

## Multi-feature-based plaque characterization in *ex vivo* MRI trained by registration to 3D histology

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2012 Phys. Med. Biol. 57 241

(<http://iopscience.iop.org/0031-9155/57/1/241>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 143.107.161.170

The article was downloaded on 14/08/2012 at 19:41

Please note that [terms and conditions apply](#).

# Multi-feature-based plaque characterization in *ex vivo* MRI trained by registration to 3D histology

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

<sup>1</sup> Biomedical Imaging Group Rotterdam, Departments of Medical Informatics and Radiology, Erasmus MC, Rotterdam, The Netherlands

<sup>2</sup> Imaging Science and Technology, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

<sup>3</sup> Department of Biomedical Engineering, Erasmus MC, Rotterdam, The Netherlands

<sup>4</sup> Department of Radiology, Erasmus MC, Rotterdam, The Netherlands

<sup>5</sup> Department of Nuclear Medicine, Erasmus MC, Rotterdam, The Netherlands

<sup>6</sup> Department of Vascular Surgery, Erasmus MC, Rotterdam, The Netherlands

<sup>7</sup> Department of Computer Science, University of Copenhagen, Denmark

E-mail: [a.vanengelen@erasmusmc.nl](mailto:a.vanengelen@erasmusmc.nl)

Received 14 July 2011, in final form 10 November 2011

Published 7 December 2011

Online at [stacks.iop.org/PMB/57/241](http://stacks.iop.org/PMB/57/241)

## Abstract

We present a new method for automated characterization of atherosclerotic plaque composition in *ex vivo* MRI. It uses MRI intensities as well as four other types of features: smoothed, gradient magnitude and Laplacian images at several scales, and the distances to the lumen and outer vessel wall. The ground truth for fibrous, necrotic and calcified tissue was provided by histology and  $\mu$ CT in 12 carotid plaque specimens. Semi-automatic registration of a 3D stack of histological slices and  $\mu$ CT images to MRI allowed for 3D rotations and in-plane deformations of histology. By basing voxelwise classification on different combinations of features, we evaluated their relative importance. To establish whether training by 3D registration yields different results than training by 2D registration, we determined plaque composition using (1) a 2D slice-based registration approach for three manually selected MRI and histology slices per specimen, and (2) an approach that uses only the three corresponding MRI slices from the 3D-registered volumes. Voxelwise classification accuracy was best when all features were used ( $73.3 \pm 6.3\%$ ) and was significantly better than when only original intensities and distance features were used (Friedman,  $p < 0.05$ ). Although 2D registration or selection of three slices from the 3D set slightly decreased accuracy, these differences were non-significant.

(Some figures may appear in colour only in the online journal)

## 1. Introduction

Cerebral infarction is one of the most important causes of death and the largest cause of disability in the western world. Approximately 20–30% of all cerebral infarctions can be related to atherosclerotic plaque rupture in the carotid artery (Petty *et al* 1999). Evidence suggests that plaques prone to rupture are characterized more by plaque composition than by the plaque size or the degree of stenosis (Seeger *et al* 1995, Shaalan *et al* 2004). Magnetic resonance imaging (MRI) is able to non-invasively visualize different plaque components. Both *in vivo* (Toussaint *et al* 1996, Hatsukami *et al* 2000, Yuan *et al* 2001, Mitsumori *et al* 2003, Chu *et al* 2004, Saam *et al* 2005) and *ex vivo* (Morrisett *et al* 2003, Shinnar *et al* 1999), the accuracy of MRI to distinguish plaque components has been demonstrated by comparison with histological sections. Relating plaque composition to plaque rupture or clinical events showed that a large lipid or necrotic core (Seeger *et al* 1995) covered by a thin fibrous cap (Yuan *et al* 2002) is associated with a high risk of rupture, while fibrous tissue and calcifications (Shaalan *et al* 2004) have a stabilizing effect. In addition, the presence of intraplaque haemorrhage (Takaya *et al* 2006) and plaque ulceration (Park *et al* 1998) is related to an increased number of cerebrovascular events. Most studies analyzed the relation between the presence of plaque components and outcome. However, the relative size of the components is also important for risk prediction. Segmentation and quantification of plaque components in MR images is therefore relevant. In addition, segmentations can be used to make biomechanical models of plaques (Akyildiz *et al* 2011). Automatic methods may reduce the inter- and intraobserver variability and are more practical in large studies. Our study focuses on the automatic segmentation of plaque components in *ex vivo* MRI.

Automated plaque-segmentation methods have previously been presented by several authors, both in *ex vivo* (Clarke *et al* 2003, 2006, Ronen *et al* 2007, Itskovich *et al* 2004, Karmonik *et al* 2009) and *in vivo* MRI (Adame *et al* 2004, Liu *et al* 2006, Hofman *et al* 2006). Most of these methods are based on voxelwise statistical classification. For *ex vivo* MR images of plaque specimens, segmentations of different components have been obtained with a maximum-likelihood classifier (Clarke *et al* 2006), a nearest-mean classifier (Clarke *et al* 2003), k-means clustering (Karmonik *et al* 2009) and k-means clustering in combination with an iterative approach to minimize pixel discontinuity (Itskovich *et al* 2004). The probability of misclassification was similar for a normal density-based approach and nonparametric Parzen window estimation (Ronen *et al* 2007). Similar approaches have been introduced for the analysis of *in vivo* MR images. In a comparison, a standard Bayesian classifier performed better than k-nearest neighbour, a feedforward neural network and a Bayesian classifier in combination with a Parzen classifier (Hofman *et al* 2006). In addition, fuzzy clustering followed by contour smoothing (Adame *et al* 2004) and a maximum-likelihood Bayesian classifier that besides intensities also included the distance of each voxel to the lumen and the local wall thickness (Liu *et al* 2006) have been investigated. Except for this last method (Liu *et al* 2006), all these methods are based only on original MRI voxel intensities. This paper investigates whether additional image features contain information that improves classification accuracy, validated by a histological ground truth.

To evaluate plaque segmentation, the resulting areas or volumes can be compared to those resulting from manual delineations in the MRI (Adame *et al* 2004), to volumes obtained from histology (Hofman *et al* 2006, Karmonik *et al* 2009) or to the visual grading of the American Heart Association (AHA) (Itskovich *et al* 2004). A disadvantage is that these approaches use global measures and do not take account of regional correspondence between the classification result and the ground truth. Other studies used co-registration between histology and MRI to perform the voxelwise analysis (Clarke *et al* 2003, 2006, Liu *et al* 2006, Ronen *et al*

**Table 1.** MRI scan parameters.

Sequence	Resolution (mm)	Repetition time (ms)	Echo time (ms)	Flip angle (degrees)
3D-T1w GRE	0.1×0.1×0.1	32.6–57.1	3.6–4.9	60
2D-T2w SE	0.1×0.1×0.6	2300–13 820	53.9–67.0	90
2D-PDw SE	0.1×0.1×0.6	7700–14 860	17.9–18.3	90

2007). Corresponding histology and MRI slices were then manually selected prior to in-plane registration, and voxelwise classification was evaluated for a selection of on average one to five slices per specimen. This manual selection may positively bias classification accuracy as the matched MRI and histology slices may appear more similar and have higher image quality. On the other hand, with 2D in-plane registration it is not possible to account for differences in out-of-plane slice orientation between the histological and MRI slices. Therefore, we aim to develop a classification method that is based on registration of MRI to 3D histology (Groen *et al* 2010). We expect this to provide a more accurate and objective ground truth, and it allows the inclusion of MRI slices with a lower image quality in both training and evaluation.

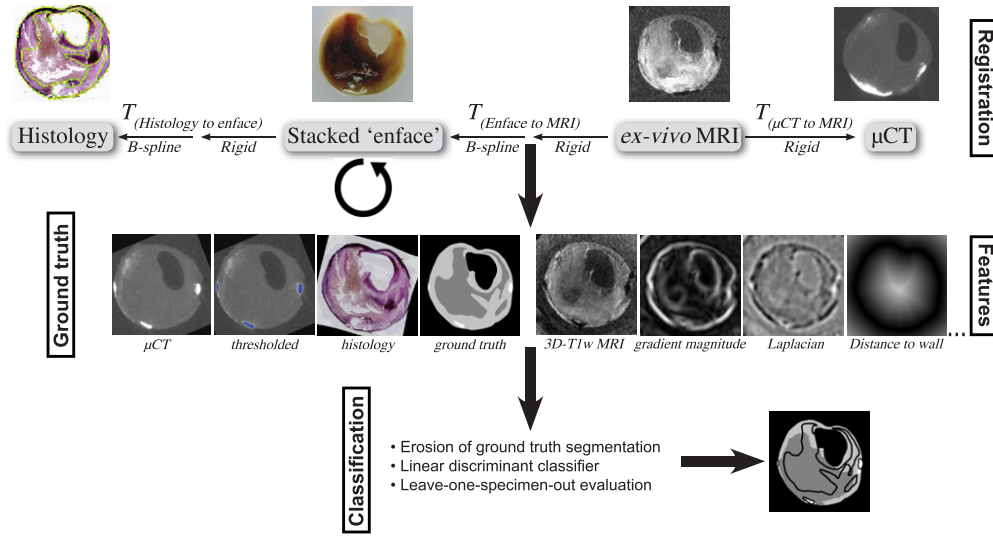
This study presents a new method for segmenting calcification, fibrous and lipid-rich necrotic tissue in MR images obtained from *ex vivo* carotid artery specimens. As well as using MRI intensity and distances to the lumen and outer vessel wall, we evaluated Gaussian smoothed intensities, gradient magnitude and Laplacian images for their contribution to classification accuracy. A 3D registration of MRI with histology was used for classifier training and evaluation. This approach was compared to the more standard method of 2D in-plane registration of selected MRI and histology slices. Preliminary results of this study have previously been published in a conference proceedings (van Engelen *et al* 2011).

## 2. Materials and methods

### 2.1. Data

Fifteen patients scheduled for carotid endarterectomy (CEA) were included in the study (nine had an ischaemic stroke, five had a transient ischaemic accident and one was asymptomatic). During the CEA procedure, the plaque was removed intact as much as possible to keep its 3D shape. The *ex vivo* specimens were scanned in a 3T clinical MRI scanner (Signa Excite, GE Healthcare, Milwaukee, USA). Three MRI contrast weightings were obtained: 3D-T1 weighted (T1w) gradient echo (GRE), 2D-T2 weighted (T2w) spin echo (SE) and 2D-proton-density weighted (PDw) SE. Scan parameters are provided in table 1. To allow accurate calcification quantification, the specimens were also scanned with a  $\mu$ CT scanner (Skyscan 1072, Skyscan, Belgium, resolution  $18 \times 18 \times 18 \mu\text{m}$ ).

After imaging, the specimens were decalcified and embedded in paraffin for histological processing. At every 1 mm interval, axial slices of  $5 \mu\text{m}$  thickness were obtained. In addition, digital photographs of the cutting plane, called ‘enface’, were taken to allow reconstruction of the 3D histology volume. An Elastica von Gieson stain was applied (EvG, Merck, Germany) to visualize cell nuclei, elastin and collagen fibres. On the high-resolution digitized histologic slices, contours of the lumen and outer vessel wall and for regions of lipid-rich necrotic tissue were drawn manually. Calcified areas were obtained from the  $\mu$ CT by intensity thresholding. The  $\mu$ CT scanner was not calibrated and values could not be related to standard Hounsfield units. Therefore, one threshold value was chosen that gave visually satisfying results for all specimens. As the resolution and contrast for calcification are both quite high in  $\mu$ CT,



**Figure 1.** Flowchart of the method. Registration: images before registration are shown; the arrows indicate what registration steps are applied and point from the fixed to the moving images. The large arrow points down from the domain in which all images are registered with each other and classification is done. Here  $\mu$ CT and MRI are only rigidly deformed, and the histology is non-rigidly deformed to match the *ex vivo* MRI. The in-plane resolution equals the ‘enface’ resolution, which is  $0.017 \times 0.017$  mm. Ground truth: deformed  $\mu$ CT (original and thresholded) and histology are shown with the resulting ground truth. Contours are shown in the undeformed histology slice in the upper row. Features: a slice of the transformed 3D-T1w MRI with its gradient magnitude and Laplacian image and the distance to the outer vessel wall for each voxel (black = 0) are shown. Classification: a segmentation result is shown (white = calcification, light grey = fibrous tissue, dark grey = necrotic tissue).

the exact threshold value is not very critical. The remaining areas were considered to be fibrous tissue, resulting in a ground truth with three components (calcification, fibrous tissue and lipid-rich necrotic tissue). Three specimens were excluded for classification due to a low quality of histology (two cases) or an incomplete MRI protocol (one case), leaving 12 specimens. We excluded slices for which it was not possible to obtain a ground truth due to histology disruption. The number of histology slices with the corresponding MRI included was  $11 \pm 4$  per specimen (range 5–17). Figure 1 shows a flowchart of the methods in this paper. It includes an example for the ground truth obtained in one image slice.

## 2.2. Registration of MRI with histology

**2.2.1. 3D registration.** The 3D registration procedure followed a pipeline that was developed to match *in vivo* CTA with histology, via  $\mu$ CT, and *ex vivo* MRI, using the same settings as published previously (Groen *et al* 2010). This procedure includes multiple steps using both rigid and deformable registration, and involves several manual annotations of anatomical landmarks, lumen and outer vessel-wall contours in each image modality. Figure 1 illustrates the different stages in registration. Here, a more detailed description of each step is given.

- *Enface stacking*: the ‘enface’ images taken during histology slicing were stacked by the 2D in-plane point-based rigid registration of manually annotated landmarks to form a 3D

volume. Each 2D image was registered to its preceding image ( $8 \pm 3$  landmarks per 2D registration were used).

- $T_{(Histology\ to\ enface)}$ : the 2D histology slices were registered to the enface stack to obtain a 3D histology stack, allowing for non-rigid in-slice deformations. Manually annotated landmarks ( $5 \pm 1$  per slice) were used for rigid initialization, and the B-spline model (Rueckert *et al* 1999) was used to maximize the mutual information (MI) (Viola and Wells 1997, Thevenaz and Unser 2000) between the original images, manual lumen segmentations and manual outer-wall segmentations in both the ‘enface’ and histology images.
- $T_{(Enface\ to\ MRI)}$ : this registration was initialized with a rigid registration based on manually annotated landmarks in the ‘enface’ and 3D-T1w MRI ( $8 \pm 2$  landmarks per specimen). A rigid registration of the MRI with the 3D histology stack was then performed with the same MI maximization and similarity measure as before for which manual lumen and outer vessel wall segmentations were made on the MRI. The registration was refined by B-spline deformable registration with the same similarity measure, in which only in-plane deformations were allowed while maintaining continuity along the slice direction. Since the specimen did not move between acquisitions, the different MR images were assumed to be aligned with each other.
- $T_{(\mu CT\ to\ MRI)}$ : to register the  $\mu$ CT to the *ex vivo* MRI, a point-based rigid registration without scaling was performed using annotated landmarks in the  $\mu$ CT and 3D-T1w image ( $7 \pm 2$  landmarks per specimen).

All registration steps were performed using the registration toolbox *elastix* (Klein *et al* 2010). Classification was performed in the space in between the *ex vivo* MRI and stacked histology in figure 1 indicated by the large arrow pointing down from the registration part. All images were transformed to this domain. This means that only the B-spline transformation of  $T_{(Enface\ to\ MRI)}$  was applied to the 3D histology stack, and that the inverse rigid transformation was applied to the *ex vivo* MRI. The in-plane resolution in this domain was chosen equal to the resolution of the enface images,  $0.017 \times 0.017$  mm, with a slice distance of 1 mm. In this way, interpolation between histology slices was avoided.

**2.2.2. 2D registration.** To study whether the match between MRI and histology with either 3D or 2D registration affects the classification results, two additional datasets were created. For that purpose, an observer manually matched 2D image slices from the MRI volumes to all histology slices of each specimen, using a procedure similar to the method used in several papers (Clarke *et al* 2003, 2006, Ronen *et al* 2007). Matching was achieved using the known distance between slices in both MRI and histology, and the shape of the lumen and outer vessel wall. The selected MRI slices had the same position in each of the three contrast weightings, as the specimen did not move between acquisitions. As the studies cited had used fewer slices per plaque for classification (1.25, 3 and 3–5), for this study the observer selected three slices from each specimen. The slices that had the best histology and MRI quality and covered as much as possible of the length of the specimen were selected, excluding adjacent slices. The selected MRI slices were registered in-plane with the corresponding histology, as in  $T_{(Enface\ to\ MRI)}$  for 3D registration: using a point-based rigid initialization ( $6 \pm 2$  landmarks per slice) followed by rigid and B-spline deformations that optimized the MI between the original images, manual lumen segmentations and manual wall segmentations in both images.

To better compare classification results in these 3D- and 2D-registered datasets, a third dataset was created. This consisted of the three MRI slices out of the 3D-registered dataset that corresponded to the three histology slices that had been selected for the 2D-registered dataset.

### 2.3. Voxel classification method

To remove intensity nonuniformities in the original MR images, we used the nonparametric nonuniform intensity normalization (N3) (Sled *et al* 1998). For intensity normalization between patients, the intensity values were scaled with respect to the average intensity of the fluid in the tube surrounding the plaque, which was set to 1000. To account for slight misregistrations and partial-volume effects, both in classifier training and evaluation, a morphological erosion with a disk-structuring element was applied to the segmentations. For calcifications, the diameter of the structuring element was set to 3 voxels (0.05 mm), for fibrous and necrotic tissue to 5 voxels (0.09 mm). The reason for this difference is that the ground truth for calcification was based on the registration of  $\mu$ CT and MRI, which was assumed to be more accurate than that of histology and MRI. The erosion was performed in separate 2D slices in the high-resolution domain for classification (indicated in figure 1) and deleted on average 21% of the voxels in the original segmentations. For every voxel, the following features were obtained.

- Intensities in the normalized MR images as obtained with the three different sequences, and after convolution with a Gaussian kernel with  $\sigma=0.1$  mm,  $\sigma=0.2$  mm and  $\sigma=0.3$  mm to obtain smoothed images (12 features)

$$IS(\mathbf{x}, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}^3} e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}} * I \quad (1)$$

with  $I$  the normalized image, and  $\mathbf{x}=[x \ y \ z]$ .

- The gradient magnitude of all smoothed images (nine features)

$$GM = \sqrt{\left(\frac{\partial(IS(\mathbf{x}, \sigma))}{\partial x}\right)^2 + \left(\frac{\partial(IS(\mathbf{x}, \sigma))}{\partial y}\right)^2 + \left(\frac{\partial(IS(\mathbf{x}, \sigma))}{\partial z}\right)^2}. \quad (2)$$

- The Laplacian of all smoothed images (nine features)

$$L = \frac{\partial^2(IS(\mathbf{x}, \sigma))}{\partial x^2} + \frac{\partial^2(IS(\mathbf{x}, \sigma))}{\partial y^2} + \frac{\partial^2(IS(\mathbf{x}, \sigma))}{\partial z^2}. \quad (3)$$

- The shortest (Euclidean) distance to the lumen and to the outer vessel wall (two features).

This resulted in a total set of 32 features. The calculation of the smoothed, gradient magnitude and Laplacian features was done in the original 3D MRI volumes before registration. The distances were calculated in-plane in the 2D slices after registration. Classification was performed with a linear discriminant classifier which finds the optimal linear boundaries between the classes under the assumption that the data are normally distributed with equal covariance matrices for the three classes. The discriminant function is as follows (Hastie *et al* 2003):

$$\rho_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k. \quad (4)$$

Here  $k$  are the classes,  $\Sigma$  the covariance matrix,  $\mu_k$  the class means,  $\pi_k$  the class prior probabilities, and  $x$  the feature vector to classify.

### 2.4. Evaluation

For both the 3D and 2D registration approach, the accuracy of the final registration of histology with *ex vivo* MRI was evaluated. We calculated the distance between the annotated landmark points, and the distance between the corresponding vessel wall contours. For the contour distance, the shortest Euclidean distance from each point on the lumen or outer vessel wall contour in the MR images to the corresponding contour in histology was taken, and the other way around. The mean of all these distances was calculated for each slice. The registration of



*ex vivo* MRI with  $\mu$ CT was evaluated based on the landmark distance after rigid transformation of the  $\mu$ CT.

The classifiers were evaluated using leave-one-out cross-validation in which one specimen per time was kept out of the training set and used for testing. One per cent of the voxels in the remaining 11 specimens was randomly selected to form the training set. Due to the large number of voxels in the images used for classification, an average of 5000 voxels of the smallest component (calcification) were still present in the training set. Both classifier training and testing were performed using the Matlab toolbox 'PRTools' (Duin *et al* 2007), version 4.1.9.

Classifier performance was evaluated on the basis of the voxelwise classification accuracy averaged over the 12 specimens. For feature evaluation, eight combinations of features were evaluated in the complete 3D-registered dataset. Firstly, the original MRI intensities as used previously (Clarke *et al* 2003, 2006, Ronen *et al* 2007) and the original images combined with either smoothed intensities, distances (as in Liu *et al* (2006)) or gradient magnitude and Laplacian images were considered to study the effect of the new features separately on their added value. In addition, a combination of original and smoothed intensities and distances was evaluated, and this combination with the gradient magnitude features, the Laplacian features or both added. Differences in accuracy between classifiers using these different feature sets were tested for significance with a Friedman test, a paired nonparametric for multiple comparisons. Individual comparisons were done using the Tukey–Kramer method.

To study whether accuracy is affected by 2D or 3D registration and slice selection, classification was performed in the three different datasets: the complete 3D-registered set with  $11 \pm 4$  slices per specimen, the 2D-registered dataset with three slices per specimen, and the selection of the same three slices per specimen from the 3D-registered dataset. Classification was performed using all features and the same leave-one-out procedure as mentioned before. Friedman's analysis was again used to test for significant differences.

### 3. Results

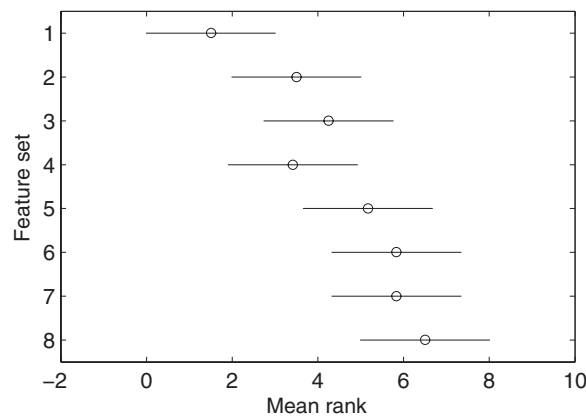
#### 3.1. Registration accuracy

For 3D registration, the distance between the annotated landmarks in histology and *ex vivo* MRI after all registration steps was  $1.3 \pm 0.7$  mm. The distance between lumen contours was on average  $0.5 \pm 0.3$  mm and between outer vessel wall contours  $0.3 \pm 0.2$  mm. Using 2D registration, the distance between landmarks was  $1.8 \pm 1.7$  mm, between lumen contours  $0.4 \pm 0.2$  mm and between outer vessel wall contours  $0.3 \pm 0.3$  mm. The distance between the landmarks on *ex vivo* MRI and  $\mu$ CT after applying the rigid transformation was  $0.3 \pm 0.7$  mm.

#### 3.2. Performance of features

The average leave-one-out accuracy, sensitivity and specificity for different combinations of features are shown in table 2. To compare, random class assignment according to the class priors (4% for calcification, 63% fibrous and 33% necrotic tissue) would result in an accuracy of 51%. A combination of the three original image intensities yielded an accuracy of  $60.8 \pm 16.1\%$ . Smoothed intensities, distances and the gradient magnitude together with Laplacian images all improved accuracy when separately added to the original intensities, although none of them statistically significant. When all features were included, the accuracy was the highest





**Figure 2.** Friedman's analysis. For each feature set, numbered in table 2, the mean rank is shown.

**Table 2.** Classification results for different feature combinations. The overall voxelwise classification accuracy is given as an average over the 12 specimens. Sensitivity and specificity for the three components are calculated after combining the 12 confusion matrices and given as one value for all specimens together. The last column indicates which feature sets had significantly different results (Friedman's analysis,  $p < 0.05$ ). I = original intensities (three features), IS = smoothed intensities (nine features), D = distances (two features), GM = gradient magnitude (nine features) and L = Laplacian (nine features).

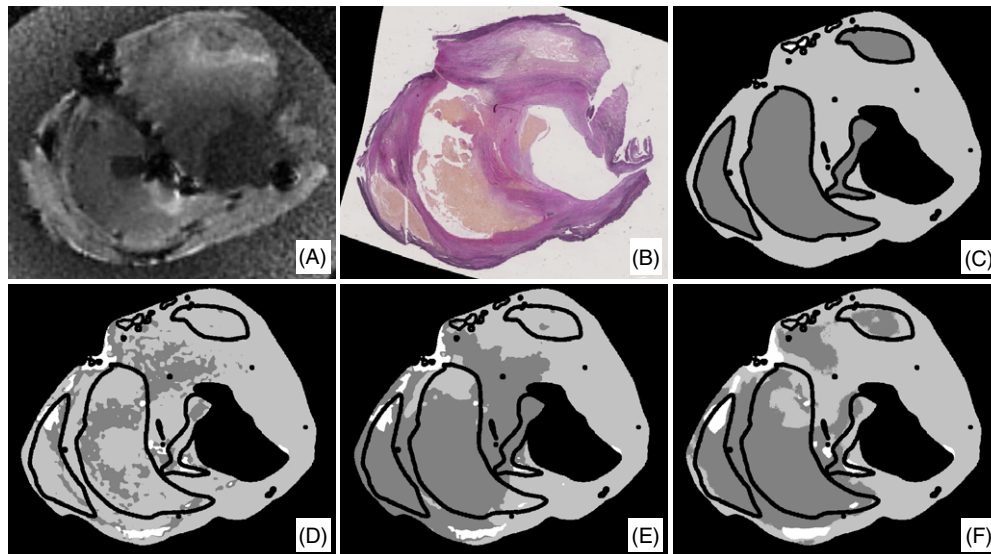
Feature set	Accuracy (%)	Calcification (Sens., Spec.)	Fibrous (Sens., Spec.)	Necrotic (Sens., Spec.)	Friedman ( $p < 0.05$ )
(1) I	$60.8 \pm 16.1$	81, 96	91, 23	10, 94	5, 6, 7, 8
(2) I, IS	$68.1 \pm 7.8$	79, 96	83, 52	42, 87	
(3) I, GM, L	$70.1 \pm 6.4$	80, 96	82, 56	48, 86	
(4) I, D	$70.5 \pm 6.8$	79, 96	84, 55	46, 88	8
(5) I, IS, D	$72.6 \pm 6.4$	78, 97	85, 59	51, 88	1
(6) I, IS, D, L	$72.9 \pm 6.6$	79, 97	85, 59	52, 88	1
(7) I, IS, D, GM	$73.0 \pm 6.2$	80, 96	85, 60	52, 88	1
(8) All 32 features	$73.3 \pm 6.3$	81, 97	85, 60	52, 89	1, 4

( $73.3 \pm 6.3\%$ ), and significantly better than classifiers that have been studied before which used only original intensities, or original intensities and distances (figure 2).

Figures 3 and 4 show two examples of a slice that was classified using only the original intensities, original intensities and distances, and all features. In figure 3, the classification accuracy increased from 62.3% to 74.9% using all features instead of only original intensities. In figure 4, the accuracy increased from 30.5% to 61.6%. The main improvement in both examples, and in most un-shown cases, is due to a more accurate detection of lipid-rich necrotic tissue.

### 3.3. Performance of the best classifier

The performance of the best classifier (i.e. including all features) is evaluated in more detail. Table 3 shows the confusion matrix of the different plaque components. As also shown in table 2, calcification can accurately be classified with high sensitivity (81%) and specificity (97%). The specificity for fibrous tissue is lower (60%, with a sensitivity of 85%) while for necrotic tissue the sensitivity is lower (52%, with a specificity of 89%). This indicates that



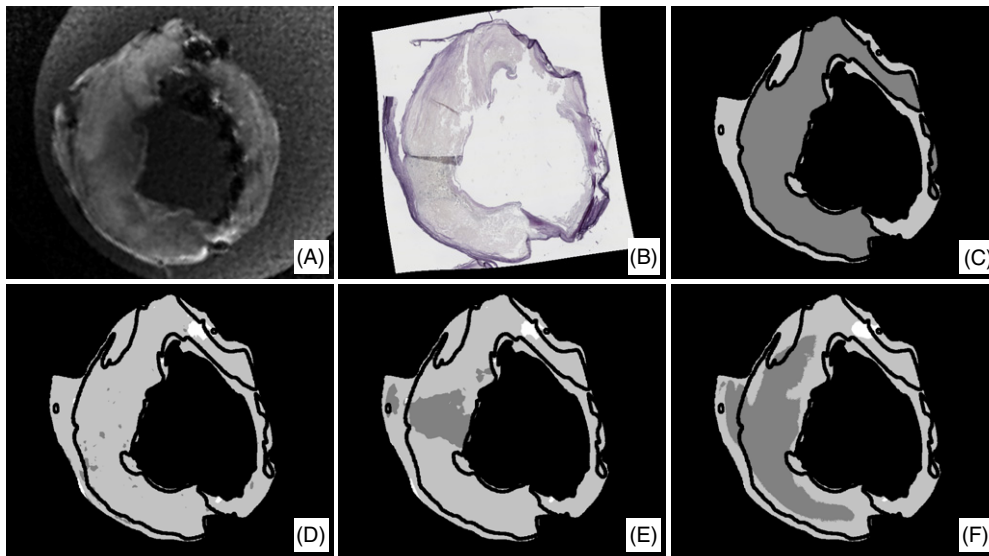
**Figure 3.** Example of classification in one slice. (A) The 3D-T1w MRI slice; (B) the registered histology and (C) the ground truth. The classification results are shown using only the three original intensities (D, 62.3% accuracy), original intensities with distances (E, 67.3% accuracy), and all 32 features (F, 74.9% accuracy). In (C–F), white represents calcification, light grey fibrous and dark grey lipid-rich necrotic tissue. Black lines and spots are due to the erosion, and not included in the analysis.

**Table 3.** Confusion matrix for classification using all features, summed over all specimens.

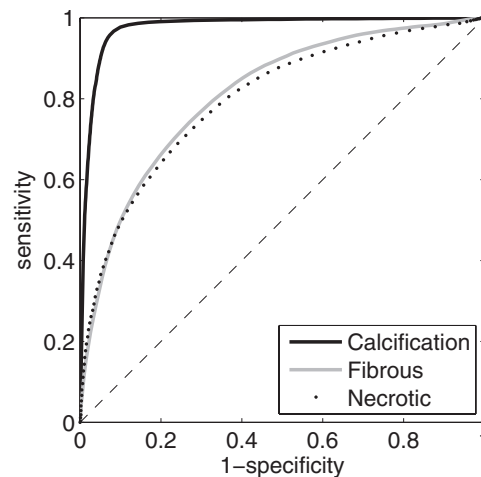
		Classification result		
		Calcification	Fibrous	Lipid-rich necrotic
Ground truth	Calcification	3.6%	0.5%	0.4%
	Fibrous	2.0%	53.5%	7.3%
	Lipid-rich necrotic	1.3%	14.4%	17.0%

the differentiation between fibrous and necrotic tissue is more difficult and, as table 3 shows, may lead to the misclassification of necrotic tissue as fibrous tissue. The accuracies per class, considering the other two classes as one, are 96% for calcification, 76% for fibrous tissue and 77% for necrotic tissue. The receiver-operating curves (figure 5) also show that a very good sensitivity and specificity can be achieved for calcification (area under the curve (AUC) of 0.98), with lower areas under the curve for fibrous (AUC of 0.82) and necrotic tissue (AUC of 0.81).

Figure 6 shows correlations between the fractions of the different plaque components found in the ground truth and by classification. Values for Spearman's correlation coefficients ( $\rho$ ) were reasonable to good (0.86 for calcification, 0.71 for fibrous tissue and 0.72 for necrotic tissue). On average, the percentage of calcification in the ground truth is  $2.3 \pm 2.4\%$  lower than in the classification result. Fibrous tissue is  $7.7 \pm 11.6\%$  overestimated and necrotic tissue  $10.1 \pm 12.2\%$  underestimated. For illustration, figure 7 shows four examples: (A) a slice with high classification accuracy (83.6%), (B and C) with average accuracy (74.0% and 69.2%) and (D) with low accuracy (52.4%). These examples again show the accurate detection of calcification and the more difficult separation between fibrous and necrotic tissue.



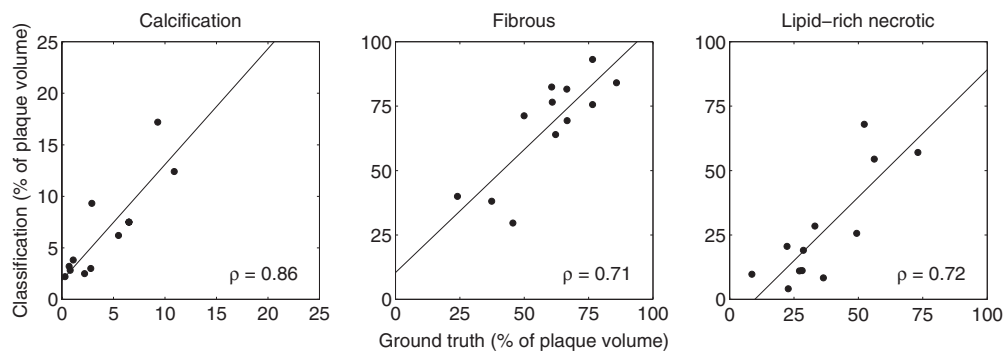
**Figure 4.** Example of classification in one slice. (A) The 3D-T1w MRI slice, (B) the registered histology and (C) the ground truth. The classification results are shown using only the three original intensities (D, 30.5% accuracy), original intensities with distances (E, 41.0% accuracy) and all 32 features (F, 61.6% accuracy). In (C–F), white represents calcification, light grey fibrous and dark grey lipid-rich necrotic tissue. Black lines and spots are due to the erosion, and not included in the analysis.



**Figure 5.** ROC curves for the three components studied. The area under the curve for calcification is 0.98, for fibrous tissue 0.82 and for lipid-rich necrotic tissue 0.81.

### 3.4. 2D versus 3D registration

The classification results in the 2D-registered and 3D-registered datasets that contained three slices per specimen are shown in table 4. The overall accuracy in these datasets is slightly lower ( $71.9 \pm 11.3\%$  using 3D registration and  $70.1 \pm 9.0\%$  using 2D registration) than when using the complete 3D-registered dataset. The differences are, however, not statistically significant.



**Figure 6.** Correlation between the relative size of plaque components in the ground truth and classification results. Values for Spearman's  $\rho$  are 0.86 for calcification, 0.71 for fibrous tissue and 0.72 for lipid-rich necrotic tissue.

**Table 4.** Classification results for the three different datasets. The overall voxelwise classification accuracy is given as an average over the 12 specimens. Sensitivity and specificity for the three components are given as the total sensitivity and specificity across all 12 specimens combined. Differences are not significant.

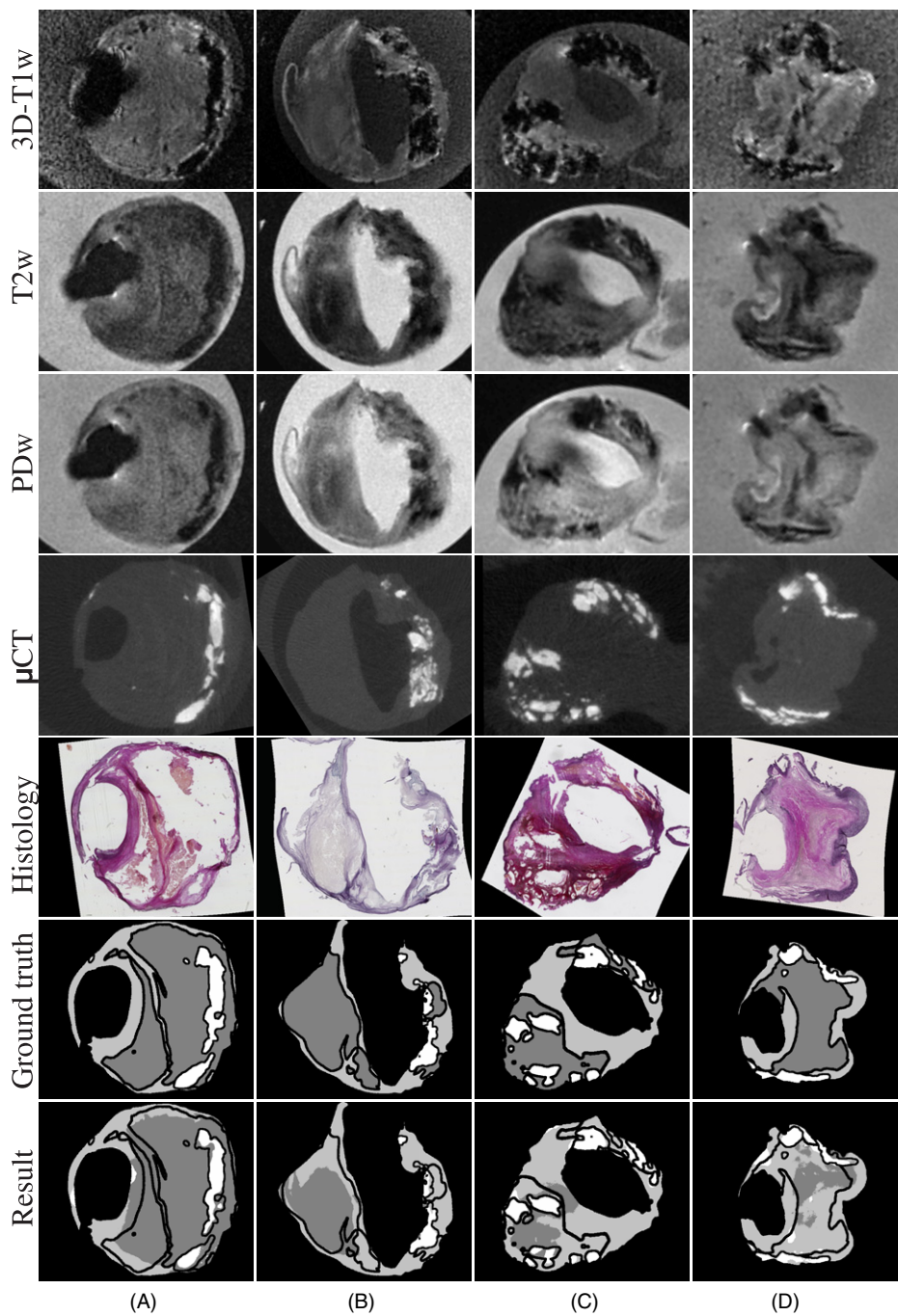
Dataset	Accuracy (%)	Calcification (Sens., Spec.)	Fibrous (Sens., Spec.)	Necrotic (Sens., Spec.)
Complete 3D set	73.3 $\pm$ 6.3	81, 97	85, 60	52, 89
Selection from 3D set	71.9 $\pm$ 11.3	81, 98	82, 62	57, 85
2D set	70.1 $\pm$ 9.0	70, 96	78, 66	61, 83

The smaller datasets show a slightly higher sensitivity for necrotic tissue and a lower sensitivity for fibrous tissue. The average rotation angle needed to align the MRI volumes with the 3D histology stack was  $7 \pm 4^\circ$ .

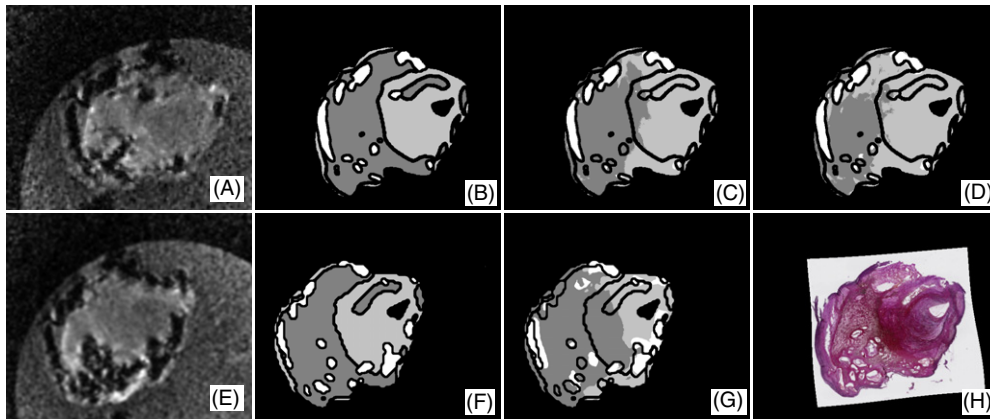
Figures 8 and 9 show classification of two slices that are part of all three datasets. While the matched 2D MRI slice and the corresponding slice from the 3D-registered MRI volume in figure 8 may look similar, the MRI volume was rotated for  $10.7^\circ$  for the 3D registration. This example indeed shows the highest classification accuracy after training on the complete 3D set, followed by the selection from the 3D set and the 2D set. Due to differences in non-rigid deformation, the ground truth images after 2D and 3D registration may differ slightly. As the  $\mu$ CT volume is rotated together with the MRI, the location of calcification varies. Figure 9 shows an example in which the rotation of the MRI volume for 3D registration was larger ( $13.4^\circ$ ). This example also shows the misclassification of fibrous and necrotic tissue as calcification caused by the presence of air in the specimen which results in black spots in the MRI.

#### 4. Discussion

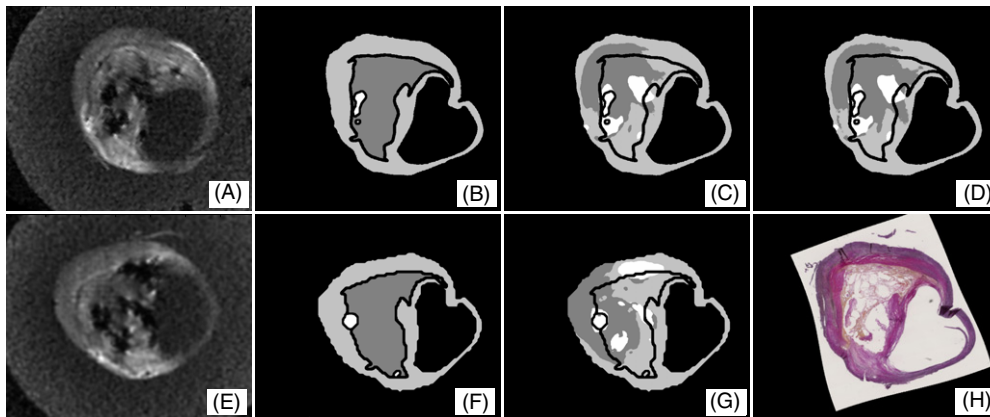
We presented an automated method for voxelwise classification of carotid plaque composition in *ex vivo* MRI. Classification accuracy was significantly improved by features that describe the local image structure and the location within the vessel wall, in addition to image intensity. Using 3D registration of MRI to histology with as many slices included for training and evaluation as possible ( $11 \pm 4$ ) gave slightly but not significantly better results than when only three slices per specimen were included, or when 2D registration was used to match MRI



**Figure 7.** Four examples of slices classified using all features. The slice in (A) has a high classification accuracy (83.6%), the slices in (B and C) an average accuracy (74.0% and 69.2%), and the slice in (D) has a low classification accuracy (52.4%). In the ground truth and results, calcification is white, fibrous tissue light grey and lipid-rich necrotic tissue dark grey. For visualization, the MRI and  $\mu$ CT images have been individually scaled relative to the minimum and maximum intensity value to obtain an optimal contrast per image.



**Figure 8.** Example showing the same histology slice (H) classified in the three different datasets. (A) MRI matched to this histology slice using 3D registration; (E) MRI matched using 2D registration. (B) (3D) and (F) (2D), the corresponding ground truth. (C) The classification result using the complete 3D-registered set for training (accuracy 76.5%), (D) using a selection from the 3D-registered set (74.5%), and (G) using the 2D-registered set (61.6%).



**Figure 9.** Example showing the same histology slice (H) classified in the three different datasets. (A) MRI matched to this histology slice using 3D registration; (E) MRI matched using 2D registration. (B) (3D) and (F) (2D), the corresponding ground truth. (C) The classification result using the complete 3D-registered set for training (accuracy 66.2%), (D) using a selection from the 3D-registered set (67.2%), and (G) using the 2D-registered set (52.2%).

and histology. The method showed high classification accuracy for calcification (96%) and reasonable accuracy for fibrous (76%) and lipid-rich necrotic tissue (77%).

Four types of features were used in addition to original intensities: the distances to the lumen and outer vessel wall, and smoothed, gradient magnitude and Laplacian images at several scales. The positive effect of including distance information has been shown previously (Liu *et al* 2006); it is relevant because the necrotic core is usually located in the centre of a plaque, surrounded by fibrous tissue. Because smoothed images reduce noise and include information from neighbouring voxels, they may therefore improve performance. Gradient magnitude and Laplacian features are useful to detect small structures; in this study, their main effect was an improved sensitivity to calcification detection. Although the improvement of adding gradient magnitude and Laplacian images was small as calcification was already accurately classified,



only the classifier that used all features was significantly better than the classifier that used only the original intensities and distances.

The component that is detected with the highest accuracy is calcification, which appears dark in MRI compared to the rest of the plaque. There is a larger overlap in intensity between fibrous and necrotic tissue in the MR images, which makes the differentiation between these components more difficult. Another reason for the difference in accuracy is that registration of *ex vivo* MRI to  $\mu$ CT is slightly more accurate than that of *ex vivo* MRI to histology due to deformations that occur during histological sectioning. Histology yielded the ground truth for fibrous and lipid-rich necrotic tissue and  $\mu$ CT for calcification. We accounted for small registration errors by erosion of the ground truth areas. The erosion was a little smaller (0.17 mm between fibrous and necrotic regions) than the average contour distance ( $0.5 \pm 0.3$  mm for lumen and  $0.3 \pm 0.2$  mm for the outer wall using 3D registration, and  $0.4 \pm 0.2$  mm for lumen and  $0.3 \pm 0.3$  mm for the outer wall using 2D registration). Although misregistration may have affected results in areas with larger errors, most smaller misregistrations were accounted for and visual inspection does not suggest misregistration to be a problem. Differentiation between fibrous and necrotic tissue is important to separate vulnerable from stable plaques. Spearman's rank correlation coefficients indicate that when the plaques are ranked on the relative volume per component, the result corresponds moderately to well with the ground-truth ranking (figure 6). It is therefore possible to use these plaque component sizes in studies that relate plaque composition to prognosis or outcome and possibly use it for the detection of vulnerable plaques.

3D registration of MRI to histology seemed to result in slightly, although not significantly, higher classification accuracies than when using 2D registration. The rotation angles for 3D registration were quite small ( $7 \pm 4^\circ$ ). Although in this study the difference in classification using either 3D registration or 2D registration was small as well, we believe that 3D registration can provide more accurate training data. This will especially be the case for datasets in which the angle between the slicing direction of the histology specimen and the MRI scan direction is larger, for example, in *in vivo* studies. In addition, registration of the complete volume at once is more consistent and reduces the chance that a single slice will be misregistered. The difference found when many slices or when a selection of slices is used may be explained by the fact that the selected slices had a slightly different composition (57% fibrous and 38% necrotic tissue in the subset from the 3D-registered set and 55% fibrous and 40% necrotic tissue in the 2D-registered set, against 63% and 33% in the complete dataset). Classification accuracy may therefore be higher in the complete dataset because it contains slices that consist of more fibrous tissue, which may be easier to classify.

These results are comparable to those of previous studies that performed voxelwise plaque classification in *ex vivo* MRI (Clarke *et al* 2003, 2006, Ronen *et al* 2007). These studies used only the original MRI intensities and had resolutions slightly lower than in the present study. When eight contrast weightings were used Clarke *et al* reported an overall accuracy of 73.5% (Clarke *et al* 2003) or 78% when either these eight scans or a combination of T1w, T2w and diffusion weighting (Dw) was used (Clarke *et al* 2006). In Clarke's study, a combination of T1w, T2w and PDw, as used in this study, resulted in an accuracy of 67% (Clarke *et al* 2006). Another study that misclassified more than 40% of the voxels when five components were distinguished and that misclassified 20–30% when fibrous and loose connective tissue, and haemorrhage and necrotic tissue were combined, also yielded slightly better results when PDw images were replaced with Dw images (Ronen *et al* 2007). Our study did not include Dw images; classification accuracies might therefore improve when this contrast weighting is added. Using only the original intensities, we obtained an accuracy of 61%, which is slightly lower than in previous studies. Differences with these previous studies may be related to image quality, MRI



protocols, the image slices selected and possibly other factors. However, we expect the added value of the studied features to be valid on other data as well. As shown in table 1, MRI scanning parameters varied between specimens in our protocol, but we did not observe a correlation between deviation from the mean scanning parameters and classification accuracy. The studies by Clarke *et al* classified fibrous, loose connective, calcified and necrotic tissue (Clarke *et al* 2003) and fibrous, necrotic, calcified and haemorrhagic tissue (Clarke *et al* 2006). 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, and combining them therefore improved the reliability of our ground truth. Haemorrhage was not included in our study as this plaque component cannot be identified on histological cross-sections with Elastica von Gieson stains. If a ground truth for haemorrhage is present, it can easily be added to the framework.

## 5. Conclusion

We presented a new system for the voxelwise classification of calcification, fibrous and lipid-rich necrotic tissue in *ex vivo* MRI of carotid specimens. Our approach uses 3D registration of MRI with a histology stack to train and evaluate classifiers. Features used for classification are original MRI intensities as well as smoothed intensities, gradient magnitude images, Laplacian images and distances to the lumen and outer vessel wall. This yields results that are significantly better than when only original intensities and distances are used. The classification results slightly improved by 3D registration compared to 2D registration.

## Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine ([www.ctmm.nl](http://www.ctmm.nl)), project PARISK (grant 01C-202), and supported by the Netherlands Heart Foundation. WN, SK and MB were financially supported by the Netherlands Organisation for Scientific Research (NWO). HG was financially supported by the Interuniversity Cardiology Institute of the Netherlands. Kim van Gaalen is acknowledged for her help with the histological processing and creating the 3D histology.

## References

- Adame I M, van der Geest R J, Wasserman B A, Mohamed M A, Reiber J H C and Lelieveldt B P F 2004 Automatic segmentation and plaque characterization in atherosclerotic carotid artery MR images *Magn. Reson. Mater. Phys. Biol. Med.* **16** 227–34
- Akyildiz A, Speelman L, van Brummelen H, Gutierrez M, Virmani R, van der Lugt A, van der Steen A, Wentzel J and Gijssen F 2011 Effects of intima stiffness and plaque morphology on peak cap stress *Biomed. Eng. OnLine* **10** 25
- Chu B, Kampschulte A, Ferguson M S, Kerwin W S, Yarnykh V L, O'Brien K D, Polissar N L, Hatsukami T S and Yuan C 2004 Hemorrhage in the atherosclerotic carotid plaque: a high-resolution MRI study *Stroke* **35** 1079–84
- Clarke S E, Beletsky V, Hammond R R, Hegele R A and Rutt B K 2006 Validation of automatically classified magnetic resonance images for carotid plaque compositional analysis *Stroke* **37** 93–7
- Clarke S E, Hammond R R, Mitchell J R and Rutt B K 2003 Quantitative assessment of carotid plaque composition using multicontrast MRI and registered histology *Magn. Reson. Med.* **50** 1199–208
- Duin R, Juszczak P, Paclik P, Pekalska E, de Ridder D, Tax D and Verzakov S 2007 *PRTTools4.1: A Matlab Toolbox for Pattern Recognition* Delft University of Technology
- Groen H C *et al* 2010 Three-dimensional registration of histology of human atherosclerotic carotid plaques to *in vivo* imaging *J. Biomech.* **43** 2087–92
- Hastie T, Tibshirani R and Friedman J H 2003 *The Elements of Statistical Learning* corrected edn (Berlin: Springer)

- Hatsukami T S, Ross R, Polissar N L and Yuan C 2000 Visualization of fibrous cap thickness and rupture in human atherosclerotic carotid plaque *in vivo* with high-resolution magnetic resonance imaging *Circulation* **102** 959–64
- Hofman J, Branderhorst W, ten Eikelder H, Cappendijk V, Heeneman S, Kooi M, Hilbers P and ter Haar Romeny B 2006 Quantification of atherosclerotic plaque components using *in vivo* MRI and supervised classifiers *Magn. Reson. Med.* **55** 790–9
- Itskovich V V, Samber D D, Mani V, Aguinaldo J G S, Fallon J T, Tang C Y, Fuster V and Fayad Z A 2004 Quantification of human atherosclerotic plaques using spatially enhanced cluster analysis of multicontrast-weighted magnetic resonance images *Magn. Reson. Med.* **52** 515–23
- Karmonik C, Basto P, Vickers K, Martin K, Reardon M J, Lawrie G M and Morrisett J D 2009 Quantitative segmentation of principal carotid atherosclerotic lesion components by feature space analysis based on multicontrast MRI at 1.5 T *IEEE Trans. Biomed. Eng.* **56** 352–60
- Klein S, Staring M, Murphy K, Viergever M and Pluim J 2010 elastix: a toolbox for intensity-based medical image registration *IEEE Trans. Med. Imaging* **29** 196–205
- Liu F, Xu D, Ferguson M S, Chu B, Saam T, Takaya N, Hatsukami T S, Yuan C and Kerwin W S 2006 Automated *in vivo* segmentation of carotid plaque MRI with morphology-enhanced probability maps *Magn. Reson. Med.* **55** 659–68
- Mitsumori L M, Hatsukami T S, Ferguson M S, Kerwin W S, Cai J and Yuan C 2003 *In vivo* accuracy of multisequence MR imaging for identifying unstable fibrous caps in advanced human carotid plaques *J. Magn. Reson. Imaging* **17** 410–20
- Morrisett J, Vick W, Sharma R, Lawrie G, Reardon M, Ezell E, Schwartz J, Hunter G and Gorenstein D 2003 Discrimination of components in atherosclerotic plaques from human carotid endarterectomy specimens by magnetic resonance imaging *ex vivo* *Magn. Reson. Imaging* **21** 465–74
- Park A E, McCarthy W J, Pearce W H, Matsumura J S and Yao J S 1998 Carotid plaque morphology correlates with presenting symptomatology *J. Vasc. Surg.* **27** 872–9
- Petty G W, Brown R D J, Whisnant J P, Sicks J D, O'Fallon W M and Wiebers D O 1999 Ischemic stroke subtypes: a population-based study of incidence and risk factors *Stroke* **30** 2513–6
- Ronen R R, Clarke S E, Hammond R R and Rutt B K 2007 Carotid plaque classification: defining the certainty with which plaque components can be differentiated *Magn. Reson. Med.* **57** 874–80
- Rueckert D, Sonoda L, Hayes C, Hill D, Leach M and Hawkes D 1999 Nonrigid registration using free-form deformations: application to breast MR images *IEEE Trans. Med. Imaging* **18** 712–21
- Saam T, Ferguson M, Yarnykh V, Takaya N, Xu D, Polissar N, Hatsukami T and Yuan C 2005 Quantitative evaluation of carotid plaque composition by *in vivo* MRI *Arterioscler. Thromb. Vasc. Biol.* **25** 234–9
- Seeger J M, Barratt E, Lawson G A and Klingman N 1995 The relationship between carotid plaque composition, plaque morphology, and neurologic symptoms *J. Surg. Res.* **58** 330–6
- Shaalán W E, Cheng H, Gewertz B, McKinsey J F, Schwartz L B, Katz D, Cao D, Desai T, Glagov S and Bassiouny H S 2004 Degree of carotid plaque calcification in relation to symptomatic outcome and plaque inflammation *J. Vasc. Surg.* **40** 262–9
- Shinnar M, Fallon J T, Wehrli S, Levin M, Dalmacy D, Fayad Z A, Badimon J J, Harrington M, Harrington E and Fuster V 1999 The diagnostic accuracy of *ex vivo* MRI for human atherosclerotic plaque characterization *Arterioscler. Thromb. Vasc. Biol.* **19** 2756–61
- Sled J G, Zijdenbos A P and Evans A C 1998 A nonparametric method for automatic correction of intensity nonuniformity in MRI data *IEEE Trans. Med. Imaging* **17** 87–97
- Takaya N *et al* 2006 Association between carotid plaque characteristics and subsequent ischemic cerebrovascular events: a prospective assessment with MRI—initial results *Stroke* **37** 818–23
- Thevenaz P and Unser M 2000 Optimization of mutual information for multiresolution image registration *IEEE Trans. Image Process.* **9** 2083–99
- Toussaint J F, LaMuraglia G M, Southern J F, Fuster V and Kantor H L 1996 Magnetic resonance images lipid, fibrous, calcified, hemorrhagic, and thrombotic components of human atherosclerosis *in vivo* *Circulation* **94** 932–8
- van Engelen A, de Bruijne M, Klein S, Verhagen H, Groen H, Wentzel J, van der Lugt A and Niessen W 2011 Plaque characterization in *ex vivo* MRI evaluated by dense 3D correspondence with histology *Proc. SPIE* **7963** 796329
- Viola P and Wells W I 1997 Alignment by maximization of mutual information *IEEE Int. Conf. on Comput. Vision* p 16
- Yuan C, Mitsumori L M, Ferguson M S, Polissar N L, Echelard D, Ortiz G, Small R, Davies J W, Kerwin W S and Hatsukami T S 2001 *In vivo* accuracy of multispectral magnetic resonance imaging for identifying lipid-rich necrotic cores and intraplaque hemorrhage in advanced human carotid plaques *Circulation* **104** 2051–6
- Yuan C, Zhang S x, Polissar N L, Echelard D, Ortiz G, Davis J W, Ellington E, Ferguson M S and Hatsukami T S 2002 Identification of fibrous cap rupture with magnetic resonance imaging is highly associated with recent transient ischemic attack or stroke *Circulation* **105** 181–5