AUTOMATIC SEGMENTATION AND CARDIOPATHY CLASSIFICATION IN CARDIAC MRI IMAGES BASED ON DEEP NEURAL NETWORKS

Yakun Chang[†]*, Baoyu Song*[†]*, Cheolkon Jung, and Liyu Huang*

School of Electronic Engineering, Xidian University, Xian, Shaanxi 710071, China
zhengzk@xidian.edu.cn

ABSTRACT

Segmentation of cardiac MRI images plays a key role in clinical diagnosis. In the traditional diagnostic process, clinical experts manually segment left ventricle (LV), right ventricle (RV) and myocardium to obtain guideline for cardiopathy diagnosis. However, manual segmentation is time-consuming and labor-intensive. In this paper, we propose automatic segmentation and cardiopathy classification in cardiac MRI images based on deep neural networks. First, we perform object detection based on a YOLO-based network to get region of interest (ROI) from the whole sequence of diastolic and systolic MRI. Then, we obtain a pixel-wise segmentation mask automatically by feeding ROI into fully convolutional neural networks (FCN). Finally, we construct a fully connected network for cardiopathy diagnosis to decide a heart disease from the given MRI. Experimental results show that the proposed method successfully segments LV, RV and myocardium as well as achieves 90% accuracy in heart disease classification.

Index Terms— Cardiac MRI, cardiopathy diagnosis, convolutional neural networks, medical image analysis, segmentation

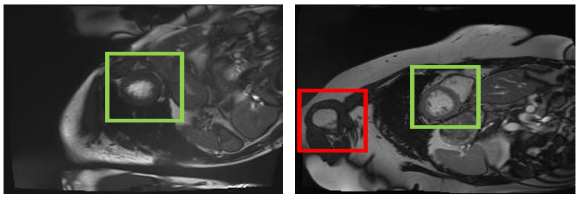
1. INTRODUCTION

Cardiovascular diseases are leading cause of death worldwide, and 17.3 million deaths are caused by it every year. The number is expected to grow up to more than 23.6 million by 2030 [1]. More and more people place importance on the cardiopathy diagnosis and prevention. Cardiac MRI offers key information for cardiovascular diagnosis by enabling quantitative assessment of functional parameters such as myocardium thickness, volume of LV and RV, ejection fraction (EF) [2]. Thus, cardiac MRI segmentation has become an emerging medical imaging issue. However, due to the special characteristics of cardiac MRI, segmentation of the heart is a challenging task. For instance, the brightness of the LV intracavity is heterogeneous, and the signals of the other organs are similar to that of the heart, the severe malformations caused by heart disease, and the complexity of the apical and basal slice images. Thus, cardiac MRI segmentation is still a problem worth studying. Many existing methods only segment LV or RV [3, 4, 5, 6]. However, when clinical experts diagnose heart diseases such as hypertrophic cardiomyopathy, and abnormal right ventricle, segmentation of both LV and RV is required. There are also some researches on dual LV and RV segmentation [7, 8, 9], but the accuracy still needs improvement. In recent years, with the advent of deep learning methods, more and more researchers have been trying deep learning on cardiac MRI segmentation. Avendi et al. [3] segmented LV within two steps. First, they down-sampled images and obtained ROI for accurate segmentation. Then, they performed segmentation of LV on the ROI. However, through observations, we found that some images do not contain objects in a whole sequence of cardiac MRI. Thus, prior to generating ROI, the existence of an object should be determined. Moreover, this approach [3] only worked for LV segmentation. Similar to [3], Kheradvar et al. [10] first detected its location and then performed RV segmentation using convolutional neural networks (CNN). Since the heart localization eliminates interference of irrelevant information such as similar signals of some other organs, it is an important issue to improve the accuracy of cardiopathy classification.

In this paper, we propose automatic segmentation and cardiopathy classification in cardiac MRI images based on deep neural networks. The proposed segmentation consists of two main steps: 1) Object detection before obtaining ROI, 2) Segmentation on the ROI. First, we detect objects and obtain ROI based on You Only Look Once (YOLO) [11]. Then, we perform automatic segmentation of LV cavity, RV cavity and myocardium using an FCN-based network. Next, we extract features such as parameters on the volumes of LV, EF from the segmentation mask, and finally classify heart diseases by a fully connected network. The classification accuracy is 90%, which can be used as an indicator for clinic. Compared with existing methods, main contributions of the proposed method are as follows:

- We perform YOLO-based object detection to generate ROI because some slices of cardiac MRI don't contain objects.
- We simultaneously segment LV cavity, RV cavity and

[†] Authors contributed equally. This work was supported by the National Natural Science Foundation of China (Nos. 61271298 and U1401255) and the International S&T Cooperation Program of China (No. 2014DFG12780).



(a) (b)

Fig. 1: Samples from ACDC database. Green bounding boxes in (a) and (b) are organs which segmentation is needed, while red one is the interference signal that cause wrong segmentation.

myocardium using FCN.

- We perform cardiopathy classification for heart disease diagnosis with the cardiac segmentation.

2. MATERIALS AND METHODS

2.1. Database

In this work, we use the Automated Cardiac Diagnosis Challenge (ACDC) database in MICCAI challenge 2017. The database consists of cardiac MRI images from 150 different patients, we use 100 for training and 50 for testing. It is divided into five evenly distributed subgroups according to etiology: normal subjects (NOR), previous myocardial infarction (MINF), dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), abnormal right ventricle (ARV). Furthermore, each patient comes with the following additional information : weight, height, as well as the diastolic and systolic phase instants.

2.2. Methods

2.2.1. Detection and Localization

The short axis of the cardiac MRI database has a wide range of observation that covers the heart and other surrounding organs. Thus, cardiac MRI contains a lot of irrelevant information, and sometimes similar signals from other organs would interfere segmentation. For example, as shown in Fig. 1, the cardiac regions are marked by green bounding boxes in (a) and (b). However, the region marked by the red bounding box in (b) is similar to the LV in (a). Thus, machine may consider it as LV, which causes wrong segmentation. To automatically detect an object and identify its location, we construct a small network based on YOLO [11]. First, training pairs for detection and localization are generated by 100 training samples of the cardiac MRI. Although MRI consists of a whole sequence of images, ground truth data are available only at the end of diastole (ED) and end of systole (ES). Thus, only 1840 image pairs are available for training. For each bounding box, we get

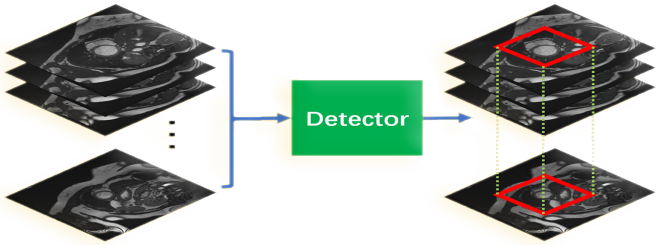


Fig. 2: A whole sequence of diastolic and systolic MRI is detected at the same time.

Table 1: Details of detector. Each convolutional layer is connected by Leaky Relu, max pooling and batch normalization [12]. Padding for each convolutional layer is set to be 'same'.

Type	K-size	Out-channels	Stride
conv.	3	16	1
conv.	3	32	1
conv.	3	64	1
conv.	3	128	1
conv.	3	256	1
conv.	3	512	1
fc		4096	
Output		7*7*30	

a vector [*x, y, w, h*], where *x*, *y* are coordinates of the center point; and *w*, *h* are width and height, respectively. When training is started, these vectors are the final ground truth of the detector. As shown in Fig. 2, when we detect objects on testing data, a whole sequence of images at ED and ES are detected at the same time. For a sequence of images $[I_1, I_2, ..., I_N]$, one object is mostly detected in each image, and (x_i, y_i) is predicted bounding box center of I_i. One image does not need to be segmented if the detector does not find the object. For most cardiac MRI, hearts are accurately detected. To ensure accuracy, the center of segmentation object in I_i is determined by:

$$\begin{cases} (x_i, y_i) \leftarrow (x_i, \quad y_i); if\ |x_i - \overline{x}| \leq 10, |y_i - \overline{y}| \leq 10 \\ (x_i, y_i) \leftarrow (\overline{x}, \quad \overline{y}); otherwise \end{cases}$$

where $(\overline{x}, \overline{y})$ is the mean of a whole sequence MRI at ED and ES. This scheme makes a more accurate detection and localization. The network architecture of the detector is shown in Table 1. The input image is divided into 7×7 grids: for each grid, we predict 5 bounding boxes, while for each bounding box, we get a confidence, a probability of heart, and [*x, y, w, h*]. Thus, the third dimension of the output is $(1 + 1 + 4) \times 5 = 30$ [11].

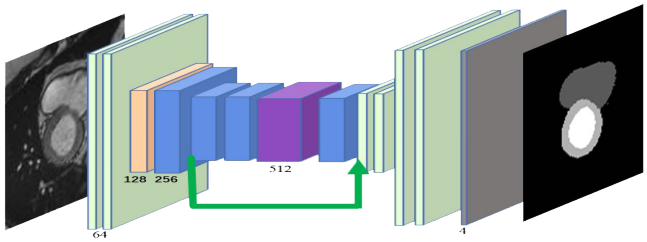


Fig. 3: Network architecture of the proposed method. Batch normalization and pooling layers are omitted. Features in the same channels are marked and painted by the same color. For more dense prediction, we add a link that combines feature1 with features at a high layer. The network outputs are four channel predictions of LV cavity, RV cavity, myocardium and background.

2.2.2. *Segmentation*

Fig. 3 shows the network architecture for segmentation based on FCN. The center of ROI is obtained by the detector, and all images are cropped to 100 × 100 according to the central pixels. The overall network is formed by: Conv(*k* = 7, *out* = 64, *s* = 1) × 2 + BN + Maxpooling → Conv(2, 128, 1) + Conv(2, 256, 1) + BN + Maxpooling (feature1) → Conv(2, 256, 1) × 2+Conv(1, 512, 1) → Deconv(5, 256, 1) (+feature1) → Deconv(5, 64, 2) + Deconv(5, 64, 1) + BN → Deconv(5, 64, 2) + Deconv(7, 64, 1) + Deconv(7, 4, 1) + BN → Softmax, where *k* denotes kernel size, *out* denotes channels of outputs, *s* denotes stride, BN is the batch normalization layer to make better gradient propagation. Denote an input image as I, X: [$x_1, x_2, \cdots, x_i, \cdots, x_M$] as the collection of all pixels of I, the collection of labels as C: [$c_1, c_2, \cdots, c_j, \cdots, c_N$], and one hot label of the pixel x_i as y_i. The features of *I* are extracted by convolutional layers; after then, we use deconvolutinal layers to reduce dimension and make a prediction for each pixel. As illustrated in Long et al.'s work [13], combining coarse and high layer information with fine low layer information refines the spacial precision of segmentation results. Thus, we combine the higher layers with lower layers to capture better spatial information. The final deconvolutional layer produces a four-channel feature map. For the pixel x_i, the corresponding location at the channel j is obtained as follows:

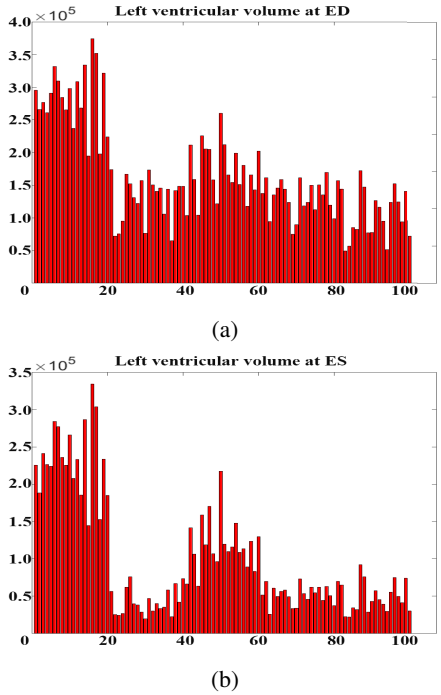
$$p(x_i = c_j) = \frac{1}{Z} \times \exp[v(c_j)] \tag{1}$$

where $p(x_i = c_j)$ is the probability of c_j, $v(c_j)$ is the value of c_j, Z is normalization term. The prediction of pixel x_i is performed as follows:

$$\hat{y_i} = \text{argmax}[p(x_i = c_j)] \tag{2}$$

The loss function is defined as follows:

$$Loss = -\frac{1}{MN} \sum_i \sum_j y_{ij} \ln[p(x_i = c_j)] \tag{3}$$



(a)

(b)

Fig. 4: Volume of LV: (a) ED, (b) ES. From left to right in (a) and (b): DDCM (1-20), HCM (21-40), MINF (41-60), NOR (61-80), and ARV (81-100).

2.2.3. *Cardiopathy Classification*

As aforementioned in Section 2.1, the data is evenly divided into five groups: NOR, MINF, DCM, HCM, and ARV. In addition to the ground truth of diastolic and systolic segmentation, pixel spatial resolution and inter slice gap are given. Thus, we obtain the volume of LV and RV, and extract ten features: Volume of RV and LV at ED and ES, EF of LV and RV, the volume ratio between RV and LV at ED and ES, myocardial volume at ED and ES. Besides, the height and weight of patients are also given. Totally, twelve features are used for the classification.

Fig. 4 shows the distribution of EF of LV in the training data. It can be observed that LV volume of patients with DCM is much larger than others, LV EF of patients with HCM is higher than the others, and ventricular volume at ES is usually lower than the normal samples. Although two distributions are given, it is obvious that the distribution for each disease follows inherent laws. To learn this pattern, we construct a tiny fully connected neural network to diagnose heart diseases. The fully connected network is formulated as: fc (n-unit = 500, dropout = 0.8) → fc (n-unit = 100) → fc (n-unit = 5) → softmax.

3. EXPERIMENTAL RESULTS

For the FCN-based model, we use Adam optimizer and set learning rate to 0.0001. Fig. 5 shows segmentation results

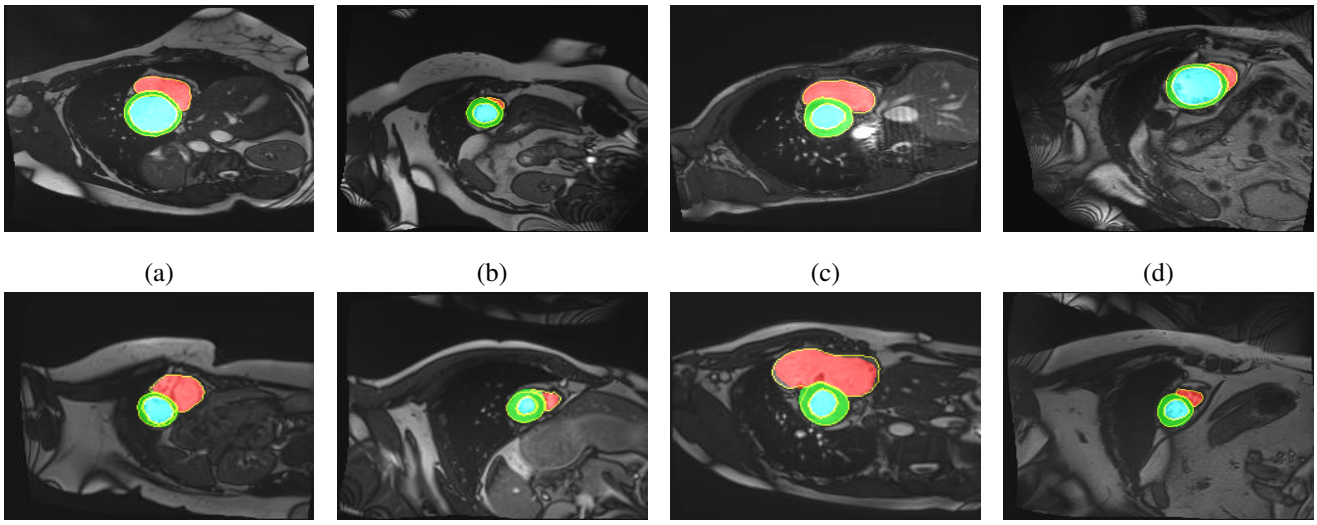


Fig. 5: Segmentation results from different patients. Cyan: LV cavity. Green: Myocardium. Red: RV cavity. Yellow lines: Ground truth.

from different patients. It can be observed that the proposed method accurately segments LV, RV and myocardium close to the ground truth. As shown in Fig. 5(d), the proposed method successfully distinguishes intraventricular regions and myocardium even in complex intra ventricle. In (f) and (h), the LV is also successfully segmented in basal slice image. All segmentation results are evaluated in terms of Dice coefficient (Dice), Hausdorff distance (HD) and sensitivity. Mean scores of Dice and HD are listed in Table 2. Dice for LV cavity is higher than others. It seems that the contrast between cavity and myocardium is higher, while myocardium and RV have relatively lower contrast over their surroundings. The sensitivity of LV and myocardium are similar, and higher than that of RV. This is because RV is more various in shape. For the cardiopathy classification, the proposed method achieves an accuracy of 90% on 50 testing samples. Notice that the ACDC database is relatively small for the heart disease classification task. We use twenty samples per disease for training, and only five groups of samples are given. However, cardiopathy diagnosis is more various in clinic and the heart disease is sometimes very complex. For instance, a patient with DCM may also have ARV and its common reasons of dilated LV are ischemic heart disease, hypertension, alcoholic cardiomyopathy, dilated cardiomyopathy, and valvulopathy. Thus, more studies are needed until being applied to the clinic.

Table 2: Performance evaluation of LV cavity, RV cavity and myocardium in terms of Dice coefficient (Dice), Hausdorff distance (HD) and sensitivity

Measures	Dice	HD	Sensitivity
LV cavity	0.9193	10.452	0.9085
RV cavity	0.8692	10.517	0.8614
Myocardium	0.8787	9.857	0.9043

4. CONCLUSION

We have proposed a fully automatic method for cardiac MRI segmentation and cardiopathy diagnosis based on deep neural networks. First, we have fed a whole sequence of MRI at ED and ES into YOLO-based detector to detect ROI and prevent interference from other organs. Then, we have performed automatic segmentation of cardiac structures on the detected ROI. Besides, we have conducted cardiopathy classification using a fully connected network. Experimental results demonstrate that the proposed method produces good segmentation results close to the ground truth and achieves 90% accuracy in cardiopathy classification. Therefore, it provides a promising solution to diagnosing cardiopathy from cardiac MRI.

5. REFERENCES

[1] Liset Vázquez Romaguera, Francisco Perdigón Romero, Cicero Ferreira Fernandes Costa Filho, and Marly Guimarães Fernandes Costa, "Left ventricle segmentation in cardiac mri images using fully convolutional neural networks," in *Proceedings of SPIE Medical Imaging*. International Society for Optics and Photonics, 2017, pp. 101342Z–101342Z.

[2] Jelmer M Wolterink, Tim Leiner, Max A Viergever, and

Ivana Isgum, "Automatic segmentation and disease classification using cardiac cine mr images," *arXiv preprint arXiv:1708.01141*, 2017.

[3] M. R. Avendi, Arash Kheradvar, and Hamid Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri," *Medical Image Analysis*, vol. 30, pp. 108–119, 2016.

[4] Marie Pierre Jolly, Nicolae Duta, and Gareth Funkalea, "Segmentation of the left ventricle in cardiac mr images," in *Proceedings of the IEEE International Conference on Computer Vision*, 2001, pp. 501–508.

[5] M. R. Kaus, Berg J Von, J Weese, W Niessen, and V Pekar, "Automated segmentation of the left ventricle in cardiac mri," *Medical Image Analysis*, vol. 8, no. 3, pp. 245–254, 2004.

[6] Caroline Petitjean, Maria A Zuluaga, Wenjia Bai, Jean-Nicolas Dacher, Damien Grosgeorge, Jérôme Caudron, Su Ruan, Ismail Ben Ayed, M Jorge Cardoso, Hsiang-Chou Chen, et al., "Right ventricle segmentation from cardiac mri: a collation study," *Medical Image Analysis*, vol. 19, no. 1, pp. 187–202, 2015.

[7] Jyrki Lötjönen, Sari Kivistö, Juha Koikkalainen, Daniel Smutek, and Kirsi Lauerma, "Statistical shape model of atria, ventricles and epicardium from short-and long-axis mr images," *Medical Image Analysis*, vol. 8, no. 3, pp. 371–386, 2004.

[8] Hui Sun, Alejandro F Frangi, Hongzhi Wang, Federico M Sukno, Catalina Tobon-Gomez, and Paul A Yushkevich, "Automatic cardiac mri segmentation using a biventricular deformable medial model," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2010, pp. 468–475.

[9] Dwarikanath Mahapatra, "Cardiac mri segmentation using mutual context information from left and right ventricle," *Journal of Digital Imaging*, vol. 26, no. 5, pp. 898–908, 2013.

[10] Michael R Avendi, Arash Kheradvar, and Hamid Jafarkhani, "Automatic segmentation of the right ventricle from cardiac mri using a learning-based approach," *Magnetic Resonance in Medicine*, 2017.

[11] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.

[12] Sergey Ioffe and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the International Conference on Machine Learning*, 2015, vol. 37, pp. 448–456.

[13] Jonathan Long, Evan Shelhamer, and Trevor Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.