Ischemic Stroke Lesion Segmentation

www.isles-challenge.org

Proceedings

5th October 2015

Munich, Germany

UNIVERSITÄT ZU LÜBECK

Technische Universität München

UNIVERSITÄT BERN

MICCAI 15 MUNICH

Preface

Stroke is the second most frequent cause of death and a major cause of disability in industrial countries. In patients who survive, stroke is generally associated with high socioeconomic costs due to persistent disability. Its most frequent manifestation is the ischemic stroke, whose diagnosis often involves the acquisition of brain magnetic resonance (MR) scans to assess the stroke lesion's presence, location, extent, evolution and other factors. An automated method to locate, segment and quantify the lesion area would support clinicians and researchers alike, rendering their findings more robust and reproducible.

New methods for stroke segmentation are regularly proposed. But, more often than desirable, it is difficult to compare their fitness, as the reported results are obtained on private datasets. Challenges aim to overcome these shortcomings by providing (1) a public dataset that reflects the diversity of the problem and (2) a platform for a fair and direct comparison of methods with suitable evaluation measures. Thus, the scientific progress is promoted.

With ISLES, we provide such a challenge covering ischemic stroke lesion segmentation in multi-spectral MRI data. The task is backed by a well established clinical and research motivation and a large number of already existing methods. Each team may participate in either one or both of two sub-tasks:

SISS Automatic segmentation of ischemic stroke lesion volumes from multi-spectral MRI sequences acquired in the sub-acute stroke development stage.

SPES Automatic segmentation of acute ischemic stroke lesion volumes from multi-spectral MRI sequences for stroke outcome prediction.

The participants downloaded a set of training cases with associated expert segmentations of the stroke lesions to train and evaluate their approach, then submitted a short paper describing their method. After reviewing by the organizers, a total of 17 articles were accepted and compiled into this volume. At the day of the challenge, each teams' results as obtained on an independent test set of cases will be revealed and a ranking of methods established.

For the final ranking and more information, visit WWW.ISLES-CHALLENGE.ORG.

Oskar Maier, Universität zu Lübeck

Mauricio Reyes, University of Bern

Björn Menze, TU Munich

August 2015

Version: 2015-09-30

Organizers

Oskar Maier, Universität zu Lübeck, Germany

Mauricio Reyes, University of Bern, Switzerland

Björn Menze, TU Munich, Germany

Sponsoring Institutions

Institute of Medical Informatics, Universität zu Lübeck, Germany

Institute for Surgical Technology & Biomechanics, University of Bern, Switzerland

Computer Science, TU Munich, Germany

Version: 2015-09-30

ISLES Challenge 2015: A voxel-wise, cascaded classification approach to stroke lesion segmentation

David Robben[1,3], Daan Christiaens[1,3], Janaki Raman Rangarajan[1,3], Jaap Gelderblom[1,3], Philip Joris[1,3], Frederik Maes[1,3], and Paul Suetens[1,2,3]

[1] KU Leuven, Dept. of Electrical Engineering, ESAT/PSI, Medical Image Computing
[2] iMinds, Medical IT Department
[3] UZ Leuven, Medical Imaging Research Center
Herestraat 49 – 7003, 3000 Leuven, Belgium

Abstract. We propose a supervised method based on cascaded extremely randomised forests for lesion segmentation, and evaluate the pipeline in the MICCAI Ischemic Stroke Lesion Segmentation (ISLES) challenge.

1 Introduction

In ischemic stroke, reduced blood flow to part of the brain results in localised tissue damage and eventual necrosis. Automated localisation and segmentation of the stroke lesion in patients is of great interest to clinicians and researchers alike, enabling them to differentiate potentially salvageable and permanently damaged tissue, identify effective treatments, and follow progression of the ischemic lesion [5]. The MICCAI Ischemic Stroke Lesion Segmentation (ISLES) challenge aims to evaluate and compare state-of-the-art methods, by providing two multi-modal MRI datasets for sub-acute ischemic stroke lesion segmentation (SISS) and for acute stroke outcome and penumbra estimation (SPES).

In this paper, we propose a supervised method based on cascaded extremely randomised forest classifiers for stroke lesion segmentation, and describe a single pipeline to be used for both datasets. After nested cross-validation on the training data, we obtained an average Dice score of 57 % for the SISS data and 82 % for the SPES dataset, which is on par with other contestants.

2 Method

2.1 Preprocessing

At first, the non-parametric images in both datasets were corrected for RF inhomogeneity. We estimated the bias field on the T1w-images using FSL FAST [2], using a 3-tissue model and a bias field smoothing filter of 40 mm full-width half maximum. The elevated smoothing parameter (default is 20 mm) was chosen to improve robustness to the pathology. The estimated bias field was subsequently applied to correct all T1w- and T2w-images, as well as the Flair and DWI images in the SISS dataset. The ADC images and the perfusion measures in the SPES dataset were not corrected, as these images are already normalised or assumed to be in physical units.

Secondly, cross-subject histogram normalisation was done for each dataset and each modality. To this end, we used a linear intensity rescaling based on two percentile intensities of the histogram. These were heuristically determined based on the histogram profile of a given modality across all subjects of each dataset. For SISS we used 20 % and 99 % for T1, T2, and DWI, and 30 % and 90 % for Flair. For SPES we used 30 % and 90 % for T1, 20 % and 99 % for T2, 20 % and 90 % for DWI (ADC), and 20 % and 50 % for TTP. No intensity normalisation was applied to Tmax, CBF, and CBV.

Additionally, we wish to include spatial features in the classifier as well. Therefore, we registered all subjects T1w-images to the MNI152 template using a 12 degrees of freedom affine transformation and normalised mutual information, as implemented in FSL FLIRT [2]. The resulting transformation matrices are converted to (affine) deformation fields which provide, for each voxel in native space, the corresponding coordinate in MNI space. As such, no image interpolation is needed and the subsequent classifier training can be done in native space.

2.2 Classifier

We decided to use a voxel-wise classification approach for both segmentation tasks. That is, we build a classifier that, given a set of features of a voxel, estimates the probability that this voxel is part of a lesion. To increase computational efficiency and spatial consistency, we use a cascaded approach. First, the to-be-classified voxel is given to a classifier that uses a limited set of features. If this classifier decides with very high probability that the voxel is non-lesion, then this probability is the final answer. Else, the voxel is given to the second classifier that uses a large set of features. Then, the voxels which were not classified as non-lesion with very high probability, are given to a third classifier. This last classifier uses the same features as the second classifier and additionally the earlier computed probabilities of that voxel and its neighbouring voxels.

We use extremely randomised trees [1] as a base classifier. This classifier builds an ensemble of decision trees, but by randomising the selection of cut-point in the decision tree nodes, its training is significantly faster than the training of random forests while achieving comparable accuracy. We use the implementation provided by scikit-learn [4].

2.3 Features

Since the classifier is the same for both challenges, the features are constructed in a similar fashion.

For the SPES sub-challenge, the first cascade uses the T1c intensity. The second cascade uses the intensity of the T1c, T2, TTP, Tmax and DWI images smoothed with a sigma of $0 - 6$ mm. It also has for TTP and Tmax the 0.5, 0.75

and 0.9 percentiles and for DWI the 0.1 and 0.25 percentiles of its neighbourhood for varying radii (4 – 12 mm). Finally, it has the MNI-coordinates. The third cascade uses the same features as the latter and additionally it has the earlier estimated probabilities smoothed with a sigma of 0 – 8 mm and the 0.5, 0.75 and 0.9 percentiles of it neighbourhood for varying radii (4 – 8 mm).

For the SISS sub-challenge, the first cascade uses the T1 intensity. The second cascade uses the intensity of the T1, T2, Flair and and DWI images smoothed with a sigma of 0 – 8 mm. It also has for Flair and DWI the 0.5, 0.75 and 0.9 percentiles of its neighbourhood for varying radii (4 – 8 mm). Finally, it uses the MNI-coordinates. The third cascade uses the same features as the latter and additionally it has the earlier estimated probabilities smoothed with a sigma of 0 – 8 mm and the 0.5, 0.75 and 0.9 percentiles of it neighbourhood for varying radii (4 – 8 mm).

2.4 Probability threshold

After the voxel-wise classification, we have for every voxel a probability of belonging to a lesion. However, the challenge requires a binary segmentation and hence we need to threshold the resulting probabilities. Instead of using a fixed threshold for all images, we use a novel technique to find the optimal threshold.

A voxel $\boldsymbol{x}$ is part of the lesion with probability $P(\boldsymbol{x})$, as estimated by the classifier. Given that the probability estimates are correct, the Dice score obtained with threshold P_t will be:

$$Dice(P_t) = \frac{2\,|GT \cap segmentation|}{|GT| \cup |segmentation|} = \frac{2\sum_{\boldsymbol{x}} I[P(\boldsymbol{x}) > P_t]P(\boldsymbol{x})}{\sum_{\boldsymbol{x}} P(\boldsymbol{x}) + \sum_{\boldsymbol{x}} I[P(\boldsymbol{x}) > P_t]}, \qquad (1)$$

with I the indicator function. We exhaustively search for the optimal threshold.

3 Results

The performance of the proposed segmentation method is evaluated in the online submission system of the challenge, and relies on the average symmetric surface distance (ASSD), the Dice overlap coefficient, and the Hausdorff distance. Additionally, precision and recall are reported to discriminate between over- and under-segmentation respectively. The results of cross-validation on the training data are reported in Table 1. Example segmentations of median and maximum overlap are shown in Fig. 1.

4 Discussion and Conclusion

We presented a supervised method for stroke lesion segmentation, based on cascaded extremely randomised forests. The cascaded approach showed strong improvement over a single voxel-wise classifier, and allows to take neighbourhood information into account while still limiting the number of features and the

Table 1. Segmentation results on the training data, reported as average symmetric surface distance (ASSD), Dice coefficient, Hausdorff distance, precision, and recall.

	ASSD (mm)		Dice		Hausdorff (mm)		Precision		Recall	
	avg	std	avg	std	avg	std	avg	std	avg	std
SISS	9.36	13.85	0.57	0.28	53.88	34.58	0.58	0.33	0.68	0.21
SPES	2.03	1.35	0.82	0.07	44.29	27.59	0.81	0.14	0.85	0.07

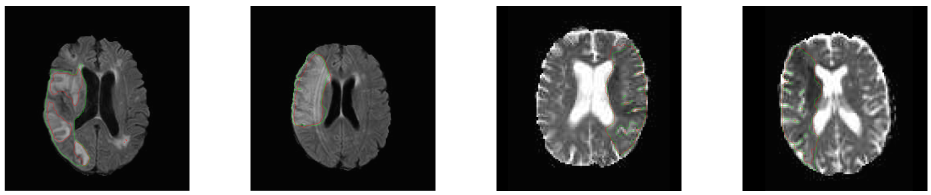

Fig. 1. Comparison between the ground-truth labels (*green*) and the predicted segmentation (*red*), shown for selected examples with median and maximum Dice coefficient. SISS Flair dataset on the left; SPES DWI (ADC) dataset on the right.

required computation time. The method works well on both datasets, although the inter-subject variability is rather large in the SISS data. Given that this is the case for other contestants as well, it would be interesting to have access to the inter-observer variability of the ground-truth segmentations.

Future work may improve upon this method by revising the histogram normalisation. A threshold-based classifier such as ours is sensitive to the intensity scaling, and the current linear approach is sub-optimal. More advanced, non-linear approaches such as Meier et al. [3] could help in this regard.

References

1. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Machine Learning 63(1), 3–42 (Mar 2006)
2. Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M.: FSL. NeuroImage 62(2), 782–790 (aug 2012)
3. Meier, D.S., Guttmann, C.R.: Time-series analysis of MRI intensity patterns in multiple sclerosis. NeuroImage 20(2), 1193–1209 (oct 2003)
4. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830 (2011)
5. Rekik, I., Allassonnière, S., Carpenter, T.K., Wardlaw, J.M.: 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 1(1), 164–178 (2012)