Medical Image Analysis 35 (2017) 159–171

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

# 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[a,1], Zhi Lu[b,1], Gustavo Carneiro[c,*]

[a] Vietnam National University of Agriculture, Vietnam
[b] The University of South Australia, Australia
[c] Australian Centre for Visual Technologies, The University of Adelaide, Australia

ARTICLE INFO

*Article history:*
Received 13 July 2015
Revised 20 May 2016
Accepted 20 May 2016
Available online 26 May 2016

*Keywords:*
Deep learning
Level set method
Segmentation of the left ventricle of the heart
Cardiac cine magnetic resonance

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. Introduction

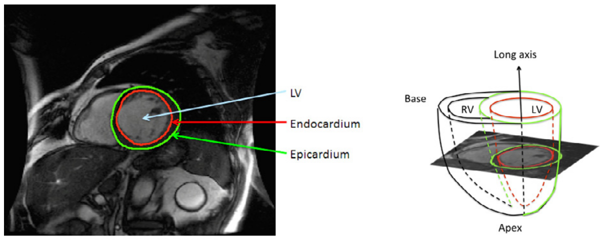
Medical image analysis segmentation problems are unique in the sense that they require highly accurate results, but at the same time provide relatively small annotated training sets. A typical example is the segmentation of the endocardium and epicardium from the left ventricle (LV) of the heart using 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

* Corresponding author.
*E-mail address:* gustavo.carneiro@adelaide.edu.au (G. Carneiro).
[1] Developed this work while they were with the University of Adelaide.

http://dx.doi.org/10.1016/j.media.2016.05.009
1361-8415/




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

**Fig. 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).

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 Vapnik, 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