IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 5, 2018, accepted January 7, 2019, date of publication January 14, 2019, date of current version February 14, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2892965

MA-Shape: Modality Adaptation Shape Regression for Left Ventricle Segmentation on Mixed MR and CT Images

JINYU CONG[1], YUANJIE ZHENG[1,2], WUFENG XUE[3,4], BOFENG CAO[5], AND SHUO LI[3]

[1]School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China
[2]Key Lab of Intelligent Computing and Information Security in Universities of Shandong, Shandong Provincial Key Laboratory for Novel Distributed Computer Software Technology, Institute of Biomedical Sciences, Shandong Normal University, Jinan 250358, China
[3]Department of Medical Imaging, Western University, London, ON N6A 3K7, Canada
[4]School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen 518060, China
[5]Medical Imaging Department, Yantai Yuhuangding Hospital, Yantai 264000, China

Corresponding author: Yuanjie Zheng (zhengyuanjie@gmail.com)

This work was supported in part by the National Nature Science Foundation of China under Grant 61572300, Grant 81871508, and Grant 61773246, in part by the Taishan Scholar Program of Shandong Province of China under Grant TSHW201502038, and in part by the Major Program of Shandong Province Natural Science Foundation under Grant ZR2018ZB0419.

ABSTRACT Left ventricle (LV) segmentation is essential to clinical quantification and diagnosis of cardiac images. While most existing LV segmentation methods focus on cardiac images of single modality or multi-modality, few have been devoted to images of mixed-modality. By Mixed-Modality, we mean that different modalities exist in the database, while for every subject, there is only one modality. In this paper, we propose a newly invented LV segmentation method from mixed-modality images: modality adaptation shape regression (MA-Shape). Compared to single-modality or multi-modality methods, the proposed MA-Shape can 1) be applied to images of new modality during the test phase, which improves the generalization of the learned methods, and 2) take advantage of existing samples of different modalities, which alleviates the high demand for multi-modality data. To achieve this, we propose a modality adaptation module to enhance the shape consistency between the MR and CT, and therefore improve the generalization of the model learned in one modality to new modalities. The experiments on a dataset with MR sequences of 145 subjects and CT scans of 96 subjects validate that the proposed MA-shape can achieve excellent performance by learning common shape information from images of mixed modality and improve the cross-modality generalization of shape regression model learned on images of one modality. These advantages not only provide an efficient way of utilizing mix-modalities data during model learning but also enables an effective and flexible way of applying automated cardiac function assessment in clinical practice.

INDEX TERMS Left ventricle segmentation, mixed-modality images, modality adaptation shape regression.

I. INTRODUCTION

Cardiovascular disease (CVDS) is one of the Asia-pacific region's biggest non-communicable diseases challenges [1], [2]. For clinical practice, diverse medical imaging technologies (MRI, CT, PET) are crucial support technologies for cardiovascular diseases diagnosis. Left ventricle (LV) segmentation in short-axis sequences is one of the prerequisites and essential steps to clinical cardiac contractile function quantification, which includes estimation of ejection fraction and the calculation of left ventricle (LV) end-diastolic and end-systolic volumes.

Manual segmentation by radiologists is a sterile, time-consuming and inefficient work which spend about 20 minutes for the MR sequence of one subject, according to the statistics. Meanwhile, there exist high variabilities intra- and inter-observers due to clinical experience and the interpretation of the myocardium. Automated left ventricle segmentation captures researchers' attention because it can free radiologists from the predicament of tedious, inefficiency, and inconsistency work caused by the enormous amounts of medical images from different modalities [3]–[6]. However, obtaining exact and correct boundary of epicardium

(epi) and endocardium (endo) is still an extremely challenging ambition due to the flow of blood, the variability of shape triggered by expansion and contraction of the heart, the overlap of intensity, the difference between subjects, the impact of disease, papillary muscles and inherent noise, etc.

Existing research regarding LV segmentation in short-axis sequences can generally be categorized as two groups:

- Single-modality based methods. Single-modality based methods are mainly consist of MR-based methods and CT-based methods. For the MR-based methods, different intensity thresholding based model [7], pixel or voxel classification [8], [9] and active contours based models [10]–[17] are employed for cardiac MR images. Then the strong prior information has consisted in automated LV segmentation for best effect on MR images. Reference [18] introduced the early application of the global shape to LV segmentation on MR images. Lekadir *et al.* [19] proposed a framework for LV segmentation that is based on shape extraction and interpretation in 4-D CMR images. A novel alternate bottom-up modeling method was presented to segment LV on MR images [20]. A partial sparse shape framework was proposed in [21] that express the shape of the left ventricle on MR images. In the work of Avendi *et al.* [22], a combined deep-learning and deformable-model approach were implemented to segment the LV in short axis MR images. The research proposed by Romaguera *et al.* [23] tackle the problem of automated LV segmentation through fully convolutional neural networks which are trained end to end on MR images. Tran [24] employ the deep FCN to perform left ventricle segmentation on MR images. Meanwhile, many researchers pay attention to cardiac CT images. Reference [25] propose an active contours based method which combining the localizing region and edge-based intensity to segment left ventricle in cardiac CT images. Ecabert *et al.* [26] propose a model-based approach by progressively increasing the degrees-of-freedom of the allowed deformations for the fully automatic segmentation in 3-D CT images. Marginal space learning and steerable features are employed on the four-chamber heart segmentation from 3-D CT volumes [27]. Dahiya *et al.* [28] proposed aa highly customized parametric model to segment myocardial in CT images. Zhuang *et al.* [29] propose a multi-atlas whole heart segmentation on CT images. Besides, there are many other methods which are designed on the single modality images. However, unlike the previous methods, they can apply to different modality images through a new training stage on the images of new modality images. For instance, Weng *et al.* [30] proposed a threshold based method on a likelihood measure for ventricle detection application to cardiac MR and CT images. Reference [31] describes a segmentation technique which contains the global localization step and local deformation step, and then it has been tested on the cardiac MR and CT images. Bernier *et al.* [32] employ graph cut to segment left ventricle on the MR and echocardiographic images. Reference [33] propose a snake model based on the MR images, CT images and ultrasound images. Li *et al.* [34] present a semi-automated segmentation method which performs well on CT and MR images.
- Multi-modality based methods. Unlike the above single-modality based methods, some researchers propose the multi-modality based methods. These methods need the different modality images which collected on the same patients. For instance, Bertone *et al.* [35] use the cine-MR segmentation on the same patient to assist the segmentation of ultrasound cardiac images. First, they find the close registration of the MR images for the ultrasound volume. Then the ultrasound segmentation process can reference the shape constraint from cine-MR.

Although these previous studies have achieved good performance, they still ignore the following questions due to the modality of images, which can prevent them from effectively utilizing medical images of different modality and generalizing well to new modalities once learned from one specific modality.

- Lack of generalization to a new modality. Single-modality based method and multi-modality application based method cannot be applied directly and accurately to novel modalities because they are different from the training samples. Due to the different imaging procedure, the appearance of cardiac in MR images and CT images differ a lot from each other, as can be observed in Fig.1. The boundary of the myocardium, the intensity of the cavity and myocardium, demonstrate the clear difference between CT images and MR images. All these differences make a learning method trained on samples of one modality invalid for new modality. The single-modality based method need to be retrained on new

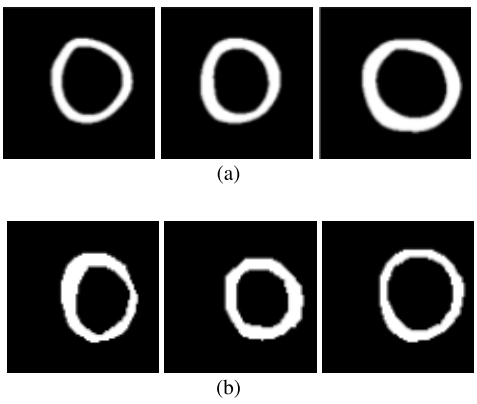

FIGURE 1. The shape of myocardium on MR images and CT images demonstrate subtle difference, which makes it possible for modality adaptation shape regression. (a) Shape of myocardium on cardiac MR images (b) Shape of myocardium on cardiac CT images.

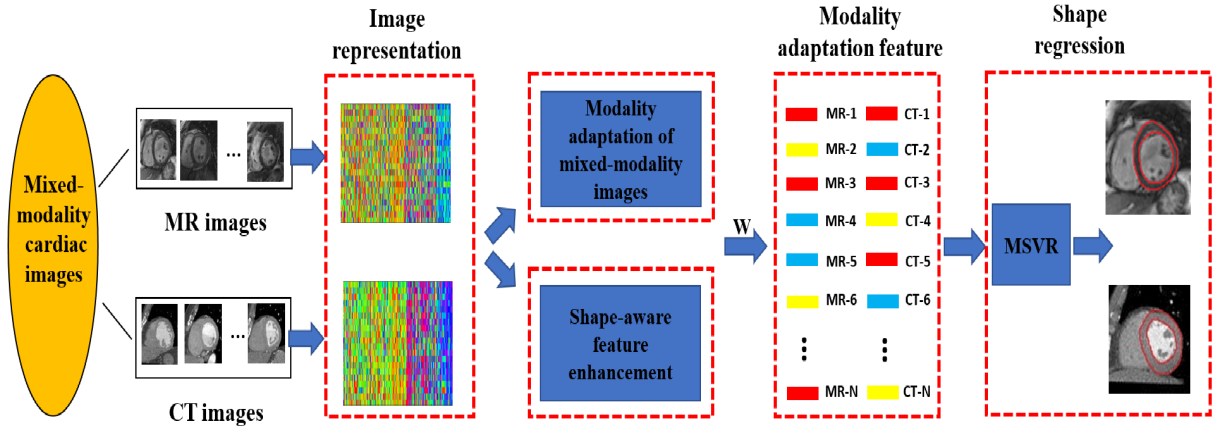

FIGURE 2. Modality adaptation shape regression for left ventricle segmentation on mixed MR and CT images.

samples before being applied to images of new modality, which again increase the demand of manually-labeled training samples.

- High cost of multi-modality data. The multi-modality based methods, such as [35], combine complementary information from multiple modalities of the same subject to improve the reliability. However, multi-modality data can not always be obtained in routine clinical practice and adds extra cost for patients. Besides, the obtained method is still restricted to the modalities contained in the training samples.

To overcome these problems, we propose a newly-invented LV segmentation method MA-Shape for images of mixed-modality. By Mixed-Modality, we mean that more than one modality exists in the training samples, while for every subject, only one modality exists. Compared to single-modality or multiple-modality methods, the proposed mixed-modality method can 1) be applied to images of new modality during the test phase without retrained on new samples, which improves the generalization of the learned methods, and 2) take advantage of existing samples of different modalities during the model learning phase, which alleviates the high demand for data for multi-modality.

To achieve this, we propose for mixed-modality cardiac images a novel adaptation module which can improve the consistency of MR images and CT images concerning the shape of cardiac myocardium. The underlying assumption of modality adaptation is that the shape of cardiac myocardium share similarities for different modalities image, while the appearance change doesn't affect the shape of the myocardium, as can be observed in Fig.1 and Fig.2. With this modality adaptation module, we propose an effective LV shape regression method MA-Shape that can successfully handle images of novel modality without retrained on new samples with the learned segmentation model.

The proposed MA-Shape is constituted of four successive steps: 1) cardiac image representation, which aims to capture the shape and appearance information of cardiac; 2) modality adaptation(MA) of mixed-modality images, which improves the shape consistency between the representation of MR and CT images and reduces the inconsistent information with dimension reduction; 3) shape-aware feature enhancement(SFE), which further enhance the shape awareness of the feature with supervised shape information, and 4) shape regression, which directly obtains the final myocardium contour from the shape-aware feature with multi-output support vector regression (MSVR).

We validate the proposed method on a mixed-modality cardiac database, which includes MR scans of 145 subjects and CT scans of 96 subjects. The average dice metric(DM), epicardium error and endocardium error are 91.1%, 1.74 mm, 1.46 mm on CT images, and 90.9%, 2.22 mm, 1.54 mm on MR images, respectively. Particularly, as demonstrated in the experiments, the proposed modality adaptation module improves the generalization of the learned model to new modalities and improves the performance of single modality method with the incorporation of training samples of new modalities.

In summary, the main contributions of our paper include:

- We propose, for the first time, a novel modality adaptation module for the mixed-modality database, which is capable of generalizing a machine learning model learned from data of one modality to data of novel modalities. It alleviates the requirements of training samplings in the learning phase and also enables deploying machine learning models to new modalities without the need for retraining.
- We propose an effective shape regression model MA-Shape for LV segmentation, which combines the advantages of modality adaptation and supervised feature learning, and obtains accurate segmentation performance on Mixed-Modality images of 145 MR scans and 96 CT scans.

The remaining of the paper is as follows. In Section II, the modality adaptation of shape regression for left ventricle segmentation on mixed MR and CT images is presented.

Then we give the database, experiment details, results and discussions in Section III. Section IV gives the conclusion.

II. MODALITY ADAPTATION OF SHAPE REGRESSION

In this paper, the left ventricle segmentation on mixed-modality images are modeled as the multi-output shape regression problem, with four consecutive steps (as illustrated in Fig. 3): image representation, modality adaptation(MA) of mixed-modality images, shape-aware feature enhancement(SFE) and shape regression. Hereafter, the mixed-modality image dataset can be denoted by $I_{M_1} = \{I_{M_1}^{1,1} \dots I_{M_1}^{a_1,b_1}\}$ and $I_{M_2} = \{I_{M_2}^{1,1} \dots I_{M_2}^{a_2,b_2}\}$, where M_1 and M_2 are the first modality and second modality, a_1 and a_2 are the number of subjects in M_1 and M_2, b_1 and b_2 are the frame number in one subject. Besides, we equally sample the N_1 and N_2 pixels in the endocardium and epicardium and use their coordinates as the ground truth for each image: $\mathbf{y} = \{(x,y)_{endo_1} \dots (x,y)_{endo_{N1}}, (x,y)_{epi_1} \dots (x,y)_{epi_{N2}}\}$. For the different modalities image I_{M1} and I_{M2}, the ground truth can be constructed by $Y_{M1} = \{\mathbf{y}_{M1}^{1,1} \dots \mathbf{y}_{M1}^{a,m}\}$ and $Y_{M2} = \{\mathbf{y}_{M2}^{1,1} \dots \mathbf{y}_{M2}^{b,n}\}$.

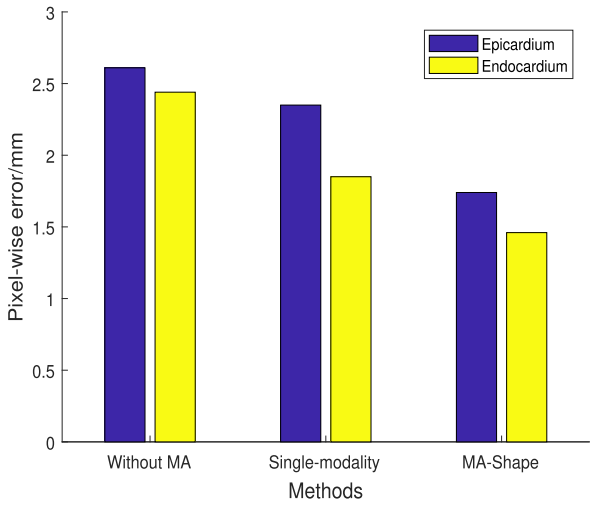

FIGURE 3. The shape regression error of MA-Shape on mixed-modality images, and comparison with settings of single modality and mixed-modality without modality adaption. The M_1 is CT in mixed-modality scenario.

A. IMAGE REPRESENTATION

Considering the mixed-modality images with different appearance and similar shape, we first obtain a representation which can model the shape of myocardium and dominate spatial structure. In this paper, we employ the GIST [36] to achieve the image representation of our mixed-modality dataset due to its capability of capturing both the texture information and the spatial layout.

First, the input image is segmented by a k by k grid. Then 32 Garbor filters are employed at 4 scales, 8 orientations, producing 32 feature maps on every grid. Then the Gist descriptor of an input image can be obtained by concatenating the averaged values of all 32 feature maps, results a vector of length $l = 32k^2$. For a dataset of mixed-modality containing the short-axis view in MR cardiac images I_{M_1} and CT cardiac images I_{M_2}, their representation can be denoted as $X_{M_1} = \{X_{M_1}^{1,1} \dots X_{M_1}^{a_1,b_1}\} \in \mathcal{R}^{a_1 b_1 \times l}$ and $X_{M_2} = \{X_{M2}^{1,1} \dots X_{M2}^{a_2,b_2}\} \in \mathcal{R}^{a_2 b_2 \times l}$, respectively.

B. MODALITY ADAPTATION OF MIXED-MODALITY IMAGES

Given the observation that cardiac images of different modalities share similar myocardium shape and show different texture information, we propose a shared matrix decomposition of the MR and CT representations X_{M_1} and X_{M_2} to improve their shape consistency while reducing the inconsistent texture information.

$$\underset{W,U_{M_1},U_{M_2}}{\arg\min} \|X_{M_1} - U_{M_1} W\|_F^2 + \|X_{M_2} - U_{M_2} W\|_F^2, \quad (1)$$

where U_{M_1}, U_{M_2} are new representations of X_{M_1}, X_{M_2}, and rows of $W \in \mathcal{R}^{l' \times l}$ acts as the bases in the transformed space, like dictionary in sparse coding, DCT basis, wavelets, etc. To ensure the uniqueness of the decomposition, each row of W is conditioned as $\|W_i\| = 1, i = 1, 2, \dots, l'$. In the following, we aim at obtaining a robust shape-aware transformation matrix W, which is capable of capturing the common shape characteristic of MR and CT images, while eliminating the different texture information.

1) SHAPE REPRESENTATION WITH GRAPH SPECTRUM

We first extract the shape information of each modality with graph spectrum, due to its role of identifying the boundary of different parts in image segmentation. For an image I, we can construct a graph $G = (V, \mathcal{E})$, where V is the set of nodes, i.e., features in each position of the image, $\mathcal{E}$ denotes the connectivity of the graph with an adjacent matrix S, and $S(i,j)$ is the similarity measurement between features of two positions. S is computed as:

$$S(i,j) = h(f(I(i)), f(I(j))) \quad (2)$$

where f extracts the image feature in each position, and $h()$ computes the similarity between two features and can be implemented as Gaussian kernels or inner product. The spectrum of a graph is by definition the spectrum of the Laplacian matrix:

$$\begin{aligned} L &= D - S, \\ D &= diag(d_1, d_2, \dots, d_n) \end{aligned} \quad (3)$$

with $d_i = \sum_j S(i,j)$. The eigenvectors E of L contain shape information of objects from large to small size, as indicated by the nonzero eigenvalues in ascending order.

2) SHAPE-AWARE MODALITY ADAPTATION

To extract the common shape information of two modalities, the shape information of each modality is first obtained from

the average adjacent matrix of all images:

$$S_M(i,j) = \frac{1}{num}\sum_r h(f(I_M^r(i)), f(I_M^r(j))), \quad M \in \{M_1, M_2\} \tag{4}$$

where r index the image in one modality, the num index the number of images in one modality($num = a_1b_1$ for first modality M_1 or $num = a_2b_2$ for second modality M_2). If inner product is used for h and mean-removed GIST feature is used for f, then the adjacent matrix becomes the covariance matrix of the image representations.

$$S_M = (X_M - \bar{X}_M)^T(X_M - \bar{X}_M) \tag{5}$$

where $\bar{X}_M$ is the mean GIST vector of all images in modality M. Then we have

$$L_M = D_M - S_M, \quad M \in \{M_1, M_2\} \tag{6}$$

Therefore, the global shape information of M_1 and M_2 can be obtained from the eigenvectors of L_{M_1} and L_{M_2}, which are denoted as E_{M_1} and E_{M_2}.

We formulate extraction of common shape information from two modalities as optimizing W to minimize the shape difference of between of M_1 and M_2:

$$\arg\min_W \|W(E_{M_1} - E_{M_2})\|_F^2 \tag{7}$$

Besides, we also make the dimension of the adapted representation l' less than l to eliminate irrelevant texture information.

C. SHAPE-AWARE FEATURE ENHANCEMENT

To model the relationship among different images, and enhance the shape awareness of the feature, a further SFE for the modality-adapted representations U_{M_1}, U_{M_2} is proposed. The underlying assumption is that if image I^i and I^j have similar shapes $\mathbf{y}^i$, $\mathbf{y}^j$, then their representations should also be close to each other. In this paper, we model the mutual relationship among different samples by constructing a second proximity graph $G^{shape} = (V^{shape}, \mathcal{E}^{shape})$ based on the ground truth shapes of all mixed modality images. Then our shape-aware feature enhancement is modeled as:

$$\arg\min_U \sum_{m,n} ||U^m - U^n||_F^2 S_{m,n}^{shape} \tag{8}$$

Here $U = \{U_{M_1}, U_{M_2}\}$ is the combination of the new representation for the two modalities. The $S_{m,n}^{shape}$ is the similarity of two different shapes $\mathbf{y}^m$ and $\mathbf{y}^n$, which is calculated by the Gaussian kernel.

$$S_{m,n}^{shape} = exp(\frac{-||\mathbf{y}^m - \mathbf{y}^n||^2}{2\sigma^2}) \tag{9}$$

for $m, n \in \{1, \ldots a_1b_1 + a_2b_2\}$.

D. SOLUTIONS VIA OPTIMIZATION

We integrated the Eqs. 1,7, and 8 to obtain the final objective function of our shape-aware modality adaptation:

$$\begin{aligned}\arg\min_{W,U_{M_1},U_{M_2}} & \frac{\|X_{M_1} - U_{M_1}W\|_F^2}{a_1b_1} + \frac{\|X_{M_2} - U_{M_2}W\|_F^2}{a_2b_2} \\ & + \alpha\|W(E_{M_1} - E_{M_2})\|_F^2 + \beta\sum_{m,n}||U^m - U^n||_F^2 S_{m,n}^{shape}\end{aligned} \tag{10}$$

The α and β are the regularization parameter($\alpha + \beta = 1$). When $l' < l$, the first two terms are the reconstruction error of low rank decomposition. Assume that rows of W are orthogonal to each other, i.e, $WW^T \approx I$, then we have $U_{M_1} = X_{M_1}W^T$, $U_{M_2} = X_{M_2}W^T$. The function in Eq. 10 becomes:

$$\begin{aligned}&\arg\min_{W,WW^T=I} \alpha\|W(E_{M_1} - E_{M_2})\|_F^2 \\ &\quad + \beta\sum_{m,n}||(X^m - X^n)W^T||_F^2 S_{m,n}^{shape} \\ &\quad = Tr(W(\alpha\Delta E\Delta E^T + \beta X^T L^{shape} X)W^T))\end{aligned} \tag{11}$$

where $Tr()$ computes the trace of a matrix, $\Delta E = E_{M_1} - E_{M_2}$, $X = \{X_{M_1}, X_{M_2}\}$, and L^{shape} is the Laplacian matrix of G^{shape}, i.e, $L^{shape} = D^{shape} - S^{shape}$.

We adopt the truncated Singular Value Decomposition (SVD) to solve this minimization problem. The pseudocode for our modality adaptation algorithm is illustrated in Algorithm 1.

Algorithm 1 Modality Adaptation Shape Regression

Input: Mixed modality images I_{M1} and I_{M2}, and their label Y_{M1} and Y_{M2};

Output: A robust shape-aware transformation matrix W, New feature matrix U_{M1} and U_{M2} of mixed modality images I_{M1} and I_{M2};

1: Compute the image representation X_{M1} and X_{M2} of the images I_{M1} and I_{M2} using GIST as the description in Section II.C.
2: Construct the objective function (10) and calculate a robust shape-aware transformation matrix W by solving the objective function (10) using (11)
3: Compute the new feature matrix U_{M1} and U_{M2} as $U_{M_1} = X_{M_1}W^T$ and $U_{M_2} = X_{M_2}W^T$.

E. SHAPE REGRESSION OF LEFT VENTRICLE

We formulated the left ventricle segmentation as a shape regression problem. Given the set of feature U_M extracted by the proposed MA-Shape and the output Y_M with dimensionality $P = N1 + N2$, we use the MSVR for the shape regression of left ventricle by finding the regressors $\mathbf{w}^{\mathbf{v}}$ and $\mathbf{z}^v$ for every output that minimize the following function,

$$L_p(w,z) = \frac{1}{2}\sum_{v=1}^{P}||w^v||^2 + C\sum_{g=1}^{o}L_{u_g} \tag{12}$$

where

$$\begin{aligned} \mathbf{z} &= [z^1, \ldots, z^P]^T. \\ u_g &= ||e_g|| = \sqrt{e_g^T e_g} \\ \mathbf{e}_g^T &= \mathbf{y}_g^T - \varphi(x_g)^T w - \mathbf{z}^T \end{aligned} \tag{13}$$

The function $\varphi()$ represents a nonlinear transformation to the feature space, C represents a hyper parameter which determines the trade-off between the regularization and the error reduction term. The ϵ-insensitive quadratic loss function $L(u)$ is defined as:

$$L(u) = \begin{cases} 0, & u < \varepsilon \\ u^2 - 2u\varepsilon + \varepsilon^2, & u \geq \varepsilon \end{cases} \tag{14}$$

III. EXPERIMENTS AND ANALYSIS

A. DATASET

The proposed modality adaptation of shape regression has been evaluated on the mixed-modality cardiac images. The mixed-modality cardiac images contain 2D short-axis cine MR and CT cardiac images. Cardiac MR images are collected from three hospitals affiliated with two health care centers(London Health Care Center and St.Josephs Health Care). It contains 2900 images, which is collected from 145 subjects, each subject has 20 images throughout a cardiac cycle. The average age of all the subjects is 58.9 years old(range from 16 yes to 97 yes). The average pixel spacings of the MR images is 1.5625 mm/pixel (range from 0.6836mm/pixel to 2.0833 mm/pixel). The cardiac CT images are collected from St.Josephs Health Care. It contains 288 images, which is collected from 96 subjects, each subject has three images throughout a cardiac cycle. The pixel spacing of the CT images range from 0.5456 mm/pixel to 0.5968 mm/pixel). All the cardiac MR and CT images are aligned with the size of 80×80 by the following preprocessing step, labeling, ROI extracting and resizing. The ground truth labeled by manual containing the boundary of endocardium and epicardium is double checked by two radiologists.

B. QUANTITATIVE METRICS FOR EVALUATION

The pixel-wise error of epicardium (E_{epi}), endocardium (E_{endo}) and dice metric (DM) are employed to evaluate the performance of our method.

The pixel-wise error is the distance between the estimated boundary and the ground truth which can be defined as:

$$E_{epi/edno} = \frac{1}{pq} \sum_{s=1}^{p} \sum_{t=1}^{q} ||Y_{st} - Y_{st}^{Pre}||; \tag{15}$$

where the p is the number of subjects and the q is the number of images in one subjects. The Y^{Pre} is the estimated boundary, and the Y is the ground truth.

The dice metric is used to measure the overlap between the ground truth and predicted left ventricle area, it can be defined as:

$$DM = \frac{1}{pq} \sum_{s=1}^{p} \sum_{t=1}^{q} \left(\frac{2|A_{st} \bigcap G_{st}|}{|A_{st}| + |G_{st}|}\right); \tag{16}$$

where A is the estimated left ventricle, and G is the ground truth.

C. CONFIGURATIONS

To test the proposed method in a mixed-modality scenario, we employ a modified 5-fold cross-validation scheme under the context of mixed modality: the subject of the first modality (M_1) is split into training set (80%) and test set (20%), and the subject of the second modality M_2 is added into the training set to help during model training. When a single modality scenario is used, the standard 5-fold cross-validation based on subjects is employed.

D. PERFORMANCE OF MA-SHAPE AND EFFECTIVENESS OF MODALITY-ADAPTION

In this section, we evaluate the performance of MA-Shape for the left ventricle segmentation on mixed-modality images and the effectiveness of the proposed modality-adaption module. To this end, LV segmentation performance is examined on three different settings: single modality, mixed-modality without MA, mixed-modality with MA (MA-shape). TABLE 1 and TABLE 2 show the results when M_1 is CT and MR, respectively. Fig.4 and Fig.5 show the bar plots of the shape regression error for both cases. It can be drawn from these two tables that: 1) MA-Shape delivers excellent performance for LV segmentation in for both CT ($E_{epi} = 1.74 \pm 0.96$ *mm*, $E_{endo} = 1.46 \pm 0.99$ *mm* and $DM = 91.1\%$) and MR ($E_{epi} = 2.22 \pm 0.97$ *mm*, $E_{endo} = 1.54 \pm 0.94$ *mm* and $DM = 90.9\%$) images; 2) Using dataset of a second modality without modality-adaption leads to obvious decrement of segmentation performance for both CT and MR images; 3) The proposed modality-adaption module effectively helps improve the segmentation performance by using dataset of M_2 and extracting the common shape features

TABLE 1. The performance of MA-Shape on mixed-modality images, and comparison with settings of single modality and mixed-modality without modality adaption. The M_1 is CT in mixed-modality scenario.

	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Single-modality	2.35 ± 1.17	1.85 ± 1.19	89.2
Without MA	2.61 ± 1.44	2.44 ± 1.69	86.4
MA-Shape	1.74 ± 0.96	1.46 ± 0.99	91.1

TABLE 2. The performance of MA-Shape on mixed-modality images, and comparison with settings of single modality and mixed-modality without modality adaption. The M_1 is MR in mixed-modality scenario.

	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Single-modality	2.67 ± 1.24	1.77 ± 1.05	89.7
Without MA	2.95 ± 1.44	1.90 ± 1.12	89.1
MA-Shape	2.22 ± 0.97	1.54 ± 0.94	90.9

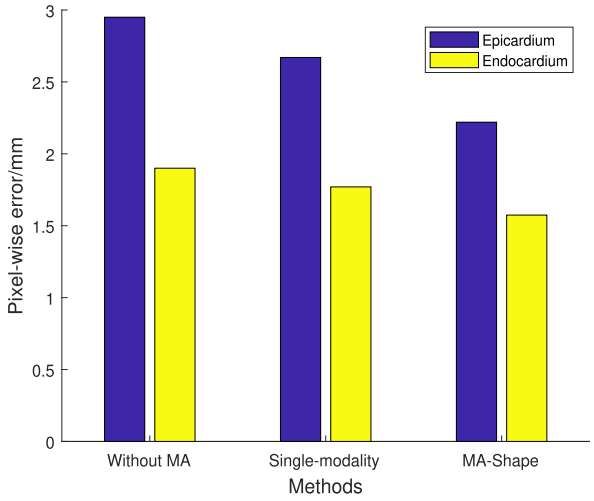

FIGURE 4. The shape regression error of MA-Shape on mixed-modality images, and comparison with settings of single modality and mixed-modality without modality adaption. The M_1 is MR in mixed-modality scenario.

from the two modality. Visualization of the segmentation results can be seen in Fig. 6(a,b,c) and Fig. 7(a,b,c). Obviously, the MA-Shape which focuses on the shape similarity of mixed-modality images can assist well the left ventricle segmentation.

E. IMPROVEMENT OF ACROSS MODALITIES GENERALIZATION FOR SINGLE MODALITY METHOD

In this section, the benefit of the proposed MA module is demonstrated regarding across-modality generalization: the M_1 is used for model training, and the M_2 is used for model testing. TABLE 3 and TABLE 4 show the results

TABLE 3. Cross-modality generalization from MR to CT with the proposed modality adaption module. The proposed modality-adaption module greatly improves the generalization of the segmentation method from MR to CT.

Cross-modality	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Without MA	4.25 ± 1.75	3.64 ± 1.84	82.0
With MA	2.71 ± 1.32	2.34 ± 1.56	87.3

TABLE 4. Cross-modality generalization from CT to MR with the proposed modality adaption module.The proposed modality-adaption module greatly improves the generalization of the segmentation method from CT to MR.

Cross-modality	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Without MA	3.58 ± 1.13	3.43 ± 1.52	80.9
With MA	2.37 ± 1.09	2.06 ± 1.32	87.9

when M_1 is CT and MR, respectively. It can be drawn that in the cross-modality scenario, the proposed modality-adaption module greatly improves the generalization of the segmentation method from MR to CT or from CT to MR. Fig.8 and Fig.9 show the bar plots of the shape regression error for both cases. When generalization from MR to CT, modality adaption helps reduce the error of epicardium and endocardium by 36.24% and 35.71%, and improve the dice metric with 5.3 percents. When generalization from CT to MR, modality adaption helps reduce the error of epicardium and endocardium by 33.80% and 39.94%, and improve the dice metric with 7 percents.

F. ROLE OF SHAPE-AWARE FEATURE ENHANCEMENT

The benefit of the SFE is also validated in the mixed-modality scenario by comparing MA-shape with its variant where the

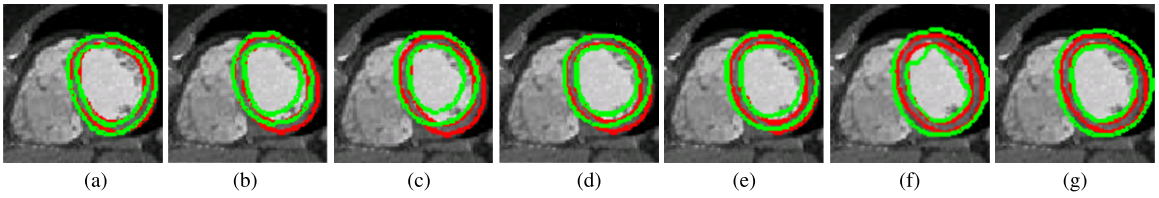

FIGURE 5. Visualization results of different methods on CT images, MA-Shape, single-modality (CT), Mixed-modality without MA, Mixed-modality without SFE, SDL, Level Set and U-Net. The green curves are the automated segmentation results of different method, and the ground truth are depicted by red curves. (a) MA-Shape. (b) Single-modality. (c) Without MA. (d) Without SE. (e) SDL. (f) Level Set. (g) U-Net.

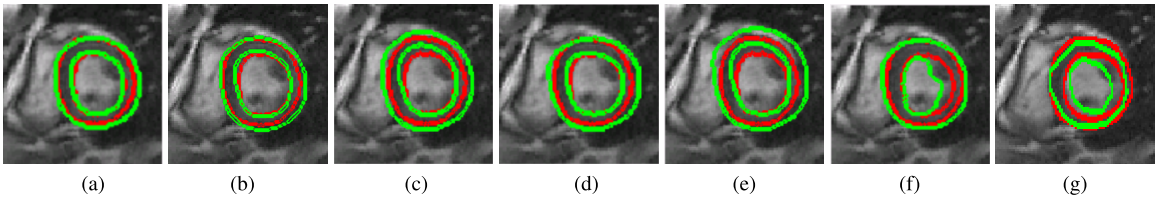

FIGURE 6. Visualization results of different methods on MR images, MA-Shape, Single-modality(MR), Mixed-modality without MA, Mixed-modality without SFE, SDL, Level Set and U-Net. The green curves are the automated segmentation results of different method, and the ground truth are depicted by red curves. (a) MA-Shape. (b) Single-modality. (c) Without MA. (d) Without SE. (e) SDL. (f) Level Set. (g) U-Net.

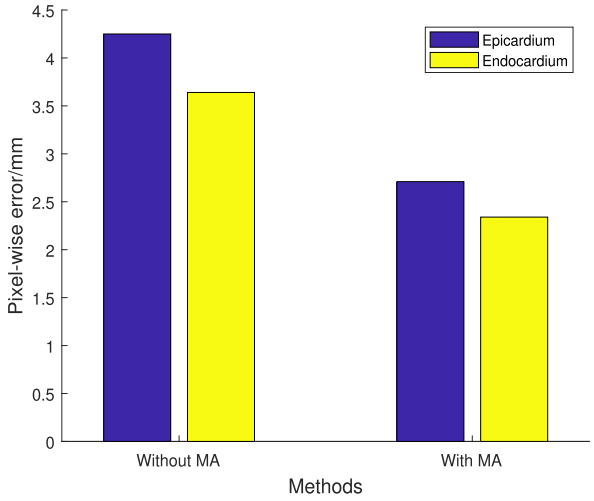

FIGURE 7. Cross-modality generalization from MR to CT with the proposed modality adaption module. The proposed modality-adaption module greatly improves the generalization of the segmentation method from MR to CT.

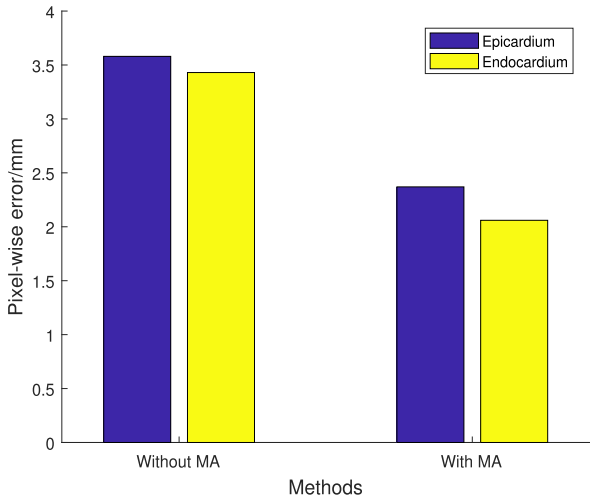

FIGURE 8. Cross-modality generalization from CT to MR with the proposed modality adaption module. The proposed modality-adaption module greatly improves the generalization of the segmentation method from CT to MR.

TABLE 5. The performance of and shape-aware feature enhancement on CT images. It can further improve the segmentation performance on CT images.

	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Without MA	2.61 ± 1.44	2.44 ± 1.69	86.4
Without SFE	2.08 ± 1.09	1.84 ± 1.13	89.4
MA-Shape	1.74 ± 0.96	1.46 ± 0.99	91.1

SFE module is removed. TABLE 5 and TABLE 6 show the results when M_1 is CT and MR, respectively. Fig.10 and Fig.11 show the bar plots of the shape regression error for both cases. It's obvious that after modality adaption, the SFE module can further improve the segmentation performance. The segmentation results are visualized in Fig.6(a,c,d) and Fig.7(a,c,d).

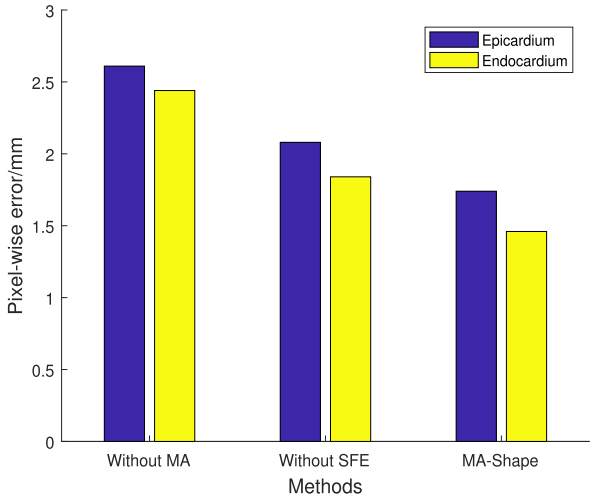

FIGURE 9. The performance of SFE on CT images. It can further improve the segmentation performance on CT images.

TABLE 6. The performance of and shape-aware feature enhancement on MR images. It can further improve the segmentation performance on MR images.

	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Without MA	2.95 ± 1.44	1.90 ± 1.12	89.1
Without SFE	2.37 ± 1.03	1.83 ± 1.20	89.5
MA-Shape	2.22 ± 0.97	1.54 ± 0.94	90.9

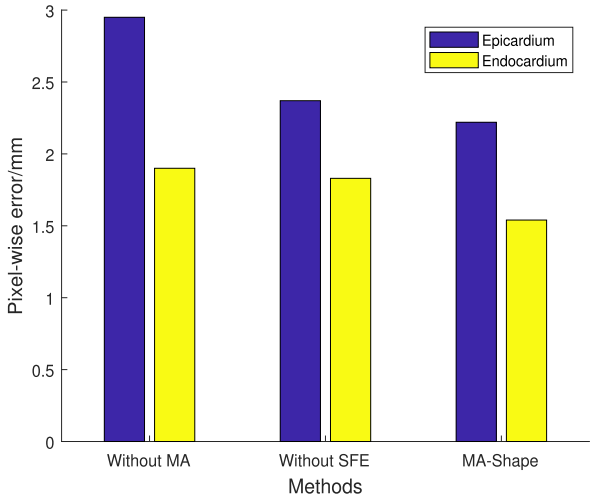

FIGURE 10. The performance of SFE on MR images. It can further improve the segmentation performance on MR images.

G. COMPARISON WITH EXISTING METHODS

The proposed MA-Shape outperforms existing methods on LV segmentation. The level-set based segmentation method, the shape regression-based method employing SDL and U-Net are used for comparison. Table 7 and 8 demonstrate the results when M1 is CT and MR, respectively. Fig.12 shows the dice metric of MA-Shape comparing with SDL, Level set and U-Net on CT images and MR images. MA-shape outperforms level set and U-Net regarding dice metric, and outperforms the direct shape regression method concerning endocardium and epicardium error and dice metric,

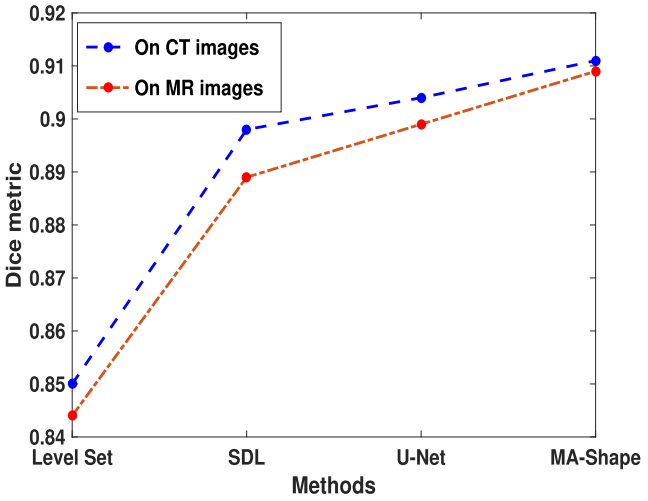

FIGURE 11. The performance of MA-Shape comparing with SDL, Level Set and U-Net on CT images and MR images. MA-shape achieves the best performance than the SDL, Level Set and U-Net both on the CT images and MR images.

TABLE 7. Comparison between MA-Shape and existing segmentation methods on CT images. MA-shape achieves the best performance than the SDL, Level Set and U-Net on the CT images.

	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Level Set	–	–	85.0
SDL	2.27 ± 0.93	1.70 ± 0.86	89.8
U-Net	–	–	90.4
MA-Shape	1.74 ± 0.96	1.46 ± 0.99	91.1

TABLE 8. Comparison between MA-Shape and existing segmentation methods on MR image. MA-shape achieves the best performance than the SDL, Level Set and U-Net on the MR images.

	$E_{epi}(mm)$	$E_{endo}(mm)$	$DM_{LV}(\%)$
Level Set	–	–	84.4
SDL	2.77 ± 1.29	2.05 ± 1.32	88.9
U-Net	–	–	89.9
MA-Shape	2.22 ± 0.97	1.54 ± 0.94	90.9

demonstrating the effectiveness of the proposed MA-shape. The segmentation results are visualized in Fig.6(a,e,f,g) and Fig.7(a,e,f,g).

IV. CONCLUSION

In this paper, we proposed a modality adaptation shape regression method (MA-Shape) for LV segmentation on mixed-modality images. The LV shape was estimated by regression from image representations, which underwent a newly-designed modality adaption module and a supervised feature enhancement module. The MA module effectively improved the performance of LV shape regression and the cross-modality generalization of MA-shape by extracting common and intrinsic shape feature in each modality. The SFE module further brought error reduction by enhancing the expressiveness in the image representation. Experimental results on a mixed-modality database with MR sequences 145 subjects and CT scans of 96 subjects showed that our method could achieve high accuracy for left ventricle segmentation on both of the CT and MR images and lead to better cross-modality generalization. Besides, MA-Shape outperformed its competitors, the shape regression method SDL and the segment method Level set. These advantages reveal the efficiency of utilizing training data of mixed-modalities during model learning and the flexibility of cross-modality deployment in the practical application of the proposed MA-shape.

ACKNOWLEDGMENT

The authors would like to thank G. Ding at Yantai Yuhuangding Hospital for the invaluable suggestions regarding ventricle anatomies.

REFERENCES

[1] E. J. Benjamin *et al.*, "Heart disease and stroke statistics—2017 update: A report from the American Heart Association," *Circulation*, vol. 135, no. 10, pp. e146–e603, 2017.

[2] W.-Y. Low, Y.-K. Lee, and A. L. Samy, "Non-communicable diseases in the Asia-Pacific region: Prevalence, risk factors and community-based prevention," *Int. J. Occupational Med. Environ. Health*, vol. 28, no. 1, pp. 20–26, 2015.

[3] W. Xue, A. Lum, A. Mercado, M. Landis, J. Warringto, and S. Li. (2017). "Full quantification of left ventricle via deep multitask learning network respecting intra- and inter-task relatedness." [Online]. Available: https://arxiv.org/abs/1706.01912

[4] W. Xue, I. B. Nachum, S. Pandey, J. Warrington, S. Leung, and S. Li, "Direct estimation of regional wall thicknesses via residual recurrent neural network," in *Proc. Int. Conf. Inf. Process. Med. Imag.*, 2017, pp. 505–516.

[5] W. Xue, A. Islam, M. Bhaduri, and S. Li, "Direct multitype cardiac indices estimation via joint representation and regression learning," *IEEE Trans. Med. Imag.*, vol. 36, no. 10, pp. 2057–2067, Oct. 2017.

[6] W. Xue, G. Brahm, S. Pandey, S. Leung, and S. Li, "Full left ventricle quantification via deep multitask relationships learning," *Med. Image Anal.*, vol. 43, pp. 54–65, Jan. 2017.

[7] A. Goshtasby and D. A. Turner, "Segmentation of cardiac cine MR images for extraction of right and left ventricular chambers," *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 56–64, Mar. 1995.

[8] W. Bai, W. Shi, C. Ledig, and D. Rueckert, "Multi-atlas segmentation with augmented features for cardiac MR images," *Med. Image Anal.*, vol. 19, no. 1, pp. 98–109, 2015.

[9] 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, 2013.

[10] N. Paragios, "A variational approach for the segmentation of the left ventricle in cardiac image analysis," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 345–362, 2002.

[11] J. Woo, P. J. Slomka, C.-C. J. Kuo, and B.-W. Hong, "Multiphase segmentation using an implicit dual shape prior: Application to detection of left ventricle in cardiac MRI," *Comput. Vis. Image Understand.*, vol. 117, no. 9, pp. 1084–1094, 2013.

[12] J. Schaerer, C. Casta, J. Pousin, and P. Clarysse, "A dynamic elastic model for segmentation and tracking of the heart in mr image sequences," *Med. Image Anal.*, vol. 14, no. 6, pp. 738–749, 2010.

[13] J. Ringenberg *et al.*, "Automated segmentation and reconstruction of patient-specific cardiac anatomy and pathology from *in vivo* MRI," *Meas. Sci. Technol.*, vol. 23, no. 12, p. 125405, 2012.

[14] D. Wei, Y. Sun, S.-H. Ong, P. Chai, L. L. Teo, and A. F. Low, "Three-dimensional segmentation of the left ventricle in late gadolinium enhanced MR images of chronic infarction combining long- and short-axis information," *Med. Image Anal.*, vol. 17, no. 6, pp. 685–697, 2013.

[15] F. Khalifa, G. Beache, G. Gimel'farb, G. A. Giridharan, and A. El-Baz, "Accurate automatic analysis of cardiac cine images," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 445–455, Feb. 2012.

[16] J.-K. Koo, B.-S. Sohn, and B.-W. Hong, "Segmentation of left ventricle in cardiac MRI via contrast-invariant deformable template," *J. Med. Imag. Health Inform.*, vol. 7, no. 8, pp. 1682–1688, 2017.

[17] M. R. Kaus, J. von Berg, J. Weese, W. Niessen, and V. Pekar, "Automated segmentation of the left ventricle in cardiac MRI," *Med. Imag. Anal.*, vol. 8, no. 3, pp. 245–254, Sep. 2004.

[18] S. F. Roohi and R. A. Zoroofi, "4D statistical shape modeling of the left ventricle in cardiac MR images," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 8, no. 3, pp. 335–351, 2013.

[19] K. Lekadir, N. G. Keenan, D. J. Pennell, and G.-Z. Yang, "An inter-landmark approach to 4-D shape extraction and interpretation: Application to myocardial motion assessment in MRI," *IEEE Trans. Med. Imag.*, vol. 30, no. 1, pp. 52–68, Jan. 2011.

[20] S. P. O'Brien, O. Ghita, and P. F. Whelan, "A novel model-based 3D+time left ventricular segmentation technique," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 461–474, Feb. 2011.

[21] X. Qin, Y. Tian, and P. Yan, "Feature competition and partial sparse shape modeling for cardiac image sequences segmentation," *Neurocomputing*, vol. 149, pp. 904–913, Feb. 2015.

[22] 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.

[23] 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, p. 101342Z, Mar. 2017.

[24] P. V. Tran. (2016). "A fully convolutional neural network for cardiac segmentation in short-axis MRI." [Online]. Available: https://arxiv.org/abs/1604.00494

[25] Y. Zhou *et al.*, "Active contours driven by localizing region and edge-based intensity fitting energy with application to segmentation of the left ventricle in cardiac CT images," *Neurocomputing*, vol. 156, pp. 199–210, May 2015.

[26] O. Ecabert *et al.*, "Automatic model-based segmentation of the heart in CT images," *IEEE Trans. Med. Imag.*, vol. 27, no. 9, pp. 1189–1201, Sep. 2008.

[27] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features," *IEEE Trans. Med. Imag.*, vol. 27, no. 11, pp. 1668–1681, Nov. 2008.

[28] N. Dahiya, A. Yezzi, M. Piccinelli, and E. Garcia, "Integrated 3D anatomical model for automatic myocardial segmentation in cardiac CT imagery," in *Proc. Eur. Cong. Comput. Methods Appl. Sci. Eng.* Springer, 2017, pp. 1115–1124.

[29] X. Zhuang *et al.*, "Multiatlas whole heart segmentation of CT data using conditional entropy for atlas ranking and selection," *Med. Phys.*, vol. 42, no. 7, pp. 3822–3833, 2015.

[30] J. Weng, A. Singh, and M. Y. Chiu, "Learning-based ventricle detection from cardiac MR and CT images," *IEEE Trans. Med. Imag.*, vol. 16, no. 4, pp. 378–391, Aug. 1997.

[31] M.-P. Jolly, "Automatic segmentation of the left ventricle in cardiac MR and CT images," *Int. J. Comput. Vis.*, vol. 70, no. 2, pp. 151–163, 2006.

[32] M. Bernier, P.-M. Jodoin, O. Humbert, and A. Lalande, "Graph cut-based method for segmenting the left ventricle from MRI or echocardiographic images," *Comput. Med. Imag. Graph.*, vol. 58, pp. 1–12, Jun. 2017.

[33] A. Sebbahi, A. Herment, A. de Cesare, and E. Mousseaux, "Multimodality cardiovascular image segmentation using a deformable contour model," *Comput. Med. Imag. Graph.*, vol. 21, no. 2, pp. 79–89, 1997.

[34] C. Li, X. Jia, and Y. Sun, "Improved semi-automated segmentation of cardiac CT and MR images," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Jun./Jul. 2009, pp. 25–28.

[35] C. Beitone, C. Tilmant, and F. Chausse, "Mutual cineMR/RT3DUS cardiac segmentation," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 125–128.

[36] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, 2001.

JINYU CONG was born in Shandong, China, in 1991. She received the B.E. and M.E. degrees from the Shandong University of Traditional Chinese Medicine, Jinan, Shandong. She is currently pursuing the Ph.D. degree with Shandong Normal University, Jinan, Shandong.

Her research interests include the medical image processing and machine learning.

YUANJIE ZHENG was a Senior Research Investigator with the Perelman School of Medicine, University of Pennsylvania. He is currently a Professor with the School of Information Science and Engineering, Shandong Normal University, and a Taishan Scholar of People's Government of Shandong Province, China. He is also serving as the Vice Dean of the School of Information Science and Technology and the Institute of Life Sciences, Shandong Normal University. His research interests include medical image analysis, translational medicine, computer vision, and computational photography, enhance patient care by creating algorithms for automatically quantifying and generalizing the information latent in various medical images for tasks, such as disease analysis and surgical planning through the applications of computer vision and machine learning approaches to medical image analysis tasks and the development of strategies for image-guided intervention/surgery.

WUFENG XUE was a Senior Research Investigator with the Perelman School of Medicine, University of Pennsylvania. He is currently a Professor with the School of Information Science and Engineering, Shandong Normal University, and a Taishan Scholar of People's Government of Shandong Province, China.

BOFENG CAO is currently pursuing the Ph.D. degree with the Medical Imaging Department, Yantai Yuhuangding Hospital, Yantai, China.

SHUO LI received the Ph.D. degree in computer science from Concordia University, Montreal, QC, Canada, in 2006. He is currently an Adjunct Research Professor with Western University and an Adjunct Scientist with the Lawson Health Research Institute. He is also leading the Digital Imaging Group of London as the Scientific Director. His current research interests include automated medial image analysis and visualization.