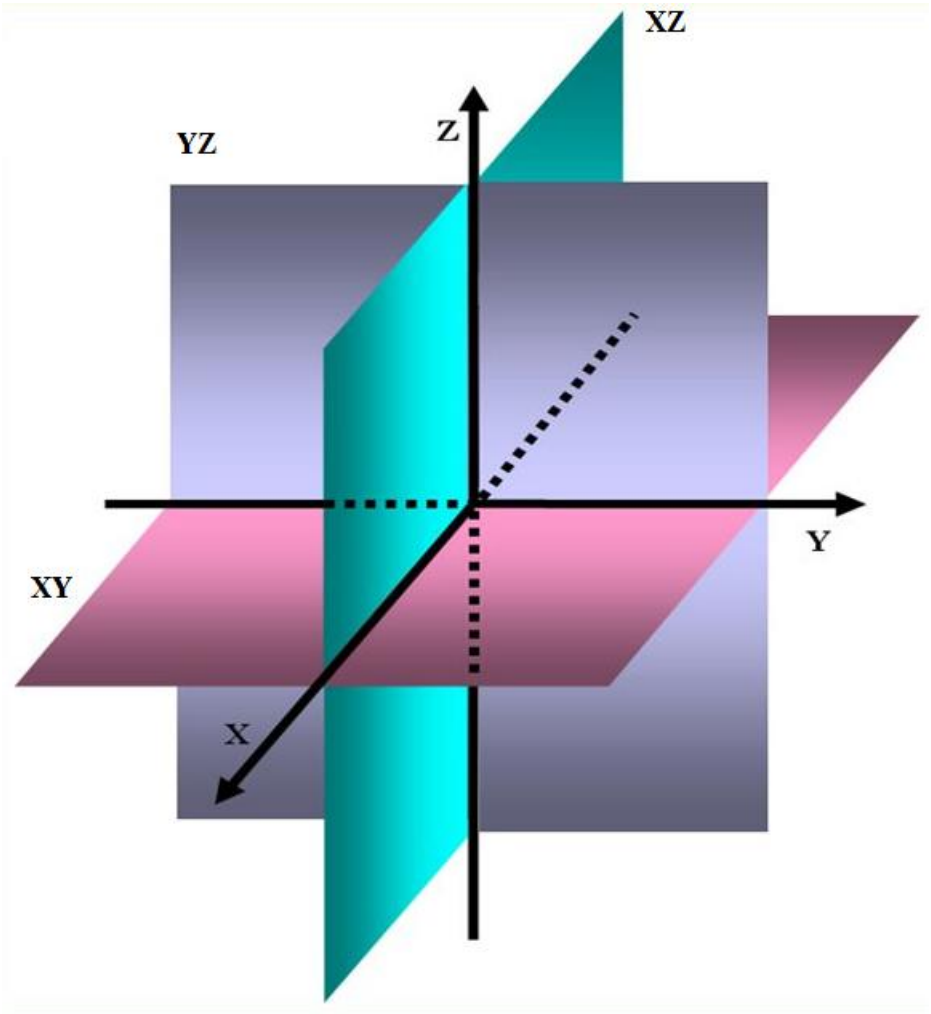

Figura 28 – Planos tradicionalmente selecionados para se fatiar uma imagem 3D.

Fonte: [I. Franca, 2016]

Utilizando o conceito criado pelo algoritmo *Bagging*, foi pensada uma maneira de se tentar diversificar e aumentar a quantidade de entradas a serem analisadas, para que se

pudesse ter respostas diferentes. Assim, haveria mais respostas e, ao se combiná-las, a chance de que a resposta teria maior precisão seria maior, assim como acontece no algoritmo *Bagging*. Contudo, para cada análise, há apenas uma única imagem 3D, que deve ser segmentada, sem mais dados adicionais. Desse modo, a ideia desenvolvida foi também segmentar a imagem 3D em fatias 2D, porém, ao invés de escolher apenas 1 dos planos citados acima, escolher todos. Desse modo, haveria mais métodos de se segmentar a mesma imagem e, no fim, obter uma resposta mais precisa, após utilizar algum método de combinação.

Quando se faz a análise de uma imagem 2D para sua posterior segmentação, é sabido que as propriedades de vizinhança são muito relevantes, e elas ajudam a influenciar diversos algoritmos para saber se determinado pixel esta dentro ou fora da área de segmentação desejada. Desse modo, ao cortar uma imagem 3D em diversas fatias 2D, acaba-se perdendo propriedades de vizinhança, dependendo do plano escolhido. Além disso, não é possível saber de maneira prévia qual vizinhança é mais relevante, de modo que pode haver perdas grandes caso se selecione o plano errado para fatiar a imagem 3D. Desse modo, o trabalho visa também atacar esse problema, tentando se aproveitar das vizinhanças diferentes que são analisadas alterando os planos das fatias. Por fim, apenas os 3 planos tradicionais podem não se mostrar suficientes para atacar os dois problemas citados. Logo, foram feitos cortes na imagem 3D em fatias em outros 6 planos, de modo que haja mais resultados diferentes para serem combinados, além de se utilizar de vizinhanças diferentes, para que se possa levar em consideração vizinhanças diferentes. As Figuras 29 a 34 mostram todos os planos selecionados. Elas estão com uma perspectiva um pouco diferente, por limitação do programa que as gerou, mas é possível visualizar todos os eixos desejados. Os limites também são arbitrários, definidos pelo programa e servem apenas para facilitar o entendimento. Foram escolhidos os planos em +45° e -45° dos 3 planos originais, gerando planos intermediários. As cores foram inseridas para facilitar a visualização, além do desenho do cubo.

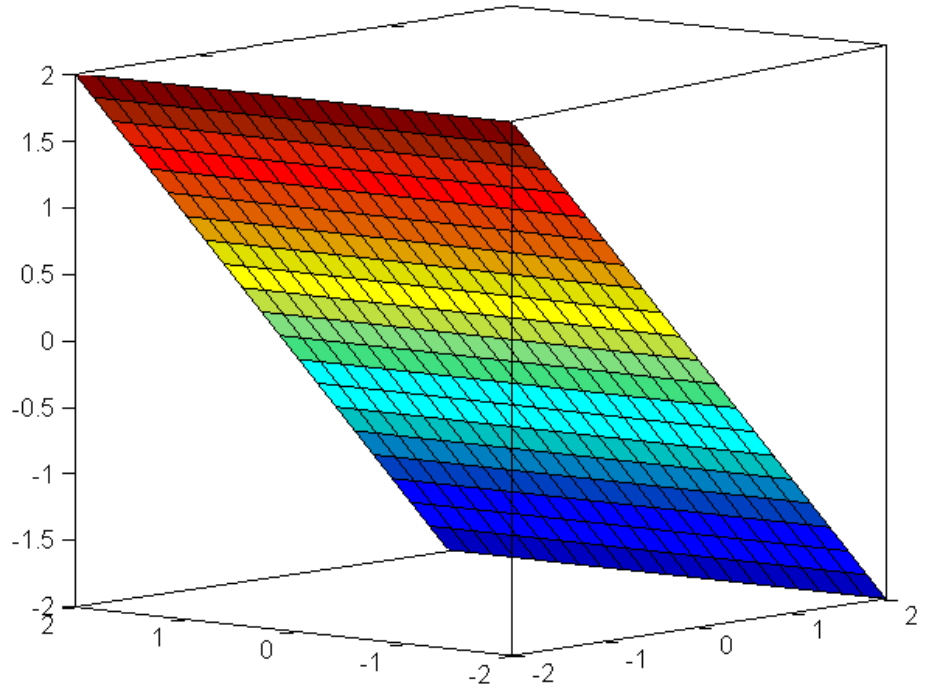

Figura 29 – Planos rotacionado +45° no plano xy.

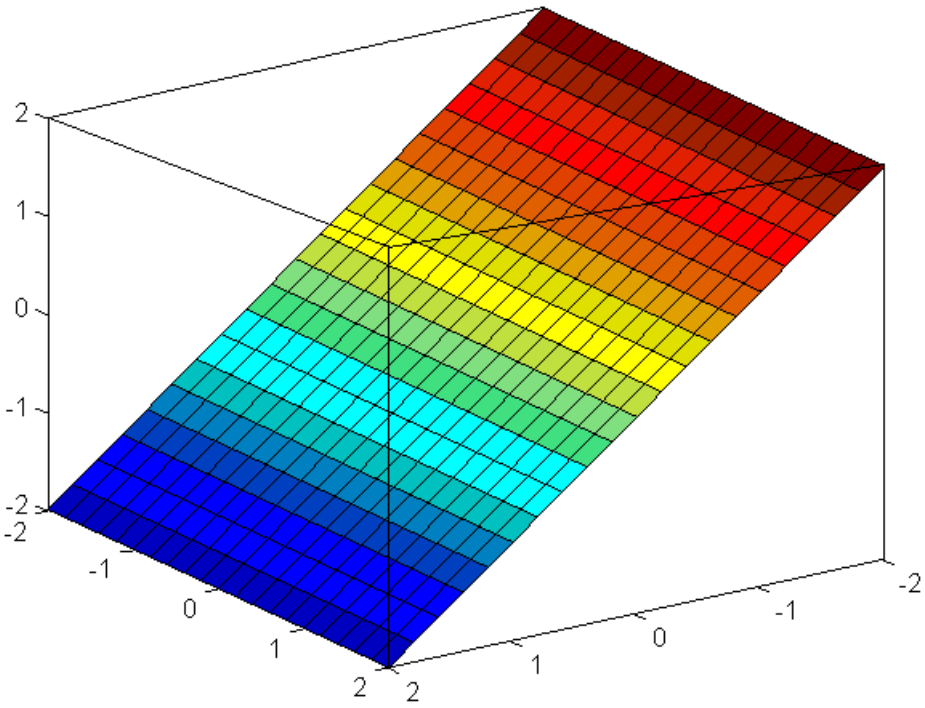

Figura 30 – Planos rotacionado -45° no plano xy.

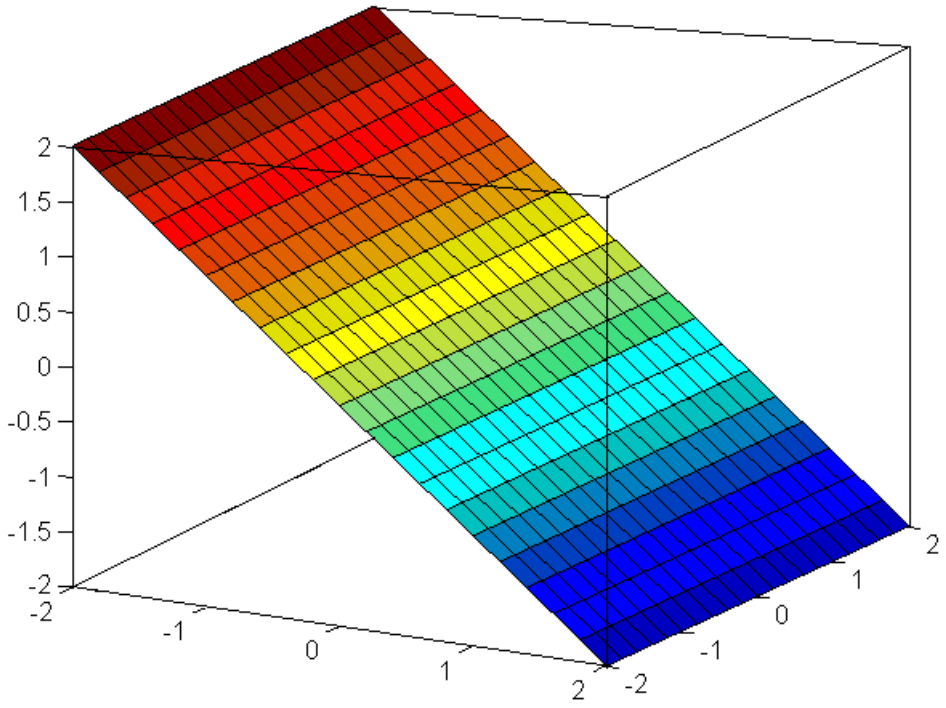

Figura 31 – Planos rotacionado +45° no plano xz.

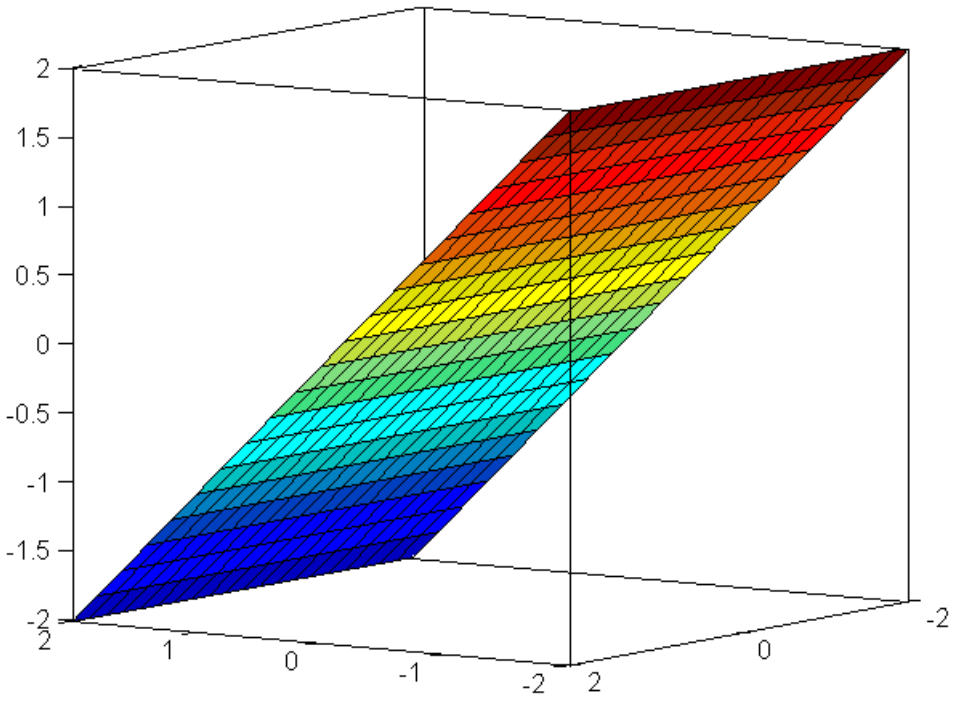

Figura 32 – Planos rotacionado -45° no plano xz.

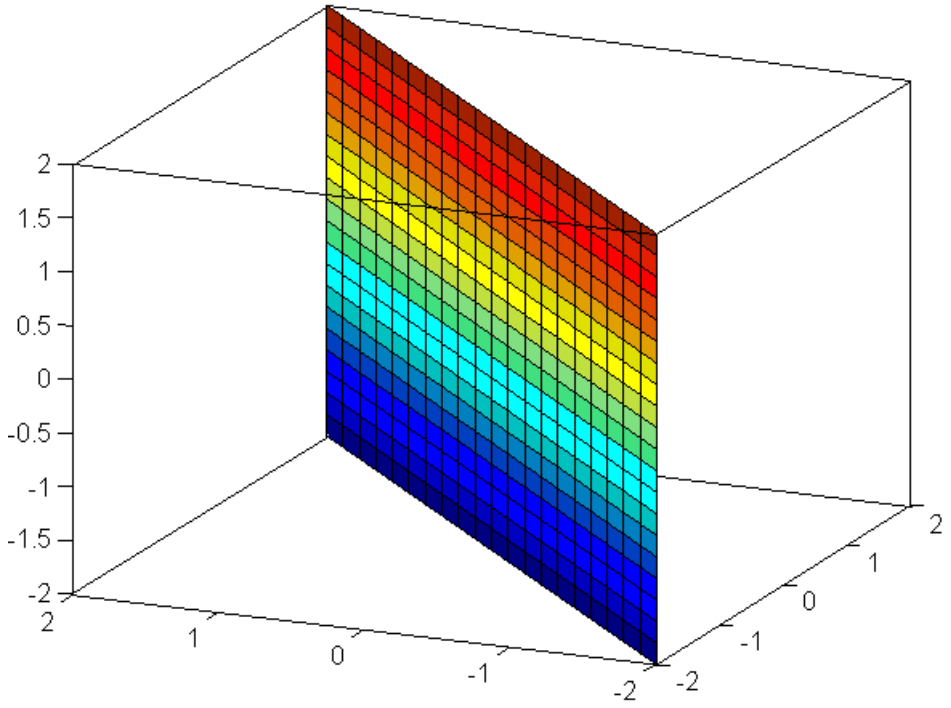

Figura 33 – Planos rotacionado +45° no plano yz.

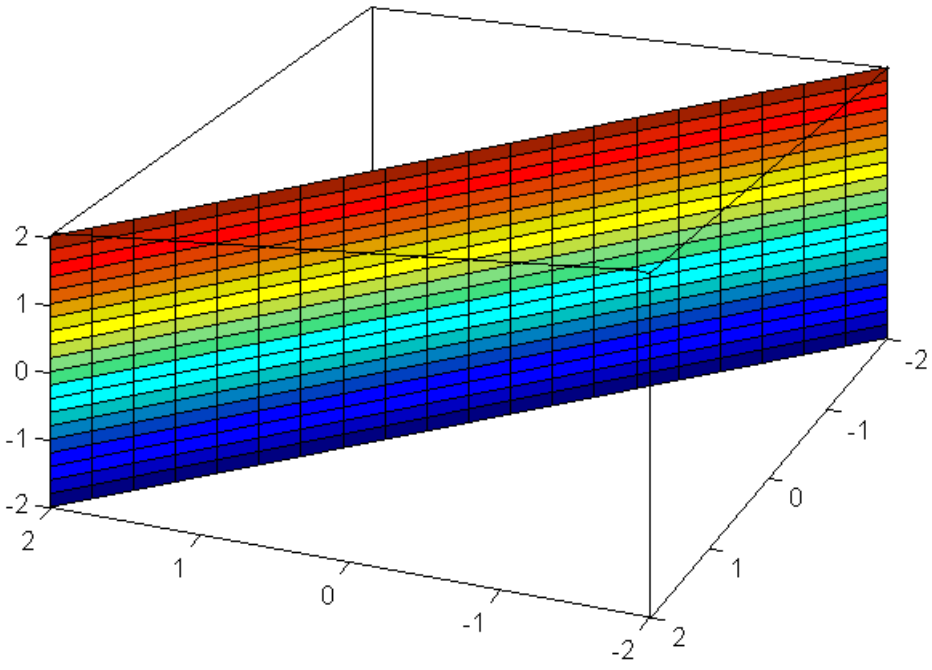

Figura 34 – Planos rotacionado -45° no plano yz.

Para demonstrar o funcionamento dessas fatias, basta visualizar a Figura 35. Ela representa as 9 fatias diferentes que o mesmo pixel (80, 128, 128) de uma imagem 3D vai estar contido. Esse pixel é representado em vermelho, para facilitar a visualização. Como pode ser facilmente observado, as vizinhanças são bastante diferentes, o que levará a segmentações também muito diferentes, aumentando assim a diversidade de respostas, que é o objetivo do trabalho proposto.

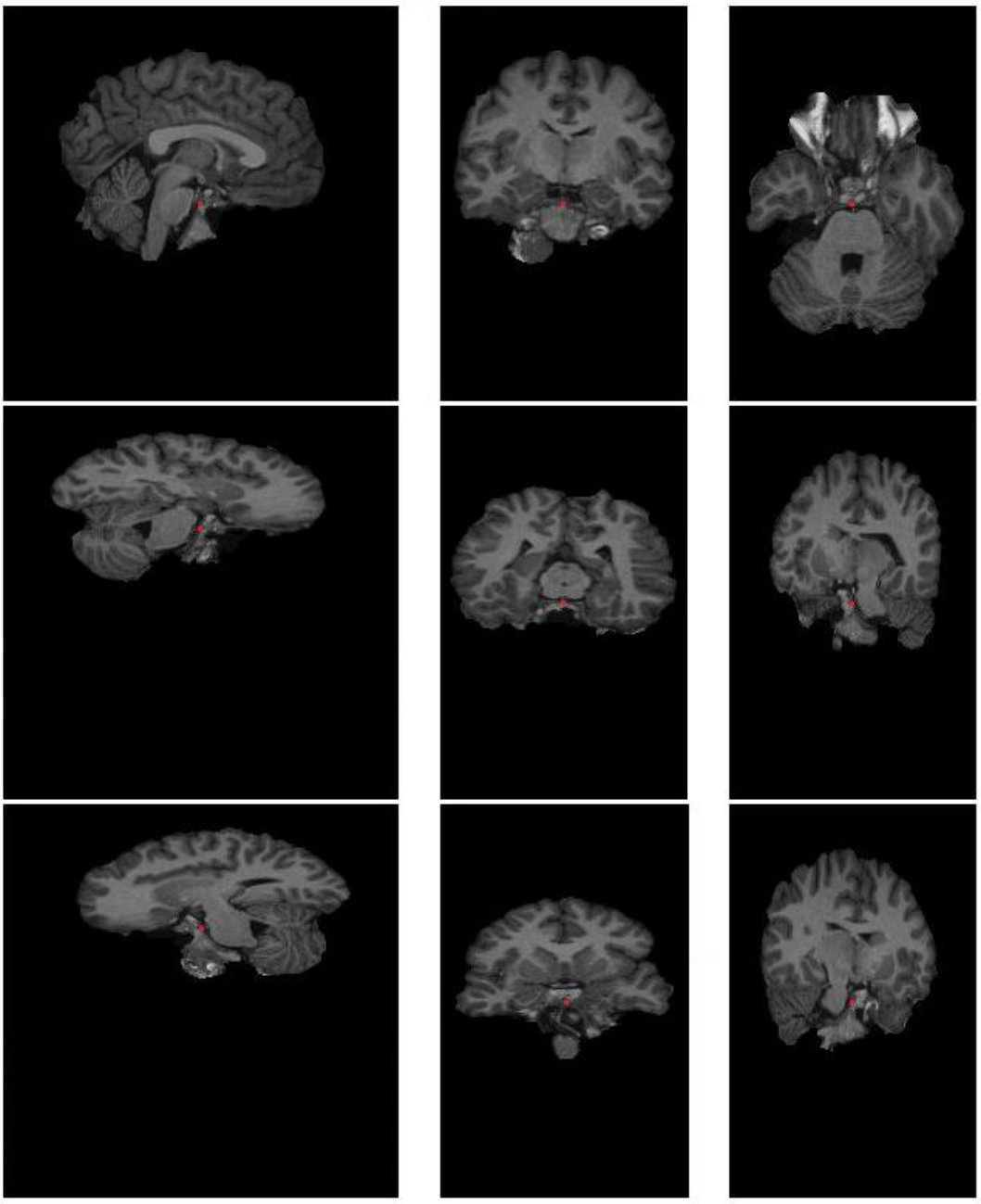

Figura 35 – Todas as fatias que o pixel (80, 128, 128) estará contido.

Assim, são feitas feita nove vezes as segmentações de cada imagem 3D, todas usando imagens 2D diferentes, mas utilizando a mesma técnica de segmentação. Cada direção do plano será escolhida isoladamente e será feita uma segmentação nessa direção, com posterior junção das fatias na mesma direção em que foram fatiadas e assim uma posterior análise. Desse modo, cada imagem 3D irá gerar 9 Imagens 3D segmentadas diferentes, que serão combinadas posteriormente através do voto majoritário. A ideia do trabalho é mostrar que, assim como o *Bagging*, tendo mais entradas para análise e fazendo a combinação dos resultados apresentados, é possível que se obtenha uma melhoria estatisticamente comprovada no resultado final da segmentação. Para verificar essa teoria, foram selecionadas das bases de dados e foram realizados experimentos e avaliações, todos detalhados no Capítulo 4.

5. DESCRIÇÃO E ANÁLISE DE EXPERIMENTOS

Nesse capítulo são mostrados os resultados dos experimentos realizados para a solução do problema proposto e uma análise sobre os experimentos é apresentada. Primeiramente, é apresentada uma validação dos algoritmos utilizados, mostrando que eles foram implementados corretamente. Após isso, são descritas as bases de dados utilizadas, e como foram feitos seus pré-processamentos. Então, são mostrados os testes realizados, tanto individualmente, quanto de maneira combinada. Por fim, é apresentada uma discussão sobre os resultados. Os algoritmos foram implementados na linguagem MatLab. Os testes foram executados em um computador com um processador Intel Core i7 3610QM, com 16GB de memória RAM e sistema operacional Windows 7.

5.1. Validação dos Algoritmos

Para validar os algoritmos, todos foram implementados em MatLab e testados usando a mesma imagem como base. Para garantir o funcionamento desejado, similar ao da base de dados utilizada nos experimentos, a imagem escolhida foi uma imagem em preto e branco, e, assim como ocorreu com a base de dados real, ela foi normalizada. Desse modo, todos os algoritmos foram executados, individualmente, e seus resultados verificados se compatíveis com o esperado. Os algoritmos utilizados foram aqueles que também foram utilizados no projeto em si: binarização de Otsu, *K-Means*, SOM e o GMM. As Figuras 35 e 36 mostram a imagem original e os resultados das segmentações.

Figura 36 – Imagem original para ser segmentada

Fonte: [C. Guyeux and J Bahi, 2010]

Figura 37 – Segmentações (a) Binarização de Otsu, (b) *K-Means*, (c) SOM e (d) GMM

Como pode ser visto, as segmentações são todas diferentes, o que mostra a diversidade dos algoritmos entre si. Desse modo, a ideia de mostrar que o modelo proposto foi testado com técnicas que apresentam resultados distintos se mostra válida.

5.2. Bases de Dados

As bases de dados utilizadas nos experimentos são bases de tomografias reais, para que não houvesse resultados gerados por bases sintéticas, diminuindo a efetividade real do modelo. As bases estão disponíveis em [MindBoogle, 2015], e estão sendo disponibilizadas com o intuito de melhorar a precisão e a consistência de classificação automática do cérebro

humano. São disponibilizadas imagens 3D de ressonâncias magnéticas, de maneira pública, além de uma segmentação feita de maneira semi-automática, com ajuda de especialistas da área, para que se tenha um *groudtruth* para que, quem use as bases, tenha uma resposta para se basear.

Foram selecionadas 42 imagens 3D de ressonância magnética, que são publicamente acessíveis sem necessidade de licenças, de pessoas saudáveis e imagens de alta qualidade para garantir uma boa reconstrução de superfície. Foram duas bases escolhidas para os experimentos: a "*Nathan Kline Institute/Rockland sample*" (NKI-RS-22) [NKI, 2016] e a "*Open access series of imaging studies test–retest* ("*reliability*") *sample*" (OASIS-TRT-20) [OASIS, 2016]. A Tabela 2 lista os dados das bases.

Tabela 2 – Dados das Bases de Dados utilizadas na segmentação.

Nome	Fonte	Número	Idade (Média, Desvio)	Sexo		Mão Principal	
				Masculino	Feminino	Direita	Esquerda
NKI-RS-22	"Nathan Kline Institute/ Rockland sample"	22	20-40 (26.0, 5.2)	12	10	21	1
OASIS-TRT-20	"Open access series of imaging studies" test–retest ("reliability") sample	20	19-34 (23.4, 3.9)	8	12	20	0

Fonte: [A. Klein and J. Tourville, 2012].

Foram selecionadas duas bases de dados para os testes, pois não há uma técnica que seja melhor para qualquer base de dados, ou seja, uma técnica pode ter um desempenho ruim na primeira base de dados, devido a problemas de contraste ou ruídos, mas um desempenho bom na segunda base, e vice versa. Desse modo, é possível observar que algumas técnicas podem ser úteis, dependendo de onde elas estiverem sendo aplicadas. As Figuras 37 e 38 mostram duas ressonâncias magnéticas, uma de cada base de dados, para mostrar a imagem original, antes de qualquer processamento. As figuras mostram essas ressonâncias nos três eixos, sendo possível observar qualquer fatia da imagem em qualquer eixo desejado do plano cartesiano. É possível perceber que elas são um pouco diferentes entre si. Isso se deve a variações tanto da pessoa quanto das máquinas utilizadas para serem feitas as ressonâncias.

Desse modo, é possível mostrar um pouco mais a robustez do trabalho desenvolvido, uma vez que não fica limitado a um tipo especifico de imagem.

Figura 38 – Imagem de ressonância magnética da base base NKI-RS-22

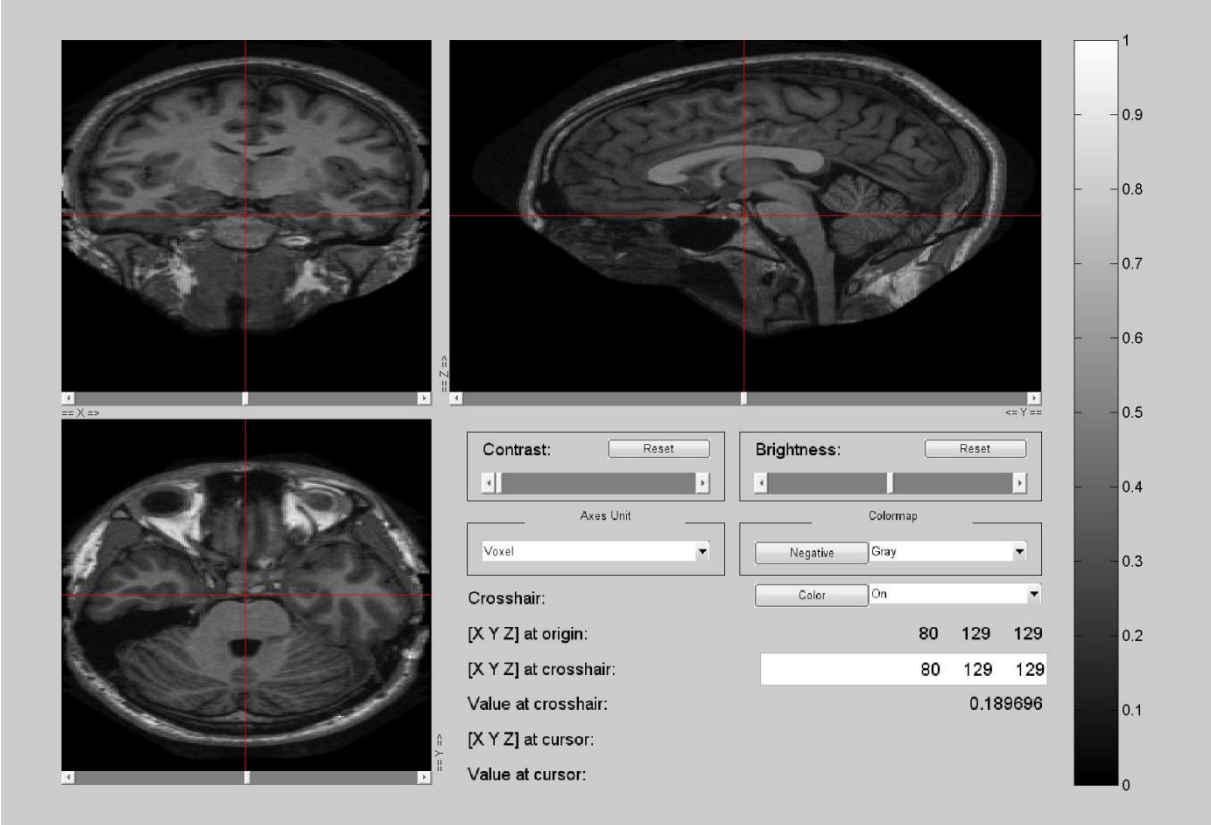

Figura 39 – Imagem de ressonância magnética da base OASIS-TRT-20

As bases utilizadas possuem 22 e 20 imagens, respectivamente. Contudo, elas vieram com uma imagem sem *groundtruth* em cada uma delas. As imagens 16 da base NKI-TRT-22 e 18 da base OASIS-TRT-20 foram descartadas do processamento. Portanto, foram analisadas 21 imagens da primeira base e 19 da segunda, totalizando 40 imagens.

Os *groundtruths* disponibilizados possuem apenas a segmentação da matéria cinza do cérebro. Desse modo, apesar de o algoritmo também realizar a segmentação da matéria branca, apenas a matéria cinzenta será levada em consideração. As imagens 39 e 40 mostram como são mostradas as matérias cinza, também como imagens 3D sendo visualizadas nos 2 eixos do plano cartesiano.

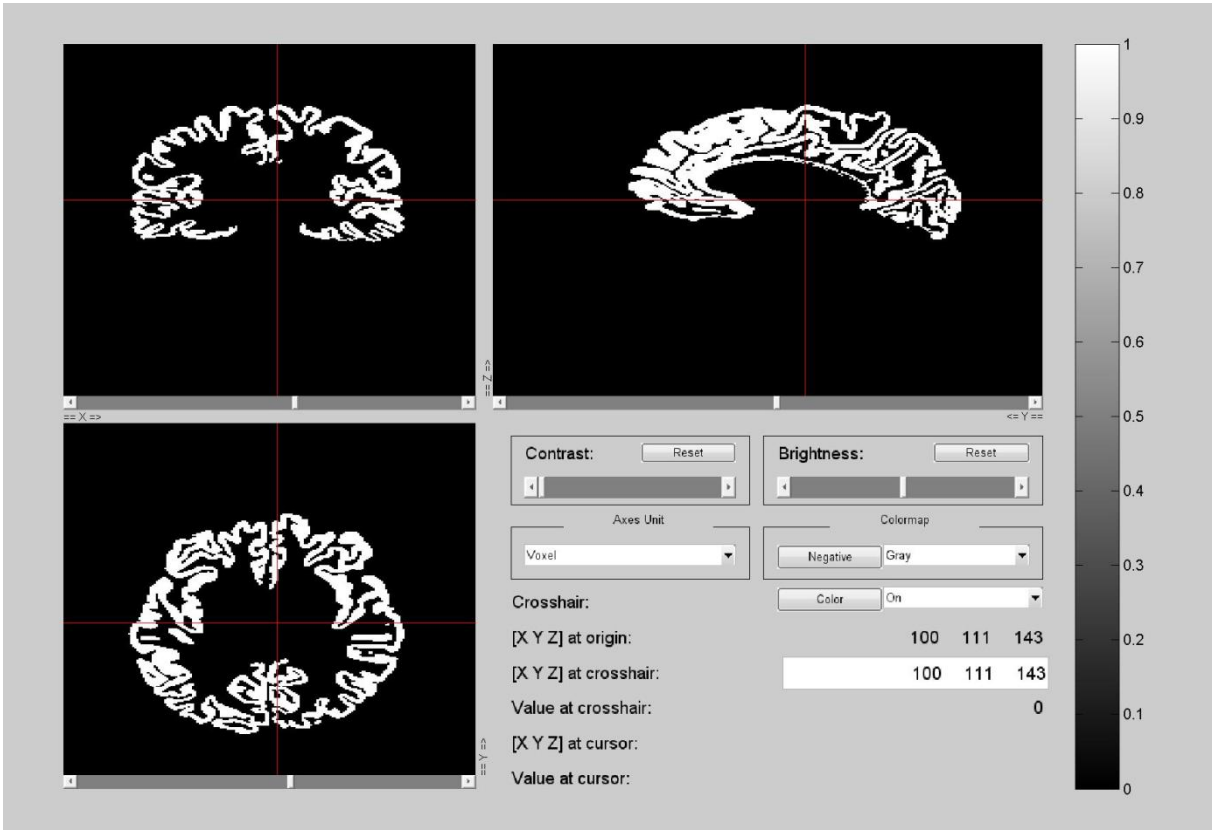

Figura 40 – *Groundtruth* da matéria cinza de imagem da base NKI-RS-22

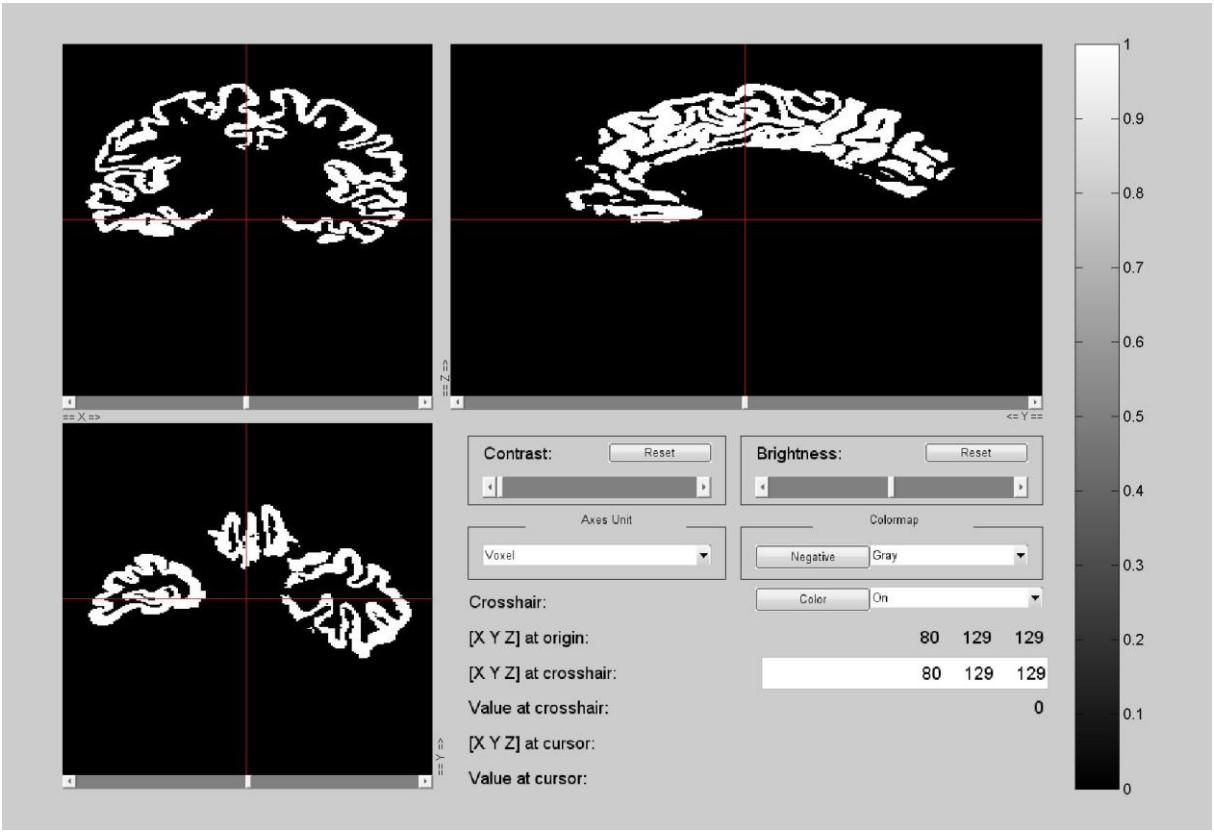

Figura 41 – *Groundtruth* da matéria cinza de imagem da base OASIS-TRT-20

5.3. Pré-Processamento

Como há muita informação desnecessária para a segmentação das matérias branca e cinza de cada cérebro, como o crânio e o cerebelo, é necessário que se faça um pré-processamento para remover essas partes indesejadas. Antes de qualquer coisa, as imagens foram normalizadas, para um intervalo entre zero e um, para que os algoritmos funcionem de maneira semelhante para todos, além de evitar discrepâncias e facilitar no cálculo de erro e acerto ao final das segmentações.

Após a normalização, o primeiro pré-processamento foi a remoção do crânio, ou extração do cérebro. Para essa etapa, foi utilizado o algoritmo BET (*Brain Extraction Tool*) [S. M. Smith, 2002], disponível dentro da biblioteca FSL no *Software* para Linux FMRIB *Software Library* v5.0 [M. W. Woolrich et al., 2009], [S. M. Smith et al., 2004], [M. Jenkinson, 2012], desenvolvido pela *Oxford Centre for Functional MRI of the Brain* [N. D. of Clinical Neurosciences, 2016]. O algoritmo apaga tecidos que não são do cérebro de uma imagem de uma cabeça completa, usando um mapeamento do cérebro humano. As mesmas imagens mostradas anteriormente do cérebro, após a utilização do BET, correspondem às Figuras 41 e 42. Como pode ser visto, agora restaram apenas os cérebros, sem os crânios.

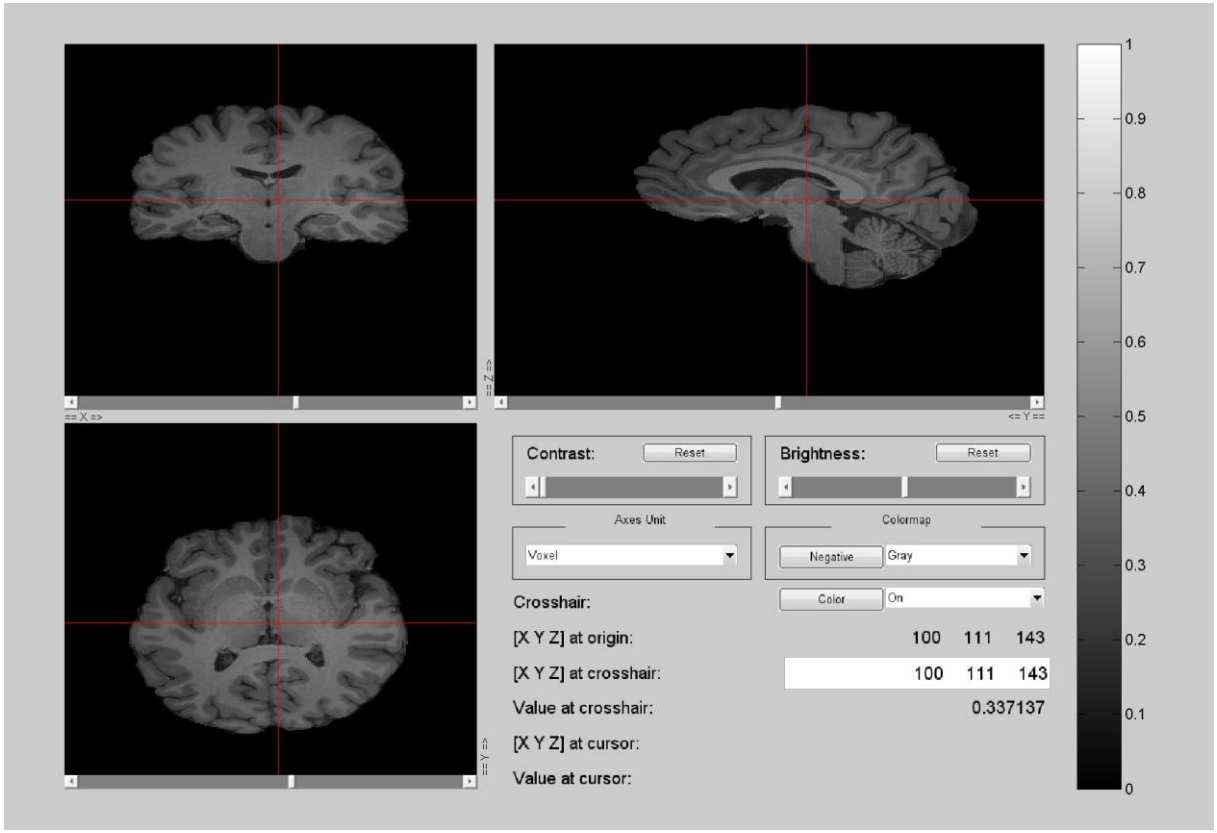

Figura 42 – Cérebro de imagem da base NKI-RS-22, sem o crânio

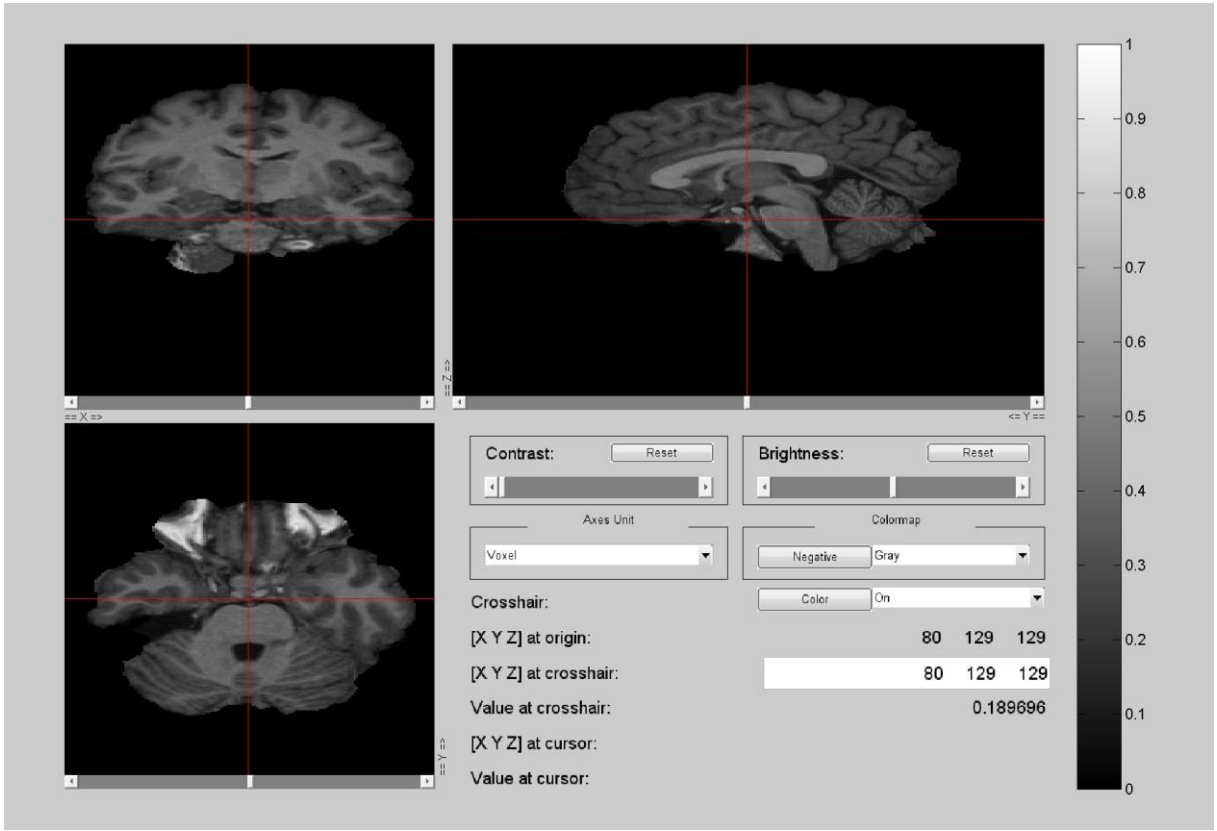

Figura 43 – Cérebro de imagem da base OASIS-TRT-20, sem o crânio

Após a normalização e a extração do cérebro, ainda falta um passo para terminar o pré-processamento: a remoção do cerebelo. Como este não faz parte da área a ser avaliada na segmentação, sua remoção em fase anterior facilita bastante a segmentação final. Para realizar essa remoção do cerebelo, foi utilizada a *Toolbox* SUIT [SUIT, 2015], [J. Diedrichsen, 2006], [J. Diedrichsen et al., 2009], [J. Diedrichsen et al., 2011], [J. Diedrichsen and E. Zotow , 2015], disponível para MATLAB. Para o uso dessa *Toolbox*, é necessária também a instalação da *Toolbox* SPM (*Statistical Parametric Mapping*) [SPM, 2015], desenvolvido pela *Wellcome Trust Centre for Neuroimaging*, da *University College London* (UCL) [W. T. C. for Neuroimaging, 2016]. O algoritmo se baseia na anatomia do cerebelo de 20 indivíduos saudáveis. Usando técnicas estatísticas para determinar os locais esperados da estrutura anatômica, a remoção é feita, preservando os detalhes das estruturas do cerebelo. As Figuras 43 e 44 mostram como ficam as imagens mostradas anteriormente após a remoção do cerebelo. Como pode ser visto, as remoções não são perfeitas, mas se mostram muito úteis para o resultado final da segmentação.

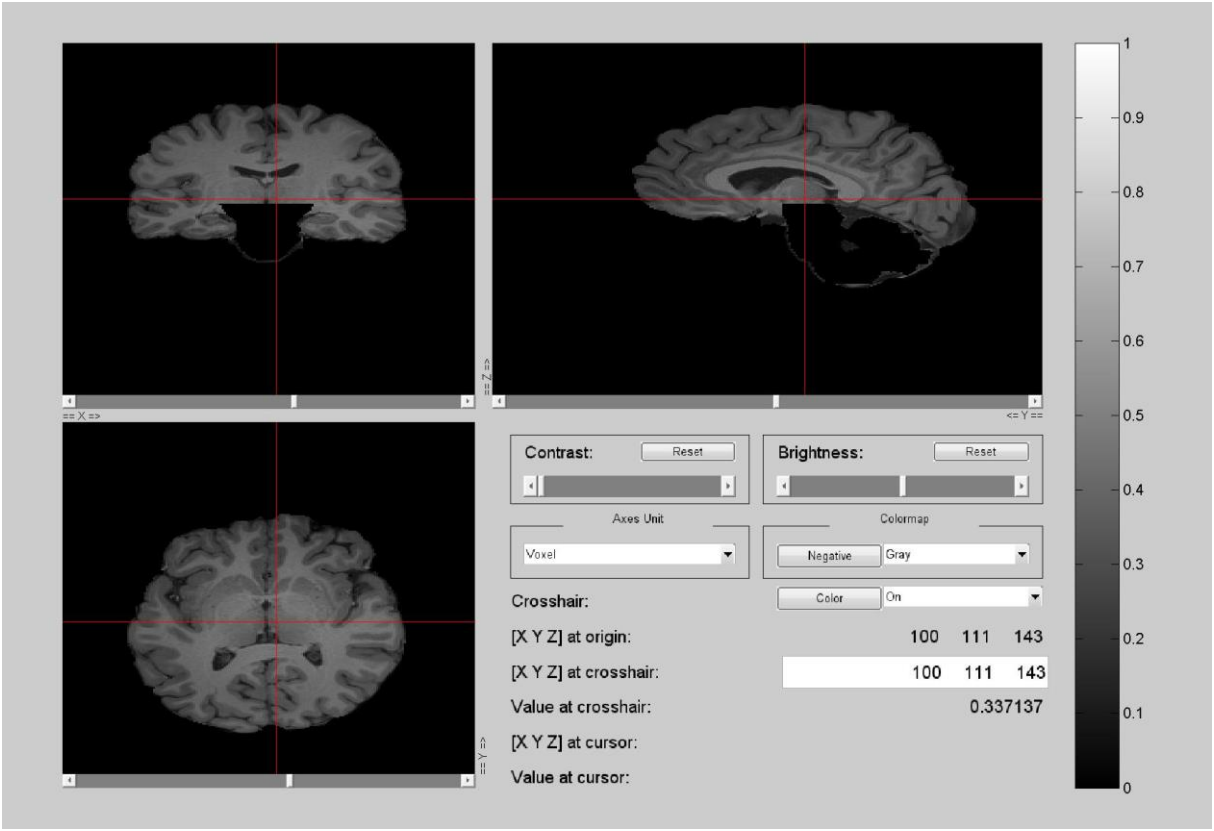

Figura 44 – Cérebro de imagem da base NKI-RS-22, sem crânio e cerebelo

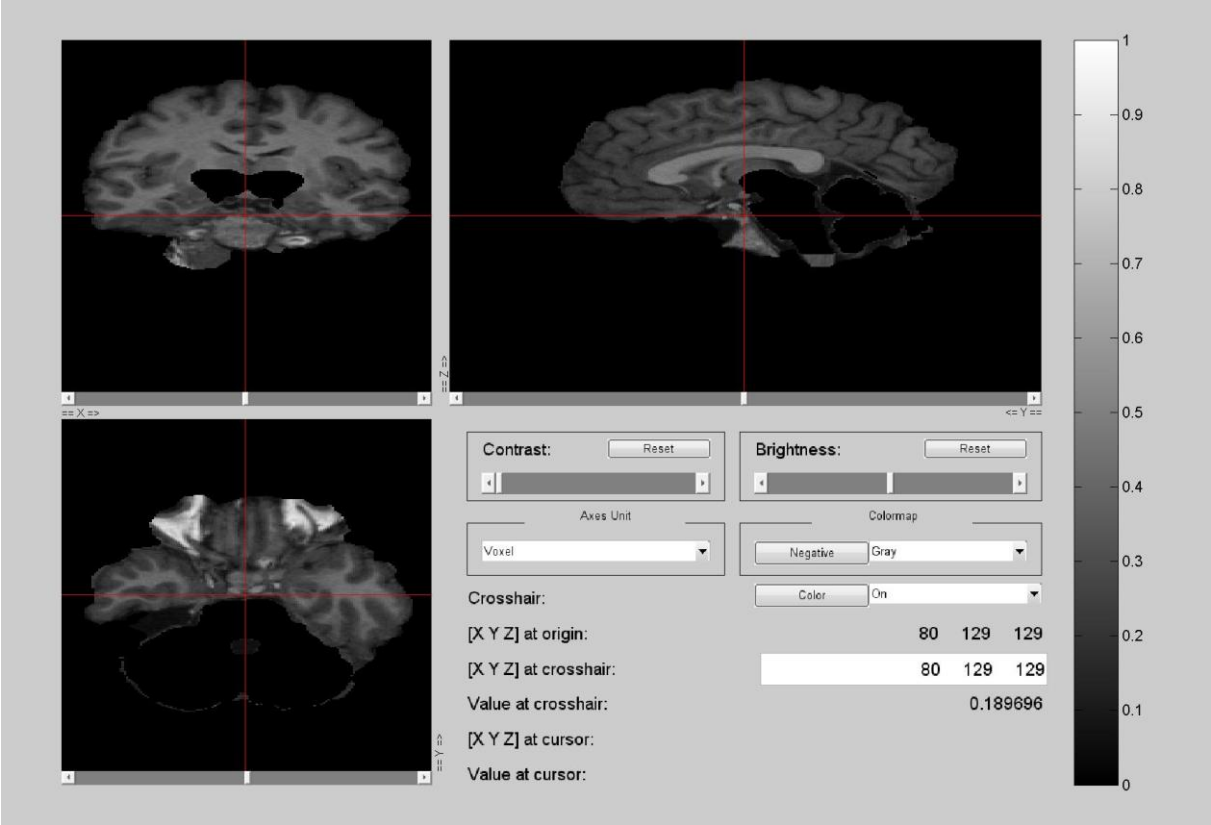

Figura 45 – Cérebro de imagem da base OASIS-TRT-20, sem crânio e cerebelo

Assim, termina-se o pré-processamento. Desse modo, a segmentação final terá um resultado mais preciso, além de possuir menor área de análise, o que diminui o tempo necessário para segmentar cada imagem.

5.4. Experimentos para Avaliar a Proposta

Utilizando os algoritmos descritos no Capítulo 4, e com as bases de dados descritas acima, pode-se iniciar os experimentos. Primeiramente, a segmentação será feita utilizando uma das técnicas mais simples para isso, a binarização de Otsu. Depois, será utilizado o algoritmo *K-Means*. Devido aos seus resultados semelhantes, eles também terão suas respostas combinadas, de modo a se alcançar melhores resultados. Na próxima segmentação, será utilizado uma rede neural, o SOM, e, por fim, a segmentação será feita usando o GMM. Nas três primeiras técnicas citadas, a base utilizada foi a OASIS-TRT-20. Contudo, o GMM utiliza uma base de dados diferente, a NKI-TRT-22, pois ele não se mostrou um bom algoritmo para a base utilizada pelas outras técnicas. Em compensação, obteve um bom

resultado nesta outra base, a qual o SOM não obteve um bom resultado, demonstrando que não há um algoritmo de segmentação que seja melhor para qualquer tipo de imagem.

O modo de verificação do desempenho dos algoritmos utilizado foi o *Dice-Coefficient*. Dado que a imagem possui boa parte como um fundo preto, é interessante utilizar uma técnica que descarte essa parte, pois isso poderia levar a um percentual de acerto maior, mas incorreto. Assim, a técnica escolhida não leva em consideração as partes que são definidas como fundo, apenas os lugares onde há informação. Desse modo, os resultados levam em consideração apenas as partes onde realmente há segmentação, mostrando o real desempenho do modelo. Essa técnica é utilizada para a avaliação em competições internacionais [MRBrainS, 2014] de segmentação do cérebro.

5.4.1. Experimentos Utilizando Binarização de Otsu

Nesse primeiro experimento, a segmentação foi feita utilizando a binarização de Otsu. Essa binarização foi feita individualmente em cada fatia 2D selecionada, em cada um dos nove eixos descritos no Capítulo 3. Desse modo, cada imagem possui limiares diferentes, já que são calculados limiares específicos para cada imagem baseados no histograma de cada imagem 2D. Como a ideia é segmentar não apenas a matéria cinza, mas também a matéria branca, mais de um limiar é calculado para cada imagem. Assim, a Figura 45 representa o resultado da combinação das segmentações com as imagens 2D sendo selecionadas em cada eixo escolhido. A base de dados utilizada foi a OASIS-TRT-20.

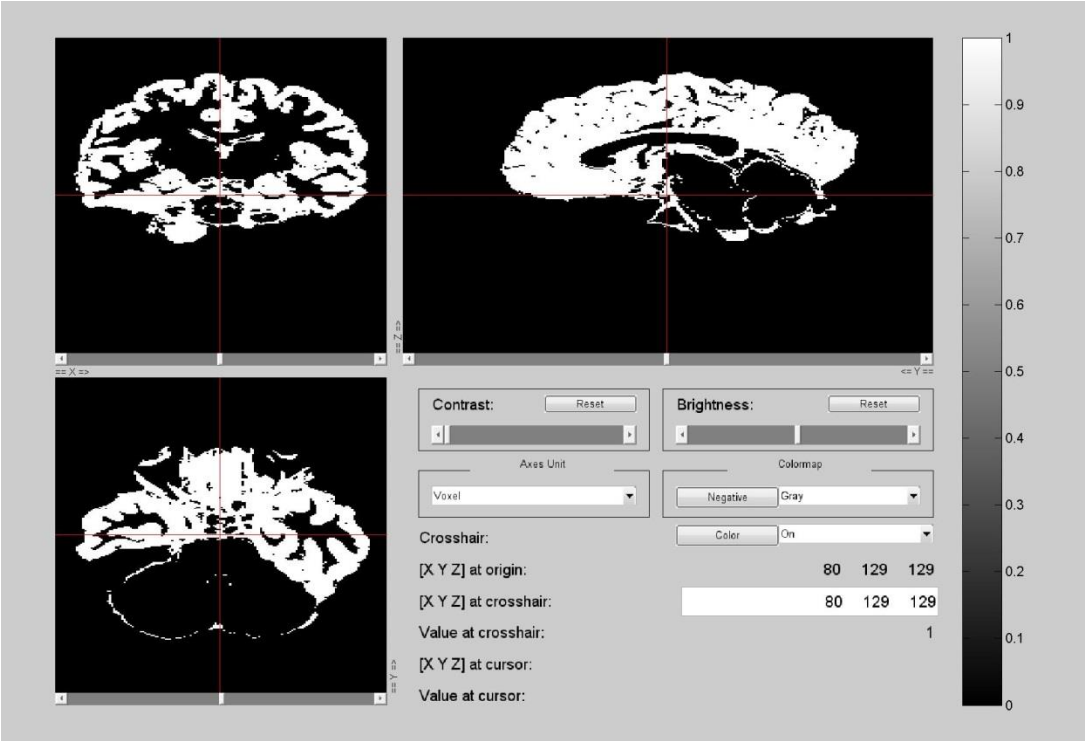

Figura 46 – Combinação das segmentações utilizando a binarização de Otsu, de imagem da base OASIS-TRT-20

Como pode ser visto na imagem, a segmentação possui erros, especialmente na parte mais inferior da cabeça. Assim, a Tabela 3 mostra o resultado do DSC de cada segmentação com o *groundtruth*, comparando os valores individuais sobre cada eixo com o resultado da combinação final entre todas elas.

Tabela 3 – Valores do DSC das segmentações usando a binarização de Otsu.

Eixo	1	2	3	4	5	6	7	8	9	V.M.
Imagem										
1	0,70677	0,71982	0,7105	0,71964	0,72259	0,72472	0,71673	0,71993	0,72492	**0,7342**
2	0,72444	0,74427	0,74031	0,73635	0,73423	0,74451	0,74461	0,73977	0,74214	**0,7548**
3	0,732	0,75581	0,73789	0,74933	0,74379	0,75628	0,74897	0,74521	0,75494	**0,7638**
4	0,70441	0,72815	0,71705	0,73051	0,73324	0,73732	0,72441	0,73248	0,7357	**0,7488**

5	0,71814	0,73556	0,73357	0,73084	0,73665	0,73994	0,73589	0,73349	0,74004	**0,7504**
6	0,70621	0,73334	0,73456	0,73285	0,72912	0,73542	0,73295	0,72873	0,73478	**0,7476**
7	0,72127	0,74024	0,73359	0,73529	0,73778	0,74011	0,73588	0,73744	0,74229	**0,754**
8	0,74307	0,76123	0,75145	0,75117	0,75393	0,75697	0,75698	0,75072	0,76073	**0,7683**
9	0,70739	0,73653	0,73338	0,73454	0,73208	0,73686	0,74087	0,73464	0,73789	**0,7523**
10	0,70085	0,72649	0,72748	0,72131	0,72087	0,72598	0,73237	0,72363	0,7239	**0,7411**
11	0,73893	0,75537	0,75207	0,74804	0,74801	0,75577	0,75927	0,75142	0,75443	**0,7633**
12	0,70598	0,73034	0,7232	0,7243	0,72688	0,73333	0,73276	0,72942	0,73138	**0,744**
13	0,74277	0,75776	0,75668	0,75752	0,75772	0,76032	0,75769	0,75748	0,76288	**0,7699**
14	0,68819	0,71491	0,70792	0,7121	0,7076	0,71569	0,71341	0,70552	0,71423	**0,7274**
15	0,70678	0,73961	0,73539	0,73197	0,73436	0,74236	0,74639	0,7392	0,7395	**0,7557**
16	0,70688	0,7259	0,72945	0,72158	0,72236	0,72868	0,72508	0,72369	0,7282	**0,7411**
17	0,71081	0,72012	0,7229	0,71982	0,72199	0,72685	0,71476	0,72111	0,72613	**0,7348**
19	0,73203	0,74837	0,74236	0,74231	0,74155	0,7503	0,75018	0,74367	0,74756	**0,7611**
20	0,68312	0,70262	0,71196	0,70307	0,70651	0,70758	0,69692	0,70001	0,70483	**0,7207**

Os resultados acima mostram que, independentemente do eixo escolhido, todas as combinações obtiveram um resultado melhor do que a segmentação individual. Além disso, não é possível saber qual eixo é o mais adequado, uma vez que, para cada imagem analisada, um eixo diferente obtém o melhor resultado, o que faz com que o resultado da combinação acabe se tornando melhor. Levando em consideração que, na grande maior parte das vezes, os eixos utilizados tradicionalmente nas segmentações são os eixos cartesianos padrões, X, Y e

Z, o resultado da combinação com esses eixos especificamente mostra-se ainda melhor. A comparação detalhada será mostrada na Seção 4.5.

5.4.2. Experimentos Utilizando *K-Means*

No segundo experimento, a segmentação foi feita utilizando o algoritmo de *K-Means*. Assim como na segmentação anterior, ela foi feita individualmente em cada fatia 2D selecionada, novamente em cada um dos nove eixos descritos no Capítulo 3. Foram definidos 3 valores de limiares, que foram ordenados para que o menor valor sempre corresponda ao fundo e o maior sempre corresponda à matéria branca, para que cada valor represente uma parte do cérebro (matéria branca, matéria cinza e o resto). Novamente, cada imagem 2D possui limiares diferentes, levando em consideração as características especificas de cada imagem. Como a segmentação usando o algoritmo *K-Means* teve desempenho inferior ao da segmentação usando a binarização de Otsu, foi feita uma junção dos resultados das duas segmentações. Essa junção foi feita, pois o tempo de execução dessas técnicas foi relativamente baixo, em comparação com outras técnicas utilizadas. Desse modo, a combinação dos dois métodos não foi computacionalmente custosa. Assim, a Figura 46 representa o resultado da combinação das segmentações com as imagens 2D tanto utilizando o algoritmo *K-Means*, quanto a binarização de Otsu, sendo selecionadas em cada eixo escolhido. A base de dados utilizada foi a OASIS-TRT-20.

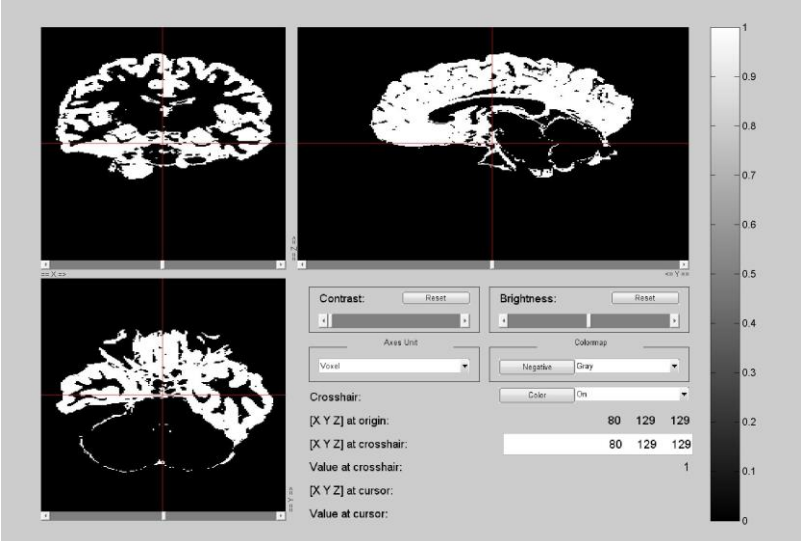

Figura 47 – Combinação das segmentações utilizando *K-Means* e a binarização de Otsu, de imagem da base OASIS-TRT-20

Como pode ser visto na imagem, a segmentação também possui erros, e parece bastante com o resultado da segmentação apenas utilizando Otsu. A Tabela 4 mostra o resultado do DSC de cada segmentação com o *groundtruth*, comparando os valores individuais sobre cada eixo utilizando o algoritmo *K-Means* com o resultado da combinação final entre todas elas e as segmentações feitas pela binarização de Otsu.

Tabela 4 – Valores do DSC das segmentações usando *K-Means* com o Voto Majoritário combinando as segmentações por *K-Means* com as por binarização de Otsu.

Eixo	1	2	3	4	5	6	7	8	9	V.M.
Imagem										
1	0,61639	0,69974	0,71224	0,65516	0,68209	0,69536	0,71754	0,67899	0,69729	**0,73658**
2	0,631	0,71642	0,72465	0,68588	0,70216	0,69035	0,73789	0,69445	0,69573	**0,75684**
3	0,69986	0,75584	0,73931	0,7287	0,72661	0,75651	0,75177	0,7402	0,75515	**0,76636**
4	0,57081	0,68341	0,66319	0,58873	0,60961	0,61662	0,71691	0,59671	0,64884	**0,74786**
5	0,63481	0,72755	0,69609	0,6785	0,68334	0,71721	0,73869	0,69148	0,71881	**0,75203**
6	0,67464	0,71953	0,73465	0,70989	0,71802	0,71253	0,73509	0,69469	0,71892	**0,74929**
7	0,61575	0,72112	0,682	0,66378	0,66156	0,70921	0,73693	0,66416	0,70566	**0,75522**
8	0,71773	0,75929	0,74884	0,75197	0,74324	0,74574	0,75753	0,74931	0,74714	**0,77281**
9	0,63297	0,7134	0,73149	0,69466	0,70919	0,71046	0,73919	0,71388	0,70463	**0,75553**
10	0,65841	0,72451	0,72993	0,71602	0,70793	0,7159	0,73446	0,71065	0,71133	**0,74261**
11	0,71374	0,74633	0,75273	0,72551	0,74174	0,74059	0,75513	0,7496	0,73701	**0,76745**
12	0,64349	0,68836	0,72476	0,69048	0,69764	0,71563	0,72648	0,72572	0,71367	**0,74718**

13	0,69245	0,71929	0,75751	0,71991	0,7515	0,72575	0,75841	0,75573	0,72292	**0,77481**
14	0,61233	0,68797	0,71054	0,67921	0,68989	0,65502	0,70869	0,67202	0,67791	**0,72924**
15	0,67348	0,74023	0,73671	0,73284	0,73127	0,72171	0,74733	0,7295	0,73552	**0,75766**
16	0,65077	0,71103	0,71241	0,67951	0,69103	0,68843	0,72665	0,70248	0,70178	**0,74386**
17	0,57971	0,69467	0,63182	0,63778	0,61027	0,64639	0,71769	0,60782	0,6419	**0,73797**
19	0,64826	0,72778	0,71706	0,66701	0,7051	0,71228	0,75068	0,69865	0,69422	**0,76472**
20	0,6422	0,68508	0,68722	0,66512	0,69078	0,67806	0,69641	0,66918	0,67483	**0,72661**

Os resultados da Tabela 4 mostram que, mais uma vez, independentemente do eixo escolhido, todas as combinações obtiveram um resultado melhor do que a segmentação individual utilizando apenas as segmentações por *K-Means* como parâmetro. O resultado se deve, principalmente, ao fato de ter sido usada a segmentação por Otsu juntamente, que obteve desempenho superior à segmentação por *K-Means*. Contudo, cabe observar que, mesmo *K-Means* obtendo um desempenho inferior a Otsu, o resultado da combinação das segmentações das duas técnicas obteve desempenho superior em 18 das 19 imagens se comparada à combinação utilizando apenas Otsu, mesmo essas diferenças sendo pequenas. Uma comparação mais detalhada entre as combinações será mostrada na Seção 4.5.

5.4.3. Experimentos Utilizando Redes Neurais (SOM)

No terceiro experimento, a segmentação foi feita utilizando uma rede neural, o SOM. Assim como nas segmentações anteriores, ela foi feita individualmente em cada fatia 2D selecionada, novamente em cada um dos nove eixos descritos no Capítulo 3. Contudo, dessa vez foi necessário um treinamento para que fossem definidos os 3 valores de limiares, para que cada valor represente uma parte do cérebro (matéria branca, matéria cinza e o resto). Então, foi selecionada uma fatia da imagem 18, que foi escolhida no mesmo eixo dos cortes feitos para a segmentação no momento, já que o treinamento do SOM é não supervisionado e não exige um *groundtruth* para ser feito. Contudo, nesse caso, os limiares eram os mesmos

para todas as fatias analisadas no eixo em questão, diferentemente dos exemplos anteriores, que cada imagem gerava seus próprios limiares. Assim, a Figura 47 representa o resultado da combinação das segmentações com as imagens 2D sendo selecionadas em cada eixo escolhido utilizando o algoritmo SOM. A base de dados utilizada foi a OASIS-TRT-20.

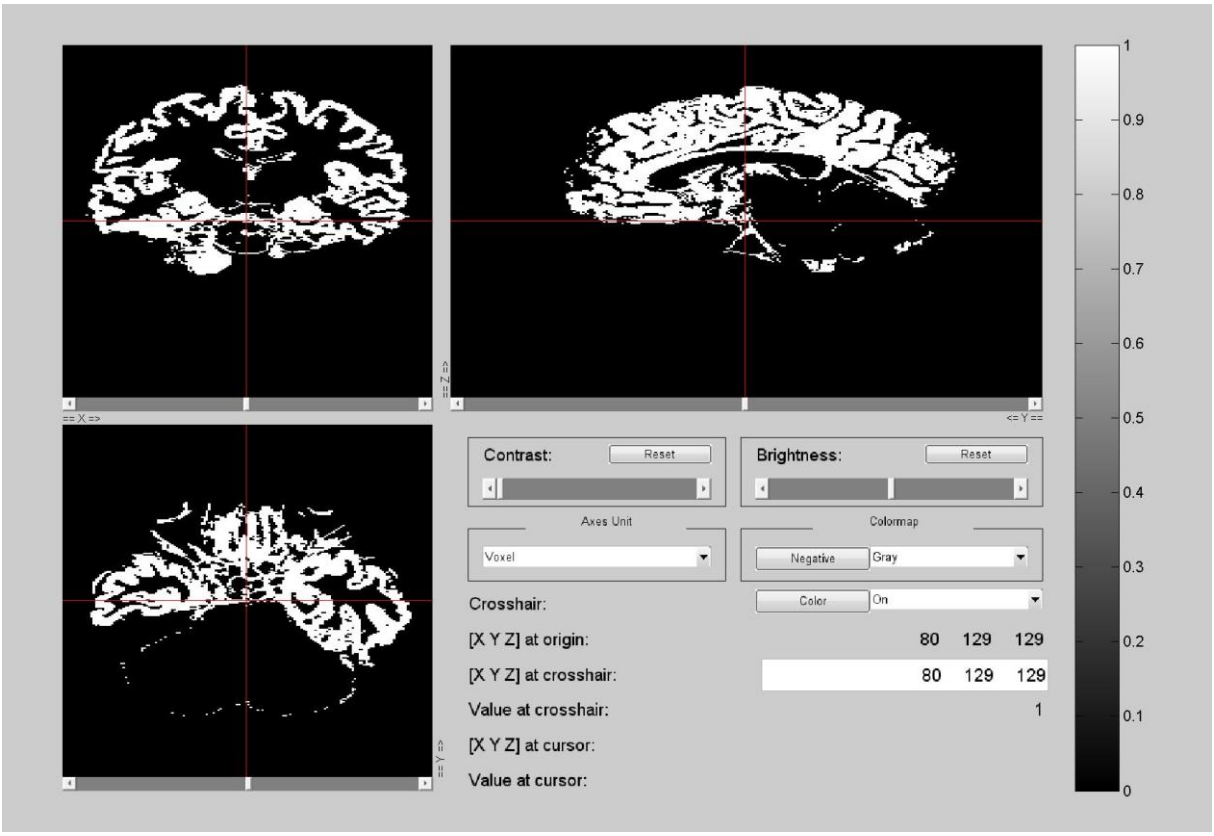

Figura 48 – Combinação das segmentações utilizando o SOM, de imagem da base OASIS-TRT-20

Como pode ser visto na imagem, a segmentação continua com boa parte dos erros e também parece com os resultados das segmentações anteriores. Contudo, é possível perceber que as linhas estão mais finas e é possível ver mais detalhes na parte inferior da primeira imagem. A Tabela 5 mostra o resultado do DSC de cada segmentação com o *groundtruth*, comparando os valores individuais sobre cada eixo com o resultado da combinação final entre todas elas, utilizando apenas o SOM.

Tabela 5 – Valores do DSC das segmentações usando o SOM.

Eixo / Imagem	1	2	3	4	5	6	7	8	9	V.M.
1	0,73618	0,79718	0,80456	0,81255	**0,81868**	0,78586	0,7698	0,70823	0,7674	0,81737
2	0,83333	0,82809	0,68501	0,72346	0,7824	0,8237	0,58673	0,49486	**0,83195**	0,79763
3	0,80839	**0,83852**	0,76934	0,79366	0,82646	0,83069	0,69634	0,61158	0,8246	0,83064
4	0,79657	**0,83355**	0,77178	0,7943	0,82243	0,8261	0,70156	0,61271	0,81756	0,82268
5	0,81251	**0,81822**	0,70107	0,73387	0,78523	0,80988	0,61259	0,52514	0,81442	0,80238
6	0,80612	0,83898	0,76928	0,79355	0,8305	0,82737	0,70175	0,62827	0,82182	**0,84352**
7	0,78042	0,82453	0,78302	0,80149	0,82833	0,81217	0,72895	0,65768	0,80207	**0,83777**
8	0,75525	0,81463	0,81346	0,82306	0,83441	0,80155	0,77428	0,71196	0,78418	**0,83768**
9	0,76913	0,81349	0,77314	0,79178	0,81757	0,80017	0,71821	0,64674	0,7899	**0,82618**
10	0,74808	0,80658	0,79138	0,80476	0,82334	0,79184	0,74711	0,68531	0,77549	**0,8307**
11	0,75	0,81631	0,8207	0,82971	0,84037	0,80146	0,7829	0,72376	0,78195	**0,84382**
12	0,66874	0,76606	**0,8231**	0,82038	0,8133	0,74609	0,81708	0,78333	0,71614	0,81614
13	0,77248	0,82585	0,81535	0,82797	0,84138	0,81426	0,76955	0,70846	0,79919	**0,84439**
14	0,65998	0,75141	**0,80444**	0,80225	0,79782	0,73187	0,79926	0,77595	0,70374	0,80235
15	0,61066	0,72128	0,79397	0,78803	0,78131	0,69445	0,80262	**0,81608**	0,66077	0,7907

16	0,81657	**0,82312**	0,70582	0,73989	0,79087	0,81889	0,61678	0,53459	0,82191	0,80334
17	0,79822	**0,82199**	0,73417	0,76117	0,8039	0,81296	0,65815	0,57558	0,81039	0,81513
19	0,79547	0,83789	0,78893	0,80962	0,83669	0,8275	0,72311	0,64296	0,81701	**0,84084**
20	0,77523	0,80795	0,74631	0,76827	0,8013	0,79675	0,68489	0,61425	0,79085	**0,81179**

Os resultados da Tabela 5 mostram que a combinação das segmentações utilizando o voto majoritário obtiveram um melhor desempenho que qualquer segmentação em 9 das 19 imagens. Contudo, é preciso observar que, a cada imagem utilizada, a melhor segmentação varia bastante, não sendo possível adivinhar de antemão qual a melhor opção para aquela imagem. O melhor eixo possível segmentando as imagens foi o melhor em apenas 5 das 19 imagens, mostrando que há, de maneira geral, um desempenho superior utilizando a combinação. Uma análise mais detalhada será feita na Seção 4.5.

5.4.4. Experimentos Utilizando GMM

No quarto experimento, a segmentação foi feita utilizando o GMM. Assim como nas segmentações anteriores, ela foi feita individualmente em cada fatia 2D selecionada, novamente em cada um dos nove eixos descritos no Capítulo 3. O GMM atua de maneira mais similar aos dois primeiros experimentos, no que diz respeito a calcular os valores de limiares para cada fatia 2D individualmente, diferentemente do SOM. Assim, a Figura 48 representa o resultado da combinação das segmentações com as imagens 2D sendo selecionadas em cada eixo escolhido utilizando o algoritmo GMM. Por análise empírica, foi visto que o algoritmo GMM não obteve bom desempenho na base de dados utilizada nos três experimentos anteriores. Contudo, na outra base ele se mostrou muito promissor, ao contrário do SOM, que não conseguiu se sair bem nessa segunda base. A base de dados utilizada nesse experimento foi a NKI-TRT-22.

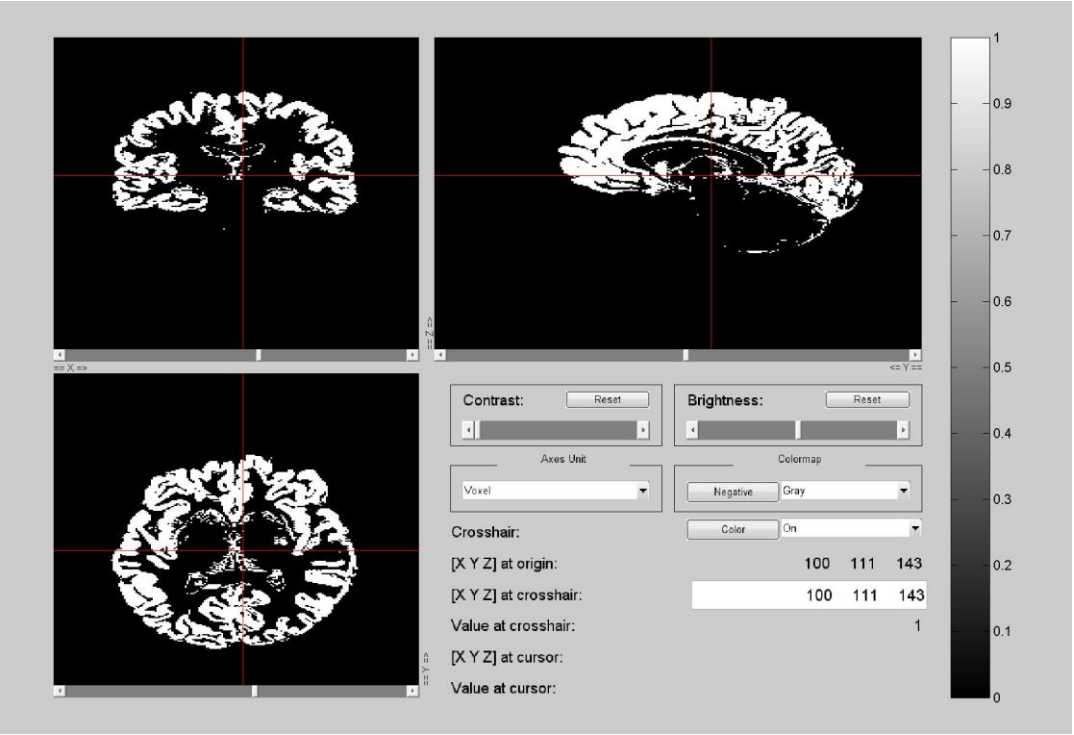

Figura 49 – Combinação das segmentações utilizando o GMM, de imagem da base NKI-TRT-22

Como pode ser visto na imagem, a segmentação ainda possui erros, mas parece ser um pouco mais refinada em relação às segmentações anteriores. As bordas parecem manter o traço mais fino, como apresentado no SOM e é possível ver mais detalhes na parte inferior da primeira imagem, além de haver menor ruído. A Tabela 6 mostra o resultado do DSC de cada segmentação com o *groundtruth*, comparando os valores individuais sobre cada eixo com o resultado da combinação final entre todas elas, utilizando apenas o GMM.

Tabela 6 – Valores do DSC das segmentações usando o GMM.

Eixo	1	2	3	4	5	6	7	8	9	V.M.
Imagem										
1	0,7538	0,79584	0,81265	0,82102	0,77444	0,80424	0,81217	0,83981	0,80188	**0,84307**
2	0,82583	0,83879	0,81167	0,84664	0,80326	0,82857	0,8199	0,8321	0,84112	**0,85131**
3	0,7046	0,77916	0,66849	0,79315	0,68668	0,72412	0,72374	0,68697	0,76099	**0,8015**

4	0,7506	0,79631	0,73013	0,78877	0,74115	0,74466	0,77401	0,76671	0,78597	**0,80598**
5	0,82177	0,81177	0,78603	0,81135	0,80282	0,80712	0,81715	0,81647	0,82877	**0,84019**
6	0,7434	0,80343	0,72753	0,80824	0,69691	0,76366	0,76321	0,74042	0,80871	**0,80797**
7	0,73635	0,73847	0,72333	0,75345	0,70359	0,7174	0,69373	**0,79009**	0,78911	0,78069
8	0,78764	0,8255	0,78947	**0,83047**	0,78602	0,80666	0,79822	0,80427	0,81718	0,82896
9	0,7677	0,79932	0,76423	0,80864	0,71707	0,76591	0,77695	0,78429	0,7881	**0,82238**
10	0,82598	0,81842	0,78972	0,80787	0,79813	0,81012	0,80074	0,81328	0,8207	**0,84915**
11	0,79337	0,81282	0,7805	**0,83649**	0,77035	0,79471	0,80134	0,79384	0,81153	0,82894
12	0,82495	0,80749	0,80522	0,81354	0,78335	0,79863	0,80183	0,83701	0,83405	**0,84345**
13	0,78343	0,81393	0,75109	0,82042	0,76253	0,7988	0,77145	0,79651	0,80704	**0,82444**
14	0,75421	0,79638	0,73098	0,79747	0,72836	0,76042	0,77185	0,7741	0,78963	**0,80138**
15	0,80804	0,82239	0,794	0,82788	0,79164	0,77704	0,7847	0,82538	0,82452	**0,83784**
16	0,78898	0,81645	0,79682	0,82565	0,76902	0,79847	0,8033	0,80824	0,81922	**0,82934**
17	0,73665	0,79179	0,75973	0,7941	0,73082	0,75587	0,72584	0,77101	0,78829	**0,79772**
19	0,81751	0,83822	0,81466	0,84184	0,81982	0,83444	0,8382	0,81802	0,83085	**0,85013**
20	0,7362	0,78145	0,71718	0,79701	0,72171	0,73634	0,73843	0,75553	0,7874	**0,80082**
21	0,70929	0,71767	0,75456	**0,76104**	0,67788	0,7058	0,70411	0,70882	0,72364	0,74917
22	0,80474	**0,83408**	0,76222	0,83194	0,80113	0,79291	0,81448	0,80083	0,81687	0,82724

Os resultados da Tabela 6 mostram que as combinações das segmentações utilizando o voto majoritário obtiveram um melhor desempenho que qualquer segmentação em 16 das 21 imagens, o que mostra uma dominância grande sobre qualquer eixo de analise. Além disso, o resultado se mostrou tão bom quanto o SOM, quando não ligeiramente superior. E seguindo todos os outros experimentos, analisando apenas os eixos, é impossível saber qual é o melhor de antemão, já que a cada imagem, algum eixo diferente apresenta um resultado superior. Uma análise mais detalhada será feita na Seção 4.5.

5.5. Análise dos Resultados

Realizados os experimentos, é preciso analisar detalhadamente os resultados das segmentações, para poder verificar se o método proposto realmente pode trazer alguma contribuição. A análise deve ser feita com base em dois pontos principais: o cálculo de acerto/erro usando o método do DSC e a comparação, utilizando métodos estatísticos para averiguar os resultados, com os modos como as segmentações são mais comumente realizadas atualmente. Os testes estatísticos foram realizados utilizando o T-teste, um método utilizado para calcular as diferenças entre duas médias para grupos relacionados, como foi descrito no Capítulo 3. Quanto às segmentações, as comparações serão realizadas de duas maneiras: analisando todos os resultados, independentemente do eixo base utilizado para a segmentação do cérebro, e uma comparação da combinação com as segmentações nos eixos mais comuns atualmente, que são os eixos cartesianos tradicionais (x, y e z). A análise também será dividida, para mostrar individualmente como cada técnica e a base de dados influenciam nos resultados das segmentações.

5.5.1. Análise da Binarização de Otsu

Como pode ser visto, a combinação teve um desempenho melhor em todas as 19 imagens da base de dados. Contudo, é preciso verificar quão melhor é esse resultado. Como citado anteriormente, não é possível saber qual eixo de análise é o melhor para aquela imagem a ser segmentada. Desse modo, a comparação principal deve ser feita com os eixos cartesianos, pois são nesses eixos que são feitas as segmentações de imagens 3D fazendo as análises de fatias 2D. Desse modo, tem-se a Figura 49, que mostra os resultados das segmentações usando as fatias 2D nos três eixos e o da combinação, a modo de comparação.

A Figura 15 mostra apenas o eixo Y de cada resultado final, para que se possa visualizar mais facilmente as diferenças.

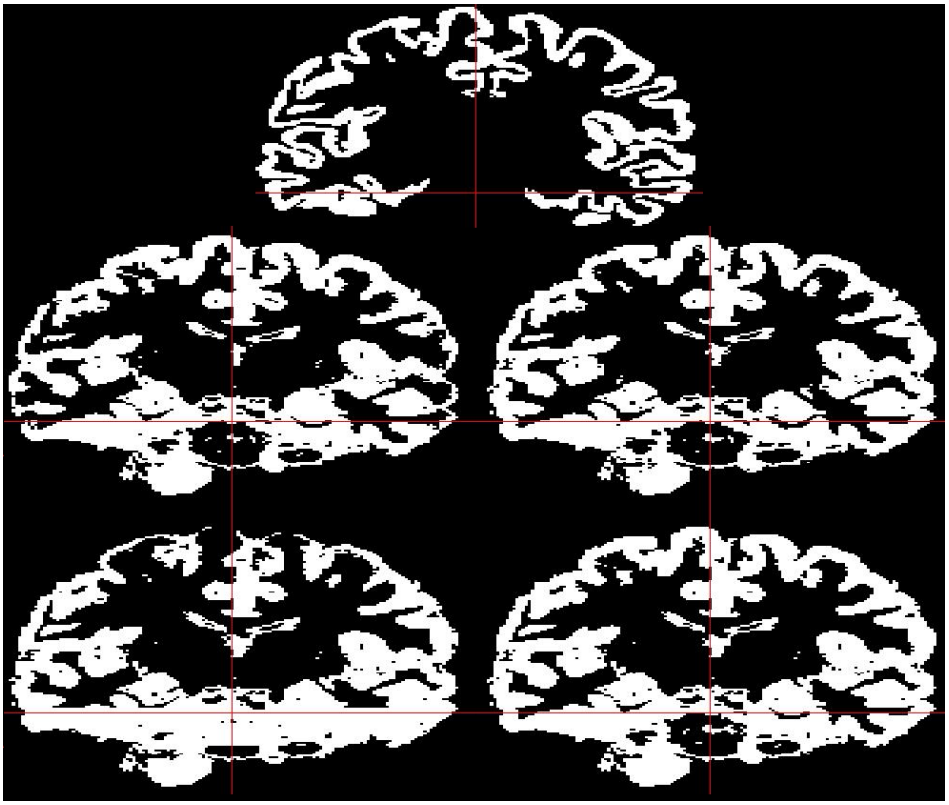

Figura 50 – Matérias cinzas de uma imagem 3D visualizadas no eixo Y de imagem da base OASIS-TRT-20, onde (a) *groundtruth*, (b) segmentação realizada no eixo X, (c) segmentação realizada no eixo Y, (d) segmentação realizada no eixo Z, (e) combinação de todos os resultados

É possível perceber que há diferenças entre as segmentações, mas não é possível dizer qual é a imagem mais parecida com o *groundtruth*. Para isso também é feita uma análise do *Dice-Coefficient*, para que se possa mensurar essa diferença. A Tabela 7 mostra os resultados dos valores do cálculo do *Dice-Coefficient* apenas para os eixos x, y e z e o resultado da combinação.

Tabela 7 – Valores do DSC das segmentações usando a binarização de Otsu.

Eixo	X	Y	Z	V.M.
Imagem				
1	0,70677	0,71982	0,7105	**0,7342**
2	0,72444	0,74427	0,74031	**0,7548**
3	0,732	0,75581	0,73789	**0,7638**
4	0,70441	0,72815	0,71705	**0,7488**
5	0,71814	0,73556	0,73357	**0,7504**
6	0,70621	0,73334	0,73456	**0,7476**
7	0,72127	0,74024	0,73359	**0,754**
8	0,74307	0,76123	0,75145	**0,7683**
9	0,70739	0,73653	0,73338	**0,7523**
10	0,70085	0,72649	0,72748	**0,7411**
11	0,73893	0,75537	0,75207	**0,7633**
12	0,70598	0,73034	0,7232	**0,744**
13	0,74277	0,75776	0,75668	**0,7699**
14	0,68819	0,71491	0,70792	**0,7274**
15	0,70678	0,73961	0,73539	**0,7557**
16	0,70688	0,7259	0,72945	**0,7411**

17	0,71081	0,72012	0,7229	**0,7348**
19	0,73203	0,74837	0,74236	**0,7611**
20	0,68312	0,70262	0,71196	**0,7207**

Como já foi dito, o resultado da combinação foi superior na segmentação para as 19 imagens da base de dados. Assim, é preciso mensurar essa superioridade. Utilizando o T-teste, foi avaliado se a combinação é estatisticamente superior, a um nível de confiança de 95%, ou seja, *alpha* = 5%. A Tabela 8 mostra as médias das segmentações em todos os eixos e na combinação.

Tabela 8 – Médias dos valores do DSC das segmentações de todas as imagens, usando a binarização de Otsu.

	Eixo X	**Eixo Y**	**Eixo Z**	**V. M.**
Média	0.7147	0.7356	0.7317	**0.7491**

Baseado nesses valores, foram calculados os valores críticos das segmentações com o resultado da combinação, que estão mostradas na Tabela 9.

Tabela 9 – Valores críticos e de testes para verificar se as amostras são estatisticamente diferentes

	Eixo X/V. M.	**Eixo Y/V. M**	**Eixo Z/V. M**
Valor Crítico	-1.7340	-1.7340	-1.7340
Variável de Teste	20.0678	17.3466	13.2676
	Estatisticamente diferentes	Estatisticamente diferentes	Estatisticamente diferentes

Assim, é possível verificar que o resultado da combinação foi superior estatisticamente a qualquer segmentação nos eixos mais utilizados, mostrando, para esse caso, a eficácia do método proposto. É possível ver que, dependendo do eixo escolhido para se fazer os cortes das imagens 2D, pode-se chegar até a mais de 3% de diferença média.

5.5.2. Análise da Binarização de Otsu + *K-Means*

Como também pode ser visto, a combinação das segmentações por Otsu combinadas com as segmentações por *K-Means* também tiveram um desempenho melhor em todas as 19 imagens da base de dados, sendo ligeiramente superior em 18 das 19 imagens se comparado apenas com a combinação utilizado apenas as segmentações por Otsu. Mais uma vez, não é possível saber qual eixo de análise é o melhor para aquela imagem a ser segmentada. Desse modo, tem-se a Figura 16, que mostra os resultados das segmentações usando as fatias 2D nos três eixos e o da combinação pelas duas técnicas, a modo de comparação. A Figura 50 mostra apenas o eixo Y de cada resultado final, para que se possa visualizar mais facilmente as diferenças.

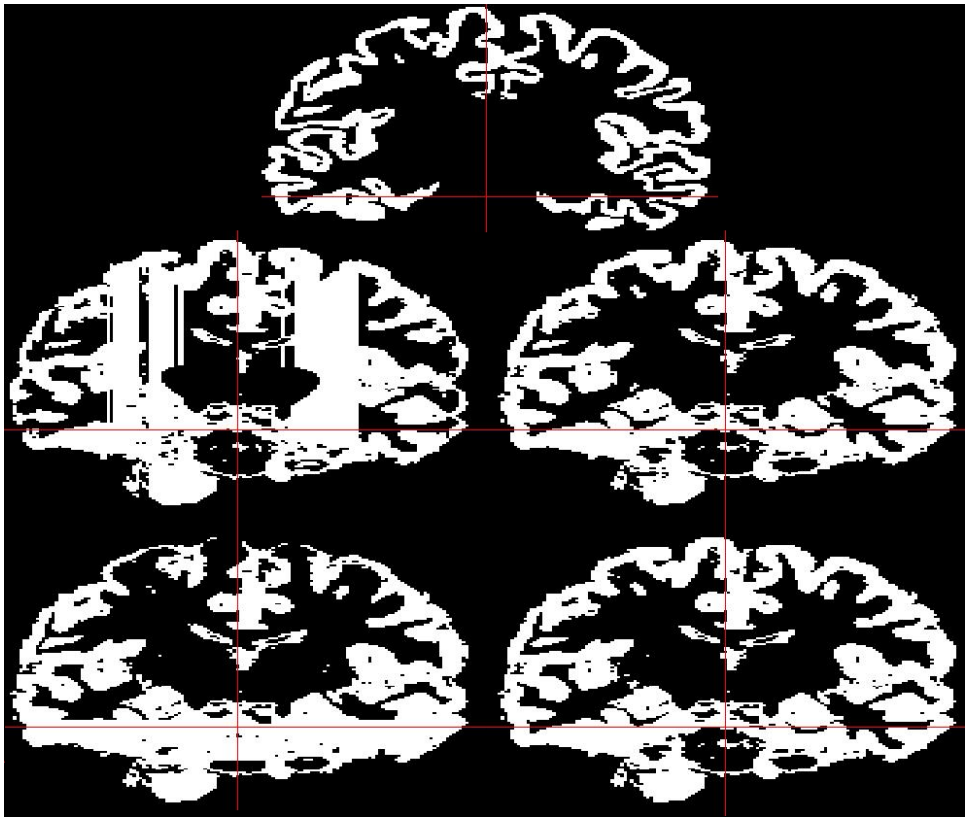

Figura 51 – Matérias cinza de uma imagem 3D visualizadas no eixo Y de imagem da base OASIS-TRT-20, onde (a) *groundtruth*, (b) segmentação realizada no eixo X, (c) segmentação realizada no eixo Y, (d) segmentação realizada no eixo Z, (e) combinação de todos os resultados usando *K-Means* com os resultados usando Otsu

É possível perceber que há diferenças entre as segmentações, especialmente na primeira segmentação, que possui linhas verticais totalmente discrepantes, mas ainda não é

possível dizer qual é a imagem mais parecida com o *groundtruth*. Para isso também é feita uma análise do *Dice-Coefficient*, para que se possa mensurar essa diferença. A Tabela 10 mostra os resultados dos valores do cálculo do *Dice-Coefficient* apenas para os eixos x, y e z, tanto da binarização de Otsu, quanto do resultado da segmentação utilizando *K-Means*, e o resultado da combinação.

Tabela 10 – Valores do DSC das segmentações usando a binarização de Otsu e *K-means*.

	Otsu			***K-Means***			
Eixo	**X**	**Y**	**Z**	**X**	**Y**	**Z**	**V.M.**
Imagem							
1	0,70677	0,71982	0,7105	0,61639	0,69974	0,71224	**0,73658**
2	0,72444	0,74427	0,74031	0,631	0,71642	0,72465	**0,75684**
3	0,732	0,75581	0,73789	0,69986	0,75584	0,73931	**0,76636**
4	0,70441	0,72815	0,71705	0,57081	0,68341	0,66319	**0,74786**
5	0,71814	0,73556	0,73357	0,63481	0,72755	0,69609	**0,75203**
6	0,70621	0,73334	0,73456	0,67464	0,71953	0,73465	**0,74929**
7	0,72127	0,74024	0,73359	0,61575	0,72112	0,682	**0,75522**
8	0,74307	0,76123	0,75145	0,71773	0,75929	0,74884	**0,77281**
9	0,70739	0,73653	0,73338	0,63297	0,7134	0,73149	**0,75553**
10	0,70085	0,72649	0,72748	0,65841	0,72451	0,72993	**0,74261**
11	0,73893	0,75537	0,75207	0,71374	0,74633	0,75273	**0,76745**

12	0,70598	0,73034	0,7232	0,64349	0,68836	0,72476	**0,74718**
13	0,74277	0,75776	0,75668	0,69245	0,71929	0,75751	**0,77481**
14	0,68819	0,71491	0,70792	0,61233	0,68797	0,71054	**0,72924**
15	0,70678	0,73961	0,73539	0,67348	0,74023	0,73671	**0,75766**
16	0,70688	0,7259	0,72945	0,65077	0,71103	0,71241	**0,74386**
17	0,71081	0,72012	0,7229	0,57971	0,69467	0,63182	**0,73797**
19	0,73203	0,74837	0,74236	0,64826	0,72778	0,71706	**0,76472**
20	0,68312	0,70262	0,71196	0,6422	0,68508	0,68722	**0,72661**

O resultado da combinação também foi superior na segmentação para as 19 imagens da base de dados. Utilizando o T-teste, foi avaliado se a combinação é estatisticamente superior, a um nível de confiança de 95%, ou seja, *alpha* = 5%. A Tabela 11 mostra as médias das segmentações em todos os eixos e na combinação.

Tabela 11 - Médias dos valores do DSC das segmentações de todas as imagens, usando a binarização de Otsu e *K-Means*.

	Otsu			***K-Means***			**.**
	Eixo X	**Eixo Y**	**Eixo Z**	**Eixo X**	**Eixo Y**	**Eixo Z**	**V. M**
Média	0.7147	0.7356	0.7317	0.6478	0.7169	0.7154	**0.7518**

Baseado nesses valores, foram calculados os valores críticos das segmentações com o resultado da combinação, que estão mostradas na Tabela 11. Como os valores foram calculados no tópico anterior e o valor da média subiu, é desnecessário calcular novamente para as segmentações utilizando a binarização de Otsu. Desse modo, a Tabela 12 mostra apenas os valores das segmentações por *K-Means*.

Tabela 12 – Valores críticos e de testes para verificar se as amostras são estatisticamente diferentes.

	Eixo X/V. M.	Eixo Y/V. M	Eixo Z/V. M
Valor Crítico	-1.7340	-1.7340	-1.7340
Variável de Teste	13.5766	10.0978	6.0762
	Estatisticamente diferentes	Estatisticamente diferentes	Estatisticamente diferentes

Assim, é possível verificar que o resultado da combinação foi superior estatisticamente a qualquer segmentação nos eixos mais utilizados, tanto utilizando a segmentação por Otsu quanto por *K-Means*. Além disso, foi possível perceber que, ao se aumentar o número de combinadores, mesmo que alguns possuam resultados inferiores (*K-Means*), é possível ainda obter pequenas melhorias de desempenho. Nesse caso, em comparação a apenas utilizar *K-Means* e escolhendo o eixo X como base, pode-se chegar até a mais de 9% de diferença média.

5.5.3. Análise do SOM

Já no caso da segmentação utilizando a rede neural SOM, não houve unanimidade no desempenho das combinações, que venceram todos os eixos em 9 das 19 imagens. Por esse motivo, é ainda mais difícil saber qual eixo de análise é o melhor a aquela imagem a ser segmentada. A Figura 51 mostra apenas o eixo Y de cada resultado final, para que se possa visualizar mais facilmente as diferenças.

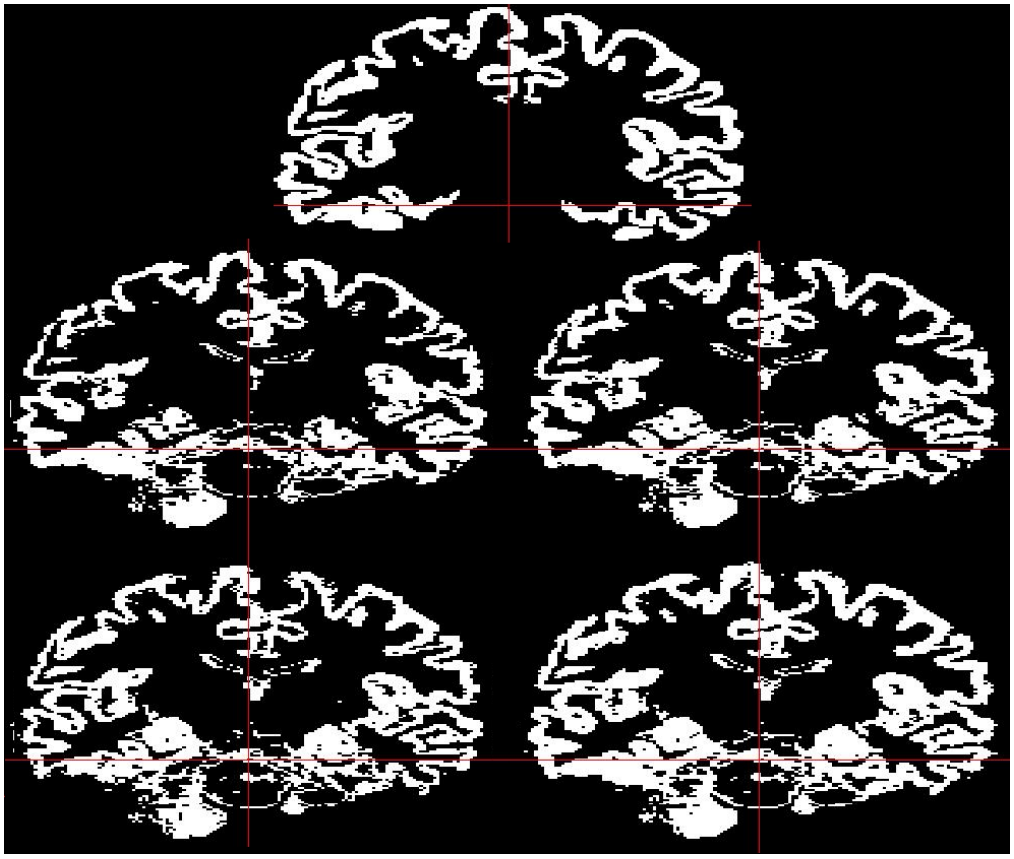

Figura 52 – Matérias cinza de uma imagem 3D visualizadas no eixo Y de imagem da base OASIS-TRT-20, onde (a) *groundtruth*, (b) segmentação realizada no eixo X, (c) segmentação realizada no eixo Y, (d) segmentação realizada no eixo Z, (e) combinação de todos os resultados usando SOM.

Como esperado, é possível perceber que há diferenças entre as segmentações, mas ainda não é possível dizer qual é a imagem mais parecida com o *groundtruth*. Para isso também é feita uma análise do *Dice-Coefficient*, para que se possa mensurar essa diferença. A Tabela 13 mostra os resultados dos valores do cálculo do *Dice-Coefficient* apenas para os eixos x, y e z, utilizando o SOM, e o resultado da combinação.

Tabela 13 – Valores do DSC das segmentações usando o SOM.

Eixo / Imagem	1	2	3	V.M.
1	0,73618	0,79718	0,80456	**0,81737**
2	**0,83333**	0,82809	0,68501	0,79763
3	0,80839	**0,83852**	0,76934	0,83064
4	0,79657	**0,83355**	0,77178	0,82268
5	0,81251	**0,81822**	0,70107	0,80238
6	0,80612	0,83898	0,76928	**0,84352**
7	0,78042	0,82453	0,78302	**0,83777**
8	0,75525	0,81463	0,81346	**0,83768**
9	0,76913	0,81349	0,77314	**0,82618**
10	0,74808	0,80658	0,79138	**0,8307**
11	0,75	0,81631	0,8207	**0,84382**
12	0,66874	0,76606	**0,8231**	0,81614
13	0,77248	0,82585	0,81535	**0,84439**
14	0,65998	0,75141	**0,80444**	0,80235
15	0,61066	0,72128	**0,79397**	0,7907
16	0,81657	**0,82312**	0,70582	0,80334

17	0,79822	**0,82199**	0,73417	0,81513
19	0,79547	0,83789	0,78893	**0,84084**
20	0,77523	0,80795	0,74631	**0,81179**

O resultado da combinação foi superior na segmentação para 9 das 19 imagens da base de dados, levando em consideração todos os 9 eixos utilizados na combinação. Contudo, se comparado apenas com os 3 principais eixos, a combinação foi superior em 10 das 19 imagens. Utilizando o T-teste, foi avaliado se a combinação é estatisticamente superior, especialmente agora que a combinação foi superior aos eixos principais em apenas 53% das imagens. A um nível de confiança de 95%, ou seja, *alpha* = 5%. A Tabela 14 mostra as médias das segmentações em todos os eixos e na combinação.

Tabela 14 – Médias dos valores do DSC das segmentações de todas as imagens, usando o SOM.

	Eixo X	Eixo Y	Eixo Z	V. M.
Média	0.7628	0.8098	0.7734	**0.8218**

Baseado nesses valores, foram calculados os valores críticos das segmentações com o resultado da combinação, que estão mostradas na Tabela 15.

Tabela 15 – Valores críticos e de testes para verificar se as amostras são estatisticamente diferentes.

	Eixo X/V. M.	Eixo Y/V. M	Eixo Z/V. M
Valor Crítico	-1.7340	-1.7340	-1.7340
Variável de Teste	4.5841	2.0426	5.9273
	Estatisticamente diferentes	Estatisticamente diferentes	Estatisticamente diferentes

Mais uma vez é possível verificar que o resultado da combinação foi superior estatisticamente a qualquer segmentação nos eixos mais utilizados, mostrando, para mais esse

caso, a eficácia do método proposto. É possível ver que, dependendo do eixo escolhido para se fazer os cortes das imagens 2D, pode-se chegar até a quase 6% de diferença média.

5.5.4. Análise do GMM

Já no caso da segmentação utilizando o GMM, assim como no SOM, também não houve unanimidade no desempenho das combinações, que venceram todos os eixos em 16 das 21 imagens. Mais uma vez, é difícil saber qual eixo de análise é o melhor a aquela imagem a ser segmentada. A Figura 52 mostra apenas o eixo Y de cada resultado final, para que se possa visualizar mais facilmente as diferenças.

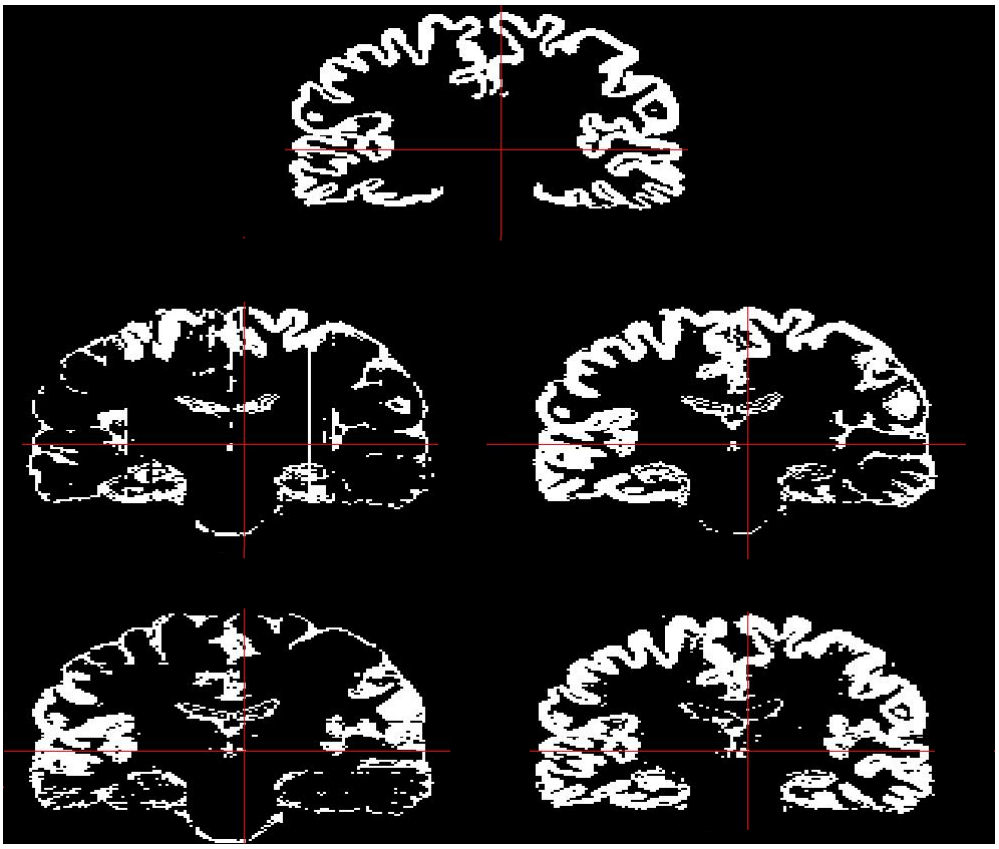

Figura 53 – Matérias cinza de uma imagem 3D visualizadas no eixo Y de imagem da base NKI-RS-22, onde (a) *groundtruth*, (b) segmentação realizada no eixo X, (c) segmentação realizada no eixo Y, (d) segmentação realizada no eixo Z, (e) combinação de todos os resultados usando GMM

Como esperado, é possível perceber que há diferenças entre as segmentações, mas ainda não é possível dizer qual é a imagem mais parecida com o *groundtruth*, apesar de parecer já que a combinação apresenta um melhor resultado. Ainda assim, também é feita

uma análise do *Dice-Coefficient*, para que se possa mensurar essa diferença. A Tabela 16 mostra os resultados dos valores do cálculo do *Dice-Coefficient* apenas para os eixos x, y e z, utilizando o GMM, e o resultado da combinação.

Tabela 16 – Valores do DSC das segmentações usando o GMM.

Eixo / Imagem	1	2	3	V.M.
1	0,7538	0,79584	0,81265	**0,84307**
2	0,82583	0,83879	0,81167	**0,85131**
3	0,7046	0,77916	0,66849	**0,8015**
4	0,7506	0,79631	0,73013	**0,80598**
5	0,82177	0,81177	0,78603	**0,84019**
6	0,7434	0,80343	0,72753	**0,80797**
7	0,73635	0,73847	0,72333	**0,78069**
8	0,78764	0,8255	0,78947	**0,82896**
9	0,7677	0,79932	0,76423	**0,82238**
10	0,82598	0,81842	0,78972	**0,84915**
11	0,79337	0,81282	0,7805	**0,82894**
12	0,82495	0,80749	0,80522	**0,84345**
13	0,78343	0,81393	0,75109	**0,82444**

14	0,75421	0,79638	0,73098	**0,80138**
15	0,80804	0,82239	0,794	**0,83784**
16	0,78898	0,81645	0,79682	**0,82934**
17	0,73665	0,79179	0,75973	**0,79772**
19	0,81751	0,83822	0,81466	**0,85013**
20	0,7362	0,78145	0,71718	**0,80082**
21	0,70929	0,71767	**0,75456**	0,74917
22	0,80474	**0,83408**	0,76222	0,82724

O resultado da combinação foi superior na segmentação para 16 das 21 imagens da base de dados, levando em consideração todos os 9 eixos utilizados na combinação. Contudo, se comparado apenas com os 3 principais eixos, a combinação foi superior em 19 das 21 imagens. Utilizando o T-teste, foi avaliado se a combinação é estatisticamente superior, a um nível de confiança de 95%, ou seja, *alpha* = 5%. A Tabela 17 mostra as médias das segmentações em todos os eixos e na combinação.

Tabela 17 – Médias dos valores do DSC das segmentações de todas as imagens, usando o GMM.

	Eixo X	**Eixo Y**	**Eixo Z**	**V. M.**
Média	0.7750	0.8019	0.7652	**0.8201**

Baseado nesses valores, foram calculados os valores críticos das segmentações com o resultado da combinação, que estão mostradas na Tabela 18.

Tabela 18 – Valores críticos e de testes para verificar se as amostras são estatisticamente diferentes.

	Eixo X/V. M.	Eixo Y/V. M	Eixo Z/V. M
Valor Crítico	-1.7340	-1.7340	-1.7340
Variável de Teste	9.6267	6.0354	9.1599
	Estatisticamente diferentes	Estatisticamente diferentes	Estatisticamente diferentes

Mais uma vez é possível verificar que o resultado da combinação foi superior estatisticamente a qualquer segmentação nos eixos mais utilizados, mostrando, para mais esse caso, a eficácia do método proposto. É possível ver que, dependendo do eixo escolhido para se fazer os cortes das imagens 2D, pode-se chegar até a mais de 5% de diferença média.

5.5.5. Análise Geral

A Tabela 19 mostra então os resultados de todas as combinações com as diferentes técnicas utilizadas, para facilitar as comparações. Como pode ser visto, em todos os casos houve uma melhora do desempenho médio. Como demonstrado anteriormente, essas diferenças foram relevantes estatisticamente, mostrando, portando, a eficiência do método proposto.

Tabela 19 – Médias dos valores do DSC das segmentações de todas as imagens, usando todas as técnicas de segmentação do trabalho.

Técnica		Eixo X	Eixo Y	Eixo Z	V. M.
Otsu	Média	0.7147	0.7356	0.7317	**0.7491**
K-Means	Média	0.7628	0.8098	0.7734	**0.8218**
SOM	Média	0.7628	0.8098	0.7734	**0.8218**
GMM	Média	0.7750	0.8019	0.7652	**0.8201**

5.6. Teste de Mann–Whitney–Wilcoxon

Além do T-teste, foi feita uma análise dos dados obtidos utilizando o teste de Mann–Whitney–Wilcoxon [Mann and Whitney, 1947]. O teste de Mann–Whitney–Wilcoxon é um teste não paramétrico, utilizado para a comparação dos valores médios de amostras emparelhadas, e é uma extensão do teste dos sinais. Ele é um teste que pode ser utilizado alternativamente ao T-teste, quando não é possível assumir que as amostras da população são normalmente distribuídas. Quando as amostras são normalmente distribuídas, possuem cerca de 95% de eficiência, se comparado com o T-teste. Já para distribuições que diferem da normalidade e com amostragem suficientemente grandes, o teste de Mann–Whitney–Wilcoxon é considerado mais eficiente que o T-teste [Conover, 1980].

Uma formulação bastante geral é assumir que:

1- Todas as observações dos dois grupos são independentes entre si;
2- As respostas são ordinais (ou seja, uma pessoa pode dizer de quaisquer duas observações qual delas é maior);
3- Sob a hipótese nula H_0, a probabilidade de uma observação da população 1 exceder uma observação da população 2 é igual a probabilidade de uma observação da população 2 exceder uma observação da população 1. Uma hipótese nula normalmente utilizada é que as distribuições das duas populações são iguais;
4- A hipótese alternativa H_1 é a probabilidade de uma observação da população 1 exceder uma observação da população 2 é diferente da probabilidade de uma observação da população 2 exceder uma observação da população 1.

Com o cálculo do *ρ value*, caso ele seja inferior a 1 menos o nível de confiança (no caso, 1-95% = 0.05), então a hipótese nula H_0 será rejeitada, mostrando que as amostras não são estatisticamente semelhantes. Abaixo seguem as tabelas dos *ρ values* calculados com as amostras obtidas nas segmentações em relação aos resultados obtidos utilizando o método proposto no trabalho:

Tabela 20 – Valores do *ρ value* para todas as segmentações. Valores abaixo de 0.05 indicam que as distribuições das duas populações não são iguais.

Técnica		Eixo X	Eixo Y	Eixo Z
Otsu	*ρ value*	1.3183e-04	0.0131	9.6965e-04
K-Means	*ρ value*	1.4798e-07	3.3885e-05	7.1726e-05
SOM	*ρ value*	4.9480e-05	0.3209	8.1044e-05
GMM	*ρ value*	1.9683e-04	0.0236	1.2027e-05

Como pode ser observado, tirando apenas o eixo Y utilizando o SOM, todos os outros *ρ values* foram inferiores a 0.05, o que levou à negação da hipótese nula H_0, mostrando que realmente as amostras são diferentes e comprovando a eficiência do método proposto. Já no caso do eixo Y utilizando o SOM, o método não confirma que as distribuições das duas populações são iguais, apenas não pode confirmar que são diferentes. Contudo, como não experimento anterior, utilizando o T-teste, foi constatada a diferença, então é muito provável que realmente sejam diferentes as distribuições.

5.7. Custo Computacional

A Tabela 20 mostra custo computacional para a execução do método proposto, utilizando cada técnica de segmentação isoladamente. Os códigos foram implementados na linguagem MatLab. Os testes foram executados em um computador com um processador Intel Core i7 3610QM, com 16GB de memória RAM e sistema operacional Windows 7, utilizando a ferramenta MatLab R2013a.

Tabela 21 – Custo computacional de cada técnica de segmentação implementada no MatLab.

Técnica	Quantidade de Imagens Segmentadas	C. C.	Custo Médio
Otsu	19	**00:59:27**	**00:03:08**
K-Means	19	**04:29:22**	**00:14:10**
SOM	19	**01:05:38**	**00:03:27**
GMM	21	**06:28:26**	**00:20:26**

6. CONCLUSÕES E TRABALHOS FUTUROS

Segmentar imagens, especialmente imagens médicas, ainda é um grande desafio em temos de métodos computacionais automáticos para os pesquisadores atualmente. E quanto maior e mais diversas essas imagens puderem ser, mais abrangente a técnica de segmentação deve ser para que se obtenha uma boa segmentação para essas imagens. Em compensação, quanto mais abrangente for a técnica, mais suscetível a erros ela será.

Uma método que apresenta bons resultados em termos de taxas de acerto e que tem sido bastante utilizada para problemas de classificação, em geral, é a combinação de classificadores, pois esta tem apresentado resultados promissores em diversas áreas diferentes. Assim, a ideia do trabalho se baseou em encontrar uma maneira nova de gerar diferentes resultados de segmentação de imagens 3D, para posteriormente fazer essa combinação de classificadores.

Para mostrar que a ideia citada é válida, o modelo proposto é um método inspirando no *Bagging*, técnica tradicional para diversificação de classificadores. Assim, ao se selecionar uma imagem médica 3D e fatiá-la em nove planos diferentes diferentes para gerar novas imagens 2D a ser segmentadas, foi possível utilizar as características diferentes de cada conjunto de imagens 2D para gerar resultados diferentes, que, por fim, acabaram sendo combinados em um único resultado final.

Com os resultados obtidos, através da escolha de nove eixos diferentes para se fatiar as imagens 3D e a escolha de quatro métodos de segmentação 2D, é possível avaliar estatisticamente que realmente há uma melhoria considerável da taxa de acerto da segmentação, se comparado com a utilização de apenas um dos eixos tradicionais, especialmente pelo fato de que não é possível saber qual o melhor dos planos entre os tradicionais, o que pode acarretar em uma queda grande da taxa de acerto caso se escolha o plano errado.

Contudo, alguns aspectos ainda não foram abordados. Um deles é o modo no qual o método proposto se comportaria com técnicas de segmentação mais robustas, uma vez que, se

o método obteve melhorias estatisticamente relevantes nas taxas de acerto com as técnicas utilizadas, é de se esperar o mesmo de outras técnicas. Outro aspecto é se a utilização de ainda mais planos diferentes para fatiar a imagem 3D pode melhorar ainda mais o resultado, pois, assim como o incremento de planos levou a melhorias nas taxas de acerto, é possível que exista um número ótimo de planos. Por fim, é importante também verificar como é possível fazer melhores combinações, de modo a se tentar aproveitar melhor todas as imagens, uma vez que o voto majoritário não é o único método de combinação existente.

6.1. Trabalhos futuros

O método proposto ainda pode ser bastante explorado, desde seu início, com uma maior quantidade de amostras, até seu término, com a combinação das segmentações geradas, como pode ser observado pelos pontos abordados anteriormente. Alguns trabalhos futuros na área são:

- Utilizar técnicas de segmentação mais robustas, como segmentações baseadas em atlas ou em modelos deformáveis, específicas para o cérebro, que levem em consideração mais características ou que utilizem modelos estatísticos mais bem modelados;
- Incrementar a quantidade de eixos utilizados para fatiar a imagem 3D, de modo a achar um número ótimo de imagens a ser combinadas para se obter o melhor resultado possível, pois, a partir de certo ponto, a melhoria pode não compensar o aumento do custo computacional;
- Utilizar outras técnicas para a combinação dos resultados, pois, apesar de apresentar bons resultados, o voto majoritário ainda é um método combinacional simples, de modo que modelos mais elaborados possam vir a apresentar melhores resultados. Foram feitas segmentações apenas com respostas não *fuzzy*. Desse modo, caso esse ponto seja alterado, seria possível utilizar métodos combinacionais conhecidos, como Produto, Máximo, Mínimo, entre outros [L. I. Kuncheva, 2004];
- Como foi visto que os planos de análise importam para a segmentação, é possível que a segmentação feita pelo especialista não seja a melhor possível, pois ela foi feita se baseando apenas em imagens fatiadas em apenas um plano. Desse modo, é de interesse construir uma base de imagens de *groundtruth*, com o auxílio de um especialista, em outros planos que não o utilizado, levando a novos resultados para o

groundtruth, talvez mais precisos, e que alterassem o resultado final da segmentação, uma vez que a combinação dos *groudtruths* gerados em planos diferentes poderia gerar um novo *groundtruth* diferente de todos os outros.

REFERÊNCIAS

G. N. Abras and V. L. Ballarín, “A weighted k-means algorithm applied to brain tissue classification,” *Journal of Computer Science & Technology*, vol. 5, 2005.

S. Angenent, E. Pichon, and A. Tannenbaum, “Mathematical methods in medical image processing,” *Bulletin of the American Mathematical Society* 43.3, pp. 365-396, 2006.

A. Assmus, "Early history of X rays," *Beam Line* 25.2: 10-24, 1995.

S. K. Bandhyopadhyay and T. U. Paul, “Automatic segmentation of brain tumour from multiple images of brain mri,” *International Journal of Application or Innovation in Engineering & Management (IJAIEM)*, vol. 2, no. 1, 2013.

M. F. Barnsley, *Fractals everywhere*. Academic press, 2014.

E. P. Bertin, *Principles and practice of X-ray spectrometric analysis*. Springer Science & Business Media, 2012.

J. C. Bezdek, L. Hall, and L. Clarke, “Review of MR Image segmentation techniques using pattern recognition.,” *Medical physics*, vol. 20, no. 4, pp. 1033-1048, 1992.

D. Bhattacharyya and T.-h. Kim, “Brain tumor detection using MRI image analysis,” in *International Conference on Ubiquitous Computing and Multimedia Applications*, pp. 307-314, Springer, 2011.

C. M. Bishop, “Pattern recognition,” *Machine Learning*, vol. 128, 2006.

R. Bitar, et al. "MR pulse sequences: What every radiologist wants to know but is afraid to ask 1." *Radiographics* 26.2: 513-537, 2006.

L. Breiman, “Bagging predictors,” *Technical Report* 421, Department of Statistics, University of California, Berkeley, 1994.

L. Breiman, “Bagging predictors,” *Machine learning*, vol. 24, no. 2, pp. 123-140, 1996.

F. W. Campbell and J. Robson, “Application of fourier analysis to the visibility of gratings,” *The Journal of physiology*, vol. 197, no. 3, p. 551, 1968.

N. D. of Clinical Neurosciences, “Oxford Centre for Functional MRI of the Brain,” disponível em: <https://www.ndcn.ox.ac.uk/divisions/fmrib>. Acesso em 07 jul. 2016.

J. M. Coggins and A. K. Jain, “A spatial filtering approach to texture analysis,” *Pattern recognition letters*, vol. 3, no. 3, pp. 195-203, 1985.

W. J. Conover, *Practical Nonparametric Statistics*, John Wiley & Sons, (2nd Edition), pp. 225–226, 1980.

D. Cornfeld, and J. Weinreb. "Simple changes to 1.5-T MRI abdomen and pelvis protocols to optimize results at 3 T." *American Journal of Roentgenology* 190.2: W140-W150, 2008.

L. Cun, B. Boser, et al., “Handwritten digit recognition with a back-propagation network,” *Advances in neural information processing systems*. 1990.

W. H. E. Day, “Consensus methods as tools for data analysis,” In H. H. Bock, editor, *Classification and Related Methods for Data Analysis*, Elsevier Science Publishers B.V. (North Holland), pp. 317–324, 1988.

R. Dass and S. Devi, "Image Segmentation Techniques 1," 2012.

L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology*, vol. 26, no. 3, pp. 297-302, 1945.

T. G. Dietterich, “Ensemble methods in machine learning,” in *International workshop on multiple classifier systems*, pp. 1-15, Springer, 2000.

J. Diedrichsen, “A spatially unbiased atlas template of the human cerebellum,” *Neuroimage*, vol. 33, no. 1, pp. 127-138, 2006.

J. Diedrichsen, J. H. Balsters, J. Flavell, E. Cussans, and N. Ramnani, “A probabilistic mr atlas of the human cerebellum,” *Neuroimage*, vol. 46, no. 1, pp. 39-46, 2009.

J. Diedrichsen, S. Maderwald, M. Küper, M. Thürling, K. Rabe, E. Gizewski, M. E. Ladd, and D. Timmann, "Imaging the deep cerebellar nuclei: a probabilistic atlas and normalization procedure," *Neuroimage*, vol. 54, no. 3, pp. 1786-1794, 2011.

J. Diedrichsen and E. Zotow, "Surface-based display of volume-averaged cerebellar imaging data," *PloS one*, vol. 10, no. 7, p. e0133402, 2015.

R. C. Dubes and A. K. Jain, "Random field models in image analysis," *Journal of applied statistics*, vol. 16, no. 2, pp. 131-164, 1989.

R. P. Duin, "The combining classifier: to train or not to train?," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, vol. 2, pp. 765-770, IEEE, 2002.

N. Duta and M. Sonka, "Segmentation and interpretation of mr brain images. An improved active shape model," *IEEE Transactions on Medical Imaging*, vol. 17, no. 6, pp. 1049-1062, 1998.

B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*. CRC press, 1994.

A. Elnakib, G. Gimelfarb, J. S. Suri, and A. El-Baz, "Medical image segmentation: a brief survey," in *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, pp. 1-39, Springer, 2011.

FECOMERCIOSP, "Saúde pública: prazo entre marcar e realizar consultas tem pior nota do paulistano," disponível em <http://www.fecomercio.com.br/noticia/saude-publica-prazo-entre-marcar-e-realizar-consultas-tem-pior-nota-do-paulistano>. Acesso em 20 jul. 2016.

I. Franca, "Z Y X O Planos Coordenados Espaço tridimensional (3D)," disponível em <http://slideplayer.com.br/slide/3680802/>. Acesso em 04 jul. 2016.

S. Franiatte, "Simple single-seeded region growing," disponível em <https://www.mathworks.com/matlabcentral/fileexchange/35269-simple-single-seeded-region-growing>. Acesso em 20 abr. 2015.

J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics Springer, Berlin, 2001.

M. L. Gonçalves, M. L. de Andrade Netto, and J. A. F. Costa, "Explorando as propriedades do mapa auto-organizável de kohonen na classificação de imagens de satélite."

C. Guyeux and J Bahi, "An improved watermarking algorithm for internet applications," *INTERNET'2010. The 2nd Int. Conf. on Evolving Internet*. 2010.

M. Haindl, "Texture synthesis, " *ERCIM*, 1993.

R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer vision, graphics, and image processing*, vol. 29, no. 1, pp. 100-132, 1985.

P. E. Hart, D. G. Stork, and R. O. Duda, "Pattern classification," *John Willey & Sons*, 2001.

S. L. Hartmann, M. H. Parks, P. R. Martin, and B. M. Dawant, "Automatic 3-d segmentation of internal structures of the head in mr images using a combination of similarity and free-form transformations. ii. validation on severely atrophied brains," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 917-926, 1999.

HIS, "The Complexities of Physician Supply and Demand: Projections from 2014 to 2025 – 2016 Update," 2016.

T. K. Ho, "Multiple classifier combination: Lessons and the next steps". In A. Kandel and H. Bunke, editors, *Hybrid Methods in Pattern Recognition*. World Scientific Publishing, pp. 171–198, 2002.

T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision combination in multiple classifier systems," *IEEE transactions on pattern analysis and machine intelligence*, vol. 16, no. 1, pp. 66-75, 1994.

S. Hojjatoleslami and F. Kruggel, "Rapid communications ," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, p. 667, 2001.

T. Honkela, "*Self-organizing maps in natural language processing*," Diss. Helsinki University of Technology, 1997.

N. C. Institute, "Cancer statistics.", disponível em <http://www.cancer.gov/about-cancer/what-is-cancer/statistics>. Acesso em 12 abr. 2015.

I. Isgum, M. Staring, A. Rutten, M. Prokop, M. A. Viergever, and B. van Ginneken, "Multi-atlas-based segmentation with local decision fusionapplication to cardiac and aortic segmentation in ct scans," *IEEE transactions on medical imaging*, vol. 28, no. 7, pp. 1000-1010, 2009.

M. A. Jaffar, et al., "Wavelet-Based Color Image Segmentation using Self-Organizing Map Neural Network," *2009 International Conference on Computer Engineering and Applications IPCSIT*. Vol. 2. 2011.

M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "Fsl," *Neuroimage*, vol. 62, no. 2, pp. 782-790, 2012.

M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 1, no. 4, pp. 321-331, 1988.

Kittipatkampa, "Image Segmentation using Gaussian Mixture Models," disponível em <https://kittipatkampa.wordpress.com/2011/02/17/image-segmentation-using-gaussian-mixture-models/>. Acesso em 12 nov. 2015.

J. Kittler, M. Hatef, R. P. Duin, and J. Matas, "On combining classifiers," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 3, pp. 226-239, 1998.

A. Klein and J. Tourville, "101 labeled brain images and a consistent human cortical labeling protocol," *Frontiers in neuroscience*, vol. 6, p. 171, 2012.

T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological cybernetics*, vol. 43, no. 1, pp. 59-69, 1982.

L. I. Kuncheva, *Combining pattern classifiers: methods and algorithms*. John Wiley & Sons, 2004.

L. Lam and C. Y. Suen, "A theoretical analysis of the application of majority voting to pattern recognition," in *Pattern Recognition, 1994. Vol. 2-Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on*, vol. 2, pp. 418-420, IEEE, 1994.

L. Lam and S. Suen, “Application of majority voting to pattern recognition: an analysis of its behavior and performance,” *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 27, no. 5, pp. 553-568, 1997.

L. Lança, “Potencial de redução de dose em Tomografia Computorizada com técnicas de reconstrução iterativa,” *52° Congresso Nacional e Internacional dos Profissionais das Técnicas Radiológicas*, 2012.

P. C. Lauterbur, “Image formation by induced local interactions: examples employing nuclear magnetic resonance,” pp. 190-191, 1973.

K. I. Laws, “Textured image segmentation,” tech. rep., DTIC Document, 1980.

B. Li, G. E. Christensen, E. A. Hoffman, G. McLennan, and J. M. Reinhardt, “Establishing a normative atlas of the human lung: intersubject warping and registration of volumetric ct images,” *Academic Radiology*, vol. 10, no. 3, pp. 255-265, 2003.

H. Li, K. R. Liu, and S.-C. Lo, “Fractal modeling and segmentation for the enhancement of microcalcifications in digital mammograms,” *IEEE transactions on medical imaging*, vol. 16, no. 6, pp. 785-798, 1997.

C. Li and P. K.-S. Tam, “An iterative algorithm for minimum cross entropy thresholding,” *Pattern Recognition Letters*, vol. 19, no. 8, pp. 771-776, 1998.

W.-C. Lin, E. C.-K. Tsao, and C.-T. Chen, “Constraint satisfaction neural networks for image segmentation,” *Pattern Recognition*, vol. 25, no. 7, pp. 679-693, 1992.

X. Lin, S. Yacoub, J. Burns, and S. Simske, “Performance analysis of pattern classifier combination by plurality voting,” *Pattern Recognition Letters*, vol. 24, no. 12, pp. 1959-1969, 2003.

J. Looman and J. Campbell, “Adaptation of sorensen's k (1948) for estimating unit affinities in prairie vegetation,” *Ecology*, vol. 41, no. 3, pp. 409-416, 1960.

S. Lynn, “Self-Organising Maps for Customer Segmentation using R,” disponível em <http://www.shanelynn.ie/self-organising-maps-for-customer-segmentation-using-r/>. Acesso em 03 mar. 2016.

J. Malik and P. Perona, “Preattentive texture discrimination with early vision mechanisms,” *JOSA A*, vol. 7, no. 5, pp. 923-932, 1990.

B. B. Mandelbrot, D. E. Passoja, and A. J. Paullay, “Fractal character of fracture surfaces of metals,” 1984.

H. B. Mann, D. R. Whitney, “On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other,” *Annals of Mathematical Statistics*. 18 (1): 50–60, 1947.

P. Mansfield and A. A. Maudsley, “Medical Imaging by NMR,” *British Journal of Radiology*, **50** (591), 188-194, 1977.

J. L. Marroquín, B. C. Vemuri, S. Botello, E. Calderon, and A. Fernandez-Bouzas, “An accurate and efficient bayesian method for automatic segmentation of brain mri,” *IEEE Transactions on Medical Imaging*, vol. 21, no. 8, pp. 934-945, 2002.

B. McCune, J. B. Grace, and D. L. Urban, *Analysis of ecological communities*, vol. 28. MjM software design Gleneden Beach, OR, 2002.

A. J. McMichael, "Environmental and social influences on emerging infectious diseases: past, present and future." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359.1447: 1049-1058, 2004.

MindBoogle, “MindBoogle Data,” disponível em: <http://www.mindboggle.info/data.html>. Acesso em 05 mar. 2015.

MRBrainS – E. F. of MR Brain Image Segmentations, “Results,” disponível em <http://mrbrains13.isi.uu.nl/results.php>. Acesso em 17 out. 2014.

W. T. C. for Neuroimaging, “WTCN Home,” disponível em <http://www.fil.ion.ucl.ac.uk/>. Acesso em 06 jul. 2016.

H. Ng, S. Ong, K. Foong, P. Goh, and W. Nowinski, “Medical image segmentation using k-means clustering and improved watershed algorithm,” in 2006 *IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 61-65, IEEE, 2006.

K. H. Nicolaides et al., “Fetal nuchal translucency: ultrasound screening for chromosomal defects in first trimester of pregnancy,” *Bmj* 304.6831: 867-869, 1992.

NKI, “Nathan Kline Institute (NKI) / Rockland Sample,” disponível em: <http://fcon_1000.projects.nitrc.org/indi/pro/nki.html>. Acesso em 20 jul. 2016.

NobelPrize, “The Nobel Prize in Physics 1901,” disponível em <https://www.nobelprize.org/nobel_prizes/physics/laureates/1901/rontgen-photo.html>. Acesso em 20 jul. 2016.

D. Nogare, “Entendendo como funciona o algoritmo de Cluster K-Means,” disponível em <http://www.diegonogare.net/2015/08/entendendo-como-funciona-o-algoritmo-de-cluster-k-means/>. Acesso em 03 jul. 2016.

A. Norouzi, M. Rahim, et al., “Medical image segmentation methods, algorithms, and applications,” *IETE Technical Review*, v. 31, n. 3, p. 199-213, 2014.

OASIS, “MRI Reliability data across the adult lifespawn,” disponível em <http://www.oasis-brains.org/app/action/BundleAction/bundle/OAS1_RELIABILITY>. Acesso em 20 jul. 2016.

N. Otsu, “A threshold selection method from gray-level histograms,” *Automatica*, vol. 11, no. 285-296, pp. 23-27, 1975.

K. R. Rao, “IMAGE ANALYSIS AND OBJECT DETECTION-RECOGNITION USING COMPUTER VISION ALGORITHMS,” *The University of Texas Arlinton*, 2015.

M. M. Reis, “Testes de Diferenças entre Médias,” disponível em: <http://www.inf.ufsc.br/~marcelo/testes2.html>. Acesso em 16 fev. 2016.

D. Reynolds, “Gaussian mixture models,” *Encyclopedia of biometrics*, pp. 827-832, 2015.

D. W. Roberts, “Ordination on the basis of fuzzy set theory,” *Vegetatio*, vol. 66, no. 3, pp. 123-131, 1986.

T. Rohlfing, R. Brandt, R. Menzel, D. B. Russakoff, and C. R. Maurer Jr, “Quo vadis, atlas-based segmentation?,” in *Handbook of Biomedical Image Analysis*, pp. 435-486, Springer, 2005.

T. Rohlfing and C. R. Maurer, “Shape-based averaging,” *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 153-161, 2007.

T. Rohlfing, D. B. Russakoff, and C. R. Maurer, “Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation,” *IEEE transactions on medical imaging*, vol. 23, no. 8, pp. 983-994, 2004.

M. Rousson and N. Paragios, “Shape priors for level set representations,” in *European Conference on Computer Vision*, pp. 78-92, Springer, 2002.

D. Ruta and B. Gabrys, “A theoretical analysis of the limits of majority voting errors for multiple classifier systems,” *Pattern Analysis & Applications*, vol. 5, no. 4, pp. 333-350, 2002.

W. S. Salem, A. F. Seddik, and H. F. Ali, “A review on brain mri image segmentation,” 2013.

Scikit-image, “Thresholding,” disponível em <http://scikit-image.org/docs/dev/auto_examples/segmentation/plot_otsu.html>. Acesso em 04 jul. 2016.

D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*, John Wiley & Sons, 2015.

M. Sezgin et al., “Survey over image thresholding techniques and quantitative performance evaluation,” *Journal of Electronic imaging*, vol. 13, no. 1, pp. 146-168, 2004.

S. Shah, “A Glance at Ensembling,” disponível em <https://saatvikshah1994.github.io/blog-post/Ensembling_1/>. Acesso em 09 dez. 2015.

L. Shapley and B. Grofman, “Optimizing group judgmental accuracy in the presence of interdependencies,” *Public Choice*, vol. 43, no. 3, pp. 329-343, 1984.

N. Sharma, and L. M. Aggarwal. “Automated Medical Image Segmentation Techniques,” *Journal of Medical Physics / Association of Medical Physicists of India* 35.1 (2010): 3–14. *PMC*. Web. 20 July 2016.

D. Shen, E. H. Herskovits, and C. Davatzikos, “An adaptive-focus statistical shape model for segmentation and shape modeling of 3-d brain structures,” *IEEE transactions on medical imaging*, vol. 20, no. 4, pp. 257-270, 2001.

W. Skarbek, A. Koschan, and Z. Veroffentlichung. "Colour image segmentation-a survey," 1994.

S. M. Smith, “Fast robust automated brain extraction,” *Human brain mapping*, vol. 17, no. 3, pp. 143-155, 2002.

S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, et al., “Advances in functional and structural mr image analysis and implementation as fsl,” *Neuroimage*, vol. 23, pp. S208-S219, 2004.

R. Smith-Bindman, D. L. Miglioretti, and E. B. Larson. “Rising Use Of Diagnostic Medical Imaging In A Large Integrated Health System: The Use of Imaging Has Skyrocketed in the Past Decade, but No One Patient Population or Medical Condition Is Responsible,” *Health affairs (Project Hope)*27.6 (2008): 1491–1502. *PMC*. Web. 20 July 2016.

M. Sonka, S. K. Tadikonda, and S. M. Collins, “Knowledge-based interpretation of mr brain images,” *IEEE transactions on medical imaging*, vol. 15, no. 4, pp. 443-452, 1996.

T. Sørensen, “A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons,” *Biologiske Skrifter*, vol. 5, pp. 1-34, 1948.

SPM, “Statistical Parametric Mapping,” disponível em <http://www.fil.ion.ucl.ac.uk/spm/#brief>. Acesso em 04 ago. 2015.

J. Stancanello, P. Romanelli, N. Modugno, P. Cerveri, G. Ferrigno, F. Uggeri, and G. Cantore, "Atlas-based identification of targets for functional radiosurgery," *Medical physics*, vol. 33, no. 6, pp. 1603-1611, 2006.

SUIT, "A spatially unbiased atlas template of the cerebellum and brainstem (SUIT)," disponível em <http://www.diedrichsenlab.org/imaging/suit.htm>. Acesso em 10 ago. 2015.

A. Tanács, E. Máté, and A. Kuba. "Application of automatic image registration in a segmentation framework of pelvic CT images," *International Conference on Computer Analysis of Images and Patterns*. Springer Berlin Heidelberg, 2005.

S. Tatiraju and A. Mehta, "Image segmentation using k-means clustering, em and normalized cuts," *University Of California Irvine*, 2008.

D. Terzopoulos, "On matching deformable models to images," in *Topical Meeting on Machine Vision Tech. Digest Series*, vol. 12, pp. 160-167, 1987.

N. Y. Times, "The Health Care Waiting Game," disponível em <http://www.nytimes.com/2014/07/06/sunday-review/long-waits-for-doctors-appointments-have-become-the-norm.html?_r=0>. Acesso em 20 jul. 2016.

A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky, "Mutual information in coupled multi-shape model for medical image segmentation," *Medical Image Analysis*, vol. 8, no. 4, pp. 429-445, 2004.

A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE transactions on medical imaging*, vol. 22, no. 2, pp. 137-154, 2003.

L. Torok, A. de Imagens, and A. Conci, "Método de otsu," 2015.

J. D. Tubbs and W. O. Alltop, "Measures of confidence associated with combining classification results," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 3, pp. 690-692, 1991.

S.-Y. Wan and W. E. Higgins, “Symmetric region growing,” *IEEE Transactions on Image processing*, vol. 12, no. 9, pp. 1007-1015, 2003.

F. Wilcoxon, “Individual comparisons by ranking methods,” *Biometrics bulletin* 1.6: pp. 80-83, 1945.

O. Wirjadi, *Survey of 3d image segmentation methods*, vol. 35. ITWM, 2007.

D. H. Wolpert and W. G. Macready, *No free lunch theorems for search*, Vol. 10. Technical Report SFI-TR-95-02-010, Santa Fe Institute, 1995.

M. W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. M. Smith, “Bayesian analysis of neuroimaging data in fsl,” *Neuroimage*, vol. 45, no. 1, pp. S173-S186, 2009.

L. Xu, A. Krzyzak, and C. Y. Suen, “Methods of combining multiple classifiers and their applications to handwriting recognition,” *IEEE transactions on systems, man, and cybernetics*, vol. 22, no. 3, pp. 418-435, 1992.

C. Xu, D. L. Pham, and J. L. Prince, “Image segmentation using deformable models,” *Handbook of medical imaging*, vol. 2, pp. 129-174, 2000.

J. Yang, L. H. Staib, and J. S. Duncan, “Neighbor-constrained segmentation with level set based 3-d deformable models,” *IEEE Transactions on Medical Imaging*, vol. 23, no. 8, pp. 940-948, 2004.

E. Yoxen, “Seeing with sound: A study of the development of medical images,” *The social construction of technological systems*, pp. 281-306, 1987.