

## Random forest active learning for AAA thrombus segmentation in computed tomography angiography images



Josu Maiora, Borja Ayerdi, Manuel Graña \*

Computational Intelligence Group, UPV/EHU, Spain

### ARTICLE INFO

Article history:  
Received 27 March 2012  
Received in revised form  
21 November 2012  
Accepted 10 January 2013  
Available online 2 August 2013

Keywords:  
Active learning  
Random forests  
CTA image segmentation  
Abdominal aortic aneurysm segmentation

### ABSTRACT

Image segmentation of 3D Computed Tomography Angiography (CTA) is affected by a variety of noise conditions that may render ineffective image segmentation procedures that have been developed and validated on a collection of training CTA data when applied on new CTA data. The approach followed in this paper to tackle this problem is to provide an Active Learning based interactive image segmentation system which will allow quick volume segmentation, with minimal intervention of a human operator. Image segmentation is achieved by a Random forest (RF) classifier applied on a set of image features extracted from each voxel and its neighborhood. An initial set of labeled voxels is required to start the process, training an initial RF. The most uncertain unlabeled voxels are shown to the human operator to select some of them for inclusion in the training set, retraining the RF classifier. The approach is applied to the segmentation of the thrombus of Abdominal Aortic Aneurysm (AAA) in CTA data (of patients), showing that the CTA volume can be accurately segmented after few iterations requiring a small labeled data sample.

© 2013 Elsevier B.V. All rights reserved.

### 1. Introduction

An **abdominal Aortic Aneurysm (AAA)** is a local dilation of the Aorta that occurs between the renal and iliac arteries. The weakening of the aortic wall leads to its deformation and the generation of a thrombus. 3D Computerized Tomography Angiography (CTA) allows minimally invasive visualization of the Aorta's lumen, thrombus and calcifications. The segmentation of the AAA thrombus is a challenging task due to the low contrast of signal intensity values between the AAA thrombus and its surrounding tissue, as can be appreciated in Fig. 1. Furthermore, the AAA thrombus shows great shape variability, both intra and inter-subjects, so that little prior information is available to guide the segmentation. General reviews of blood vessel segmentation methods are given in [1,2].

AAA thrombus segmentation methods reported in the literature need a lot of human interaction or *a priori* information one way or the other. A method based on Active Shape Models is described in [3] which needs initial manual labeled landmark points in one slice. The initial contour is propagated to neighboring slices on the basis of grayscale similarities. A classification approach that needs an initial manual segmentation of the Aorta lumen is proposed in [4]. The classifier features are grayscale profiles extracted from the normal to the lumen surface following a careful manual sampling procedure. In [5] an initial rough

specification of the aneurysm surface is refined by means of level set segmentation driven by an *a priori* model and the likelihood estimation provided by Support Vector Machine classifiers trained on voxel location, intensity and texture features. In [6] a deformable NURBS model is driven by a probability map built from a Gaussian Mixture Model trained on selected samples. This approach needs an initial manual lumen segmentation and intensity renormalization to avoid convergence mishaps of the NURBS model adaptation. In [7] the AAA thrombus after endovascular repair is detected following a radial model approach needing the specification of the lumen centerline and some manually tuned correction performed on the polar coordinate representation of the image. A graph-cut approach constrained by a geometrical model is proposed in [8], needing a previous lumen segmentation and centerline computation. The approach iterates labeling and geometric model re-estimation, which are costly processes.

The approach followed in this paper for AAA thrombus segmentation is to build a voxel<sup>1</sup> classifier into AAA thrombus or background classes [8,4,5]. Classification approaches need careful selection and labeling of training data samples from the available data. In response to this issue, Active Learning [9] tries to achieve the most accurate classification using the smallest possible training set, minimizing the user interaction needed to label the training samples. Active Learning starts with a minimal training sample, adding new labeled samples in an iterative process.

\* Corresponding author. Tel.: +34 943018044.  
E-mail address: [manuel.grana@ehu.es](mailto:manuel.grana@ehu.es) (M. Graña).

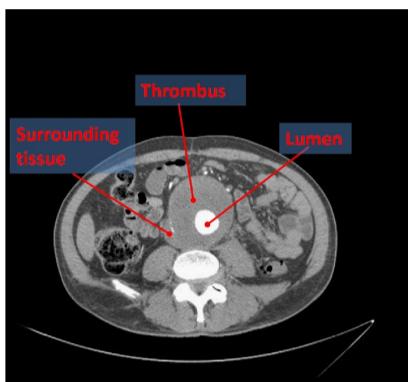
0925-2312/\$ - see front matter © 2013 Elsevier B.V. All rights reserved.  
<http://dx.doi.org/10.1016/j.neucom.2013.01.051>

<sup>1</sup> In this paper we work with 3D images, so that each image element is a voxel instead of a pixel.

1 build a voxel<sup>1</sup> classifier into AAA thrombus or background classes [8,4,5]. Classification approaches need careful selection and labeling of training data samples from the available data

2 initial manual labeled landmark points

,



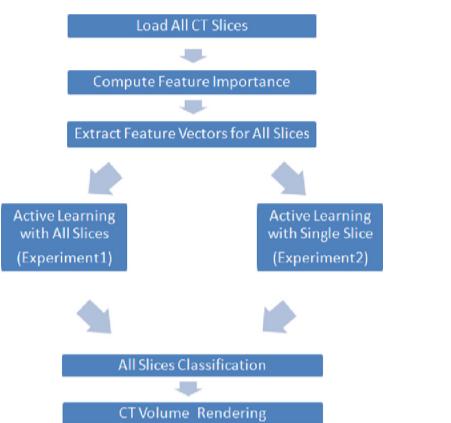
**Fig. 1.** CTA axial slice depicting thrombus and lumen of AAA. Blood in lumen is highlighted due to contrast, but thrombus intensity levels are similar to other surrounding tissues.

Aiming to provide the greatest increase in classifier accuracy [10], the additional samples are selected according to some classification uncertainty measure.<sup>2</sup> Besides its benefits in economy of computation and data labeling, Active Learning is also useful when the underlying data statistics are non-stationary, so that the classifier built at one time instant may not be optimal later on. Active Learning has been successfully applied to the classification of remote sensing images [11–13], and image retrieval based on semisupervised Support Vector Machines [14]. Active Learning inspiration for the selection of a minimal collection of training images is proposed in [15] for the development of combined generative and discriminative models in the segmentation of CT scans. An active feedback approach is used in [16] to improve the classification based annotation of radiographs. From a clinical application point of view, Active Learning may allow the expert radiologist to obtain fast and accurate segmentations with minimal interaction, despite strong changes in the CTA data.

We perform voxel classification using a Random forest (RF) classifier based on intensity features computed on each voxel's neighborhood. These features are maximum, minimum, median and Gaussian weighted average of the 2D neighborhoods of the voxel of increasing radius. The RF is convenient for this task because of its superior reported performance [17], simplicity of training and robustness, which have been shown in reported applications to medical image segmentation. Specific examples are myocardium delineation in 3D ultrasound (US) imaging of adult hearts [18], brain tissue segmentation [19,20] and segmentation of soft organs in abdominal and thoracic CT volumes [21,22]. In [23] we provided some first results of the approach proposed in this paper. Moreover, the hybridization [24,30,31] of Active Learning and RF is possible because RF allows for the straightforward definition of a classification uncertainty measure, namely, the variance of the individual tree classifiers' outputs.

The specific contributions of the approach proposed in this paper relative to the state of the art of AAA thrombus segmentation algorithms are as follows: (1) the need for human intervention in the selection of samples and labeling is reduced to a minimum by Active Learning, (2) use of RF allows quick learning and adaptation to incremental training datasets, (3) there is no requirement of a priori

<sup>2</sup> The classification uncertainty measure does not require actual knowledge of the data sample label, thus no double-dipping is incurred.



**Fig. 2.** Pipeline of the experimental setup for the Active Learning segmentation process.

information or geometric models, (4) feature extraction does not require sophisticated data processing, (5) the adaptation of the classifier to new data does not require skilful data processing, only picking the most uncertain voxels over a data visualization.

The experimental setup for validation is illustrated in Fig. 2. The paper tests two validation strategies. On one hand (Experiment 1), we train separate RF classifiers by Active Learning on each CT slice known to contain part of the thrombus. On the other hand (Experiment 2), we apply one RF classifier trained by Active Learning on the volume's central slice to the remaining slices of the volume, in order to test the generalization power of the approach. In both the experiments, the Active Learning oracle providing the samples' labels in the reported experiments is the ground truth provided by manual segmentation.

The structure of the paper is as follows: Section 2 describes the learning and feature selection methods. Section 3 describes the experimental set-up. Section 4 provides the experimental results. Finally, Section 5 provides our conclusions and some further work ideas.

## 2. Learning and feature selection

### 2.1. Random forest classifiers

Random forest (RF) algorithm is a classifier [25] that encompasses bagging [26] and random decision forests [27,28] being used in a variety of applications [17]. RF became popular due to its simplicity of training and tuning while offering a similar performance to boosting. Consider a RF as a collection of decision tree predictors, built so that they are as much decorrelated as possible, denoted as

$$\{h(\mathbf{x}; \psi_t); t = 1, \dots, T\},$$

where  $\mathbf{x}$  is a  $d$ -dimensional random sample of random vector  $X$ ,  $\psi_t$  are independent identically distributed random vectors whose nature depends on their use in the tree construction, and each tree casts a unit vote to find the most popular class of input  $\mathbf{x}$ . RF captures complex interaction structures in data, and are proposed [25] to be resistant to both over-fitting of data when individual trees are very deep and no pruned, and under-fitting when individual trees are too shallow.

Given a dataset of  $N$  samples, a bootstrapped training dataset is used to grow a tree  $h(\mathbf{x}; \psi_t)$  on a randomly selected subset of data dimensions  $d$  such that  $d \ll d$ . Decision tree growing recursively picks the best data split of each node based on these information measure of each dimension. In RF pruning is not required [25]. The independent identically distributed random vectors  $\psi_t$  determine the random dimension selection and data sample bootstrapping prior to tree training, which are the source for individual tree diversity and decorrelation of their outputs.

The trained RF can be used for classification of a new input  $\mathbf{x}$  by majority voting among the class prediction of the RF trees  $C_u(x)$ . The critical parameters of the RF classifier for the experiments reported below are the number of trees in the forest, the dimension of the random subspace, and the maximum tree depth. These parameters are set heuristically or by trial-and-error experimentation.

## 2.2. Active learning fundamentals

The performance of supervised classifiers strongly depends on the information provided by the data used to train the classifier, so that the appropriate selection and labeling of the training set may be a cumbersome task requiring extensive manual inspection and analysis of the data, typically requiring some visualization tool and labeling of each data sample. Besides, noisy samples may interfere the class statistics, which may lead to poor classification performances and/or over-fitting. For these reasons, a training set must be constructed in a smart way, meaning that it must consist of the minimal set of samples allowing to compute correctly the class boundaries, therefore it must contain the most informative data samples. In the machine learning literature this approach is known as Active Learning.

Active Learning [9,13] focuses on the interaction between the user and the classifier. Let  $X = \{\mathbf{x}_i, y_i\}_{i=1}^l$  be a training set consisting of labeled samples, with  $\mathbf{x}_i \in \mathbb{R}^d$  and  $y_i \in \{1, \dots, N\}$ . Let  $U = \{\mathbf{x}_j\}_{j=l+1}^{l+u} \in \mathbb{R}^d$  be the pool of candidates, with  $u \gg l$ , corresponding to the set of unlabeled voxels to be classified. In a given iteration  $t$ , the Active Learning algorithm selects from the pool  $U$  the  $q$  candidates that will, at the same time, maximize the gain in performance and reduce the uncertainty of the classification model when added to the current training set  $X^t$ . The selected samples  $S^t = \{\mathbf{x}_m\}_{m=1}^q \subset U$  are labeled with labels  $\{y_m\}_{m=1}^q$  by an oracle, which can be a human operator in interactive segmentation, or the available ground truth when performing cross-validation experiments. Finally, the set  $S^t$  is added to the current training set ( $X^{t+1} = X^t \cup S^t$ ) and removed from the pool of candidates ( $U^{t+1} = U^t \setminus S^t$ ). The process is iterated until a stopping criterion is met, such as the achieved accuracy reaching a preset threshold  $\theta_{max}$ . Algorithm 1 summarizes the Active Learning process.

## Algorithm 1. Active learning general algorithm.

Inputs  
 -Initial training set  $X^t = \{\mathbf{x}_i, y_i\}_{i=1}^l (X \in \mathcal{X}, t = 1)$ .  
 -Pool of candidates  $U^t = \{\mathbf{x}_j\}_{j=l+1}^{l+u} (U \in \mathcal{X}, t = 1)$ .  
 -Number of voxels  $q$  to add at each iteration (defining the batch of selected voxels  $S$ ).  
 1: **repeat**  
 2:   Train a classifier with current training set  $X^t$   
 3:   **for** each candidate in  $U^t$  **do**  
 4:     Evaluate a user-defined heuristic  
 5:   **end for**  
 6:   Rank the candidates in  $U^t$  according to the score of the heuristic  
 7:   Select the  $q$  most interesting voxels  $S^t = \{\mathbf{x}_k\}_{k=1}^q$

8:   The system assigns a label to the selected voxels  
 $S^t = \{\mathbf{x}_k, y_k\}_{k=1}^q$   
 9:   Add the batch to the training set  $X^{t+1} = X^t \cup S^t$   
 10:   Remove the batch from the pool of candidates  
 $U^{t+1} = U^t \setminus S^t$   
 11:    $t = t + 1$   
 12: **until** test accuracy  $> \theta_{max}$

**Classification uncertainty in RF classifiers.** RF classifiers allow a committee approach for the estimation of unlabeled sample uncertainty [13]: assume that we have built a committee of  $k$  base classifiers, i.e. a RF with  $k$  trees. The output of the committee members provides  $k$  labels for each candidate sample  $\mathbf{x}_i \in U$ . The data sample class label is provided by the majority voting. Our heuristic is that the standard deviation  $\sigma(\mathbf{x}_i)$  of the class labels is the measure of the classification uncertainty of  $\mathbf{x}_i$ . Let us consider an ordering of the pool of candidates  $U^* = \{\mathbf{x}_j\}_{j=l+1}^{l+u}$ , where  $\sigma(\mathbf{x}_{j_l}) > \sigma(\mathbf{x}_{j_{l+1}})$ . The standard deviation query-by-bagging heuristic selection of samples to be added to the train set is stated as the following selection:

$$S^t = \{\mathbf{x}_{j_m}\}_{m=1}^q \quad (1)$$

Standard deviation of predicted class labels is a natural multi-class heuristic measure of classification uncertainty. A candidate sample for which all the classifiers in the committee agree has a zero prediction standard deviation, thus its inclusion in the training set does not bring additional information. On the contrary, a candidate with maximum disagreement between the classifiers results in maximum standard deviation, and its inclusion will be highly beneficial.

## 2.3. Active learning for image segmentation

The goal is to classify image voxels into two classes, the target region and the background [29]. The Active Learning system returns to the user, the unlabeled voxels, whose classification outcome is most uncertain with the current classifier. After manual labeling by the user, voxels are included into the training set and the classifier is trained again [19]. The feature vector associated to each pixel for its classification is computed using information from its neighboring voxels, applying linear and/or non-linear filtering. In this paper the features initially associated with CTA voxels are: its coordinates in the data domain grid, the voxel intensity, the mean, variance, maximum and minimum of the voxel neighborhood, for different values of the neighborhood radius (1,2,4...2<sup>n</sup>). The definition of these features increases the data dimensionality and the complexity of the classifiers built on them. A first step towards the practical feasibility of the approach (meaning affordable computation times) is the selection of the most informative features, reducing data dimensionality. In this paper feature selection is done on the basis of the variable importance defined in the following subsection.

## 2.4. Feature selection based on variable importance

Let us denote  $x_j$  the  $j$ -th feature of the feature vector. The RF variable importance of  $x_j$  is defined as follows. For each tree  $h(\mathbf{x}; \psi_t)$  of the forest, consider the associated out-of-box  $OOB_t$  dataset [25] constituted by data samples not included in the bootstrap sample used to construct  $h(\mathbf{x}; \psi_t)$ . Denote  $e_{OOB_t}$  the error corresponding to the miss-classification rate for classification of the single tree  $h(\mathbf{x}; \psi_t)$  on the  $OOB_t$  dataset.

Next, randomly permute the values of  $x_j$  in  $OOB_t$  to get a perturbed sample denoted by  $\tilde{OOB}_t$  and compute  $e_{\tilde{OOB}_t}$ , the error

of  $h(\mathbf{x}; \psi_t)$  on the perturbed sample. Variable importance of  $x_j$  is then equal to

$$VI(x_j) = \frac{1}{T} \sum_t \left( e_{\text{OOB}_t} - e_{\text{OOB}_t^j} \right), \quad (2)$$

where  $T$  denotes the number of trees of the RF. For the feature selection, we order the features by decreasing value of  $VI(x_j)$ , we compute the total accumulated variable importance of the features,

$$TI = \sum_{j=1}^N VI(x_j). \quad (3)$$

### 3. Experimental setup

**Datasets.** We have performed computational experiments on 8 CTA datasets to test the proposed Active Learning based image classification approach. Each dataset consists of real human contrast-enhanced datasets of the abdominal area with  $512 \times 512$  voxel resolution on each slice. Each dataset consists of 216–560 slices and  $0.887 \times 0.887 \times 1$  mm spatial resolution corresponding to patients with Abdominal Aortic Aneurysm. The dataset collection shows a wide diversity of sizes and locations of the thrombus. Some of them have metal streaking artifacts due to the stent graft placement. Ground truth segmentations of the thrombus for each dataset, that

**Table 1**  
Feature importance ranking for the first 10 features selected, specifying the operator used (O), neighborhood radius ( $R$ ) and the variable importance (VI). Max, Med, GA correspond to maximum, median and Gaussian weighted average, respectively.

Parameter	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
O	Max	Max	Max	Med	Max	Max	Med	Med	Med	GA
R	16	4	8	8	2	1	1	4	16	4
VI	1.277	0.953	0.9531	0.803	0.762	0.759	0.741	0.740	0.732	0.725

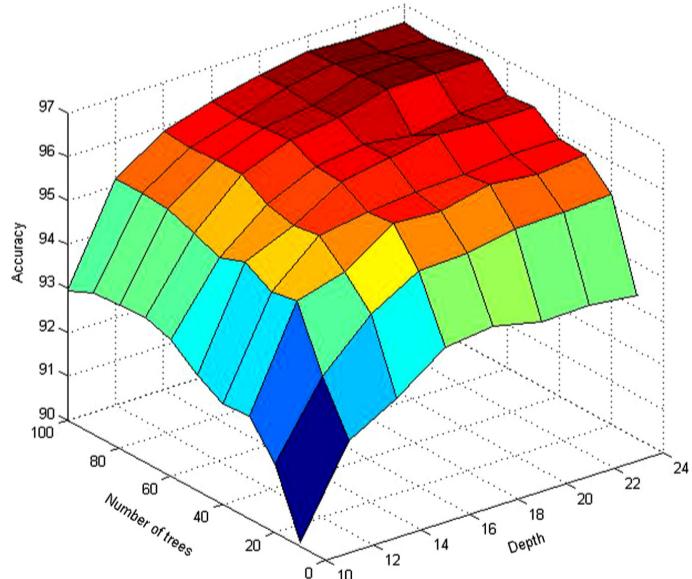
simulate a human oracle providing the labels for the voxels, were obtained manually by a clinical radiologist.

**Feature selection.** Starting with a set of 24 features computed as the maximum, minimum, median and Gaussian weighted average of 2D voxel neighborhoods of increasing radius  $R = 2^k; k = 0, \dots, 4$ . We train the RF with one central slice to perform feature selection. We compute the variable importances and total importance (Eq. (3)), discarding features falling behind the 95% of the TI value in the ordered list of features. The current list of features selected is given in Table 1. We select 10 features.

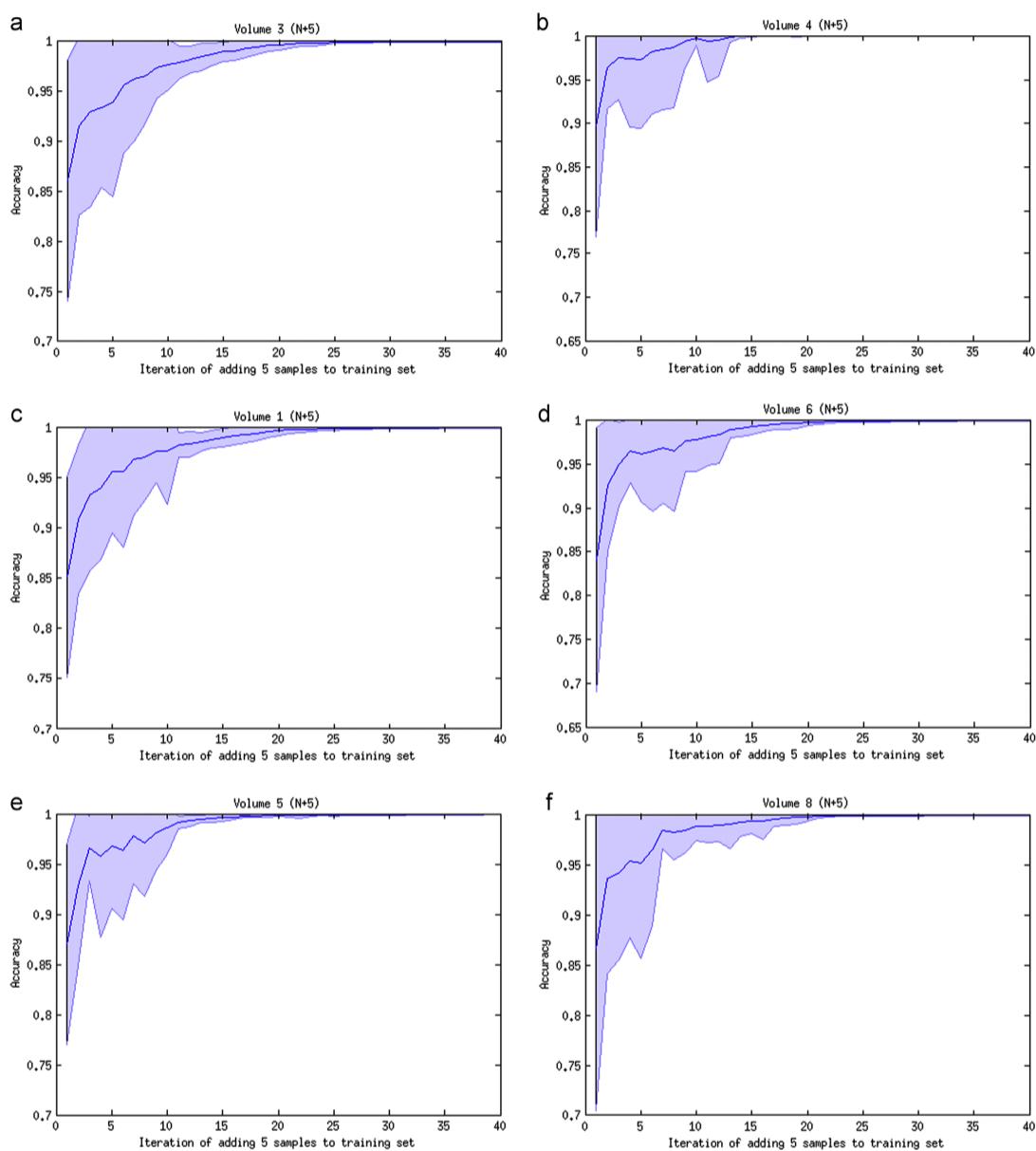
**Parameter tuning.** We train the RF classifier with a single slice to test the sensitivity of the forest parameters: the number of the trees  $T$  and their depth  $D$ . We vary  $T$  from 10 to 100 in steps of 10 and  $D$  from 10 to 24 in steps of 2. 3-fold cross-validation is carried out for each parameter combination. In Fig. 3 segmentation results are evaluated for each combination of RF parameters to compute the response surface of classification accuracy due to variations in RF parameters. The surface corresponds to the average accuracy in the 3-fold cross validation. The figure shows that, for a fixed depth, increasing the number of trees leads to a more accurate classification. The increase in performance stabilizes around 80 trees and depth equals to 20. The number of variables randomly selected for each decision tree in the RF is heuristically set to half the number of features  $d = 5$ . The stopping threshold for the Active Learning has been set to  $\theta_{\max} = 0.99$ . In our experiments we always reach this degree of accuracy on the training set.

Once we get the optimal parameters and feature set, we have designed two different experiments to test our method in the patient CT volumes as illustrated in Fig. 2:

- (1) Independent slice classifier: we build a RF classifier for each slice of the volume, and we test it with the corresponding slice.
- (2) Generalization of a single slice classifier: we build just one RF classifier from the data of the central slice of the aneurysm, and we test its generalization with the remaining slices of the CT volume.



**Fig. 3.** Accuracy of the segmentation as a function of the RF parameters: number of trees and depth.



**Fig. 4.** Average and standard deviation interval of the accuracy of Active Learning at each slice versus size of the additions to the train dataset for 6 CTA volumes of AAA patients. Active learning performed adding the 5 most uncertain unlabeled samples to the train dataset at each iteration. Initial train dataset size is 20 voxels. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

**Validation.** The performance measure of the experiments is the classification accuracy. For the first experiment, the independent slice classifiers, we average the classification accuracy obtained on each slice at each iteration of the Active Learning process. That is, we obtain the average evolution of the classification accuracy, with a corresponding variance value. For the second experiment, we plot the accuracy obtained when applying the classifier on each slice.

#### 4. Experimental results

**Fig. 4** shows the performance of the Active Learning based image segmentation algorithm for 6 CTA volumes of AAA patients. We plot the average accuracy of the RF classifiers built at each iteration of the Active Learning process on each slice, the abscissas correspond to the number of voxels being added to the training

set, 5 voxels per iteration. It can be appreciated that all plots reach convergence to 0.99 accuracy after four iterations (20 voxels added). The variance of the classification accuracy is represented by a blue region around the mean, upper and lower limits correspond to adding and subtracting twice the standard deviation. It can be appreciated that the variance drops considerably after the fourth iteration, and it is negligible after the fifth iteration. This means that the process is extremely robust.

The experiment reported in Fig. 4 is equivalent to asking the human operator to follow the Active Learning procedure on each slice, which amounts to one to two hundred slices taking into account only the CTA data imaging the AAA thrombus. Then we run the experiment in which we test all the slices with the classifier built on the training set corresponding to the image features of one single slice. If the RF classifier obtained on one slice can be applied to the remaining slices without loss of accuracy, the human operator would only need to perform once the Active Learning process to obtain the whole volume segmentation. Fig. 5 shows the overlaid plots of the accuracy obtained at each CTA volume slice applying the RF classifier trained on the thrombus central slice for each of the 8 CTA volumes in the experiment. The abscissa's zero value corresponds to the central slice, the negative abscissa values correspond to slices above the central slice, the positive values correspond to slices below the central slice. There is some variability of the plots' span, due to the different sizes of the thrombus in each patient. As can be expected, the drop in classification accuracy is symmetric. The generalization results are

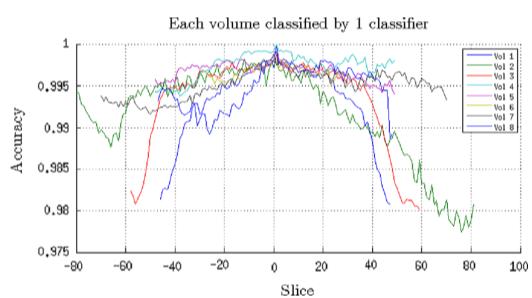
very good: the worst accuracy is above 0.98 in almost all cases. These results come with a hundred-fold reduction of learning complexity involving computer and human operator time.

A 3D volume rendering of the Aorta's lumen (green) and thrombus (red) of one patient is shown in Fig. 6 in three situations. Fig. 6(a) shows the rendering of the ground truth given by volume manual segmentation. Fig. 6(b) shows the result of the segmentation based on the Active Learning performed on each slice. The high accuracy of the segmentation is evident from the comparison with the rendering of the manual segmentation. Fig. 6(c) shows the result of the segmentation based on the RF classifier built from the thrombus' central slice. The structure of the thrombus is well delineated, most errors come from the identification as thrombus of separate structures which can be easily removed by morphological operations.

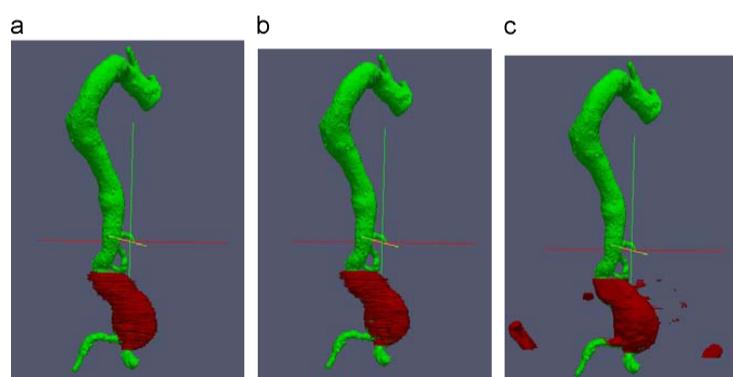
## 5. Conclusion and future works

In this paper we propose an Active Learning approach for training RF classifiers for the segmentation of the thrombus in EVAR follow-up CTA volumes. First, we apply a feature selection approach based on variable importance measured as the classification error induced by variable value randomization. The selected image features were used for the Active Learning of RF classifiers on a slice-by-slice basis. Almost perfect classification accuracy is obtained in all slices and CTA volumes tested after four iterations adding five voxels at each iteration. In a human operator based setting, this means that the segmentation of each slice can be obtained with a minimal human intervention. To obtain an assessment of the impact of further human intervention reduction, we compute the volume classification obtained applying the RF classifier trained on one slice. This approach means a hundred-fold reduction of the human intervention. Accuracy on the test slices is above 0.98 in almost all cases. Besides, most false positive errors introduced correspond to separate structures that are disconnected from the thrombus region, which can be easily filtered by conditional morphological operators. Active Learning of RF classifiers allows to obtain near optimal segmentation results with minimal intervention of the human operator in a very fast semi-automated process.

We plan to test the approach on new CTA data provided by local clinicians, as well as assessing its value through live active learning experiments with expert radiologists and/or clinicians, needing a convenient user interface and data visualization tool. The approach is going to be submitted to the consideration by



**Fig. 5.** Generalization assessment experiment. Overlayed plots of the classification accuracy obtained in each slice when the classifier learned on the central slice is applied to the remaining slices. Each plot corresponds to one experimental dataset.



**Fig. 6.** Volume rendering of aortic lumen (green) and AAA thrombus (red) obtained from the segmentation of one CTA volume. (a) Manual segmentation of the ground truth, (b) result of Active Learning RF classifiers detecting the thrombus in each slice, (c) result of generalization of the Active Learning RF classifier on the central slice applied to the remaining slices. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

clinicians and expert radiologists of the Hospital Donostia in order to assess its usefulness in the clinical environment.

Further work will address the extension of this approach to a direct volume processing, that is, dealing with the whole volume instead of 2D slices. From the computational point of view, this implies working on 3D neighborhoods and the classification of the entire volume at each Active Learning iteration, though in all great computational savings can be achieved. For a practical clinical tool, this implies the appropriate 3D visualizations of data classification uncertainty, and the ability to navigate searching for the most uncertain voxels. The approach can be generalized to multi-class classifiers, that can be used to segment the Aorta's lumen and the aneurysm thrombus as well as other surrounding anatomical structures that can be of interest for further studies.

## Acknowledgments

This work has been supported by the MICINN under Grant TIN2011-23823. The CTA images used in the computational experiments were provided by the group of Leo Joskowicz [8] of the Mount Sinai School of Medicine, New York, NY.

## References

- [1] I. Macía, M. Graña, C. Paloc, Knowledge management in image-based analysis of blood vessel structures, *Knowledge and Information Systems* 30 (2) (2012) 457–491.
- [2] D. Lesage, E.D. Angelini, I. Bloch, G. Funka-Lea, A review of 3d vessel lumen segmentation techniques: models features and extraction schemes, *Medical Image Analysis* 13 (6) (2009) 819–845.
- [3] M. de Brujin, B. van Ginneken, M.A. Viergever, W.J. Niessen, Interactive segmentation of abdominal aortic aneurysms in CTA images, *Medical Image Analysis* 8 (2) (2004) 127–138.
- [4] S. Olabarriaga, J. Rouet, M. Fradkin, M. Breeuwer, W. Niessen, Segmentation of thrombus in abdominal aortic aneurysms from CTA with nonparametric statistical grey level appearance modeling, *IEEE Transactions On Medical Imaging* 24 (4) (2005) 477–485.
- [5] F. Zhuge, G.D. Rubin, S.H. Sun, S. Napel, An abdominal aortic aneurysm segmentation method: level set with region and statistical information, *Medical Physics* 33 (5) (2006) 1440–1453.
- [6] S. Demirci, G. Lejeune, N. Navab, Hybrid deformable model for aneurysm segmentation, in: ISBI'09, 2009, pp. 33–36.
- [7] I. Macía, M. Graña, J. Maiora, C. Paloc, M. de Blas, Detection of type II endoleaks in abdominal aortic aneurysms after endovascular repair, *Computers in Biology and Medicine* 41 (10) (2011) 871–880.
- [8] M. Freeman, S.J. Esses, L. Joskowicz, J. Sosna, An iterative model-constraint graph-cut algorithm for abdominal aortic aneurysm thrombus segmentation, in: Proceedings of the 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'10), IEEE, Rotterdam, The Netherlands, 2010, pp. 672–675.
- [9] D. Cohn, L. Atlas, R. Ladner, Improving generalization with active learning, *Machine Learning* 15 (1994) 201–221.
- [10] B. Settles, *Active Learning Literature Survey*, vol. 15(2), 2010, Sciences, New York.
- [11] P. Mitra, B.U. Shankar, S.K. Pal, Segmentation of multispectral remote sensing images using active support vector machines, *Pattern Recognition Letters* 25 (9) (2004) 1067–1074.
- [12] D. Tuia, M. Volpi, L. Coppi, M. Kanevski, J. Munoz-Mari, A survey of active learning algorithms for supervised remote sensing image classification, *IEEE Journal of Selected Topics in Signal Processing* 5 (3) (2011) 606–616.
- [13] D. Tuia, E. Pasolini, W. Emery, Using active learning to adapt remote sensing image classifiers, *Remote Sensing of Environment* 115 (9) (2011) 2232–2242.
- [14] S.C.H. Hoi, R. Jin, J. Zhu, M.R. Lyu, Semisupervised SVM batch mode active learning with applications to image retrieval, *ACM Transactions on Information Systems* 27 (3) (2009) 1–29.
- [15] J. Iglesias, E. Konukoglu, A. Montillo, Z. Tu, A. Criminisi, Combining generative and discriminative models for semantic segmentation of ct scans via active learning, in: *Information Processing in Medical Imaging*, Springer, 2011, pp. 25–36.
- [16] Y. Tao, Z. Peng, B. Jian, J. Xuan, A. Krishnan, X. Sean Zhou, Robust learning-based annotation of medical radiographs, in: *Medical Content-Based Retrieval for Clinical Decision Support*, vol. 5853 of Lecture Notes in Computer Science, Springer Berlin/Heidelberg, 2010, pp. 77–88.
- [17] I. Barandiaran, C. Paloc, M. Graña, Real-time optical markerless tracking for augmented reality applications, *Journal of Real-Time Image Processing* 5 (2010) 129–138.
- [18] V. Lempitsky, M. Verhoek, J. Noble, A. Blake, Random forest classification for automatic delineation of myocardium in real-time 3d echocardiography, *Functional Imaging and Modeling of the Heart* (2009) 447–456.
- [19] E. Geremia, B. Menze, O. Clatz, E. Konukoglu, A. Criminisi, N. Ayache, Spatial decision forests for MS lesion segmentation in multi-channel MR images, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* 2010, 2010, pp. 111–118.
- [20] Z. Yi, A. Criminisi, J. Shotton, A. Blake, Discriminative, semantic segmentation of brain tissue in MR images, in: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2009*, 2009, pp. 558–565.
- [21] A. Criminisi, J. Shotton, S. Bucciarelli, Decision forests with long-range spatial context for organ localization in ct volumes, in: *MICCAI Workshop on Probabilistic Models for Medical Image Analysis*, 2009.
- [22] A. Criminisi, J. Shotton, D. Robertson, E. Konukoglu, Regression forests for efficient anatomy detection and localization in CT studies, *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging* (2011) 106–117.
- [23] J. Maiora, M. Graña, Abdominal CTA image analysis through active learning and decision random forests: application to AAA segmentation, in: *International Joint Conference on Neural Networks (IJCNN)*, 2012.
- [24] A. Abraham, E. Corchado, J. Corchado, Hybrid learning machines, *Neurocomputing* 72 (13–15) (2009) 2729–2730.
- [25] L. Breiman, Random forests, *Machine learning* 45 (1) (2001) 5–32.
- [26] L. Breiman, Bagging predictors, *Machine learning* 24 (2) (1996) 123–140.
- [27] Y. Amit, D. Geman, Shape quantization and recognition with randomized trees, *Neural computation* 9 (7) (1997) 1545–1588.
- [28] T. Ho, The random subspace method for constructing decision forests, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (8) (1998) 832–844.
- [29] M. Yaqub, M. Javid, C. Cooper, J. Noble, Improving the classification accuracy of the classic RF method by intelligent feature selection and weighted voting of trees with application to medical image segmentation, *Machine Learning in Medical Imaging* (2011) 184–192.
- [30] E. Corchado, A. Abraham, A. de Carvalho, Hybrid intelligent algorithms and applications, *Information Sciences: An International Journal* 180 (2010) 2633–2634.
- [31] E. Corchado, M. Graña, M. Wozniak, New trends and applications on hybrid artificial intelligence systems, *Neurocomputing* 75 (2012) 61–63.



**Josu Maiora** received Ph.D. degree from Universidad del País Vasco (UPV/EHU) in 2013, in electrical engineering. He is assistant professor at the Electrical Engineering department of the UPV/EHU. He belongs to the Computational Intelligence Group of the UPV/EHU. His current interests are medical image processing and bio-inspired machine learning.



**Borja Ayerdi** received his M.Sc. in 2011 from Universidad del País Vasco. He is currently pursuing his Ph.D. with a grant from the Departamento de Educación of the Gobierno Vasco. His research interests cover the application of bioinspired Machine Learning approaches to image processing, including remote sensing and medical images.



**Manuel Graña Romay** (M'94) received the M.Sc. and Ph.D. degrees from Universidad del País Vasco (UPV/EHU), Donostia, Spain, in 1982 and 1989, respectively, both in computer science. His current position is a Full Profesor (Catedrático de Universidad) with the Computer Science and Artificial Intelligence Department of the Universidad del País Vasco (UPV/EHU). He is the head of the Computational Intelligence Group (Grupo de Inteligencia Computacional). His current research interests are in applications of computational intelligence to multicomponent robotic systems, medical image in the neurosciences, multimodal human computer interaction, remote sensing image processing, content based image retrieval, lattice computing, semantic modelling, data processing, classification, and data mining.