Towards Automatic MRI Volumetry for Treatment Selection in Acute Ischemic Stroke Patients

Stefan Bauer[1,2], Pascal P. Gratz[2], Jan Gralla[2], Mauricio Reyes[1] and Roland Wiest[2]

***Abstract*— In many tertiary clinical care centers, decision-making and treatment selection for acute ischemic stroke is based on magnetic resonance imaging (MRI). The "mismatch" concept aims to segregate the infarct core from potentially salvageable hypo-perfused tissue, the so-called penumbra that is determined from a combination of different MRI modalities. Recent studies have challenged the current concept of tissue at risk stratification targeted to identify the best treatment options for every individual patient. Here, we propose a novel, more elaborate image analysis approach that is based on supervised classification methods to automatically segment and predict the extent of the tissue compartments of interest (healthy, infarct, penumbra regions). The output of the algorithm is a label image including quantitative volumetric information about each tissue compartment. The approach has been evaluated on an image dataset of 10 stroke patients and it compared favorably to currently available tools.**

I. INTRODUCTION

Stroke is the 2nd most frequent cause of death and a major cause of disability in industrial countries. In patients who survive, stroke is frequently associated with high socio-economic costs due to persistent disability [1]. An acute ischemic stroke is characterized by sudden impairment of brain function due to a reduced oxygen supply caused by an immediate vessel occlusion or embolism to the brain. Advanced neuroimaging techniques are recommended for a quick, reliable diagnosis and stratification for therapy [2]. Magnetic resonance imaging (MRI) identifies the infarct core by diffusion-weighted imaging and hypo-perfused, yet vital surrounding tissue that can be potentially rescued (i.e. the "penumbra"). Treatment options for acute stroke aim at an early and sustained revascularization of occluded vessels, preferably by thrombectomy or intravenous thrombolysis in specialized centers [3]. In clinical practice, imaging analysis and treatment decision is performed visually by the neuroradiologist and stroke neurologist. Automatic methods for medical image analysis could further aid in making fast decisions based on more detailed information otherwise neglected by visual analysis [4]. Stroke MR protocols include a wealth of information that is composed of structural (non-enhanced and enhanced T1-weighted, T2-weighted, FLAIR) and functional (perfusion- and diffusion-weighted) image datasets and vessel imaging (magnetic resonance angiography (MRA)).

These datasets could be further explored towards proper identification of patients that are most likely to benefit from endovascular stroke treatment. Imaging might also aid in the identification of patients who would benefit from a further extension of the time-window after stroke where patients are candidates for endovascular therapies. Currently, the most crucial imaging feature for this decision is the volumetric mismatch between the penumbra that can potentially be saved and the irreversibly damaged infarct core. Several studies indicated that the larger this mismatch, the more likely the patient is going to have a favorable prognosis [5]. Recently, these data have been challenged by negative multi-centric randomized trials (IMS III [6], MR-Rescue [7]).

Computer-assisted image analysis could help to quantify this mismatch, however as mentioned in a recent review paper [4], there are only few approaches available in the literature that explore this issue in a convincing way. Many works in the medical image analysis focus either on the segmentation of the infarct only (e.g. [8]) or the hypo-perfused region only (e.g. [9]). The few approaches that consider both regions simultaneously usually have limited accuracy because they rely on overly simplistic classification models (e.g. [10]).

An automated image analysis tool to identify candidates for acute stroke treatment has recently been proposed (RAPID [11]). This approach relies on diffusion- and perfusion-weighted MR images to quantify the mismatch. For identification of the ischemic core, the Apparent Diffusion Coefficient, a quantitative measure derived from diffusion images, is thresholded at values $< 600 \cdot 10^{-6}\,\mathrm{mm}^2/\mathrm{s}$. To identify the penumbra region, the $\mathrm{T}_{\max}$ map derived from the DSC-perfusion images is thresholded at values > 6 seconds. Some additional morphological constraints are applied to suppress outliers. Despite encouraging results reported by the authors, clinical trials have lead to controversial outcomes about the clinical benefit of this approach [7], [5].

We hypothesize that the currently recommended, somewhat simplified two-parameter approach to identify infarct core and penumbra could be further improved, if modern concepts from machine learning are engaged instead of simple thresholding methods. Therefore, we propose to employ a supervised classification approach for performing a multi-parametric segmentation from a number of different MRI modalities with training based on manually labeled samples.

[1,2]S. Bauer is with the Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland and with the Department of Diagnostic and Interventional Neuroradiology, Inselspital, Bern University Hospital, Switzerland stefan.bauer at istb.unibe.ch

[1]M. Reyes is with the Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland

[2]P. Gratz, J. Gralla and R. Wiest are with the Department of Diagnostic and Interventional Neuroradiology, Inselspital, Bern University Hospital, Switzerland

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

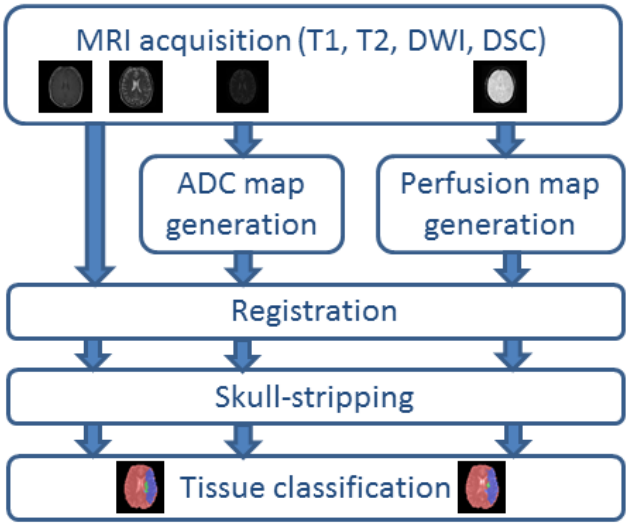

Fig. 1. Illustration of the workflow for the complete segmentation / prediction pipeline.

II. METHODS

The segmentation is based on structural and functional MR images. T1-weighted images with contrast enhancement, T2-weighted images, diffusion-weighted images and dynamic susceptibiliy contrast (DSC) perfusion-weighted images have been acquired from acute ischemic stroke patients before and after therapy. From the diffusion-weighted images, the apparent diffusion coefficient (ADC) maps were extracted and from the perfusion-weighted images, four standard perfusion maps were computed using the PMA toolbox [1]. The perfusion maps comprised cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT) and the peak time (T_{max}). All modalities (T1contrast, T2, ADC, CBF, CBV, MTT, Tmax) pre- and post-treatment were rigidly registered to the pre-treatment T1contrast image of the same patient and automatically skull-stripped [12]. The 7 pre-treatment MRI modalities (T1contrast, T2, ADC, CBF, CBV, MTT, Tmax) were used as an input for the segmentation algorithm. The complete workflow is illustrated in figure 1.

The proposed segmentation method is conceptually inspired by a work on multi-modal brain tumor segmentation [13], [14]. The segmentation task is cast as an energy minimization problem in a conditional random field context, with the energy to be minimized being expressed as

$$E = \sum_i V(y_i, \mathbf{x}_i) + \sum_{ij} W(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) \qquad (1)$$

where the first term in equation (1) corresponds to the voxel-wise singleton potentials and the second term corresponds to the pairwise potentials, modeling voxel-to-voxel interactions. $\mathbf{x}$ is a voxel-wise feature vector and y the final segmentation label. The singleton potentials are computed by a decision forest classifier [15]. A decision forest is a supervised classifier that makes use of training data for computing a probabilistic output label for every voxel based on a certain feature vector. We use a 283-dimensional feature vector $\mathbf{x}$ as an input for the classifier, consisting of the voxel-wise intensities, and multi-scale local texture, gradient, symmetry and position descriptors of each modality. These singleton potentials are computed according to equation (2), with $p(\tilde{y}_i|\mathbf{x_i})$ being the output probability from the classifier and δ is the Kronecker-δ function.

$$V(y_i, \mathbf{x}_i) = p(\tilde{y}_i|\mathbf{x_i}) \cdot (1 - \delta(\tilde{y}_i, y_i)) \qquad (2)$$

The second term in equation (1) corresponds to the pairwise potentials, introducing a spatial regularization in order to suppress noisy outputs caused by outliers. It is computed according to equation (3), where $w_s(i,j)$ is a weighting function that depends on the voxel spacing of the image in each dimension. The term $(1-\delta(y_i, y_j))$ penalizes different labels of neighboring voxels and the degree of neighborhood smoothing is regulated by the difference of the feature vectors in $\exp\left(\frac{|\mathbf{x}_i - \mathbf{x}_j|}{2 \cdot \bar{x}}\right)$

$$W(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) = w_s(i,j) \cdot (1 - \delta(y_i, y_j)) \cdot \exp\left(\frac{|\mathbf{x}_i - \mathbf{x}_j|}{2 \cdot \bar{x}}\right) \qquad (3)$$

Optimization of the energy function in equation (1) is achieved with fast primal-dual strategies from [16].

We propose two different approaches using the same basic method, but with different training sets:

- *approach A* is a classical segmentation approach, where the training is based on manual segmentations of infarct core and penumbra on the pre-treatment images (same parameters for the manual segmentation as in [11])
- *approach B* aims for prediction instead of segmentation. The training is also based on manual segmentation, but only the penumbra is defined on the pre-treatment images, whereas the infarct core is the real infarct, which is defined on the T2-weighted images from the follow-up exam after 3 months. Please note that the 3 month follow-up images are only needed for generating the training data, for running and testing the algorithm only the pre-treatment images are needed. This ensures that the approach can be used for decision-making before treatment.

III. RESULTS

We used the proposed segmentation approach for automatic MRI volumetry of 10 stroke patients, employing a leave-one-out cross-validation concept for training and testing. The ground-truth for the infarct core and penumbra regions, against which the proposed methods were compared, was defined by an expert. The infarct core was defined as the largest connected component with hyperintensity on the T2-weighted 3-month follow up images. The penumbra was defined as the connected component with Tmax>6s on the baseline image. The results of the proposed approaches A

[1] http://asist.umin.jp/index-e.htm

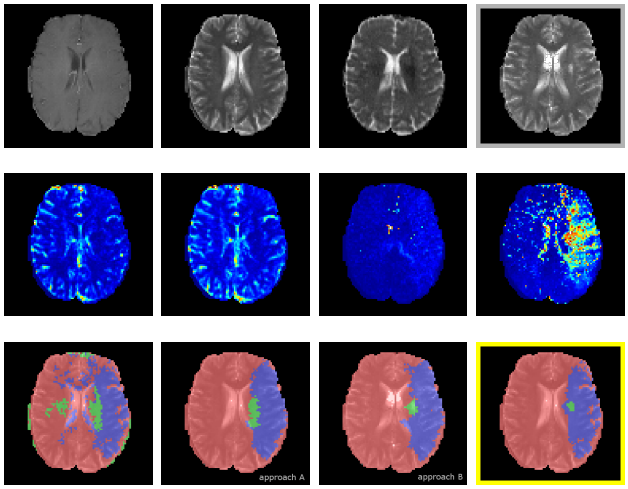

Fig. 2. Example of the results on patient W: The first row shows the T1contrast image, T2 image and ADC map before treatment on an axial slice from left to right. The image on the very right (gray frame) shows the T2 follow-up image after 3 months. The second row shows the perfusion maps before treatment (CBF, CBV, MTT, Tmax from left to right). The third row shows the segmentations, where red corresponds to healthy tissue, green is the infarct core and blue is the penumbra region. From left to right the labels resulting from RAPID and the labels resulting from our proposed method A and B are shown. The manually labeled groundtruth is depicted on the very right (yellow frame).

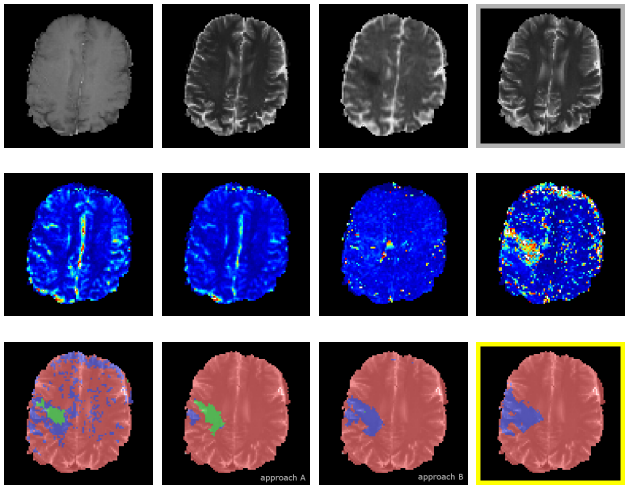

Fig. 3. Example of the results on patient E: The first row shows the T1contrast image, T2 image and ADC map before treatment on an axial slice from left to right. The image on the very right (gray frame) shows the T2 follow-up image after 3 months. The second row shows the perfusion maps before treatment (CBF, CBV, MTT, Tmax from left to right). The third row shows the segmentations, where red corresponds to healthy tissue, green is the infarct core and blue is the penumbra region. From left to right the labels resulting from RAPID and the labels resulting from our proposed method A and B are shown. The manually labeled groundtruth is depicted on the very right (yellow frame).

and B were compared with our own implementation of the state of the art RAPID approach [11].

Figure 2 shows the results for a patient who has an infarct core and a penumbra region. The first row depicts the structural images before treatment (T1contrast, T2, ADC map from left to right) and the T2 image after treatment (very right). The second row shows the calculated perfusion maps before treatment (CBF, CBV, MTT, Tmax). And the last row visualizes the segmentation results. From left to right, the RAPID segmentation, our approach A and approach B and finally the groundtruth are shown. It can be seen that the RAPID approach looks quite noisy with many outliers due to the simple thresholding procedure. The segmentations of approach A and approach B look more convincing and it can be seen that approach B does a better job at predicting the real infarct core.

Figure 3 shows the results for a patient who has no infarct core at the follow-up examination. However, both RAPID and also approach A seem to detect false positives outside the true infarct. Also here, the predictive approach B seems to do a better job because only penumbra, no infarct region is detected.

For quantitative evaluation of the results, we used the established Dice score [17], which can range between 0 and 1, where 0 indicates no overlap and 1 indicates a perfect overlap. Additionally, we also calculated the absolute volume error of the segmented regions in ml. This metric is clinically most relevant because the volumetric infarct/penumbra mismatch will support the treatment decision. In table I, we compared our own implementation of the RAPID approach [11], our proposed method A and our proposed method B with the manually defined groundtruth. Results show consistently better results for the proposed methods in terms of Dice score and also in terms of absolute volume error. The absolute volume error is best for the predictive approach B, but even approach A exhibits less than half of the volume error of the RAPID method. The relatively low values for the Dice score are caused by two factors: First, we are dealing with a prediction problem, not a simple segmentation problem because the ground-truth is defined on the 3-month follow-up images, not the pre-treatment images. And second, the low scores of the infarct region in particular are caused by the fact that this region is relatively small and the Dice score is biased by the size of the region.

TABLE I

DICE SCORES AND ABSOLUTE VOLUME ERROR RELATIVE TO THE MANUALLY DEFINED GROUNDTRUTH FOR 10 PATIENTS (MEAN±STDDEV). RESULTS ARE SHOWN FOR THE PROPOSED METHODS AND FOR OUR IMPLEMENTATION OF THE RAPID APPROACH AS A COMPARISON.

	infarct core	penumbra
Dice score RAPID	0.10±0.11	0.53±0.24
Dice score our approach A	0.21±0.20	0.57±0.24
Dice score our approach B	0.13±0.19	0.61±0.22
abs volume error RAPID	43±21 ml	143±123 ml
abs volume error our approach A	25±33 ml	51±72 ml
abs volume error our approach B	23±35 ml	44±72 ml

Computation time is crucial for decision making in stroke because “time is brain”. After pre-processing (registration and calculation of the perfusion maps), computation time for both approach A and approach B is slightly longer

than the computation time of RAPID, but always less than one minute. Pre-processing takes several minutes, mostly dominated by the time needed for computing the perfusion maps. This is currently done offline because the perfusion maps are regarded as an external input and the method is not integrated into the clinical workflow yet.

IV. DISCUSSION & CONCLUSIONS

Here, we propose a novel method for image-based segmentation and prediction of the infarct core and penumbra region in stroke patients. Integrating all the information that is available within routine MRI datasets offers advantages for treatment selection in individual patients. Our first, preliminary experience from the evaluation on 10 patients shows promising results and may serve as a proof of concept.

Both the newly proposed approaches provide meaningful results overall. The accuracy regarding the penumbra is slightly higher for approach B. However the accuracy regarding the infarct core is inconclusive: from visual inspection, the predictive approach B seems to have advantages over approach A, this is also confirmed quantitatively by the volume error. The Dice score in contrast, is lower for the infarct core using approach B. It remains to be investigated on a larger dataset, whether the segmentation-oriented approach A or the prediction-oriented approach B offers more potential. In fact, the quality of the prediction also depends on the way of treatment and the treatment success, therefore it might be useful to carry out the analysis on different patient subgroups individually.

The potential and clinical benefit of the proposed method can only be evaluated in a prospective clinical study, where the classifier has been trained on a large dataset. For accurate prediction, clinically meaningful information, such as the stroke topography, severity, the vascular supply of the hypo-perfused tissue and other prognostic factors should be identified and incorporated as modeling parameters.

REFERENCES

[1] G. A. Donnan, M. Fisher, M. Macleod, and S. M. Davis, "Stroke." *Lancet*, vol. 371, no. 9624, pp. 1612–23, May 2008. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/18468545

[2] L. M. Nentwich and W. Veloz, "Neuroimaging in acute stroke." *Emergency medicine clinics of North America*, vol. 30, no. 3, pp. 659–80, Aug. 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22974643

[3] J. Gralla, C. Brekenfeld, M. Arnold, and G. Schroth, "Acute stroke: present and future of catheter-based interventions." *Herz*, vol. 33, no. 7, pp. 507–17, Nov. 2008. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19066747

[4] I. Rekik, S. Allassonnière, T. K. Carpenter, and J. M. Wardlaw, "Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal," *NeuroImage: Clinical*, vol. 1, no. 1, pp. 164–178, Jan. 2012. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S2213158212000228

[5] M. Wintermark, P. Sanelli, and C. C. Meltzer, "Stroke imaging: diffusion, perfusion, but no more confusion!" *AJNR. American journal of neuroradiology*, vol. 34, no. 11, p. 2053, Nov. 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23907248

[6] J. P. Broderick, Y. Y. Palesch, A. M. Demchuk, S. D. Yeatts, P. Khatri, M. D. Hill, E. C. Jauch, T. G. Jovin, B. Yan, F. L. Silver, R. von Kummer, C. a. Molina, B. M. Demaerschalk, R. Budzik, W. M. Clark, O. O. Zaidat, T. W. Malisch, M. Goyal, W. J. Schonewille, M. Mazighi, S. T. Engelter, C. Anderson, J. Spilker, J. Carrozzella, K. J. Ryckborst, L. S. Janis, R. H. Martin, L. D. Foster, and T. a. Tomsick, "Endovascular therapy after intravenous t-PA versus t-PA alone for stroke." *The New England journal of medicine*, vol. 368, no. 10, pp. 893–903, Mar. 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23390923

[7] C. S. Kidwell, R. Jahan, J. Gornbein, J. R. Alger, V. Nenov, Z. Ajani, L. Feng, B. C. Meyer, S. Olson, L. H. Schwamm, A. J. Yoo, R. S. Marshall, P. M. Meyers, D. R. Yavagal, M. Wintermark, J. Guzy, S. Starkman, and J. L. Saver, "A trial of imaging selection and endovascular treatment for ischemic stroke." *The New England journal of medicine*, vol. 368, no. 10, pp. 914–23, Mar. 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23394476

[8] J. Mitra, P. Bourgeat, J. Fripp, S. Ghose, S. Rose, O. Salvado, A. Connelly, B. Campbell, S. Palmer, G. Sharma, S. Christensen, and L. Carey, "Classification Forests and Markov Random Field to Segment Chronic Ischemic Infarcts from Multimodal MRI," in *Multimodal Brain Image Analysis*. Springer LNCS, 2013, pp. 107–118.

[9] K. Mouridsen, K. Nagenthiraja, K. Y. Jónsdóttir, L. R. Ribe, A. B. Neumann, N. Hjort, and L. Ostergaard, "Acute stroke: automatic perfusion lesion outlining using level sets." *Radiology*, vol. 269, no. 2, pp. 404–12, Nov. 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23687176

[10] O. Wu, W. J. Koroshetz, L. Ostergaard, F. S. Buonanno, W. a. Copen, R. G. Gonzalez, G. Rordorf, B. R. Rosen, L. H. Schwamm, R. M. Weisskoff, and a. G. Sorensen, "Predicting Tissue Outcome in Acute Human Cerebral Ischemia Using Combined Diffusion- and Perfusion-Weighted MR Imaging," *Stroke*, vol. 32, no. 4, pp. 933–942, Apr. 2001. [Online]. Available: http://stroke.ahajournals.org/cgi/doi/10.1161/01.STR.32.4.933

[11] M. Straka, G. W. Albers, and R. Bammer, "Real-time diffusion-perfusion mismatch analysis in acute stroke." *Journal of magnetic resonance imaging : JMRI*, vol. 32, no. 5, pp. 1024–37, Nov. 2010. [Online]. Available: www.ncbi.nlm.nih.gov/pubmed/21031505

[12] S. Bauer, T. Fejes, and M. Reyes, "A Skull-Stripping Filter for ITK," *Insight Journal*, 2012. [Online]. Available: http://hdl.handle.net/10380/3353

[13] S. Bauer, L.-P. Nolte, and M. Reyes, "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization." in *MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, ser. Lecture Notes in Computer Science, G. Fichtinger, A. Martel, and T. Peters, Eds., vol. 14, no. Pt 3. Toronto: Springer Berlin Heidelberg, Jan. 2011, pp. 354–61. [Online]. Available: http://www.springerlink.com/index/10.1007/978-3-642-23626-6 http://www.ncbi.nlm.nih.gov/pubmed/22003719

[14] S. Bauer, T. Fejes, J. Slotboom, R. Wiest, L.-P. Nolte, and M. Reyes, "Segmentation of Brain Tumor Images Based on Integrated Hierarchical Classification and Regularization," in *Miccai Brats Workshop*, 2012.

[15] A. Criminisi and J. Shotton, Eds., *Decision Forests for Computer Vision and Medical Image Analysis*. London: Springer London, 2013. [Online]. Available: http://link.springer.com/10.1007/978-1-4471-4929-3

[16] N. Komodakis, G. Tziritas, and N. Paragios, "Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies," *Computer Vision and Image Understanding*, vol. 112, no. 1, pp. 14–29, Oct. 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1077314208000982

[17] W. R. Crum, O. Camara, and D. L. G. Hill, "Generalized overlap measures for evaluation and validation in medical image analysis." *IEEE transactions on medical imaging*, vol. 25, no. 11, pp. 1451–61, Nov. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/17117774