Evaluation of Ischemic Stroke Hybrid Segmentation in a Rat Model of Temporary Middle Cerebral Artery Occlusion using Ground Truth from Histologic and MR data

Celina Imielinska[1,2], Yinpeng Jin[3], Xin Liu[1], Joel Rosiene[5], Brad E. Zacharia[4] , Ricardo J. Komotar[4] , J Mocco[4], Michael E. Sughrue[4], Bartosz Grobelny[4], Alex Sisti[4], Josh Silverberg[4], Joyce Khandji[4], Hillary Cohen[4], Sander Connolly[4] Jr., Anthony L. D'Ambrosio[4]

[1]Department of Biomedical Informatics, Columbia University, New York, NY
[2]Department of Computer Science, Columbia University, New York, NY
[3]Department of Biomedical Engineering, Columbia University, New York, NY
[4]Department of Neurological Surgery, Columbia University, New York, NY
[5]Department of Mathematics and Computer Science, Eastern Connecticut State University, CT

1. ABSTRACT

A segmentation method that quantifies cerebral infarct using rat data with ischemic stroke is evaluated using ground truth from histologic and MR data. To demonstrate alternative approach to rapid quantification of cerebral infarct volumes using histologic stained slices that requires scarifying animal life, a study with MR acquire volumetric rat data is proposed where ground truth is obtained by manual delineations by experts and automated segmentation is assessed for accuracy. A framework for evaluation of segmentation is used that provides more detailed accuracy measurements than mere cerebral infarct volume. Our preliminary experiment shows that ground truth derived from MRI data is at least as good as the one obtained from the histologic slices for evaluating segmentation algorithms for accuracy. Therefore we can develop and evaluate automated segmentation methods for rapid quantification of stroke without the necessitating animal sacrifice.

Keywords: magnetic resonance, segmentation, quality, neurological

1. INTRODUCTION

1.1 Background

In vivo models of focal cerebral ischemia have been developed to investigate stroke therapy. These models necessitate rapid quantification of cerebral infarct volumes using vital stains with tetrazolium salts to delineate the extent of neuronal death. 2,3,5-Triphenyltetrazolium chloride (TTC) is the current gold standard used to stain cerebral infarct volume [1]. TTC, however, is extremely time intensive and requires animal sacrifice. To further streamline our protocol and preserve animal life, we sought a more efficient method of quantifying infarct size in our rat model of focal cerebral ischemia.

1.2 Purpose

The purpose of this study is to demonstrate that we can evaluate automated segmentation methods for accuracy using ground truth derived from manually delineated, stroke regions from volumetric MR data of rats with ischemic stroke. The manual segmentations were done by several experts repeating the same task three times and using the histologic stained data, that were available in this perspective study, to guide the manual segmentation. We must note that the experts were trained technicians who dealt with low-resolution rat brain data and histologic stained slices obtained from a rat that did not exactly correspond to the slices in the volumetric MR data.

To evaluate our segmentation algorithms, we use a methodology for evaluating image segmentation algorithms [2], that takes under consideration three factors: precision (reproducibility), accuracy (agreement with truth, validity), and efficiency (time taken). To assess precision, we need to choose a figure of merit, repeat segmentation considering all sources of variation, and determine variations in figure of merit via statistical analysis. It is impossible usually to establish true segmentation. Hence, to assess accuracy, we need to choose a surrogate of true segmentation and proceed as for precision. To assess efficiency, both the computational and the user time required for algorithm and operator training and for algorithm execution should be measured and analyzed. Precision, accuracy, and efficiency are interdependent. It is difficult to improve one factor without affecting others. Segmentation methods must be compared based on all three factors. The weight given to each factor depends on application. In our experiment, that was based on retrospective data, we focus on evaluation of accuracy only using two type of ground truth: from histologic and MR data.

2. THE METHODOLOGY

2.1 Data acquisition.

The volumetric retrospective MR data was acquired from 8 rats with ischemic stroke and in addition to histologic slices stained with 2,3,5-Triphenyltetrazolium chloride (TTC) to delineate the extent of neuronal death and display cerebral infarct regions. It should be noted that there is no strict correspondence between the MR slices and hostologic slices, as demonstrated in Fig.1, due to different slice thickness.
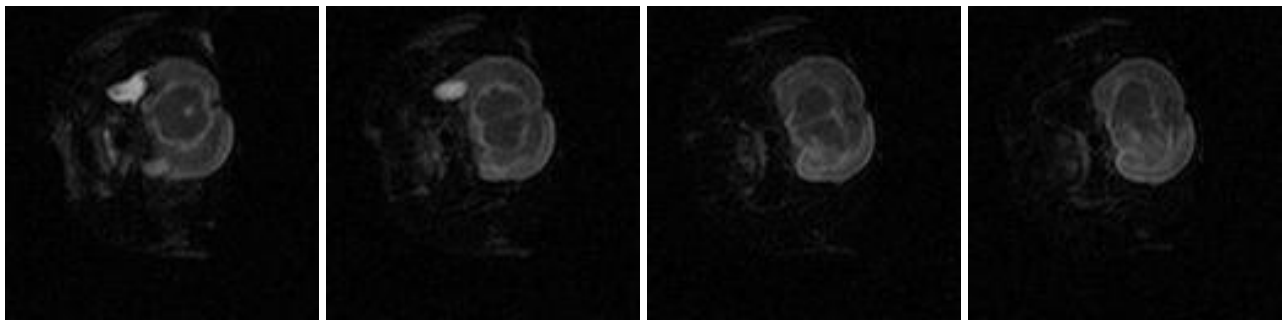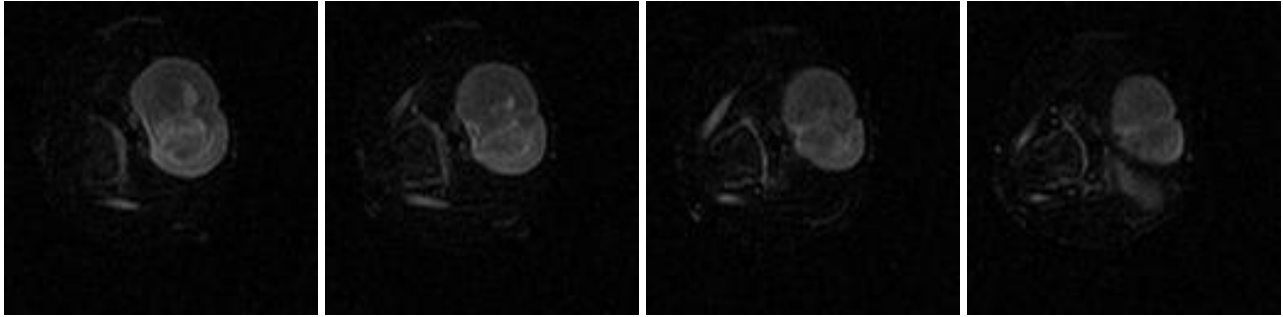
(a)

(b)

Fig.1 Data obtained from a rat with ischemic stroke: (a) MR slices in the sub-volume enclosing the ischemic stroke region, (b) histologic stained slices.

2.2 Generating Surrogate of Ground Truth

We used our framework for evaluation of image segmentation algorithms, and we tested the segmentation methods for accuracy using delineation ground truth derived from the histologic and MR data. Three experts performed the manual tracing of the stroke regions for each rat, following a strict protocol, and repeating it three times. A surrogate of "true delineation" was derived both for histologic, Fig2(a) and MR data, Fig.2 (c), respectively. Because of the low quality of the MR data, the histologic ground truth, Fig.2 (b), was used by the experts to guide manual delineation of stroke regions in the MR images, as shown in Fig.2 (d). The multiple delineations, nine for each rat, represent intra- and –inter expert variability. From the nine delineation a surrogate of ground truth

for the stroke region is derived and presented as a fuzzy set, where each pixel (in a slice), or voxel (in a volume) has assigned value between 0 and 1 that corresponds to the relative number of experts that labeled it as "inside the stroke region". We call such a representation a fuzzy object.

The following ground truth delineated from one dataset using both digitized histologic data and corresponding MR images:

      (a)   (b)   (c)   (d)
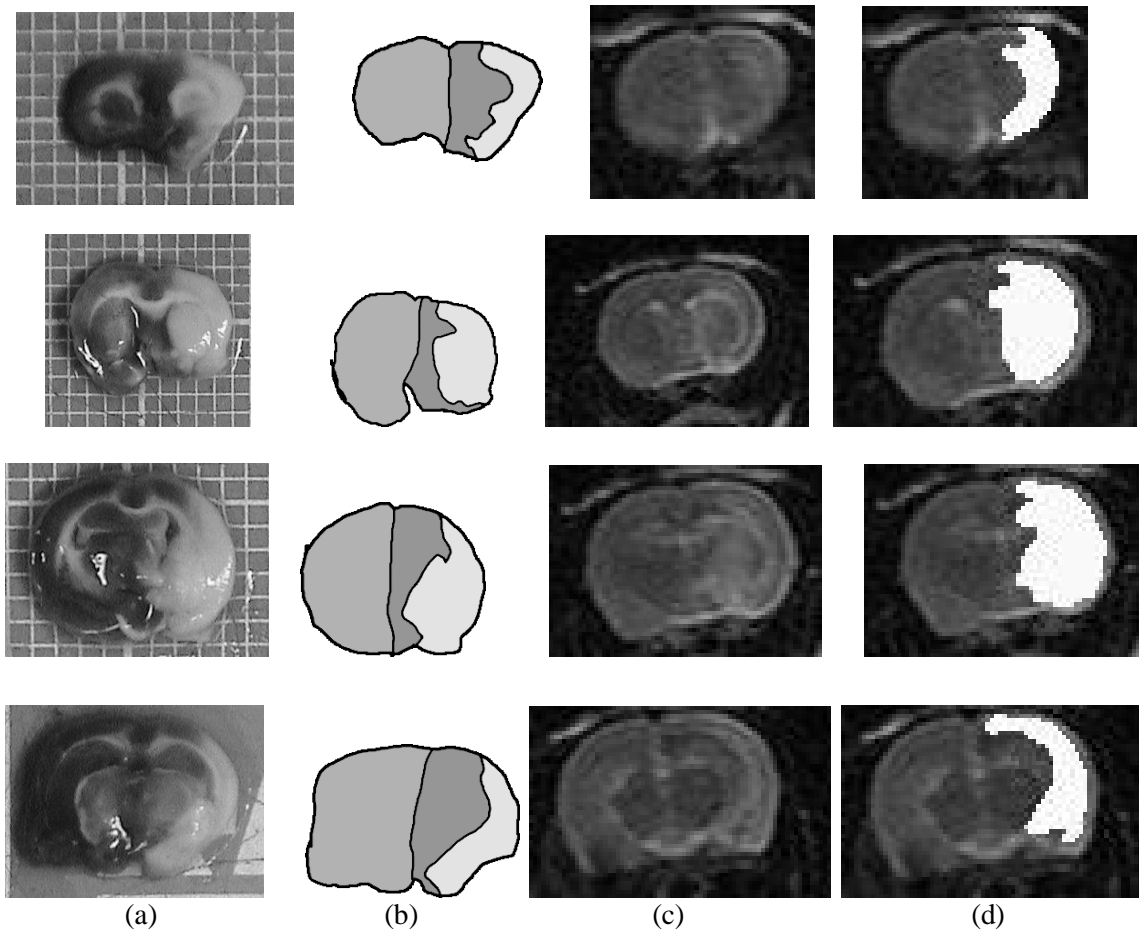
Fig.2 Histologic stained slices and MRI data a rat with ischemic stroke: (a) Histologic data, (b) ground truth delineated from histologic data, (c) MR data, (d) ground truth derived from MR data.

2.3 Assessment of accuracy.
To measure accuracy of a segmentation method M against a "true" delineation, we compute three parameters: True Positive Volume Fraction (TPVF), False Positive Volume Fraction FPVF) , and False Negative Volume Fraction (FNVF) [2], as depicted in Fig.3.
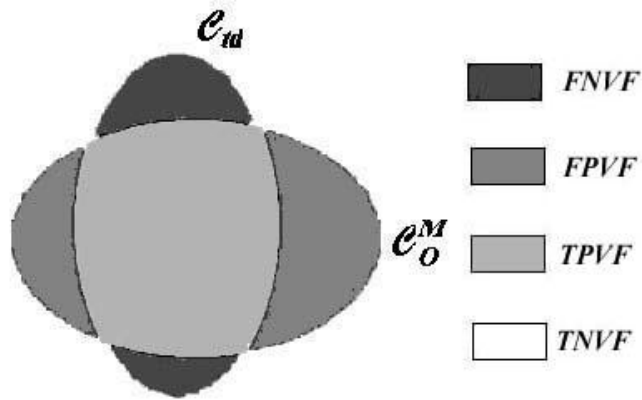
Fig.3 A geometric illustration of the three accuracy factors for delineation.

Using notation from [2], Let $\mathcal{C}_{td}$ be the "true" delineation for the stroke region of a rata data. For any $\mathcal{C}^M$ representing a segmentation obtained by using a segmentation method M, the following measures are defined to characterize the accuracy of segmentation method M for delineation of stroke region.

$$\text{False Negative Volume Fraction: } FNVF_M^d = \frac{|\mathcal{C}_{td} - \mathcal{C}_M|}{\mathcal{C}_{td}} \quad . \tag{1}$$

$$\text{False Positive Volume Fraction: } FPVF_M^d = \frac{|\mathcal{C}_M - \mathcal{C}_{td}|}{|\mathcal{C}_{td}|} \quad . \tag{2}$$

$$\text{True Positive Volume Fraction: } TPVF_M^d = \frac{|\mathcal{C}_M \bigcap \mathcal{C}_{td}|}{|\mathcal{C}_{td}|} \quad . \tag{3}$$

2.4 Preprocessing

Because of poor quality of the MR rat brain data, segmentation of the ischemic stroke is preceded first by segmentation of the brain-only regions from the MR data, followed by the estimation of symmetry and computation of highlighted stroke regions about axis of symmetry using a version Relative Difference Map (RDM) method [8].

Estimation of symmetry
The technique identifies the centroid of the anatomical region of interest (ROI), in our case, the brain-only region, Fig.4 (b) in each slice of the input MRI data, Fig.4 (a) . After the centroid, depicted in the color map in Fig.4 (c), is computed, the axis of reflection symmetry, Fig.4 (d) is derived as described in [3].
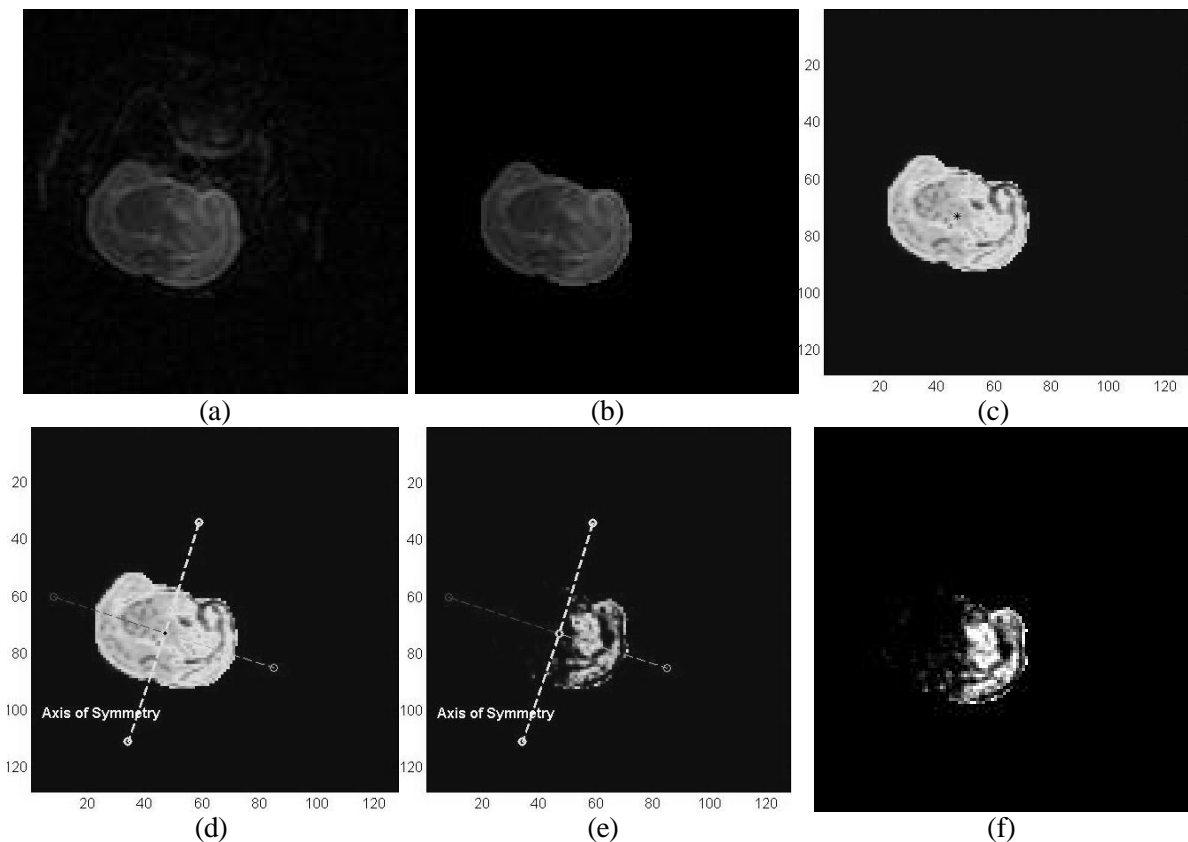
Fig.4 Preprocessing of rat stroke data using a variation of the RDM method: (a) input MR slice with stroke region, (b) brain-only region, (c) color map with centroid point, (d) axis of symmetry passing through the centroid, (e) RDM color map that highlights intensities of asymmetric regions, with respect of the axis of symmetry, (f) output image with highlighted stroke region.

RelativeDifference Map

Subject to a line of symmetry, we use the Relative Difference Map (RDM) technique [8-9] to highlight asymmetric areas in the stroke region. We fix a pair of symmetric windows, being ordered sets of points, of size M by M, about the axis of symmetry and compute set of relative differences with the associated statistics (mean and variance) while scanning simultaneously both hemispheres. We highlight regions, retain the original pixel intensities, that show relative differences (in the stroke region) and suppress regions, setting the pixel values to 0, that express symmetric similarity, as shown in the color RDM map in Fig3.(e) and corresponding gray scale preprocessed image in Fig3.(f). Such pre-segmented data is ready for segmentation with our hybrid segmentation method.

2.5 Segmentation

Two state-of-the-art image segmentation techniques were tested in this research for automatic quantification of stroke. A homogeneity based levelset methods, as proposed in [3], improves the traditional gradient based levelset segmentation by incorporates an adaptive speed term and prevents the "leakage" at weak edges. The hybrid segmentation method [4] begins with a fuzzy connectedness segmentation [5] to generate the homogeneity operators, which provide the

classification criteria for Voronoi Diagram segmentation [6]. The classification results were further processed using a volume-preserved diffusion [7] to generate the delineation boundaries.

3. RESULTS

To evaluate a segmentation algorithm, we compare the segmented stroke region with ground truth derived from the multiple delineations of the stroke region in the MR data,. Three experts were used who repeated the delineation of each rat's stroke region three times, yielding nine hand segmentation for each slice in the brain MR volume that enclose the stroke regions Fig.5 (c)-(k). Each manual delineation produces a binary mask, and a fuzzy object representing surrogate of true delineation is derived from all nine manual segmentations, as illustrated in Fig.5 for one single slice in the MR volumetric data.
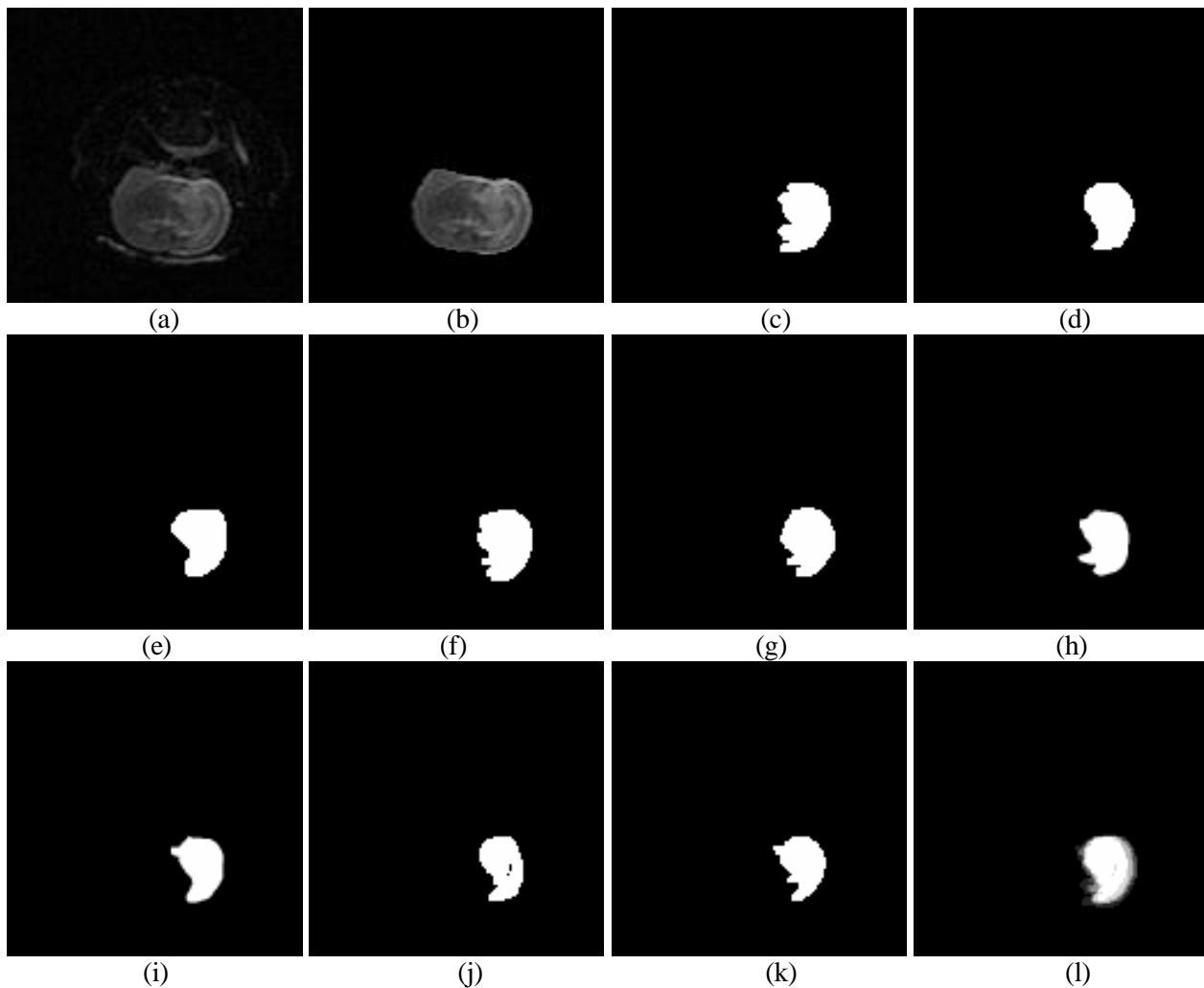
Fig.5 Surrogate of ground truth derived from nine manual delineations: (a) input MR slice, (b) brain-only region, (c)-(k) nine manual delineations of the stroke region performed by three experts and repeated three times, (l) fuzzy object representing surrogate of ground truth derived from (c)-(k).

Having surrogate of ground truth derived for each rat data from multiple delineation by the experts, we compare automatically segmented stroke region that was obtained from our hybrid segmentation method. In Fig.6, we illustrate steps to process of a single slice containing a stroke region: Fig.6.(a) input MR data; Fig.6 (b) brain-only MR image; Fig.6 (c) RDM-preprocessed highlighted stroke region, (d) stroke region segmented with the automated hybrid segmentation method.

(a)   (b)

(c)   (d)
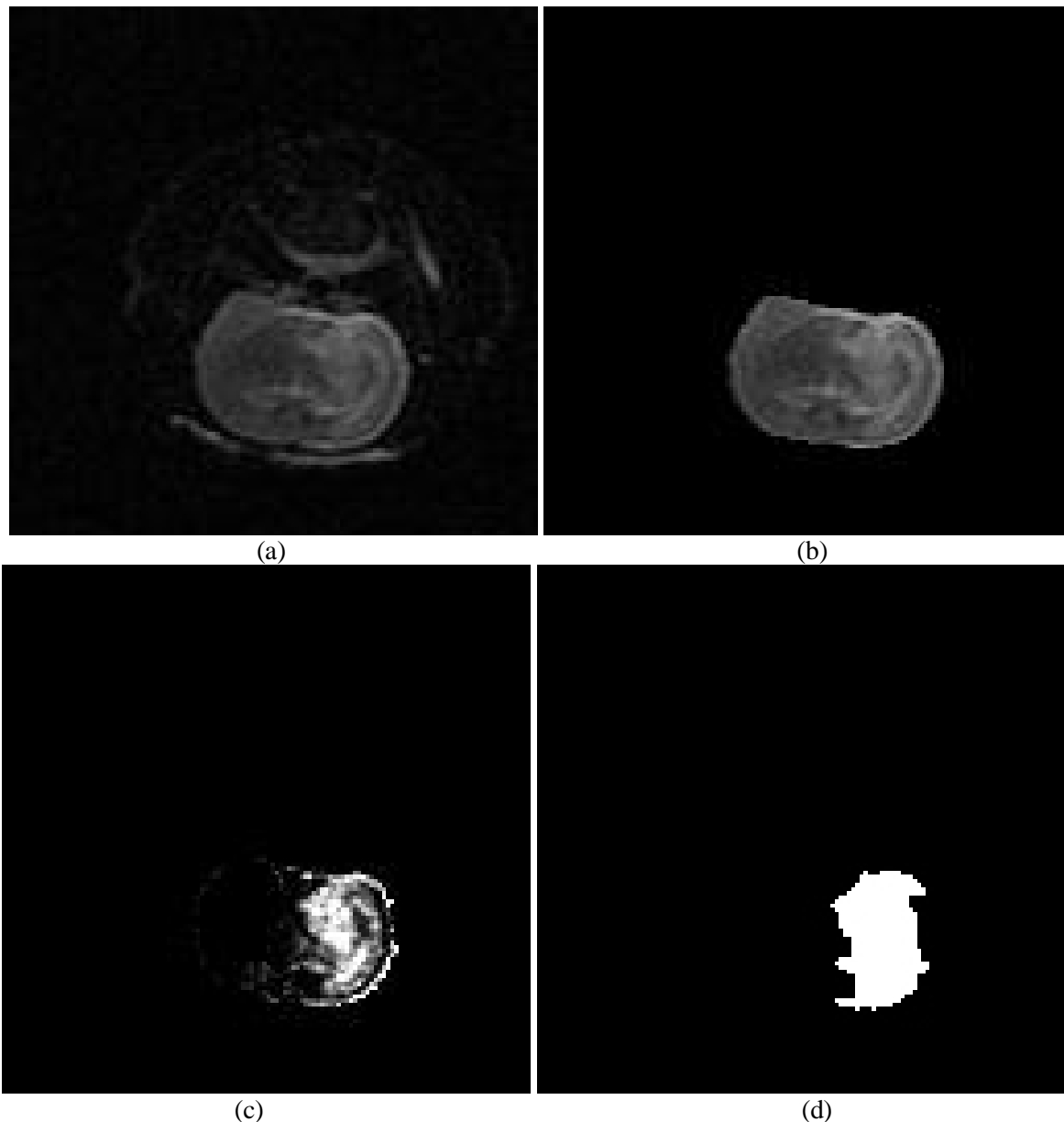
Fig.6 Segmentation of stroke region demonstrated on a single slice: (a) input MR slice with stroke region, (b) brain-only region, (c) preprocessing with the RDM method: highlighted stroke region, (d) segmented region using an automated hybrid method.

The accuracy of the segmented stroke region from the slice depicted in Fig.6(a)-(b) has been assessed using three parameters described in the subsection 2.3: FNVF, FPVF and TPVF, and is presented below.

Accuracy Measurement:

Ground Truth Area:	486.361
Segmenation Area:	554
Area Differency:	13.9072%
False Negative Volume Fraction:	0.113206
False Positive Volume Fraction:	0.252278
True Positive Volume Fraction:	0.886794

We would like to point to the fact that the "experts" used in this project were trained technicians who showed noticeable intra-operator discrepancy. This definitely affected the quality of the computed surrogate ground truth data. This prompted us to look in the future into an assessment of the quality of hand-segmenters who are not experts. We could improve the computation of the surrogate ground truth by eliminating (or suppressing) hand delineations that are visibly (statistically) sub-standard and that would contribute to an overall error of accuracy measurement of the segmentation method. This will be the topic addressed in the follow-up paper.

CONCLUSIONS

Our preliminary experiment, based on retrospective rat data, shows that surrogate for ground truth derived from MRI data is at least as good as the one obtained from the histologic slices for evaluating segmentation algorithms for accuracy. Moreover, through the histologic stained data can only capture absolute volume of the stroke region. The histologic stained slices are not aligned and can't be compared directly to automatically segmented 3D stroke region obtained from the MR images. Only absolute volume between the segmented and hitologically estimated stroke regions can be compared, that is a weak assessment of accuracy. On the other hand, having a surrogate of ground truth derived from hand-segmented delineations, we can use TPFV, FPVF and FNVF parameters to evaluate accuracy of the automated segmentation, a much stronger measurement of validity of the method. Therefore we can develop and evaluate automated segmentation methods for rapid quantification of stroke without the necessitating animal sacrifice.

REFERENCES

[1]   Bederson JB, Pitts LH, Germano SM, Nishimura MC, Davis RL, Bartkowski HM., " Evaluation of 2,3,5-triphenyltetrazolium chloride as a stain for detection and quantification of experimental cerebral infarction in rats.", Stroke. 1986 Nov-Dec;17(6):1304-8.

[2]   J.K. Udupa, V.R. Leblanc, H. Schmidt, C. Imielinska, K.P. Saha, G.J. Grevera, Y. Zhuge, P. Molholt, L. Currie, Y. Jin, "A Methodology for Evaluating Image Segmentation Algorithms", *SPIE* Medical Imaging, San Diego, 2002.

[3]   Y. Jin, A. Laine, and C. Imielinska, "An Adaptive Speed term based on Homogeneity for Level-set Segmentation" SPIE Conference on Medical Imaging, San Diego CA.,vol. 4684 (1), pp. 383-390, 2002.

[4]   C. Imielinska, D. Metaxas, J. Udupa, Y. Jin, and T. Chen, "Hybrid Segmentation of Anatomical Data." Proceedings of The Fourth International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI 2001), Utrecht Netherlands, pp. 1058-1066, 2001.

[5]   J. Udupa and S. Samarasekera, "Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation." *Graphical Models and Image Processing*, vol. 58 (3), pp. 246-261, 1996.

[6]   C. Imielinska, M. Downes, and W. Yuan, "Semi-automated color Segmentation of Anatomical Tissue." *Computerized Medical Imaging and Graphics*, vol. 24, pp. 173-180, 2000.

[7]   J. Sethian, *Level set methods and fast marching methods*, 1999.

[8]   C. Imielinska, X. Liu, M. Sughrue, E. Hagiwara, E.S. Connolly, A. D'Ambrosio, "Objective Quantification of Perfusion-Weighted Tomography in the Setting of Acute Aneurysmal Subarachnoid Hemorrhage", *Computer Assisted Radiology and Surgery,* pp. 43-43, June 2004.

[9]   Imielinska, X. Liu, J. Rosiene, M. Sughrue, R.J. Komotar, J. Mocco, E.R. Ransom, A. Lignelli, B. E. Zacharia, E.S. Connoly, A. D'Ambrosio, "Towards Objective Quantification of Perfusion-Weighted Computed Tomography in the Setting of Subarachnoid Hemorrhage: Quantification of Symmetry and Automated Delineation of Vascular Territories", accepted to *Journal of Academic Radiology* Special Issue on Medical Imaging, 2005.