DBAN: Adversarial Network with Multi-Scale Features for Cardiac MRI Segmentation

Xinyu Yang, Yuan Zhang*, *Senior Member*, *IEEE*, Benny Lo, *Senior Member*, *IEEE*, Dongrui Wu, *Senior Member*, *IEEE*, Hongen Liao, *Senior Member*, *IEEE* and Yuan-Ting Zhang, *Fellow*, *IEEE*

Abstract—With the development of medical artificial intelligence, automatic magnetic resonance image (MRI) segmentation method is quite desirable. Inspired by the power of deep neural networks, a novel deep adversarial network, dilated block adversarial network (DBAN), is proposed to perform left ventricle, right ventricle and myocardium segmentation in short-axis cardiac MRI. DBAN contains a segmentor along with a discriminator. In the segmentor, the dilated block (DB) is proposed to capture and aggregate multi-scale features. The segmentor can produce segmentation probability maps while the discriminator can differentiate the segmentation probability map and the ground truth at the pixel level. In addition, confidence probability maps generated by the discriminator can guide the segmentor to modify segmentation probability maps. Extensive experiments demonstrate that DBAN has achieved the state-of-the-art performance on the ACDC dataset. Quantitative analyses indicate that cardiac function indices from DBAN are similar to those from clinical experts. Therefore, DBAN can be a potential candidate for short-axis cardiac MRI segmentation in clinical applications.

Index Terms—Cardiac MRI, Medical Image Processing, Automatic Segmentation Method, Adversarial Network.

I. Introduction

CARDIOVASCULAR disease is the most fatal diseases in the world even if the diagnostic procedures have been facilitated thanks to the development of cardiac imaging technologies [1]. The short-axis cardiac magnetic resonance image (MRI) in particular, as a non-invasive imaging modality, is widely used by physicians for the diagnosis of cardiovascular diseases [2]. The important reason is that short-axis cardiac MRI with high spatial and contrast resolution supports accurate evaluation of cardiac function. In clinical cardiology, clinicians manually delineate the left ventricle (LV), right ventricle (RV), and myocardium (MYO) in short-axis cardiac MRI. In Fig. 1, examples are presented to show the delineated short-axis cardiac MRI. As the prerequisite, the delineated short-axis cardiac MRI is used to calculate cardiac function indices such as myocardium mass in end-diastolic phase, ventricular volume and ejection fraction. These cardiac function indices are important references to the diagnosis of cardiovascular diseases. However, manual delineation is both time-consuming and prone to subjective errors due to the large number of MRI slices and ambiguous boundaries. Therefore, automatic and accurate short-axis cardiac MRI segmentation approaches are demanded urgently.

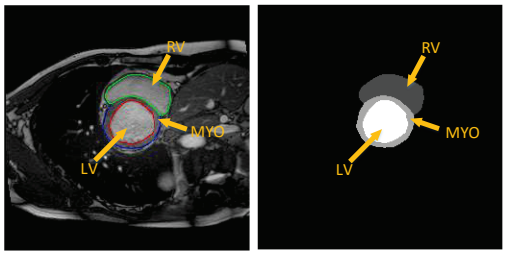


Fig. 1: Left: manually delineated short-axis cardiac MRI. Right: mask of manually delineated short-axis cardiac MRI.

In order to perform short-axis cardiac MRI segmentation, some machine learning methods were proposed. However, many of them were semi-automatic methods [3]. In order to facilitate the development of automatic short-axis cardiac MRI segmentation technologies, Automated Cardiac Diagnosis Challenge (ACDC) was launched in Medical Image Computing and Computer Assisted Intervention (MICCAI) 2017 where the ACDC dataset was released [4]. Note that the ACDC is still active. Hence it is reasonable to believe that approaches from ACDC leaderboard represent state-of-the-art level.

In this paper we propose a novel deep adversarial network which consists of a segmentor and a discriminator. Under the constraint of discriminator, the segmentor can produce accurate segmentation results automatically. The main contributions of the work are summarized as follows:

1) A novel network, dilated block adversarial network (DBAN), is proposed for the automatic LV, RV and MYO segmentation. Within the segmentor, the dilated block (DB) is designed to capture and aggregate multi-scale features without information loss or increasing too many

Xinyu Yang and Yuan Zhang are with the Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, College of Electronic and Information Engineering, Southwest University, Chongqing, China (corresponding author: Yuan Zhang).

Benny Lo is with the Hamlyn Centre, Imperial College London, London, UK.

Dongrui Wu is with the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China.

Hongen Liao is with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China.

Yuanting Zhang is with the Department of Mechanical and Biomedical Engineering, City University of Hong Kong, Hong Kong, China.

Manuscript received April 19, 2005; revised August 26, 2015.

parameters. This innovation, to the best of our knowledge, is a pioneering study in combining DB with adversarial network to perform short-axis cardiac MRI segmentation.

2) In the discriminator, a fully convolutional scheme is adopted to differentiate the segmentation probability map from the ground truth at the pixel level. Moverover, our discriminator can produce confidence probability maps helping the segmentor refine the segmentation results.

The paper is structured as follows: The related works are reviewed in Section II. The proposed method is detailed in Section III. In Section IV, we present and analyze the experiment results, and finally the conclusion is drawn in Section V.

II. Related Work

Starting from the 21st century, machine learning approaches have been used for segmentation. For example, a region growing method was proposed by Codella et al. [5] to segment the LV. However, the seed points needed to be picked manually. Pluempitiwiriyawej et al. [6] presented a novel stochastic active contour scheme (STACS) which could overcome some unique challenges in cardiac MRI, such as adverse effect of papillary muscles on segmentation. Unfortunately, this method was less sensitive to the contour and the accuracy was still limited. In addition, Zhang et al. [7] proposed a novel external force named as gradient vector flow over manifold (GVFOM) for active contours to perform image segmentation. The GVFOM active contours presented good results with respect to cardiac image segmentation. However, this method required an appropriate choice of parameters. Some researchers also used a priori probabilistic atlas to perform ventricular segmentation [8], which primarily required a priori knowledge about shape and appearance. The performance of this approach relied on whether prior knowledge was sufficient [9]. In general, there are still some drawbacks in machine learning methods, such as manual interaction or limited accuracy.

With the rapid development of deep learning technology, many automatic segmentation approaches have been proposed in the last few years. Jonathan Long et al. [10] proposed fully convolutional networks (FCN) which opened a new period of the field of automatic segmentation. Based on FCN, U-Net [11] was proposed which can produce reasonable segmentation result with a small training dataset. By adding skip connections between the contracting and expanding paths, this network architecture is very effective in medical image processing. Except for the above classic segmentation architectures, Ngo et al. [12] introduced a new methodology that combined deep learning and level set to train small datasets for the automated segmentation of the LV. Wang et al. [13] developed the cascaded segmentation and regression network (CSRNet) which firstly generated contours of the LV and then estimated the desired LV metrics. In order to improve segmentation performance, some researchers also used dilated convolution in neural networks to aggregate features [14] [15]. Generally, the emergence of deep learning based method further improves the performance of automatic segmentation.

In addition to the works mentioned above, Goodfellow et al. [16] proposed generative adversarial networks (GAN) which borrowed the idea from game theory. Based on GAN, conditional generative adversarial nets (CGAN) was raised by Mirza et al. [17]. This architecture consisting of a generator and a discriminator could produce particular data, which provided researchers with new ideas. Luc et al. performed segmentation using adversarial networks firstly [18]. Since then, many segmentation methods based on adversarial networks have emerged [19]. For example, Zhang et al. introduced a GAN-based method with a multi-scale feature fusion (MSFF) module to perform cardiac MRI segmentation [20]. Moeskops et al. developed a GAN-based methods with dilated convolution to conduct brain MRI segmentation [21]. Singh et al. combined adversarial networks with a convolutional neural network (CNN) to perform segmentation and classification [22]. GAN-based methods bring the development of network structures, which provides the possibility of further improving segmentation accuracy.

Overall, machine learning approaches are effective, but their performance is limited. Deep learning bring the improvement for the biomedical image segmentation accuracy and GAN-based methods provide the possibility of further improvment for segmentation performance. However, there still exist some limitations in current GAN-based methods. For the segmentor, multi-scale features are still not fully exploited for segmentation [22]. Even if some researchers have realized this issue and tried to fix it, new problems arise, such as information loss due to the gridding problem or dramatic parameter increase [20] [21]. In addition, the discriminator in most GAN-based method can only perform identification at the image level, which may limit the update of segmentor parameters [18] [19] [20] [21] [22]. Therefore, DBAN is proposed to address these issues. In the segmentor, DBs are designed to capture and aggregate multi-scale features without information loss or increasing too many parameters. In the discriminator, a fully convolutional scheme is adopted to perform identification at the pixel level. Moverover, confidence probability maps produced by our discriminator can guide the segmentor to refine the segmentation results. In the next section, we will detail the proposed DBAN.

III. Method

A. Network Architecture

The block diagram of the proposed method is shown in Fig. 2. Firstly we implement adversarial training on the proposed DBAN, as shown in TABLE I, after which the segmentation results are available from the model.

1) Segmentor: In deep learning, classical neural networks such as FCN commonly use large-scale features to perform segmentation, which is not able to handle the variation of morphology appropriately [23].There are

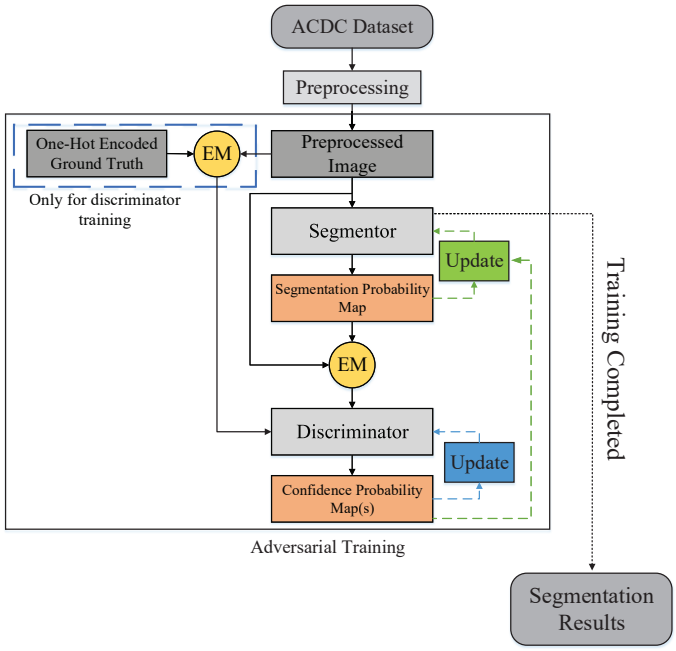


Fig. 2: The block diagram of the proposed method. EM denotes element-wise multiplication. During the training stage, the segmentor and the discriminator are trained alternately. When training the discriminator, only the parameters of the discriminator are updated, and the same scheme is applied on the segmentor. After training, the segmentation results can be obtained from the model.

TABLE I: Steps of adversarial training for DBAN

Algorithm1 Training of DBAN. $\bigotimes$ denotes element-wise multiplication.

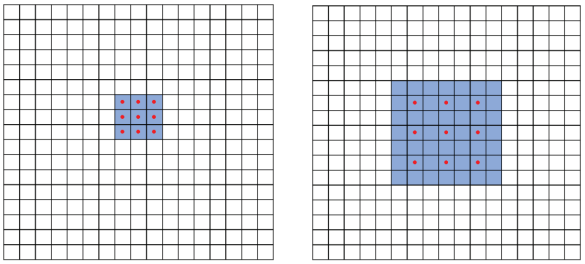
for number of training epochs **do**
 for number of iterations **do**
 - the preprocessed data is put into the segmentor
 - the preprocessed image $\bigotimes$ the segmentation probability map from the segmentor to construct Input I
 - the ground truth is one-hot encoded
 - the preprocessed image $\bigotimes$ the one-hot encoded ground truth to construct Input II
 - Input I and II are put into the discriminator, respectively
 - the discriminator is updated by minimizing discriminator loss: $Loss_D$
 - the preprocessed data is put into the segmentor
 - the preprocessed image $\bigotimes$ the segmentation probability map from the segmentor to construct Input III
 - Input III is put into the discriminator
 - the segmentor is updated by minimizing segmentor loss: $Loss_S$

 end for
end for

various variations on the shape of the LV and RV due to different breath hold during MRI acquisition. Such situation needs multi-scale features to eliminate ambiguity and suppress the false prediction. Therefore, multi-scale features are crucial for improving the segmentation performance of short-axis cardiac MRI [24]. In general, the scale of the learned features is determined by the size of the receptive field, while the size of the receptive field is determined by the size of the convolution kernel [25]. However, calculation costs and the number of parameters will be increased dramatically if large convolution kernels are applied to learn large-scale features. Such situation is easy to causes overfitting on a limited dataset [26]. To the best of our knowledge, there are two options to alleviate the above problem. One solution is to stack convolution layers with small convolution kernels to obtain large receptive fields. Another one is to adopt dilated convolution which can expand the receptive fields without increasing too many parameters and calculation costs. The dilated convolution operation is defined as [24]:

$$D(p) = \sum_{s+lt=p} F(s)k(t) \tag{1}$$

where $F : \mathbb{Z}^2 \to \mathbb{R}$ is a discrete function which denotes the input of dilated convolution. $\Omega_r = [-r, r]^2 \bigcap \mathbb{Z}^2$ and $k : \Omega_r \to \mathbb{R}$ represents a convolution kernel whose size is $(2r + 1)^2$. $D(\cdot)$ denotes the output of dilated convolution operation, p is an element of $D(\cdot)$ and l represents the dilation rate of convolution.



(a) dilation rate = 1 (b) dilation rate = 2

Fig. 3: Examples of dilated convolution. Blue region is the receptive field and red dots denote convolution kernel.

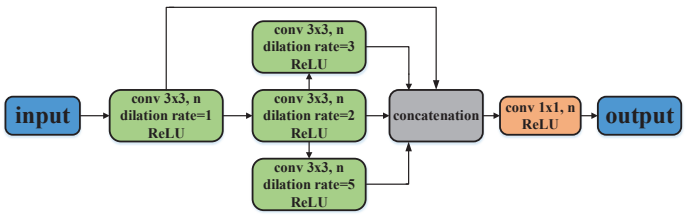


Fig. 4: The structure of DB. n denotes the number of convolutional kernels, which is equal to the number of the output channels. Note that convolutions in DB will not change the size of feature maps.

Examples of dilated convolution operation are presented in Fig. 3. Fig. 3 (a) and Fig. 3 (b) illustrate dilated convolutions with 3×3 kernel at different dilation rates. Fig. 3 (a) is a 1-dilated convolution. Through (a), each element of the output has a 3×3 receptive field. Fig. 3 (b) denotes a 2-dilated convolution. If we perform (b) after (a), the element in the output of (b) has a 7×7 receptive field. To obtain the same size of the receptive field, three stacked convolution layers with 3×3 kernel

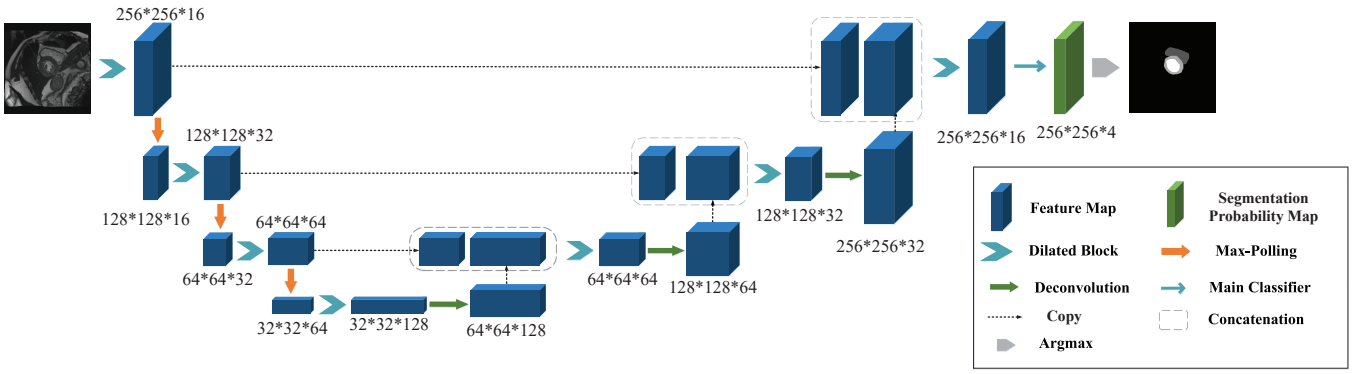


Fig. 5: The structure of the segmentor in DBAN.

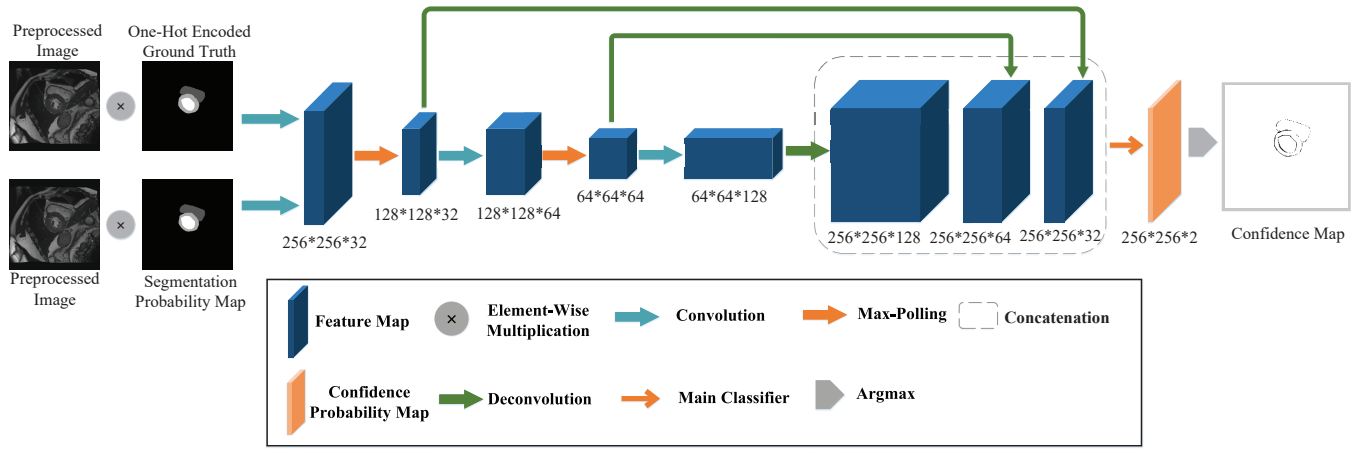


Fig. 6: The structure of the discriminator in DBAN.

are required. Therefore, compared to stacked convolution layers with small convolution kernels, dilated convolution is able to obtain large receptive field through fewer layers. Enlightened by this observation, we design the DB within the segmentor to extract multi-scale features for the LV, RV and MYO segmentation.

The structure of DB is displayed in Fig. 4. In the DB, dilated convolution with 3×3 kernels at different dilation rates is adopted so that multi-scale features can be extracted. However, there is a theoretical issue in the dilated convolution called gridding problem: The sample from the input can be very sparse, which is detrimental to learning because: 1) local information is omitted completely; 2) the information can be irrelevant across large distances. Besides, the consistency of local information may be impaired. To fix this issue, we follow the HDC scheme [27] to stack dilated convolution layers and to set the dilation rates to 1, 2, 3 and 5 respectively. In this way, reception fields with the size of 3×3, 7×7,13×13 and 17×17 can be acquired to capture multi-scale features. Then feature maps produced by dilated convolution layers are concatenated. Multi-scale features in the output of the concatenation layer are aggregated by the 1×1 convolution layer which can be applied for different feature fusion [26]. Compared with the classical convolution operations, the proposed DB is capable of obtaining multi-size receptive fields and capturing multi-scale features through fewer layers. Fewer layers mean fewer parameters, which is significant for biomedical image segmentation with very limited dataset.

As shown in Fig. 5, Dilated Block U-Net (DB U-Net) is designed as the segmentor of DBAN. As for the overall architecture, the U-Net scheme is followed. We replace the convolution layers with the seven DBs to extract and fuse multi-scale featrues. In the DB, n is equal to the number of output channels. For example, n in the first DB is equal to 16. Besides, three down-sampling layers (max-pooling) and three up-sampling layers (deconvolution) are adopted, which is diffierent from U-Net. The main reason is that U-Net is designed for microscopy images with the size of 512×512, while the size of cardiac MRI is only half of the microscopy images (256×256). To obtain the segmentation results, we apply 1×1 convolution layer and softmax function as the main classifier to refine the aggregated features and produce segmentation probability map. In the end, segmentation results can be viewed through argmax function.

2) Discriminator: For the training of adversarial networks, the discriminator input is crucial. In this work, we adopt element-wise multiplication (EM) to construct the discriminator input. EM is the pixel-level multiplication between the preprocessed image and the segmentation

probability map or the ground truth [28]. To make the number of ground truth channel consistent with that of segmentation probability map channel, we implement the one-hot encoding on the ground truth before performing EM. The advantage of this method is that the segmentation probability map and the preprocessed image can be fused at the initial stage. Therefore, the information from both the segmentation probability map and the preprocessed image can be involved in the discriminator's decision making.

Fig. 6 displays the structure of the discriminator. In the previous studies, discriminator is usually used to distinguish whether the whole image is from the segmentor or from the ground truth [18] [19] [20] [21] [22]. In this work, we apply the fully convolutional scheme for the discriminator to differentiate the segmentation probability map from the ground truth at the pixel level. Compared to the previous image-level classification scheme, the fully convolutional scheme can make the discriminator implement identification in greater detail (at the pixel level), which is beneficial to the segmentor parameter update. Due to the limited dataset, only a few hidden layers are applied in the discriminator where three convolution layers with 3×3 kernels are employed. ReLU is used as activation function. In addition, two max-polling layers and three deconvolution layers are adopted as down-sampling layers and up-sampling layers, respectively. Concatenation operation is implemented to combine feature maps. In the output layer, 1×1 convolution operation as well as softmax function is used as the main classifier to identify the input pixels as two classes: segmentation results or ground truth. The output of the main classifier is the confidence probability map which can help the segmentor modify the segmentation probability map through back propagation. Confidence map can be viewed by the argmax function as the visualization production of the confidence probability map. Each pixel p of the confidence map represents that the discriminator identifies whether the corresponding pixel in the input is from the ground truth ($p = 1$) or from the segmentor ($p = 0$).

B. Loss Function

During the training stage of DBAN, segmentor and the discriminator are trained respectively. Therefore, different loss functions are required.

1) Discriminator Loss: To train the discriminator, the loss function $Loss_D$ is defined as follow:

$$\begin{aligned} \text{Loss}_D = \sum_{m=1}^{M} [&L_b(D(S(X_m) \otimes X_m), Y_0) \\ &+ L_b(D(Y_m \otimes X_m), Y_1)] \end{aligned} \tag{2}$$

where M denotes the number of preprocessed images, L_b represents binary-class cross-entropy loss, X_m is the preprocessed image and Y_m denotes the one-hot encoded ground truth. $S(\cdot)$ is the segmentation probability map and $D(\cdot)$ represents the pixel-wise identification result. $\otimes$ denotes element-wise multiplication. Y_0 is the pixel-wise label of the discriminator input which is from segmentation probability map and Y_1 is the pixel-wise label of the input which is from the ground truth.

In the Equation (2):

$$\begin{aligned} L_b(D(S(X_m) \otimes X_m), Y_0) = -\sum_{w,h} [&y_0 log(D_{seg}^{(w,h)}) \\ &+ (1 - y_0) log(1 - D_{seg}^{(w,h)})] \end{aligned} \tag{3}$$

where $y_0 = 0$ is the pixel-wise label, indicating the pixel of the input at location (w, h) is from the segmentation probability map. $D_{seg}^{(w,h)}$ denotes the identification result (confidence probability from the confidence probability map) of the segmentor output at location (w, h). Therefore, Equation (3) can be written as:

$$L_b(D(S(X_m) \otimes X_m), Y_0) = -\sum_{w,h} log(1 - D_{seg}^{(w,h)}) \tag{4}$$

Similarly, $L_b(D(Y_m \otimes X_m), Y_1)$ can be written as:

$$\begin{aligned} L_b(D(Y_m \otimes X_m), Y_1) &= -\sum_{w,h} [y_1 log(D_{gt}^{(w,h)}) \\ &\quad + (1 - y_1) log(1 - D_{gt}^{(w,h)})] \\ &= -\sum_{w,h} log(D_{gt}^{(w,h)}) \end{aligned} \tag{5}$$

where y_1 is the pixel-wise label and $y_1 = 1$ denotes the pixel of the input at location (w, h) is from the ground truth. $D_{gt}^{(w,h)}$ denotes the identification result of the ground truth at location (w, h).

As a summary, discriminator loss function can be written as:

$$\begin{aligned} Loss_D = -\sum_{m=1}^{M} \sum_{w,h} [&log(D_{gt}^{(w,h)}) \\ &+ log(1 - D_{seg}^{(w,h)})] \end{aligned} \tag{6}$$

By training the discriminator, this discriminator loss function will become smaller and smaller, indicating the gradually improved discrimination ability.

2) Segmentor Loss: the segmentor is trained by minimizing the segmentor loss $Loss_S$:

$$\begin{aligned} \text{Loss}_S = \sum_{m=1}^{M} [&L_m(S(X_m), Y_m) \\ &+ \lambda L_b(D(S(X_m) \otimes X_m), Y_1)] \end{aligned} \tag{7}$$

where L_m is multi-class cross-entropy loss. λ is a hyper-parameter.

In the Equation (7):

$$\begin{aligned} &L_m(S(X_m) \otimes Y_m) \\ &= -\sum_{w,h} \sum_{c \in C} Y_m^{(w,h,c)} log(S(X_m)^{(w,h,c)}) \end{aligned} \tag{8}$$

where C is the class number. $Y_m^{(w,h,c)}$ is the pixel-wise label from the one-hot encoded ground truth at the location (w, h, c) and $S(X_m)^{(w,h,c)}$ is the pixel of the segmentation probability map at the location (w, h, c).

$$L_b(D(S(X_m) \otimes X_m), Y_1) = -\sum_{w,h} log(D_{seg}^{(w,h)}) \tag{9}$$

In $L_b(D(S(X_m)\otimes X_m), Y_1)$, the pixel-wise label is 1 so that the segmentor can be guided to fool the discriminator.

Therefore, the $Loss_S$ can be written as:

$$Loss_S = -\sum_{m=1}^{M}[\sum_{w,h}\sum_{c\in C} Y_m^{(w,h,c)} log(S(X_m)^{(w,h,c)}) + \lambda \sum_{w,h} log(D_{seg}^{(w,h)})] \quad (10)$$

By minimizing this loss function, segmentor can produce the segmentation results whose quality is good enough to fool the discriminator. Hyperparameter λ is used for balancing the adversarial training, which will be further discussed in the next section.

IV. Experiment and Analysis

A. Datasets

In this work, ACDC dataset acquired from two MRI scanners (1.5T and 3.0T, University Hospital of Dijon, France) was used [4]. This dataset consists of 150 subjects' short-axis cardiac MRIs (3D data) which are categorized into 5 groups: normal case, heart failure with infarction, dilated cardiomyopathy, hypertrophic cardiomyopathy and abnormal RV. In ACDC dataset, 100 subjects' cardiac MRIs (ACDC training dataset) are provided with the corresponding ground truth which cover both the end-diastolic (ED) phase and the end-systolic (ES) phase. The remaining 50 subjects (ACDC testing dataset) without the ground truth are for testing. Challenge participants can upload the testing results to the ACDC evaluation platform for assessment.

B. Preprocessing

1) Slicing: The slice-thickness (5 to 10 mm) in 3D data is too large, which results in insufficient connectivity information between adjacent slices [29]. Considering the above, we converted the 3D MRI data into 2D images by slicing them.

2) Resizing: In deep learning, batch training allows a smaller memory footprint, which can also be employed to improve machine throughput [30]. However, batch training requires a uniform size of input data. In this work, 2D short-axis cardiac MRIs have a wide range of sizes from 154×224 to 428×512. Considering that large size of images might limit the batchsize, we resized all images to 256×256. For images whose width or height was smaller than 256, we padded residual regions with minimum grey value of each image. As for images whose width or height was larger than 256, we cropped it to 256.

3) Data Augmentation: To avoid overfitting, rotation transformation from 0° to 120° with unique interval of 15° was implemented to augment these resized 2D images.

4) Data Normalization: Cardiac MR images in ACDC dataset have a wide range of voxel intensity due to different scanner types or acquisition protocols, which might impede the segmentation performance. Hence, we adoptted the Z-Score method to normalize the voxel intensity of MRIs.

C. Implementation Details

We trained, validated and tested DBAN on the ACDC dataset. Among 100 subjects (ACDC training dataset) with the ground truth, 70 of them were used as training set and the remaining 30 subjects were validation set. As for the 50 subjects (ACDC testing dataset) without the ground truth, we used them as the testing set. The results based on testing set had been uploaded to the ACDC evaluation platform to test our model. The proposed approach was trained on a Nvidia GTX 1080Ti using tensorflow framework. In order to train DBAN, Adam optimizer was adopted where the learning rate was set to 10^{-4} and the drop-out rate was 0.1. During the training stage, the segmentor and the discriminator were trained alternately. When DBAN was trained 600 epochs, the pixel classification accuracy of the discriminator on the training set fluctuated at 0.5, indicating the segemntor could fool the discriminator and the training was completed.

D. Exploration of λ

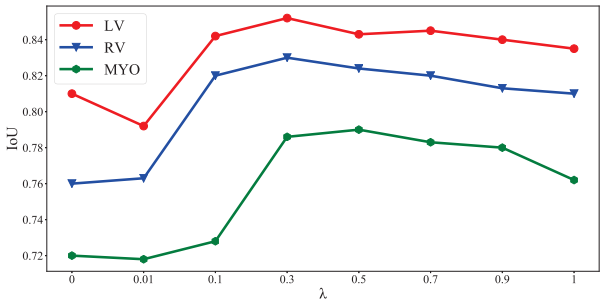


Fig. 7: The IoU for different λ in the proposed DBAN.

In the proposed DBAN, hyperparameter λ is used to balance the adversarial training. λ was set to 0, 0.01, 0.1, 0.3, 0.5, 0.7, 0.9 and 1 respectively to compare the performance of DBAN on the validation set. Intersection over union (IoU) was adopted to evaluate the segmentation results. As shown in the Fig. 7, different λ causes different results in terms of LV, RV and MYO segmentation. Overall, the proposed DBAN achieves the best performance when $\lambda = 0.3$.

E. Experiment Result and Analysis

1) Distance error metrics: Following the ACDC rules, we adopted mean dice similarity coefficient (DSC) and mean Hausdorff Distance (HD) to evaluate the segmentation performance. The formulas of these two metrics are defined as follows:

$$DSC = \frac{2\,|S_R \cap S_{GT}|}{|S_R| + |S_{GT}|} \quad (11)$$

where S_R represents the segmentation result and S_{GT} stands for the ground truth. This evaluation metrics mainly assesses the ratio of pixel overlap.

$$HD = MAX(MAX_{X\subset C_R} MIN_{X\subset C_{GT}} d(x,y), MAX_{X\subset C_{GT}} MIN_{X\subset C_R} d(x,y)) \quad (12)$$

TABLE II: Ablation experiments to explore the function of DBAN components.

Methods	LV				RV				MYO			
	DSC		HD		DSC		HD		DSC		HD	
	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES
FCN [10]	0.88	0.85	9.7	13.4	0.84	0.82	14.4	15.1	0.81	0.80	14.4	13.9
FCN+CD	0.89	0.86	10.1	13.7	0.85	0.83	13.9	14.3	0.81	0.82	13.3	12.8
FCN+our discriminator	0.92	0.86	9.5	11.9	0.88	0.86	12.5	13.7	0.85	0.84	11.9	11.5
U-Net [11]	0.91	0.87	9.2	12.6	0.86	0.85	13.9	13.5	0.82	0.85	11.0	12.7
U-Net+CD	0.92	0.85	8.9	12.1	0.88	0.86	13.1	13.4	0.83	0.85	10.6	10.8
U-Net+our discriminator	0.94	**0.88**	8.6	11.0	0.89	0.88	12.2	13.3	0.85	**0.89**	9.8	9.3
DB U-Net (our segmentor)	0.93	0.86	8.5	13.2	0.88	0.85	11.3	14.2	0.83	0.84	10.4	10.2
DB U-Net+CD	0.94	0.86	8.9	11.1	0.87	0.86	11.9	13.3	0.85	0.88	10.6	9.4
DBAN	**0.96**	**0.88**	**7.6**	**10.8**	**0.92**	**0.90**	**10.3**	**12.7**	**0.87**	**0.89**	**8.2**	**8.8**

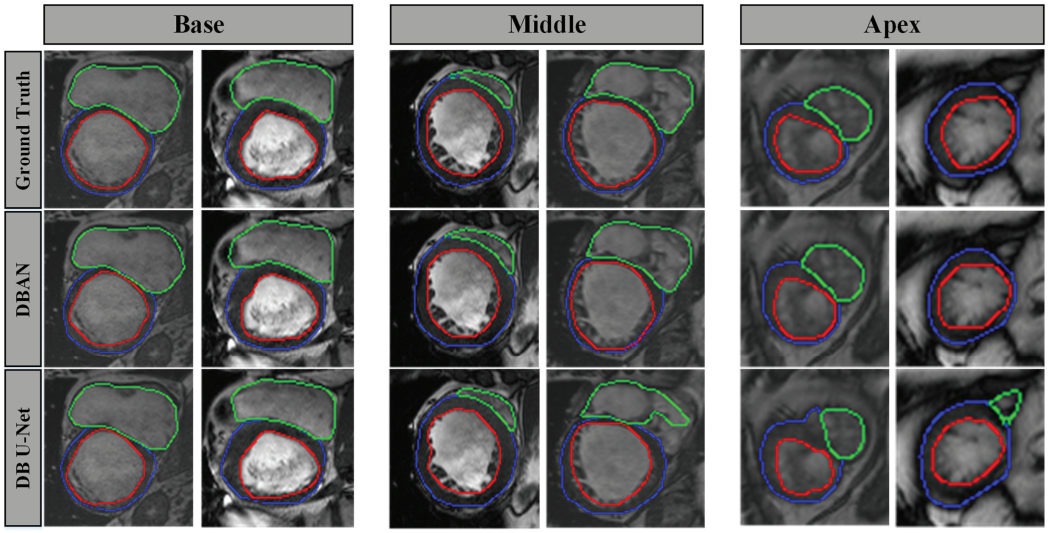


Fig. 8: Visual comparison between DBAN and DB U-Net (our segmentor), which illustrates the effectiveness of the proposed discriminator.

where C_R means the contour of segmentation result, C_{GT} is the contour of ground truth, and $d(\cdot,\cdot)$ denotes the Euclidean Distance between two points. Hausdorff Distance is a measure to evaluate difference between two sets of points. The smaller value of HD demonstrates higher similarity between the segmentation results and the ground truth.

The first step is to investigate the function of DBAN components. Therefore, we conducted ablation studies. FCN, U-Net and our segmentor (DB U-Net) were trained based on the training set and their models were run on the validation set. As indicated in TABLE II, DB U-Net is superior to FCN and U-Net in terms of 7 metrics out of 12 metrics in total, demonstrating the effectiveness of our segmentor for short-axis cardiac MRI segmentation.

Then the proposed discriminator was combined with FCN, U-Net and our segmentor (DB U-Net) respectively so that FCN+our discriminator, U-NET+our discriminator and DBAN were built. Compared with FCN, U-Net and DB U-Net, results of FCN+our discriminator, U-Net+our discriminator and DBAN are improved, respectively. This process means that our discriminator can improve the segmentation performance in terms of LV, RV and MYO. Additationly, Fig. 8 illustrates the visual comparison between DBAN and DB U-Net with respect to apex, middle and base of the heart. Visual comparison reveals that DBAN produces less false positives and false negatives at the edge of target regions. Thanks to the confidence probability map from the discriminator, the segmentor can be guided to produce better segmentation results.

Finally, we conducted experiments to compare the effect between the proposed discriminator and the other common discriminator. In adversarial networks, CNN-based discriminator is the most common architecture to distinguish whether the whole input is from the segmentor or from the ground truth (at the image leavel) [18] [19] [20] [21] [22]. We call it convolution discriminator (CD). For comparison, we replaced our discriminator with CD to construct FCN+CD, U-Net+CD and DB U-Net+CD architectures, respectively. In this work, three convolution layers with the 3×3 kernels, two max-pooling layers, three fully connected layers and a softmax layer were adopted in the CD. Note that the hidden layer number of the

TABLE III: Comparison (based on ACDC testing dataset) between DBAN and the top 10 ACDC methods in terms of distance error metrics.

Methods	LV				RV				MYO			
	DSC		HD		DSC		HD		DSC		HD	
	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES
DBAN	**0.96**	**0.90**	**6.7**	**8.1**	**0.94**	**0.89**	**10.6**	**12.6**	**0.85**	**0.88**	**8.8**	**8.7**
Isensee [31]	**0.97**	**0.93**	**5.6**	**6.3**	**0.95**	**0.90**	**8.8**	**11.4**	**0.90**	**0.92**	**7.6**	**7.1**
Zotti [32]	0.96	0.91	6.2	8.4	0.93	0.89	11.0	12.6	0.89	0.90	9.6	9.3
Painchaud [33]	0.96	0.91	6.1	8.3	0.93	0.88	13.7	13.3	0.88	0.90	8.6	9.6
Khened [34]	0.96	0.91	8.1	9.0	0.93	0.88	14.0	13.9	0.89	0.90	9.8	12.6
Baumgartner [35]	0.96	0.91	6.5	9.2	0.93	0.88	12.7	14.7	0.89	0.90	8.7	10.6
Wolterink [36]	0.96	0.92	7.5	9.6	0.93	0.87	11.9	13.4	0.87	0.89	11.1	10.7
Rohé [37]	0.96	0.90	7.5	10.7	0.92	0.84	14.0	15.9	0.87	0.87	11.5	13.0
Zotti [38]	0.96	0.90	6.6	8.7	0.94	0.88	10.3	14.0	0.88	0.90	8.7	9.3
Jain [39]	0.95	0.89	8.2	10.9	0.91	0.82	13.5	18.7	0.88	0.90	9.8	11.3
Grinias [40]	0.95	0.85	8.9	12.9	0.89	0.77	19.0	24.2	0.80	0.78	12.3	14.6

TABLE IV: The contents in this table present the ranking of segmentation results produced by DBAN in the ACDC leaderboard. For example, the corresponding value of LV and DSC_ED in the table is 2, indicating that LV segmentation in the ED phase rank the second in terms of DSC.

	DSC_ED	DSC_ES	HD_ED	HD_ES
LV	**2**	7	6	**2**
RV	**2**	**2**	**3**	**2**
MYO	10	9	5	**2**

TABLE V: Generalization performance of our model on the ACDC dataset and the RVSC dataset.

Datasets	DSC		HD	
	ED	ES	ED	ES
ACDC Dataset	0.94	0.89	10.6	12.6
RVSC Dataset	0.89	0.85	12.2	13.9

TABLE VI: Generalization performance of our model on the ACDC dataset and the Sunnybrooke dataset.

Datasets	DSC		HD	
	ED	ES	ED	ES
ACDC Dataset	0.96	0.90	6.7	8.1
Sunnybrooke Dataset	0.93	0.89	9.1	10.9

CD is the same as that of the proposed discriminator. As shown in TABLE II, our discriminator brings forth more improvements than CD because of the powerful discrimination ability of the fully convolutional scheme adopted in our discriminator. The minor differences between the segmentation probability map and the ground truth can be identified in greater detail (at the pixel level), beneficial to the segmentor parameter update. Therefore, the results appear more close to the ground truth. To sum up, the above ablation experiments demonstrate that the proposed segmentor and discriminator are effective and DBAN outperforms other methods in TABLE II.

The second step is to test the performance of the proposed model. We submitted the segmentation results on the ACDC testing set to the ACDC evaluation platform, as shown in TABLE III. In addition, the performance of top ten solutions in ACDC is presented in the table for comparison. Note that there are 7 U-Net based methods in the top ten of ACDC leaderboard. For each metric, DBAN can be among the top 10 of the ACDC leaderboard. As shown in TABLE IV, segmentation results of DBAN are among top 3 in terms of 7 metrics out of 12 metrics in total. Also, the difference between DBAN and the first position (Isensee's methods [31]) is no more than 0.05 and 1.8mm in terms of DSC and HD, respectively. Therefore, we believe that DBAN has reached the state-of-the-art level for LV, RV and MYO segmentation.

The final step is to test the generalization ability of the proposed model. We ran it (without finetuning) on the training dataset of the Right Ventricle Segmentation Challenge (RVSC) dataset [42] and the Sunnybrooke dataset [43]. Note that Sunnybrooke dataset is a short-axis cardiac MRI dataset for LV segmentation and RVSC dataset is a short-axis cardiac MRI dataset for RV segmentation. TABLE V and TABLE VI indicate that the results based on these two datasets are close to those on the ACDC testing dataset, thus demonstrating the good generalization ability of DBAN.

2) Clinical indice metrics: It is important to test whether DBAN can be employed in the clinical applications or not. In the clinical practice, end diastolic volume (EDV), end systolic volume (ESV), myocardium mass (MM) in ED phase and ejection fraction (EF) are wildly used to analyize the cardiac function. In the ACDC leaderboard, correlation coefficient (CC), bias and standard deviation (SD) are calculated as the clinical indice metrics to analyse left ventricle EF, left ventricle EDV, right ventricle EF, right ventricle EDV, MM in ED phase and myocardium ESV. TABLE VII presents the comparison between our method and the top eleven methods from

TABLE VII: Comparison (based on ACDC testing dataset) between DBAN and the top 11 ACDC methods in terms of clinical indice metrics.

Methods	LV						RV						MYO					
	EF			EDV			EF			EDV			ESV			MM_ED		
	CC	bias	SD	CC	bias	SD	CC	bias	SD	CC	bias	SD	CC	bias	SD	CC	bias	SD
DBAN	**0.991**	**−0.9**	**3.0**	**0.994**	**−6.0**	**8.6**	**0.890**	**0.1**	**6.6**	**0.988**	**−14.1**	**11.2**	**0.973**	**0.6**	**13.3**	**0.980**	**6.8**	**10.8**
Isensee [31]	**0.992**	0.3	**2.8**	**0.997**	1.6	5.8	**0.925**	-3.0	**5.1**	**0.991**	**2.1**	**9.0**	**0.989**	-3.4	**7.9**	0.986	-4.0	8.6
Zotti [32]	0.990	-0.5	3.1	0.997	3.7	**5.1**	0.869	-0.9	6.8	0.986	2.4	11.5	0.980	1.2	10.9	0.986	-1.8	8.6
Painchaud [33]	0.990	-0.5	3.2	0.997	3.8	5.2	0.865	-0.9	6.9	0.986	2.1	11.4	0.979	0.3	11.1	0.987	-2.9	8.4
Khened [34]	0.989	-0.5	3.4	0.997	**0.6**	5.5	0.858	-2.2	6.9	0.982	-2.9	12.6	0.979	-2.6	11.0	**0.990**	-2.9	**7.5**
Baumgartner [35]	0.988	0.6	3.4	0.995	1.4	7.6	0.851	1.2	7.3	0.977	-2.3	15.1	0.983	-9.6	9.9	0.982	-6.9	9.8
Wolterink [36]	0.988	-0.5	3.4	0.993	3.0	8.7	0.852	-4.6	6.9	0.980	3.6	15.2	0.971	**0.9**	13.4	0.963	**−1.0**	14.6
Rohé [37]	0.989	**−0.1**	3.2	0.993	4.2	8.6	0.781	-0.7	9.9	0.983	7.3	13.4	0.955	5.1	16.8	0.967	-3.4	13.3
Zotti [38]	0.987	-1.2	3.6	0.997	9.6	6.4	0.872	-2.2	6.8	0.991	-3.7	9.2	0.960	-7.8	15.2	0.984	-12.4	9.0
Jain [39]	0.971	1.7	5.5	0.997	9.9	6.7	0.791	6.8	8.1	0.945	5.6	22.2	0.986	-4.5	9.1	0.989	-11.6	8.1
Grinias [40]	0.970	-1.7	5.5	0.992	2.4	11.1	0.756	**−0.2**	9.7	0.916	11.9	27.8	0.890	-1.7	27.6	0.950	-19.6	21.3
Yang [41]	0.926	1.5	8.7	0.894	12.2	32.0	0.576	8.8	23.2	0.789	47.3	41.9	N/A	N/A	N/A	N/A	N/A	N/A

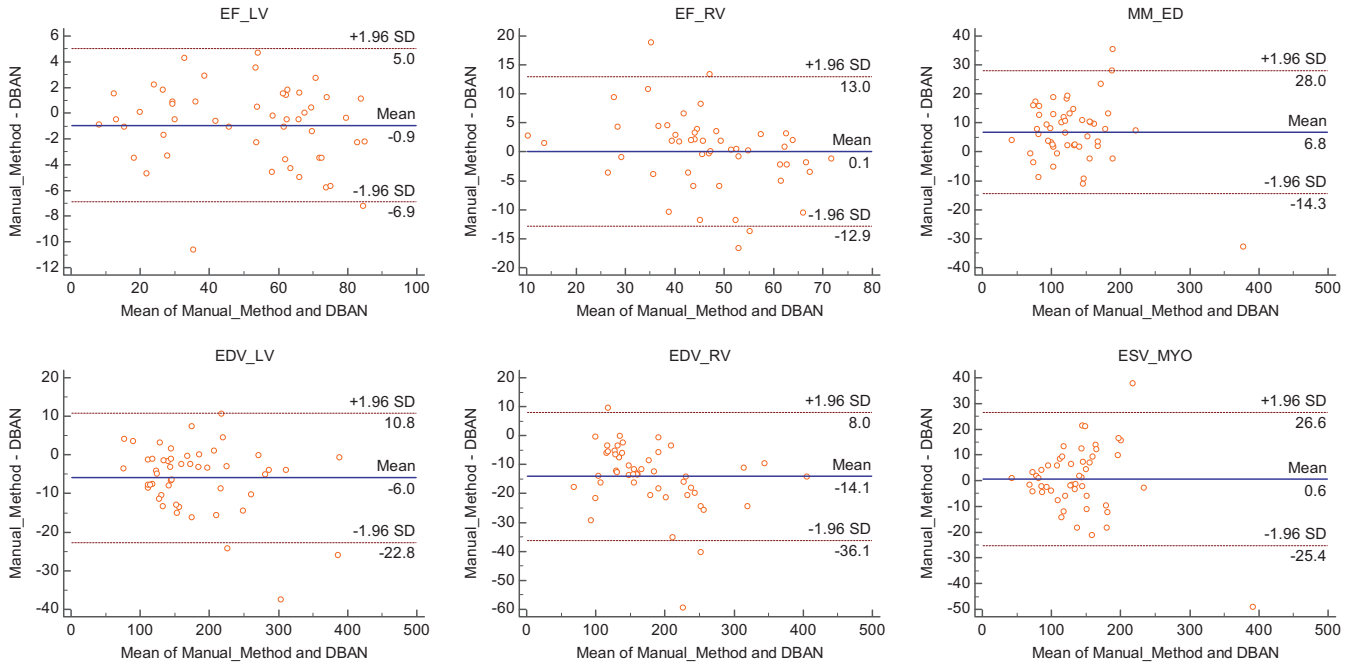


Fig. 9: Bland-Altman plots for identifying the difference between DBAN and clinical experts (mannual methods) in terms of cardiac function indices. Note that EF_LV represents the left ventricular ejection fraction, EDV_LV represents left ventricular volume in the ED phase and MM_ED indicates the myocardium mass in the ED phase. Other metrics are named similarly.

ACDC in terms of the clinical indice metrics. As can be seen from TABLE VII, all DBAN results are among the top eleven. In particular, bias of right ventricle EF and myocardium ESV rank the first, respectively. Both correlation coefficient and standard deviation computed from left/right ventricle EF get the second position. Correlation coefficient and standard deviation of right ventricle EDV rank the third. Therefore, we believe that DBAN has potential practical value.

Then in order to assess the difference and similarity between DBAN and manual delineation method of clinical experts, we performed Bland-Altman plots [44], linear regression and correlation analysis based on the results from ACDC evaluation platform (totally 50 subjects).

As shown in Fig. 9, the differences between DBAN and clinical experts are presented in Bland-Altman plots. The differences within 95% limits of agreement (mean ± 1.96 standard deviation) are not clinically important and the two methods could replace each other [44]. Among 50 patients, a maximum of 48 patients' cardiac function indices (EF_LV, MM_ED and ESV_MYO) and a minimum of 46 patients' cardiac function indices (EF_RV) from DBAN segmentation results are within 95% limits of agreement (i.e. they are acceptable for clinical application).

Fig. 10 shows the P-value, correlation coefficient r and linear regression. In the Fig. 10, all P-value are less than 0.001, which demonstrates that DBAN has a very significant correlation with manual method. The closer the correlation coefficient is to 1, the higher the correlation between two methods is. Almost all the correlation

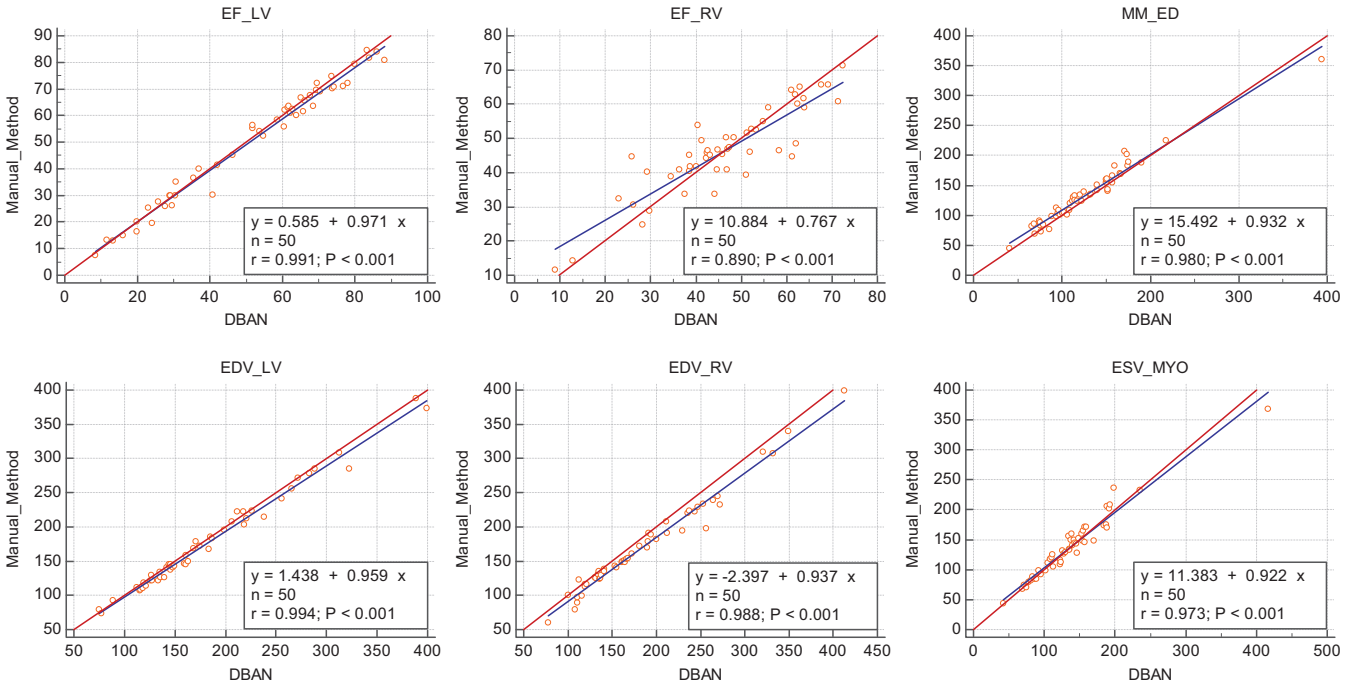


Fig. 10: P-value, correlation coefficient and linear regression for analyzing the correlation between DBAN and clinical experts (mannual methods) in terms of cardiac function indices.

coefficients in the Fig. 10 are very closed to 1 except for that of the right ventricular EF (0.890). However, as shown in the TABLE VII, CC of the right ventricular EF from DBAN ranks the second in ACDC leaderboard. It is due to that the calculation of right ventricle EF is based on RV segmentation in ED phase and ES phase, while the change of RV shape and surrounding unrelated tissues may render a poor segmentation performance. We believe that these errors can be alleviated by adding prior knowledge and increasing the diversity of data. In linear regression plots, the red lines are regression lines which are close to the blue lines (identity function: y = x). Overall, the above correlation analysis demonstrate that DBAN behaves similarly to clinical experts in terms of cardiac function indices.

V. Conclusion

In this paper, we propose DBAN, an automatic segmentation approach, for short-axis cardiac MRI segmmentation. DBAN contains two main sections: a segmentor and a discriminator. In the segmentor, DB is proposed to capture and aggregate multi-scale features. The segmentor can produce segmentation probability maps, while the discriminator adopting fully convolutional scheme can distinguish whether the input is from the segmentation probability map or from the ground truth at the pixel level. Moreover, confidence probability map produced by our discriminator can guide the segmentor to produce more accurate segmentation results. On the ACDC dataset, the proposed DBAN outperforms majority of approaches and reaches the state-of-the-art level in terms of DSC and HD. In additation, DBAN behaves similarly to clinical experts in terms of cardiac function indices. Therefore, DBAN is a promising method for LV, RV and MYO segmentation in the clinical application. In future work, we plan to cooperate with a local hospital to obtain more short-axis cardiac MRIs with different modalities so that DBAN can gain more comprehensive training to handle various situations in clinic. Then, the interpretability issue of the neural network will be explored for quantitative explanation of the DBAN prediction structure. In this way, DBAN segmentation results can be interpreted quantitatively to be more aligned with clinical demand.

Acknowledgment

The authors acknowledge supports from National Natural Science Foundation of China (81427803, 81771940), National Key Research and Development Program of China (2017YFC0108000), Beijing National Science Foundation (7172122, L172003), and Introduced Talent Program of Southwest University under Grant SWU020008.

References

[1] W. H. Organization, “Cardiovascular diseases (cvds) - key facts,” [Online], https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), [Accessed: March 20, 2020].

[2] D. Ripley, T. Musa, L. Dobson, S. Plein, and J. Greenwood, “Cardiovascular magnetic resonance imaging: what the general cardiologist should know,” Heart, vol. 102, no. 19, pp. 1589–1603, 2016.

[3] D. A. Auger, X. Zhong, F. H. Epstein, E. M. Meintjes, and B. S. Spottiswoode, “Semi-automated left ventricular segmentation based on a guide point model approach for 3d cine dense cardiovascular magnetic resonance,” Journal of Cardiovascular Magnetic Resonance, vol. 16, no. 1, p. 8, 2014.

[4] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P.-A. Heng, I. Cetin, K. Lekadir, O. Camara, M. A. G. Ballester et al., "Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved?" IEEE transactions on medical imaging, vol. 37, no. 11, pp. 2514–2525, 2018.

[5] N. C. Codella, J. W. Weinsaft, M. D. Cham, M. Janik, M. R. Prince, and Y. Wang, "Left ventricle: automated segmentation by using myocardial effusion threshold reduction and intravoxel computation at mr imaging," Radiology, vol. 248, no. 3, pp. 1004–1012, 2008.

[6] C. Pluempitiwiriyawej, J. M. Moura, Y.-J. L. Wu, and C. Ho, "Stacs: New active contour scheme for cardiac mr image segmentation," IEEE transactions on medical imaging, vol. 24, no. 5, pp. 593–603, 2005.

[7] Z. Zhang, C. Duan, T. Lin, S. Zhou, Y. Wang, and X. Gao, "Gvfom: a novel external force for active contour based image segmentation," Information Sciences, vol. 506, pp. 1–18, 2020.

[8] S. C. Mitchell, B. P. Lelieveldt, R. J. Van Der Geest, H. G. Bosch, J. Reiver, and M. Sonka, "Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac mr images," IEEE Transactions on medical imaging, vol. 20, no. 5, pp. 415–423, 2001.

[9] M. Lorenzo-Valdés, G. I. Sanchez-Ortiz, A. G. Elkington, R. H. Mohiaddin, and D. Rueckert, "Segmentation of 4d cardiac mr images using a probabilistic atlas and the em algorithm," Medical Image Analysis, vol. 8, no. 3, pp. 255–265, 2004.

[10] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.

[12] T. A. Ngo, Z. Lu, and G. Carneiro, "Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance," Medical image analysis, vol. 35, pp. 159–171, 2017.

[13] W. Wang, Y. Wang, Y. Wu, T. Lin, S. Li, and B. Chen, "Quantification of full left ventricular metrics via deep regression learning with contour-guidance," IEEE Access, vol. 7, pp. 47 918–47 928, 2019.

[14] X. Du, S. Yin, R. Tang, Y. Zhang, and S. Li, "Cardiac-deepied: Automatic pixel-level deep segmentation for cardiac bi-ventricle using improved end-to-end encoder-decoder network," IEEE journal of translational engineering in health and medicine, vol. 7, pp. 1–10, 2019.

[15] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Dilated convolutional neural networks for cardiovascular mr segmentation in congenital heart disease," in 1st International Workshop on Reconstruction and Analysis of Moving Body Organs, RAMBO 2016 and 1st International Workshop on Whole-Heart and Great Vessel Segmentation from 3D Cardiovascular MRI in Congenital Heart Disease, HVSMR 2016. Springer, 2017, pp. 95–102.

[16] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in neural information processing systems, 2014, pp. 2672–2680.

[17] M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784, 2014.

[18] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, "Semantic segmentation using adversarial networks," in NIPS Workshop on Adversarial Training, 2016.

[19] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," Medical image analysis, p. 101552, 2019.

[20] H. Zhang, X. Cao, L. Xu, and L. Qi, "Conditional convolution generative adversarial network for bi-ventricle segmentation in cardiac mr images," in Proceedings of the Third International Symposium on Image Computing and Digital Medicine, 2019, pp. 118–122.

[21] P. Moeskops, M. Veta, M. W. Lafarge, K. A. Eppenhof, and J. P. Pluim, "Adversarial training and dilated convolutions for brain mri segmentation," in Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, 2017, pp. 56–64.

[22] V. K. Singh, H. A. Rashwan, S. Romani, F. Akram, N. Pandey, M. M. K. Sarker, A. Saleh, M. Arenas, M. Arquez, D. Puig et al., "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," Expert Systems with Applications, vol. 139, p. 112855, 2020.

[23] K. Gao, S. Niu, Z. Ji, M. Wu, Q. Chen, R. Xu, S. Yuan, W. Fan, Y. Chen, and J. Dong, "Double-branched and area-constraint fully convolutional networks for automated serous retinal detachment segmentation in sd-oct images," Computer methods and programs in biomedicine, vol. 176, pp. 69–80, 2019.

[24] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," arXiv preprint arXiv:1511.07122, 2015.

[25] W. Luo, Y. Li, R. Urtasun, and R. Zemel, "Understanding the effective receptive field in deep convolutional neural networks," in Advances in neural information processing systems, 2016, pp. 4898–4906.

[26] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[27] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, and G. Cottrell, "Understanding convolution for semantic segmentation," in 2018 IEEE winter conference on applications of computer vision (WACV). IEEE, 2018, pp. 1451–1460.

[28] Y. Zhang, L. Yang, J. Chen, M. Fredericksen, D. P. Hughes, and D. Z. Chen, "Deep adversarial networks for biomedical image segmentation utilizing unannotated images," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2017, pp. 408–416.

[29] Y. Jang, Y. Hong, S. Ha, S. Kim, and H.-J. Chang, "Automatic segmentation of lv and rv in cardiac mri," in International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, 2017, pp. 161–169.

[30] D. Masters and C. Luschi, "Revisiting small batch training for deep neural networks," arXiv preprint arXiv:1804.07612, 2018.

[31] F. Isensee, P. F. Jaeger, P. M. Full, I. Wolf, S. Engelhardt, and K. H. Maier-Hein, "Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features," in International workshop on statistical atlases and computational models of the heart. Springer, 2017, pp. 120–129.

[32] C. Zotti, Z. Luo, A. Lalande, and P. Jodoin, "Convolutional neural network with shape prior applied to cardiac mri segmentation." IEEE journal of biomedical and health informatics, vol. 23, no. 3, pp. 1119–1128, 2019.

[33] N. Painchaud, Y. Skandarani, T. Judge, O. Bernard, A. Lalande, and P.-M. Jodoin, "Cardiac mri segmentation with strong anatomical guarantees," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2019, pp. 632–640.

[34] M. Khened, V. A. Kollerathu, and G. Krishnamurthi, "Fully convolutional multi-scale residual densenets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers," Medical image analysis, vol. 51, pp. 21–45, 2019.

[35] C. F. Baumgartner, L. M. Koch, M. Pollefeys, and E. Konukoglu, "An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation," in International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, 2017, pp. 111–119.

[36] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Automatic segmentation and disease classification using cardiac cine mr images," Statistical Atlases and Computational Models of the Heart, vol. 10663, p. 101, 2018.

[37] M.-M. Rohé, M. Sermesant, and X. Pennec, "Automatic multi-atlas segmentation of myocardium with svf-net," in International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, 2017, pp. 170–177.

[38] C. Zotti, Z. Luo, O. Humbert, A. Lalande, and P.-M. Jodoin, "Gridnet with automatic shape prior registration for automatic mri cardiac segmentation," in International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, 2017, pp. 73–81.

[39] J. Patravali, S. Jain, and S. Chilamkurthy, "2d-3d fully convolutional neural networks for cardiac mr segmentation," in In-

ternational Workshop on Statistical Atlases and Computational Models of the Heart. Springer, 2017, pp. 130–139.
[40] E. Grinias and G. Tziritas, "Fast fully-automatic cardiac segmentation in mri using mrf model optimization, substructures tracking and b-spline smoothing," in International Workshop on Statistical Atlases and Computational Models of the Heart. Springer, 2017, pp. 91–100.
[41] X. Yang, C. Bian, L. Yu, D. Ni, and P.-A. Heng, "Class-balanced deep neural network for automatic ventricular structure segmentation," in International workshop on statistical atlases and computational models of the heart. Springer, 2017, pp. 152–160.
[42] C. Petitjean, M. A. Zuluaga, W. Bai, J.-N. Dacher, D. Grosgeorge, J. Caudron, S. Ruan, I. B. Ayed, M. J. Cardoso, H.-C. Chen et al., "Right ventricle segmentation from cardiac mri: a collation study," Medical image analysis, vol. 19, no. 1, pp. 187–202, 2015.
[43] P. Radau, Y. Lu, K. Connelly, G. Paul, A. Dick, and G. Wright, "Evaluation framework for algorithms segmenting short axis cardiac mri," The MIDAS Journal-Cardiac MR Left Ventricle Segmentation Challenge, vol. 49, 2009.
[44] J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," The lancet, vol. 327, no. 8476, pp. 307–310, 1986.

Xinyu Yang received his B.S. in Communication Engineering in 2015 from University of Jinan, China. Currently, He is with the College of Electronic and Information Engineering, Southwest University, China, and also with the Shandong Provincial Key Laboratory of Network Based Intelligent Computing as a graduate student. His current research interests include biomedical image/signal processing and deep learning.

Yuan Zhang (SM'14) received the M.S. degree in communication systems and the Ph.D. degree in control theory and engineering both from Shandong University, China, in 2003 and 2012, respectively. He is currently a Professor with the College of Electronic Information Engineering, Southwest University, China. Dr. Zhang was a visiting professor with the Computer Science Department, Georgia State University, USA, in 2014. His research interests include wearable sensing for smart health, machine learning for auxiliary diagnosis, and biomedical big data analytics. He is a member of IEEE EMBS Wearable Biomedical Sensors and Systems Technical Committee. He is also an associate editor for IEEE Reviews in Biomedical Engineering, IEEE Open Journal of Engineering in Medicine and Biology and IEEE Access. He has served as Leading Guest Editor for six special issues of IEEE, Elsevier, Springer and InderScience publications. He has served on the technical program committee for numerous international conferences. Dr. Zhang is a Senior Member of ACM.

Dr. Benny Lo is a Senior Lecturer of the Hamlyn Centre and the Department of Surgery and Cancer, Imperial College London. Dr Lo's research mainly focuses on body sensor networks and wearable technologies for health care and well-being. Much of his research has translated for clinical applications, and he has also led a few pioneering clinical trials of novel sensing technologies, such as the clinical validation of a new spectroscopic sensor for blood perfusion monitoring for microvascular surgeries. He is a Senior Member of IEEE, an Associate Editor of the IEEE Journal of Biomedical Health Informatics, Chair of the IEEE EMBS Wearable Biomedical Sensors and Systems Technical Committee, and a member of the IEEE EMBS Standards Committee.

Dongrui Wu (S'05-M'09-SM'14) received the BE degree in automatic control from the University of Science and Technology of China in 2003, the ME degree in electrical engineering from the National University of Singapore in 2005, and the PhD degree in electrical engineering from the University of Southern California in 2009. He is now Professor in the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China, and Deputy Director of the Key Laboratory of Image Processing and Intelligent Control, Ministry of Education. His research interests include affective computing, brain-computer interface, computational intelligence, and machine learning. He was/is an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS (2011-2018), the IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS (2014-), the IEEE COMPUTATIONAL INTELLIGENCE MAGAZINE (2017-), and the IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING (2019-).

Hongen Liao is Professor and Vice Chairman of the Department of Biomedical Engineering, School of Medicine, Tsinghua University, China. He received his Ph.D. degrees in biomedical precision engineering from the University of Tokyo, Tokyo, Japan in 2003. He was a Research Fellow of Japan Society for the Promotion of Science (JSPS). From 2004, he was a faculty member at the Graduate School of Engineering, The University of Tokyo, where he became an Associate Professor in 2007. He has been selected as National "Thousand Talents" Distinguished Professor, National Recruitment Program of Global Experts, China since 2010. He is Director and Founder of the Advanced Theranostics and 3-D Imaging Laboratory, Tsinghua University. Professor Liao's research interests include 3D medical image, image-guided surgery, medical robotics, computer-assisted surgery, and fusion of these innovative healthcare technologies for minimally invasive precision diagnosis and therapy. He has also been involved in long viewing distance autostereoscopic display and 3D visualization. Prof. Liao is an Associate Editor of IEEE Engineering in Medicine and Biology Society Conference, the General Chair, Program Chair and Organization Chair of multiple international conferences including MIAR, MICCAI, and ACCAS. He is a President of Asian Society for Computer Aided Surgery and Secretary of Asian-Pacific Activities Working Group, International Federation for Medical and Biological Engineering.

Dr. Yuan – Ting Zhang (Fellow, IEEE) received the B.S. and M.S. degrees in telecommunication from the Department of Electronics, Shandong University, China, in 1976 and 1981, respectively, and the Ph.D. degree in electrical engineering from the University of New Brunswick, Canada, in 1990. He is the first Head of Division of Biomedical Engineering and Professor of the Department of Electronic Engineering at the Chinese University of Hong Kong (CUHK), Hong Kong. He serves concurrently as the Director of the Key Lab of Health Informatics of the Chinese Academy of Sciences (HICAS), and the founding Director of the CAS-SIAT Institute of Biomedical and Health Engineering. His research interests include wearable medical devices, body sensor networks, neural engineering, etc. Dr. Zhang was previously the Vice-President of the IEEE-EMBS in 2000–2001. He is currently the Editor-in-Chief of the IEEE Reviews in Biomedical Engineering, and was the Editor-in- Chief for IEEE Transactions on Information Technology in Biomedicine and the founding Editor-in-Chief of IEEE Journal of Biomedical and Health Informatics. He is a Fellow of the International Academy of Medical and Biological Engineering and the American Institute of Medical and Biological Engineering.