

# AUTOMATIC VENTRICLE DETECTION IN COMPUTED TOMOGRAPHY PULMONARY ANGIOGRAPHY

*Sara Rodríguez-López,<sup>1</sup> Daniel Jimenez-Carretero,<sup>1</sup> Raúl San José Estépar,<sup>2</sup> Eduardo Fraile Moreno,<sup>4</sup> Kanako K. Kumamaru,<sup>3</sup> Frank J. Rybicki,<sup>3</sup> Maria J. Ledesma-Carbayo,<sup>1</sup> Germán González<sup>2,5</sup>*

<sup>1</sup>Biomedical Image Technologies. Universidad Politécnica de Madrid, Madrid, Spain & CIBER-BBN

<sup>2</sup>Surgical Planning Laboratory, Brigham and Womens Hospital, Boston, MA, USA

<sup>3</sup>Applied Imaging Science Laboratory. Brigham and Womens Hospital, Boston, MA, USA

<sup>4</sup>Unidad Central de Radiodiagnóstico, Madrid, Spain

<sup>5</sup>Madrid-MIT M+Visión Consortium. Massachusetts Institute of Technology. Cambridge, MA, USA

## ABSTRACT

Automated medical image analysis requires methods to localize anatomic structures in the presence of normal inter-patient variability, pathology, and the different protocols used to acquire images for different clinical settings. Recent advances have improved object detection in the context of natural images, but they have not been adapted to the 3D context of medical images. In this paper we present a 2.5D object detector designed to locate, without any user interaction, the left and right heart ventricles in Computed Tomography Pulmonary Angiography (CTPA) images. A 2D object detector is trained to find ventricles on axial slices. Those detections are automatically clustered according to their size and position. The cluster with highest score, representing the 3D location of the ventricle, is then selected. The proposed method is validated in 403 CTPA studies obtained in patients with clinically suspected pulmonary embolism. Both ventricles are properly detected in 94.7% of the cases. The proposed method is very generic and can be easily adapted to detect other structures in medical images.

**Index Terms**— Heart Ventricle, Detection, CTPA, HOG

## 1. INTRODUCTION

Object detection is an essential preprocessing step for fully automatic segmentation or registration algorithms, as well as for the automatic computation of image-based biomarkers and computer-aided detection (CAD) systems [1]. The growing number of digitized images and distributed repositories poses a new demand of this type of methods, both for application in large clinical studies as well as for assisting specialists to perform their reporting faster and more efficiently. Robust identification of anatomical structures

This project has been financially supported by the Comunidad de Madrid, Spain through the M+Visión Consortium. Daniel Jimenez was supported by an FPU grant by the Spain's Ministry of Education.

in biomedical imaging is particularly challenging, since structures and even organs can vary greatly in appearance.

Computed Tomography (CT) is the reference standard imaging modality for the diagnosis and exclusion of Pulmonary Embolism (PE). The study is commonly referred to as CT Pulmonary Angiography (CTPA). Iodinated contrast is used to enhance the pulmonary arteries and reveal emboli as filling defects within the pulmonary arteries. CT acquires data volumetrically; resulting 3D images include the entire heart, and thus can assess the status of the right ventricle. This is particularly important for patients with acute PE because right ventricular strain, characterized by an enlarged right ventricular cavity when compared to that of the left, predicts a poor prognosis and is used to decide those patients who will benefit from a more aggressive treatment plan. Such aggressive treatments are associated with life-threatening complications, emphasizing the need for accurate reporting of the sizes of both cardiac ventricles when a patient is diagnosed with acute PE. [2].

Automatic ventricle detection in CTPA is challenging, due to inter-subject variability and inconsistency in image quality from factors related to either the patient (e.g. obesity) or the CT acquisition. Pathology such as atelectasis and cardiomyopathy make the appearance of the heart variable, as does differences in CT settings and variation in the timing of the iodinated contrast delivery. Also, while almost all CT studies of the heart use electrocardiogram gating to minimize motion, CTPA are done without gating to simplify the acquisition and reduce patient radiation exposure. The summation of these effects cause spatial blur and high contrast variability within the heart. Furthermore, the cranio-caudal field of view varies widely, e.g. including/excluding cervical or abdominal information. All of these factors compromise the ability of simple heuristics to locate the ventricles and justify the need of the following more complex solution for clinical implementation

Object recognition has been widely studied in computer

vision, with techniques that range from simple models, such as rigid templates [3] and bag-of-features [4], to richer models such as pictorial structures [5]. These techniques are designed to be robust to changes on shape and location of the objects detected. Some of these methods have been previously applied to the detection of anatomical structures in medical images [6, 7]. However, to our knowledge, there is no previous study focused on the detection of heart chambers from CTPA data.

## 2. METHOD

In this work we extend one of the best performing object detectors [8] to detect 3D objects in medical images by combining it with a clustering mechanism that selects the subset of the 2D detections that form the 3D ventricles. The right and left ventricles are detected independently in each CTPA scan, since they have different image characteristics and may be visible at different axial slices. For each ventricle, we follow the workflow described in Fig. 1. The detailed workflow below focuses on left ventricle detection; the process is identical for the right ventricle.

### 2.1. Training Dataset

A training dataset of 80 CTPA axial images from 40 patients of Unidad Central de Radiodiagnóstico in Madrid, Spain was assembled from 3D CTPA images. The positive training set was the axial slices of each study where the ventricle had the highest diameter. For each slice a bounding box that fits the ventricle was defined. The negative training was 40 images, one of each patient, in which neither ventricle is visible. The training dataset thus consists of  $P = 40$  positive axial slices per ventricle and  $N = 40$  negative axial slices.

### 2.2. Model Creation

The left ventricular model was learned using Histograms of Oriented Gradients (HOG) as features, efficient matching algorithms for deformable part-based models (pictorial structures) and discriminative learning with latent variables (latent SVM), as described in [8].

The ventricles in the positive training data have different aspect ratios according to their shape. In some patients the ventricles are wide structures, with an aspect ratio close or above one, while other patients have thinner ventricles. In order to detect properly both cases, each final ventricle model was created by merging two different models obtained for different aspect ratios: 0.8 and 1. Those aspect ratios were chosen from the positive training set. The models are firstly trained with a random subset of locations of the negative images as negative training samples. The learned model is later evaluated in the negative training images. Then the model is retrained with the same method but using as

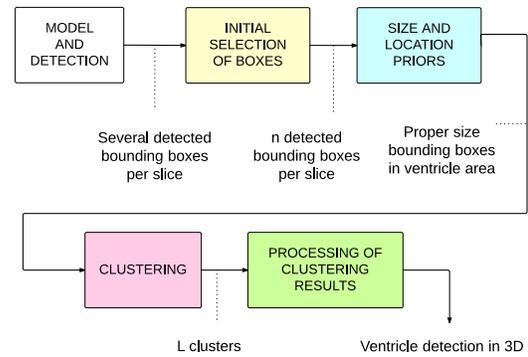


Fig. 1. Method workflow.

negative samples a combination of random locations on the negative training images and false positive detections found on the negative training images (“hard negatives”). These hard negatives usually correspond to areas around the trachea, the superior region of heart and near aorta or liver region. At the end of the process, there are two single models, one for each ventricle.

### 2.3. Detection

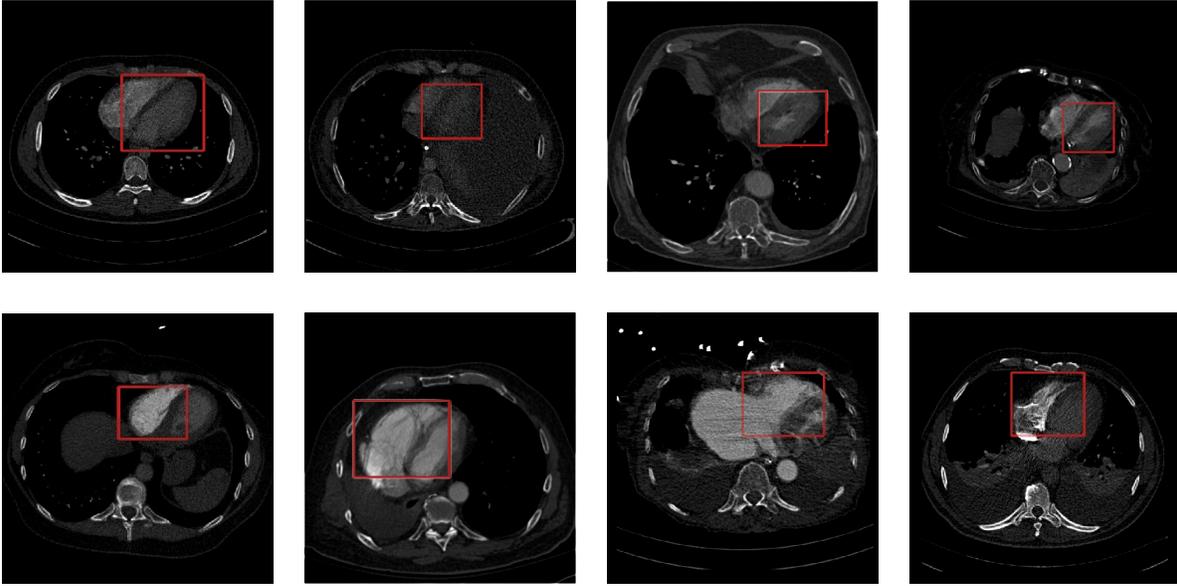
For each axial slice, a Gaussian pyramid is created and the right and left ventricles are detected separately using their respective models. Since the detection algorithm was developed for 2D image detection, each axial slice is processed independently. The method returns several detection boxes per slice, each with an associated score  $sc(i)$  representing the goodness of fit of the image to the model. Almost all axial slices had at least one detection. The bounding boxes were not only in the ventricles, but also in other structures not related to the heart. For such reason a post-processing is needed, to find the set of boxes, one per slice, that better describes the heart.

### 2.4. Post-Processing

Four main steps are performed to obtain the 3D detection of the ventricle from the set of 2D detections: initial selection of boxes, size and location constraints, clustering and a final processing of clustering results. The sequence of steps are shown in Fig. 1.

#### Initial selection of boxes

A selection of boxes is needed to reduce processing time. Selecting the box with highest scoring in each slice is insufficient since the highest scored bounding box does not always correspond to the ventricle. Consequently, for each axial slice we select the  $n$  bounding boxes with highest score.  $n$  was adjusted to 25 after few tests as a good tradeoff between correct detected boxes and processing time.



**Fig. 2.** Examples of successful ventricle detections imaged at the middle axial slice of the detection. Each image is taken from a different patient. Please note inter-patient variability. Top row: left ventricle. Bottom row: right ventricle. The first column represents images with high contrast and clearly delimited chambers. The second column shows patients with collapsed lungs. The third column shows hearts whose image properties differ from the rest of the dataset: top row: abnormal distance between the heart and the sternum, bottom row: dilated right atrium. The fourth column represents variability in contrast and location.

### Size and location constraints

Size priors are established in order to remove bounding boxes that due to size (extremely big or small) can not be part of a ventricle. A minimum and a maximum value are established for the width and the height of detected boxes. These values are different for right and left ventricles, and are computed from the bounding boxes of the training dataset. Location priors are established to remove bounding boxes that are far away from the heart. In order to deal with inter-subject variability, we position the location prior relative to a landmark that is very constant for all patients and CTPAs. The bone tissues were the most stable structure in the training dataset. We therefore select as reference point for distances computation the center of mass of the bone-thresholded CTPA projection in the Z axis. We use the distances between bounding boxes centers and CTPA reference points of the training dataset to establish the maximum and minimum distance allowed between the center of the ventricular candidate detected box and the reference point of the CTPA.

### Clustering

Some remaining bounding boxes are placed in the heart region, but instead of being in the ventricles, they are placed near the aorta and the superior vena cava. A clustering method is used to assign bounding boxes with different features to different groups. The clustering method used was mean shift

clustering, a non-parametric feature space analysis technique that does not require prior knowledge of the number of clusters. The features used for clustering are the position of the center and the area of the boxes, thereby the clusters consists of groups boxes with similar size and located close to each other.

### Cluster selection

The clustering process returns  $L$  clusters, each one including bounding boxes with similar features. In order to select a single cluster, a score is assigned to each cluster. The selected cluster  $\hat{l}$  is the one that maximizes:

$$\hat{l} = \underset{1 \leq l \leq L}{\operatorname{argmax}} \left( \frac{\sum_{i=1}^{nb(l)} sc(i)}{nb(l) + 2} \right) \quad (1)$$

where  $nb(l)$  stands for the number of boxes that are in the cluster and  $sc(i)$  is the detection score associated with the box  $i$ . The cluster score therefore depends on the individual scores of each box and on the number of boxes of the cluster. Equation 1 penalizes clusters made of few boxes that have high detection score and encourages clusters with high number of boxes with high detection score. This criterion is very useful because mistaken boxes are often grouped in clusters with low number of boxes. Consequently, despite having high individual scores, such clusters will receive a low score and it will not be selected. The output of the algorithm is the collection of 2D bounding boxes associated with cluster

$\hat{l}$ , which define the 3D region where the ventricle is located.

### 3. RESULTS

We test the algorithm in two datasets. The first one consists of 199 CTPA scans positive for PE obtained from the Applied Imaging Science Laboratory, Department of Radiology, of Brigham and Women’s Hospital in Boston, USA. A description of the population study and other details of this database can be found in [9]. The second dataset consists of 204 CTPA studies obtained from Unidad Central de Radiodagnóstico in Madrid, Spain. None of the 403 studies was used for training. The algorithm was applied to all 403 patients. The detection was defined as correct if all its 2D detections enclose the ventricle in the axial slice of detection. Two experts jointly examined the results, with no disagreements between them. Table 1 presents the number of correct ventricular detections in the right ventricle, the left ventricle and both ventricles in the same patient.

	BWH		UCR	
RV	192/199	(96.5 %)	194/204	(95.1%)
LV	196/199	(98.5 %)	194/204	(95.1%)
Both	191/199	(96 %)	191/204	(93.6%)

**Table 1.** Correct detections on the datasets.

Detection mistakes occurred when the ventricle is detected on the veins and arteries above the heart (n=16), in the atria (n=3), when the detection is in the ventricle but does not cover it completely (n=4) or when the detection is not placed close to the heart (n=7). On average, the method analyzes a CTPA examination with 500 axial slices in 120 seconds, being most of the time spent in ventricle detection (118s) and only 2 seconds in the post-processing.

### 4. DISCUSSION AND CONCLUSIONS

This paper describes a robust 2.5D method to detect the right and the left heart ventricles in CTPA studies. Both ventricles were correctly detected in 94.7% of 403 CTPA studies acquired from routine clinical practice. Few of the challenging correct detections are shown in Fig. 2. Most of detection errors are located in anatomy with similar contours as the ventricles and in close proximity to the heart. The performance of the method with only 40 training subjects proves that the clustering step and the size and location constraints in particular, are critical to deal with the high variability in shape and appearance of human hearts.

The two minute processing time is compatible with the interpretation time for clinical CTPA workflow. Specifically, automatic segmentation can be concurrent with the radiologists clinical interpretation of the study for the presence or absence of acute PE. For those patients

who have PE, these methods can be used to quantify the right-to-left ventricular (RV/LV) diameter ratio, a proven method for prognosis and an important step for clinical risk stratification [10].

While the method has been tested in database of CTPA images, the workflow could be adapted to localize automatically heart ventricles or other structures generated by other imaging modalities by changing the training dataset and adjusting the size and location priors.

### 5. REFERENCES

- [1] Kunio Doi, “Computer-aided diagnosis in medical imaging: historical review, current status and future potential,” *Computerized medical imaging and graphics*, vol. 31, no. 4-5, pp. 198–211, 2007.
- [2] Vittorio Pengo, Anthonie WA Lensing, et al., “Incidence of chronic thromboembolic pulmonary hypertension after pulmonary embolism,” *New England Journal of Medicine*, vol. 350, no. 22, pp. 2257–2264, 2004.
- [3] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1, pp. 886–893.
- [4] Eric Nowak, Frédéric Jurie, and Bill Triggs, “Sampling strategies for bag-of-features image classification,” in *Computer Vision–ECCV 2006*, pp. 490–503. Springer, 2006.
- [5] Pedro F Felzenszwalb and Daniel P Huttenlocher, “Pictorial structures for object recognition,” *International Journal of Computer Vision*, vol. 61, no. 1, pp. 55–79, 2005.
- [6] Yefeng Zheng, Xiaoguang Lu, Bogdan Georgescu, Arne Littmann, Edgar Mueller, and Dorin Comaniciu, “Automatic left ventricle detection in MRI images using marginal space learning and component-based voting,” in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2009, pp. 725906–725906.
- [7] Marius Erdt, Oliver Knapp, Klaus Drechsler, and Stefan Wesarg, “Region detection in medical images using HOG classifiers and a body landmark network,” in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2013, pp. 867004–867004.
- [8] Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan, “Object detection with discriminatively trained part-based models,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [9] K. K. Kumamaru, A. R. Hunsaker, N. Wake, M. T. Lu, J. Signorelli, A. Bedayat, and F. J. Rybicki, “The variability in prognostic values of right ventricular-to-left ventricular diameter ratios derived from different measurement methods on computed tomography pulmonary angiography: a patient outcome study,” *J Thorac Imaging*, vol. 27, no. 5, pp. 331–336, Sep 2012.
- [10] JH Reid and JT Murchison, “Acute right ventricular dilatation: a new helical CT sign of massive pulmonary embolism,” *Clinical radiology*, vol. 53, no. 9, pp. 694–698, 1998.