
Article:
Albà, X., Lekadir, K., Pereañez, M. et al. (3 more authors) (2018) Automatic initialization and quality control of large-scale cardiac MRI segmentations. Medical Image Analysis , 43. pp. 129-141. ISSN 1361-8415

https://doi.org/10.1016/j.media.2017.10.001

Automated Segmentation Initialization and Quality Control for Population Cardiac MRI Analysis

Xènia Albà[a,*], Karim Lekadir[a], Marco Pereañez[b], Pau Medrano-Gracia[c], Alistair A. Young[c], Alejandro F. Frangi[b]

[a] *Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), Universitat Pompeu Fabra, Barcelona, Spain*
[b] *Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), Electronic and Electrical Engineering Department, University of Sheffield, Sheffield, UK*
[c] *Department of Anatomy and Medical Imaging, University of Auckland, Auckland, NZ*

Abstract

Continuous advances in imaging technologies enable ever more comprehensive phenotyping of human anatomy and physiology. Concomitant reduction of imaging costs has resulted in widespread use of imaging in large clinical trials and population imaging studies. Magnetic Resonance Imaging (MRI), in particular, offers one-stop-shop multidimensional biomarkers of cardiovascular physiology and pathology. A wide range of analysis methods offer sophisticated cardiac image assessment and quantification for clinical and research studies. However, most methods have only been evaluated on relatively small databases often not accessible for open and fair benchmarking. Consequently, published performance indices are not directly comparable across studies and their translation and scalability to large clinical trials or population imaging cohorts is uncertain. Most existing techniques still rely on considerable manual intervention for the initialization and quality control of the segmentation process, becoming prohibitive when dealing with thousands of images.

The contributions of this paper are three-fold. First, we propose a fully automatic method for initializing cardiac MRI segmentation, by using image features and random forests regression to predict an initial position of the heart and key anatomical landmarks in an MRI volume. In processing a full imaging database, the technique predicts the optimal corrective displacements and positions in relation to the initial rough intersections of the long and short axis images. Second, we introduce for the first time a quality control measure capable of identifying incorrect cardiac segmentations with no visual assessment. The method uses statistical, pattern and fractal descriptors in a random forest classifier to detect failures to be corrected or removed from subsequent statis-

*Corresponding author
Email address: xenia.alba@upf.edu (Xènia Albà)

[1] The work of X. Albà was supported by an FPU grant by the Spanish Ministry of Education Culture and Sport.

tical analysis. Finally, we validate these new techniques within a full pipeline for cardiac segmentation applicable to large-scale cardiac MRI databases. The results obtained based on over 1200 cases from the Cardiac Atlas Project show the promise of fully automatic initialization and quality control for population studies.

Keywords: Cardiac segmentation, statistical shape models, magnetic resonance imaging, large-scale studies, automatic image segmentation

1. Introduction

Continuous advances in imaging technology, which enable ever more comprehensive phenotyping of human anatomy and physiology and concomitant reduction of imaging costs, have resulted in widespread use of imaging in large clinical trials and population imaging studies (Murdoch and Detsky, 2013). There has been an emergence of large-scale population imaging databases (Rueckert et al., 2016), opening up challenges and opportunities for the understanding of disease phenotypes, and for the delivery of precision imaging (Frangi et al., 2016). In the area of cardiovascular imaging, for example, the exploitation of large-scale population image data is expected to influence the characterization of cardiovascular phenotypes like never before (Suinesiaputra et al., 2015). This, however, calls for developing of new techniques for cardiac image analysis that can handle the scale and variability associated with large imaging studies (Medrano-Gracia et al., 2015).

Among existing imaging techniques, cardiovascular Magnetic Resonance Imaging (MRI) has established itself as the one-stop-shop approach for non-invasive examination of cardiac morphology and function (Ponikowski et al., 2016). To enable high throughput analysis of imaging data automatically and reliably, a first step is the delineation of the myocardial boundaries followed by the estimation of various cardiac functional indices (Peng et al., 2016). However, practical problems arise when dealing with large MRI studies. Manual expert input becomes unfeasible and there is a need for fast and scalable methods to parse image data under large variability of anatomy, physiology and image quality.

In the existing literature, a wide range of approaches for automatic and semi-automatic cardiac MRI segmentation methods have been proposed (some reviews have been published (Frangi et al., 2001; Petitjean and Dacher, 2011; Tavakoli and Amini, 2013; Peng et al., 2016). These methods include a wide range of techniques including image-based classification (Jolly, 2006; Katouzian et al., 2006; Lu et al., 2009b; Cousty et al., 2010), pixel classification (Lynch et al., 2006; Pednekar et al., 2006; Nambakhsh et al., 2013), deformable models (Berbari et al., 2007; Billet et al., 2009; Cordero-Grande et al., 2011; Queirós et al., 2015), cardiac atlases (Lorenzo-Valdés et al., 2004; Lötjönen et al., 2004; Bai et al., 2015), statistical models (van Assen et al., 2006; Ordas et al., 2007; Zhu et al., 2010; Lu et al., 2011; Grosgeorge et al., 2011; Zhang et al., 2010),

Table 1: Summary of state of the art methods for automatic and semi-automatic LV segmentation with large databases.

Reference (year)	Views	# Datasets	# Phases	Pathologies	Interaction	Method
Lin et al. (2006)	SA Cine	330	25	CAD, PAD, MI	none	Combination of temporal Fourier analysis with contour detection
Zhu et al. (2007)	SA Cine	225	1	n/a	none	Deformable model with intensity and volume constraint of myocardium
Lu et al. (2011)	SA Cine	100	20	n/a	none	Joint LV-RV model combined both spatial and temporal context
Jolly et al. (2012)	SA Cine	100†	20	CAD	none	Combined deformable registration method with gray level based shortest path segmentation algorithm
Margeta et al. (2012)	SA Cine	100†	20	CAD	none	Supervised voxel-wise classification with layered spatio-temporal forests
Eslami et al. (2013)	SA Cine	104	18-25	N, DCM, HCM, MI, HF	manual seeds	Segmentation by retrieval with guided random walks
Tsadok et al. (2013)	LA Cine	126◇	1	CAD	manual mitral valve landmarks	Combination of a shortest path algorithm and a non-rigid registration
Lu et al. (2013)	SA Cine	133	20	N, HCM, MI, HF	none	Combination of optimal thresholding, fast Fourier transform and multiple seeds region growing

n/a = not available; CAD = Coronary Artery Disease, PAD = peripheral arterial disease, MI = myocardial infarction
N = Normal, DCM = dilated cardiomyopathy, HCM = hypertrophic cardiomypathy, HF = heart failure
† STACOM'11 Challenge (Suinesiaputra et al., 2012); ◇ 100 from DETERMINE database (Kadish et al., 2009)

and learning-based approaches (Margeta et al., 2012; Eslami et al., 2013; Tsadok et al., 2013; Avendi et al., 2016).

Most published methods have been developed and validated typically using a few dozen image datasets and image databases not openly accessible. Consequently, published performance indeces are not directly comparable across studies and, more important, they have questionable translation and scalability as a measure of usability and performance in large clinical trials or population imaging cohorts. Most existing techniques still rely on considerable manual intervention of some sort either during initialization or correction of cardiac image segmentation. The quality control of the segmentation process and results is also manual or visual, which becomes prohibitive when dealing with hundreds or thousands of image volumes. One problem is solved by segmentation challenges on publicly available image databases and evaluation protocols and measures (Medrano-Gracia et al., 2015). Two relevant challenges for LV segmentation, for instance, are the 2009 challenge (Radau et al., 2009) organized during the MICCAI Conference; or the 2011 challenge (Suinesiaputra et al., 2014) organized as part of the STACOM Workshop, also during MICCAI. These two challenges, however, were tested only on 30 and 100 cardiac MRI datasets for which manual annotations were provided as ground-truth. Given the small size of the datasets used for testing, whether the reported results can be generalized to larger cohorts remains highly questionable. A summary of the methods for semi-automatic and automatic LV segmentation in MRI using a large database (more than 100 datasets) of the last decade can be found in Table 1.

Existing tools often rely on manual user intervention for the initialization (Eslami et al., 2013) of image segmentation or definition of key anatomical landmarks (Tsadok et al., 2013). This approach becomes unfeasible when dealing with hundreds or thousands of MRI studies. Very few works deal with the automatic detection of key ventricular landmarks in large size databases (Medrano-Gracia et al., 2015). A list of automatic landmark detection methods is provided in Table 2, with sizes varying between 80 and 338 datasets. Zheng et al. (2009), who restrict their work to long-axis (LA) images, evaluate their landmark detection method with 163 cases by using a marginal space learning approach. As an indication of the difficulty of the task, only two participants submitted their results in the Landmark Detection Challenge of STACOM 2012. The first technique by Mahapatra (2013) required prior image segmentation, which is generally unavailable. Lu and Jolly (2013) proposed context modeling for LV landmark detection based on Lu et al. (2009a) and Lu et al. (2010). Both are supervised methods relying on large datasets with manually annotated ground truth for training (Medrano-Gracia et al., 2015).

A final limitation of existing approaches to analyze cardiac MRI data is their reliance on visual verification of the segmentation results. It is important to identify/remove grossly incorrect segmentations before they are used to derive anatomical or functional image biomarkers (e.g. ejection fraction, left ventricular volume, wall thickness, etc.). This also becomes impractical in large population imaging studies and needs to be automated. Ideally, one would want a method that automatically initializes the segmentation process and that

Table 2: Automatic Landmark Detection Methods in cardiac MRI.

Reference (year)	Views	# Datasets	# Phases	Landmark	Method
Lu et al. (2009a)	LA Cine	116	1	Apex and MV	Joint context based approach under a learning-based object detection framework
Zheng et al. (2009)	LA Cine	163	1	Apex and MV	Marginal Space Learning and Component-Based Voting
Lu et al. (2010)	SA / LA Cine	188 (LA) 338 (SA)	1	RVI, MV, Apex and RVL	Context Modeling
Mahapatra (2013)	SA / LA Cine	80†	1	RVI, BACA and MV	Morphological operators and Random Forests with low level features
Lu and Jolly (2013)	SA / LA Cine	100†	1	RVI, BACA and MV	Joint context based approach

MV = Mitral valve points ; RVI = RV insert points ; RVL = RV lateral point; BACA = Base-to-apex central axis points

† Data from CAP as part of the STACOM Challenge 2012 (Fonseca et al., 2011)

subsequently scores the quality of the segmentation results.

The contributions of this paper are three-fold. First, we propose a fully automatic method for initializing cardiac MRI segmentation, by using image features and random forests regression to predict an initial position of the heart and key anatomical landmarks in an MRI volume. In processing a full imaging database, the technique predicts the optimal corrective displacements and positions in relation to the initial rough intersections of the long and short axis images. Secondl, we introduce a measure capable of identifying incorrect or failed segmentations without the need for any visual assessment. To the best of our knowledge, this is the first technique proposed for automatic quality control of cardiac MRI segmentation, which remains a user-dependent task (Slomka et al., 2017). The method uses statistical, pattern and fractal descriptors in a random forest classifier to detect failures to be corrected or removed from subsequent statistical analysis. Finally, we integrate these new techniques within a full pipeline for cardiac segmentation applicable to large-scale cardiac MRI databases. Specifically, we illustrate the capability of the pipeline with a previously published Active Shape Model (ASM) cardiac segmentation method.

To evaluate the proposed framework, we use MRI datasets obtained from the Cardiac Atlas Project (CAP) (Fonseca et al., 2011). To the best of our knowledge, this is the first attempt to parse more than 1200 cardiac MRI datasets from both normal and abnormal cases using a fully automatic pipeline.

2. Method

2.1. Overview

Our automatic MRI segmentation framework has three main components: regression-based organ detection, model-based organ segmentation, and classification-based quality control. We illustrate them schematically in Fig. 1.

The goal of the proposed initialization is to estimate, with no user interaction, the location of the left ventricle (LV) and key anatomical landmarks (basal and apical points). From these points, it is possible to compute the initial position, height and size of the heart that feed the subsequent segmentation algorithm. The basic idea behind the proposed technique is to learn a regressor to model the displacements from arbitrary incorrect positions to the correct localization of ventricular axes and anatomical landmarks. More specifically, we start with a rough estimation of the pose of the heart based on the positions of the short- and long-axis images. We then refine this position by predicting via regression the displacements of the ventricular axis at each short-axis slice to a more suitable position. The same strategy is used for the detection of other landmarks such as the apex and mitral valve points.

To build the pose correction regression model, the training phase is illustrated in Fig. 2. The idea is to generate synthetic displacements by sampling arbitrary positions around the ground truth (*e.g.* ventricle's center and the correct positions of the salient anatomical landmarks). Therefore, instead of learning the appearance around a particular landmark, we learn the features

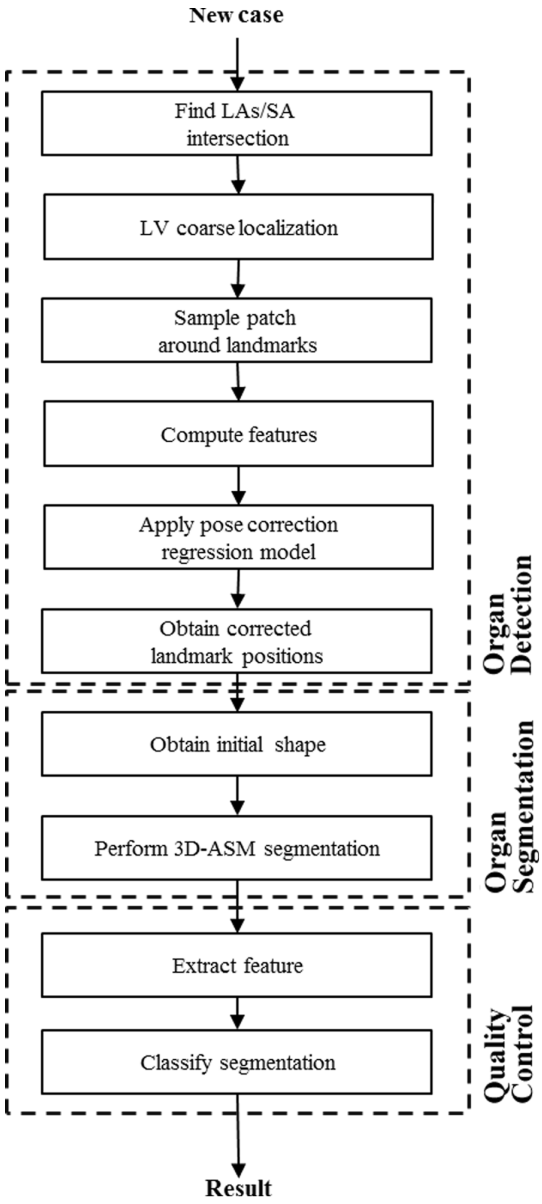


Figure 1: Pipeline of the LV automatic segmentation. The first step is a key landmark detector, which is used to obtain an initial shape. A 3D-ASM segmentation step provides the final delineation of the ventricular boundaries. A quality control step indicates the confidence of the obtained results.

related to the corrections from arbitrary locations. Based on the characteristic appearance around an initial position (e.g. the middle of the image or a location known to fall within the blood pool), the regression model will propose a positional correction.

After the initialization, a segmentation method is applied to obtain the final delineation of the ventricular boundaries. The segmentation is performed based on a sparse ASM paradigm (SPASM) (van Assen et al., 2006), which uses statistical shape models (SSMs) of the heart to encode prior statistical knowledge about cardiac shape and MRI intensity in a training set.

In view of using the proposed image segmentation methodology in population imaging, visual inspection of the segmentation results is as unpractical as the segmentation itself. Hence, we wish to endow the segmentation method with a mechanism of self-verification. This is highly desirable and could open the way to self-correcting segmentation methods. The training set is built by creating synthetic displacements and deformations from the ground truth left ventricle segmentation to learn a mapping between the appearance of correct and failed segmentations.

All the steps are explained in the next section and shown in Fig. 1.

2.2. Automatic Key Landmark Detection

2.2.1. Landmark Set Definition

To enable automatic initialization, an accurate identification of key anatomical landmarks is an essential step in cardiac MRI segmentation, as these points will estimate the initial pose, scale and rotation of the model. Generally, as shown in Table 2, relevant landmark points include the mitral valve hinge points (defining the base line position), the apex and the central axis position, which are often used to initialize the segmentation process (Tobon-Gomez et al., 2012; Tsadok et al., 2013; Piazzese et al., 2016). These points are usually defined manually, which is impractical for large population studies.

In cardiac MRI, the mitral valve (located between the left atrium and the LV) is shown in long axis (LA) view, because it contains both the atrium and LV. The line connecting the mitral valve hinge points is the base plane, which defines the border of LV at base. The left ventricular apex is at the bottom of the LV, opposite the base of the heart. The central axis location, which is not an actual anatomical feature, is defined in this paper at the centroid of the LV cavity on each SA slice. The axis location will be exploited to define the LV position for the models.

In this work, all these key anatomical landmarks, as well as the LV axis points on each SA slice are identified automatically as described in the following subsections.

2.2.2. Coarse LV Localization

As an initial step in the automatic initialization, we identify the location of the LV through a rough estimation of its axis. We do this by estimating the intersection between the long axis (LA) and short axis (SA) slices of each

CMR study. The main MRI planes used for cardiac function (Kramer et al., 2013) typically include three long axis (LA) views (horizontal long axis − HLA, vertical long axis − VLA, and left ventricular outflow tract − LVOT), and a complete short axis (SA) stack covering the LV and RV. The series of SA views are generated starting from the LV base at the level of the mitral valve, all the way down to the apex. The VLA is orthogonal to the SA axis and passes approximately through the apex and center of the mitral valve. The HLA is aligned orthogonally to the VLA, also passing through the apex and center of the mitral valve. In this paper, the intersection between the HLA, the VLA and the SA images is exploited to obtain a rough estimation of LV position. Fig. 3 illustrates the intersection of the two LA images and the SA images, providing a point on each SA image and a set of points on the LA images. However, the result is approximate due to scanning inaccuracies and changes in breath-hold positions, which lead to mis-registration between slices. The next step tries to refine the pose estimation and to obtain a more accurate estimation of the LV axis to increase the robustness of the later LV segmentation.

2.2.3. Key Landmarks Detection

To refine the location of the LV axis, and to identify the positions of the key anatomical landmarks of the LV, we use a regressor that refines the localization of key LV landmark points using image features that characterize the local appearance of these landmarks as detected in SA/LA views. The landmark regressors are trained by creating a large synthetic training set of random displacements that emulate localization errors from the ground-truth. The regressors learn the mapping between these offsets and the known ground-truth positions using Random Forests (RF). The main steps in the landmark detection are illustrated in Fig. 1.

Training Set Generation. To train the landmark regressors, we first emulate displacements errors as random displacements from the ground-truth landmark positions available from expert annotations. Specifically, each landmark is perturbed randomly from its correct position denoted as the vector $\mathbf{d}_i$. To each displacements, we associate image features $\mathbf{f}_i$ sampled from their neighborhood of the simulated erroneous position. Let us denote all these pairs in the training sets as $\mathbf{P}_i = (\mathbf{f}_i, \mathbf{d}_i)$. In our case, the displacements are generated as random vectors with a norm between 2 and 15 mm with uniform angular distribution. Then the goal is to learn a mapping function M from $\mathbf{f}_i$, the image feature space (around the incorrect position), to $\mathbf{d}_i$, the displacement vector space that would correct the position estimate. The main steps of the training are shown in Fig. 2.

Random Forests Regressor. In this paper, we use the well-known Random Forest (RF) (Breiman, 2001) to compute a regression model for landmark detection. RFs are widely considered as one of the most robust classifiers and regressors. The algorithm is based on decision trees ensembles exploiting two mechanisms: building a tree ensemble via bagging with replacement, whereby any example

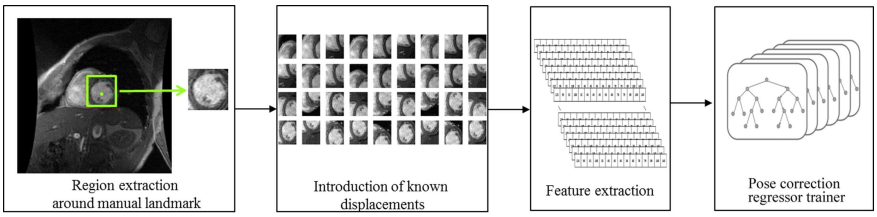
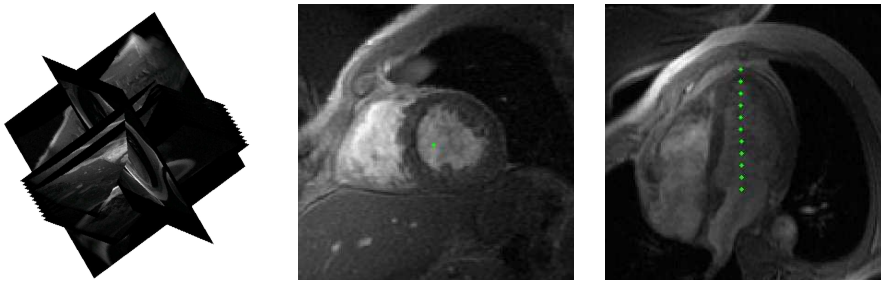


Figure 2: In each training dataset, we sample a patch around the landmark (e.g. centroid of the LV) and we produce arbitrary displacements from the real position provided by the manual annotation of the expert. We then calculate image features from the patches and we learn a mapping function from the feature space to the displacement space.



(a) (b) (c)

Figure 3: The intersection of the two LA and the SA images (a) provides rough positions for the LV, which is shown as a point on each SA image (b) and a set of points on the LA images (c).

selected from the training set can be used again; and random feature selection at each tree node, *i.e.* random selection of a small fraction of features and splitting using the best feature from the current set. During training, the RF model calculates a response variable by creating different regression trees (the forest of trees) and then processing each object to be modeled down each tree. Once the RF is trained, it is treated as a complete ensemble of base learning trees. Each leaf contains a distribution for the continuous output variable/s. Leaves are associated with different degrees of confidence. To perform the regression, each learner produces a prediction individually and then predictions of all learners in the ensemble are combined to generate a prediction of the ensemble. The number of trees can be adapted to find the desired trade-off between accuracy and computational efficiency of the detection process. In this paper, we constructed a RF with 50 trees with a maximum decision depth of 15; values were selected based on experiments (see Results section).

Predictive Landmark Correction. Once the displacement regression model is trained, given a new slice, we sample a patch around the initial rough position $\mathbf{z}_s$ obtained as described in Sec. 2.2.2 from the intersection of the LA and SA planes. From this patch, we calculate an image feature vector $\mathbf{f}_s$. Then through

the trained mapping M, we can calculate a predicted displacement $\mathbf{d}'_s = M(\mathbf{f}'_s)$, and then $\mathbf{z}_s + \mathbf{d}'_s$ becomes the prediction of the correct landmark position. We use a distinct displacement regression model for each landmark described in Sec. 2.2.3.

Feature Extraction. The last step of the automatic initialization defines the imaging features $\mathbf{f}_i$. Using raw image intensities directly, as in deep learning, is an option. However, we have designed more specialized image features that explicitly relate to the orientation and extent of deformation required to correct landmark positions. While deep learning overcomes the need for feature handcrafting, it does so at the expense of handcrafting the network architecture that could be equally arbitrary and less directly associated to the nature of the problem. In this work, we have used two complementary descriptors, i.e. the Histogram of Oriented Gradients (HOG) and Gabor Filters (GFs). This choice is motivated by the need for descriptors that can predict both the extent and orientation of the displacements.

- *Histogram of Oriented Gradients (HOG)*– The HOG descriptor (Dalal and Triggs, 2005) is a local statistic of the orientations of the image gradients, thus describing the local appearance. It is characterized by its invariance to rotation and illumination changes. Moreover, its computation technique is simple and fast. The main idea behind this descriptor is that local object appearance and shape can often be characterized rather well by the distribution of local intensity gradients or edge directions. The HOG feature divides the patch into many cells, with each cell location (x, y), represented through its unsigned gradient orientation angle θ and gradient magnitude ρ. A histogram counts the occurrences of pixels orientations given by their gradients, *i.e.* each cell (x, y) votes according to its magnitude into the bin corresponding to its gradient orientation. The final HOG descriptor is then built as a combination of these histograms. The local histograms are contrast-normalized by calculating a measure of the intensity across different regions and then using this value to normalize all cells.

- *Gabor Filters*– The Gabor filter (Gabor, 1946) is an established texture descriptor in image analysis. It is applied to extract features by analyzing the frequency domain of the image. A GF is basically a Gaussian function modulated by complex sinusoidal of frequency and orientation. It can filter both in the spatial and frequency domains and can be extended to many dimensions. These filters are more desirable since they provide for finer characterization of different textures. In this paper, we applied a two-dimensional (2D) Gabor filter since the data sets collected are sparse collections of 2D images. Two-dimensional Gabor filter banks decompose an image into feature maps related to different scales and orientations, capturing visual properties such as spatial localization, orientation selectivity, and spatial frequency. The 2D GF consists of a complex exponential centered at a frequency and modulated by a Gaussian envelope. To obtain

the Gabor-filtered image $\mathbf{I}_g$ of an image $\mathbf{I}$ the 2D convolution operation is performed:

$$\mathbf{I}_g(x, y) = G(x, y, \sigma_x, \sigma_y, \omega, \theta) * I(x, y), \quad (1)$$

where G is the Gabor function, and σ_x and σ_y are the standard deviations of the Gaussian envelope along the x and y axes. The parameters ω and θ are, respectively, the central frequency and the rotation of the Gabor filter. The values of these parameters σ_x, σ_y, ω, θ are chosen empirically based on tests. In this paper, the Gabor features are extracted by applying Gabor filters with three scales and four orientations.

In summary, the feature variables for each patch are thus obtained as $\mathbf{f}_i = \{\mathbf{f}_{HOG}, \mathbf{f}_{GF}\}$, denoting HOG and GF, respectively.

2.3. Three-dimensional Model-based Cardiac Segmentation

The cardiac segmentation of the LV is performed with a modified 3D-SPASM segmentation method (van Assen et al., 2006). The essential components of an standard ASM (Cootes et al., 1995) scheme are a Point Distribution Model (PDM), an Intensity Appearance Model (IAM), and a model matching algorithm. We use Principal Component Analysis for the PDM, normalized intensity gradient, like in standard ASM, for the IAM, and an adapted fitting algorithm, SPASM, for 3D model to sparse image matching.

2.3.1. Shape Model

The point distribution model (PDM) is a statistical template of the organ of interest represented by corresponding anatomical points or nodes. In our case, the PDM is a surface mesh representing the LV, which includes endocardial and epicardial surfaces. The PDM is built during training by applying Principal Component Analysis to a set of aligned shapes (Cootes et al., 1995) and maintaining eigenvectors corresponding to a predefined percentage of shape variability. The learned shape variability can be modeled as:

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{\Phi}\mathbf{b}, \quad (2)$$

where $\hat{\mathbf{x}}$ is a vector representing the shape, $\bar{\mathbf{x}}$ is the mean shape, $\mathbf{\Phi}$ is the eigenvector matrix and $\mathbf{b}$ is a vector of scaling values for each principal component. By modifying $\mathbf{b}$, different shapes can be defined.

2.3.2. Intensity Model

The IAM learns the appearance around the boundaries of the target organ. The IAM is created at each heart mesh node by sampling the intensity around each node over the image training set. In our case, the IAM models the intensity distribution that characterizes the myocardial boundaries. This is done by sampling a 1D profile of gray-level values for each node along the perpendiculars to the boundary. During training, from voxels sampled for each node i, a mean grey level profile $\bar{\mathbf{g}}_i$ and its corresponding covariance matrix $\mathbf{S}_{\mathbf{g}_i}$ are estimated and stored. We have used a region-based approach for appearance modeling,

i.e., all grey-level profiles of the nodes of the same regions (using a 17-segment model (Cerqueira et al., 2002)) are combined, which leads to models that encode larger variability.

2.3.3. Model Initialization

To determine the initial translation, rotation and scaling of the shape model, we rely on the anatomical landmarks computed in Sec. 2.2.

These landmarks are used (i) to suitably place the initial shape inside the image volume (translation), (ii) to scale the initial shape to the image dataset (scaling) based on the LV apex-to-base distance along the axis of the heart; and (iii) to define the initial orientation of the heart based on the relative position of the valve (rotation). We estimate these initial pose parameters by registering the obtained landmarks to their corresponding points on the mean shape $\bar{\mathbf{x}}$. We then obtain the initial shape $\mathbf{x}^0$.

2.3.4. Matching Algorithm

The last element of the segmentation process is the matching algorithm, with the role of deforming the mesh to match the image data $\mathbf{I}$. Our approach is based on the sparse fitting method SPASM (van Assen et al., 2006). The ASM segmentation searches for optimal boundary points, with the best match between candidate boundary points and a model of intensity features subject to global anatomical statistical constraints (Cootes et al., 1995).

Our algorithm, which is based on the SPASM technique (van Assen et al., 2006), first finds the intersections of the shape with the imaging planes, which could be in any arbitrary orientation in 3D space. These model-plane intersections generate a stack of 2D contours. Each node in these contours is used to compute the closest boundary point in the image plane it belongs to. The candidate boundary point, or feature point, $\mathbf{y}_i$ is estimated by searching along a normal profile to the original model surface at the contour point. Our algorithm uses an intensity model where each candidate point is obtained by selecting the minimal Mahalanobis distance between the sampled profiles $\mathbf{g}_i(\mathbf{y}_i))$ and the mean profiles $\bar{\mathbf{g}}_i$ of the intensity model IAM as:

$$\underset{\mathbf{y}_i}{\operatorname{argmin}} \; (d_i^2(\bar{\mathbf{g}_i}, \mathbf{g}_i(\mathbf{y}_i))). \tag{3}$$

The Mahalanobis distance is computed as:

$$d_i^2(\mathbf{g}_i, \bar{\mathbf{g}}_i) = (\mathbf{g}_i - \bar{\mathbf{g}}_i)^T \mathbf{S}_{\mathbf{g}_i}^{-1} (\mathbf{g}_i - \bar{\mathbf{g}}_i), \tag{4}$$

where $\mathbf{g}_i$ represents the grey-level profile for node i.

The obtained the feature points $(\mathbf{y}_1, ..., \mathbf{y}_n)$ operate as deformation forces propagated to neighboring nodes with a weighting function, normalized by a Gaussian kernel as

$$w(\lambda, \omega) = e^{-\frac{||\lambda - \omega||^2}{2\sigma_p^2}}, \tag{5}$$

where $(||\lambda - \omega||^2)$ is the geodesic distance between source and receiving nodes, and σ_p is the width of the normalizing Gaussian kernel.

Finally, the deformed mesh obtained after propagations is used to obtain a valid shape based on the reference shape model PDM. The parameter vector $\mathbf{b}$ controlling model deformation is calculated by computing an adjustment regarding the previous iteration that best fit the shape to the PDM. The allowed shape instances are limited by the statistical shape description from the training set.

The mesh is deformed for several iterations until the best-fit location is found.

2.4. Segmentation Quality Control

Incorrect segmentation results sometimes are inevitable when processing large population studies due to varying image quality, sub-optimal segmentation parameters or other algorithmic failures. Therefore, when dealing with large amounts of data, it is desirable to have pipeline self-verification capabilities to automatically detect incorrect results either to reprocess them, or to disregard them altogether from later data analysis.

In this paper, the automated segmentation quality control uses an anatomically-motivated definition of correct vs. incorrect LV delineations. Independently of its shape and size, the LV in our CMR studies is expected to have a set of geometrical and appearance characteristics:

- The blood pool has bright MRI contrast.
- The myocardium has dark MRI contrast.
- The blood pool has a quasi-circular shape in SA.
- The myocardial wall thickness changes smoothly.

Similar to Sec. 2.2, we use a database of manual ground-truth segmentations alongside a set of incorrect segmentations generated under controlled conditions to develop a quality control system based on a statistical model of the appearance and geometry of the LV. We learn intensity features associated to these anatomical regions (blood pool and myocardium), distinguishing between incorrect and accurate segmentations using the ground-truth segmentations as a reference.

2.4.1. Blood Pool and Myocardial Feature Description

Based on the previous premises, we define statistical features computed directly on image intensity values of the blood pool and myocardium. In the MRI sequences of the databases used in our experiments, the average intensity is expected to be higher in the blood pool and lower in the myocardial region. The dispersion of the intensity values shown by the standard and absolute deviations is related to texture regularity. Other statistical moment-based features we consider in this work are skewness and kurtosis, which account for higher order information of the intensity probability distribution inside the LV regions. For the statistical descriptors, image intensity ranges were normalized to enable comparison across subjects using a max-min normalization strategy.

Table 3: A summary of the image feature descriptors and their type.

Descriptor	Type
Mean	Statistical
Standard deviation	Statistical
Absolute deviation	Statistical
Skewness	Statistical
Kurtosis	Statistical
Energy	Pattern/GLCM
Entropy	Pattern/GLCM
Contrast	Pattern/GLCM
Homogeneity	Pattern/GLCM
Fractal dimension mean	Fractal dimension
Fractal dimension standard deviation	Fractal dimension
Fractal dimension lacunarity	Fractal dimension

In order to study the gray-scale spatial patterns, we compute gray level co-occurrence matrices (GLCM's) (Haralick et al., 1973). GLCM's characterize the texture of an image by calculating how often pairs of pixels with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. We generate several GLCM's using four orientations with 1 pixel of distance between the pixel of interest and its neighbor, comprising the 4-connected pixels. Statistics on the summation of these four matrices are then used to provide information about the regularity of patterns occurring (as the energy or the entropy) and some information about the patterns themselves (contrast and homogeneity). Inhomogeneous patches have high first order entropy, while a homogeneous one has a low entropy. The final measures of patterns complexity used are based on the fractal dimension (FD) of the image, which measures a ratio of the change in detail to the change in scale. While the run length features could work in only one dimension at a time, the FD works in two dimensions. Using a differential box-counting approach (Bisoi and Mishra, 2001), the FD at each pixel in a slice is computed, resulting in the FD image, and these are aggregated in the mean, standard deviation, and lacunarity (how densely the fractal fills the space).

A summary of the image feature descriptors used can be found in Table 3. We extract these descriptors from the blood pool $\mathbf{f}_i^{BP}$ and the myocardium $\mathbf{f}_i^{M}$ regions separately. The feature vector is composed by both vectors, as $\mathbf{f}_i = \mathbf{f}_i^{BP}, \mathbf{f}_i^{M}$. These features are expected to differ between correct and incorrect segmentations according to the previous assumptions, and they can be used to automatically and robustly evaluate the quality of the end segmentations in population samples.

2.4.2. Training Set Generation

In order to create a training set to learn to discriminate between correct and incorrect segmentation results, we proceed similarly to Sec. 2.2.3. We generate training samples by producing correct and incorrect segmentations as perturbations about the manual ground truth.

There are two main types of possible failures in the LV segmentation: incorrect positioning of the initial shape due to a failure in the initialization and large contour errors due to a segmentation failure. To generate training samples, we try to reproduce these two failures. On the one hand, we translate the contours to reproduce incorrect initializations and, on the other hand, we deform the contours to reproduce incorrect segmentations.

To generate exemplars of correct segmentations, we translate the contours arbitrarily using a random vector $\mathbf{v}$ with uniform distribution and norm under 3 mm. To generate exemplars of incorrect segmentations, we do the same but with random perturbations with norms in the range between 6 and 10 mm. Incorrect segmentations also include deformed contours. To deform the segmentation contours and create incorrect segmentation contours, we fit a spline to the gold-standard contours and then we randomly perturb the control points introducing a displacement with a norm in the range between 3 and 6 mm. These values were chosen empirically based on observations in the experiments. Once the generation of the training samples is finalized, we estimate the feature descriptors for the original and generated exemplars.

2.4.3. Classification of Correct and Incorrect Segmentations

Similar to the automatic initialization, we use a RF classifier (Breiman, 2001) to perform the detection of incorrect segmentations. Given the training labeled database, the RF classifier builds a set of classification trees, which predicts the class label (correct or incorrect) of a sample using a set of features. During training, each branch node of a tree learns features and threshold that results in the best split of the training samples into its child nodes. The splitting process continues recursively until the maximum tree depth is reached. At this time, a leaf node is created and the class distribution of the training samples reaching the leaf node is used to predict the class label of unseen samples. The final classification is based on the majority votes from individually developed tree classifiers in the forest. More details on RF can be found in (Breiman, 2001). In this paper, the random forests for quality control were constructed with 100 trees, with each tree having a maximum depth of 15.

3. Data Description

3.1. Images

Validation is performed using more than 1200 MRI datasets obtained from the Cardiac Atlas Project (CAP) (Fonseca et al., 2011) (see Table 4). CAP is a web-accessible resource (www.cardiacatlas.org), which provides a resource for cardiac image data sharing and atlas-based shape analysis for population

studies. The datasets used in this work are part of two cohorts: asymptomatic volunteers (AV) and patients with myocardial infarction (MI). Imaging protocols are described in (Zhang et al., 2014).

3.1.1. Population Set 1 - asymptomatic volunteers (AV)

The AV group used fast gradient-recalled echo (GRE) imaging with 10-12 short-axis slices with typical parameters: 6 mm thickness, 4 mm gap, field of view 360-400 mm, 256 × 160 matrix, flip angle 20°, TE 3-5 ms, TR 8-10 ms, and pixel size from 1.4 to 2.5 mm/pixel depending on patient size.

3.1.2. Population Set 2 - patients with myocardial infarction (MI)

The MI group used retrospectively gated SSFP cine imaging with typical short-axis slice parameters either a 6 mm slice thickness with 4 mm gap or 8 mm slice thickness with 2 mm gap and a field of view 360-400 mm. Image size was ranging from 138 × 192 to 512 × 512 pixels and pixel size from 0.7-2.5 mm/pixel depending on the patient.

For both groups, sufficient short-axis slices were expected to be acquired to cover the whole heart, and long-axis slices in the four chamber, two chamber and LV outflow tract views.

3.2. Ground Truth

To evaluate the segmentation results, manual contours were also provided by the Cardiac Atlas Project. To evaluate the landmark detection algorithm, landmarks were manually marked for 500 patients by using the interactive medical image processing software GIMIAS v.1.6 (www.gimias.org) (Larrabide et al., 2009).

3.3. Data Preprocessing

Prior to any analysis, from the data provided by the Cardiac Atlas Project, we have discarded incomplete datasets, which included:

- sets with incomplete or inconsistent DICOM tags,
- sets with inconsistent SA image geometry (different image dimensions of slices in a stack)
- sets with missing data (SA or LA series),
- sets with missing manual contours.

A summary of the data used in this work is shown in Table 4.

4. Results

The performance of the landmark detection framework has been evaluated on a subset of group AV, described in Table 4. The segmentation framework has been evaluated in two groups. To study the robustness of the presented method, we did the training exclusively on images from group AV and tested them on part of AV and, also, on an independent dataset, the group MI.

Table 4: Data available for evaluation.

	Complete Datasets* + Manual Contours	Manual Landmarks
Asymptomatic (AV)	1055 cases (855 test + 200 train)	500 cases (300 test + 200 train)
Myocardial infarction (MI)	193 cases (all test)	-

*Complete datasets include SA slices covering the whole heart, 2-chamber and 4-chamber LA view slices with all required DICOM tags.

4.1. Landmark Detection Evaluation

To enable automatic initialization, we have detected four landmarks including the mitral valve hinge points (left and right), the apex and the central axis position, which are described in Sec. 2.2.1. There are 500 cases with manual landmarks from the AV group, as described in the data section. To evaluate the landmark detection, we conducted Monte Carlo Cross-Validation also known as repeated random sub-sampling cross-validation, which consists of repeated rounds of validation conducted on a fixed dataset (Shao, 1993). In each validation round, we randomly split the dataset into training (200 sets) and validation (300 sets) data. Ten validation rounds were performed and the automatic landmarks were compared with the manual ones using the Euclidean distance.

We tested the performance of the Random Forest (RF) regression as a function of the number of trees based on a fixed depth of 15. Our results are shown in Fig. 4, which indicate that stable error rates for the RF are achieved from 50 trees, which is the value fixed in the remaining experiments.

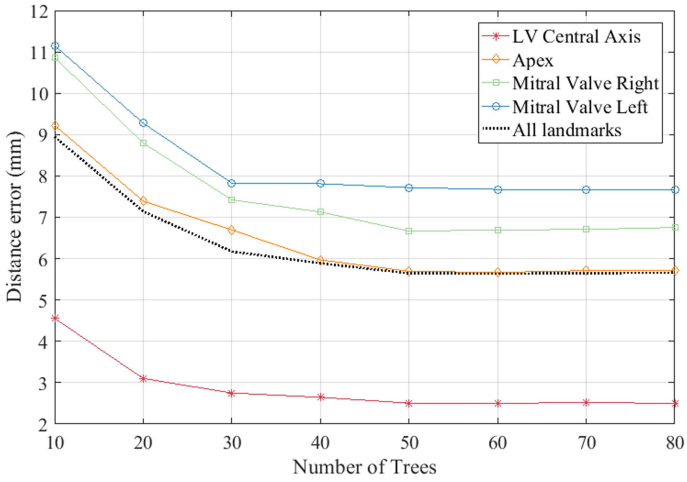


Figure 4: Performance results of increasing the number of trees in the forest for landmark detection.

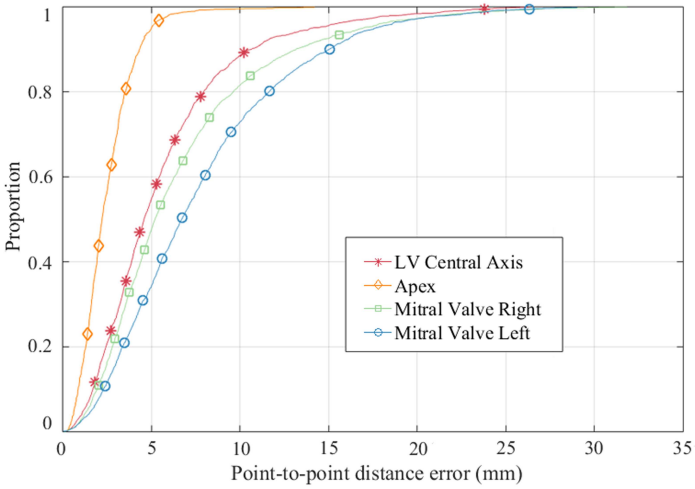


Figure 5: Cumulative error distribution curves for each landmark detection.

The statistics of the landmark detection error for each landmark are summarized in Table 5, which includes the median, average, and standard deviation (SD) values. In the case of the LV central axis, the LV centroid position is computed on each SA slice defining the axis, the result on Table 5 is the average distance error. It can be seen that the median errors vary between 2.19 mm for the ventricular axis to 6.69 mm for the more challenging mitral valve points. To have an indication of the relative distribution of the errors in the sample, cumulative percentages corresponding to the percentage of test images for which the error is less than a specific value are produced and shown in Fig 5. The percentage values show that more than 90 percent of the landmarks are identified within 1.5 cm error. Specifically, the central axis is detected with an error less than a 5 mm in more than 90% of the cases. In these experiments, we have not discarded the cases finally detected as 'fails' by the final automated segmentation quality control.

Table 5: Point-to-point (P2P) distance error statistics in landmark detection (subset of group AV).

	Median (mm)	Mean (mm)	SD (mm)
LV Central axis	2.19	2.47	1.48
Apex	4.56	5.60	4.13
Mitral Valve Right	5.19	6.57	4.90
Mitral Valve Left	6.69	7.76	5.08

4.2. Quality Control Evaluation

The performance of the proposed segmentation quality control method is evaluated using both data groups, AV and MI. For quantitative evaluation, the accuracy, sensitivity and specificity of the detection were calculated. The performance measurement metrics used is defined as:

$$\begin{aligned} Accuracy &= \frac{TP+TN}{TP+FP+TN+FN}, \\ Sensitivity &= \frac{TP}{TP+FP}, \\ Specificity &= \frac{TN}{TN+FN}, \end{aligned}$$

where TP (true positive) represents the number of times the classifier guessed a correct positive value; TN (true negative) is the number of times the classifier correctly guessed a negative value; FP (false positive) is the number of times the classifier predicted a positive value, but the correct value was negative; and FN (false negative) is the inverse of FP. In this failure detection problem, TP has been defined as the case with less than a certain segmentation error and FN with a larger error. We tested the performance of the classifier as a function of the threshold defined as a wrong segmentation. We have considered as failed segmentations those with an error over 4 mm regarding the ground truth.

The results are summarized in Table 6, indicating over 96% accuracy for the quality control. The obtained results are consistent for both groups. To assess the performance of the quality control, it was evaluated as a function of the different texture descriptors and their combinations as shown in Table 6. Statistical descriptors perform with better accuracy and sensitivity than pattern and fractal descriptors but have low specificity. Combining statistical and pattern descriptors improves the accuracy, sensitivity and specificity measures to over 0.90, while combining all three types of descriptors further enhances the overall performance to over 0.96.

Finally, during the development of the framework, we evaluated various machine learning and classification approaches to optimize the quality control. We found SVM (Cortes and Vapnik, 1995) and RF to perform the best for our application, with RF producing the best results as shown in Table 7.

4.3. Fully Automatic Segmentation Errors

The anatomical landmarks detected in the previous section are then used to initialize the 3D SPASM model, which provides a 3D final segmentation mesh. For evaluation, 2D contours are obtained for each slice by using the intersection between the mesh and the SA slices. To show the flexibility and robustness of the technique to different databases, we have evaluated our method using the AV and MI cases. We calculated for accuracy evaluation the point-to-point (P2P) errors between the manual and the automatic segmentations. We report the results of segmenting the cases described in Sec. 3, which are summarized in Table 4.

Table 6: Performance of segmentation quality control using different feature types.

	Group AV			Group MI		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Statistical	0.87	0.88	0.75	0.90	0.95	0.66
Fractal	0.58	0.53	0.88	0.63	0.59	0.83
Pattern	0.72	0.69	0.92	0.55	0.46	1
Statistical + Fractal	0.88	0.89	0.80	0.92	0.95	0.73
Statistical + Pattern	0.92	0.91	0.95	0.89	0.89	0.9
Fractal + Pattern	0.74	0.71	0.92	0.79	0.75	1
All	**0.96**	**0.96**	**0.95**	**0.96**	**0.98**	**0.93**

Table 7: Performance of segmentation quality control using a Support Vector Machine (SVM) or a Random Forest (RF) classifier.

		Accuracy	Sensitivity	Specificity
SVM	Group AV	0.90	0.91	0.88
	Group MI	0.93	0.94	0.90
RF	Group AV	0.96	0.96	0.95
	Group MI	0.96	0.98	0.93

To assess the validity of our approach, the results of the segmentation using our algorithm are compared against the manual contours of the LV. P2P segmentation error statistics are given in Table 8, including the median, average, and standard deviation errors for the initialization step and the final segmentation (with and without quality control of segmentation). Over the full test set, Table 8 shows that the proposed method has a final median error of 2.51 mm and 2.82 mm, respectively, for endocardial and epicardial segmentations for group AV and 2.41 mm and 2.47 mm for group MI. The mean and the median of the segmentation with the failures included are different, probably due to some failed initializations. Thanks to the automatic quality control, these cases are identified and discarded. We have also included in the table the errors for the initialization step, indicating a median error around 4 mm for both databases, then refined by the segmentation step. Fig. 6 shows the effect of the quality control, most cases have less than 4 mm error after discarding failures.

To specifically evaluate the relevance of the automatic initialization, Table 9 presents the point-to-point segmentation errors obtained by using (1) manual initialization (available for 300 cases) vs. (2) the proposed automatic initialization. The errors are relatively similar, and slightly better by using automatic initialization after segmentation quality control, showing the accuracy and relevance of our method. The mean and median of the results for manual initialization are even closer, showing that maximum errors have less influence than in the case of the automatic initialization. This is indicative of the importance of an automatic quality control in automatic image segmentation.

Table 8: P2P error statistics for automatic image segmentation.

		Before Automatic Quality Control						After Automatic Quality Control					
		Initialization			Segmentation			Initialization			Segmentation		
		Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD
		mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm
Group AV	Endocardium	3.81	4.32	1.88	2.60	3.14	1.83	3.77	4.08	1.42	2.51	2.55	0.41
	Epicardium	4.16	4.75	1.94	2.95	3.47	1.74	4.05	4.40	1.53	2.82	2.89	0.53
Group MI	Endocardium	3.92	4.42	1.82	2.53	3.46	3.03	3.72	3.92	1.12	2.41	2.33	0.56
	Epicardium	4.23	4.63	1.92	2.59	3.63	3.06	3.78	3.94	1.18	2.47	2.46	0.51

Table 9: P2P error statistics for automatic image segmentation in the subset of group AV (manual versus automatic initialization).

	Before Automatic Quality Control						After Automatic Quality Control					
	Segmentation with manual init			Segmentation with auto init			Segmentation with manual init			Segmentation with auto init		
	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD
	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm
Endocardium	2.62	2.87	1.33	2.63	3.28	1.81	2.58	2.58	0.34	2.51	2.57	0.35
Epicardium	3.03	3.27	1.18	3.12	3.60	1.62	2.98	2.99	0.49	2.92	2.96	0.50

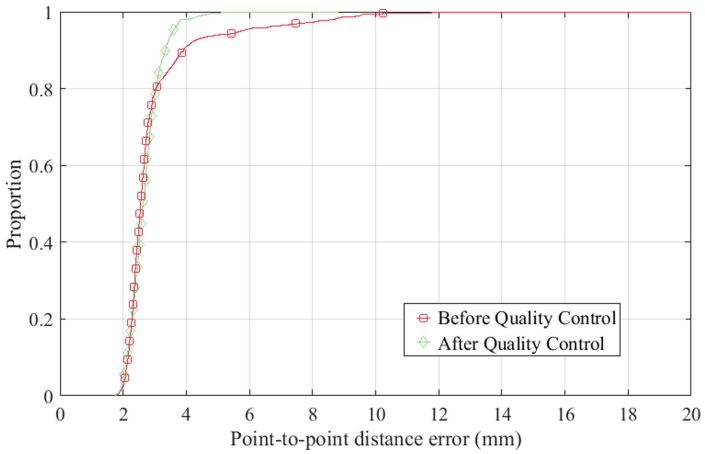


Figure 6: Cumulative error distribution curves for myocardial segmentation error, before and after the quality control.

4.4. Illustrative Results

Fig. 7 shows three examples of segmentation and landmark detection results for AV group. First and second columns show the landmarks detected by our method (red) with the manually defined points (green). The final segmentation results are also shown, both the automatic (red) and the manual (green) delineations. The satisfactory quantitative analysis (Tables 5 and 8) are confirmed with our visual analysis. Larger errors are produced in the mitral valve and the central axis is detected with lower error. The segmentation of the endocardium is performed slightly better than the epicardium. This can be explained by the high variability of the surrounding tissues and the lack of contrast on some parts of the boundary, while the endocardium only interfaces with the myocardium. The errors on the endocardium are produced mainly around the papillary muscles.

Fig. 8 shows four examples of segmentation results from the MI database. The first thing to observe is the high variability of the images in appearance and shape between them and, also, compared to the AV examples. Although we have trained our method with sets from group AV, similar results have been obtained.

Fig. 9 shows four examples of wrong segmentation results for both databases. These cases were detected as failures by the quality control and discarded from the final evaluation results. Note that the main problem is that the RV is considered part of the LV.

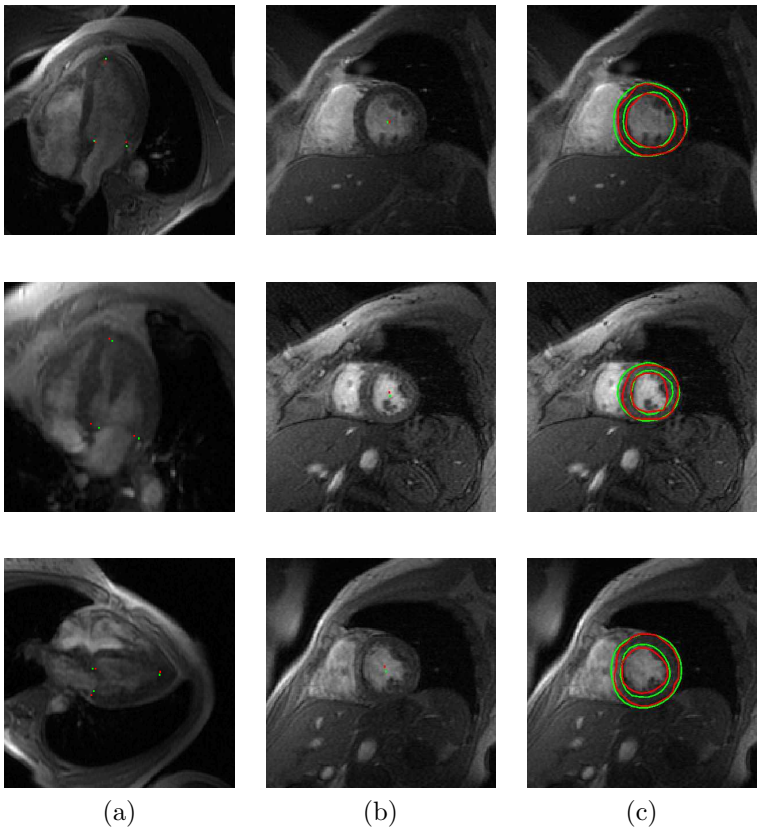


Figure 7: Examples of landmark detection and segmentation results for AV datasets: landmark detection, (a) mitral valve points, apex and (b) the central axis and (c) the automatic segmentation contours with manual contours. Our method is marked as red and the ground truth as green.

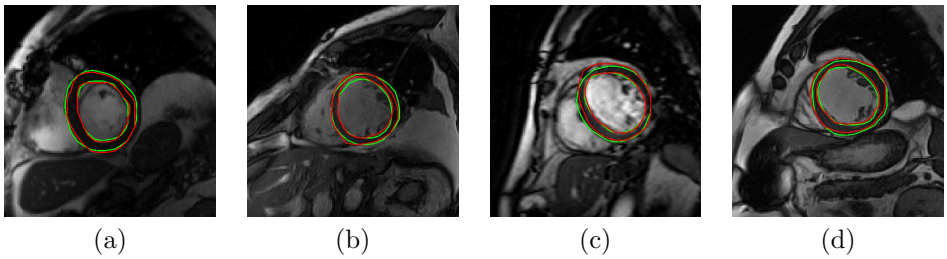
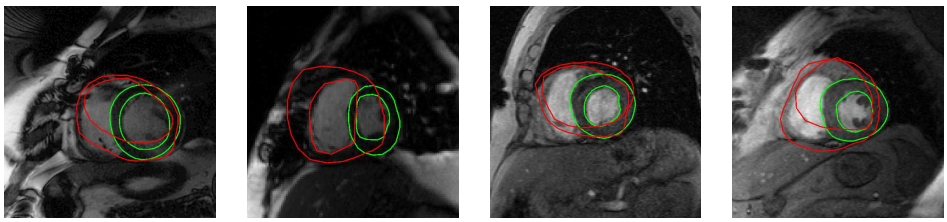


Figure 8: Examples of segmentation results for MI datasets, the automatic segmentation contours with manual contours. Our method is marked as red and the ground truth as green.



(a) (b) (c) (d)

Figure 9: Examples of discarded segmentation results, thanks to the quality control, for groups (a,b) MI and (c,d) AV. Automatic segmentation contours are shown with manual contours, our method is marked as red and the ground truth as green.

5. Discussion and Conclusion

We described and validated a fully automatic approach to perform the analysis of cardiac MRI in large MRI studies, which cannot be processed through user interaction. The technique combines an RF-based landmark detector, a 3D-SPASM segmentation algorithm and an anatomically-driven classifier for automated segmentation quality control. Different to existing techniques, our method predicts the correction of the initial ventricular axis and anatomical landmarks by using a pose correction regression model. Subsequently, the landmarks are used to initialize a three-dimensional LV segmentation based on ASM search. After segmentation, a quality control scores the quality of the segmentations and discards failures from later data analytics.

The novelty of the technique lies in using a regression/classification approach based on large-scale synthetic examples of correct and incorrect LV characteristics. It is easy to implement by using RFs and feature descriptors widely used in medical image analysis, and thus are well-suited for cardiovascular population studies. A complete validation based on two large-scale public databases, groups AV and MI, demonstrates the potential of our fully automatic segmentation for large-scale populations. To the best of our knowledge, this is the first attempt to parse more than 1200 MRI datasets from both normal and abnormal cases using a fully automatic pipeline.

The algorithm demonstrated good computational performance; the average running time to produce all the anatomical landmarks is 6.87 s per subject and the full segmentation procedure with the score takes on average 28 s per subject. All experiments were performed using MATLAB (The Mathworks Inc., Nantucket, MA) with non-optimized code running on a 2.4 GHz Windows PC with 8 GB of RAM.

We believe that the concept behind our approach (i.e. using artificial false examples and patch/texture-based regression) can be applied to other problems in cardiac MRI analysis. For example, automatic correction of slice misalignments, which is an important challenge, can be modelled by similarly building a synthetic training sample of misalignments (by artificially displacing SA slices)

and by learning the required corrections through patch-based regression.

The global measures used for quality control and the corresponding performance suggest these could be incorporated into the segmentation process to further improve the segmentation results. However, their use within our Active Shape Model framework is more complicated as it corresponds to an inherently local search problem, which requires different descriptors, such as the local normal appearance profiles typically used to fit the boundaries locally to the image evidence.

Future perspectives for this work include (i) applying it to other public large databases (e.g. the UK Biobank), which would further state the scalability and robustness of the method and (ii) use of the results to compute and analyze cardiac measures as EF or volume and extract new clinical knowledge of MRI-based phenotypes of cardiovascular health and disease.

References

van Assen, H.C., Danilouchkine, M.G., Frangi, A.F., Ordas, S., Westenberg, J.J., Reiber, J.H., Lelieveldt, B.P., 2006. SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data. Medical Image Analysis 10, 286–303.

Avendi, M., Kheradvar, A., Jafarkhani, H., 2016. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Medical Image Analysis 30, 108–119.

Bai, W., Shi, W., Ledig, C., Rueckert, D., 2015. Multi-atlas segmentation with augmented features for cardiac MR images. Medical Image Analysis 19, 98–109.

Berbari, R.E., Bloch, I., Redheuil, A., Angelini, E., Mousseaux, E., Frouin, F., Herment, A., 2007. An automated myocardial segmentation in cardiac MRI, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), pp. 4508–4511.

Billet, F., Sermesant, M., Delingette, H., Ayache, N., 2009. Cardiac motion recovery and boundary conditions estimation by coupling an electromechanical model and Cine-MRI data, in: Proceedings of the International Conference on Functional Imaging and Modeling of the Heart (FIMH), Springer. pp. 376–385.

Bisoi, A.K., Mishra, J., 2001. On calculation of fractal dimension of images. Pattern Recognition Letters 22, 631–637.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Cerqueira, M.D., Weissman, N.J., Dilsizian, V., Jacobs, A.K., Kaul, S., Laskey, W.K., Pennell, D.J., Rumberger, J.A., Ryan, T., Verani, M.S., 2002. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. Circulation 105, 539–542.

Cootes, T., Taylor, C.J., Cooper, D.H., Graham, J., 1995. Active shape models - their training and application. Computer Vision and Image Understanding 61, 38–59.

Cordero-Grande, L., Vegas-Sanchez-Ferrero, G., de-la Higuera, P.C., San-Roman-Calvar, J.A., Revilla-Orodea, A., Martin-Fernendez, M., Alberola-Lopez, C., 2011. Unsupervised 4D myocardium segmentation with a markov random field based deformable model. Medical Image Analysis 15, 283–301.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning 20, 273–297.

Cousty, J., Najman, L., Couprie, M., Clm̀ent-Guinaudeau, S., Goissen, T., Garot, J., 2010. Segmentation of 4D cardiac MRI: Automated method based on spatio-temporal watershed cuts. Image and Vision Computing 28, 1229–1243.

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886–893.

Eslami, A., Karamalis, A., Katouzian, A., Navab, N., 2013. Segmentation by retrieval with guided random walks: Application to left ventricle segmentation in MRI. Medical Image Analysis 17, 236–253.

Fonseca, C.G., Backhaus, M., Bluemke, D.A., Britten, R.D., Chung, J.D., Cowan, B.R., Dinov, I.D., Finn, J.P., Hunter, P.J., Kadish, A.H., Lee, D.C., Lima, J.A., Medrano-Gracia, P., Shivkumar, K., Suinesiaputra, A., Tao, W., Young, A.A., 2011. The Cardiac Atlas Project-an imaging database for computational modeling and statistical atlases of the heart. Bioinformatics 27, 2288–2295.

Frangi, A.F., Niessen, W.J., Viergever, M.A., 2001. Three-dimensional modeling for functional analysis of cardiac images: a review. IEEE Transactions on Medical Imaging 20, 2–25.

Frangi, A.F., Taylor, Z.A., Gooya, A., 2016. Precision imaging: more descriptive, predictive and integrative imaging. Medical Image Analysis 33, 27–32.

Gabor, D., 1946. Theory of communication. Journal of the Institute of Electrical Engineers 93, 429–459.

Grosgeorge, D., Petitjean, C., Caudron, J., Fares, J., Dacher, J.N., 2011. Automatic cardiac ventricle segmentation in MR images: a validation study. International Journal of Computer Assisted Radiology and Surgery 6, 573–581.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics 3, 610–621.

Jolly, M.P., 2006. Automatic segmentation of the left ventricle in cardiac MR and CT images. International Journal of Computer Vision 70, 151–163.

Jolly, M.P., Guetter, C., Lu, X., Xue, H., Guehring, J., 2012. Automatic segmentation of the myocardium in cine MR images using deformable registration, in: Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM), Springer. pp. 98–108.

Kadish, A.H., Bello, D., Finn, J.P., Bonow, R.O., Schaechter, B., Subacius, H., Albert, C., Daubert, J.P., Fonseca, C.G., Goldberger, J.J., 2009. Rationale and design for the defibrillators to reduce risk by magnetic resonance imaging evaluation (DETERMINE) trial. Journal of Cardiovascular Electrophysiology 20, 982–987.

Katouzian, A., Konofagou, E., Prakash, A., 2006. A new automated technique for left and right-ventricular segmentation in magnetic resonance imaging, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), pp. 3074–3077.

Kramer, C.M., Barkhausen, J., Flamm, S.D., Kim, R.J., Nagel, E., 2013. Standardized cardiovascular magnetic resonance (CMR) protocols 2013 update. Journal of Cardiovascular Magnetic Resonance 15, 91.

Larrabide, I., Omedas, P., Martelli, Y., Planes, X., Nieber, M., Moya, J., Butakoff, C., Sebastian, R., Camara, O., De Craene, M., Bijnens, B., Frangi, A.F., 2009. GIMIAS: an open source framework for effcient development of research tools and clinical prototypes, in: Proceedings of the International Conference on Functional Imaging and Modeling of the Heart (FIMH), Springer. pp. 417–426.

Lin, X., Cowan, B., Young, A., 2006. Automated detection of left ventricle in 4D MR images: Experience from a large study, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 728–735.

Lorenzo-Valdés, M., Sanchez-Ortiz, G.I., Elkington, A.G., Mohiaddin, R.H., Rueckert, D., 2004. Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. Medical Image Analysis 8, 255–265.

Lötjönen, J., Kivistö, S., Koikkalainen, J., Smutek, D., Lauerma, K., 2004. Statistical shape model of atria, ventricles and epicardium from short- and long-axis MR images. Medical Image Analysis 8, 371–386.

Lu, X., Georgescu, B., Jolly, M.P., Guehring, J., Young, A., Cowan, B., Littmann, A., Comaniciu, D., 2010. Cardiac anchoring in MRI through context modeling, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 383–390.

Lu, X., Georgescu, B., Littmann, A., Mueller, E., Comaniciu, D., 2009a. Discriminative joint context for automatic landmark set detection from a single cardiac MR long axis slice, in: Proceedings of the International Conference on Functional Imaging and Modeling of the Heart (FIMH), Springer. pp. 457–465.

Lu, X., Jolly, M.P., 2013. Discriminative context modeling using auxiliary markers for LV landmark detection from a single MR image, in: Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM), Springer. pp. 105–114.

Lu, X., Wang, Y., Georgescu, B., Littman, A., Comaniciu, D., 2011. Automatic delineation of left and right ventricles in cardiac MRI sequences using a joint ventricular model, in: Proceedings of the International Conference on Functional Imaging and Modeling of the Heart (FIMH), Springer. pp. 250–258.

Lu, Y., Connelly, K.A., Dick, A.J., Wright, G.A., Radau, P.E., 2013. Automatic functional analysis of left ventricle in cardiac cine MRI. Quantitative Imaging in Medicine and Surgery 3, 200–209.

Lu, Y., Radau, P., Connelly, K., Dick, A., Wright, G.A., 2009b. Segmentation of left ventricle in cardiac cine MRI: An automatic image-driven method, in: Proceedings of the International Conference on Functional Imaging and Modeling of the Heart (FIMH), Springer. pp. 339–347.

Lynch, M., Ghita, O., Whelan, P., 2006. Automatic segmentation of the left ventricle cavity and myocardium in MRI data. Computational Biology in Medicine 36, 389–407.

Mahapatra, D., 2013. Landmark detection in cardiac MRI using learned local image statistics, in: Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM), Springer. pp. 115–124.

Margeta, J., Geremia, E., Criminisi, A., Ayache, N., 2012. Layered spatio-temporal forests for left ventricle segmentation from 4D cardiac MRI data, in: Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM), Springer. pp. 109–119.

Medrano-Gracia, P., Cowan, B.R., Suinesiaputra, A., Young, A.A., 2015. Challenges of cardiac image analysis in large-scale population-based studies. Current Cardiology Reports 17, 9.

Murdoch, T., Detsky, A., 2013. The inevitable application of big data to health care. JAMA 309, 1351–1352.

Nambakhsh, C.M., Yuan, J., Punithakumar, K., Goela, A., Rajchl, M., Peters, T.M., Ayed, I.B., 2013. Left ventricle segmentation in MRI via convex relaxed distribution matching. Medical Image Analysis 17, 1010–1024.

Ordas, S., Oubel, E., Leta, R., Carreras, F., Frangi, A.F., 2007. A statistical shape model of the whole heart and its application to model-based segmentation, in: Proceedings SPIE Medical Imaging.

Pednekar, A., Kurkure, U., Muthupillai, R., Flamm, S., Kakadiaris, I., 2006. Automated left ventricular segmentation in cardiac MRI. IEEE Transactions on Biomedical Engineering 53, 1425–1428.

Peng, P., Lekadir, K., Gooya, A., Shao, L., Petersen, E.S., Frangi, A.F., 2016. A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. Magnetic Resonance Materials in Physics, Biology and Medicine , 1–41.

Petitjean, C., Dacher, J.N., 2011. A review of segmentation methods in short axis cardiac MR images. Medical Image Analysis 15, 169–184.

Piazzese, C., Carminati, M.C., Colombo, A., Krause, R., Potse, M., Auricchio, A., Weinert, L., Tamborini, G., Pepi, M., Lang, R.M., Caiani, E.G., 2016. Segmentation of the left ventricular endocardium from magnetic resonance images by using different statistical shape models. Journal of Electrocardiology 49, 383–391.

Ponikowski, P., Voors, A.A., Anker, S.D., Bueno, H., Cleland, J.G.F., Coats, A.J.S., Falk, V., González-Juanatey, J.R., Harjola, V.P., Jankowska, E.A., Jessup, M., Linde, C., Nihoyannopoulos, P., Parissis, J.T., Pieske, B., Riley, J.P., Rosano, G.M.C., Ruilope, L.M., Ruschitzka, F., Rutten, F.H., van der Meer, P., 2016. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. European Heart Journal .

Queirós, S., Vilaça, J.L., Morais, P., Fonseca, J.C., Dhooge, J., Barbosa, D., 2015. Fast left ventricle tracking in CMR images using localized anatomical affine optical flow, in: Proceedings SPIE Medical Imaging.

Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A., Wright, G., 2009. Evaluation framework for algorithms segmenting short axis cardiac MRI. The MIDAS Journal - Cardiac MR Left Ventricle Segmentation Challenge .

Rueckert, D., Glocker, B., Kainz, B., 2016. Learning clinically useful information from images: Past, present and future. Medical Image Analysis 33, 13–18.

Shao, J., 1993. Linear model selection by cross-validation. Journal of the American Statistical Association 88, 486–494.

Slomka, P.J., Dey, D., Sitek, A., Motwani, M., Berman, D.S., Germano, G., 2017. Cardiac imaging: working towards fully-automated machine analysis & interpretation. Expert Review of Medical Devices 14, 197–212.

Suinesiaputra, A., Cowan, B.R., Al-Agamy, A.O., Elattar, M.A., Ayache, N., Fahmy, A.S., Khalifa, A.M., Medrano-Gracia, P., Jolly, M.P., Kadish, A.H., Lee, D.C., Margeta, J., Warfield, S.K., Young, A.A., 2014. A collaborative

resource to build consensus for automated left ventricular segmentation of cardiac MR images. Medical Image Analysis 18, 50–62.

Suinesiaputra, A., Cowan, B.R., Finn, J.P., Fonseca, C.G., Kadish, A.H., Lee, D.C., Medrano-Gracia, P., Warfield, S.K., Tao, W., Young, A.A., 2012. Left ventricular segmentation challenge from cardiac MRI: A collation study, in: Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges (STACOM), Springer. pp. 88–97.

Suinesiaputra, A., Medrano-Gracia, P., Cowan, B., Young, A., 2015. Big heart data: Advancing health informatics through data sharing in cardiovascular imaging. IEEE Journal of Biomedical and Health Informatics 19, 1283–1290.

Tavakoli, V., Amini, A.A., 2013. A survey of shaped-based registration and segmentation techniques for cardiac images. Computer Vision and Image Understanding 117, 966–989.

Tobon-Gomez, C., Sukno, F.M., Butakoff, C., Huguet, M., Frangi, A.F., 2012. Automatic training and reliability estimation for 3D ASM applied to cardiac MRI segmentation. Physics in Medicine and Biology 57, 4155–4174.

Tsadok, Y., Petrank, Y., Sarvari, S., Edvardsen, T., Adam, D., 2013. Automatic segmentation of cardiac MRI cines validated for long axis views. Computerized Medical Imaging and Graphics 37, 500–511.

Zhang, H., Wahle, A., Johnson, R.K., Scholz, T.D., Sonka, M., 2010. 4-D cardiac MR image analysis: left and right ventricular morphology and function. Transactions on Medical Imaging 29, 350–364.

Zhang, X., Cowan, B.R., Bluemke, D.A., Finn, J.P., Fonseca, C.G., Kadish, A.H., Lee, D.C., Lima, J.A.C., Suinesiaputra, A., Young, A.A., Medrano-Gracia, P., 2014. Atlas-based quantification of cardiac remodeling due to myocardial infarction. PLoS ONE 9, e110243.

Zheng, Y., Lu, X., Georgescu, B., Littmann, A., Mueller, E., Comaniciu, D., 2009. Automatic left ventricle detection in MRI images using marginal space learning and component-based voting, in: Proceedings SPIE Medical Imaging, pp. 1–12.

Zhu, Y., Papademetris, X., Duncan, J., Sinusas, A., 2007. Cardiac MR image segmentation with incompressibility constraint, in: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 185–188.

Zhu, Y., Papademetris, X., Sinusas, A.J., Duncan, J.S., 2010. Segmentation of the left ventricle from cardiac MR images using a subject-specific dynamical model. IEEE Transactions on Medical Imaging 29, 669–687.