Enhancing Active Learning Computed Tomography Image Segmentation with Domain Knowledge

Borja Ayerdi, Josu Maiora, Manuel Graña
UPV/EHU

September 12, 2013
HAIS 2013

Enhancing Active Learning Computed Tomography Image Segmentation with Domain Knowledge

- Introduction
- Methods
    - Random Forest
    - Active learning
    - Domain Knowlegde
- Experimental setup
- Experimental Results
- Conclusion and future works

Introduction

- Several segmentation methods for vascular structures have been developed but segmentation of abdominal images, specially the AAA thrombus, is still a challenging task.

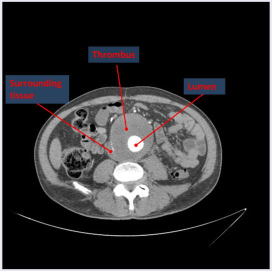
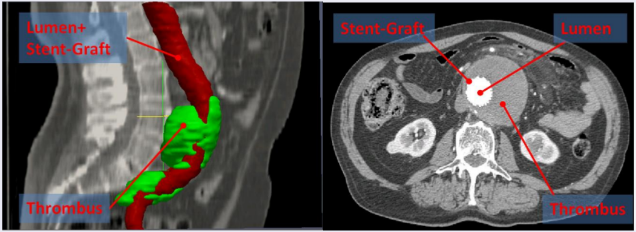


Figure : Slice example

Introduction

- Abdominal Aortic Aneurysms (AAA) is a focal dilation of the aorta in the abdominal region.
- Endovascular prosthesis for aneurysm repair (EVAR): effective technique to reduce the pressure and rupture risk of aneurysm.

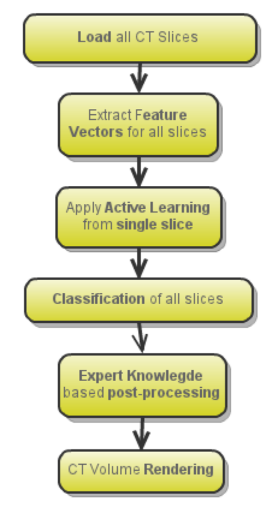
Introduction



Introduction: Our approach

- Initially a **multiclass classification** of voxel samples
  - Lumen.
  - Thrombus.
  - Bones .
  - Background.
- **Simplify to binary classification**.
  - Aortic thrombus. (Lumen + Thrombus)
  - Background. (Bones + Background)
- We perform the classification with a supervised method: **Random Forest**.
- We build the training data in an iterative **active learning** process .
- **Expert knowlegde** based **post-processing**.

Introduction: Our approach



Methods: Active learning

So, how are we building the training set?

- Small.
- Focused on those pixels effectively improving the performance of the model .
  1. The model returns to the user the pixels whose classification outcome is the most **uncertain**.
  2. Accurate **labeling** by the user.
  3. Pixels are **included** into the training set.

Methods: Decision tree

- The **decision trees** are constructed by recursively partitioning the data space. We define a separation based on a single attribute recursively for each node of the tree, from the root to the leaves with some discretion.

Methods: Random Forest

- The **Random Forest** are ensembles of decision trees where each decision tree is constructed from a subset of the data set to train and a random subset of input variables.
- Try to get better classification results combining **weak classifiers** and **diversified**.

Methods: Domain Knowlegde

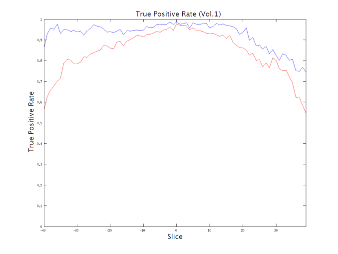
- Allows to post-processing the results of the classification in order to remove spurious detections.
  - At each axial slice, the thrombus is composed of only one **connected component**. We can remove all connected components disconnected from the one that is more likely to be the thrombus, which is identified by the following rules.
  - Thrombus has a roughly **circular shape** in any axial cut of the volume.
  - In succesive slices moving away from the thrombus middle slice the **radius** of the thrombus region **decreases**.
  - In successive axial slices, the thrombus **region overlap** is large (between 80% and 90% of the area).
  - The 2D coordinates of the **centroid of the thrombus region** have a small (smooth) **variation** between successive slices.

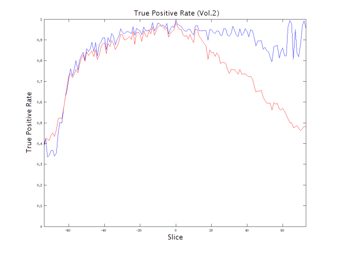
Experimental setup

- **6 real** human contrast-enhanced **datasets** of the abdominal area.
- **Two-class segmentation** problem.
- We get the optimal parameters and feature set for the RF. The increase in performance stabilizes around number of trees=80 and depth=20.
- The performance measure results of the experiments are the post-processing average **True Positive Rate** (TPR).

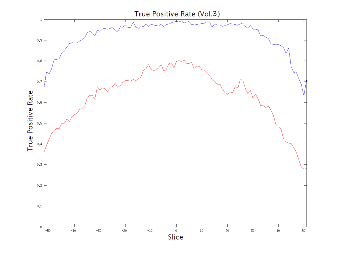
Experimental results

- **Red curves** corresponds to **RF results trained with Active Learning**, and **blue curves** to the **Domain Knowledge post-processing**.

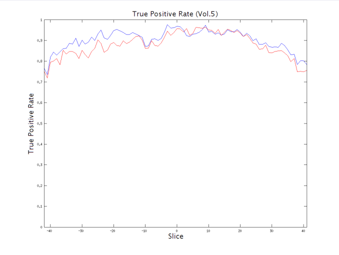


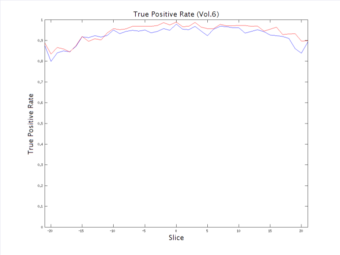


Experimental results



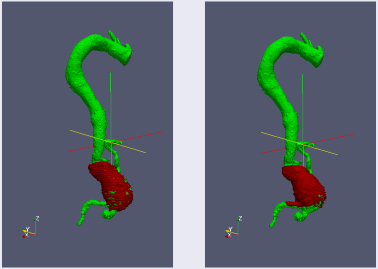
Experimental results





Volume rendering

- Volume rendering of aortic lumen (green) and thrombus (red) obtained from the segmentation of one CT volume. (a) manual segmentation of the ground truth, (b) result of post-processing rules.



Conclusions

- With **active learning** we build quickly and efficiently specific classifiers trained on the volume to be segmented with a minimum number of training samples requiring labeling.
- This results can be improved by using specific **Domain Knowledge** of the structure being segmented.
- We have transformed such rules in **heuristic post-processing rules**.
- The results show that in some cases, the application of such heuristics can provide dramatic **increase in performance, maintaining all the advantages of the Active Learning approach.**

Conclusions

- Thank you for your attention!