Combining Deep Learning and Level Set for the Automated Segmentation of the Left Ventricle of the Heart from Cardiac Cine Magnetic Resonance

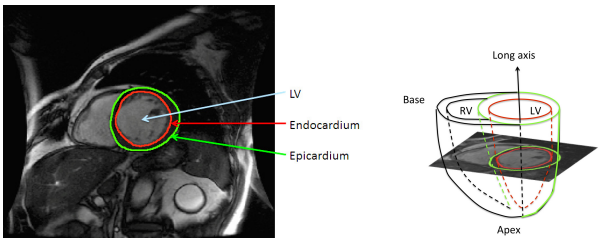
Tuan Anh Ngo[*,1], Zhi Lu[†,1] and Gustavo Carneiro[‡,2]

[*] *Vietnam National University of Agriculture, Vietnam*
[†] *The University of South Australia, Australia*
[‡] *Australian Centre for Visual Technologies, The University of Adelaide, Australia*

Abstract

We introduce a new methodology that combines deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance (MR) data. This combination is relevant for segmentation problems, where the visual object of interest presents large shape and appearance variations, but the annotated training set is small, which is the case for various medical image analysis applications, including the one considered in this paper. In particular, level set methods are based on shape and appearance terms that use small training sets, but present limitations for modelling the visual object variations. Deep learning methods can model such variations using relatively small amounts of annotated training, but they often need to be regularised to produce good generalisation. Therefore, the combination of these methods brings together the advantages of both approaches, producing a methodology that needs small training sets and produces accurate segmentation results. We test our methodology on the MICCAI 2009 left ventricle segmentation challenge database (containing 15 sequences for training, 15 for validation and 15 for testing), where our approach achieves the most accurate results in the semi-automated problem and state-of-the-art results for the fully automated challenge.

[1] T. Ngo and Z. Lu developed this work while they were with the University of Adelaide.
[2] Corresponding author.



(a) mid-ventricular image (b) heart model

Figure 1: LV segmentation from cardiac cine MR imaging (Radau et al. (2009)) (a), and a 3-D model of the heart with respective MR image, representing one of the volume slices (b).

Keywords: Deep learning, Level set method, Segmentation of the Left Ventricle of the Heart, Cardiac Cine Magnetic Resonance.

1. Introduction

Medical image analysis segmentation problems are unique in the sense that they require highly accurate results, but at the same time provide relatively small annotated training sets. A typical example is the segmentation of the endocardium and epicardium from the left ventricle (LV) of the heart using cardiac cine Magnetic Resonance (MR), as shown in Fig. 1. The LV segmentation is necessary for the assessment of the cardiovascular system function and structure and needs to be accurate for a precise diagnosis, but current public databases do not present large annotated training sets (Petitjean and Dacher (2011); Radau et al. (2009)). Therefore, one of the main research topics in this field is how to obtain the precision required with these small training sets.

The main techniques being explored for the automated segmentation of the endocardium and epicardium from cardiac cine MR are based on active contour models, machine learning models, and integrated active contour and machine learning models. Active contour models (Kass et al. (1988); Osher and Sethian (1988)) represent one of the most successful methodologies in the field, and they are based on an optimisation that minimises an energy functional that varies the shape of a contour using internal and external constraints. The energy to

bend, stretch or shrink a contour is represented by the internal constraints, while the external constraints use the observed data (e.g., image) to move the contour towards (or away from) certain appearance features (such as edges). These constraints are usually designed by hand based on shape and appearance priors that use small or no annotated training sets. Although successful, active contour models are based on low-complexity shape and appearance models that are usually unable to robustly model all variation present in the visual object of interest studied in several medical image analysis problems.

The advent of machine learning methods to medical image analysis (Cootes et al. (1995); Georgescu et al. (2005); Zheng et al. (2008)) has addressed this issue by estimating more complex shape and appearance models using annotated training sets. However, the accuracy requirements found in medical image analysis applications usually mean that these models need to be quite complex in order to allow the learning of all appearance and shape variations found in the annotated training set, and as a consequence, this training set has to be large and rich. The issue in machine learning based models then becomes centred on the acquisition of comprehensive annotated training sets, which is a particularly complicated task in medical image analysis. Therefore, in order to reduce the model complexity and consequently, the need for large and rich training sets, a natural idea is to combine the prior information of active contour models with the learned information of machine learning models. The most dominant approach in this direction is the integration of active contour models and Markov random fields (Cobzas and Schmidt (2009); Huang et al. (2004); Tsechpenakis and Metaxas (2007)), but the main issue of these approaches is that these models are in general quite complex, and as a result they still require large amounts of training data.

In this paper, we propose a new automated segmentation approach for the endocardial and epicardial borders of the left ventricle (LV) from all slices of the end diastole (ED) and end systole (ES) cardiac phases of an MR cine study, where the ED and ES volumes are manually selected by the user. This proposed approach combines an active contour model (distance regularised level sets) (Li

et al. (2010)) with a machine learning approach (deep belief network) (Hinton and Salakhutdinov (2006)). This is a sensible combination because this problem does not usually have comprehensive training sets available, but still requires high segmentation accuracy (Radau et al. (2009)). Specifically, we explore the fact that the prior information explored by the level set method reduces the need of using highly complex machine learning models (requiring large training sets), but the limitations of this prior information indicates the need of a machine learning method that can reliably model the shape and appearance of the LV. However, this method must be able to be robustly trained with a limited number of annotated training images, which is the exactly one of the advantages behind deep belief network training (Carneiro et al. (2012); Carneiro and Nascimento (2013)). We show that this combination leads to competitive segmentation accuracy results on the MICCAI 2009 LV segmentation challenge database (Radau et al. (2009)), which does not contain a large training set and that has been tested by several different methodologies. Specifically, our experiments show that our approach produces the best result in the field when we rely on a semi-automated segmentation (i.e., with manual initialisation). Also, our fully automated approach produces a result that is on par with the current state of the art on the same database (Jolly (2009)).

1.1. Contributions

The main contributions of our approach are the following: 1) structured output for the region of interest (ROI) of the LV using a deep belief network (DBN), 2) structured output for the delineation of the endocardial and epicardial borders using another DBN, and 3) extension to the distance regularised level set method (DRLS) (Li et al. (2010)) that takes the estimated ROI from innovation (1) (above) to initialise the optimisation process and the delineation from innovation (2) to constrain the level set evolution. One advantage of using DBN models lies in the need of smaller training sets (Hinton and Salakhutdinov (2006)) compared to other machine learning methods (Cobzas and Schmidt (2009); Huang et al. (2004); Tsechpenakis and Metaxas (2007); Cortes and Vap-

nik (1995); Freund and Schapire (1995)). Another advantage of our method is the improved accuracy brought by the integration of the DBN and DRLS, when compared to the accuracy of the DBN and DRLS independently. Finally, compared to our preliminary papers (Ngo and Carneiro (2013, 2014)), this work presents the following contributions: 1) detection and segmentation of the epicardial border, and 2) comparison of our epicardium segmentation results (in addition to the endocardium segmentation already presented in (Ngo and Carneiro (2013, 2014))) with the state of the art.

2. Literature Review

We focus this work on the segmentation of the endocardial and epicardial borders of the LV from short axis cine MR images (see Fig. 1), so we explore the literature for this application, but in principle our proposed methodology is general enough to be extended to other applications (this extension is out of the scope of this paper). This segmentation has several challenges, which include the lack of gray level homogeneity of LV among different cases (due to blood flow, papillary muscles and trabeculations) and the low resolution of the apical and basal images (Petitjean and Dacher (2011)). According to (Petitjean and Dacher (2011)), current LV segmentation approaches can be classified based on three characteristics: 1) segmentation method (region and edge based, pixel classification, deformable models, active appearance and shape models), 2) prior information (none, weak, and strong), and 3) automated localisation of the heart (time-based or object detection). Furthermore, their analysis (Petitjean and Dacher (2011)) of the MICCAI 2009 challenge results (Radau et al. (2009)) indicates that image-based methodologies (Lu et al. (2009); Huang et al. (2009); Uzunbaş et al. (2012)) (e.g., thresholding, or dynamic programming applied to image segmentation results) produce the highest accuracy, but have the drawbacks of requiring user interaction and of being unable to assess the ventricular surface in all cardiac phases. More sophisticated methodologies (O'Brien et al. (2009); Schaerer et al. (2010); Jolly (2009)) demonstrate how to handle these

challenges, but they show slightly less accurate results. Also, by making the technique specific to the LV segmentation, some methodologies (Lu et al. (2009); Huang et al. (2009); Constantinides et al. (2012); Uzunbaş et al. (2012)) present more accurate results when compared to more general approaches (O'Brien et al. (2009); Wijnhout et al. (2009)). The main conclusion reached by the authors of the review (Petitjean and Dacher (2011)) is that the methodology presented by (Jolly (2009)) is the most competitive because it is fully automatic and offers the best compromise between accuracy and generalisation. Therefore, we regard Jolly's approach (Jolly (2009)) as our main competitor for the fully automated case. For the semi-automated case, the most competitive methods in the MICCAI 2009 challenge was developed by (Huang et al. (2009)) and (Uzunbaş et al. (2012)), so we consider them to be our main competitors for the semi-automated case.

Structured inference and learning is the classification problem involving a structured output (BakIr (2007)), such as the case for segmentation tasks, where the classification is represented by a multi-dimensional binary vector. Although most of the current work in computer vision and machine learning is focused on the large margin structured learning formulation (Tsochantaridis et al. (2005)), one of the most natural ways to represent a structured learning is with a multilayer perceptron (MLP), where the output layer consists of a multi-dimensional binary vector denoting the segmentation (Collins (2002)). One of the recent breakthroughs in the field was the discovery of an efficient learning algorithm for training DBN (Hinton and Salakhutdinov (2006)), which allowed the development of structured inference and learning with DBN, as demonstrated by several works recently proposed in the field (Fasel and Berry (2010); Farabet et al. (2012); Ngo and Carneiro (2013, 2014)). The method proposed by (Farabet et al. (2012)) shows a method to parse a scene into several visual classes. Fasel et al. (Fasel and Berry (2010)) propose a DBN that takes as input an ultrasound image of the mouth and outputs a segmentation of the tongue, and (Ngo and Carneiro (2013, 2014)) propose the segmentation of the endocardium of the LV from cardiac MR cine study (please recall that the contributions of

this paper compared to our previos works (Ngo and Carneiro (2013, 2014)) are mentioned in Sec. 1.1).

3. Methodology

3.1. Notation

A cardiac cine MR series consists of a sequence of K volumes $\{V_i\}_{i=1}^{K}$, each representing a particular cardiac phase. In turn, each volume comprises a set of L images $\{I_i\}_{i=1}^{L}$ (also known as volume slices), where each image is represented by $I : \Omega \rightarrow \mathbb{R}$, with $\Omega \subseteq \mathbb{R}^2$ denoting the image coordinate space. We assume to have annotation only at the ED and ES cardiac phases (i.e., only two out of the K phases available) for all L images in these two volumes. In each of these annotated images, the explicit endocardial and epicardial contour representations are denoted by $\mathbf{c}_{\text{ENDO}} : [0,1] \rightarrow \Omega$ and $\mathbf{c}_{\text{EPI}} : [0,1] \rightarrow \Omega$, respectively. The implicit contour representation is formed with the zero level set of an Euclidean signed distance function $\phi : \Omega \rightarrow \mathbb{R}$, represented by $\mathcal{C} = \{\mathbf{x} \in \Omega | \phi(\mathbf{x}) = 0\}$, where points inside the contour have $\phi(\mathbf{x}) < 0$ and outside, $\phi(\mathbf{x}) > 0$. Assume that a set of annotated sequences is represented by $\mathcal{D} = \{(I, \mathbf{c}_{\text{ENDO}}, \mathbf{c}_{\text{EPI}}, i, q)_s\}_{i \in \{1,...,N_s\}, s \in \{1,...,S\}, q \in \{\text{ED,ES}\}}$, where $i \in \{1, ..., N_s\}$ is an index to an image within the sequence s, $q \in \{\text{ED,ES}\}$ is the annotation of the cardiac phase, $s \in \{1, ..., S\}$ is an index to a sequence and S is the number of sequences in $\mathcal{D}$. A segmentation map is represented by $\mathbf{y}_{\text{ENDO}} : \Omega \rightarrow \{0,1\}$ (or $\mathbf{y}_{\text{EPI}} : \Omega \rightarrow \{0,1\}$), where 1 represents foreground (i.e., the region inside the contour $\mathbf{c}_{\text{ENDO}}$ or $\mathbf{c}_{\text{EPI}}$) and 0 denotes background (region outside the contour). For the explanation of our methodology below, please assume that we run our segmentation slice by slice in each of the ED and ES volumes, using a sequence of steps displayed in Fig. 2, where the ED and ES volumes are manually selected by the user.

3.2. Endocardium Segmentation

The endocardium segmentation is divided into two steps, with the first step comprising the ROI detection using structured inference on a DBN, which pro-

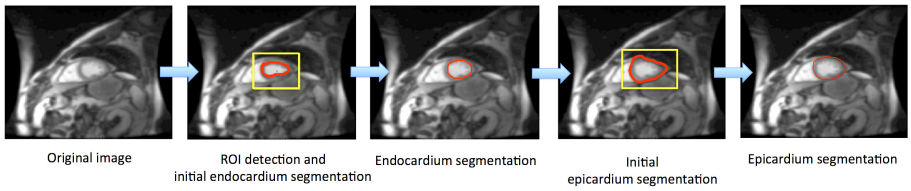


Figure 2: All steps involved in our methodology - Fig. 3 depicts each step in more detail.

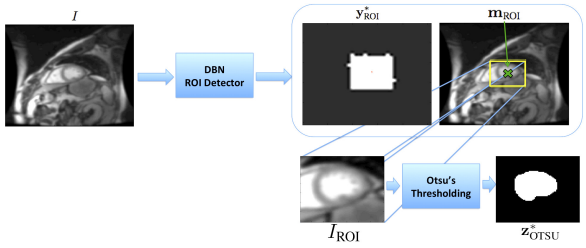
duces a rectangular region. Using this region as input, an initial endocardium segmentation is produced using Otsu's thresholding (Otsu (1975)) (Fig. 3-(a)). Note that Otsu's thresholding (Otsu (1975)) is a method that binarizes a gray-level image using a threshold value that is estimated in order to minimise the intra-class variance of the grey values, where the classes are defined by the pixel values above and below this threshold. The second step uses this initial segmentation to initialise an optimisation using the distance regularised level set method (DRLS) (Li et al. (2010)), which is based on an energy functional using length, area, shape prior and DBN-based appearance terms (Fig. 3-(b)). We give details about both steps below.

3.2.1. ROI DBN Detection and Initial Endocardium Segmentation

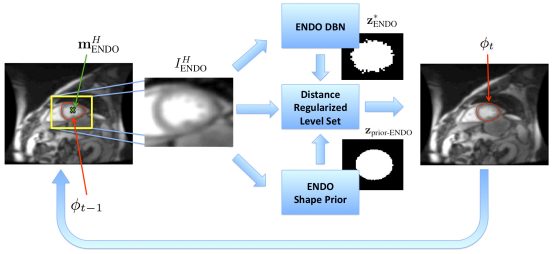
The ROI detection is based on a structured output inference using a DBN, which is a generative model composed of several layers of unsupervised networks, known as restricted Boltzmann machines (RBM). These RBMs have connections between layers but not between units within each layer, which facilitates the training procedure (Hinton and Salakhutdinov (2006)). The visible layers in this DBN are composed of the input image and the segmentation map (see Fig. 4). The ROI DBN detection is based on the maximisation of the following joint probability function representing a DBN model:

$$\mathbf{y}^{*}_{\mathrm{ROI}} = \arg\max_{\mathbf{y}} \int_{\mathbf{h}_1} ... \int_{\mathbf{h}_K} P(\mathbf{v}, \mathbf{h}_1, ..., \mathbf{h}_K, \mathbf{y}; \Theta_{\mathrm{ROI}}) d\mathbf{h}_1 ... d\mathbf{h}_K, \qquad (1)$$

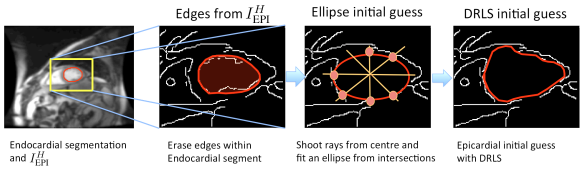
where $\mathbf{h}_k \in \{0,1\}^{|\mathbf{h}_k|}$ represents the $|\mathbf{h}_k|$ hidden nodes of layer $k \in \{1,..,K\}$ of the DBN, $\mathbf{v}$ is a vector representation of the input image I, $\mathbf{y} : \Omega \rightarrow \{0,1\}$, and



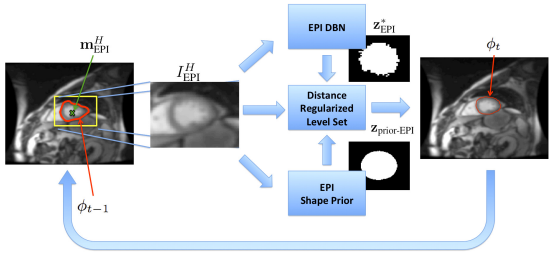
(a) ROI Detection and Initial Endocardium Segmentation



(b) Endocardium Segmentation



(c) Initial Epicardium Segmentation



(d) Epicardium Segmentation

Figure 3: Models of the ROI detection and initial endocardium segmentation (a), final endocardium segmentation (b), initial epicardium segmentation (c) and final epicardium segmentation (d).

Θ_{ROI} denotes the DBN parameters (weights and biases). The probability term

in (1) is computed as

$$P(\mathbf{v}, \mathbf{h}_1, ..., \mathbf{h}_K, \mathbf{y}) = P(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y}) \left(\prod_{k=1}^{K-2} P(\mathbf{h}_{k+1}|\mathbf{h}_k) \right) P(\mathbf{h}_1|\mathbf{v}), \quad (2)$$

where $P(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y}) \propto \exp\{-\mathcal{E}_{\text{RBM}}(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y})\}$ with

$$\begin{aligned} \mathcal{E}_{\text{RBM}}(\mathbf{h}_K, \mathbf{h}_{K-1}, \mathbf{y}) = &- \mathbf{b}_K^\top \mathbf{h}_K - \mathbf{a}_{K-1}^\top \mathbf{h}_{K-1} - \mathbf{a}_y^\top \mathbf{y} - \\ &(\mathbf{h}_K)^\top \mathbf{W}_K \mathbf{h}_{K-1} - (\mathbf{h}_K)^\top \mathbf{W}_\mathbf{y} \mathbf{y}, \end{aligned} \quad (3)$$

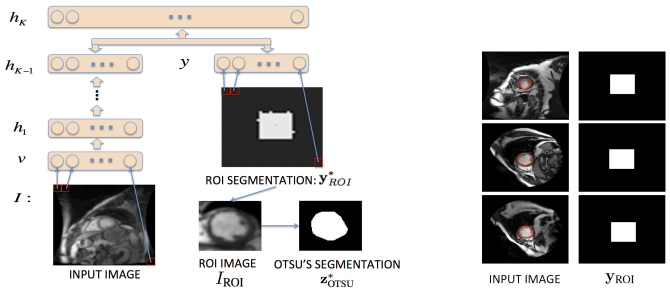
representing the energy function of an RBM (Hinton and Salakhutdinov (2006)), where $\mathbf{b}_K, \mathbf{a}_{K-1}, \mathbf{a}_y$ denote the bias vectors and $\mathbf{W}_K, \mathbf{W}_\mathbf{y}$ are the weight matrices. In (2), we also have

$$P(\mathbf{h}_{k+1}|\mathbf{h}_k) = \prod_j P(\mathbf{h}_{k+1}(j) = 1|\mathbf{h}_k), \quad (4)$$

with $P(\mathbf{h}_{k+1}(j) = 1|\mathbf{h}_k) = \sigma(\mathbf{b}_{k+1}(j) + \mathbf{h}_k^\top \mathbf{W}_{k+1}(:, j))$, $P(\mathbf{h}_1(j) = 1|\mathbf{v}) = \sigma(\mathbf{b}_1(j) + \mathbf{v}^\top \mathbf{W}_1(:, j))$ [3], where $\sigma(x) = \frac{1}{1+e^{-x}}$, the operator (j) returns the j^{th} vector value, and $(:, j)$ returns the j^{th} matrix column.

The estimation of the DBN parameter in (1) uses a training set comprising images I and their respective ROI segmentation maps $\mathbf{y}_{\text{ROI}}$. This annotation is automatically built from the manual endocardial border delineations $\mathbf{c}_{\text{ENDO}}$ (from $\mathcal{D}$, defined in Sec. 3.1), by producing a segmentation map with 0's everywhere except at a square of 1's with size M_{ROI}, centred at the centre of gravity of the annotation $\mathbf{c}_{\text{ENDO}}$ (see training samples in Fig. 4-(b)). The training process is based on the initial unsupervised bottom-up training of each pair of layers, where the DBN parameters are estimated in order to build an auto-encoder, and the top RBM is trained with an additional input containing the segmentation map $\mathbf{y}_{\text{ROI}}$ (Hinton and Salakhutdinov (2006)). The main algorithm used in this training process is the contrastive divergence, which is an approximation to gradient descent (Hinton and Salakhutdinov (2006)). Note that the DBN is a generative model, so the inference process to produce a segmentation map

[3] That is, we assume Gaussian visible units for the DBN with mean zero and standard deviation one.



(a) ROI DBN & Otsu's segmentation (b) Training samples

Figure 4: ROI DBN Model and Otsu's segmentation (a) and training samples for the ROI DBN (b).

given an input image is based on the generation of a segmentation map when the input $\mathbf{v}$ is clamped at this input image values. More specifically, using the input image at the bottom layer, bottom-up inferences are realised with mean-field approximation until reaching the top two layers, which form an RBM. The segmentation map layer is then initialised at $\mathbf{y} = \mathbf{0}$ and we then run Gibbs sampling on the layers $\mathbf{y}$ and $\mathbf{h}_K$ until convergence (Hinton and Salakhutdinov (2006)), with $\mathbf{h}_{K-1}$ clamped from the mean-field approximation. The stable vector for the layer $\mathbf{y}$ is labelled $\mathbf{y}^*_{\text{ROI}}$.

After estimating the ROI segmentation map $\mathbf{y}^*_{\text{ROI}}$, a rough endocardial border delineation is estimated by first applying the following function:

$$(I_{\text{ROI}}, \mathbf{m}_{\text{ROI}}, \mathbf{z}_{\text{ROI}}) = f_R(\mathbf{y}^*_{\text{ROI}}, I, M_{\text{ROI}}), \tag{5}$$

where $\mathbf{m}_{\text{ROI}}$ is the centre of gravity of $\mathbf{y}^*_{\text{ROI}}$ computed as $\mathbf{m}_{\text{ROI}} = \int_\Omega \mathbf{x} h(\mathbf{y}^*_{\text{ROI}}) d\mathbf{x}$, with $h(\mathbf{y}^*_{\text{ROI}}) = \frac{H(\mathbf{y}^*_{\text{ROI}})}{\int_\Omega H(\mathbf{y}^*_{\text{ROI}}) d\mathbf{x}}$ and $H(.)$ denoting the Heaviside step function, the binary map $\mathbf{z}_{\text{ROI}} : \Omega \rightarrow \{0, 1\}$ from $\mathbf{y}^*_{\text{ROI}}$ is computed with

$$\mathbf{z}_{\text{ROI}}(\mathbf{x}) = \begin{cases} 1, & \mathbf{y}^*_{\text{ROI}}(\mathbf{x}) > 0.5 \\ 0, & \text{otherwise} \end{cases}, \tag{6}$$

and I_{ROI} is a sub-image of size $M_{\text{ROI}} \times M_{\text{ROI}}$ extracted with $I_{\text{ROI}} = I(\mathbf{m}_{\text{ROI}} \pm M_{\text{ROI}}/2)$. Then, Otsu's thresholding (Otsu (1975)) is run on sub-image I_{ROI}, where the convex hull of the connected component linked to the centre $M_{\text{ROI}}/2$

is returned as the rough endocardial border delineation with $\mathbf{z}^*_{\text{OTSU}} = f_O(I_{\text{ROI}})$, as displayed in Fig. 4-(a). This segmentation is used to form the initial signed distance function, as follows:

$$\phi_0 = f_\phi(\mathbf{z}^*_{\text{OTSU}}, \mathbf{m}_{\text{ROI}}, M_{\text{ROI}}, I), \tag{7}$$

where we first create a temporary binary map $\mathbf{z} : \Omega \rightarrow \{0, 1\}$ with a map of the size of I containing only zeros, as in $\mathbf{z} = \mathbf{0}_{size(I)}$ (the function $size(i)$ returns the size of the image), then we fill this map with the result from $\mathbf{z}^*_{\text{OTSU}}$ centred at $\mathbf{m}_{\text{ROI}}$, with $\mathbf{z}(\mathbf{m}_{\text{ROI}} \pm M_{\text{ROI}}/2) = \mathbf{z}^*_{\text{OTSU}}(M_{\text{ROI}}/2 \pm M_{\text{ROI}}/2)$. Finally, the signed distance function $\phi_0 : \Omega \rightarrow \mathbb{R}$ in (7) is computed with

$$\phi_0(\mathbf{x}) = \begin{cases} -d(\mathbf{x}, \Omega^{out}), & \text{if } \mathbf{x} \in \Omega^{in} \\ +d(\mathbf{x}, \Omega^{in}), & \text{if } \mathbf{x} \in \Omega^{out} \end{cases}, \tag{8}$$

where $\Omega^{in} = \{\mathbf{x} \in \Omega | \mathbf{z}(\mathbf{x}) = 1\}$, $\Omega^{out} = \{\mathbf{x} \in \Omega | \mathbf{z}(\mathbf{x}) = 0\}$, and $d(\mathbf{x}, \Omega) = \inf_{\mathbf{y} \in \Omega} \|\mathbf{x} - \mathbf{y}\|_2$.

3.2.2. Endocardium Segmentation Combining DRLS and DBN

Given the initial segmentation ϕ_0 defined in (7), we run an optimisation algorithm to estimate the final endocardial border using the distance regularised level set (DRLS) formulation (Li et al. (2010)), where the energy functional is represented by

$$\mathcal{E}(\phi) = \mu \mathcal{R}_p(\phi) + \mathcal{E}_{\text{ext}}(\phi, \phi_{\text{ENDO-DBN},q}, \phi_{\text{ENDO-PRIOR},q}), \tag{9}$$

where $\mathcal{R}_p(\phi) = \int_\Omega p(|\nabla\phi|) d\mathbf{x}$ (with $p(s) = 0.5(s-1)^2$) is a regularisation term that guarantees $|\nabla\phi| \approx 1$; and $\mathcal{E}_{\text{ext}}(\phi)$ is defined as (Ngo and Carneiro (2013)):

$$\begin{aligned} &\mathcal{E}_{\text{ext}}(\phi, \phi_{\text{ENDO-DBN},q}, \phi_{\text{ENDO-PRIOR},q}) = \\ &\qquad \lambda \mathcal{L}(\phi) + \alpha \mathcal{A}(\phi) + \beta \mathcal{S}(\phi, \phi_{\text{ENDO-DBN},q}) + \gamma \mathcal{S}(\phi, \phi_{\text{ENDO-PRIOR},q}), \end{aligned} \tag{10}$$

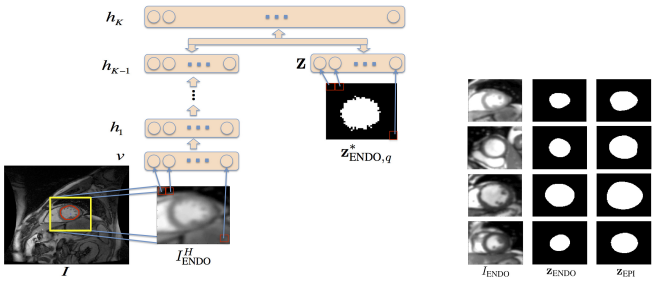
where the length term $\mathcal{L}(\phi) = \int_\Omega g\delta(\phi)|\nabla\phi| d\mathbf{x}$ (with $\delta(.)$ denoting the Dirac delta function and $g = \frac{1}{1+|\nabla G_\sigma * I|}$ representing the edge indicator function), the area $\mathcal{A}(\phi) = \int_\Omega gH(-\phi) d\mathbf{x}$, and $\mathcal{S}(\phi, \phi_\kappa) = \int_\Omega (\phi(\mathbf{x}) - \phi_\kappa(\mathbf{x} + \mathbf{m}_\phi))^2 d\mathbf{x}$

(with $\kappa \in \{(\text{ENDO-DBN}, q), (\text{ENDO-PRIOR}, q)\}$, and $q \in \{\text{ED,ES}\}$) represents the shape term that drives ϕ either towards the shape $\phi_{\text{ENDO-DBN},q}$ inferred from the ENDO DBN (described below in Sec. 3.2.3) or towards the shape prior $\phi_{\text{ENDO-PRIOR},q}$ estimated from the training set (see Sec. 3.4 below). Notice that the shape term $\mathcal{S}(\phi, \phi_\kappa)$ matches the two signed distance functions using the translation invariance by intrinsic alignment (Cremers et al. (2006)), where $\mathbf{m}_\phi = \int_\Omega \mathbf{x} h(\phi(\mathbf{x}))d\mathbf{x}$ with $h(\phi) = \frac{H(-\phi)}{\int_\Omega H(-\phi)d\mathbf{x}}$ is the centre of gravity of the segmentation from ϕ, and assuming that the shape prior represented by ϕ_κ has its centre of gravity at the origin. Note that this translation aligns the centre of gravity of ϕ_κ and ϕ. It is important to mention that when $\kappa \in \{\text{ENDO-PRIOR}, q, \text{EPI-PRIOR}, q\}$, then $\phi_\kappa(\mathbf{x} + \mathbf{m}_\phi)$ is essentially the same signed distance function translated according to $\mathbf{m}_\phi$, but when $\kappa \in \{\text{ENDO-DBN}, q, \text{EPI-DBN}, q\}$, the shape of the signed distance function changes as a function of $\mathbf{m}_\phi$. This happens because the result from the DBN segmentation changes as a function of where it is applied in the input image. The gradient flow of the energy $\mathcal{E}(\phi)$ is then defined as follows:

$$\begin{aligned}\frac{\partial \phi}{\partial t} =& \mu \text{div}(d_p(|\nabla\phi|)\nabla\phi) + \lambda\delta(\phi)\text{div}(g\frac{\nabla\phi}{|\nabla\phi|}) + \alpha g\delta(\phi) + \\ & 2\beta(\phi(\mathbf{x}) - \phi_{\text{ENDO-DBN},q}(\mathbf{x} + \mathbf{m}_\phi)) + \\ & 2\gamma(\phi(\mathbf{x}) - \phi_{\text{ENDO-PRIOR},q}(\mathbf{x} + \mathbf{m}_\phi)),\end{aligned} \tag{11}$$

where div(.) denotes the divergence operator, $\phi(\mathbf{x})$ denotes the current level set function, $\phi_{\text{ENDO-DBN},q}(\mathbf{x}+\mathbf{m}_\phi)$ denotes the translated signed distance function produced by the ENDO-DBN (similarly for ENDO-PRIOR), and $d_p(.)$ denotes the derivative of the function $p(.)$ defined in (9).

The estimated final endocardium segmentation is obtained from the minimisation of the energy functional in (9). In practice, the segmentation is obtained from the steady solution of the gradient flow equation (Li et al. (2010)) $\frac{\partial \phi}{\partial t} = -\frac{\partial \mathcal{E}}{\partial \phi}$, where $\partial\mathcal{E}/\partial\phi$ is the Gâteaux derivative of the functional $\mathcal{E}(\phi)$ and $\frac{\partial \phi}{\partial t}$ is defined in (11). The main idea of the DRLS (Li et al. (2010)) is then to iteratively follow the steepest descent direction (11) until convergence, resulting in the final steady solution $\phi^*_{\text{ENDO},q}$.



(a) ENDO DBN (b) Training samples

Figure 5: Graphical model for the ENDO DBN (a) and respective training samples (b).

3.2.3. ENDO DBN

The ENDO DBN used at this stage is similar to the ROI DBN from Sec. 3.2.1, but with the following differences: 1) instead of using the whole image I as the input, we use the sub-image I^H_{ENDO} of size M_{ENDO} (centred at $\mathbf{m}^H_{\text{ENDO}}$) extracted with $(I^H_{\text{ENDO}}, \mathbf{m}^H_{\text{ENDO}}, \mathbf{z}^H_{\text{ENDO}}) = f_R(H(-\phi_{t-1}), I, M_{\text{ENDO}})$, where $f_R(.)$ is defined in (5), $H(-\phi_{t-1})$ is a binary image containing the estimation for the endocardium map from DRLS (at iteration $t-1$), I denotes the original image, and $\mathbf{z}^H_{\text{ENDO}}$ represents the binary segmentation map (of size M_{ENDO}) from $H(-\phi_{t-1})$ on sub-image I^H_{ENDO}. We estimate the parameters of two distinct DBNs, one to segment images for q = ES phase and another for $q =$ ED phase of the cardiac cycle, where the training set is formed by samples $\{(I_{\text{ENDO}}, \mathbf{z}_{\text{ENDO}}, i, q)_s\}_{i\in\{1,...,N_s\},s\in\{1,...,S\},q\in\{\text{ED,ES}\}}$ extracted from the original training set with $(I_{\text{ENDO}}, \mathbf{m}_{\text{ENDO}}, \mathbf{z}_{\text{ENDO}}) = f_R(\mathbf{y}_{\text{ENDO}}, I, M_{\text{ENDO}})$, where $f_R(.)$ is defined in (5), and $\mathbf{y}_{\text{ENDO}}$ is the binary map formed from the original endocardium annotation $\mathbf{c}_{\text{ENDO}}$ (see Sec. 3.1). The segmentation from ENDO DBN is obtained with (see Fig. 5):

$$\mathbf{z}^*_{\text{ENDO},q} = \arg\max_{\mathbf{z}} \int ... \int P(\mathbf{v}, \mathbf{h}_1, ..., \mathbf{h}_K, \mathbf{z}; \Theta_{\text{ENDO,q}}) d\mathbf{h}_1 ... d\mathbf{h}_K, \qquad (12)$$

which is defined in (1), with $\mathbf{v}$ receiving the vectorised sub-image I^H_{ENDO}. The segmentation $\mathbf{z}^*_{\text{ENDO},q}$ can then be used to define the signed distance function $\phi_{\text{ENDO-DBN},q}$ in (9) with $\phi_{\text{ENDO-DBN},q} = f_\phi(\mathbf{z}^*_{\text{ENDO},q}, \mathbf{m}^H_{\text{ENDO}}, M_{\text{ENDO}}, I)$, with

$f_\phi(.)$ defined in (7). The training and inference processes for these ENDO DBNs for $q \in \{\text{ES}, \text{ED}\}$ are the same as described for the ROI DBN in Sec. 3.2.1.

3.3. Epicardium Segmentation

The epicardium segmentation also follows two steps, comprising an initial epicardium segmentation, which produces a square region containing the epicardium and an initial estimation of its border, similarly to the approach in Sec. 3.2.1 (Fig. 3-(c)). The second step involves an optimisation with DRLS (Li et al. (2010)), similar to the one presented above in Sec. 3.2.2 (Fig. 3-(d)).

3.3.1. Initial Epicardium Segmentation

The epicardium segmentation process is initialised with a rough delineation based on the endocardium detection (see Figure 3-(c)). Specifically, after the endocardium segmentation is finalized, we estimate the borders of the epicardium segmentation by first running the Canny edge detector (Canny (1986)) that outputs the edges within the window I^H_{EPI}, produced with $(I^H_{\text{EPI}}, \mathbf{m}^H_{\text{EPI}}, \mathbf{z}^H_{\text{EPI}}) = f_R(H(-\phi^*_{\text{ENDO},q}), I, M_{\text{EPI}})$, where $\phi^*_{\text{ENDO},q}$ represents the result from the DRLS, described in Sec. 3.2.2, and $f_R(.)$ is defined in (5). The edges lying in the region where $H(-\phi^*_{\text{ENDO},q})$ equals to one (this region represents blood pool found by the endocardium segmentation) are then erased and then, by "shooting" 20 rays (18 degrees apart from each other) from the centre $\mathbf{m}^H_{\text{EPI},q}$ and recording the intersection position between each ray and the first edge it crosses, we form a set of points that are likely to belong to the endocardial border. At this stage, since it is expected that the endocardial border will be relatively close to the epicardial border, we only record the points that are within a limited range from the original endocardial border (specifically, we expect the epicardial border to be within 1.05 and 1.1 of the length of the ray from $\mathbf{m}^H_{\text{EPI}}$ to the endocardial border; otherwise no point is recorded - these numbers are estimated from the 95% confidence interval of the distance between the endocardium and epicardium annotations from the training set). Finally, by fitting an ellipse to these points and running a small number of iterations of the original DRLS (Li et al. (2010))

(which is the model in (9)-(10) with $\beta = \gamma = 0$), we form the initial epicardium segmentation that is represented by a map $\mathbf{z}^*_{\text{EPI-initial}}$, which is then used to form the initial signed distance function $\phi_0 = f_\phi(\mathbf{z}^*_{\text{EPI-initial}}, \mathbf{m}^H_{\text{EPI}}, M_{\text{EPI}}, I)$, as defined in (7).

3.3.2. Epicardium Segmentation Combining DRLS and DBN

Using the initial epicardium segmentation ϕ_0 from Sec. 3.3.1 above, we run the optimisation function as defined in (9), but with the following external energy function: $\mathcal{E}_{\text{ext}}(\phi, \phi_{\text{EPI-DBN},q}, \phi_{\text{EPI-PRIOR},q})$, with $q \in \{\text{ED,ES}\}$, where $\phi_{\text{EPI-DBN},q}$ and $\phi_{\text{EPI-PRIOR},q}$ are defined below. The final steady solution of this optimisation is represented by $\phi^*_{\text{EPI},q}$.

3.3.3. EPI DBN

The EPI DBN runs similarly to the network defined above in Sec. 3.2.3, where the input sub-image I^H_{EPI} (centred at $\mathbf{m}^H_{\text{EPI}}$) of size M_{EPI} is extracted with $(I^H_{\text{EPI}}, \mathbf{m}^H_{\text{EPI}}, \mathbf{z}^H_{\text{EPI}}) = f_R(H(-\phi_{t-1}), I, M_{\text{EPI}})$, defined in (5). We can estimate the parameters of two DBNs for $q \in \{\text{ED,ES}\}$ with the following training set $\{(I_{\text{EPI}}, \mathbf{z}_{\text{EPI}}, i, q)_s\}_{i\in\{1,\ldots,N_s\}, s\in\{1,\ldots,S\}, q\in\{ED,ES\}}$ also extracted from the original training set with $(I_{\text{EPI}}, \mathbf{m}_{\text{EPI}}, \mathbf{z}_{\text{EPI}}) = f_R(\mathbf{y}_{\text{EPI}}, I, M_{\text{EPI}})$, with $\mathbf{y}_{\text{EPI}}$ representing the binary map computed from the epicardium annotation $\mathbf{c}_{\text{EPI}}$. The inference process is the same as the one defined in (12), resulting in $\mathbf{z}^*_{\text{EPI},q}$. The signed distance function is then defined by $\phi_{\text{EPI-DBN},q} = f_\phi(\mathbf{z}^*_{\text{EPI},q}, \mathbf{m}^H_{\text{EPI}}, M_{\text{EPI}}, I)$.

3.4. Shape Prior

The shape priors are computed with the mean of the manual annotations $\mathbf{z}_{\text{ENDO}}$ and $\mathbf{z}_{\text{EPI}}$, respectively, as follows: $\bar{\mathbf{z}}_{\text{ENDO-PRIOR}}(j) = \frac{1}{SN_s}\sum_{s=1}^{S}\sum_{i=1}^{N_s}\mathbf{z}_{\text{ENDO}}(j)$, where the index j represents as specific pixel address in the window $\mathbf{z}_{\text{ENDO}}$ of size $M_{\text{ENDO}} \times M_{\text{ENDO}}$. Assuming that each element of the mean map $\bar{\mathbf{z}}_{\text{ENDO}}$ is between 0 and 1, the shape prior is computed as

$$\mathbf{z}_{\text{ENDO-PRIOR}}(j) = \begin{cases} 1, & \text{if } \bar{\mathbf{z}}_{\text{ENDO-PRIOR}}(j) > 0.5 \\ 0, & \text{if } \bar{\mathbf{z}}_{\text{ENDO-PRIOR}}(j) \leq 0.5 \end{cases}. \tag{13}$$

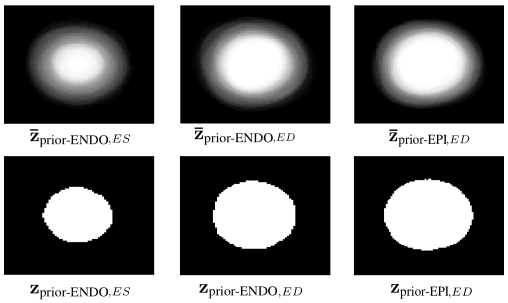


Figure 6: Shape priors for the endocardium and epicardium segmentation in ES and ED cardiac cycles. Note that for the epicardium case, the MICCAI 2009 challenge database (Radau et al. (2009)) does not contain training samples for the ES cycle.

Fig. 6 shows $\bar{\mathbf{z}}_{\text{ENDO-PRIOR}}$ and $\mathbf{z}_{\text{ENDO-PRIOR}}$ for the ED and ES cycles (and also the epicardium prior for the ED cycle). The signed distance function for the endocardium segmentation at cardiac cycle $q \in \{\text{ED,ES}\}$ is then defined by $\phi_{\text{ENDO-PRIOR},q} = f_\phi(\mathbf{z}_{\text{ENDO-PRIOR},q}, \mathbf{m}^H_{\text{ENDO}}, M_{\text{ENDO}}, I)$. This process works in the same way for the case of epicardial shape prior.

4. Experiments

4.1. Data Set and Evaluation Measures

The accuracy of the endocardium and epicardium segmentation results produced by our methodology is assessed using the database and the evaluation proposed in the MICCAI 2009 LV segmentation challenge (Radau et al. (2009)), obtained from the Sunnybrook Health Sciences Centre, Toronto, Canada. In total, 45 cardiac short axis (SAX) cine-MR data sets are available, which are divided into three sets (online, testing and training sets) of 15 sequences, where each sequence contains four ischemic heart failures, four non-ischemic heart failures, four LV hypertrophies and three normal cases. Each of those sequences has been acquired during a 10-15 second breath-hold, with a temporal resolution of 20 cardiac phases over the heart cycle, starting from the ED cardiac phase, and containing six to 12 SAX images obtained from the atrioventricular

ring to the apex (thickness=$8mm$, gap=$8mm$, FOV=$320mm \times 320mm$, matrix= 256×256). Expert annotations are provided for endocardial contours in all slices at ED and ES cardiac phases, and for epicardial contours only at ED cardiac phase. The evaluation proposed for assessing the algorithms submitted to the MICCAI 2009 LV segmentation challenge is based on the following three measures: 1) percentage of "good" contours, 2) the average Dice metric (ADM) of the "good" contours, and 3) average perpendicular distance (APD) of the "good" contours. A segmentation is classified as good if APD $< 5mm$.

During the MICCAI 2009 LV Segmentation Challenge (Radau et al. (2009)), the organisers first released the training and testing sets, where the training set contained the manual annotation, but the testing set did not include the manual annotation. The online dataset only became available on the challenge day, so that the participants could submit their segmentation results for assessment. The challenge organisers reported all segmentation results for all datasets that were available from the participants. Currently all three data sets with their respective expert annotations are publicly available. Given that most of the results from the challenge participants are available for the training and testing sets, we decided to use the training set to estimate all DBN parameters, the online set for validating some DBN parameters (e.g., number of layers and number of nodes per layer), and the testing set exclusively for testing (since this is the set which has the majority of results from the participants).

4.2. Experimental Setup

The training set is used for estimating the ROI DBN, ENDO DBN and EPI DBN parameters (network weights and biases), the shape priors (as described in Sec. 3.4) and for estimating the weights of the DRLS method (i.e., $\mu, \lambda, \alpha, \beta, \gamma$ in (9) and (10)); while the online set is used for the model selection of the DBNs (i.e., estimation of the number of DBN hidden layers and number of nodes per layer). Specifically, we use the online set for cross validating the number of hidden layers (we test from two to four hidden layers), and the number of nodes per hidden layer (we consider ranges from 100 to 2000 nodes per layer in intervals

of 100 nodes). Table 1 shows the network structures learned for each DBN used in this paper, where for the ROI DBN, the input image is resized from 256×256 to 40×40 using standard blurring and downsampling techniques, for the ENDO DBN, the input image is resized from $M_{\text{ENDO}} \times M_{\text{ENDO}}$ to 40×40, and for the EPI DBN, the input image is resized from $M_{\text{EPI}} \times M_{\text{EPI}}$ to 40×40. Note that all these DBN models are trained using an augmented training set, where for each annotated training image, we generate additional ones by translating the original image (and its annotation) within a range of ± 10 pixels. More specifically, we have 105 ED images and 75 ES annotated training images (from the 15 training volumes), and in addition to the original image, we generate 40 additional images with the translations mentioned above. Therefore, in total we have 105x41=4305 annotated images for training the ED endocardial DBN and epicardial DBN, and 75x41=3075 annotated images for training the ES endocardial DBN. The segmentation accuracy on training saturates with this augmented training data (i.e., adding more translated training images no longer improves the training results).

Table 1: ROI DBN, ENDO DBN and EPI DBN learned structures.

Network	input layer	$\mathbf{h}_1$	$\mathbf{h}_2$	output layer
ROI DBN	40×40	1300	1500	40×40
ENDO DBN (ED cycle)	40×40	1000	1000	40×40
ENDO DBN (ES cycle)	40×40	700	1000	40×40
EPI DBN (ED cycle)	40×40	1000	1000	40×40

The level set weights in (9) learned with the training set for the endocardium segmentation are $\Delta t = 2$ (time step in the level set formulation), $\mu = \frac{0.24}{\Delta t} = 0.12, \lambda = 4, \alpha = -2, \beta = 0.02$, and $\gamma = 0.001$; and for the epicardium segmentation, we have $\Delta t = 2, \mu = \frac{0.24}{\Delta t} = 0.12, \lambda = 4, \alpha = -4, \beta = 0.015$, and $\gamma = 0.001$. Note that we follow the recommendation by (Li et al. (2010)) in defining the values for Δt, and μ (the recommendations are $\Delta t > 1$ and $\mu < \frac{0.25}{\Delta t}$). For the inference procedure,the number of level set (DRLS) iterations is $T = 10$,

the size of the sub-windows are set as $M_{\text{ROI}}, M_{\text{ENDO}}, M_{\text{EPI}} = 100$. We found that the segmentation results are stable if these constants are within the ranges: $T \in [5, 20]$, $M_{\text{ROI}}, M_{\text{ENDO}}, M_{\text{EPI}} \in [80, 120]$. Finally, given that the proposed method can only segment the ED and ES volumes, we assume that these volumes are manually selected by the user.

4.3. Results of Each Stage of the Proposed Methodology

The role of each stage of our algorithm for the endocardium segmentation is presented in Table 2. The "Initial endocardium segmentation" shows the result produced by the zero level set ϕ_0 in (7) (i.e., the result from the ROI detection, followed by the initial endocardium segmentation). The "ENDO DBN alone" displays the accuracy results of the endocardium segmentation produced by the ENDO DBN (Sec. 3.2.3) alone. The "Model without DBN/shape prior" represents the energy functional in (10) with $\beta = \gamma = 0$, which effectively represents our model without the influence of the shape prior and the ENDO DBN. Similarly the "Model without DBN" denotes the case where the functional in (10) has $\beta = 0$ (i.e., with no influence from ENDO DBN) and the "Model without shape prior" has $\gamma = 0$ (no influence from the shape prior). Finally, the "Proposed model" displays the result with all steps described in Sec. 3.2, and "Proposed model (semi)" represents our model using a manual initialisation instead of the automated initialisation described in Sec. 3.2.1. This manual initialisation consists of a circle, where the centre is the manual annotation centre of gravity and the radius is the minimum distance between the manual annotation and this centre. Table 3 shows the result of the "initial epicardium segmentation" explained in Sec. 3.3.1, and the result of the segmentation produced by the complete model described in Sec. 3.3.2 (labelled as "Proposed model"). We also show the result of the semi-automated epicardium segmentation with manual initialisation (defined in the same way as the manual initialisation above for the endocardium segmentation), labelled as "Proposed model (semi)". Another important question about our proposed model is the need of separate classifiers for the ED and ES phases of the cardiac cycle, so we trained and tested a version of

our fully automated method using a model that does not discriminate between these two phases, which is labeled in Tab. 2 as Proposed model (joint ED ES). Note that we do not show all steps in Table 3 because the results are similar to the initial epicardium segmentation. We also show that the combination of DBN and DRLS provides an accuracy improvement by running the independent two-sample t-test for the three measures considered in this paper (i.e., the good percentage, APD and ADM) for the endocardium segmentation, where the first experiment compares the measures from the proposed model (combining DBN and DRLS) and from a method consisting only of the level set without the DBN, and the second experiment compares the proposed model and the segmentation result produced by the DBN segmentation alone. In both experiments and for all measures, the null hypothesis that the measures are drawn from independent samples from normal distributions with equal means is rejected at the 5% significance level. Finally, another important question about the ROI DBN is how it compares with more standard boosted cascade detectors Viola and Jones (2004), so we implemented that approach for the endocardium detection using integral image and Haar wavelet features, where the output is a window of size $M_{\mathrm{ROI}} \times M_{\mathrm{ROI}}$, where the center is computed from the result of the boosted cascade detector. Table 4 shows the Dice metric between the $M_{\mathrm{ROI}} \times M_{\mathrm{ROI}}$ ground truth window (computed from the centre of gravity of the annotation $\pm M_{\mathrm{ROI}}$ in each direction) and the detected window at the centre of gravity of the detection $\pm M_{\mathrm{ROI}}$ in each direction.

4.4. Comparison with the State of the Art

Tables 5 and 6 shows a comparison between our methodology (labelled "Proposed model") and the state of the art for the endocardium segmentation problem, while Tables 7 and 8 displays a similar comparison for the epicardium problem for different subsets of the MICCAI 2009 challenge databases (Radau et al. (2009)). Most of the approaches on that table are based on active contour models (Constantinides et al. (2012); Huang et al. (2009, 2011); Jolly (2009); Lu et al. (2009); Marak et al. (2009)), machine learning models (O'Brien et al.

Table 2: Quantitative experiments on the MICCAI 2009 challenge database (Radau et al. (2009)) showing the influence of each step of the proposed methodology for the endocardium segmentation. Each cell is formatted as "mean (standard deviation) [min value - max value]".

Method	"Good" Percentage	Endocardium ADM	Endocardium APD
Testing set (15 sequences)			
Proposed model (semi)	100(0)[100 − 100]	0.91(0.03)[0.83 − 0.95]	1.79(0.36)[1.28 − 2.75]
Proposed model	95.91(5.28)[84.62 − 100]	0.88(0.03)[0.82 − 0.93]	2.34(0.46)[1.62 − 3.24]
Proposed model (joint ED ES)	90.29(12.73)[61.11 − 100]	0.88(0.03)[0.80 − 0.93]	2.42(0.36)[1.75 − 2.96]
Model without shape prior	95.71(6.96)[78.95 − 100]	0.88(0.03)[0.83 − 0.93]	2.34(0.45)[1.67 − 3.14]
Model without DBN	85.89(18.00)[36.84 − 100]	0.84(0.04)[0.77 − 0.92]	2.77(0.58)[1.73 − 3.74]
Model without DBN/shape prior	84.49(18.31)[36.84 − 100]	0.84(0.04)[0.78 − 0.92]	2.78(0.58)[1.72 − 3.81]
ENDO DBN alone	18.31(19.46)[0 − 100]	0.87(0.02)[0.84 − 0.89]	3.81(0.64)[2.97 − 4.88]
Initial endocardium segmentation	85.18(15.83)[47.37 − 100]	0.85(0.04)[0.79 − 0.92]	2.81(0.47)[2.07 − 3.58]
Training set (15 sequences)			
Proposed model (semi)	100(0)[100 − 100]	0.91(0.03)[0.85 − 0.95]	1.63(0.40)[1.29 − 2.70]
Proposed model	97.22(3.16)[91.67 − 100]	0.88(0.05)[0.76 − 0.95]	2.13(0.46)[1.27 − 2.73]
Proposed model (joint ED ES)	92.08(6.11)[82.35 − 100.00]	0.88(0.04)[0.81 − 0.93]	2.42(0.50)[1.60 − 3.24]
Model without shape prior	97.42(4.63)[83.33 − 100]	0.88(0.04)[0.76 − 0.95]	2.14(0.43)[1.28 − 2.63]
Model without DBN	89.42(11.83)[61.11 − 100]	0.85(0.06)[0.71 − 0.93]	2.61(0.66)[1.74 − 3.65]
Model without DBN/shape prior	88.11(13.84)[50.00 − 100]	0.84(0.06)[0.70 − 0.93]	2.57(0.62)[1.72 − 3.53]
ENDO DBN alone	48.09(38.42)[0 − 100]	0.86(0.05)[0.73 − 0.90]	3.23(0.44)[2.70 − 4.05]
Initial endocardium segmentation	89.61(11.57)[55.56 − 100]	0.85(0.06)[0.71 − 0.93]	2.71(0.57)[1.78 − 3.49]

(2009); Wijnhout et al. (2009)), or a combination of both models (Hu et al. (2012); Uzunbaş et al. (2012)). Furthermore, Tables 5-8 also show a semi-automated version of our method (labelled "Proposed model (semi)") using the same initial guess described above in Sec. 4.3. In Tab. 9, we show the ADM and APD measures of our proposed method using all images in each dataset (i.e., not only the images that "survived" the "Good" percentage test, described above in

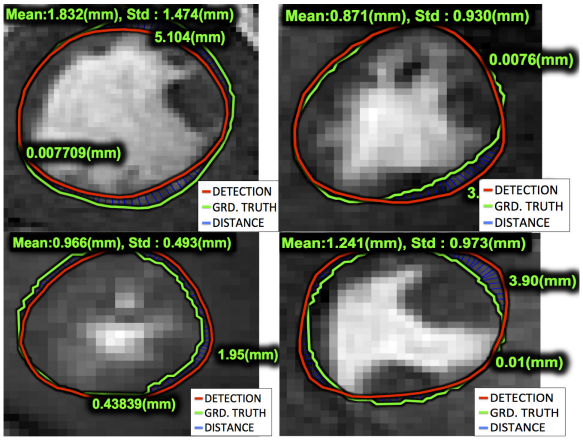
Table 3: Quantitative experiments on the MICCAI 2009 challenge database (Radau et al. (2009)) compared different versions of the proposed methodology for the epicardium segmentation. Each cell is formatted as "mean (standard deviation) [min value - max value]".

Method	"Good" Percentage	Epicardium ADM	Epicardium APD
Testing set (15 sequences)			
Proposed model (semi)	100(0)[100 − 100]	0.94(0.01)[0.92 − 0.97]	1.73(0.28)[1.16 − 2.17]
Proposed model	94.65(6.18)[85.71 − 100]	0.93(0.02)[0.88 − 0.96]	2.08(0.60)[1.27 − 3.74]
Initial epicardium segmentation	94.65(6.18)[85.71 − 100]	0.93(0.02)[0.88 − 0.96]	2.19(0.58)[1.32 − 3.68]
Training set (15 sequences)			
Proposed model (semi)	100.00(0.00)[100 − 100]	0.94(0.01)[0.91 − 0.96]	1.64(0.34)[1.17 − 2.47]
Proposed model	98.52(5.74)[77.78 − 100]	0.93(0.02)[0.89 − 0.96]	1.99(0.46)[1.35 − 3.13]
Initial epicardium segmentation	96.83(6.92)[77.78 − 100	0.93(0.02)[0.89 − 0.95]	1.99(0.40)[1.46 − 3.14]

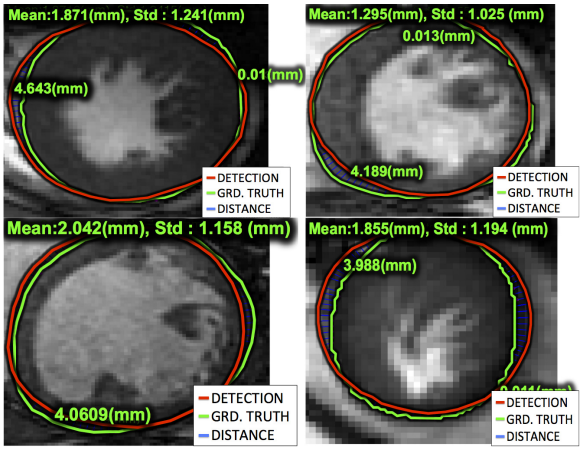
Table 4: Average Dice metric between the proposed DBN ROI detector and the boosted cascade detector Viola and Jones (2004). The cell formatting is the same as in Tab. 2.

Dataset	DBN ROI	Boosted Cascade
Online	0.88(0.06)[0.70−0.96]]	0.83(0.05)[0.76−0.91]
Testing	0.90(0.04)[0.84−0.94]]	0.86(0.04)[0.77−0.91]
Training	0.96(0.01)[0.95−0.97]	0.84(0.05)[0.74−0.92]

Sec. 4.1). Fig. 7 shows a few endocardium and epicardium segmentation results produced by our approach for challenging cases, such as with images from apical and basal slice images and presenting papillary muscles and trabeculations (please see supplementary material for more results). Finally, Fig. 8 shows a few unprocessed 3-D models of the endocardial and epicardial borders obtained with our proposed methodology.



a) Results of endocardium segmentation on the testing set



b) Results of epicardium segmentation on the testing set

Figure 7: Epicardium and endocardium segmentation results with challenging cases, such as images from apical and basal slice images and presenting papillary muscles and trabeculations. The red contour denotes the automated detection, and green shows the manual annotation. For more results, please see the supplementary material.

5. Discussion and Conclusions

The role of each stage of our methodology for the endocardium segmentation becomes clear with the results presented in Table 2. For instance, the DRLS method alone (Li et al. (2010)) (i.e., without the prior and ENDO DBN terms)

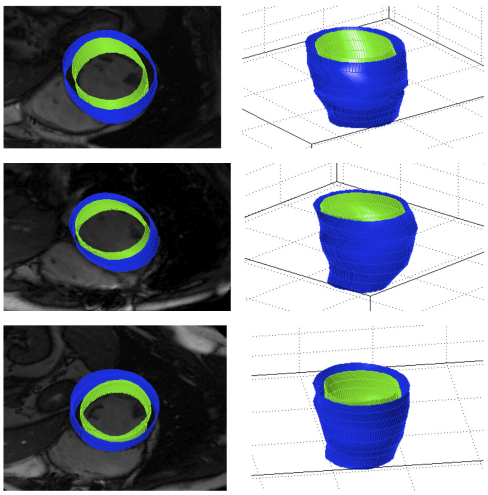


Figure 8: 3D Model formed by linking the slice by slice results for the endocardial (green surface) and epicardial (blue) borders in ED cycle (note that we focus on ED cardiac cycle because no annotated epicardium training set is available for the ES cycle from the MICCAI 2009 challenge database (Radau et al. (2009))).

is not able to improve significantly the result from the initial endocardium segmentation. The addition of the shape prior term improves slightly the accuracy (see row "Model without DBN"), but not significantly so; therefore we can removed it from the framework in order to obtain small gains in terms of efficiency. ENDO DBN (see row "Model without shape prior") is the term that provides the largest gain in terms of accuracy, even though its performance as a stand alone segmentation system is not competitive. This indicates that the results produced by ENDO DBN complements the results from DRLS using the information available (and automatically learned) from the training set. Putting all terms together, the "Proposed model" displays the best performance of our method, which is shown to be statistically significantly superior to both the DRLS and DBN methods. It is important to notice the relative small accuracy differences between the training and testing sets, which indicates good generalisation capabilities of our method (even with the relatively small training set of the MICCAI 2009 challenge database (Radau et al. (2009)). The use

of separate models for the ED and ES phases of the cardiac cycles appears to provide small accuracy gains, but the simplicity of having a single model may outweight these gains, depending on the system requirements. Alternatively, a single model could be implemented based on the methodology described by Carneiro and Nascimento (Carneiro and Nascimento (2013)) that automatically combines two models: one for the systolic interval and another for the diastolic interval, but the training process for this method would require a dataset that has annotations available for randomly selected frames from all stages of the cardiac cycle, in addition to the ED and ES frames available from the MICCAI 2009 LV segmentation challenge (Radau et al. (2009)). Moreover, our proposed DBN ROI detector shows in Tab. 4 better detection results compared to more standard methods in the field, such as the boosted cascade detector Viola and Jones (2004). Finally, by using a manual initialization, note that we obtain the best result in the field.

Table 3 shows that the initial epicardium segmentation already produces a result that is close to the result produced by our proposed model. Therefore, even though we notice that the use of the EPI DBN also improves the result, it is only a slight improvement that mostly happens on the training set. Furthermore, similarly to the endocardium segmentation, the use of manual initialisation also shows the best result in the field. Finally, given the similar appearance of the endocardium and epicardium images, it is important to justify the need for learning two separate DBN models, that is the ENDO and EPI DBNs, instead of a single one estimated with all training sets. The main reason for these two models lies in the empirical evidence that they produce more accurate segmentation results, as shown in Tab. 5-6, where the rows labelled by **Proposed model (semi)** show the results with the two separate DBNs, while the rows labelled by Ngo and Carneiro (2013) display results using a single classifier.

The comparison with the state of the art in terms of the endocardium segmentation in Tables 5-6 and the epicardium segmentation in Tables 7-8 shows that our approach produces the best results in the field for the semi-automated

segmentation problem. For the fully automated segmentation problem our results is on par with the result of the method proposed by (Jolly (2009)), which is considered to be the current state of the art by a recent review paper by (Petitjean and Dacher (2011)). In general, for the endocardium segmentation, our results are better in terms of "Good" percentage than other methods but comparable to the best ones with respect to ADM and APD. For the epicardium segmentation our results are comparable to the method proposed by (Jolly (2009)), but better than all others. Note that while some approaches are more accurate in terms of APD or ADM (Lu et al. (2009)), they also present low values for "Good" percentage, which means that these methods also produce a large number of segmentations with APD larger than $5mm$, but the few ones that survive the "Good" percentage test are reasonably accurate. Another important observation is the relatively worse performance of the fully automated compared to semi-automated segmentation (not only for our proposed method, but other methods proposed in the literature), indicating that there is still room for improving the accuracy of the initial endocardium and epicardium segmentations. It is also important to mention that our approach runs on (mean) average in 175 ± 35 seconds for the endocardium segmentation and 119 ± 20 seconds for the epicardium segmentation using a *non-optimised Matlab program*, which is slower or comparable to other aproaches that run between one minute (Constantinides et al. (2012); Jolly (2009); Uzunbaş et al. (2012); Wijnhout et al. (2009)) and three minutes (Hu et al. (2012); Lu et al. (2009)).

There are several points that can be explored in order to improve the results of the endocardium and epicardium segmentation. First, instead of running the segmentation algorithm slice by slice, we can run it over the whole volume and use a 3-D shape model to constrain the search process. Second, we can also use a motion model as another constraint for the segmentation process. Third, if new training sets become available in the field, we can train more complex DBN models that can potentially produce more accurate segmentation results. Finally, we can decrease the running time of our approach by parallelizing the segmentation processes since the segmentation of each slice is done

independently of all others (roughly this means that we can in principle make our approach 10 times faster).

Acknowledgements

This work was partially supported by the Australian Research Council's Discovery Projects funding scheme (project DP140102794). Tuan Anh Ngo acknowledges the support of the 322 Program - Vietnam International Education Development, Ministry of Education and Training (VIED-MOET).

References

BakIr, G., 2007. Predicting structured data. MIT press.

Canny, J., 1986. A computational approach to edge detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on , 679–698.

Carneiro, G., Nascimento, J.C., 2013. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. Pattern Analysis and Machine Intelligence, IEEE Transactions on 35, 2592–2607.

Carneiro, G., Nascimento, J.C., Freitas, A., 2012. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. Image Processing, IEEE Transactions on 21, 968–982.

Cobzas, D., Schmidt, M., 2009. Increased discrimination in level set methods with embedded conditional random fields, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE. pp. 328–335.

Collins, M., 2002. Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms, in: Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, Association for Computational Linguistics. pp. 1–8.

Constantinides, C., Roullot, E., Lefort, M., Frouin, F., 2012. Fully automated segmentation of the left ventricle applied to cine mr images: Description and results on a database of 45 subjects, in: Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, IEEE. pp. 3207–3210.

Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., 1995. Active shape models-their training and application. Computer vision and image understanding 61, 38–59.

Cortes, C., Vapnik, V., 1995. Support vector machine. Machine learning 20, 273–297.

Cremers, D., Osher, S.J., Soatto, S., 2006. Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. International Journal of Computer Vision 69, 335–351.

Farabet, C., Couprie, C., Najman, L., LeCun, Y., 2012. Scene parsing with multiscale feature learning, purity trees, and optimal covers. arXiv preprint arXiv:1202.2160 .

Fasel, I., Berry, J., 2010. Deep belief networks for real-time extraction of tongue contours from ultrasound during speech, in: Pattern Recognition (ICPR), 2010 20th International Conference on, IEEE. pp. 1493–1496.

Freund, Y., Schapire, R.E., 1995. A desicion-theoretic generalization of on-line learning and an application to boosting, in: Computational learning theory, Springer. pp. 23–37.

Georgescu, B., Zhou, X.S., Comaniciu, D., Gupta, A., 2005. Databased-guided segmentation of anatomical structures with complex appearance, in: CVPR.

Hinton, G., Salakhutdinov, R., 2006. Reducing the dimensionality of data with neural networks. Science 313, 504–507.

Hu, H., Liu, H., Gao, Z., Huang, L., 2012. Hybrid segmentation of left ventricle in cardiac mri using gaussian-mixture model and region restricted dynamic programming. Magnetic resonance imaging .

Huang, R., Pavlovic, V., Metaxas, D.N., 2004. A graphical model framework for coupling mrfs and deformable models, in: Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, IEEE. pp. II–739.

Huang, S., Liu, J., Lee, L., Venkatesh, S., Teo, L., Au, C., Nowinski, W., 2009. Segmentation of the left ventricle from cine mr images using a comprehensive approach. The MIDAS Journal 49.

Huang, S., Liu, J., Lee, L.C., Venkatesh, S.K., San Teo, L.L., Au, C., Nowinski, W.L., 2011. An image-based comprehensive approach for automatic segmentation of left ventricle from cardiac short axis cine mr images. Journal of digital imaging 24, 598–608.

Jolly, M., 2009. Fully automatic left ventricle segmentation in cardiac cine mr images using registration and minimum surfaces. The MIDAS Journal 49.

Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: Active contour models. International journal of computer vision 1, 321–331.

Li, C., Xu, C., Gui, C., Fox, M.D., 2010. Distance regularized level set evolution and its application to image segmentation. Image Processing, IEEE Transactions on 19, 3243–3254.

Lu, Y., Radau, P., Connelly, K., Dick, A., Wright, G., 2009. Automatic image-driven segmentation of left ventricle in cardiac cine mri. The MIDAS Journal 49.

Marak, L., Cousty, J., Najman, L., Talbot, H., et al., 2009. 4d morphological segmentation and the miccai lv-segmentation grand challenge, in: MICCAI 2009 Workshop on Cardiac MR Left Ventricle Segmentation Challenge, pp. 1–8.

Ngo, T.A., Carneiro, G., 2013. Left ventricle segmentation from cardiac mri combining level set methods with deep belief networks, in: Image Processing (ICIP), 2013 20th IEEE International Conference on, IEEE. pp. 695–699.

Ngo, T.A., Carneiro, G., 2014. Fully automated non-rigid segmentation with distance regularized level set evolution initialized and cosntrained by deep-structured inference, in: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, IEEE.

Osher, S., Sethian, J.A., 1988. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. Journal of computational physics 79, 12–49.

Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica 11, 23–27.

O'Brien, S., Ghita, O., Whelan, P., 2009. Segmenting the left ventricle in 3d using a coupled asm and a learned non-rigid spatial model. The MIDAS Journal 49.

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

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

Schaerer, J., Casta, C., Pousin, J., Clarysse, P., 2010. A dynamic elastic model for segmentation and tracking of the heart in mr image sequences. Medical Image Analysis 14, 738–749.

Tsechpenakis, G., Metaxas, D.N., 2007. Crf-driven implicit deformable model, in: Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, IEEE. pp. 1–8.

Tsochantaridis, I., Joachims, T., Hofmann, T., Altun, Y., Singer, Y., 2005. Large margin methods for structured and interdependent output variables. Journal of Machine Learning Research 6.

Uzunbaş, M.G., Zhang, S., Pohl, K.M., Metaxas, D., Axel, L., 2012. Segmentation of myocardium using deformable regions and graph cuts, in: Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, IEEE. pp. 254–257.

Viola, P., Jones, M.J., 2004. Robust real-time face detection. International journal of computer vision 57, 137–154.

Wijnhout, J., Hendriksen, D., Assen, H., der Geest, R., 2009. Lv challenge lkeb contribution: Fully automated myocardial contour detection. The MIDAS Journal 43.

Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D., 2008. Four-chamber heart modeling and automatic segmentation for 3-d cardiac ct volumes using marginal space learning and steerable features. Medical Imaging, IEEE Transactions on 27, 1668–1681.

Biography

Tuan Anh Ngo received the B.S. degree in mathematics from Hanoi National University of Education, Hanoi, Vietnam, the M.S. degree in computer science from Asian Institute of Technology, Bangkok, Thailand, and the Ph.D. degree from the School of Computer Science of the University of Adelaide. Currently, he is a lecturer at Department of Computer Science, Faculty of Information Technology, Vietnam National University of Agriculture. His research interests include medical image analysis and optimisation methods based on deep learning and level sets.

Zhi Lu received the Ph.D degree in computer science from City University of Hong Kong, Hong Kong, China, in 2014. Previously, Dr. Zhi Lu received

his M.Sc. and M.Phil. degrees in computer science from the same university in 2008 and 2011, respectively. Currently, he is a postdoctoral researcher at the School of Computer Science of the University of Adelaide in Australia. His research interests include medical image analysis and pattern recognition.

Gustavo Carneiro received the BS and MSc degrees in computer science from the Federal University of Rio de Janeiro and the Military Institute of Engineering, Brazil, in 1996 and 1999, respectively, and the PhD degree in computer science from the University of Toronto, Canada, in 2004. Currently, he is an associate professor at the School of Computer Science of the University of Adelaide, Australia. Previously, he worked at the Instituto Superior Tecnico of Lisbon, from 2008 to 2011 as a visiting researcher and assistant professor, and from 2006 to 2008, he worked at Siemens Corporate Research in Princeton, New Jersey. His research interests include medical image analysis, image feature selection and extraction, content-based image retrieval and annotation, and general visual object classification. He has authored more than 60 peer reviewed publications in international journals and conferences.

Table 5: Quantitative experiments on the **training and testing sets** of the MICCAI 2009 challenge databases (Radau et al. (2009)) comparing the performance of our proposed approach with the state of the art on the **endocardium segmentation problem**. Notice that the methods are classified into fully or semi-automated. The cell formatting is the same as in Tab. 2, but note that '?' means that the result is not available in the literature. The top performance for each measure and dataset is highlighted.

Method	"Good" Percentage	Endocardium ADM	Endocardium APD
Testing set (15 sequences)			
Semi Automated			
Proposed model (semi)	**100(0)[100 − 100]**	**0.91(0.03)[0.83 − 0.95]**	**1.79(0.36)[1.28 − 2.75]**
Ngo and Carneiro (2013)	96.58(9.58)[63.15 − 100]	0.89(0.03)[0.83 − 0.93]	2.22(0.46)[1.69 − 3.30]
Huang et al. (2009)	?	0.89(0.04)[?−?]	2.10(0.44)[?−?]
Uzunbaş et al. (2012)	?	0.82(0.06)[?−?]	2.98(0.88)[?−?]
Fully Automated			
Proposed model	**95.91(5.28)[84.62 − 100]**	0.88(0.03)[0.82 − 0.93]	2.34(0.46)[1.62 − 3.24]
Jolly (2009)	94.33(9.93)[62.00 − 100]	0.88(0.03)[0.84 − 0.94]	2.44(0.62)[1.36 − 3.68]
Wijnhout et al. (2009)	86.47(11.00)[68.4 − 100]	**0.89(0.03)[0.82 − 0.94]**	2.29(0.57)[1.67 − 3.93]
Lu et al. (2009)	72.45(19.52)[42.11 − 100]	**0.89(0.03)[0.84 − 0.94]**	**2.07(0.61)[1.32 − 3.77]**
Marak et al. (2009)	?	0.86(0.04)[?−?]	?
O'Brien et al. (2009)	?	0.81(?)[?−?]	?

Training set (15 sequences)			
Semi Automated			
Proposed model (semi)	**100(0)[100 − 100]**	**0.91(0.03)[0.85 − 0.95]**	**1.63(0.40)[1.29 − 2.70]**
Ngo and Carneiro (2013)	98.45(3.11)[91.66 − 100]	0.90(0.03)[0.84 − 0.94]	1.96(0.35)[1.43 − 2.55]
Huang et al. (2009)	?	**0.90(0.04)**[?−?]	2.03(0.34)[?−?]
Fully Automated			
Proposed model	**97.22(3.16)[91.67 − 100]**	**0.88(0.05)[0.76 − 0.95]**	2.13(0.46)[1.27 − 2.73]
Jolly (2009)	96.93(7.59)[72 − 100]	**0.88(0.06)[0.75 − 0.95]**	**2.09(0.53)[1.35 − 3.23]**

Table 6: Quantitative experiments on the **online and full sets** of the MICCAI 2009 challenge databases (Radau et al. (2009)) comparing the performance of our proposed approach with the state of the art on the **endocardium segmentation problem**. Notice that the methods are classified into fully or semi-automated. The cell formatting is the same as in Tab. 2, but note that '?' means that the result is not available in the literature. The top performance for each measure and dataset is highlighted.

Method	"Good" Percentage	Endocardium ADM	Endocardium APD
Online set (15 sequences)			
Semi Automated			
Proposed model (semi)	**100(0)[100 − 100]**	**0.91(0.03)[0.85 − 0.96]**	**1.78(0.49)[1.17 − 3.15]**
Ngo and Carneiro (2013)	98.71(3.66)[86.66 − 100]	0.90(0.04)[0.83 − 0.95]	2.04(0.35)[1.53 − 2.67]
Fully Automated			
Proposed model	**90.54(14.40)[46.67 − 100]**	**0.89(0.03)[0.82 − 0.94]**	**2.17(0.46)[1.62 − 3.46]**

Full set (45 sequences)			
Semi Automated			
Proposed model (semi)	**100(0)[100 − 100]**	**0.91(0.03)[0.83 − 0.96]**	**1.73(0.31)[1.17 − 3.15]**
Ngo and Carneiro (2013)	97.91(6.18)[63.15 − 100]	0.90(0.03)[0.83 − 0.95]	2.08(0.40)[1.43 − 3.30]
Constantinides et al. (2012)	91.00(8.00)[61 − 100]	0.89(0.04)[0.80 − 0.96]	1.94(0.42)[1.47 − 3.03]
Fully Automated			
Proposed model	**94.55(9.31)[46.67 − 100]**	0.88(0.04)[0.76 − 0.95]	2.22(0.46)[01.27 − 3.46]
Constantinides et al. (2012)	80.00(16.00)[29 − 100]	0.86(0.05)[0.72 − 0.94]	2.44(0.56)[1.31 − 4.20]
Hu et al. (2012)	91.06(9.42)[?−?]	**0.89(0.03)**[?−?]	2.24(0.40)[?−?]
Huang et al. (2011)	79.20(19.00)[?−?]	**0.89(0.04)**[?−?]	**2.16(0.46)**[?−?]

Table 7: Quantitative experiments on the **training and testing sets** of the MICCAI 2009 challenge databases (Radau et al. (2009)) comparing the performance of our proposed approach with the state of the art on the **epicardium segmentation problem**. Notice that the methods are classified into fully or semi-automated. The cell formatting is the same as in Tab. 2, but note that '?' means that the result is not available in the literature. The top performance for each measure and dataset is highlighted.

Method	"Good" Percentage	Epicardium ADM	Epicardium APD
Testing set (15 sequences)			
Semi Automated			
Proposed model (semi)	**100(0)[100 − 100]**	**0.94(0.01)[0.92 − 0.97]**	**1.73(0.28)[1.16 − 2.17]**
Huang et al. (2009)	?	**0.94(0.01)**[?−?]	1.95(0.34)[?−?]
Uzunbaş et al. (2012)	?	**0.91(0.03)**[?−?]	1.78(0.35)[?−?]
Fully Automated			
Proposed model	94.65(6.18)[85.71 − 100]	0.93(0.02)[0.88 − 0.96]	2.08(0.60)[1.27 − 3.74]
Jolly (2009)	**95.60(6.90)[80.00 − 100]**	0.93(0.02)[0.90 − 0.96]	2.05(0.59)[1.28 − 3.29]
Wijnhout et al. (2009)	94.20(7.00)[80.00 − 100]	0.93(0.01)[0.90 − 0.96]	2.28(0.39)[1.57 − 2.98]
Lu et al. (2009)	81.11(13.95)[57.14 − 100]	**0.94(0.02)[0.90 − 0.97]**	**1.91(0.63)[1.06 − 3.26]**

Training set (15 sequences)			
Semi Automated			
Proposed model (semi)	**100.00(0.00)[100 − 100]**	**0.94(0.01)[0.91 − 0.96]**	**1.64(0.34)[1.17 − 2.47]**
Huang et al. (2009)	?	0.93(0.02)[?−?]	2.28(0.42)[?−?]
Fully Automated			
Proposed model	98.52(5.74)[77.78 − 100]	**0.93(0.02)[0.88 − 0.96]**	1.99(0.46)[1.35 − 3.13]
Jolly (2009)	**99.07(3.61)[86.00 − 100]**	**0.93(0.01)[0.91 − 0.95]**	**1.88(0.40)[1.20 − 2.55]**

Table 8: Quantitative experiments on the **online and full sets** of the MICCAI 2009 challenge databases (Radau et al. (2009)) comparing the performance of our proposed approach with the state of the art on the **epicardium segmentation problem**. Notice that the methods are classified into fully or semi-automated. The cell formatting is the same as in Tab. 2, but note that '?' means that the result is not available in the literature. The top performance for each measure and dataset is highlighted.

Method	"Good" Percentage	Epicardium ADM	Epicardium APD
Online set (15 sequences)			
Semi Automated			
Proposed model (semi)	**100.00(0.00)[100 − 100]**	**0.94(0.02)[0.88 − 0.96]**	**1.90(0.53)[1.22 − 3.16]**
Fully Automated			
Proposed model	**84.32(23.45)[12.50 − 100]**	**0.93(0.03)[0.84 − 0.95]**	**2.05(0.61)[1.39 − 3.63]**

Full set (45 sequences)			
Semi Automated			
Proposed model (semi)	**100(0)[100 − 100]**	**0.94(0.02)[0.88 − 0.97]**	**1.76(0.40)[1.16 − 3.16]**
Constantinides et al. (2012)	91.00(10.00)[70 − 100]	0.92(0.02)[0.84 − 0.95]	2.38(0.57)[1.28 − 3.79]
Fully Automated			
Proposed model	**92.49(15.31)[12.50 − 100]**	0.93(0.02)[0.84 − 0.96]	**2.04(0.55)[1.27 − 3.70]**
Constantinides et al. (2012)	71.00(26.00)[0 − 100]	0.91(0.03)[0.81 − 0.96]	2.80(0.71)[1.37 − 4.88]
Hu et al. (2012)	91.21(8.52)[?−?]	**0.94(0.02)**[?−?]	2.21(0.45)[?−?]
Huang et al. (2011)	83.90(16.80)[?−?]	0.93(0.02)[?−?]	2.22(0.43)[?−?]

Table 9: Average Dice metric and perpendicular distance on **all images of the MICCAI 2009 challenge databases** (Radau et al. (2009)) showing the performance of our proposed fully-automated approach for the **endocardium and epicardium segmentation problems**. The cell formatting is still the same as in Tab. 2.

Dataset	Endocardium ADM	Endocardium APD	Epicardium ADM	Epicardium APD
Online	0.86(0.09)[0.57−0.94]	3.04(2.36)[1.62−11.22]	0.89(0.09)[0.56−0.94]	3.38(3.38)[1.76−15.29]
Testing	0.86(0.05)[0.76−0.93]	2.67(0.89)[1.62−5.12]	0.92(0.03)[0.83−0.96]	2.59(1.35)[1.27−6.94]
Training	0.87(0.05)[0.74−0.95]	2.24(0.49)[1.27−2.83]	0.93(0.02)[0.89−0.96]	2.04(0.49)[1.35−3.13]
Full	0.86(0.06)[0.57−0.95]	2.65(0.51)[1.27−11.22]	0.91(0.02)[0.56−0.96]	2.67(0.38)[1.27−15.29]