Ground truth for evaluation of ischemic stroke hybrid segmentation in a rat model of temporary middle cerebral artery occlusion — Source link [↗]

Celina Imielinska, Joel A. Rosiene, Yinpeng Jin, Xin Liu ...+12 more authors

Institutions: Columbia University, University of Pennsylvania

Related papers:

- Evaluation of ischemic stroke hybrid segmentation in a rat model of temporary middle cerebral artery occlusion using ground truth from histologic and MR data

- Segmenting Hemorrhagic and Ischemic Infarct Simultaneously From Follow-Up Non-Contrast CT Images in Patients With Acute Ischemic Stroke

- Generative Model-Based Ischemic Stroke Lesion Segmentation

- Objective Ventricle Segmentation in Brain CT with Ischemic Stroke Based on Anatomical Knowledge.

- Joint Segmentation of Intracerebral Hemorrhage and Infarct from Non-Contrast CT Images of Post-treatment Acute Ischemic Stroke Patients

Share this paper: [f] [twitter] [in] [✉]

View more about this paper here: https://typeset.io/papers/ground-truth-for-evaluation-of-ischemic-stroke-hybrid-ihh1v99uks

Ground Truth for Evaluation of Ischemic Stroke Hybrid Segmentation in a Rat Model of Temporary Middle Cerebral Artery Occlusion

C. Imielinska[a,b,*], J. Rosiene[e], Y. Jin[c], X. Liu[a], J. Udupa[f],
B. Zacharia[d], R. Komotar[d], J. Mocco[d], M. Sughrue[d], B. Grobelny[d],
A. Sisti[d], J. Silverberg[d], J. Khandji[d], H. Cohen[d], E.S. Connolly Jr.[d], A. D'Ambrosio[d],

[a]*Department. of Biomedical Informatics, Columbia University, New York, NY, USA*
[b]*Department of Computer Science, Columbia University, New York, NY, USA*
[c]*Department. of Biomedical Engineering,, Columbia University, New York, NY, USA*
[d]*Department of Neurological Surgery, Columbia University, New York, NY, USA*
[e]*Eastern Connecticut State Univ., Willimantic, CT, USA*
[f]*Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA*

Abstract. In vivo rodent models of focal cerebral ischemia have been developed to investigate stroke therapy. Typically these models require rapid quantification of cerebral infarct volumes using vital stains with tetrazolium salts to delineate the extent of neuronal death. To avoid animal sacrifice, we sought a study with MR acquired volumetric rata data where surrogate of ground truth is obtained by repeated manual delineation by experts, and an automated hybrid segmentation is evaluated for accuracy. We propose a rating system for the expert delineations that captures intra- and inter-expert discrepancy. Our preliminary results show that surrogate ground truth derived from MR data is at least as good as the one derived from histologic stained slices. Hence animal sacrifice is not necessary to evaluate ischemic stroke automated segmentation in a rat model of temporary middle cerebral artery occlusion.

Keywords: Segmentation, Evaluation, Ground Truth, Stroke, Magnetic Resonance, Quality

1. Introduction

Previously, we have reported a framework for evaluating of automated image segmentation [2] that would provide a ranking of experts generating delineations for surrogate ground truth. The current report presents recent progress, illustrated by processing of MR data acquired from rats with experimentally produced ischemic stroke. To achieve this, we have utilized expert assessment and delineation of histological cerebral infarct to generate surrogate ground truth, "true" delineation of MR ischemic

---

* Corresponding author. Email address: ci42@columbia.edu

regions. While the use of human experts introduces a potential source of variation, human analysis of histologic stained slices currently represents the accepted gold standard for quantification of infarct volume, and seems a reasonable starting point for development of a ground truth from MR data.

2. Material and methods

2.1 The Rat Model of Focal Reperfused Cerebral Ischemia

All procedures performed on animals were approved by the Columbia University IACUC. 8 adult male Wistar rats weighing 225-300g were anesthetized with halothane in a mix of 70% nitrous oxide/30% oxygen and animal core temperature was maintained at 37 C during surgery, occlusion, and for 90 minutes post-reperfusion [1]. A midline neck incision was created to expose the right carotid sheath under the operating. MCAo was accomplished by advancing a 25 mm 4-0 nylon suture with a 0.38mm outer diameter cast rubber tip until the tip occluded the MCA. The occluding suture was removed after 120 minutes of ischemia. To confirm cerebral ischemia, transcranial measurements of cerebral blood flow (CBF) were made using laser-Doppler flowmetry, using 15-mm flexible fiberoptic Doppler probes attached to the skull over the MCA territory. Reduction of LDF readings to at least 40% of baseline was defined as adequate CBF drop-off. MRI analysis was performed on all rats on day 7 post ischemia. Animals were anesthetized and scans were performed using a 1.5T MRI and a Knee coil. T2W images were stored in DICOM format. Immediately following MRI imaging, rats were sacrificed and brains were removed intact and placed in a brain matrix for 2-mm coronal sectioning. Sections were immersed in 2% TTC at 4 C. Intact, viable brain tissue stains red, and infarcted brain was identified as the area of unstained tissue.

2.2 Development of Expert Based Ground Truth

Three experts performed the manual tracing of the stroke regions for each rat, following a strict protocol, and repeating it three times, using the histologic stained slices to guide delineation in the MR images, Fig.1. The collection of MR data $\{I_{\underline{r}} : \underline{r} \in ZxZxZ\}$ has been hand segmented K times by K/3 individual experts, leading to a classification that partitions the samples into two sets "abnormal" (stroke) $\overline{N}_k = \{\underline{r} : M_k \text{ points selected as abnormal by expert k}\}$ and "normal" (background). The K non-normal sets are combined to generate a "fuzzy" set. A fuzzy set is one in which membership to the set is assigned a probability, while a crisp set is one where the membership function is binary, true or false. The following *crisp sets* are easy to interpret, $\chi_1 = \bigcap_k \overline{N}_k$, the set of abnormal points common to all of the experts, and the set

$\chi_\varepsilon = \bigcup_k \overline{N}_k, \varepsilon > 0$, which is the *support of the set of abnormal points*, at least one of the experts classified the point as abnormal. To assign a probability to the inclusion of a pixel into the set of abnormal points is to count the number of experts which classified

the point as abnormal, $\Pr(\underline{r}) = \sum_k \left(1, \text{ if } \underline{r} \in \overline{N}_k, 0 \text{ otherwise}\right)/K$, hence to define a fuzzy set with the associated membership function $\Pr(r)$. The α-level characteristic function χ_α is the set of pixels that are considered abnormal with a probability greater than α. We can select probability-based threshold α to define the fuzzy set at a selected value of α, corresponding to the crisp set of α-level characteristic function, with all the points associated with probability greater or equal to α. The following figures of merit are calculated. Setting $C_{td} = \chi_\varepsilon$, where C_{td} corresponds to "true" delineation [2], and letting C_k be the segmentation for expert k we have a figure of merit $F_0 = \left(\sum_r p(\underline{r})\right)\left(K/(K-1)\right)/|C_{td}|$, $\underline{r} \in C_{td} - C_k$ that reduces to the False Negative Volume Fraction in [2], $FNVF_k^d$ if all other experts have classified the region as stroke. A similar modification of False Positive Volume Fraction, $FPVF_k^d$ has a metric which reduces to the second figure of merit $F_1 = \left(\sum_r (1 - p(\underline{r}))\right)/|C_{td}| = \left(|C_k - C_{td}| - \sum_r p(\underline{r})\right)/|C_{td}|$, $\underline{r} \in C_k - C_{td}$, reduces to the False Positive Volume Fraction in [2], $FPVF_k^d$ if *no other* expert have classified similar points. The True Positive Volume Fraction in [2], $TPVF_k^d = |C_k \cap \chi_e|/|\chi_e|$ which reflects the area where the expert k is in agreement with the support of abnormal points, provides us with the third figure of merit, the average $TPVF_k^d$, $\underline{r} \in C_m \cap \chi_\varepsilon, \varepsilon > 0$, $F_2 = \left(\sum_r p(r)\right)/|\chi_\varepsilon|$. These numbers can be combined into a single value by weighting them, $F = w_0 F_0 + w_1 F_1 + w_2 F_2, w_0 + w_1 + w_2 = 1$, with the weights based on the cost of under or over segmentation. In the paper, we set $F = w_0 F_0 + w_1 F_1 + w_2 F_2, w_0 = 0, w_1 = 0, w_2 = 1$.
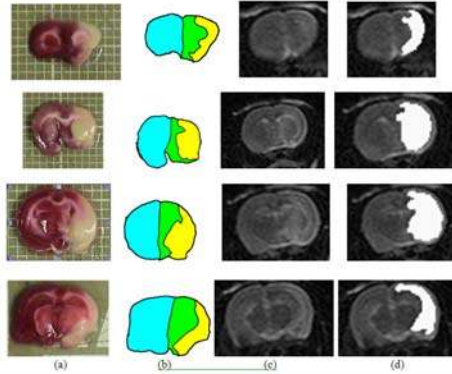
Fig.1 Delineation of stroke area in MR data using histologic slices for guidance: (a) histologic data, (b) delineated stroke regions in histologic data, (c) MR data, (d) delineated stroke regions.

2.3. Design of α-Optimal Classifiers

The above fuzzy set representing the region of a cerebral infarct can be used to create an optimal classier based on a desired feature. In our work we consider the following:

the magnitude only classifier. Using only magnitude we can design the optimal classifier relative to a give crisp set of α-level characteristic function.

2.4 Segmentation

Two state-of-the-art image segmentation technique is tested in this research for automatic quantification of stroke. The hybrid segmentation method [4] begins with a fuzzy connectedness segmentation [5] that provides the classification criteria for Voronoi Diagram segmentation [6], followed by a volume-preserved diffusion [7] to generate the delineation boundaries.
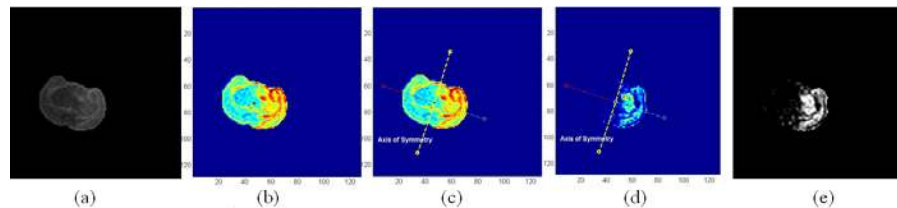
2.4.1 Preprocessing

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

Estimation of Symmetry

We convert planar symmetry detection problem into the global minimum of the reflectional symmetry specification. Once the unwrapped image of the object is obtained, the bilateral symmetry identification turns into the rectangular window comparison and correlation and an axis of symmetry is generated Fig.2(b)-(c).

Relative Difference Map

Subject to a line of symmetry, we calculate the left to right RDM [8]. 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 differences with the associated statistics while scanning simultaneously both hemispheres. We highlight regions that show relative differences and suppress regions that express symmetric similarity Fig.2(d)-(e).


Fig. 2 Preprocessing of rat stroke data: (a) input brain image, (b) color map with centroid, (c) axis of symmetry passing through centroid, (d) RDM, (e) output image with highlighted stroke region.

3. Results

To measure accuracy of a segmentation method against a "true" delineation, a surrogate of ground truth, we compute three parameters: TPVF, FPVF, and FNVF as in

[2]. Having surrogate of ground truth derived for each rat data, Fig.3, we compare automatically segmented ischemic region using our segmentation method. The segmentation of a single slice as illustrated in Fig.4 starts with RDM-preprocessed highlighted stroke region, Fig.4 (b), followed by the hybrid segmentation, Fig.4 (c). The accuracy of the segmented stroke region, Fig.4. (c), is presented below, using the surrogate of ground truth in Fig.4.(d).
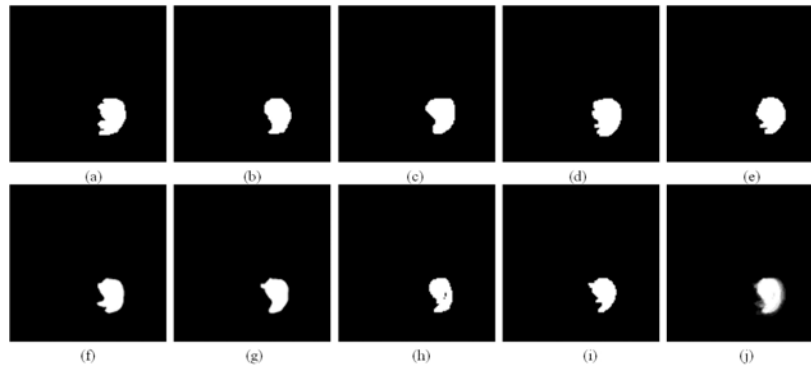
Fig.3 Surrogate ground truth derived from hand delineations for a single slice, from Fig.2(a): (a)-(i) nine S1-S9 hand-segmentations generated by experts, (j) fuzzy object representing surrogate ground truth.
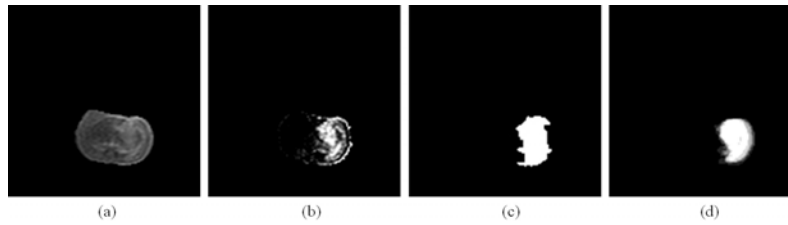
Fig.4 Segmentation of stroke region in a single slice: (a) input brain region, (b) RDM-highligheted stroke region, (c) hybrid segmentation of stroke region, (d) surrogate of ground truth.

Accuracy Measurement:

Ground Truth Area (GTA):	486.361
Segmentation Area (SA):	554
Area Difference (AD):	13.9072%
FNVF:	0.113206
FPVF:	0.252278
TPVF:	0.886794

We have also assessed the agreement with the surrogate of ground truth of each of the nine delineations presented in Fig.3.(a)-(i).

	S1	S2	S3	S4	S5	S6	S7	S8	S9
GTA:	476.3								
SA:	534	475	514	577	666	459	419	390	396
AD:	12.1%	0.29%	7.91%	21.13%	39.8%	6.39%	12.0%	18.1%	16.8%
FNVF:	0.064	0.10	0.066	0.057	0.0	0.127	0.162	0.202	0.22

FPVF:	0.186	0.098	0.146	0.257	0.398	0.091	0.041	0.021	0.052
TPVF:	0.935	0.899	0.933	0.954	1	0.872	0.838	0.798	0.779
Figure of merit F:	0.533	0.669	0.620	0.509	0.558	0.656	0.803	0.852	0.865

We can see a discrepancy between the hand delineations that affected the resulting surrogate of ground truth, and the outcome of assessing accuracy of the segmentation. The "experts" used in this project were trained technicians who showed noticeable intra- and inter- operator discrepancy.

4. Conclusion

Our preliminary experiment shows that surrogate for ground truth derived from MRI data is at least as good as the one obtained from the histologic slices. Because the histologic stained data can only capture absolute volume of the stroke region, only the absolute volume between the segmented and histologically estimated infarct regions can be compared, which is a weak assessment of accuracy. Having a surrogate of ground truth derived from hand delineations, we can use FNVF, FPVF, TPVF parameters to evaluate accuracy of segmentation, a much stronger measurement of validity of the method. In our retrospective rat data of focal cerebral ischemia, the coronal sectioning of brain tissue was not obtained with the same thickness as the MR slices. Having that, we could correlate better the histologic stained slices with corresponding MR slices, improve the resulting intra- and inter- expert discrepancy of the delineations, and in turn, provide better quality of the surrogate ground truth. We demonstrated that we can develop and evaluate automated segmentation methods for rapid quantification of cerebral infarct volumes without the necessitating animal sacrifice. This latter goal is particularly desirable in long term studies of functional outcome which generally require months of animal survival to truly demonstrate neurotherapeutic efficacy.

References

[1] Bederson JB, Pitts LH, et. al., " 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, et. al.,"A Methodology for Evaluating Image Segmentation Algorithms", SPIE Medical Imaging, San Diego, 2002.
[3] J. Rosiene, X. Liu, C. Imielinska "Ray Casting approach for boundary extraction and Fourier shape descriptor characterization", SPIE 2005, *Electronic Imaging.*
[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. et al., "Objective Quantification of Perfusion-Weighted Tomography in the Setting of Acute Aneurysmal Subarachnoid Hemorrhage", *CARS,* pp. 43-43, 2004.