

# A Combined Multi-scale Deep Learning and Random Forests Approach for Direct Left Ventricular Volumes Estimation in 3D Echocardiography

Suyu Dong<sup>1</sup>, Gongning Luo<sup>1</sup>, Guanxiong Sun, Kuanquan Wang<sup>1</sup>, Henggui Zhang<sup>1,2</sup>

<sup>1</sup>Harbin Institute of Technology, Harbin, China

<sup>2</sup>University of Manchester, Manchester, UK

## Abstract

*Estimation of left ventricular (LV) volumes from 3D echocardiography (3DE) is a popular clinical approach in accurate assessment of left ventricular function for the diagnosis of cardiac disease. The segmentation of 3DE volumes is a crucial step in traditional methods. Nevertheless, segmentation itself is an extremely challenging problem due to the presence of speckle noise and discontinuous edges. Therefore, direct left ventricular volumes estimation methods without the segmentation become attractive in cardiac function analysis. The aim of this paper is to present a fully learning framework to estimate the left ventricular volume in 3DE. The proposed method combined unsupervised multi-scale convolutional deep network and random forests. The multi-scale convolution deep network adopted multi-scale convolutional filters to represent features of unlabeled end-diastolic and end-systolic 3DE volumes (EDV and ESV). And then we formulated left ventricular volume estimation as a regression problem and used random forests for efficient volume estimation. The experiments results suggested that our proposed method is feasible and can achieve higher accuracy, even in case of echocardiography images with irregular geometry.*

## 1. Introduction

Assessment of cardiac function in cardiac diseases is an important topic in clinical practice and scientific research. Left ventricle (LV) volume estimation is one of the most important focus attracting wide attentions.

So far, there are a large number of imaging modalities for assessing cardiac functions. Magnetic resonance imaging (MRI), for its high spatial resolution, is known as the gold standard for the LV volume estimation. However, it is expensive and has low temporal resolution, which are crucial limitations for cardiac imaging. Instead, cardiac computed tomography (CT) has high spatial resolution, allowing a better distinction between myocardium and blood, as well as a high temporal

resolution to adapt to cardiac beat. However, as it needs injection of contrast agents, cardiac CT is not employed for cardiac imaging frequently. Due to the characteristics of low cost and real-time imaging, echocardiograph has become the typical modality for cardiac function detection in clinical diagnosis. Therefore, it is the most promising imaging modality that is widely used in clinical practice.

Based on echocardiograph, several different methods have been proposed to estimate the LV volumes [1-8]. Among them deformable models (such as active contours and level sets) are the most popular segmentation methods for estimating the LV volume. With the active contour method, segmentation was achieved based on optimization of energy function which consists of internal energy and external energy [1, 4]. Level sets use the zero level-set to represent the contour [2, 9, 10]. Moreover, active shape model (ASM) [5, 11] and active appearance model (AAM) [8, 12, 13], which are based on the prior statistical information, are used for segmentation of the contour. ASM and AAM use the manual labelled contour and joint distribution to estimate the parameters representing the shape or appearance of the LV. Classification methods [7, 14, 15], such as cluster, depend on a large data set. According to the features acquiring from dataset, these methods train classifiers and then classify the objects into different classes.

Note that the segmentation of LV is the crucial step of above methods. Nevertheless, segmentation itself is an extremely challenging problem due to the presence of speckle noise and discontinuous edges of images. Therefore, direct left ventricular volumes estimation methods without the segmentation still remains a challenge for cardiac function analysis [16]. The aim of this study was to develop a fully learning framework to estimate the left ventricular volume based on 3D echocardiographic images.

## 2. Methods

The flowchart of the proposed method is showed in Figure 1. The proposed method combines unsupervised

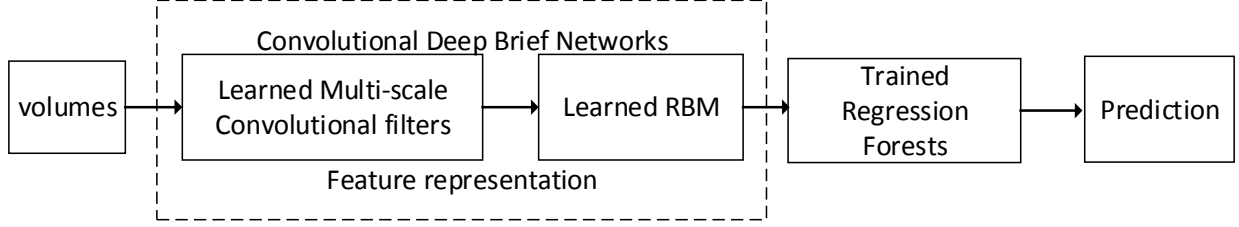


Figure 1. The proposed flowchart using unsupervised feature representation and random forest regression.

multi-scale convolutional deep network and random forests. First, the multi-scale convolution deep network adopted multi-scale convolutional filters and restricted Boltzmann machine (RBM) to represent features of unlabeled end-diastolic and end-systolic 3DE volumes (EDV and ESV). And then left ventricular volume estimation was formulated as a regression problem and random forests were used for efficient volume estimation.

## 2.1. Feature representation

Features, the crucial representations of cardiac images, will directly affect the results of regression forests. The proposed method used multi-scale convolutional deep network to acquire features of cardiac images.

### 2.1.1 Restricted Boltzmann machines

Restricted Boltzmann Machine (RBM), which contains a visible layer  $V$  and a hidden layer  $H$ , is a completely connected bipartite graph. The visible layer and the hidden layer are connected by the weight matrix  $W$ . However, there is no connection between any units which are among the same layer. So, if the visible layer is given, the hidden units are conditionally independent, and vice versa. Otherwise, the RBM is optimized by contrastive divergence approximation.

### 2.1.2 Convolutional Restricted Boltzmann machines

Convolutional Restricted Boltzmann Machines (CRBM) combine the RBM and convolutional filters [13]. Similar to RBM, it also has a visible layer  $V$  and a hidden layer  $H$ . But, the weights which connect the visible layer and hidden layer become shared weights  $W$  using the same convolutional filter for every group of hidden layer.  $N_V * N_V$  array denotes the input layer. There are  $k$  groups of hidden layer and every  $N_H * N_H$  array denotes a group of hidden layer. In our work, the size of every convolutional filter is different and it is denoted by  $N_K * N_K$ . Therefore, we had a multi-scale convolutional Restricted Boltzmann Machine. Besides, the visible layer has a bias  $c$  and every hidden group has a bias  $b_k$ . The definition of probabilistic semantics  $P(v, h)$  and energy function  $E(v, h)$  are :

$$P(v, h) = \frac{1}{Z} \exp(-E(v, h)) \quad (1)$$

$$E(v, h) = - \sum_{k=1}^K \sum_{i,j=1}^{N_H} \sum_{r,s=1}^{N_K} h_{ij}^k W_{rs}^k v_{i+r-1, j+s-1} - \sum_{k=1}^K b_k \sum_{i,j=1}^{N_H} h_{ij}^k - c \sum_{i,j=1}^{N_V} v_{ij} \quad (2)$$

Where  $Z$  denotes partition function.

### 2.1.2 Convolutional Deep Belief Networks

Based on the CRBM and probabilistic max-pooling, the Convolutional Deep Belief Networks (CDBN) was proposed [17]. Probabilistic max-pooling is a generative model. The energy of probabilistic max-pooling-CRBM is defined as:

$$E(v, h) = - \sum_k \sum_{i,j} \left( h_{i,j}^k \left( w_{i,j}^k * v \right) + b_k h_{i,j}^k \right) - c \sum_{i,j} v_{i,j} \quad (3)$$

CDBN consists of several probabilistic max-pooling-CRBM and uses the method which is same to DBN to train the networks. The training method treats the each layer of DBN as an RBM. And make the output of trained layer as the input of the next layer. Figure 2 shows the feature maps of CDBN.

## 2.2. Random forests

Random forests [6], a typical machine learning method, consist of many binary trees. Simple random sampling with replacement is used to select samples, which are treated as the training subset of every tree. Analogously to the selection of training subset, there are parts of features used as split features for every tree. The features are randomly selected without replacement for subtree for split. According to the judgement of the split node, the data in the current split node is sent to right or left child node. In our work, we used random forests as a regression method to estimate the LV volume. The input of random forests is the feature maps acquiring by CDBN, and the output is the LV volume.

## 3. Results

### 3.1. Datasets and setting

In our experiments, the datasets contain a training set and a validation set. For the training set, 60 unlabeled

echocardiography volumes from 30 patients are used for unsupervised feature learning, 120 labeled echocardiography volumes from 60 patients for the training of random forests. For the validation set, 120 labeled echocardiography volumes from another 60 patients for the validation of left volume estimation. For the training set and validation set, there are half patients with cardiac dysfunction and others are healthy. Due to the valve of base part and the muscle tissue of the apical part are not benefit for the volume estimation, we acquired one-third volume from the middle of the full volume for every volume to represent the original volume. Furthermore, we sliced one-third volume into ten frames uniformly.

For the proposed method, we used a CDBN and random forests. The CDBN followed the original work [17], the connections between all layers are undirected. And there are three hidden layers for CDBN. There are 16, 64 and 128 groups for the first layer, second layer and third layer, respectively. The sizes of the corresponding group are 51\*51, 7\*7 and 3\*3. The dimension of input of the CDBN is 10\*128\*128.

Random forests are used for the analysis of regression of LV volume. Due to the range of LV volume is from 1 to 600ml, we design 600 leaf nodes for every tree. As we know, the number of trees is crucial for the random forests, it has large effect on the effectiveness and efficiency of the experiments results. Therefore, we compared the experiments results of different number of trees to select the appropriate number of trees for the experiments.

### 3.2. Metrics

We adopted the EDV, ESV and EF to evaluate the prediction results of the proposed method. EF is the important clinical index and is calculated by the following formula:

$$EF = \frac{EDV - ESV}{EDV} \quad (4)$$

Otherwise, we compared the estimated volumes with the ground truth from cardiologist in terms of correlation coefficients and mean errors on EDV, ESV, and EF.

### 3.3. Experiments results and discussion

The experiments results suggested the proposed method is effective. Figure 3 shows the effect of the number of trees of random forests. With the increase of number of trees, the correlation coefficient increased but with a small fluctuation. However, the cost of computation increased as well. We used 200 trees in our experiments which can effectively balance the accuracy and the cost of computation. This also indicates that if we ignored the cost computation or used high-performance

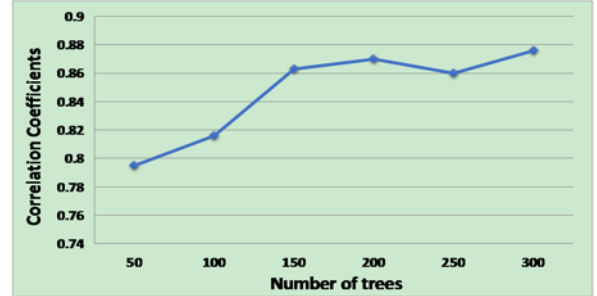


Figure 3. The performance of random forests with different numbers of trees.

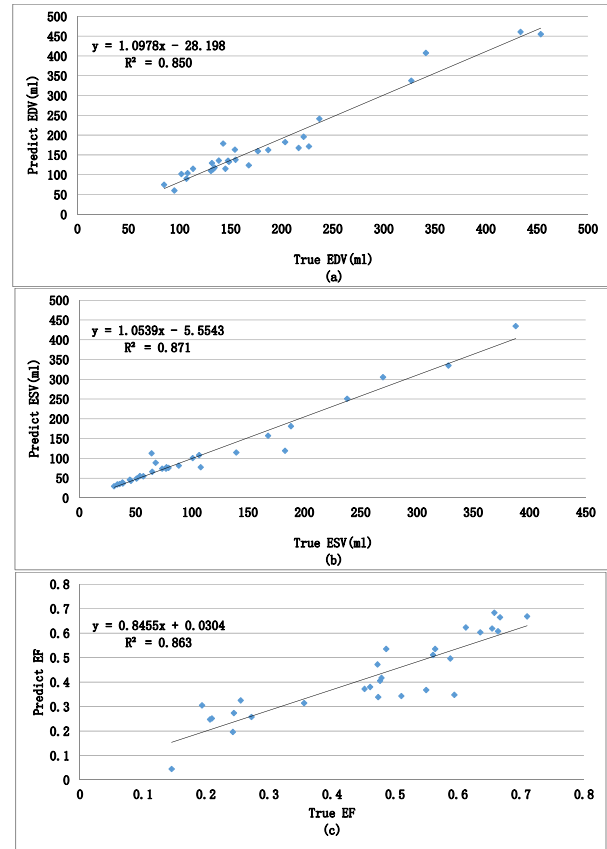


Figure 4. Correlation graphs. The correlation between predicted EDV and true EDV, predicted ESV and true ESV, predicted EF and true EF, respectively.

machine, we can use more trees to achieve higher accuracy.

Figure 4 shows the correlation between the predicted results and the ground truth. The R values for EDV, ESV and EF are 0.850, 0.871 and 0.863. We also used Bland-Altman (BA) to analyze the differences distribution between our method and ground truth. Figure 5 shows the difference distribution between the predicted results and ground truth, the means for EDV, ESV and EF are -8.4ml, 0.4ml, 0.04, the confidence intervals are (-59.9-43ml), (-39.9-40.6ml) , (-0.19-0.12). We found that the results of Bland-Altman analysis is worse, which can be related to

the lack of adequate dataset to train and validate the model. If we have more data, it is possible to achieve

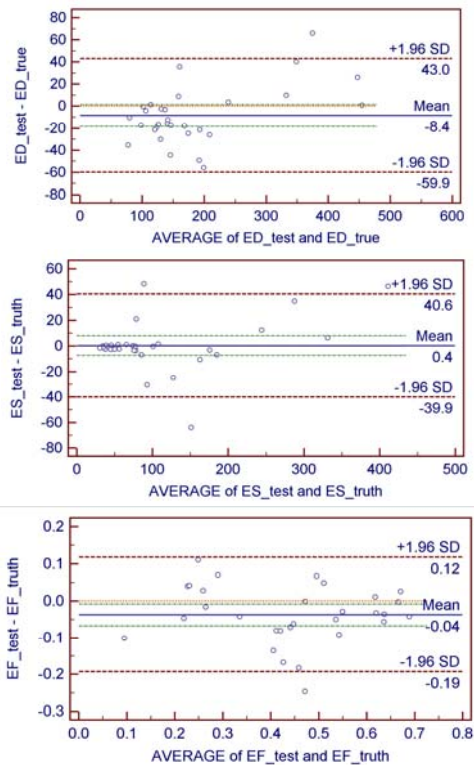


Figure 5. The performance of random forests with different numbers of trees.

better agreement and reduce the overfitting. However, the high correlation shows clinical indexes derived from the estimated volume were in good agreement with ground truth from manual evaluation, and our method has potential clinical applicability for the LV volume estimation.

#### 4. Conclusion

In this paper, we proposed a method for direct estimation of LV volume using a fully learning method. The method avoids the segmentation of 3DE volumes, a crucial step in traditional methods. On the whole, this method used the CDBN to capture features from original volumes and the trained random forests to estimate the LV volume. The experiments results demonstrate the method is feasible and can accurately evaluate the LV volume, even in case of irregular geometry of images.

#### Acknowledgements

This work is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61571165, No. 61572152.

#### References

[1] De Alexandria AR, Cortez PC, Bessa JA, Felix JHD, de Abreu JS, de Albuquerque VHC. pSnakes: A new radial

active contour model and its application in the segmentation of the left ventricle from echocardiographic images. *Comput Meth Prog Bio.* 2014;116(3):260-73.

[2] Dietenbeck T, Barbosa D, et al. Whole myocardium tracking in 2D-echocardiography in multiple orientations using a motion constrained level-set. *Med Image Anal.* 2014;18(3):500-14.

[3] Saini K, Dewal ML, Rohit M. A Fast Region-Based Active Contour Model for Boundary Detection of Echocardiographic Images. *J Digit Imaging.* 2012;25(2):271-8.

[4] Kass M WA, Terzopoulos D. Snakes: Active contour models. *International journal of computer vision.* 1988; 1(4): 321-31.

[5] Cootes T F TCJ, Cooper D H, et al. . Active shape models-their training and application. *Computer vision and image understanding.* 1995; 61(1):38-59.

[6] L B. Random forests. *Machine learning.* 2001; 45(1):5-32.

[7] Georgescu B ZXS, Comaniciu D, et al. Database-guided segmentation of anatomical structures with complex appearance. *IEEE Computer Vision and Pattern Recognition (CVPR'05);* ; 2005. p. 429-36.

[8] Barbosa D, Dietenbeck T, et al. Fast and fully automatic 3-d echocardiographic segmentation using B-spline explicit active surfaces: feasibility study and validation in a clinical setting. *Ultrasound Med Biol.* 2013;39(1):89-101.

[9] Lin N, Yu W, Duncan JS. Combinative multi-scale level set framework for echocardiographic image segmentation. *Med Image Anal.* 2003;7(4):529-37.

[10] Dietenbeck T, Alessandrini M, Barbosa D, D'hooge J, Friboulet D, Bernard O. Detection of the whole myocardium in 2D-echocardiography for multiple orientations using a geometrically constrained level-set. *Med Image Anal.* 2012;16(2):386-401.

[11] Hansgård J OF, Rabben S I. Real-time active shape models for segmentation of 3D cardiac ultrasound. *International Conference on Computer Analysis of Images and Patterns: Springer Berlin Heidelberg;* 2007. p. 157-64.

[12] Mitchell SC, Lelieveldt BP, et al. Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images. *IEEE Trans Med Imaging.* 2001;20(5):415-23.

[13] Barbosa D, Dietenbeck T, Schaerer J, D'Hooge J, Friboulet D, Bernard O. B-spline explicit active surfaces: an efficient framework for real-time 3-D region-based segmentation. *IEEE transactions on image processing* 2012;21(1):241-51.

[14] Carneiro G, Nascimento JC, Freitas A. The Segmentation of the Left Ventricle of the Heart From Ultrasound Data Using Deep Learning Architectures and Derivative-Based Search Methods. *Ieee T Image Process.* 2012;21(3):968-82.

[15] Luo GN, An R, Wang KQ, Dong SY, Zhang HG. A deep learning network for right ventricle segmentation in short-axis MRI. *Computing in Cardiology: IEEE;* 2016.

[16] Luo GN, Sun GX, et al. A novel left ventricular volumes prediction method based on deep learning network in cardiac MRI. *Computing in Cardiology: IEEE;* 2016.

[17] Lee H GR, Ranganath R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. *Proceedings of the 26th annual international conference on machine learning ACM* 2009. p. 609-16.