SUPERVISED IN-VIVO PLAQUE CHARACTERIZATION INCORPORATING CLASS LABEL UNCERTAINTY

Arna van Engelen[1], Wiro J Niessen[1,6], Stefan Klein[1], Harald C Groen[2,3,4], Hence JM Verhagen[5], Jolanda J Wentzel[2], Aad van der Lugt[3] and Marleen de Bruijne[1,7]

[1]Biomedical Imaging Group Rotterdam, Depts. of Medical Informatics & Radiology
Depts. of [2]Biomedical Engineering, [3]Radiology, [4]Nuclear Medicine and [5]Vascular Surgery
[6]Imaging Science and Technology, Fac. of Applied Sciences, Delft University of Technology
[7]Department of Computer Sciences, University of Copenhagen, Denmark
[1−5]Erasmus MC, Rotterdam, [1−6]the Netherlands

ABSTRACT

We segment atherosclerotic plaque components in *in-vivo* MRI and CT data using supervised voxelwise classification. The most reliable ground truth can be obtained from histology sections, however, it is not straightforward to use this for classifier training as the registration with *in-vivo* data often shows misalignments. Therefore, for training we incorporate uncertainty in the ground truth via "soft" labels that indicate a probability for each class. Soft labels are created by Gaussian blurring of the original "hard" segmentations, and weighted by the registration accuracy. Classification is evaluated on the relative volumes for fibrous, lipid-rich necrotic and calcified tissue. Using conventional hard labels, the differences between the ground truth and classification result per subject are -0.4±3.6% for calcification, +7.6±14.9% for fibrous and -7.2±14.5% for necrotic tissue. Using the new approach accuracy is improved: for calcification -0.6±1.6%, fibrous +3.6±16.8% and necrotic tissue -2.9±16.1%.

Index Terms— Segmentation, classification, pattern recognition, atherosclerosis, histology

1. INTRODUCTION

Rupture of an atherosclerotic plaque in the carotid artery may lead to embolisation of plaque material and/or thrombus into the intracranial circulation, causing cerebral infarction. It has been shown that plaques prone to rupture differ from stable plaques in composition [1]. *In-vivo* quantification of plaque components is therefore important for risk stratification in patients with atherosclerosis. Both MRI [2] and CT-angiography (CTA) [3, 4] are able to visualize different plaque components. Few studies have developed methods that automatically segment plaque components in *in-vivo* MRI [5, 6, 7] and CTA [3, 4]. To develop such methods using supervised pattern classification, a known ground truth is needed. This may be obtained from manual segmentations, but these may be inaccurate due to overlapping intensities between classes and inter- and intra-observer variability [5, 6]. Histology sections are considered to be more objective. It is, however, difficult to accurately align those sections with *in-vivo* scans due to tissue deformation. Registration has mostly been done by manually selecting corresponding slices prior to 2D registration [3, 5, 6], but 3D registration has also been used [4, 8] to allow rotation of the *in-vivo* image direction with respect to the histology slicing direction. Although this does allow for the correct rotation, still the vessel is often not well aligned due to large deformations caused by plaque excision and histological processing.

In this paper we combine MRI and CTA scans for plaque component classification. To allow supervised classification, 1) semi-automatic registration of *in-vivo* images with histology is performed and 2) a framework is developed in which "soft" labels are derived from original hard manual segmentations in histology. This framework takes into account the registration accuracy per slice, and uncertainty in segmentation around region boundaries in the ground truth.

2. METHODS

2.1. Data

For 13 patients that underwent carotid endarterectomy, *in-vivo* MRI (precontrast 3D-T1w, T1w, PDw, and TOF, and postcontrast 3D-T1w) and CTA scans were acquired. To facilitate registration to histology, *ex-vivo* MRI (3D-T1w) and μCT scans of the excised plaque were made. Histology sections were taken every 1-mm interval and stained with Elastica von Gieson (Merck, Germany) staining. To obtain ground truth segmentations, the vessel wall in histology was manually segmented into fibrous and lipid-rich necrotic regions. The ground truth for calcification was obtained by thresholding the μCT. After registration to histology, manual annotations of the lumen and outer vessel wall were made

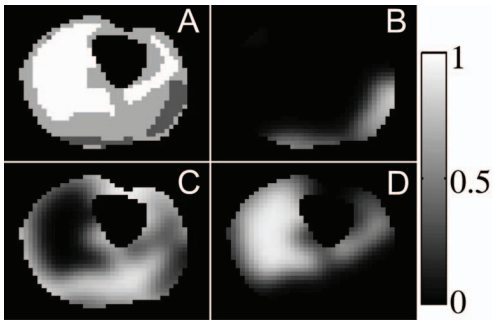


Fig. 1. Example of soft labels: A shows an original hard segmentation based on histology and μCT with calcification in dark gray, fibrous tissue in light gray and necrotic tissue in white. B shows the soft label for calcification, C for fibrous tissue and D for necrotic tissue.

on the *in-vivo* data. Based on histology quality, 11±4 image slices per subject were included (range 3-17 slices).

2.2. Registration

3D registration of *ex-vivo* MRI, μCT and CTA with histology has been described by Groen et al. [8]. We added *in-vivo* MRI to this framework by a rigid registration of all MRI scans to the postcontrast 3D-T1w scan, and rigidly registered this scan to the CTA, based on mutual information of image intensity. In addition, compared with Groen et al. the registration of histology and *in-vivo* data in this work is refined by 1) replacing isotropic scaling of CTA to match μCT with a thin-plate spline deformation and 2) deforming the *ex-vivo* MRI with a B-spline to match the *in-vivo* MRI after deformation to the histology domain, based on the combined mutual information of image intensity, and manual segmentations of the lumen and outer vessel wall. The resulting transformation was applied to the ground truth segmentations (histology and μCT). For classification all images were resampled to 0.25x0.25 mm in-plane such that their resolution is in the order of the *in-vivo* MRI and CTA.

2.3. Classifier development

All MR images were corrected for intensity inhomogeneities [9], and normalized by setting the mean intensity to 0 and the standard deviation to 100. A set of 23 features was calculated for each voxel: the intensities in the five image sequences, these images blurred with a Gaussian filter and the gradient magnitude and Laplacian of these blurred images (σ=1 mm) [10], the original CTA scan and the Euclidean distances to the lumen and outer vessel wall. For training, the distances were based on the deformed histology segmentation, and for testing on the distance to *in-vivo* contours.

To account for registration inaccuracies, we accounted for uncertainties in the ground truth segmentations by blurring them with a Gaussian filter (σ=0.87 mm). This creates soft labels that indicate a probability of belonging to each of the three components (figure 1). In addition, since the reliability of the label depends on the registration accuracy, the Dice overlap between the vessel wall segmentation in histology and MRI was calculated for each slice. The soft labels were multiplied with their corresponding Dice^n to obtain the final soft labels. In our experiments we determine the optimal value for n. This assigns a weight to the samples, such that samples close to region boundaries or from slices with a low registration accuracy contribute less to the classifier than samples with a more certain ground truth. A linear discriminant classifier was trained using the soft labels on all voxels that are within the vessel wall in both the histology segmentation, and the *in-vivo* wall segmentation, using the Matlab toolbox *prtools* [11].

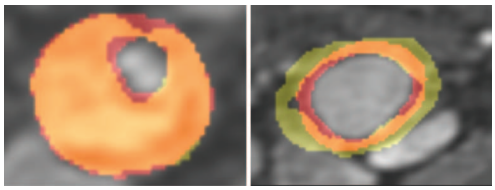


Fig. 2. Two slices after registration. In yellow the deformed histology segmentation is shown, in red the *in-vivo* segmentation, and in the orange regions they overlap. The left image has a Dice overlap of 0.93, the right image of 0.53.

2.4. Experiments

Classification was performed on all voxels within the *in-vivo* segmented vessel wall with a leave-one-subject-out approach. Results are evaluated on the difference in the relative volumes of the three plaque components between the ground truth and the classification result. Training on hard and soft labels is compared using Wilcoxon signed ranks test. Spearman rank correlations are calculated as well.

3. RESULTS

The refinement of the registration proposed in [8], by deforming the *ex-vivo* vessel wall to match the *in-vivo* vessel wall, increased the Dice overlap between the histology and *in-vivo* data from 0.57±0.18 (range 0.11-0.87) to 0.67±0.16 (range 0.22-0.94) (figure 2).

Both per slice and per subject volume estimates improved by using the soft labels. Using hard labels for training, the difference between the relative volume (%) in the result and the ground truth (deformed histology) per slice was -0.3±3.6% for calcification, +9.5±19.5% for fibrous tissue and -9.2±19.3% for necrotic tissue. The smallest differences were found when the blurred labels were multiplied with

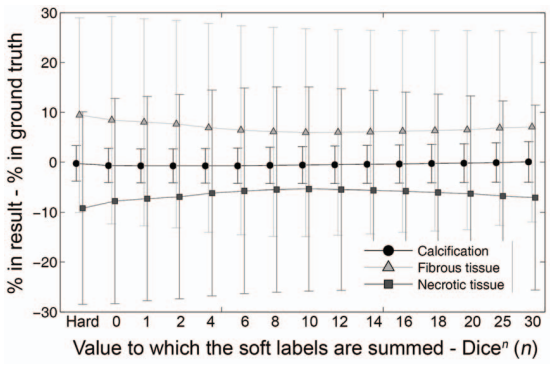


Fig. 3. The difference in the relative volume per slice (%) for each component between the ground truth and classification result, as a function of n. Here n indicates the value with which the soft labels are multiplied (Dice^n).

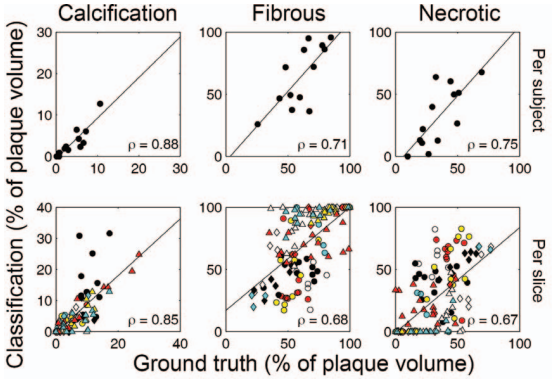


Fig. 4. Correlations between the ground truth and classification result per subject for all three components. The top row shows the results per subject and the bottom row per slice. Here each symbol denotes a different subject. Spearman rank correlation values (ρ) are given.

Dice^{10}: -0.6±3.7% for calcification, +6.0±20.8% for fibrous tissue and -5.4±20.4% for necrotic tissue (figure 3). Per subject this changed the difference in calcification from -0.4±3.6 to -0.6±1.6%, for fibrous tissue from +7.6±14.9 to +3.6±16.8% and for necrotic tissue from -7.2±14.5% to -2.9±16.1%. These differences are statistically significant ($p<0.5$), however, when comparing the absolute differences, these are not significantly different.

Using n=10 for further experiments, Spearman rank correlation coefficients (ρ) of the relative volumes did not change considerably: per slice 0.85 for calcification, 0.67 for fibrous and 0.66 for necrotic tissue for training on hard labels, and 0.85, 0.68 and 0.67 for training on the soft labels. Per subject the correlations were 0.86, 0.71 and 0.80 (hard labels) and 0.88, 0.71 and 0.75 (soft labels) (figure 4).

Classification in four example slices is shown in figure 5. Two show a good accuracy, one shows reasonable volume estimations but shows the necrotic core at a displaced location and one shows poor segmentation as the necrotic tissue is missed.

4. DISCUSSION AND CONCLUSION

We segmented different components of atherosclerotic plaques using both *in-vivo* MRI and CTA images. Our approach accounts for training a voxelwise classifier on labels obtained from registered histology in the presence of misregistrations. The use of soft labels in general improved volume estimations for fibrous and lipid-rich necrotic tissue compared to using the original labels for training. The final results show per subject 3.6±16.8% more fibrous tissue and 2.9±16.1% less lipid-rich necrotic tissue compared to the ground truth. The slight underestimation of calcification becomes larger (-0.6±1.6%), probably due to the blurring of the ground truth. The calcification spots are smaller and thinner than fibrous and necrotic regions, and therefore may gain less weight in the training phase. Although only 13 subjects are studied, Spearman rank correlations look good: 0.88 for calcification, 0.71 for fibrous and 0.75 for necrotic tissue. The fact that the volume differences do not change significantly when looking at absolute values, indicates that there is a change of bias. In the minority of cases where fibrous tissue is underestimated instead of overestimated, this underestimation becomes larger, and vice versa for necrotic tissue.

Two previous studies automatically segmented plaque components in *in-vivo* MRI and compared their results with histology. These show Pearson correlation values (r^2) of 0.83 for calcifications, 0.78 for necrotic tissue, 0.41 for loose matrix and 0.82 for fibrous tissue [5], and (r) 0.41 for calcifications, 0.75 for lipid, 0.61 for hemorrhage and 0.67 for fibrous tissue [6]. Two other studies compared segmentation results in CTA with histology [3, 4], both based on intensity thresholding. Both these studies can accurately segment calcifications, but obtain lower accuracies for fibrous and lipid-rich necrotic tissue due to overlap in Hounsfield units. We decided not to differentiate between fibrous and loose connective tissue, as they are both stable components and often occur mixed. This makes it difficult to distinguish them in histology. Hemorrhage was not included in our study as this plaque component cannot be identified on histological cross-sections with Elastica von Gieson stains.

In our training approach we give less weight to samples from slices with a low registration accuracy between histology and *in-vivo* data. This may be similar to including only samples for training that come from slices with a Dice overlap larger than 0.7 for example. Additional experiments show that although this does improve over training on hard labels of all samples from all slices, this gives less accurate results than our approach that includes soft labels. When the effects of

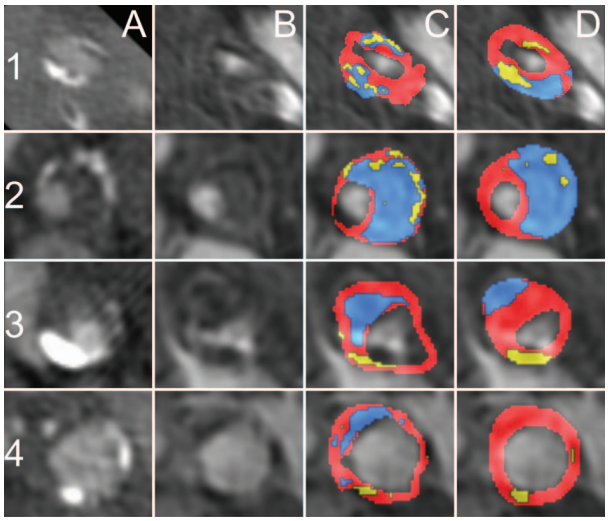


Fig. 5. Four classification examples. A: CTA, B: postcontrast 3D-T1w MRI, C: ground truth and D: resulting classification. Rows 1 and 2 show reasonably good results, row 3 shows the detection of necrotic tissue but wrongly placed and row 4 shows the failed detection of necrotic tissue (calcification=yellow, fibrous=red and necrotic=blue).

ground truth blurring and scaling the soft labels with respect to the Dice overlap value are evaluated independent of each other, both show to improve the segmentation accuracy.
Ideally, the classification accuracy would be evaluated in a voxelwise manner. We could not do this because the ground truth shows misalignment with the *in-vivo* data and can therefore not be considered accurate when looking at single voxels. We chose plaque component volumes as they may provide a clinical biomarker for prediction of plaque vulnerability. Whether our accuracy is good enough for this purpose, however, will have to be evaluated in future research. In conclusion, a classifier trained on labels that account for registration accuracies, by incorporating the uncertainty at the edges of ground truth segmentations and the registration accuracy per slice, can be used for characterization of plaque composition. In future research, the size of the Gaussian filter to blur the original segmentations may be optimized or additional factors may be taken into account for creating the soft labels. In addition, it would be interesting to study which features are most relevant, and to compare segmentations based on only MRI or CTA with the combined approach.

5. ACKNOWLEDGEMENTS

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project PARISk (grant 01C-202), and supported by the Dutch Heart Foundation. WN, SK and MdB were financially supported by the Netherlands Organisation for Scientific Research (NWO). HG was financially supported by the Interuniversity Cardiology Institute of the Netherlands. We thank Kim van Gaalen for her technical assistance.

6. REFERENCES

[1] J. M. Seeger, et al, "The relationship between carotid plaque composition, plaque morphology, and neurologic symptoms," *Journal of Surgical Research*, vol. 58, no. 3, pp. 330 – 336, 1995.

[2] T. Saam, et al, "Quantitative evaluation of carotid plaque composition by in vivo MRI," *Arterioscler Thromb Vasc Biol*, vol. 25, no. 1, pp. 234–239, 2005.

[3] T. T. de Weert, et al, "In vivo characterization and quantification of atherosclerotic carotid plaque components with multidetector computed tomography and histopathological correlation," vol. 26, no. 10, pp. 2366–2372, 2006.

[4] M. Wintermark, et al, "High-resolution CT imaging of carotid artery atherosclerotic plaques," vol. 29, no. 5, pp. 875–882, May 2008.

[5] F. Liu, et al, "Automated in vivo segmentation of carotid plaque MRI with morphology-enhanced probability maps," *Magnetic Resonance in Medicine*, vol. 55, no. 3, 2006.

[6] J.M.A. Hofman, et al, "Quantification of atherosclerotic plaque components using in vivo MRI and supervised classifiers," *Magnetic Resonance in Medicine*, vol. 55, no. 4, 2006.

[7] I. M. Adame, et al, "Automatic segmentation and plaque characterization in atherosclerotic carotid artery mr images," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 16, pp. 227–234, 2004.

[8] H. C. Groen, et al, "Three-dimensional registration of histology of human atherosclerotic carotid plaques to in-vivo imaging," *Journal of Biomechanics*, vol. 43, no. 11, pp. 2087 – 2092, 2010.

[9] J. G. Sled, et al, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, 1998.

[10] Arna van Engelen, et al, "Multi-feature-based plaque characterization in ex vivo MRI trained by registration to 3D histology," *Physics in Medicine and Biology*, vol. 57, no. 1, pp. 241–256, 1 2012.

[11] R.P.W. Duin, et al, *PRTools4.1, A Matlab Toolbox for Pattern Recognition*, Delft University of Technology, 2007.