Evaluation Method for MRI Brain Tissue Abnormalities Segmentation Study

SHAFAF IBRAHIM, NOOR ELAIZA ABDUL KHALID, MAZANI MANAF
Faculty of Computer Science and Mathematics
University Technology MARA
Shah Alam, Selangor
MALAYSIA
elaiza@tmsk.uitm.edu.my

Abstract: - Segmentation poses one of the most challenging problems in medical imaging. Segmentation of Magnetic Resonance Imaging (MRI) images is an important part of brain imaging research as it can facilitates the neurological diseases diagnosis. However, there are few limitations in evaluating the segmentation accuracy due to difficulties in obtaining the ground truth. This research proposes an evaluation method for brain tissue abnormalities segmentation study. Controlled experimental data called mosaic images are used as the testing data. The data is designed which that prior knowledge of the size of the abnormalities is known. It is done by cutting various shapes and sizes of various abnormalities and pasting it onto normal brain tissues, where the tissues and the background are divided into three different intensities. The knowledge of the size of abnormalities by number of pixels are then used as the ground truth to compare with the various segmentation results. The validation of segmentation was done with fifty data of each category using methods of Particle Swarm Optimization (PSO), Adaptive Network-based Fuzzy Inference System (ANFIS) and Fuzzy c-Means (FCM), where the evaluation for each technique exhibits some variation of results. Therefore, the proposed evaluation method of ground truth formation called image mosaicing is found to be reasonable and acceptable to use as it produces potential solutions to the current difficulties in evaluating the brain tissue abnormalities segmentation outcome.

Key-Words: - Mosaicing,Evaluation method, Medical imaging, Segmentation, Magnetic Resonance Imaging (MRI).

1 Introduction

Segmentation is the labeling of objects in image data and has been a crucial stage in many medical imaging processing tasks for operation planning, radio therapy or diagnostics, and studying the differences of healthy subjects and subjects with tumor. Its purpose is to subdivide an image into meaningful non-overlapping regions which analysis, interpretation or quantification can be performed [1]. In the past years, a large number of researches have focus on the development medical image segmentation methods as accurate segmentation of biomedical images can contribute to improve diagnosis, surgical planning and prognosis [2]. These leads to the increasing number in investigations of applications and considerable effort is needed to find reliable and accurate algorithms to solve the difficulties in evaluating the segmentation accuracy.

In past several years, medical image segmentation problems has been approached with several solution methods by different levels of automation and range of applicability such as Particle Swarm Optimization [3], Genetic Algorithm [4], Region Growing [5], [7], Adaptive Network-based Fuzzy Inference System (ANFIS) [6], Self Organizing Map (SOM) [8] and Fuzzy c-Mean (FCM) [9]. However, segmenting brain internal structures remains a challenging task due to their small size, partial volume effects, anatomical variability, and the lack of clearly defined edges [10]. Thus, a thorough evaluation of its performance is necessary with some quantifiable measurement of its accuracy and variability [11].

Evaluation is not only used in evaluating the performance of segmentation algorithms. It could also be used in combining the results of several segmentation results [12], and acted as a guide in selecting appropriate segmentation algorithms [13]. Nevertheless, evaluation of segmentation performance has been very subjective that leaves the researcher in tricky situation [14]. Therefore, it may leads to difficulties in judging the effectiveness of the techniques implemented. Chabrier et al. [15] found that it is difficult to evaluate the segmentation methods accuracy and efficiency on a single method as no one being optimal in all cases.

The common standard used for validating segmentation results is the manual segmentation results done by medical experts [16], [17]. A main issue is that obtaining these validation data and comparison metrics for segmentation are difficult tasks due to the lack of reliable ground truth [18], [19]. Unfortunately, lack of reliability and reproducibility of manual segmentation method should also be addressed [20]. Thus, even if a rich set of manual segmentations are available, they may not reflect the ground truth and the true gold standard may need to be estimated [21].

Another typical technique is the use of phantoms. For segmentation purpose, phantoms are usually refers to synthetic images for which the true segmentation is known [22], [23]. A physical object can also be used as a phantom ground truth where the phantom is measured and imaged. The original true measurement and segmentation measurements are then compared and performance is thus assessed [24]. However, for many medical problems, phantom studies are considered insufficient for validation and it is exceedingly difficult to design phantoms that appropriately imitate the living tissues [11].

From the reviews done, it can be summarize that one of the biggest challenges in the medical imaging domain is to accurately and reliably quantify and optimize the clinical performance of the image segmentation algorithms and outcomes. Therefore, this study proposed evaluation method that able to generate synthetic ground truth of MRI images called image mosaicing that exhibit comparable segmentation challenges to real MRI brain tissue abnormalities. The validation of the proposed image mosaicing for evaluation of brain abnormalities segmentation is then performed using three methods of Particle Swarm Optimization (PSO), Adaptive Network-based Fuzzy Inference System (ANFIS) and Fuzzy c-Means (FCM).

The organization of the rest of this paper is as follows: Section 2 presents our methods, including image mosaicing description, PSO, ANFIS and FCM algorithms structure. Section 3 discusses our results and discussions. Finally, we present our conclusion in Section 4.

2 Methods

In several previous studies done by researchers, image mosaicing strategy has been successfully employed in many research and application areas, as well as various representations of methods for image segmentation such as texture mapping [25], texture segmentation [26], edge detection [27], texturing three-dimensional modeling [28] and texture registration [29].

This research emphasizes on the development of evaluation method in solving one of the difficulties of brain tissue abnormalities segmentation study. In this paper, controlled experimental data as the testing data called mosaic images is proposed. The data is designed which that prior knowledge of the size of the abnormalities are known. This is done by cutting various shapes and sizes of various abnormalities and pasting it onto normal brain tissues, where it consists of three basic steps. The pictorial representation of the proposed mosaic image generation process is illustrated in Fig. 1.

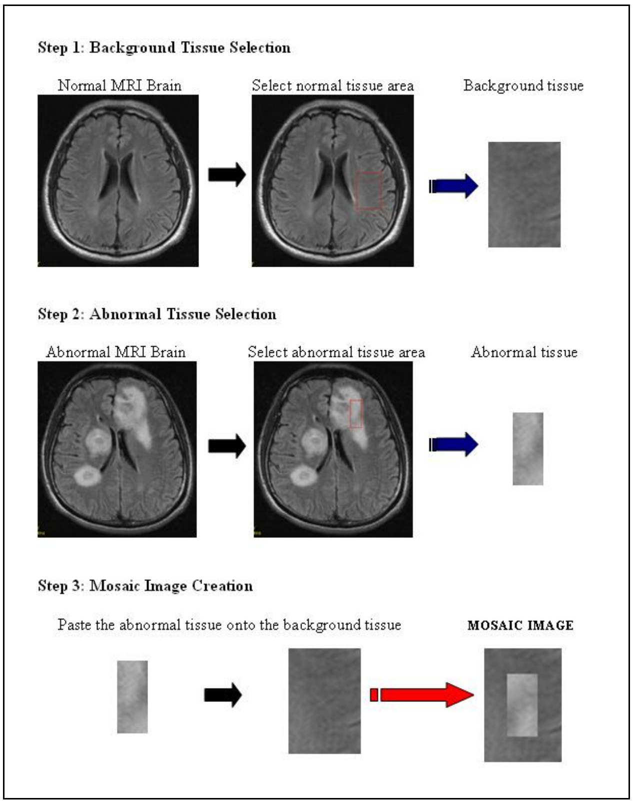
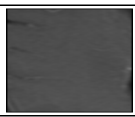


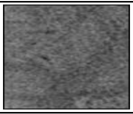
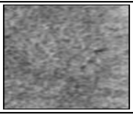
Fig. 1 Mosaic image generation process

Step 1: Background tissue selection

The background tissues are selected from normal area of brain tissue or so-called as membrane. There are three different categories of background tissues intensities which are low, medium and background as tabulated in Table 1.

Table 1. Background Images

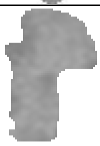
Background	Intensity	Min	Max	Size in pixel
	Low	30	114	12144

	Medium	39	145	12144
	High	56	202	12144

Step 2: Abnormal tissue selection

The abnormal tissues are picked out from the abnormalities area. There are three possible shapes of selecting the abnormalities such as square, oval or irregular shapes as illustrated in Table 2.

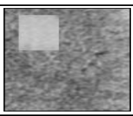
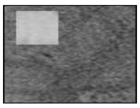
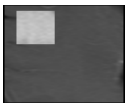
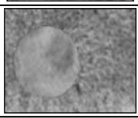
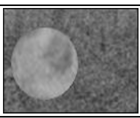
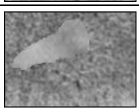
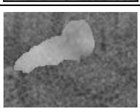
Table 2. Samples of Abnormalities Images

Shape	Image	Size in pixel
Square		1472
Oval		2010
Irregular		1901

Step 3: Paste the abnormal tissue onto the background tissue

Mosaic images are then created by pasting the selected abnormality tissues onto the different background tissues as shown in Table 3. This is used to test out the performance and accuracy of the segmentation outcome.

Table 3. Mosaic Images

Background / Abnormality	High	Medium	Low
Square			
Oval			
Irregular			

3 Results and Discussion

The numbers of pixels of the raw MRI brain images are compared with the segmented abnormality area. The segmentation accuracy results are then measured by considering the value of false positive, false negative, true positive and true negative. Four conditions areas of false positive, false negative, true positive and true negative are illustrates in Fig. 2.

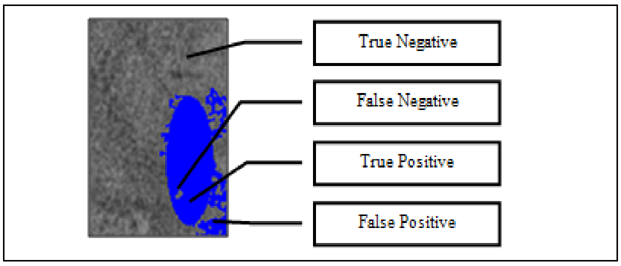


Fig. 2. Sample of mosaic image of oval abnormality within medium background intensity after segmentation

The four primary conditions are used for quantifying the qualities of segmentation outcomes. The descriptions, states and calculations for each condition are explained in Table 4.

Table 4. Conditions of accuracy

Condition	Description	State	Calculation
False Positive	normal areas that are incorrectly detected as abnormality	If segmented area > abnormality area	(segmented area - abnormality area) / background size in pixels
		If segmented area <= abnormality area	0
False Negative	abnormality areas that are not detected	If segmented area >= abnormality area	0
		If segmented area < abnormality area	(abnormality area - segmented area) / background size in pixels
True Positive	abnormality areas that are correctly detected	If segmented area >= abnormality area	1
		If segmented area < abnormality area	1 - (abnormality area - segmented area) / background size in pixels
True Negative	normal areas that are correctly undetected	If segmented area > abnormality area	1 - (segmented area - abnormality area) / background size in pixels
		If segmented area <= abnormality area	1

ISBN: 978-1-61804-019-0

Table 5 shows a few samples of segmentation outcomes tested to square, oval and irregular shapes of abnormalities onto the high, medium and low background tissue intensities. Three techniques had been chosen for the testing purpose which are Particle Swarm Optimization (PSO), Adaptive Network-based Fuzzy Inference System (ANFIS) and Fuzzy c-Means (FCM). Then, every mean value of false positive, false negative, true positive and true negative is evaluated by relating the results to any certain circumstances.

Table 5. PSO vs ANFIS vs FCM segmentation

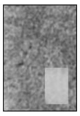
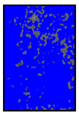
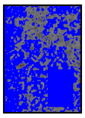
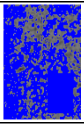
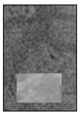
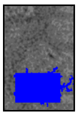
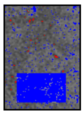
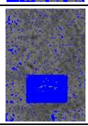
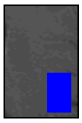
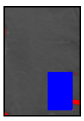
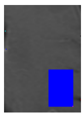
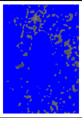
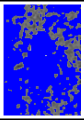
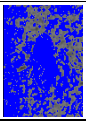
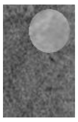
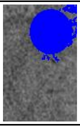
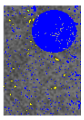
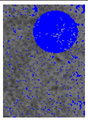
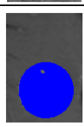
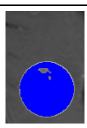
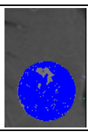
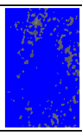
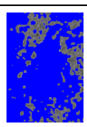
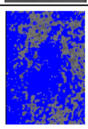
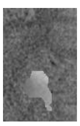
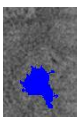
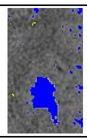
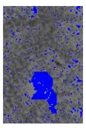
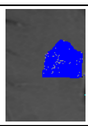
Shape	B/Ground Intensity	Mosaic Image	PSO	ANFIS	FCM
Square	High				
	Medium				
	Low				
Oval	High				
	Medium				
	Low				
Irregular	High				
	Medium				
	Low				

Table 6 tabulates the summary of Receiver Operating Characteristic (ROC) analysis for PSO, ANFIS and FCM segmentation results which includes the entire four primary conditions. These statistical values are used to quantify the segmentation quality and the level of accuracy for each technique in three different types of background intensities which are high, medium and low background grey level value.

Table 6. Summary of ROC analysis for PSO, ANFIS and FCM

Technique	B/Ground Grey Level Value	Mean of False Positive	Mean of False Negative	Mean of True Positive	Mean of True Negative
PSO	High	0.837	0	1	0.173
	Medium	0.029	0	1	0.971
	Low	0.009	0.009	0.991	0.991
ANFIS	High	0.532	0	1	0.468
	Medium	0.006	0.006	0.994	0.994
	Low	0	0.027	0.973	1
FCM	High	0.469	0	1	0.531
	Medium	0.038	0.001	0.999	0.962
	Low	0	0.011	0.989	1

As seen from the Table 6, PSO shows the most excellent segmentation result in low background grey level value intensity. The statistics show that the combination of abnormality within the low background grey level value produced the highest mean values for both true positive and true negative which are the most important condition in producing good quality of segmentation. The combination of abnormality within the medium background grey level value also cannot be underestimated since it produced high mean values for both true positive and true negative. However, small occurrence of mean value of false positive is observed. The combination of abnormality within the high background grey level value is seen to produce poor performance as it appears the highest mean value of false positive compared to the medium and low background grey level value. This is found to be caused by the texture similarity for both abnormality and high background grey level value that leads the neighboring pixels to grow beyond the abnormality areas.

In contrast, the ANFIS produced a good performance of segmentation in medium background grey level value intensity as it produced the highest mean values for both true positive and true negative. The performance of segmentation in low background grey level value returns good segmentation outcome as it displays excellent mean value of true negative. However, there is slightly lower in mean of true positive value since it is affected by the small occurrence of false negative. Same as PSO, the segmentation of abnormality within the high background grey level value performed unsatisfactorily since it produced the highest mean value of false positive compared to the medium and low background grey level value.

ISBN: 978-1-61804-019-0

As PSO, the FCM demonstrates the best segmentation outcomes in low background grey level value intensity too. Followed by the medium and high background grey level value as it appears low values for both true positive and true negative respectively.

The second analysis method employed is Pearson's correlation. Pearson's correlation is widely used to reflect the degree of linear relationship between two variables. In this paper, the Pearson correlation value of three categories between the original abnormalities area vs PSO, original abnormalities area vs ANFIS, and original abnormalities area vs FCM segmentation pixels value are measured so that the variation of results obtained can be clearly monitored. The Pearson's correlation value for each categories mentioned is presented in Table 7:

Table 7. Pearson's correlation for PSO, ANFIS and FCM

Description	B/Ground	Correlation value
Original vs PSO Correlation	High	0.872
	Medium	0.993
	Low	0.999
Original vs ANFIS Correlation	High	0.894
	Medium	0.999
	Low	0.999
Original vs FCM Correlation	High	0.927
	Medium	0.998
	Low	0.507

From the table above, it clearly noticed that PSO and ANFIS correlation values are almost excellent in abnormalities segmentation regardless of background. However, the correlation values of the FCM shows a slightly lower value in low background tissue intensity.

4 Conclusion

This paper has presented an evaluation method for brain tissue segmentation study. The application to a variety of MRI brain medical data has been successful. The techniques of Particle Swarm Optimization (PSO), Adaptive Network-based Fuzzy Inference System (ANFIS) and Fuzzy c-Means (FCM) had been tested for the evaluation purpose, and exhibit some variation of results. Therefore, it can be conclude that the proposed evaluation method of image mosaicing is found to be reasonable and acceptable to use as it produces potential solutions to the current difficulties in evaluating and validating the brain tissue abnormalities segmentation outcome.

Acknowledgment

The authors acknowledge with gratitude the great support from the management of Research Management Institute (RMI), UiTM and financial support from E-Science Fund (06-01-01-SF0306) under the Minister of Science, Technology and Innovation (MOSTI), Malaysia.

References:

[1] N. R. Pal, and S. K. Pal, A Review on Image Segmentation Techniques, *Pattern Recognition*, Vol. 26, Issue 9, pp. 1277-1294. 1993.

[2] H. Nayak, M. M. Amini, P. T. Bibalan, and N. Bacon, *Medical Image Segmentation*, June 12, 2008.

[3] S. Ibrahim, N.E.A Khalid, and M. Manaf, Empirical Study of Brain Segmentation using Particle Swarm Optimization, *International Conference on Information Retrieval and Knowledge Management*, CAMP'10, pp. 235-239, 2010.

[4] N.E.A. Khalid, S. Ibrahim, M. Manaf, and U.K. Ngah, Seed-Based Region Growing Study for Brain Abnormalities Segmentation, *International Symposium on Information Technology 2010 (ITSim 2010)*, Vol. 2, pp. 856 - 860 2010.

[5] R. Ganesan and S.Radhakrishnan, Segmentation of Computed Tomography Brain Images using Genetic Algorithm, *International Journal of Soft Computing*, Vol. 4, Issue 4, pp. 157-161, 2009.

[6] N.M. Noor, N.E.A. Khalid, R. Hassan, S. Ibrahim, and I.M. Yassin, Adaptive Neuro-Fuzzy Inference System for Brain Abnormality Segmentation, *2010 IEEE Control and System Graduate Research Colloquium*, ICSGRC 2010, 2010.

[7] C. Wasserthal,,K. Engel, K. Rink, and A. Brechman, Automatic segmentation of the cortical grey and white matter in MRI using Region Growing Approach based on Anatomical Knowledge, *Springer Berlin Heidelberg*, ISBN978-3-540-78639-9. 2008.

[8] K. M. Iftekharuddin, J. Zheng, M. A. Islam, R. J., and F. Lanningham, Brain tumor detection in MRI: technique and statistical validation, *Fortieth Asilomar Conference on Signals, Systems and Computers*, Oct. 29 2006-Nov. 1 2006, pp. 1983-1987, 2006.

[9] S. Shen, W. Snadham, M. Granat, and A. Sterr., MRI Fuzzy segmentation of brain tissue using neighborhood attraction with neural network optimization, *IEEE*

Transactions on Information Technology in Biomedicine, Vol. 9, Issue 3, Sept. 2005, pp. 459-467, 2005.
[10] M. G. Linguraru, M. Á. G. Ballester, N. Ayache, Deformable Atlases for the Segmentation of Internal Brain Nuclei in Magnetic Resonance Imaging, *International Journal of Computers, Communications & Control*, 2007(II): pp. 26–36, 2007.
[11] S. Bouix, M. Martin-Fernandez, L. Ungar, M. Nakamura, M. Koo, R. McCarley, and M. Shenton, On evaluating brain tissue classifiers without a ground truth, *Neuroimage* 07, 447458, 2007.
[12] C. Reosenberger, and K. Chehdi, Genetic Fusion: Application to Multi-components Image Segmentation, *Proceedings ICASSP*, 4: pp. 2223-2226, 2000.
[13] Y.J. Zhang, and H.T. Luo Optimal Selection of Image Segmentation Algorithms based on Performance Evaluation, Optical Engineering, Vol. 39, Issue 6, pp. 1450-1456, 2000.
[14] R. Unnikrishnan, C. Pantofaru, and M. Hebert, A Measure for Objective Evaluation of Image Segmentation Algorithms, *Workshop on Empirical Evaluation Methods in Computer Vision, IEEE Conference on Computer Vision and Pattern Recognition (CVPR '05)*, June 2005, 2005.
[15] S. Chabrier, H. Lauren, and B. Emile, Performance Evaluation of Image Segmentation Application to Parameters Fitting, *Proceedings of European Signal Processing*, 2005.
[16] V. Grau, A.U.J. Mewes, M. Alcaniz, R. Kikinis, S.K. Warfield, Improved watershed transform for medical image segmentation using prior information, *IEEE Trans. Med. Imaging*, Vol. 23, Issue 4, pp. 447-458, 2004.
[17] D.E Rex, D.W. Shattuck, R.P. Woods, K.L. Narr, E. Luders, K. Rehm, S.E. Stolzner, D.A. Rottenberg, A.W. Toga, A meta-algorithm for brain extraction in MRI. *NeuroImage 23*, pp. 625–627, 2004.
[18] M. Prastawa, E. Bullita, G. Gerig, Simulation of Brain Tumors in MR Images for Evaluation of Segmentation Efficacy, Medical Image Analysis, *Elsevier*, Vol. 13, Issue 2, April 2009, pp. 29-311, 2009.
[19] X. Lin, T. Qiu, F. Nicolier, S. Ruan, Automatic Hippocampus Segmentation from Brain MRI Images, *International Journal of Computer Information Systems and Industrial Management Applications*, Vol. 2, pp. 1-10, 2010.
[20] A.P. Zijdenbos, R. Forghani, A.C. Evans, Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis, *IEEE Trans. Med. Imag.*, Vol. 21, Issue 10, pp. 1280-1291, 2002.
[21] S.K. Warfield, K.H. Zou, W.M. Wells, Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation, *IEEE Trans. Med. Imaging*, Vol. 23, Issue 7, pp. 903-921, 2004.
[22] Y. Zhang, M. Brady, S. Smith, Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm, *IEEE Trans. Med. Imag*, Vol. 20, Issue 1, pp. 45-57, 2001.
[23] J. Ashburner, and K. Friston, Spatial normalization using basis functions, In: R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, K.J. Friston, C.J. Price, S. Zeki, J. Ashburner, and W. Penny, (Eds.), *Human Brain Function*, 2nd edition, Academic Press, 2003.
[24] J.D. Klingensmith, R. Shekhar, and D.G. Vince, Evaluation of threedimensional segmentation algorithms for the identification of luminal and medial–adventitial borders in intravascular ultrasound images, *IEEE Trans. Med. Imag.*, Vol.19, Issue 10, pp. 996-1011, 2000.
[25] S.L. Victor V.I. Denis, Seamless Mosaicing of Image-Based Texture Maps, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.
[26] Hofmann, J. Puzicha, J.M. Buhmann, An Optimization Approach to Un- supervised Hierarchical Texture Segmentation, *ICIP*, 1997.
[27] L. Wolf, Xiaolei Huang, Ian Martin, Dimitris Metaxas, Patch-Based Texture Edges and Segmentation, *Proceedings of the European Conference on Computer Vision (ECCV)*, May 2006.
[28] A. Baumberg, Blending Images for Texturing 3D Models, *Proceedings of the British Machine Vision Conference (BMVC)*, 2002.
[29] H. Lensch,W. Heidrich, H. Seidel. A Silhouette-Based Algorithm for Texture Registration and Stitching. *Journal of Graphical Models*, Vol. 63, No. 4, pp. 245-262, 2001.

ISBN: 978-1-61804-019-0