J Digit Imaging (2014) 27:794–804
DOI 10.1007/s10278-014-9705-0

Automatic Cardiac Segmentation Using Semantic Information from Random Forests

Dwarikanath Mahapatra

Published online: 4 June 2014

Abstract We propose a fully automated method for segmenting the cardiac right ventricle (RV) from magnetic resonance (MR) images. Given a MR test image, it is first oversegmented into superpixels and each superpixel is analyzed to detect the presence of RV regions using random forest (RF) classifiers. The superpixels containing RV regions constitute the region of interest (ROI) which is used to segment the actual RV. Probability maps are generated for each ROI pixel using a second set of RF classifiers which give the probabilities of each pixel belonging to RV or background. The negative log-likelihood of these maps are used as penalty costs in a graph cut segmentation framework. Low-level features like intensity statistics, texture anisotropy and curvature asymmetry, and high level context features are used at different stages. Smoothness constraints are imposed based on semantic information (importance of each feature to the classification task) derived from the second set of learned RF classifiers. Experimental results show that compared to conventional method our algorithm achieves superior performance due to the inclusion of semantic knowledge and context information.

Keywords Automatic segmentation · MRI · Right ventricle · Graph cut · Semantic information

Introduction

The leading cause of death in the Western world is attributed to cardiovascular diseases [1]. Numerous medical imaging modalities are being used for their diagnosis and treatment, like echocardiography, computed tomography (CT), coronary angiography and magnetic resonance imaging (MRI). MRI has emerged as the preferred diagnostic modality because of its non-invasive nature. It also provides reliable information on morphology, muscle perfusion, tissue viability, and blood flow. These parameters are obtained by segmenting the left ventricle (LV) and right ventricle (RV) from cardiac MR images.

Majority of the works on cardiac segmentation deal with the LV because of its importance in determining physiological parameters. However, segmentation of the RV has recently gained attention because of new findings that confirm the relationship between RV function and a number of cardiac diseases such as heart failure and RV myocardial infarction [2]. Shors et al. [3] show that MRI can provide an accurate quantification of RV mass. In this work we propose a fully automated cardiac RV segmentation method that combines learned image statistics with graph cut optimization to segment the RV endocardium and epicardium.

Manual segmentation of the RV is tedious and prone to intra- and inter-observer variability. This has necessitated the development of computer-aided segmentation algorithms. Although cardiac RV segmentation methods are few, an excellent review of cardiac LV segmentation algorithms is given in [4]. Any automated/semiautomated RV segmentation algorithm has to overcome the challenges of: (1) its complex crescent shape; (2) low resolution and noisy MRI; (3) thinner structure than LV; and (4) presence of papillary muscles.

Previous Work on MRI RV Segmentation

Lapp et al. [5] combine active appearance models (AAMs) and deformable registration for RV segmentation. However, such an approach has difficulty in defining a proper energy function to drive the curve evolution to the boundary because

D. Mahapatra (✉)
Department of Computer Science, Swiss Federal Institute of Technology, CAB E65.1, Universitatstrasse 6, Zurich 8092, Switzerland
e-mail: dmahapatra@gmail.com

of the complex and variable RV shape [6]. Other works have used registration based approaches [7] and active shape models (ASM) with inter profile modeling [8] for RV segmentation.

Ou et al. in [9] used mutual saliency in a multi-atlas based registration method. Zuluaga et al. [10] used a coarse to fine segmentation and label propagation in a multi-label fusion framework. Graph cut (GC)-based methods were used with convex relaxation and distribution matching [11], statistical principal components as shape priors [12] and region merging [13].

Knowledge-Based Cardiac Segmentation

Most cardiac MR images show poor contrast between RV blood pool and the surrounding myocardium wall, thus giving minimal edge information. Additionally, similar intensity distributions in different regions make segmentation of the RV a very challenging task when using only low-level information (e.g., intensity, gradient, etc.). In such a scenario, high level knowledge (e.g., shape priors) is included for segmentation.

Paragios et al. [14] used a signed distance map to incorporate prior shape knowledge in a level set framework for LV segmentation. A probability density function of the shape or appearance from training data was used in level sets [15] and graph cuts [16], while orientation histograms were used as shape priors for graphcut segmentation of the LV [17, 18] and registration information [19, 20]. Ayed et al. employed level sets to match overlap priors of myocardium and blood pool. Mutual context information from the LV and RV was used to segment each other in [21]. Some of the works on LV segmentation also show results for RV segmentation using deformable models [22, 23] and atlas-based methods [24].

Our Contribution

While shape priors significantly increase the segmentation accuracy of cardiac images, they fail when the RV (or LV) has complex deformations that cannot be handled by the shape prior. To overcome such challenges, we propose to use a machine-learning approach which learns image statistics of the object of interest (OOI), e.g., RV blood pool, and the surrounding myocardial wall. This enables us to distinguish the OOI from the non-interesting regions.

Our approach has two stages. First, we identify a region of interest (ROI) that contains the RV and second, differentiate between the RV blood pool, myocardium, and background within this ROI. A given image is first oversegmented into superpixels and each supervoxel is classified using random forest (RF) classifiers for the presence of the RV. The superpixels containing parts of the RV constitute the ROI. A second set of RF classifiers output probability maps for every ROI pixel indicating its likelihood of belonging to blood pool, myocardium or background. The probability maps are integrated into a second order Markov random field (MRF) cost function and the final labels are obtained using graph cut optimization.

RF classifiers are used because: (i) they allow us to extract semantic information after the training step in the form of relative importance of different features to the classification task. This is important in the segmentation of the RV. (ii) RF classifiers allow for a probabilistic interpretation of the classification of test samples which aids in designing an appropriate cost function for segmentation.

This paper makes the following contributions in terms of technical novelty. First, we develop a hierarchical framework using superpixel segmentation and trained RF classifiers to define a ROI containing the RV region. The RF classifiers also discard medically irrelevant regions away from the RV. Our second contribution lies in the use of semantic information to guide the segmentation process. We derive semantic information from the trained RF classifiers of the second stage to quantify the importance of different features in generating probability maps. The importance measures are used to weigh features in the smoothness cost function for greater segmentation accuracy. The rest of the paper is structured as follows: "Methods" describes different parts of our method. We describe our dataset and discuss our results in "Experiments and Results" and conclude with "Conclusion and Future Work".

Methods

Method Overview

An overview of our method is given in Algorithm 1. A ROI containing the RV is automatically determined by superpixel segmentation and their classification by RF classifiers. Probability values of each pixel within the ROI are calculated using a second set of RF classifiers. Intensity information is not used directly as regions in the RV neighborhood have similar intensity. The negative log-likelihood of the probability map is used as the penalty cost in a second-order MRF cost function. Furthermore, the second set of trained RF classifiers provides semantic information about the importance of different features in the classification task. These importance measures are used to weigh each image feature differently in the smoothness cost. The final class labels

are obtained by optimizing the cost function using graph cuts to obtain a regularized solution.

Algorithm 1 MR Cardiac RV Segmentation
Input Image with *N* pixels.
Output Segmented RV endocardium and epicardium
Sequence of Steps:
Superpixel segmentation
Classification of superpixels to get ROI
Generate probability maps of ROI pixels
Calculate penalty cost and smoothness cost of pixels
Obtain final segmentation labels using graph cuts

Random Forest Classifiers

Random forests [25] are increasingly used by many medical applications like cancer classification, tissue segmentation, [26–28], to detect abnormalities in mammograms [29], identify coronary artery stenoses [30], and analyze Crohn's disease [31–33]. They are computationally efficient for large training data, can solve multiclass classification problems, and the learned knowledge can be extracted and interpreted to get a deeper insight into the training procedure. An RF is an ensemble of decision trees, where each tree is typically trained with a different subset of the training set ("bagging"), thereby improving the generalization ability of the classifier. Samples are processed along a path from the root to a leaf in each tree by performing a binary test at each internal node along this path. A test compares a certain feature with a threshold. Training a forest amounts to identifying the set of tests that best separate the data into the different training classes. At each internal node, the feature space is searched for a test that maximizes the reduction of class impurity, typically measured with the class entropy.

Rather than inspecting the full space of features at each node, a random subset is probed, and the best one is selected. Even if this choice renders the individual trees weaker, it decreases the correlation between their outputs, increasing the performance of the forest as a whole. Each training sample is sent to the corresponding child depending on the result of the test, and the process is recursively repeated until the number of samples in a node falls below a threshold, a predefined maximum tree depth is reached, or all the samples belong to the same class. In that case, the node becomes a leaf, and the most frequent class of the training data at the node is stored for testing.

During testing, a new sample is processed by applying respective tests according to the path from the root node to the leaf it traverses. When a leaf node is reached, the tree casts a vote corresponding to the class assigned to this node in the training stage. The final decision for a test sample is obtained by selecting the class with the majority of votes. The class probability of a test sample is estimated as the fraction of votes for that class cast by all trees.

Image Features

This section describes the features used in our method—intensity statistics, texture and curvature anisotropy, and spatial context features. Context features are a combination of intensity, texture, and curvature values sampled using a template. ROI identification requires classification of superpixels for which we use intensity, texture and curvature features, (excluding context information) to ensure fast feature extraction and subsequent classification, as well as good generalization of the classifier. For generating ROI probability maps, we employ the complete set of features (including context information).

Intensity Statistics

It is not always easy to identify the edges of the RV blood pool and surrounding myocardium wall. Psychophysical experiments have established that the human visual system (HVS) is sensitive only to image features of the first and second order (mean and variance) [34–38]. However, MR images commonly contain regions that do not form distinct spatial patterns but differ in their higher order statistics, e.g. boundaries of some malignant tumors are diffuse and invisible to the naked eye [39]. Therefore, in addition to features processed by the HVS, we propose to investigate features that are not discernible by the human eye but may provide discriminating information for our task. For every sub-region (superpixel or pixel neighborhood) we calculate the mean, variance, skewness, and kurtosis of the intensity values.

Texture Anisotropy

Texture is modeled as patterns distinguished by a high concentration of localized spatial frequencies. 2D Gabor filter banks are used to generate texture maps for each image. Gabor filters have optimal joint localization in the spatial and frequency domains. Their multi-scale and multi orientation structure conforms to the receptive field profiles of simple cortical cells [40], and captures rich visual properties such as spatial localization, orientation selection and spatial frequency characteristics. Since Gabor filters incorporate Gaussian smoothing, they are robust to noise.

Texture maps are partitioned into nine equal parts corresponding to nine sectors of a circle, and the entropy of texture values is calculated for each sector. Figure 1a shows a template of the sectors. More sectors led to fewer samples per sector which jeopardizes estimation of stable higher order statistics. Too few sectors reduce the discriminative power as

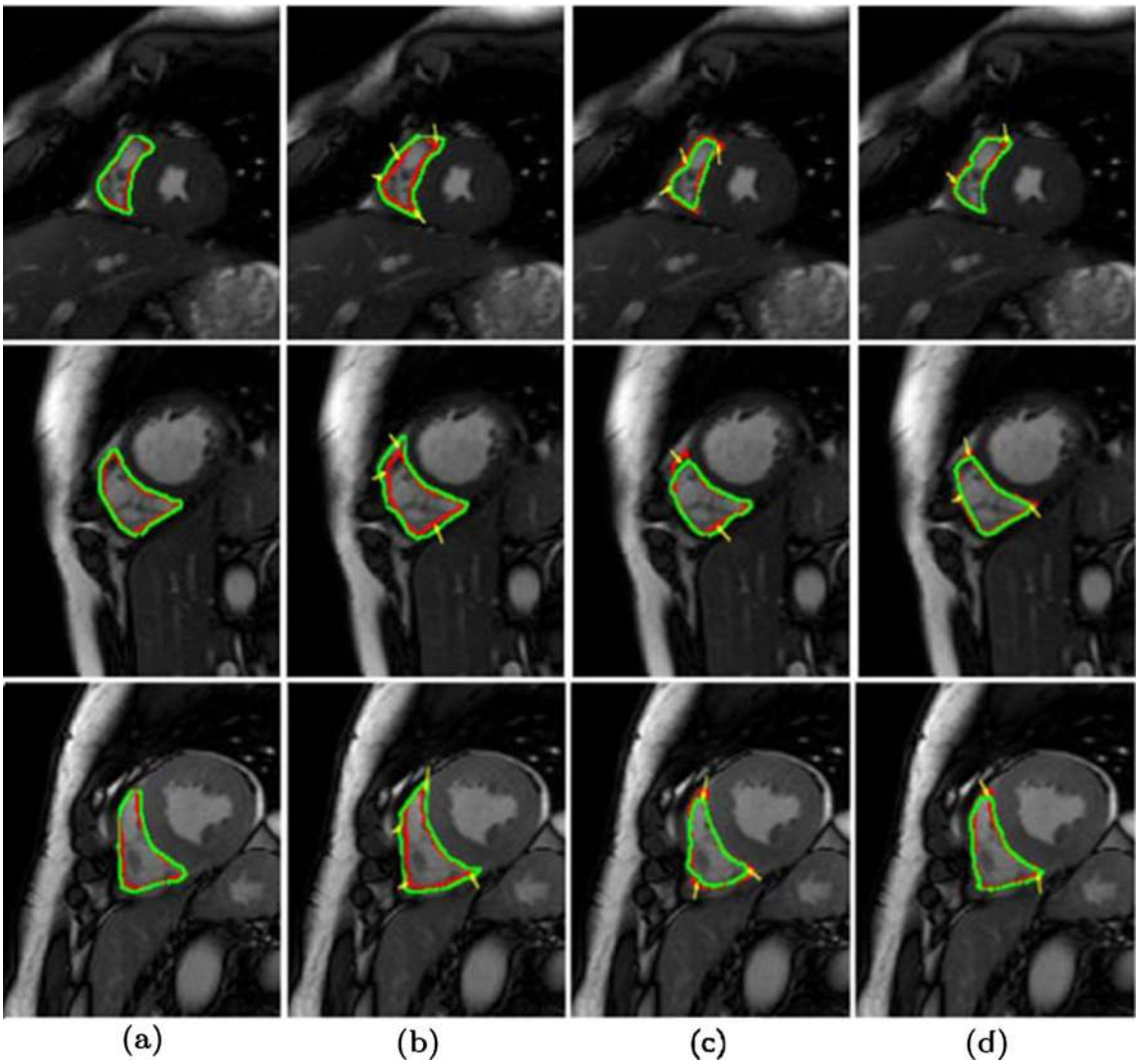

Fig. 1 **a** Sectors of circle for calculating texture and curvature anisotropy; **b** sample locations for deriving context information

many different texture profiles are grouped in one sector. A high entropy value *indicates* wide distribution of texture values (hence higher asymmetry or anisotropy), while low entropy *indicates* low asymmetry or anisotropy. The texture anisotropy for sector r is

$$\chi^{r}_{\text{ani}} = -\sum_{\text{tex}} p^{r}_{\text{tex}} \log p^{r}_{\text{tex}} \quad (1)$$

p^{r}_{tex} denotes the probability distribution of texture values in sector r. This procedure is repeated for all the eight texture maps (four orientations and two scales) to extract a (8×9=) 72 dimensional feature vector.

Curvature Anisotropy

We extend the concept of anisotropy to 2D curvature features. Curvature anisotropy is calculated in a similar manner as texture anisotropy. The entropy of curvature values is determined from nine sectors of an image. If the curvature values have a wide distribution it indicates greater anisotropy, leading to a higher entropy value. On the other hand, low entropy values indicates less anisotropy. The anisotropy measure for a sector r is given by

$$\text{Curv}^{r}_{\text{ani}} = -\sum_{\theta} p^{r}_{\theta} \log p^{r}_{\theta} \quad (2)$$

p^{r}_{θ} denotes the probability distribution of curvature values in sector r, θ denotes the curvature values. Similar to texture anisotropy, the curvature asymmetry measure is also a nine-dimensional feature vector for a region. The curvature asymmetry vector is denoted as Curv. Intensity, texture and curvature features combined give a 85 dimensional feature vector.

Spatial Context Features

Since the human anatomy displays a high degree of regularity with only moderate variations, presence of one organ provides

a strong cue about the presence of another organ in medical images. Through appropriately designed features we aim to capture the contextual relationship between RV and other tissues in the image. Context features have been used to segment brain structures in MRI [41], prostate from CT and MR images [42, 43], cardiac structures from MRI [21], localizing anatomical structures [44] and segmenting the cardiac chamber [45]. Basically, context features derive information of one set of objects from another set of objects.

Since contextual information depends on relative orientation and distance we sample regions at fixed positions from a pixel. Figure 1b shows an illustration of the sampling scheme where the circle center is the pixel in question and the sampled points are identified by a red "X". At each point corresponding to a "X", we extract a 5×5 region and calculate the mean intensity, texture and curvature values. The texture values were derived from the texture maps at 90° orientation and scale 1. The "X"s are located at distances of 3, 8, 15, 22 pixels from the center, and the angle between consecutive rays is 45°. The values from the 32 regions are concatenated into a 96 dimensional feature vector.

Region of Interest Identification

ROI identification is an important part of our method because: (1) it reduces the total computation time since we need not classify each pixel; and (2) the ROI gives an initial selection of likely RV pixels and reduces false positives in subsequent analysis. Intensity inhomogeneity correction was performed using the nonparametric nonuniform intensity normalization (N3) method of [46]. This method performs well without requiring a model of the tissue classes present. The intensities were normalized using the method in [47]. First a "standard" intensity histogram is learnt from a subset of the training images. The training intensity histograms are roughly bimodal and parameters such as minimum and maximum percentile intensities (p_1, p_2), and the second mode of the histogram (μ) are determined. For a given test image, the intensities are rescaled using the following formula

$$x' = s_1 + \frac{x - p_1}{p^2 - p^1}(s_2 - s_1) \tag{3}$$

where x' is the new intensity obtained from the original intensity value x; s_1, s_2 are the minimum and maximum intensities of the test image. This approach leads to good contrast of the different images.

The normalized images are oversegmented into superpixels using the simple linear iterative clustering (SLIC) superpixel algorithm [48]. The desired number of superpixels is specified, which is the number of initial superpixel centers assigned in the image. These centers are initially equally spaced, and then moved to the lowest gradient position. The pixels are clustered based on intensity similarity and spatial proximity to the nearest supervoxel centers. After every iteration, the cluster centers are updated based on the pixels assigned to that cluster. The iterative procedure continues till the superpixel centers do not change.

Superpixel classification for ROI identification needs to be fast and accurate. From the training images, we identify superpixels that contain RV and background voxels, and extract relevant features from them. For every superpixel class, we train RF classifiers with 50 trees. Each superpixel of the test image is labeled by the RF classifier as "RV" or "background" with "RV" superpixels making up the ROI. In "ROI identification", we discuss our strategy of removing false positives.

Probability Maps and Graph Cut Segmentation

Probability maps are generated for all ROI pixels using a second set of RF classifiers. An approximately equal number of samples from RV and background pixels are taken from the training datasets. Intensity, texture, curvature, and context features derived from these samples were used to train a RF classifier (different from the one trained on superpixel features). The features were extracted from a 31×31 neighborhood of each pixel. The training set varies with each round of cross-validation. The trained classifier is used to generate probability maps for every pixel within the identified ROI. Each pixel has three probability values corresponding to the bloodpool, myocardium, and background. The probability maps serve as penalty costs in a second-order MRF cost function. Figure 2a–b show the probability maps of the RV blood pool and myocardium region, and the final segmentation output is shown in Fig. 2c.

The final segmentation is obtained by optimizing a second-order MRF energy function which is written as

$$E(L) = \sum_{s \in P} D(L_s) + \lambda \sum_{(s,t) \in N} V(L_s, L_t), \tag{4}$$

where P denotes the set of pixels, N is the set of neighboring pixels for pixel s L_s is the label of s, and L is the set of all labels. The cost function is optimized using graph cuts [49]. λ is a weight that determines the relative contribution of penalty cost (D) and smoothness cost (V). $D(L_s)$ is given by

$$D(L_s) = -\log(Pr(L_s) + \epsilon), \tag{5}$$

where Pr is the likelihood (from probability maps) previously obtained using RF classifiers and $\epsilon = 0.00001$ is a very small value to ensure that the cost is a real number. The higher the probability for a class, the lower is the corresponding data penalty for that class.

V ensures a smooth solution by penalizing spatial discontinuities. The RF classifier returns a measure of the importance

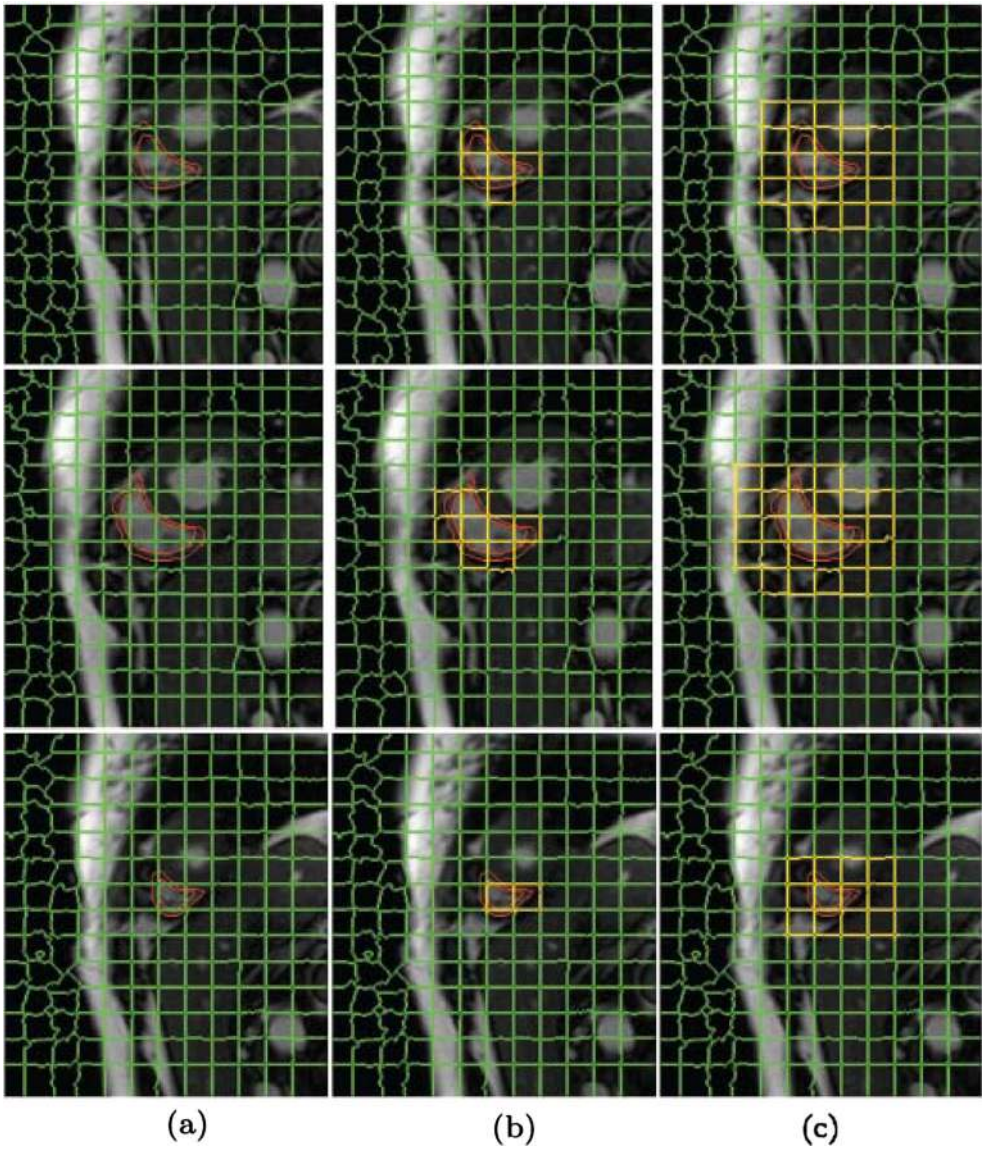

Fig. 2 **a** RV endocardium; **b** RV epicardium; **c** final segmentation. Higher values indicate greater likelihood of belonging to that region. *Red* indicates maximum probability while *blue* indicates zero probability

of each dimension in the feature vector to the classification task. In spite of the multiple dimensional feature vector, the features can be classified into three types—intensity, texture, and curvature. The context information is a combination of the three. By aggregating the importance values of each feature category and normalizing them we obtain the relative importance of each feature in the classification task. This provides the necessary semantic information by quantifying the importance of each feature in classifying a pixel into different categories. Let the weight of the different features be w_I (intensity), w_T (texture) and w_C (curvature), where $w_I+w_T+w_C=1$. The smoothness cost V is given by

$$V(L_s, L_t) = \left\{ \begin{array}{ll} w_I V_I + w_T V_T + w_C V_C, & L_s \neq L_t, \\ 0 & L_s = L_t \end{array} \right\} \quad (6)$$

where V_I, V_T, V_C are the individual contributions to the smoothness by intensity, texture and curvature. V_I is defined as

$$V_I(L_s, L_t) = e^{-\frac{(I_s - I_t)^2}{2\sigma^2}} \cdot \frac{1}{\|s-t\|}, \quad (7)$$

I is the intensity. V_T and V_C are similarly defined using texture and curvature. Note that the weights (or importance measures) depend upon the training set. Since we use different volumes for training (as in a cross-validation setting) we obtain different weight values. However, after training the weights take the following values w_I=0.19–0.22,w_T=0.3–0.33, and w_C=0.43–0.48. This indicates that the relative importance of the different features is the same in all cases.

Table 1 Illustration of calculating importance measures

Low level						Context		
Int	Tex0	Tex45	Tex90	Tex135	Curv	Int	Tex90	Curv
147	98	81	94	85	212	196	130	489

With the help of Table 1, we illustrate how the weights are derived from semantic information. Table 1 shows the importance measures of different features after training. For convenience, we have grouped the measures under two categories—"Low-level" and "Context" features. Low-level features indicate intensity, texture and curvature features, while context features are a combination of the three values sampled at fixed points from a pixel. Note that there are four columns for texture under "low-level" features corresponding to the four orientations of Gabor filters, while "context" has only one value for texture.

The importance measures are not normalized. First we sum up all the values under respective feature categories namely intensity, texture and curvature. Therefore the importance measure of intensity is 343 (Col 1+Col 7). Similarly, the importance measure of Tex is 488 (Col 2+Col 3+Col 4+Col 5+Col 8) and for Curv is 701 (Col 6+Col 9). The sum of all the importance measures is 1532. Dividing the individual importance measures with the sum gives the final normalized importance measures as follows: Int—0.22(w_I), Tex—0.32(w_T) and Curv—0.46(w_C). For the particular round of cross-validation these are the values for the weights.

Experiments and Results

Cardiac MR examinations were performed at 1.5 T (Symphony Tim, Siemens Medical Systems, Erlangen, Germany) using a eight-element phased-array cardiac coil and repeated breath-holds of 10–15 s. A total of 8–12 contiguous cine short axis slices were performed. Sequence parameters were as follows: TR=50 ms; TE=1.7 ms; flip angle=55; slice thickness=7 mm; matrix size=256×216; field of view=360–420 mm; 20 images per cardiac cycle; spatial resolution of 0.75 mm/pixel. There were 32 datasets and we use a leave-one-out strategy to evaluate our method.

Table 2 Change in segmentation accuracy with different values of λ (Eq. 4)

λ	10	5	1	0.5	0.1	0.01	0.02	0.001
DM	74.2	74.8	77.2	80.8	83.7	84.6	89.6	89.7

MRF Regularization Strength λ (Eq. 4)

To choose the MRF regularization strength λ we adopt the following steps. We choose a separate group of seven patient volumes, and perform segmentation using our method but with λ taking different values from 10 to 0.001. The results are summarized in Table 2. The maximum average dice metric (DM) was obtained for λ=0.02 and this parameter value was chosen for the subsequent parameter λ=0.02. Note that these seven datasets were *not part* of the test dataset used for evaluating our algorithm.

ROI Identification

Errors in ROI identification lead to inaccurate RV segmentation. ROI detection is a classification problem where we need to identify those superpixels that contain RV regions. If a superpixel has even one RV pixel, it is denoted as RV, while background superpixels have all background pixels. Table 3 summarizes the superpixel classification performance for different feature combinations using fivefold cross-validation. Over the 32 patients we have 630 RV blood pool superpixels, 612 myocardium superpixels and 653 background superpixels. All Features indicates the combination of intensity, texture and curvature features. Note that context features were not derived from superpixels.

As expected the accuracy for the individual features are lower than their combinations. The combination of texture and curvature features produces results closest to AllFeat. However, this does not indicate that intensity information is unimportant. Conducting a *t* test on the values for Tex+Curv

Table 3 Quantitative measures for superpixel classification under different feature combinations

	Int	Tex	Curv	Tex+Int	Curv+Int	Curv+Tex	AllFeatures
Acc_{BP} (%)	71.6±2.4	75.7±2.5	76.1±2.3	77.9±1.6	78.6±2.7	81.9±1.7	89.9±2.2
Acc_{Myo} (%)	70.8±1.7	72.8±2.7	73.8±2.3	75.9±2.5	76.2±2.9	79.7±2.2	87.9±2.7
Acc_B (%)	70.6±1.9	72.9±2.0	73.5±2.7	75.1±2.2	75.7±1.9	78.9±2.4	87.5±3.1

AccBP classification accuracy for blood pool superpixels, *AccMyo*, classification accuracy for myocardium superpixels, *AccB* classification accuracy for background superpixels, *Int* intensity, *Tex* texture, *Curv* curvature

and AllFeat gives $p<0.017$ which indicates statistically different results. Further, we also conduct t tests for features Tex versus Tex–Int, and Curv versus Curv–Int. In all cases, we find that $p<0.025$, thus clearly showing that inclusion of intensity statistics improves classification accuracy without much extra computational cost.

In any classification scheme, it is difficult to get 100 % classification/detection accuracy. Misclassification of RV superpixels occurs when the number of RV pixels in a superpixel is very low. Hence, the extracted features are more representative of the background. Note that by RV superpixels we mean both blood pool and myocardium superpixels. To overcome this shortcoming, we adopt the following strategy. After classification, we choose the largest cluster of superpixels, thus obtaining "RV" superpixels and avoiding false positives. Then, we change the labels of all their neighboring superpixels (irrespective of their originally assigned labels). This allows us to include some "RV" superpixels that may have been missed in the initial classification.

Figure 3 shows an example where this strategy is particularly effective. The first column shows images from a patient where the RV segmentation is shown in red and the superpixels are shown in green. The second column shows the detected ROI in each slice (shown by the yellow superpixels) overlaid on the manually annotated RV (in red). Small RV regions are missed by the superpixel classification scheme. However, when we include all the neighboring superpixels, the ROI encompasses all possible RV pixels (third column).

Table 3 (AllFeat) gives quantitative measures for superpixel classification before changing the labels of neighboring superpixels. With the change of labels of neighboring superpixels, the following values are obtained: Acc_{BP}= 98.5 %, Acc_{Myo}=97.9.4 % and Acc_B=80.5 %. The labeling of neighboring superpixels increases "false positives" (background superpixels labeled RV). Correspondingly, it also reduces the classification accuracy of background superpixels. However, it also reduces the false negatives (RV superpixels labeled background) and hence the "sensitivity" (RV labeled as RV) also increases. The overall accuracy increases due to a higher improvement in Acc_{BP} and Acc_{Myo} than decrease in Acc_B.

Table 4 Quantitative measures for different superpixel sizes

N	300–600	600–900	900–1,200
Acc_{BP} (%)	77.9±1.6	89.8±2.9	83.5±2.7
Acc_{Myo} (%)	75.1±2.8	88.5±2.3	80.8±3.2
Acc_B (%)	74.7±2.8	89.3±2.8	80.0±2.7

N number of pixels in a superpixel

Effect of Superpixel Size

Smaller superpixels are more homogeneous and the extracted features are representative of a single class. However, they may not always provide sufficient number of pixels to estimate stable features. Larger superpixels contain more pixels to calculate features but may contain pixels from more than one class. Consequently, the extracted features may not be representative of one class. Table 4 summarizes the performance for different superpixel sizes in terms of Acc_{BP}, Acc_{Myo}, and Acc_B. Our experiments show that a good trade-off between accuracy and homogeneous samples is achieved when the number of pixels in a superpixel is in the range 600–900. Depending upon the image dimensions, we set superpixel parameters to get a superpixel of appropriate size.

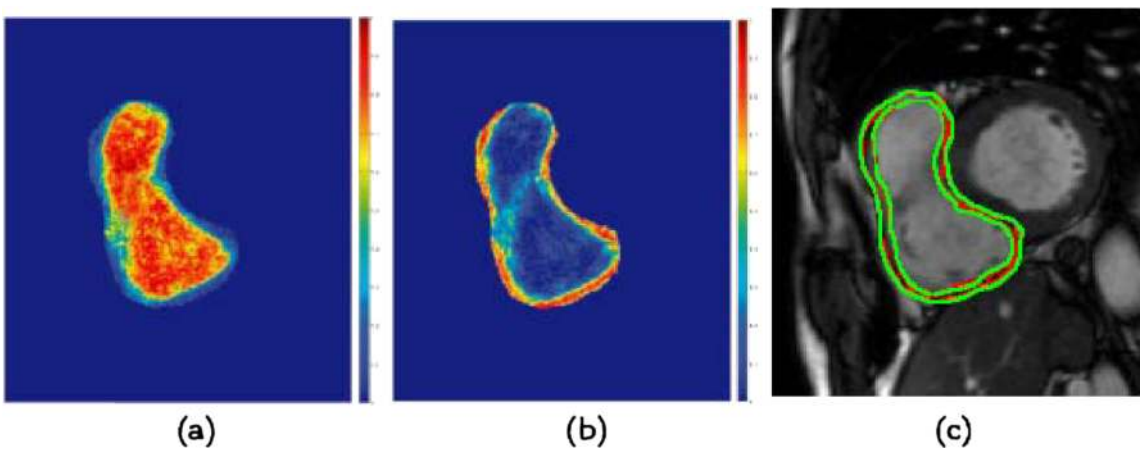

Fig. 3 ROI detection. *First column* shows the superpixels (*green*) and manually annotated RV region (*red*). *Second column* shows the identified RV superpixels (*in yellow*), and *third column* shows the final ROI after neighboring superpixels of Column 2 are changed labels. Results are shown for different slices of the same patient

Table 5 Quantitative measures for RV (blood pool and myocardium) segmentation accuracy

	RF	RF_{nV}	RF_{nC}	RF_{nVI}	RF_{nVT}	RF_{nVC}	MAH
DM	93.2	84.1	81.2	89.7	87.2	89.9	87.4
HD	6.7	11.3	13.2	8.1	9.2	7.9	9.0

DM dice metric in percent, *HD* is Hausdorff distance millimeters

Table 6 Effect of number of trees in RF classifiers (N_T) on segmentation accuracy and training time

N_T	5	7	10	20	50	70	100	150
DM	79.3	81.0	83.4	85.2	90.4	90.5	90.7	90.7
HD	0.1 T	0.1 T	0.3 T	0.5 T	T	1.4 T	2.1 T	3.4 T

Segmentation Results

Table 5 summarizes the performance of our method under different conditions. RF: Our proposed method using RF classifiers. RF_{nC}: RF without context information from images for training the RF classifier; RF_{nV}: RF without semantic context in *V*. $w_I=w_T=w_C=0.33$; RFnVI: RF with $w_I=0$; RF_{nVT}: RF with $w_T=0$; RF_{nVC}: RF with $w_C=0$; MAH: the shape prior based LV segmentation method in [17] applied to our RV datasets. Quantitative evaluation of segmentation performance is given in terms of DM and Hausdorff distance (HD) measures. DM gives a measure of overlap between the reference manual segmentations and the algorithm segmentations. HD gives an idea on the distance between the boundaries of the two segmentations.

RF performs better than MAH. The mean DM values by RF are higher and the mean HD values are lower than other methods. This indicates a more consistent segmentation of images from different phases. The consistently better performance of RF can be attributed to two factors: (1) use of machine learning techniques to identify most discriminant features; and (2) incorporating semantic information into the smoothness cost.

RF_{nC} gives the worst result out of all methods due to exclusion of context information. A *t* test between the values of RF and RF_{nC} gives $p_{RF-RFnC}<0.001$ indicating a large drop in performance without context information. $p_{RF-RFnV}<0.02$ indicates that semantic information is also an important contributor to the segmentation accuracy. Comparing all the other methods with RF gives $0.028<p<0.04$. This shows that while different components of the smoothness cost (like curvature, texture, intensity) are important, semantic information is the most important aspect. The *p* values also show that among the feature maps the most important is curvature followed by texture and intensity.

Figure 4 shows segmentation results for Patients 18, 11, 4 obtained by RF, RF_{nC}, RF_{nVC}, and RF_{nVT}. To ensure clarity, we show only the results for endocardium (or the blood pool). The results reflect the values in Table 5. The low DM values for RF_{nC} highlights the fact that context plays a very important role in our method. Although all the low-level features are used, without context information it is very difficult to discriminate between RV endocardium and epicardium.

Computational Cost

Our method basically consists of the following steps: superpixel segmentation and classification to get the ROI, analyzing every pixel within the ROI to identify RV pixels by generating probability maps and graphcut segmentation.

The average computation time for our entire method on a 256×216 image was 528 s. The automatic ROI identification stage including sub sampling, feature extraction, classification, segmentation, and upsampling to get the ROI took 224 s on an average. Further segmentation of the ROI took 304 s on

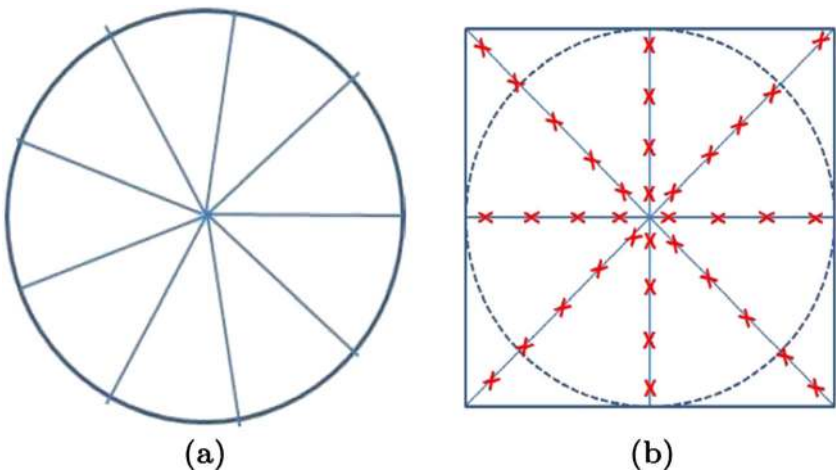

Fig. 4 RV segmentation results: **a** RF; **b** RF_{nC}; **c** RF_{nVC}; and **d** RF_{nVT}. The manual segmentations are in *red* while the algorithm segmentations are in *green*. Areas of inaccurate segmentation are highlighted by *yellow arrows*. The three rows show results for patients 18, 11, and 4

an average, inclusive of the time taken for classification and segmentation.

Influence of Number of Trees

We examine the effect of varying number of trees (N_T) in the second RF classifier (used for generating probability maps) on overall performance. The results for RF_{Sem} are summarized in Table 6 in terms of DM values. When $N_T<10$ DM<0.71. With increasing N_T DM increases alongwith the time taken for training. Table 6 shows the training time (T_{Tr}) for different N_T as a multiple of the training time for $N_T=50$. For $N_T>50$ there is no significant increase in DM ($p>0.1$) but the training time increases significantly. The best trade-off between N_T and DM is achieved for 50 trees and is the reason we have 50 trees for our RF ensemble.

Conclusion and Future Work

In this work, we have proposed a fully automated to segment the cardiac right ventricle from MR images. Given a test image we first oversegment it using superpixel segmentation, and classify each superpixel with random forests classifiers and intensity, texture and curvature features. These steps yield an approximate ROI encompassing the RV. Each pixel within the ROI is further analyzed to segment the RV blood pool and myocardium wall. Probability maps are first generated for each ROI pixel using intensity, texture, curvature, and context features. These maps give the probability of a pixel belonging to RV blood pool, myocardium or the background, and are obtained by a second set of random forest classifiers. The negative log-likelihood of the probability values is the penalty cost for a graph cut segmentation framework.

For spatial smoothness constraints, we make use of semantic information from the second set of trained RF classifiers. The RF classifiers give a measure of the importance of each feature in the classification task. The importance measures are aggregated for different feature types and normalized to get a set of weights for intensity, texture, and curvature. The smoothness cost is designed to incorporate this semantic information where intensity, texture and curvature differences of neighboring pixels are weighted by the values obtained from semantic information.

Experimental results on 32 patient datasets show that semantic information helps us achieve a high level of segmentation accuracy compared to the shape prior based segmentation method of [17]. We also analyze the importance of individual features, and conclude that context information is the most important feature followed by curvature features.

References

1. Allender S, Scarborough P, Peto V, Rayner M, Leal J, Luengo-Fernandez R, Gray A: European cardiovascular disease statistics, European Heart Network, 2008
2. Matthews JC, Dardas TF, Dorsch MP, Aaronson KD: Right sided heart failure: diagnosis and treatment strategies. Curr. Treat. Options Cardiovasc 10(4):329–341, 2008
3. Shors S, Fung C, Francois C, Finn P, Fieno D: Accurate quantification of right ventricular mass at MR imaging by using cine true fast imaging with steady state precession: study in dogs. Radiology 230(2):383–388, 2004
4. Petitjean C, Dacher J-N: A review of segmentation methods in short axis cardiac mr images. Med. Imag. Anal. 15(2):169–184, 2011
5. Lapp RM, Lorenzo-Valdes M, Daniel Rueckert: 3d/4d cardiac segmentation using active appearance models, non-rigid registration, and the insight toolkit, in Proc. MICCAI, 2004, pp. 419–426
6. Zhuang X, Rhode KS, Razavi RS, Hawkes DJ, Ourselin S: A registration based propagation framework for automatic whole heart segmentation of cardiac MRI. IEEE Trans. Med. Imag. 29(9):1612–1625, 2010
7. Lorenzo-Valdes M, Sanchez-Ortiz GI, Elkington AG, Mohiaddin RH, Rueckert D: Segmentation of 4d cardiac MR images using a probabilistic atlas and the EM algorithm. Med Image Anal. 8(3):255–265, 2004
8. ElBaz MS, Fahmy AS: Active shape model with inter-profile modeling paradigm for cardiac right ventricle segmentation, in MICCAI, 2012, pp. 691–698
9. Ou Y, Doshi J, Erus G, Davatzikos C: Multi-atlas segmentation of the right ventricle in cardiac mri, in Proc. MICCAI RV Segmentation Challenge, 2012
10. Zuluaga MA, Cardoso MJ, Ourselin S: Multi atlas fusion: Automatic right ventricle segmentation using multi-label fusion in cardiac mri, in Proc. MICCAI RV Segmentation Challenge, 2012
11. Nambakhsh CMS, Rajchl M, Yuan J, Peters TM, Ben-Ayed I: Rapid automated 3d rv endocardium segmentation in mri via convex relaxation and distribution matching, in Proc. MICCAI RV Segmentation Challenge, 2012
12. Grosgeorge D, Petitjean C, Ruan S, Caudron J, Dacher J: Right ventricle segmentation by graph cut with shape prior, in Proc. MICCAI RV Segmentation Challenge, 2012
13. Maier O, Jimenez-Carretero D, Santos A, Ledesma-Carbayo MJ: Right- ventricle segmentation with 4d region-merging graph cuts in mr, in Proc. MICCAI RV Segmentation Challenge, 2012
14. Paragios N: A variational approach for the segmentation of the left ventricle in cardiac image analysis. Intl. J. Comp. Vis. 50(3):345–362, 2002
15. Lynch M, Ghita O, Whelan P: Left ventricle myocardium segmentation using a coupled level set with a-priori knowledge. Comput. Med. Imag. Graph. 30(4):255–262, 2006
16. Lin X, Cowan B, Young A: Model based graph cut method for segmentation of the left ventricle, in In Proc: EMBC, 2005, pp. 3059–3062
17. Mahapatra D, Sun Y: Orientation histograms as shape priors for left ventricle segmentation using graph cuts, in In Proc: MICCAI, 2011, pp. 420–427
18. Mahapatra D: Cardiac image segmentation from cine cardiac mri using graph cuts and shape priors. J. Digit. Imaging 26(4):721–730, 2013
19. Mahapatra D, Sun Y: Joint registration and segmentation of dynamic cardiac perfusion images using mrfs., in Proc. MICCAI, 2010, pp. 493–501
20. Mahapatra D, Sun Y: Integrating segmentation information for improved elastic registration of perfusion images using an mrf framework. IEEE Trans. Imag. Proc. 21(1):170–183, 2012

21. Mahapatra D: Cardiac LV and RV segmentation using mutual context information, in Proc. MICCAI-MLMI, 2012, pp. 201–208
22. Pluempitiwiriyawej C, Moura JMF, Wu YL, Ho C: STACS: new active contour scheme for cardiac MR image segmentation. IEEE Trans. Med. Imag. 24(5):593–603, 2005
23. Billet F, Sermeanst M, Delingette H, Ayache N: Cardiac motion recovery and boundary conditions estimation by coupling an electromechanical model and cine-MRI data, in Functional Imaging nad modeling of the heatt (FMIH), 2009, pp. 376–385
24. Lotjonen J, Kivisto S, Koikkalainen J, Smutek D, Lauerma K: Statistical shape model of atria, ventricles and epicardium from short- and long-axis mr images. Med Image Anal. 8(3):371–386, 2004
25. Breiman L: Random forests. Machine Learning 45(1):5–32, 2001
26. Mahapatra D, Schueffler P, Tielbeek J, Makanyanga JC, Stoker J, Taylor SA, Vos FM, Buhmann JM: Automatic detection and segmentation of crohn's disease tissues from abdominal mri. IEEE Trans. Med. Imaging 32(12):1232–1248, 2013
27. Mahapatra D, Buhmann JM: Analyzing training information from random forests for improved image segmentation., In press IEEE Trans. Imag. Proc
28. Mahapatra D, Schüffler P, Tielbeek J, Vos FM, Buhmann JM: Crohn's disease tissue segmentation from abdominal mri using semantic information and graph cuts, in Proc. IEEE ISBI, 2013, pp. 358–361
29. Berks M, Chen Z, Astley S, Taylor C: Detecting and classifying linear structures in mammograms using random forests, in IPMI, 2011, pp. 510–524
30. Kelm BM, Mittal S, Zheng Y, et al: Detection, grading and classification of coronary stenoses in computed tomography angiography, in MICCAI, 2011, pp. 25–32
31. Mahapatra D, Schüffler PJ, Tielbeek J, Buhmann JM, Vos FM: A supervised learning based approach to detect crohn's disease in abdominal mr volumes, in Proc. MICCAI workshop Computational and Clinical Applications in Abdominal Imaging (MICCAI-ABD), 2012, pp. 97–106
32. Schffler PJ, Mahapatra D, Tielbeek JAW, Vos FM, Makanyanga J, Pends DA, Nio CY, Stoker J, Taylor SA, Buhmann JM: A model development pipeline for crohns disease severity assessment from magnetic resonance images, in In Proc: MICCAI-ABD, 2013, pp. 1–10
33. Mahapatra D, Schüffler P, Tielbeek J, Vos FM, Buhmann JM: Semi-supervised and active learning for automatic segmentation of crohn's disease, in Proc. MICCAI, Part 2, 2013, pp. 214–221
34. Julesz B, Gilbert EN, Shepp LA, Frisch HL: Inability of humans to discriminate between visual textures that agree in second-order statistics-revisited. Perception 2(4):391–405, 1973
35. Vos FM, Tielbeek FM, Naziroglu R, Li Z, Schüffler P, Mahapatra D, Alexander Wiebel, Lavini C, Buhmann J, Hege H, Stoker J, van Vliet L: Computational modeling for assessment of IBD: to be or not to be?, in Proc. IEEE EMBC, 2012, pp. 3974–3977
36. Mahapatra D, Saini MK, Sun Y: Illumination invariant tracking in office environments using neurobiology-saliency based particle filter, in IEEE ICME, 2008, pp. 953–956
37. Mahapatra D, Sun Y: Registration of dynamic renal mr images using neurobiological model of saliency, in Proc. ISBI, 2008, pp. 1119–1122
38. Mahapatra D, Sun Y: Nonrigid registration of dynamic renal MR images using a saliency based MRF model, in Proc. MICCAI, 2008, pp. 771–779
39. Petrou M, Kovalev VA, Reichenbach JR: Three-dimensional nonlinear invisible boundary detection. IEEE Trans. Imag. Proc 15(10): 3020–3032, 2006
40. De Valois RL, Albrecht DG, Thorell LG: Spatial-frequency selectivity of cells in macaque visual cortex. Vis. Res. 22(5):545–559, 1982
41. Tu Z, Bai X: Auto-context and its application to high-level vision tasks and 3d brain image segmentation. IEEE Trans. Patt. Anal. Mach. Intell. 32(10):1744–1757, 2010
42. Li W, Liao S, Feng Q, Chen W, Shen D: Learning image context for segmentation of prostate in ct-guided radiotherapy, in MICCAI, 2011, pp. 570–578
43. Mahapatra D, Buhmann JM: Prostate mri segmentation using learned semantic knowledge and graph cuts. In press IEEE Trans. Biomed. Engg 61(3):756–764, 2014
44. Criminsi A, Shotton J, Bucciarelli S: Decision forests with long range spatial context for organ localization, in MICCAI workshop on Probabilistic Models for Medical Image Analysis (MICCAI-PMMIA), 2009
45. Zheng Y, Barbu A, Beorgescu B, Scheuering M, Comaniciu D: Four chamber heart modeling and automatic segmentation for 3D cardiac CT volumes using marginal space learning and steerable features. IEEE Trans. Med. Imag. 27(11):1668–1681, 2008
46. Sled JG, Zijdenbos AP, Evans AC: A nonparametric method for automatic correction of intensity nonuniformity in mri data. IEEE Trans. Med. Imag. 17(1):87–97, 1998
47. Nyl LG, Udupa JK: On standardizing the mr image intensity scale. Magnetic resonance in medicine 42(6):1072–1081, 1999
48. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Ssstrunk S: Slic superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Patt. Anal. Mach. Intell. 34(11):2274–2282, 2012
49. Boykov Y, Veksler O: Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23:1222–1239, 2001