IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 26, 2020, accepted November 23, 2020, date of publication November 26, 2020,
date of current version December 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3040888

Left Ventricle Segmentation Based on a Dilated Dense Convolutional Networks

SHENGZHOU XU[1,2], SHIYU CHENG[1], XIANGDE MIN[3], NING PAN[4], AND HUAIFEI HU[4]

[1]College of Computer Science, South Central University for Nationalities, Wuhan 430074, China
[2]Hubei Provincial Engineering Research Center for Intelligent Management of Manufacturing Enterprises, Wuhan 430074, China
[3]Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China
[4]College of Biomedical Engineering, South Central University for Nationalities, Wuhan 430074, China

Corresponding author: Shengzhou Xu (xushengzhou@mail.scuec.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61302192, in part by the Natural Science Foundation of Hubei Province under Grant 2020CFB541, and in part by the Fundamental Research Funds for the Central Universities under Grant CZY19011.

ABSTRACT The automatic segmentation of the left ventricle in magnetic resonance (MR) images is the basis of computer-aided diagnosis systems. To accurately extract the endocardium and epicardium of the left ventricle from MR images, a method based on a dilated dense convolutional network (DDCN) has been proposed in this article. First, to reduce memory consumption, computing time and the class imbalance between the target and background, a clustering algorithm that combines the prior knowledge of the spatial relationship between the slices has been proposed to crop the region of interest (ROI). Then, the DDCN model with 8 dilated convolutional layers and dense connections, which is efficient with respect to its memory consumption and training time, has been proposed to delineate the endocardium and epicardium. To compare the DDCN model with other algorithms, 30 sequences of the MICCAI 2009 left ventricle segmentation challenge database are used to train the proposed model and the other 15 sequences are used for testing. The performance of the proposed method is evaluated by the percentage of ''good'' contours (PGC), average Dice metric (ADM) and average perpendicular distance (APD). Our results show that for the endocardial and epicardial contours, the PGCs are 99.49%±1.99% and 100%±0%, the APDs are 1.50±0.34 mm and 1.31±0.22 mm, and the ADMs are 0.93±0.03 and 0.96±0.01, respectively, which indicates that our method provides contours with great agreement with the ground truth. In addition, the comparison results show that our method exhibits outstanding performance and possesses promising potential to be used in computer-aided diagnosis systems for cardiovascular disease.

INDEX TERMS Segmentation, left ventricle, magnetic resonance image, dilated dense convolutional network.

I. INTRODUCTION

Cardiovascular diseases are the leading cause of death in many countries throughout the world [1]. The diagnosis of these pathologies is dependent on cardiac images. Magnetic resonance (MR) can provide high-quality images for the non-invasive assessment of the left ventricle [2]. The automatic segmentation of the left ventricle (LV) in MR images, which consists of delineating the epicardium (the outer wall) and endocardium (the inner wall), is the basis of computer-aided analysis. Generally, there are several tissues in the image, such as the myocardium, epicardium, endocardium, left ventricle (LV) blood pool, and right ventricle (RV) blood pool. Due to the motion of the heart, the rapid flow of blood, the interference of image noise, and boundaries often being obscured, the segmentation of the epicardium and endocardium remains a challenging task [3]. In recent years, many researchers have made extensive research on the segmentation of the LV, and many methods have been proposed. These methods can be categorized into two major categories: traditional segmentation methods and deep learning-based methods. Traditional methods mainly include boundary-based and region-based methods, as well as hybrid methods [4]–[6]. These methods can use the information of adjacent slices to improve the segmentation result; however, the segmentation process depends on artificial features, which usually requires experience and time to adjust these features. These methods usually tend to perform poorly on data originating from a database outside the training data [7].

In recent years, with the advances in convolutional neural networks (CNNs), breakthroughs have been made in the LV segmentation of MR images [8], [9]. H. Abdeltawab *et al.*

The associate editor coordinating the review of this manuscript and approving it for publication was Luca Cassano.

proposed a fully convolutional U-Net that incorporates a novel loss function to get the epicardium and endocardium from MR images [10]. Romaguera *et al.* developed a fully convolution neural network (FCN) similar to U-Net for LV segmentation [11]. They obtained good segmentation results only for the endocardium. Avendi *et al.* proposed an LV segmentation method that combines a deformable model and CNN [12], but only the endocardium was segmented.

To alleviate the class imbalance problem, some methods extracted ROI patches and used them to train FCN models [7], [13]. In some other methods, images are first processed by deep learning and then refined by traditional methods. Ngo *et al.* proposed a novel method that combines deep learning and a level set for the automated segmentation of the LV [14]. The combination of these methods brings together the advantages of both approaches. Lan *et al.* obtained the coarse region of the LV with SegNet, and then extracted the exact contour with a double snake model [15]. However, the division of the training set and the test set was not clearly described.

Deepening the network layers and widening the network structure are the common strategies adopted to improve the performance of the CNN; however, DenseNet [16] did not follow this fixed mindset. DenseNet greatly reduces the number of parameters in the network and alleviates the vanishing gradient problem through feature reuse and bypass setting. However, as the number of DenseNet features increases exponentially with the number of network layers, it consumes a lot of memory and wastes too much time in training. To solve these problems, the combination of the dilated convolution and DenseNet was used to delineate the right ventricle [17]. M. Khened *et al.* proposed a DenseNet-based FCN architecture for cardiac segmentation, which is parameter and memory efficient [18]. They validated the proposed network on three publicly available datasets: ACDC-2017, LV-2011 and the 2015 Kaggle Data Science Bowl cardiac challenge data.

In this article, a novel automated segmentation method has been proposed for delineating the epicardium and endocardium in all slices of the end-diastole (ED) and end-systole (ES) cardiac phases. First, to reduce the memory consumption, computing time and class imbalance between the target and the background, a clustering algorithm that combines the prior knowledge of the spatial relationship between the slices has been proposed to crop the region of interest (ROI). Then, the DDCN model with 8 dilated convolution layers and dense connections, which can further reduce memory consumption and training time, has been proposed to delineate the endocardium and epicardium.

II. MATERIALS AND METHODS

A. DATA SET

The LV images used in this study are from the dataset of the cardiac short-axis cine MR images published by the MICCAI 2009 grand challenge on the Internet (http://sourceforge.net/projects/cardiac-mr/files/) [19]. There are 45 cases in this dataset, which are collected from Canada's Sunnybrook Health Sciences Centre (SHSC). They are divided into three groups with an average of 15 cases in each group, including 4 heart failure with ischemia (HF-I) cases, 4 heart failure without ischemia (HF-NI) cases, 4 hypertrophy (HYP) cases and 3 normal (N) cases. These MR images were obtained during 10- to 15-second breath-holds with a temporal resolution of 20 cardiac phases over the heart cycle. Six to twelve short-axis cine images were obtained from the atrioventricular ring to the apex (matrix=256 $\times$ 256, thickness =8$\sim$10 mm, and FOV=320 mm$\times$320 mm). Expert annotations are provided for the endocardium in all slices in the ED and ES cardiac phases and for the epicardium only in the ED cardiac phase. These annotations are regarded as the ground truths of the segmentation in our study.

B. THE FRAMEWORK OF THE SEGMENTATION ALGORITHM

As shown in Figure 1, the automatic segmentation of the LV (delineating the endocardium and epicardium) in an MR image consists of two steps: ROI cropping and dilated dense convolutional network (DDCN)-based segmentation. The output of the ROI cropping is used as the input of the DDCN. The LV segmentation is a dense classification problem. In addition, the surrounding tissues of the LV significantly dominate the image. The convolutional network might be biased towards the majority class. To alleviate the problem of the class imbalance between the desired object and surrounding tissues, improve the classification (segmentation) accuracy and reduce the computation, the regions of interest (ROIs) including the LV are extracted as the input images of the network. In the following sections, the ROI cropping and the LV segmentation processes are introduced.

C. ROI CROPPING

As shown in Figure 1, the multislice MR image of the left ventricle resembles an inverted pear. The region of the blood pool in the basal slice at the top is large. In contrast, the region of the blood pool in the apical slice at the bottom is small. Furthermore, there is no significant difference between the blood pool and the surrounding tissues. Therefore, the blood pool is difficult to detect without using prior knowledge. Notice that the LV blood pool is located in the middle of the mid-slice [4]. The ROI cropping process starts from the mid-slice image of the cardiac MR image and proceeds to the direction of the basal slice and the apical slice. In this process, we make use of the strong spatial relationship between neighboring slices. The procedure of ROI cropping is shown in Figure 2.

It is found that the position of the blood pool in the mid-slice is generally located near the center of the slice. Therefore, we first automatically specify an initial ROI with a size of 110 $\times$ 110 pixels at the center of the mid-slice. Then, using the k-means clustering method, the image intensities in the initial ROI are classified into three classes by two thresholds [4]. In addition, the pixels with gray values above the higher threshold are set to be the candidate pixels of

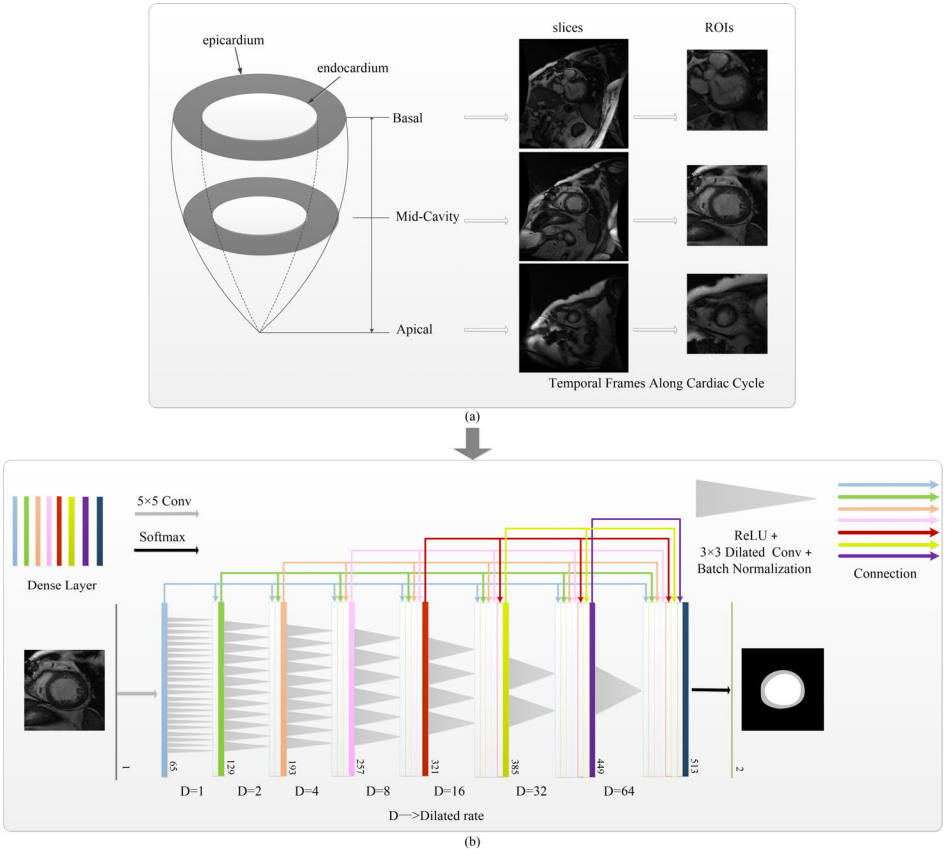

FIGURE 1. **The overall flow chart of LV automatic segmentation (For the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).**

the blood pool. To further optimize the location of the ROI, a binary mask with the same size as the ROI is specified for each slice. For the mid-slice, the mask is a circle with a radius of 10 pixels at the center of the slice. If the gray level of the center point of the ROI is lower than its mean value, then a new center point with a gray level higher than the mean value is searched up or right. For the other slices, the binary mask is the blood pool area of the previous slice. For each slice, the overlap rate between the detected blood pool area and its corresponding mask is measured. If the overlap rate is less than 0.2, the clustering threshold is decreased to obtain a larger blood pool area until the overlap rate reaches 0.2. For each slice, after the area of the blood pool is determined, the center of the blood pool area is taken as the center point of the ROI to obtain the ROI of the slice.

D. LV SEGMENTATION

As shown in Figure 1, The input of the network is ROIs which including the LV. The architecture of the proposed model for LV segmentation is made up of eight dense layers, which are represented by the colored rectangles. Each layer concatenates the feature maps of all previous layers and its own feature maps as the input of all subsequent layers (the colored arrows indicate the concatenations between layers), which greatly alleviates the vanishing gradient problem and strengthens the feature propagation. The horizontal gray arrow at the beginning of the network represents a 5×5 convolution. In addition, the horizontal black arrow at the end indicates the Softmax operation. The gray triangles between layers in the middle part of the network represent the combination of batch normalization (BN), a rectified linear unit (ReLU), and a 3×3 dilated convolution. The width of the gray triangle (variable D in Figure 1) represents the dilated rate of the dilated convolution in each layer. The number in the bottom right corner of the colored rectangle represents the total number of feature maps of the current layer, which is sum of the feature maps of the current layer and all previous layers. For example, the total number of feature maps for the first dense layer is 1+64=65, that for the second dense layer is 1+64+64=129, and so on. As the layers continue to deepen, the dilated rate correspondingly doubles, but the size of the feature map in the entire network remains unchanged at 110×110 pixels.

The dilated convolution is a special form of the traditional convolution in which the effective receptive field of kernels is increased by inserting zeros (or holes) between each pixel in the convolutional kernels [20]. The difference between the dilated convolution used in the DDCN and the traditional convolution is shown in Figure 3. From left to right are the 3×3 traditional convolution, the 3×3 convolution with a

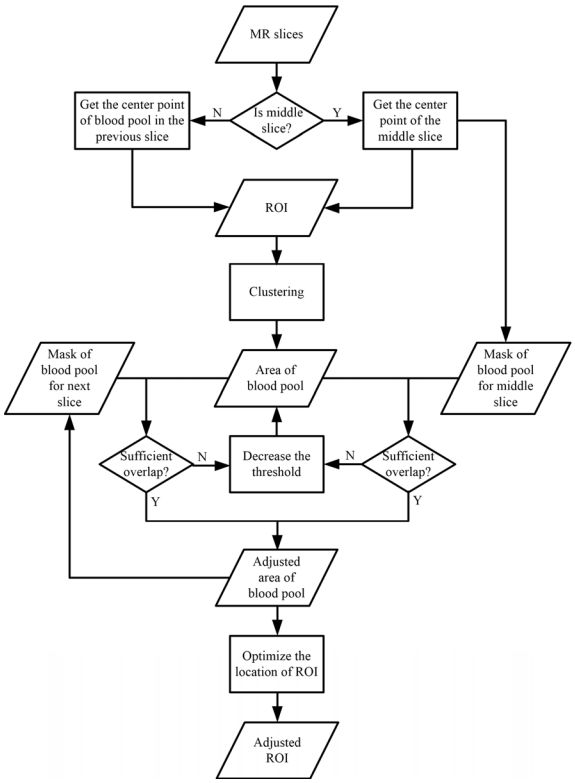

FIGURE 2. The flowchart of ROI extraction.

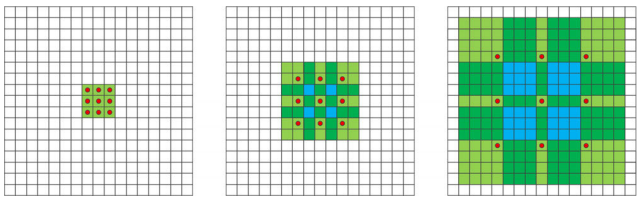

FIGURE 3. Comparison between the dilated convolution and traditional convolution. (a) 3 × 3 convolution; (b) 3 × 3 dilated convolution with a dilated rate of 2; (c) 3 × 3 dilated convolution with a dilated rate of 4 (For the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

dilated rate of 2 and the 3 × 3 convolution with a dilated rate of 4. Their corresponding receptive fields are 3 × 3, 7 × 7 and 15 × 15. The number of parameters (red dot) associated with each layer is identical. The dilated convolution supports the exponential expansion of the perceptive field without any loss of resolution or coverage [21].

Similar to other deep learning methods, data augmentation is helpful to improving model training, particularly when only a few training samples are available. As shown in Figure 4, in this study, we conduct a data augmentation strategy for the ROIs by rotating 45, 90, 135, 180, 225, 270 and 315 deg, flipping in the vertical and horizontal directions. Therefore, there are another 9 different copies of the ROIs generated from the described augmentation scheme.

E. SEGMENTATION EVALUATION

The performance of the segmentation model is evaluated by comparing its accuracy with that of the ground truth

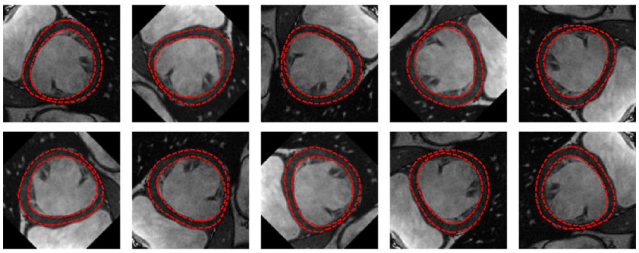

FIGURE 4. Data augmentation (For the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

(manual annotations by experts) based on the software published by the MICCAI Clinical Image Segmentation Grand Challenge Workshop. The evaluation proposed for assessing the algorithms submitted to the MICCAI 2009 LV segmentation challenge is based on the following three measures: 1) the percentage of "good" contours (PGC), 2) the average Dice metric (ADM) of the "good" contours, and 3) the average perpendicular distance (APD) of the "good" contours. A segmentation is classified as good if the APD is less than 5 mm [12]. To facilitate comparisons with other algorithms, we employed all these measures in this study. The perpendicular distance measures the distance from the autosegmented contour to the corresponding ground truth averaged over all contour points. A low APD implies that the two contours closely match. The Dice metric, $DM = 2(A_{\text{seg}} \cap A_{\text{gt}})/(A_{seg} + A_{gt})$ is a measure of contour overlap considering the autosegmented contour area A_{seg}, the ground truth contour area A_{gt}, and their intersection. A high DM indicates a better match between automated and manual segmentations. ADM is the average value of the DM per slice. The PGC is the ratio of the number of good contours to the total number of contours. The contours from automatic algorithms that are not considered to be good are excluded in the calculations of the ADM and APD. Each measure is computed slice by slice and a mean value and standard deviation for all slices were calculated.

In addition, two critical parameters for cardiac diagnosis, the left ventricle mass (LVM) and ejection fraction (EF), are computed based on segmentation results and used for correlation and Bland-Altman analyses [22]. The LVM and EF are defined as follows.

$$LVM = \left(V_{epi}^{ED} - V_{end}^{ED}\right) * 1.05 \tag{1}$$

$$EF = \frac{\left(V_{end}^{ED} - V_{end}^{ES}\right)}{V_{end}^{ED}} * 100\% \tag{2}$$

where V_{epi}^{ED} and V_{end}^{ED} represent the epicardial and endocardial volumes in the end-diastole (ED) phase, respectively, while V_{end}^{ES} represents the endocardial volume in the end-systole (ES) phase. Correlation analysis is performed to obtain the slope and intercept equation and the coefficient of determination R2. For Bland-Altman analyses, the mean value and the standard deviation (SD) of the differences between the automatic segmentation results and the ground truth

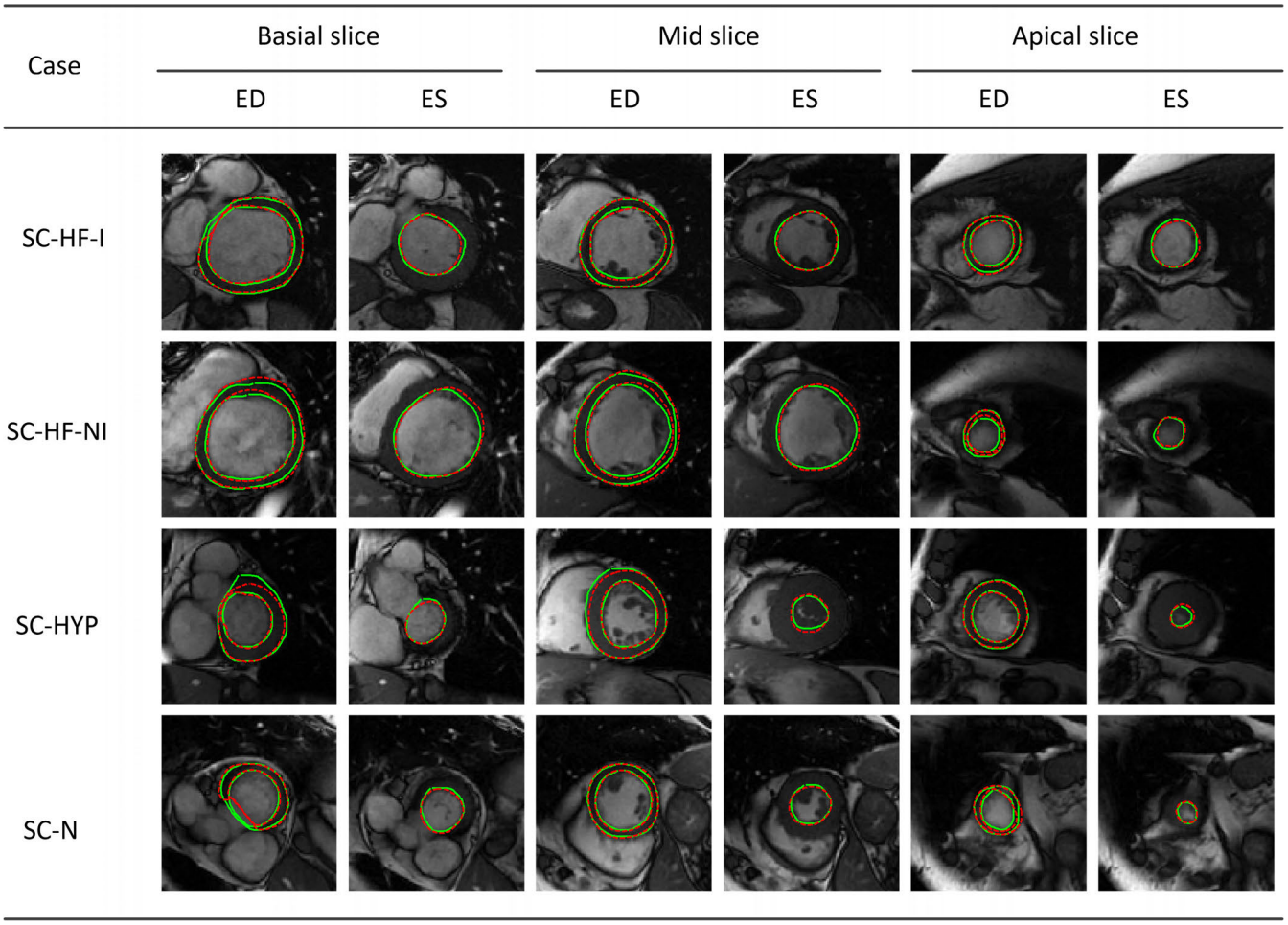

FIGURE 5. Automatic (solid green line) and manual (dotted red line) LV segmentation in the basial slice, mid-slice and apical slice for multiple cases (various patient types) of the MICCAI 2009 (For the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article). ED, end-diastole; ES, end-systole; SC-HF-I, heart failure with ischemia; SC-HF-NI, heart failure without ischemia; SC-HYP, hypertrophy; and SC-N, normal heart.

are computed. The coefficient of variation (CV), defined as the SD divided by the mean value, and the reproducibility coefficient (RPC), defined as 1.96∗SD, are computed.

III. EXPERIMENTS AND RESULTS

A. IMPLEMENTATION DETAILS

Our model is implemented on a computer with an Intel® Core® 2.6 GHz CPU, 32 GB of RAM, and the 64-bit Windows 10 operating system. The development environment is TensorFlow 1.12 with CUDA 9.0 and cuDNN 7.0. The learning rate is adjusted dynamically and linearly from 0.004 in the first epoch to 0.00001 in the 120th epoch. The dropout rate was set to 0.2. The training data are shuffled before each epoch.

To facilitate the comparison with the results of other LV segmentation algorithms, our experiments are based on 45 sequences of the MICCAI 2009 left ventricle segmentation challenge database. 30 sequences (the training and validation datasets of MICCAI 2009) are used to train the proposed model and the other 15 sequences (online dataset) are used for testing.

B. SEGMENTATION RESULTS

Figure 5 illustrates the automatic and manual LV segmentation results for four different types of patients in typical slices (basial slice, mid-slice and apical slice) in the ED and ES cardiac phases. The epicardial contour in the ES phase is omitted by clinical experts and segmentation methods because it is not used when computing the LVM and EF. Each column corresponds to one patient with heart failure with ischemia (SC-HF-I), one with heart failure without ischemia (SC-HF-NI), one with hypertrophy (SC-HYP) and one with a normal heart (SC-N). The automatic segmentation results are shown using green solid lines, and the ground truth manual segmentations drawn by experts are shown using red dotted lines. As seen from Figure 5, for all types of cases, the segmentation results of the LV epicardium and endocardium in all slices of all cases are very close to the ground truth of experts' manual annotations.

TABLE 1. Experimental results.

Patient id	endocardium			epicardium		
	PGC (%)	ADM	APD (mm)	PGC (%)	ADM	APD (mm)
SC-HF-I-05	100	0.96	1.12	100	0.97	0.94
SC-HF-I-06	100	0.95	1.41	100	0.97	1.12
SC-HF-I-07	100	0.93	1.57	100	0.97	1.22
SC-HF-I-08	100	0.95	1.49	100	0.96	1.59
SC-HF-NI-07	100	0.94	1.87	100	0.97	1.18
SC-HF-NI-11	100	0.95	1.64	100	0.97	1.37
SC-HF-NI-31	100	0.91	1.80	100	0.96	1.32
SC-HF-NI-33	100	0.91	1.72	100	0.96	1.33
SC-HYP-06	100	0.92	0.99	100	0.96	1.15
SC-HYP-07	100	0.94	0.98	100	0.97	1.4
SC-HYP-08	100	0.92	1.47	100	0.96	1.56
SC-HYP-37	92.31	0.90	2.21	100	0.94	1.73
SC-N-05	100	0.91	1.55	100	0.95	1.49
SC-N-06	100	0.94	1.10	100	0.96	1.04
SC-N-07	100	0.87	1.54	100	0.95	1.18
MEAN	99.49	0.93	1.50	100	0.96	1.31
STD	1.99	0.03	0.34	0.00	0.01	0.22

PGC, percentage of "good" contours; ADM, average Dice metric; APD, average perpendicular distance; SC-HF-I, heart failure with ischemia; SC-HF-NI, heart failure without ischemia; SC-HYP, hypertrophy; SC-N, normal heart.

To further quantitatively analyze our segmentation results, the statistics of the automatic segmentation results compared to the ground truths in 15 testing sequences are listed in Table 1. The mean±SD of the PGC (percentage of good contours), APD (average perpendicular distance) and ADM (average Dice metric) are listed for the online dataset. These results show that for the endocardium and epicardium, the percentages of good contours are 99.49%±1.99% and 100%±0.00%, respectively, which means that the results of LV epicardial contour segmentation are good (APD is less than 5 mm) for all slices and that the LV endocardial contour segmentation results are good for almost all slices. In addition, the APDs are 1.50±0.34 mm and 1.31±0.22 mm, respectively, and the ADMs are 0.93±0.03 and 0.96±0.01, respectively, which indicate that our method provides contours that have great agreement with the ground truth.

A regression and Bland-Altman analysis are used to evaluate our segmentation results. Figure 6 shows the regression (left) and Bland-Altman plots (right) for the left ventricle mass (LVM) and the ejection fraction (EF). The coefficient of determination R2 for the LVM is 0.97, the spread of the values is pretty low, and the slope is approximately 1.09. In addition, the linear regression of the EF is similar to that of the LVM. The level of agreement between our results and the ground truths is represented by the interval of the percentage difference between the mean±1.96SD numbers. The mean and confidence interval of the difference between

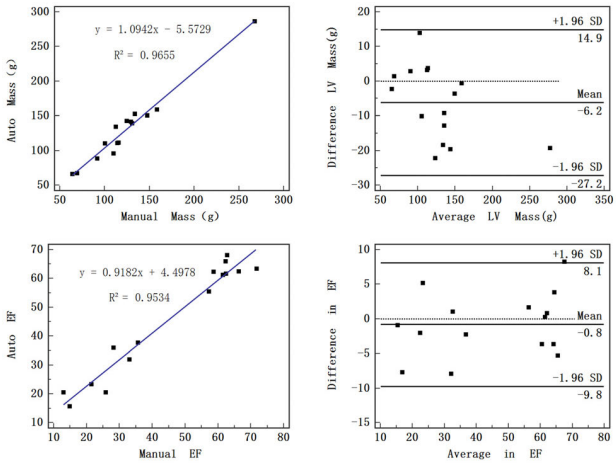

FIGURE 6. Regression curve and Bland-Altman plot for the left ventricle (LV) mass and the ejection fraction (EF); left: linear regression and Bland-Altman plots of LV mass, right: linear regression and Bland-Altman plots of EF.

the automatic and manual LVM results are -6.2 g and (-27.2 g, 14.9 g), respectively. The CV and RPC are 8.4% and 16.5%, respectively. For the EF, the coefficient of determination is 0.95, and the slope is 0.92. The mean and confidence interval of the difference between the automatic and manual EF results are -0.8 and (-9.8, 8.1), respectively. The CV and RPC

TABLE 2. Comparison of our segmentation performance with the traditional segmentation methods and state-of-the-art technique using the MICCAI database.

methods	endocardium			epicardium		
	PGC (%)	ADM	APD (mm)	PGC (%)	ADM	APD (mm)
LU's (2009)	77.63%	0.89	2.07	85.68%	0.94	1.91
Hu et al. (2014)	93.10%*	0.88*	2.38*	93.20%*	0.94*	2.31*
Ngo et al. (2017)	95.91%	0.88	2.34	94.65%	0.93	2.08
Ours'	99.49%	0.93	1.50	100%	0.96	1.31

PGC, percentage of "good" contours; ADM, average Dice metric; APD, average perpendicular distance; Symbol * denotes the p-value for the distributions of this metric between our model and the corresponding method is lower than 0.01.

are 10.0% and 19.7%, respectively. The regression and Bland-Altman analysis depict a high correlation for the two clinical cardiac indices. The high correlation between our results and the ground truths show the accuracy and clinical applicability of our framework for the automatic evaluation of LV functioning. Additionally, for the Bland-Altman scatter plot, the confidence intervals for LVM and EF are (-27.2, 14.9 g) and (-9.8, 8.1), respectively. In addition, 0.0% (0/15) and 6.7% (1/15) are outside the confidence intervals for the LVM and EF, respectively. All these results show that the results of our algorithm are in good agreement with the manually delineated results.

C. COMPARISONS

To verify the effectiveness of the segmentation algorithm proposed in this article, we compare our proposed model with Lu's method [23], Hu's method [24] and Ngo's method [14] on the same database of MICCAI 2009. Table 2 summarizes the percentages of good contours (PGCs), the average Dice metrics (ADMs) and the average perpendicular distances (APDs) of these four methods. It can be seen that the PGC of our method is not only much higher than those of the two traditional methods, but it also improves that of the state-of-the-art technique by 3.6% for the endocardium and 5.4% for the epicardium. Considering that the ADM and APD are calculated using the good contours, the higher the PGC, the more slices that we have calculated in the experiment. In this case, the ADM and APD of the endocardium with our method are 0.93 mm and 1.50 mm, respectively, which are still 0.04 mm and 0.57 mm better than the best ADM and APD reported by others. For the epicardium, its ADM and APD with our method are 0.02 mm and 0.60 mm better than the best ADM and APD reported by others. According to Student's t-test, the p-values for the distribution of the PGC, ADM and APD metrics of our method and Hu's algorithm are all <0.01, which indicates that the difference is statistically significant. Due to the lack of relevant details of Lu's method [23] and Ngo's method [14], we have not been able to reproduce these algorithms for the time being. And the p-values between our results and them cannot be given. These performances show that the results of our model are more accurate and robust than those other algorithms.

To demonstrate the advantages of our proposed method, we compared the derived contours from 5 cases (SC-HF-I-05, SC-HF-I-06, SC-HYP-06, SC-HF-NI-11, and SC-N-06) obtained with our method, Hu's method, and experts' manual annotations. Lu's and Ngo's delineations are not given because we have not been able to reproduce their algorithms for the time being. Figure 7 shows the segmentation results on 5 ROIs randomly selected from the above 5 cases. The automatic segmentation results are shown by the green solid lines and the ground truth manual segmentations drawn by experts are shown by the red dotted lines. As it can be seen, compared with Hu's algorithm, our segmentation results are closer to the manual results.

D. ABLATION STUDIES

The automatic segmentation model based on the DDCN has been tested and verified in the previous section. To evaluate each component of our proposed method, we also perform ablation studies in terms of ROI cropping, the network architecture, and the dilated convolution.

To analyze the influence of ROI cropping on the segmentation results, two different methods are used to crop the ROIs (Figure 8), which are then input into the same DDCN model. The first method is to crop ROIs that are 190×190 pixels at the center of each slice. These ROIs contain the objects (blood pool and myocardium) and almost all surrounding tissues in the slice, which significantly dominate the image. The second method is to crop ROIs that are 110×110 pixels as described earlier in this article. For these ROIs, the number of pixels in the object is roughly the same that of the surrounding tissues. Table 3 shows the comparison results. It can be seen from the table that the PGC, ADM and APD of the endocardium (epicardium) segmentation results corresponding to our 110×110 ROIs are 99.49% (100.00%), 0.93 (0.96) and 1.50 mm (1.31 mm), respectively, which are better than the results for the 190×190 ROIs of 98.60% (99.17), 0.92 (0.96) and 1.59 mm (1.56 mm), respectively. And the training time of endocardium (epicardium) of 110×110 ROIs decreased from 1271 (615) second/epoch of 190×190 ROIs to 591 (298) second/epoch.

For the second ablation study, our DDCN model with eight dense layers is compared with structures have combinations

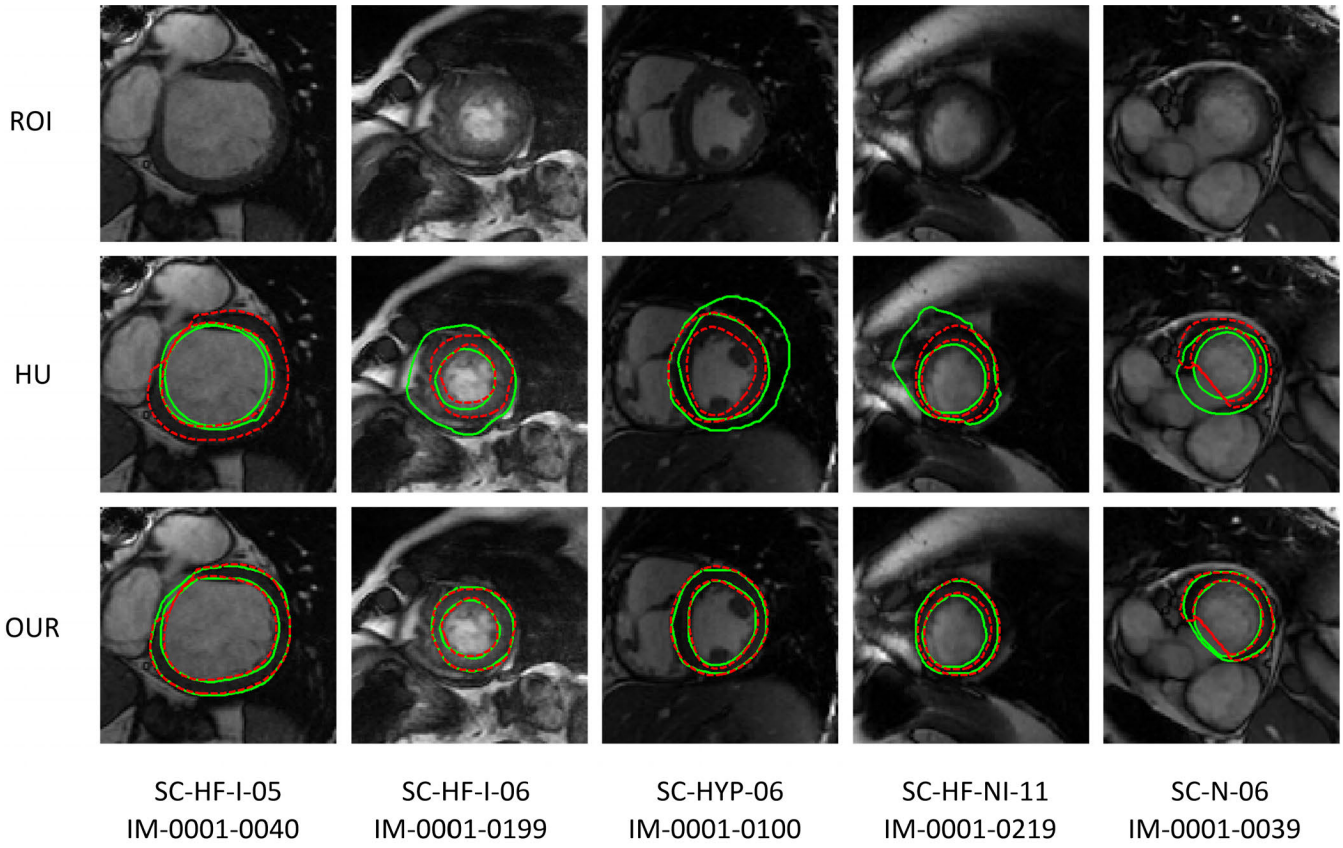

FIGURE 7. Comparison of Left Ventricular MR Image Segmentation (For the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article). SC-HF-I, heart failure with ischemia; SC-HF-NI, heart failure without ischemia; SC-HYP, hypertrophy; and SC-N, normal heart.

TABLE 3. Comparison of the ROI segmentation results of the 190 × 190 and 110 × 110 ROIs.

ROI	endocardium				epicardium			
	PGC (%)	ADM	APD (mm)	Time (s)	PGC (%)	ADM	APD (mm)	Time (s)
190×190	98.60	0.92	1.59	1271	99.17	0.96	1.56	615
110×110	99.49	0.93	1.50	591	100.00	0.96	1.31	298

PGC, percentage of "good" contours; ADM, average Dice metric; APD, average perpendicular distance.

of different numbers of blocks and layers. As seen from Table 4, when the number of layers of the model is fixed, the training time (second/epoch) of the model increases rapidly as the number of blocks increases, but the corresponding performance does not improve. When the number of dense layers is no more than 6, the overall performance of the model is increasing as the number of layers increases. However, when the number of dense layers is more than 6, the benefit of increasing the number of layers is not significant, but the training time has increased quickly. Compared with the 8-layer model, the 9-layer model has a decreased PGC and a more than tripled training time. Therefore, we chose the 8-layer model.

For the third ablation study, we compare the dilated convolution with the traditional convolution to illustrate its efficiency. Figure 9 shows the loss curves and acc curves of the dilated convolution (blue line) and traditional convolution (orange line) on the training set (left column) and validation set (right column). It can be seen that the dilated convolution has a better convergence speed and final convergence result than the traditional convolution.

IV. DISCUSSION

In this article, a segmentation schema is developed for LV segmentation in MRI. First, a clustering algorithm that can make full use of the relationship between adjacent slices to accurately locate the blood pool area is proposed to crop the ROI. Then, the DDCN model with 8 dilated convolutional layers and dense connections is proposed to delineate the endocardium and epicardium. Our proposed method is compared with other traditional algorithms and deep learning algorithms based on the widely used MICCAI 2009 left

TABLE 4. Comparison of our eight dense layer structure model with other structures.

structures		endocardium			epicardium		
layers	blocks	PGC (%)	ADM	Time (s)	PGC (%)	ADM	Time (s)
4	1	93.42	0.90	55	87.17	0.92	30
4	2	92.73	0.92	118	86.13	0.95	61
4	3	92.60	0.93	183	89.05	0.95	95
5	1	97.76	0.91	74	90.52	0.93	38
5	2	98.04	0.91	162	96.35	0.93	88
5	3	94.48	0.92	242	98.50	0.95	128
6	1	97.05	0.91	110	98.73	0.94	58
7	1	96.98	0.91	179	97.08	0.94	92
8	1	98.00	0.91	324	98.66	0.95	164
9	1	97.09	0.92	1094	98.54	0.94	550

PGC, percentage of "good" contours; ADM, average Dice metric; APD, average perpendicular distance.

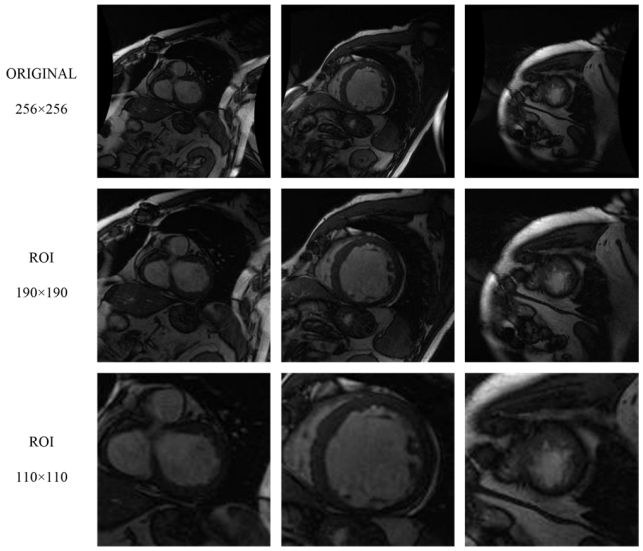

FIGURE 8. Original MR images and ROIs with sizes of 190 × 190 and 110 × 110 pixels.

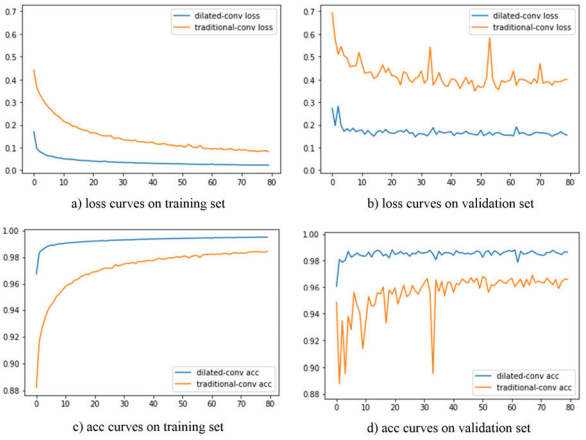

FIGURE 9. Training process of the dilated convolution and traditional convolution (For the interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

ventricle segmentation challenge database. The comparison results show that our method is very effective in the LV segmentation of MR images.

The traditional LV segmentation methods are usually based on shape and appearance terms and tend to perform poorly on images with obscured boundary, especially those slices with LV outflow tract or apical slices with very small region of blood pool. However, through annotated training, our DDCN model can achieve promising results on these images. In addition, compared with other deep learning methods, our proposed method has the following characteristics: 1) The ROI cropping process can locate the blood pool area accurately and improve the segmentation result. 2) The customized densely connected network structure avoids the problem of gradient vanishing, can effectively update the network parameters, and improve the training effect. 3) The dilated convolution expands the receptive field, and enables our model to learn the characteristics more fully to obtain precise contours. All these characteristics effectively improved the accuracy of LV segmentation.

As shown in Figure 1, the surrounding tissues of the LV significantly dominate the original MR image. To alleviate the problem of the class imbalance between the desired object and surrounding tissues, improve the classification (segmentation) accuracy and reduce the computation, the ROI containing the region of the blood pool should be cropped from the original image. In addition, the region of the blood pool is large at the top and small at the bottom. When cropping the ROI, blood pool region with various size should be contained, so the size of the ROI cannot be too small. According to the above considerations, the ROI with a size of 110 × 110 pixels has been adopted to make the pixels of object and surrounding tissues roughly balanced. The results in Table 3 verify the effectiveness of our ROI cropping. It shows that ROI cropping can not only increase the PGC but also reduce the class

imbalance and improve the segmentation result. In addition, the ROI-based method can significantly reduce the memory consumption and training time.

To verify that the DDCN model (with 1 block and 8 convolutional layers) is generalizable to variable clinical datasets with different ROI sizes, we also conduct experiments on another updated database MICCAI 2018. The ROIs with a size of 80 × 80 pixels in MICCAI 2018 are collected from 3 hospitals affiliated with two health care centers (London Healthcare Center and St. Josephs HealthCare). 15 of the 145 cases in the MICCAI 2018 have been used for testing, and the PGC, ADM and APD of the endocardium (epicardium) segmentation results are 100.00% (100.00%), 0.96 (0.97) and 1.31 mm (1.42 mm). Although the ROIs used in MICCAI 2018 and MICCAI 2009 come from different hospitals and have different sizes, our DDCN model has achieved very promising performance on both datasets, which verifies the adaptability of our model to different datasets. In the future research, we will further verify whether the optimal parameters (i.e., ROI size and convolution layer structure) of the model vary with different datasets.

The DenseNet (with 3 blocks, each block has 5 convolutional layers) with dilated convolutions was used to delineate the right ventricle by Zhao *et al.* [17]. The architecture of our DDCN model has a different structure from them. Considering the performance of LV segmentation and training time (as shown in Table 4), a structure with 1 block and 8 convolutional layers has been adopted. Each layer concatenates the feature maps of all previous layers and its own feature maps as the input of all subsequent layers, which greatly alleviates the vanishing gradient problem and strengthens the feature propagation. Therefore, it can effectively update the network parameters, and improve the training effect. This is particularly important for segmentation of LV images with obscured and incomplete boundaries.

The dilated convolution expands the receptive field of the convolution kernel without any loss of resolution or coverage while keeping the number of parameters identical. And the expanded field enables our model to learn the characteristics of endocardium and epicardium more fully to obtain precise contours. In addition, using the dilated convolution instead of the traditional convolution, the convergence speed and final convergence result of our model have been significantly improved (as shown in Figure 9).

As discussed above, ROI cropping, the network architecture and the dilated convolution ensure that our DDCN model can achieve excellent performance in LV segmentation. Developing a pure deep learning method on a larger set of clinical data is the subject of our future research.

V. CONCLUSION

To accurately extract the endocardium and epicardium of the LV from cardiac MR images, a novel method for automatic LV segmentation of cardiac MR images is proposed. First, the ROI containing the blood pool is cropped based on a clustering method that integrates the prior knowledge about the spatial relationship between the adjacent slices. The ROI is incorporated to reduce memory consumption, computing time and the class imbalance between the target and background. Then, a dilated dense convolution network that combines the dilated convolution and dense connections is proposed to delineate the endocardium and epicardium. The segmentation results on the MICCAI 2009 are compared with the results of the traditional algorithms and deep learning algorithms (state-of-the-art). Excellent agreement and a high correlation with the reference contours are obtained by our method of left ventricle segmentation.

ACKNOWLEDGMENT

We thank Sunnybrook Health Sciences Centre, London Healthcare Center and St. Josephs HealthCare for their image data.

CONFLICT OF INTEREST

This article content has no conflict of interest.

REFERENCES

[1] S. S. Virani, A. Alonso, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, A. R. Chang, S. Cheng, F. N. Delling, and L. Djousse, "Heart disease and stroke statistics—2020 update: A report from the American heart association," *Circulation*, vol. 141, pp. E139–E596, Mar. 2020.

[2] E. Maffei, G. Messalli, C. Martini, K. Nieman, O. Catalano, A. Rossi, S. Seitun, A. I. Guaricci, C. Tedeschi, N. R. Mollet, and F. Cademartiri, "Left and right ventricle assessment with cardiac CT: Validation study vs. cardiac MR," *Eur. Radiol.*, vol. 22, no. 5, pp. 1041–1049, May 2012.

[3] C. Petitjean and J.-N. Dacher, "A review of segmentation methods in short axis cardiac MR images," *Med. Image Anal.*, vol. 15, no. 2, pp. 169–184, Apr. 2011.

[4] H. Hu, H. Liu, Z. Gao, and L. Huang, "Hybrid segmentation of left ventricle in cardiac MRI using Gaussian-mixture model and region restricted dynamic programming," *Magn. Reson. Imag.*, vol. 31, no. 4, pp. 575–584, May 2013.

[5] Z. Wang, J. Lu, X. Zhang, and M. Lin, "A fusion algorithm of spatial fuzzy C-means clustering and level set for magnetic resonance image segmentation," *J. Comput. Inf. Syst.*, vol. 10, no. 39, pp. 4675–4682, Jun. 2014.

[6] S. Xu, X. Xu, H. Hu, and B. Li, "Left ventricle MRI segmentation based on developed dynamic programming," *J. Guangxi Normal Univ., Natural Sci. Ed.*, vol. 32, no. 2, pp. 35–41, Jun. 2014.

[7] L. Qi, X. Lyu, B. Yang, and L. Xu, "Segmentation of left ventricle endocardium based on transfer learning of fully convolutional networks," *J. Northeastern Univ., Natural Sci.*, vol. 39, no. 11, pp. 1577–1581 and 1592, Nov. 2018.

[8] L. K. Tan, Y. M. Liew, E. Lim, and R. A. McLaughlin, "Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences," *Med. Image Anal.*, vol. 39, pp. 78–86, Jul. 2017.

[9] C. Cong and H. Zhang, "Invert-U-net DNN segmentation model for MRI cardiac left ventricle segmentation," *J. Eng.*, vol. 2018, no. 16, pp. 1463–1467, Nov. 2018.

[10] H. Abdeltawab, F. Khalifa, F. Taher, G. Beache, T. Mohamed, A. Elmaghraby, M. Ghazal, R. Keynton, and A. El-Baz, "A novel deep learning approach for left ventricle automatic segmentation in cardiac cine MR," in *Proc. 5th Int. Conf. Adv. Biomed. Eng. (ICABME)*, Tripoli, Lebanon, Oct. 2019, pp. 1–4.

[11] L. V. Romaguera, M. G. F. Costa, F. P. Romero, and C. F. F. C. Filho, "Left ventricle segmentation in cardiac MRI images using fully convolutional neural networks," *Proc. SPIE*, vol. 10134, Mar. 2017, Art. no. 101342Z.

[12] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," *Med. Image Anal.*, vol. 30, pp. 108–119, May 2016.

[13] J. Zhang, L. Fang, Y. Chen, Z. Tian-ming, and X. Li, "Left ventricle MRI segmentation based on active contour model," *Acta Electronica Sinica*, vol. 39, pp. 2670–2673, Nov. 2011.
[14] 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," *Med. Image Anal.*, vol. 35, pp. 159–171, Jan. 2017.
[15] Y. Lan and R. Jin, "Automatic segmentation of the left ventricle from cardiac MRI using deep learning and double snake model," *IEEE Access*, vol. 7, pp. 128641–128650, 2019.
[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269.
[17] Z. Xingrong, L. Zheng, Z. Heming, and C. Xueqin, "Segmentation of right ventricular MR image based on deep neural network: Dilated DenseNet of two level losses," in *Proc. 14th IEEE Int. Conf. Signal Process. (ICSP)*, Aug. 2018, pp. 355–358.
[18] 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," *Med. Image Anal.*, vol. 51, pp. 21–45, Jan. 2019.
[19] (2009). *Evaluation of Cardiac MR Segmentation*. [Online]. Available: http://sourceforge.net/projects/cardiac-mr/files/
[20] S. Mehta, M. Rastegari, A. Caspi, L. Shapiro, and H. Hajishirzi, "Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 552–568.
[21] Y. Wang, G. Wang, C. Chen, and Z. Pan, "Multi-scale dilated convolution of convolutional neural network for image denoising," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 19945–19960, Jul. 2019.
[22] C. Bunce, "Correlation, agreement, and Bland–Altman analysis: Statistical analysis of method comparison studies," *Amer. J. Ophthalmol.*, vol. 148, no. 1, pp. 4–6, Jul. 2009.
[23] Y. Lu, P. Radau, K. Connelly, A. Dick, and G. A. Wright, "Segmentation of left ventricle in cardiac cine MRI: An automatic image-driven method," in *Proc. Int. Conf. Funct. Imag. Model. Heart*, 2009, pp. 339–347.
[24] H. Hu, Z. Gao, L. Liu, H. Liu, J. Gao, S. Xu, W. Li, and L. Huang, "Automatic segmentation of the left ventricle in cardiac MRI using local binary fitting model and dynamic programming techniques," *PLoS ONE*, vol. 9, no. 12, Dec. 2014, Art. no. e114760.

SHENGZHOU XU received the Ph.D. degree in computer science and technology from the Huazhong University of Science and Technology in 2011. He is currently an Associate Professor with the College of Computer Science, South Central University for Nationalities, Wuhan, Hubei, China. He was a Visiting Scholar with the University of North Carolina at Chapel Hill, USA, from 2017 to 2018. His current research interests include medical image processing and deep learning.

SHIYU CHENG received the B.S. degree in electronic and information engineering from Southwest Minzu University, Chengdu, Sichuan, China, in 2018. He is currently pursuing the M.S. degree in computer technology from South Central University for Nationalities, Wuhan, Hubei, China. His main research interests include deep learning and medical image processing.

XIANGDE MIN was born in Zibo, Shandong, China, in 1987. He received the B.S. degree in medical imaging from the Binzhou Medical University, Yantai, Shandong, in 2012, and the Ph.D. degree in imaging and nuclear medicine from the Huazhong University of Science and Technology, Wuhan, Hubei, in 2017.

From 2017 to 2019, he was a Diagnostic Radiology Resident. Since 2019, he has been an Attending Radiologist with the Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. He has authored more than 70 articles. His research interests include machine/deep learning and their applications to medical image analysis and magnetic resonance image processing.

NING PAN received the M.S. degree in computer software and theory from Chongqing University in 2007 and the Ph.D. degree in computer science from the Huazhong University of Science and Technology (HUST) in 2013. He was a Postdoctoral Fellow with the School of Computer Science and Technology, HUST, for two years. He is currently a Lecturer with the College of Biomedical Engineering, South Central University for Nationalities. His research interests include medical image processing, data mining, and statistical learning.

HUAIFEI HU received the M.S. and Ph.D. degrees in computer science and technology from the Huazhong University of Science and Technology in 2007 and 2012, respectively. He was a Visiting Scholar with The University of Sheffield, U.K., from 2015 to 2017. He is currently a Lecturer with the College of Biomedical Engineering, South Central University for Nationalities, Wuhan, Hubei, China. His current research interests include medical image processing and deep learning.