Computerized Medical Imaging and Graphics 71 (2019) 49–57

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Multi-sequence myocardium segmentation with cross-constrained shape and neural network-based initialization

Jie Liu[a], Hongzhi Xie[b,*], Shuyang Zhang[b], Lixu Gu[a,*]

[a] School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
[b] Department of Cardiology, Peking Union Medical College Hospital, Beijing, China

ARTICLE INFO

Article history:
Received 9 April 2018
Received in revised form
25 September 2018
Accepted 12 November 2018

Keywords:
Cardiovascular magnetic resonance imaging
Multi-sequence analysis
Cross constrained shape
Shape discrepancy compensation
Conditional generative adversarial network

ABSTRACT

For myocardial infarction (MI) patients, delayed enhancement (DE) and T2-weighted cardiovascular magnetic resonance imaging (CMR) can play significant roles in diagnosis, prognosis and therapeutic strategy evaluation. However, the non-rigid registration between different CMR sequences is particularly challenging and prevents the use of multi-sequence image analysis. In this article, we propose an approach for segmenting T2 and DE CMR simultaneously with cross-constrained shape and shape discrepancy compensation. A framework for the unified segmentation of multi-sequence images is built based on the coupled level set method. Additionally, a sparse representation-based shape model is optimized under the constraints from both sequences for complementary information sharing. Considering the myocardium shape discrepancy between the two sequences due to non-perfect registration, an error term is added to explicitly model this difference. The intensity feature is extracted with a Gaussian mixture model from each sequence. To obtain a fully automatic approach, the conditional generative adversarial network is adopted for initialization. The results are evaluated with T2 and DE images from 32 MI patients. A promising Dice similarity coefficient of the myocardium is achieved (84.97±4.15% for T2 and 78.13±6.22% for DE CMR). This approach is a pilot work toward automatic, multi-sequence CMR image analysis.

1. Introduction

With technical advances in hardware and software, cardiovascular magnetic resonance imaging (CMR) is increasingly used in clinical practice (Motwani et al., 2014; Ripley et al., 2016). For patients with myocardial infarction (MI), delayed enhancement (DE) CMR and T2-weighted CMR can play significant roles in diagnosis, prognosis and therapeutic strategy evaluation (Beek and van Rossum, 2010; Kim et al., 2009). The DE CMR is acquired 10–15 min after the intravenous administration of extracellular contrast agents, and irreversible ischemic damage is hyper-enhanced due to delayed clearance of the contrast agents. The sensitivity and accuracy of DE CMR for the detection and visualization of MI have been well validated (Amado et al., 2004; Rehwald et al., 2002; Wagner et al., 2003). Acute MI also leads to intra- and extracellular edema in the myocardium, which will cause prolongation of the T2 relaxation time. These edema regions appear bright in T2-weighted CMR and may have a correlation with the area at risk (Eitel et al., 2010; Friedrich et al., 2008). Some examples of DE and T2 images are shown in Fig. 1 to illustrate the characteristics of the two sequences.

For quantitative analysis of the MI region, segmentation of the myocardium is a prerequisite. Myocardium segmentation is mostly performed manually in clinical practice, which is time consuming and suffers from inter-observer variations. Hence, automatic myocardium segmentation is desired. However, three issues related to the intensity characteristics challenge robust segmentation of the myocardium from CMR of MI patients, especially DE CMR. Firstly, the intensity of the myocardium is heterogeneous due to the existence of MI in DE CMR and edema in T2 CMR. The assumption of a simple myocardium intensity distribution (e.g., a Gaussian distribution) (Queirós et al., 2014; Wang et al., 2015), does not fit this application. Secondly, the DE CMR has low intensity contrast between the myocardium and the surrounding tissue. For example, the intensity of the enhanced infarcted region is similar to the blood pool, and the intensity range of the suppressed normal myocardium overlaps that of the liver or lung. This arrangement leads to indistinguishable myocardium boundaries, and the methods that mainly rely on the local texture are challenged (Bai et al.,

Abbreviations: cGAN, conditional generative adversarial network; CLS, coupled level set; CMR, cardiovascular magnetic resonance imaging; DE, delayed enhancement; DSC, dice similarity coefficient; EM, expectation maximization; GMM, Gaussian mixture model; LV, left ventricle; MI, myocardial infarction.

* Corresponding authors.
E-mail addresses: xiehongzhi@medmail.com.cn (H. Xie), gulixu@sjtu.edu.cn (L. Gu).

https://doi.org/10.1016/j.compmedimag.2018.11.001
0895-6111/

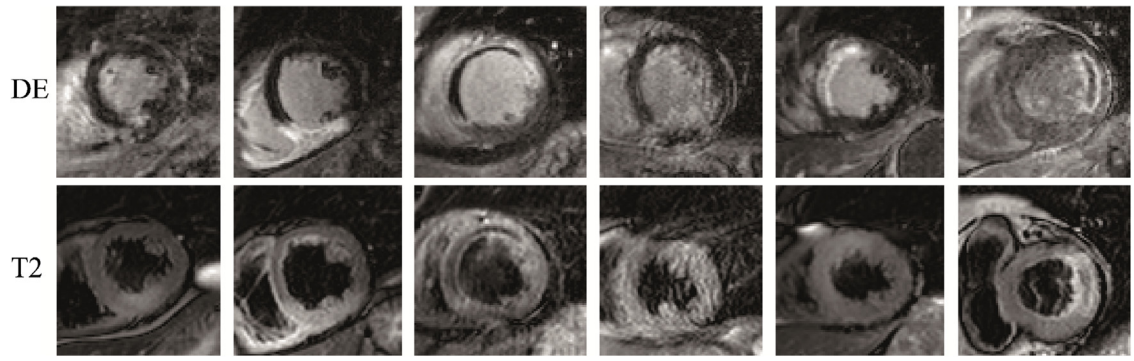

Fig. 1. Corresponding DE and T2 images of six patients. Each column shows a different patient.

2015; Mukhopadhyay et al., 2015; Tan et al., 2017). Finally, the complex patterns of infarction, whose locations and geometrics are difficult to define, also increase the difficulty of segmentation. A case of microvascular obstruction can also occur, where a hypo-enhanced region is encompassed by a hyper-enhanced region due to the lack of contrast agent uptake (Eitel et al., 2010), as shown in the third and sixth columns in Fig. 1. These intensity characteristics of DE and T2 CMR are challenges for the myocardium segmentation methods that work well in other modalities, such as computed tomography (CT) or other CMR sequences (e.g., cine CMR).

To the best of our knowledge, few studies have focused on myocardium segmentation in T2 CMR, and a few DE CMR approaches have been proposed. Some studies directly perform myocardium segmentation in DE CMR (Kurzendorfer et al., 2017; Liu et al., 2017). In the work of Kurzendorfer et al. (2017), the blood pool is approximated using circular Hough transforms and a morphological active contours approach. Then, the endocardium and epicardium are refined in polar space while considering the edge information and scar distribution. Liu et al. (2017) proposed a framework integrated coupled level set (CLS) and expectation maximization (EM). However, the lack of important information in a single sequence of DE CMR, such as the distinguishable boundaries of the myocardium, prevents further improvement of the segmentation accuracy. Most existing approaches make use of the cine CMR that is acquired in the same session (El Berbari et al., 2009). This approach is a reasonable choice when integrating information from different sequences, since the cine sequence is widely used in CMR examinations, and many segmentation methods have been proposed (Avendi et al., 2016; Hajiaghayi et al., 2017; Petitjean and Dacher, 2011). However, instead of a multi-sequence analysis, these studies take the segmentation of cine MRI as *a priori* knowledge, which is propagated and deformed toward myocardium contours in target DE images (Ciofolo et al., 2008; Merino-Caviedes et al., 2016; Tao et al., 2015; Wei et al., 2013). This type of approach is prone to error accumulation, and thus human supervision is usually needed for segmentation of the cine CMR and registration between the cine and DE images. Few recent studies have attempted multi-sequence segmentation (Liu et al., 2014; Zhuang, 2016). This type of method best makes use of the complementary information from multiple sequences. However, different sequences are assumed to be perfectly pre-registered (Liu et al., 2014) or only misalignments of the slice and the whole image volume are considered (Zhuang, 2016).

In clinical applications, the heart undergoes complex non-rigid deformation during the CMR scan. The unneglectable changes in the heart shape between two acquisitions can be caused by inevitable spatial and temporal variations, such as movement of the patient, imperfect ECG gating, and differences in the thoracic air volume among several breath holdings. Additionally, different MRI sequences have different intensity characteristics and appearances, especially in the pathological region. Thus, non-rigid registration between different CMR sequences is particularly challenging, and changes in the heart shape cannot be perfectly captured by pre-registration. Therefore, we propose to simultaneously segment T2 and DE CMR for the same patient while explicitly modeling the myocardium shape discrepancy. The major contribution is four-fold. First, a promising framework for addressing multi-sequence CMR segmentation is built. The CLS method is used to build the segmentation framework, and the energy function is carefully designed. Second, the shape of the patient's heart is chosen as information that can be shared across different sequences. A sparse representation-based shape model is optimized under the constraints from both sequences. Third, considering the existence of the myocardium shape discrepancy due to non-perfect alignment of the two sequences, an error term is added in the shape model to explicitly model this difference. Finally, to design a fully automatic approach, the conditional generative adversarial network (cGAN) technique is adopted and trained to generate the initial segmentation.

2. Materials and methods

The proposed framework is illustrated in Fig. 2. In this section, first we provide an introduction to cGAN-based initialization. Then, the sparse representation-based shape model is described in detail. Finally, the overall framework built with the CLS method is introduced. The three terms in the energy function of the CLS are described, especially the shape term, which integrates the shape model and the shape discrepancy compensation. Additionally, the evolution functions of the contours and updating of the sparse coefficient and shape discrepancy compensation term are shown.

2.1. Conditional generative adversarial network-based initialization

In recent years, deep learning is widely used for medical image segmentation, and promising results have been achieved, especially with the fully convolutional network (FCN) (Chen et al., 2017; Çiçek et al., 2016; Roth et al., 2018). However, in the standard training of FCN for semantic segmentation, valuable information about the 'structure' of the predicted label maps has not been used. Instead, the cGAN can learn a structured loss by training a generator and a discriminator simultaneously. The generator is trained to generate a label map condition on an input image, whereas the loss is learned by the discriminator, which attempts to discriminate between generated and manual annotated label maps. Since the joint configuration of the label map is penalized by the discriminator, this method holds advantages for segmentation tasks on complex structured, small datasets (Kohl et al., 2017). In the proposed method, the cGAN is adopted to initial segmentation of the

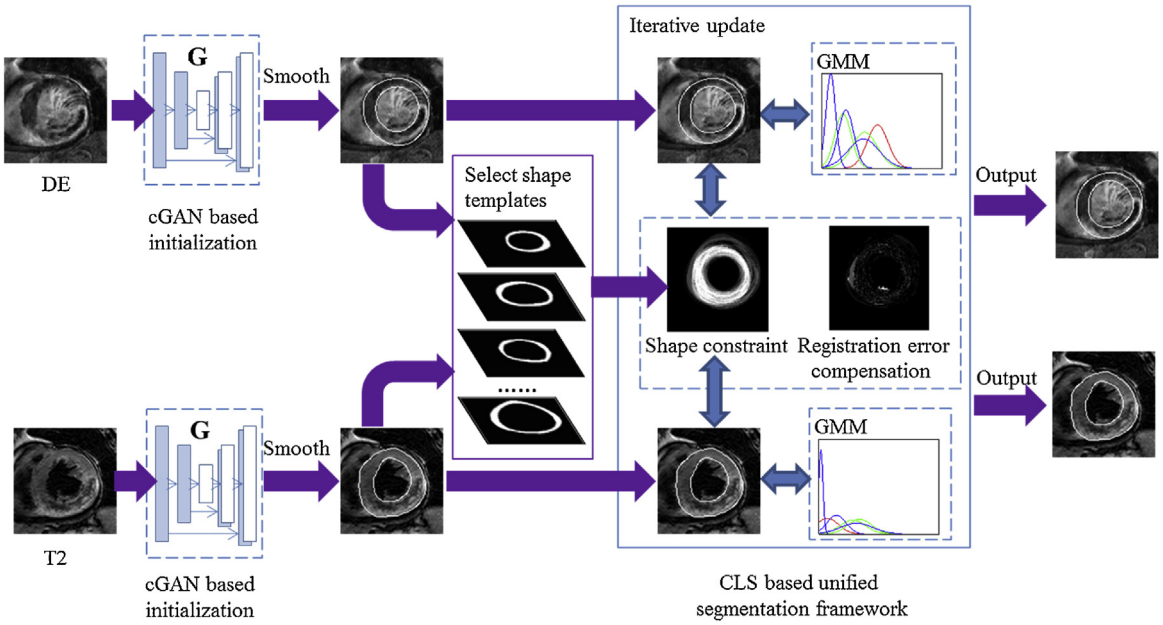

Fig. 2. Flowchart of the proposed framework.

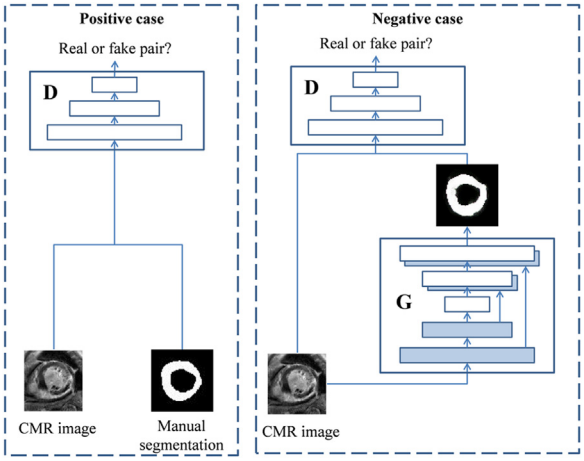

Fig. 3. Illustration of adversarial training for myocardium segmentation. The discriminator D learns to discriminate between generated and manual annotated label maps. The generator G learns to generate a label map which is indistinguishable from manual annotated ones to the discriminator.

myocardium from both the DE and T2 slices. Since the structured loss is learned by the discriminator, the overall myocardium shape information can be encoded in the trained cGAN. The cGAN used in this work is designed and trained in accordance with Isola et al. (2017), where a "U-Net"-based architecture is used as the generator, and a convolutional "PatchGAN" is used as the discriminator, as shown in Fig. 3. With this design, an output with pixel-to-pixel correspondence to the input image can be obtained. Initialization of the DE and T2 sequences are each achieved with a separate cGAN, which is trained with the patient-wise leave-one-out strategy.

Since our data set is relatively small, the cGAN results may not fit well with the elliptical shape of the endo- and epicardium. The elliptic Fourier descriptors (Kuhl and Giardina, 1982) approach is adopted to smooth the initialization, where a closed contour is characterized by calculating the Fourier coefficients and then reconstructing with the reserved number of harmonics equal to 3. Some examples are shown in Fig. 4. The left ventricle (LV) center is estimated with the resultant initialization, which is used to align the corresponding DE and T2 slices as well as the shape templates.

2.2. Sparse representation-based shape model

Since the intensity characteristics are quite different in T2 and DE CMR, we consider shape information to be a superior approach for complementary information sharing between the two sequences. A shape model is adopted to estimate the heart shape of the patients utilizing the prior shape templates. The use of a shape model is also important for achieving robust and reasonable results.

2.2.1. Building a shape template repertory

To build a shape template repertory, we take the manual segmentation of both the T2 and DE slices in the data set as the training data, except for the target patient. We represent each shape template as a binary mask, whose value is set to 1 in the myocardium and 0 elsewhere. The first template is taken as the reference, and the remaining templates are aligned with the center of the LV.

2.2.2. Sparse representation-based shape model

The shape model generates the myocardium probability map by taking the linear combination of a selected subset of the shape templates. When the target slice is given, the center of the LV is estimated with the initial segmentation, and all of the shape templates are aligned to the target slice. To reduce the computational load, a subset of the shape templates is selected according to the initial segmentation using the Dice similarity coefficient (DSC) as the metric. The DSC measures the overlap between shape template S and initialization G as $Dice(S, G) = 2|S \cap G|/(|S| + |G|)$. Templates with a DSC larger than a threshold T are selected and used to construct a target slice-specific shape dictionary $\boldsymbol{D}$, where each column represents a reorganized shape template.

Let $\boldsymbol{y}$ be the column vector of the reorganized myocardium probability map $\boldsymbol{y^I}$. We model the shape probability using a linear combination of the shape templates in $\boldsymbol{D}$ as $\boldsymbol{y} = \boldsymbol{Dw}$. Here, $\boldsymbol{w}$ is the sparse representation coefficient vector. Only shape templates that correspond to nonzero elements in $\boldsymbol{w}$ are activated in the mod-

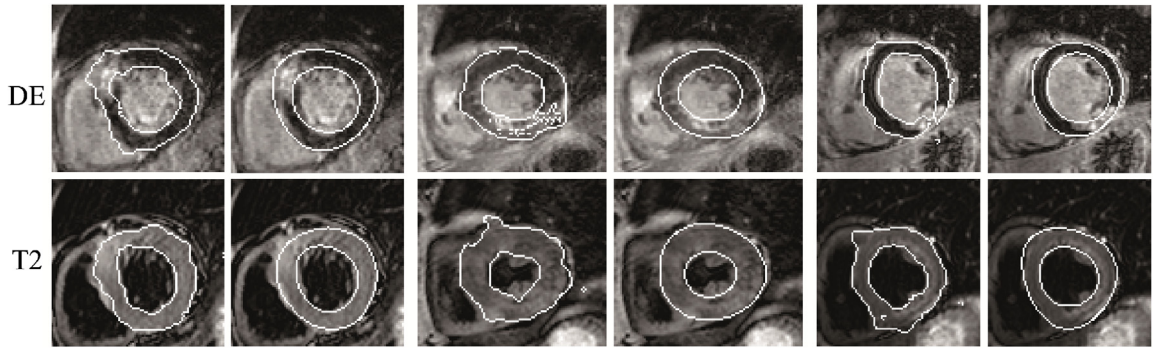

Fig. 4. Results from the conditional generative adversarial networks (cGAN) (columns 1, 3, and 5) and the corresponding results smoothed with elliptic Fourier descriptors (columns 2, 4, and 6). Both columns 1 and 2, 3 and 4, as well as 5 and 6 are from the same patient.

eling of the myocardium shape. This model can provide a realistic shape constraint while preserving the characteristics of each shape template.

2.3. Coupled level set-based framework

Level set method is widely used in medical image segmentation (Chen et al., 2018; Swierczynski et al., 2018), because of its flexibility in designing energy function and ability to handle topology change. We adopt the CLS method to build the segmentation framework. In the proposed method, the energy function of the CLS contains three terms: the shape term, the posterior probability term, and the length term. These three terms integrate the heart shape model and the shape discrepancy compensation, the intensity information in each sequence, and the smoothness constraint of the contours, respectively. The total energy function is as follows:

$$E(\phi_1,\phi_2,\phi_3,\phi_4) = \upsilon E_{shape}(\phi_1,\phi_2,\phi_3,\phi_4) + E_{prob}(\phi_1,\phi_2,\phi_3,\phi_4) + \gamma E_{length}(\phi_1,\phi_2,\phi_3,\phi_4), \quad (1)$$

where ϕ_1 and ϕ_2 are the level set functions that represent the endocardial contour and epicardial contour in the T2 slice, respectively. Additionally, ϕ_3 and ϕ_4 are level set functions of the two contours in the corresponding DE slice. The level set function is defined as positive inside and negative outside. Here, υ and γ are the weighting parameters of the shape and length terms, respectively.

2.3.1. Shape term

To achieve robust and reasonable segmentation results, we adopted the sparse representation-based shape model to estimate the heart shape of the patient, which can impose strong prior shape information on both sequences. In addition, the contours in both sequences are used to update the parameters in the shape model, making the cross-constrained shape a bridge for complementary information sharing. The shape term in the CLS energy function is designed to integrate the shape model and pull the contours of both sequences to fit the myocardium probability map.

$$E_{shape}(\phi_1,\phi_2,\phi_3,\phi_4) = \int_\Omega \left[(H(\phi_2) - H(\phi_1)) - \boldsymbol{y^I}\right]^2 dx + \int_\Omega \left[(H(\phi_4) - H(\phi_3)) - \boldsymbol{y^I}\right]^2 dx + \lambda_1 ||\boldsymbol{w}||_1, \quad (2)$$

where $H(\phi)$ is the Heaviside function, and $H(\phi_2) - H(\phi_1)$ represents the region between the two contours in the T2 slice labeled with 1, which is the myocardium, and the other tissues are labeled with 0. Similarly, $H(\phi_4) - H(\phi_3)$ represents the myocardium in the DE slice. $\boldsymbol{y^I}$ is the shape probability map obtained from the shape model with the parameter vector $\boldsymbol{w}$. The two sequences share the same shape probability map $\boldsymbol{y^I}$. A L1 norm regularization term of the representation coefficient vector $\boldsymbol{w}$ is added to impose sparsity. Here, Ω denotes the image space.

Considering the shape discrepancy between the T2 and DE slices, using the exact same shape constraint for both sequences may not generate good results. Inspired by the work of Zhang et al. (2012), where the errors in the shape caused by occlusion or missing points are modeled with a sparse vector, we explicitly model the shape discrepancy with a shape discrepancy compensation term $\boldsymbol{e^I}$. Then, the shape term in (2) becomes:

$$E_{shape}(\phi_1,\phi_2,\phi_3,\phi_4) = \int_\Omega \left[(H(\phi_2) - H(\phi_1)) - \boldsymbol{y^I}\right]^2 dx + \int_\Omega \left[(H(\phi_4) - H(\phi_3)) - \boldsymbol{y^I} - \boldsymbol{e^I}\right]^2 dx + \lambda_1 ||\boldsymbol{w}||_1 + \lambda_2 ||\boldsymbol{e}||_1, \quad (3)$$

Because the T2 and DE CMR are from the same patient and are taken at a similar phase of the cardiac cycle, we assume that the shape discrepancy is relatively small. The notation $\boldsymbol{e}$ represents a column vector reorganized from $\boldsymbol{e^I}$. Based on the small discrepancy assumption, $\boldsymbol{e}$ should be a sparse vector; thus, a L1 norm regularization of $\boldsymbol{e}$ is added.

2.3.2. Posterior probability and length terms

The posterior probability term is designed to integrate the intensity information from both sequences. This term is separately calculated on the T2 and DE slices. Because the amounts of effective intensity information contained in the two sequences are different, different weighting parameters (μ_1, μ_2) are set:

$$E_{prob}(\phi_1,\phi_2,\phi_3,\phi_4) = \mu_1 E_{probT2}(\phi_1,\phi_2) + \mu_2 E_{probDE}(\phi_3,\phi_4). \quad (4)$$

The intensity distributions of the LV, myocardium and background in a target slice are each modeled with a Gaussian mixture model (GMM), and the numbers of Gaussian distributions are set as 1, 2 and 3, respectively. The parameters in the GMM are initialized as described in Liu et al. (2017) and are updated to fit the intensity distribution of the current labeled pixels of the three classes with the EM algorithm. Then, the posterior probabilities of each class can be calculated for all pixels. The posterior probability term makes use of the log of the posterior probabilities of the LV, myocardium (Myo) and background (BG) and is formularized similar to the method in Duan et al. (2010). The T2 slice is formulated as follows:

$$E_{probT2}(\phi_1,\phi_2) = -\int_\Omega p_{LV,T2} H(\phi_1) H(\phi_2)\, dx - \int_\Omega p_{Myo,T2} [1 - H(\phi_1)] H(\phi_2)\, dx - \int_\Omega p_{BG,T2} [1 - H(\phi_1)][1 - H(\phi_2)]\, dx, \quad (5)$$

where p_L, $L \in \{LV, Myo, BG\}$ is a function of position x in the image space. For pixel i, $p_L(i)$ represents the log of the posterior probability of pixel i belonging to class L. The posterior probability term on the DE slice is the same, except that p_L is calculated with the intensity of the DE slice.

The length term imposes a smoothness constraint on all four contours:

$$E_{length}(\phi_1, \phi_2, \phi_3, \phi_4) = E_{length}(\phi_1) + E_{length}(\phi_2) + E_{length}(\phi_3) + E_{length}(\phi_4), \quad (6)$$

$$E_{length}(\phi_i) = \int_{\Omega} \delta(\phi_i) \left|\nabla\phi_i\right| dx,\ i = 1, 2, 3, 4, \quad (7)$$

where $\delta(\phi)$ is the Dirac delta function.

2.3.3. Updating the parameters

First, with the shape model and shape discrepancy compensation terms are fixed, the four contours ϕ_i, $i = 1, 2, 3, 4$ are evolved according to the corresponding Euler-Lagrange equation:

$$\frac{\partial\phi_1}{\partial t} = 2\upsilon\delta(\phi_1)[(H(\phi_2) - H(\phi_1)) - \boldsymbol{y^I}] + \mu_1\delta(\phi_1)\left\{\left(p_{LV,T2} - p_{Myo,T2}\right)H(\phi_2) - p_{BG,T2}[1 - H(\phi_2)]\right\} + \delta(\phi_1)div\left(\frac{\nabla\phi_1}{\left|\nabla\phi_1\right|}\right), \quad (8)$$

$$\frac{\partial\phi_2}{\partial t} = -2\upsilon\delta(\phi_2)[(H(\phi_2) - H(\phi_1)) - \boldsymbol{y^I}] + \mu_1\delta(\phi_2)\left\{p_{LV,T2}H(\phi_1) - \left(p_{Myo,T2} - p_{BG,T2}\right)[1 - H(\phi_1)]\right\} + \delta(\phi_2)div\left(\frac{\nabla\phi_2}{\left|\nabla\phi_2\right|}\right), \quad (9)$$

$$\frac{\partial\phi_3}{\partial t} = 2\upsilon\delta(\phi_3)\left[(H(\phi_4) - H(\phi_3)) - \boldsymbol{y^I} - \boldsymbol{e^I}\right] + \mu_2\delta(\phi_3)\left\{\left(p_{LV,DE} - p_{Myo,DE}\right)H(\phi_4) - p_{BG,DE}[1 - H(\phi_4)]\right\} + \delta(\phi_3)div\left(\frac{\nabla\phi_3}{\left|\nabla\phi_3\right|}\right), \quad (10)$$

$$\frac{\partial\phi_4}{\partial t} = -2\upsilon\delta(\phi_4)\left[(H(\phi_4) - H(\phi_3)) - \boldsymbol{y^I} - \boldsymbol{e^I}\right] + \mu_2\delta(\phi_4)\left\{p_{LV,DE}H(\phi_3) - \left(p_{Myo,DE} - p_{BG,DE}\right)[1 - H(\phi_3)]\right\} + \delta(\phi_4)div\left(\frac{\nabla\phi_4}{\left|\nabla\phi_4\right|}\right). \quad (11)$$

Then, we fix the contours in the CLS and update the sparse representation coefficient $\boldsymbol{w}$ and shape discrepancy compensation term $\boldsymbol{e}$. Let $\boldsymbol{y}_{T2}^{CLS}$, $\boldsymbol{y}_{T2}^{CLS}$ denote the column vectors reorganized from $H(\phi_2) - H(\phi_1)$, $H(\phi_4) - H(\phi_3)$, respectively. Then, the shape term in the CLS energy function can be recast as:

$$\left\|\boldsymbol{y}_{T2}^{CLS} - \boldsymbol{Dw}\right\|_2^2 + \left\|\boldsymbol{y}_{DE}^{CLS} - \boldsymbol{Dw} - \boldsymbol{e}\right\|_2^2 + +\lambda_1\|\boldsymbol{w}\|_1 + \lambda_2\|\boldsymbol{e}\|_1. \quad (12)$$

When minimizing this term with $\boldsymbol{w}$ and taking $\boldsymbol{e}$ as a fixed parameter, the problem becomes:

$$\min_{\boldsymbol{w}} 2\left\|\frac{1}{2}(\boldsymbol{y}_{T2}^{CLS} + \boldsymbol{y}_{T2}^{CLS} - \boldsymbol{e}) - \boldsymbol{Dw}\right\|_2^2 + \lambda_1\|\boldsymbol{w}\|_1 + const. \quad (13)$$

Similarly, when minimizing with $\boldsymbol{e}$, the problem becomes:

$$\min_{\boldsymbol{e}}\left\|\boldsymbol{y}_{DE}^{CLS} - \boldsymbol{Dw} - \boldsymbol{e}\right\|_2^2 + \lambda_2\|\boldsymbol{e}\|_1 + const. \quad (14)$$

Eq. (13) is treated as the standard LASSO problem, which is solved with a downloaded toolbox (SPAMS: http://spams-devel.gforge.inria.fr/). Since the dimension of the shape discrepancy compensation term $\boldsymbol{e}$ is relatively large, Eq. (14) is solved with the gradient descent method for better efficiency.

The whole framework is solved by alternatively minimizing/descending the total energy function with respect to the level set functions ϕ_i, $i = 1, 2, 3, 4$, the sparse coefficient $\boldsymbol{w}$, the shape discrepancy compensation term $\boldsymbol{e}$, and the parameters in the GMM in the posterior probability term.

The segmentation results from this framework usually contain part of the papillary muscles, which are located close to the myocardium and have a similar intensity. However, for medical diagnostic purposes, the papillary muscles are usually excluded from the myocardium. Thus, we compute the convex hull of the inner contour and perform interpolation to include the papillary muscles in the LV class.

2.4. Experiment

2.4.1. Data

We obtained T2 and DE images from 32 MI patients with approval from the institutional review board. The resolution of the images was 0.7292 mm×0.7292 mm×5 mm for DE CMR and 1.3194 mm×1.3194 mm×12 mm for T2 CMR. In the experiment, the DE images were transformed into the space of the T2 images and down-sampled to the same resolution as the T2 images. The segmentation was performed on paired DE and T2 images from the same patient.

For the T2 and DE images of 32 subjects, manual segmentation was performed by an expert and was taken as the ground truth.

2.4.2. Parameter settings and speed

The proposed framework was tested on all 32 subjects. The whole procedure was fully automatic. Because the slice thickness was large in CMR, the segmentation was performed in a slice-by-slice manner.

The data used to train the cGAN was the same data descripted in Section 2.4.1. Each patient had 3–7 images for both T2 and down-sampled DE sequences, which added up to 134 images for each sequence in total. The original image size was 256×256. To remove black background and some remote organs in each image, the center region with size of 128×128 was cropped. In the training of cGAN, data augmentation was applied by resizing the input images to 286×286 and then randomly cropping back to size 256×256. Random left-right flip was also used. The networks were trained for 200 epochs with batch size 1 using the patient-wise leave-one-out strategy. The Adam optimizer is used with learning rate 0.0002, exponential decay rate 0.5 for the 1 st moment estimates and 0.999 for the 2nd moment estimates.

The parameters used in the presented results were $\upsilon = 5$, $\mu_1 = 1.5$, $\mu_2 = 0.5$, $\gamma = 0.5$, $\lambda_1 = 2$, and $\lambda_2 = 5$. Additionally, the threshold used to select the slice-specific dictionary was set to $T = 0.3$.

In our experiments, a TensorFlow implementation of cGAN was downloaded from the web (https://github.com/affinelayer/pix2pix-tensorflow) and run in single Pascal Titan X GPU. The initial results for each patient were achieved within one second. The rest of the framework was built on Matlab and performed on a personal computer equipped with dual core Intel i5-4200U processor and 4GB RAM. The runtime for the segmentation of T2 and DE images of one subject (including multiple paired slices) was 15.2 ± 4.6 s.

2.4.3. Experimental design

To better demonstrate the effectiveness of unified segmentation in the proposed framework, we performed four comparison methods:

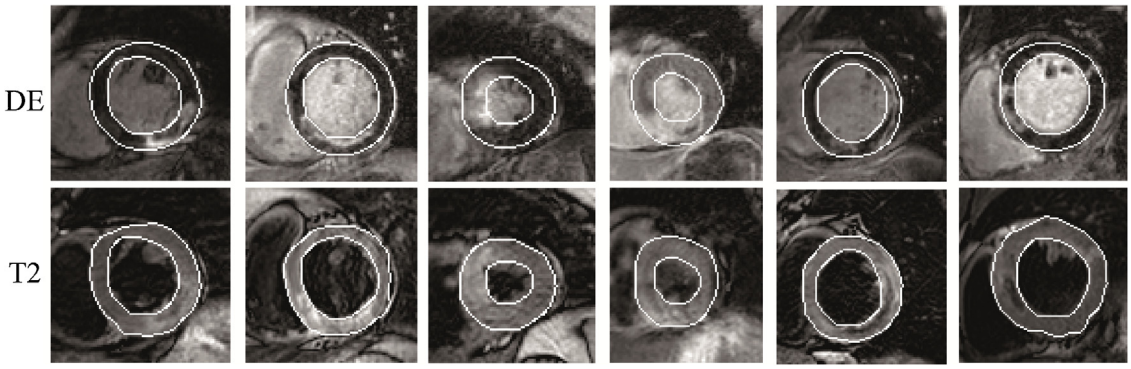

Fig. 5. Segmentation results of the corresponding DE and T2 images. Each column shows a different patient.

Table 1
Myocardium segmentation accuracy evaluated with DSC (%) for the endocardium (endo), myocardium (myo), and epicardium (epi).

		Proposed	Separate	NoDiff	cGAN	uniGAN
DE	Endo	87.91 ± 4.47	87.76 ± 4.50	87.90 ± 4.66	76.41 ± 19.90	80.22 ± 17.16
	Myo	78.13 ± 6.22	74.45 ± 6.28	78.09 ± 6.32	72.50 ± 9.90	73.17 ± 8.50
	Epi	92.62 ± 2.40	91.46 ± 2.41	92.64 ± 2.34	90.95 ± 4.70	90.86 ± 2.91
T2	Endo	90.48 ± 3.82	88.85 ± 4.67	89.59 ± 4.63	86.13 ± 6.54	87.56 ± 4.68
	Myo	84.97 ± 4.15	82.51 ± 4.98	83.26 ± 4.75	82.21 ± 5.05	84.30 ± 4.33
	Epi	93.53 ± 2.22	92.17 ± 2.97	92.64 ± 2.60	92.63 ± 2.23	93.78 ± 1.70

Separate: For the T2 and DE slices, each has a different sparse coefficient $\boldsymbol{w}$ in the shape model, and no shape discrepancy compensation is used. In this method, segmentation of the DE and T2 slices is totally separated, and no shape information is shared across the two sequences.

NoDiff: The T2 and DE slices share the same sparse coefficient $\boldsymbol{w}$, but the shape discrepancy term $\boldsymbol{e}$ is removed from the energy function. In this method, the T2 and DE slices share exactly the same shape constraint.

cGAN: This method is the cGAN-based initialization stage in the proposed method, and we list the results to establish a baseline for comparison. The results are smoothed with elliptic Fourier descriptors. In addition, the method is tested with the patient-wise leave-one-out strategy separately for the DE and T2 sequences.

uniGAN: In this method, the cGAN network design is the same as the cGAN method. However, the corresponding T2 and DE slices are combined as the input, and the manual segmentation of both slices is combined for training. The elliptic Fourier descriptors are used for smoothing, and the patient-wise leave-one-out strategy is also used.

3. Results

Some exemplary results are shown in Fig. 5 for visual examination. To quantitatively evaluate the segmentation results, the DSC of the myocardium was calculated. The DSC of the proposed method and the comparison methods are shown in Table 1. The typical results of three subjects are illustrated in Fig. 6.

4. Discussion

4.1. Study of the effectiveness of unified segmentation

The results show that the proposed method achieves significant improvement in both the T2 and DE images compared to the **cGAN** method ($p < 0.01$) and the best DSC for the myocardium among all of the compared methods. This finding demonstrates the overall strength of the proposed framework.

Compared to the **Separate** method, the proposed method achieves a significant improvement in both T2 and DE segmentation ($p < 0.01$). This finding is due to the cross-constrained shape, which can better estimate the real myocardium shape with information from both sequences. The use of complementary information also makes the segmentation framework more robust. As shown in the first and second columns in Fig. 6, the improvement mainly occurred in the regions where no clear boundaries could be observed.

Although the improvement of the proposed method is relatively minor in the T2 sequence compared to the **NoDiff** method, the improvement is significant ($p < 0.01$). As illustrated in columns 1 and 3 in Fig. 6, forcing the shape model to fit the information in both sequences directly may result in a non-efficient shape constraint due to the existence of a shape discrepancy and misalignment. The difference term can not only handle the shape discrepancy between the two sequences but can also strengthen the probability shape map from the sparse representation model. While, no significant difference is detected between the results of the proposed method and the **NoDiff** method in DE segmentation. This finding shows that with the low tissue contrast, the DE sequence benefits more from the multi-sequence analysis approach. In addition, the T2 sequence better visualizes the right ventricle wall, which has the same intensity as the myocardium of the LV, and can be wrongly segmented into the LV myocardium. Thus the myocardium segmentation of T2 relies more on the shape contrast to retain a circular shape.

In the **uniGAN** method, the T2 and DE slices are concatenated as the input, and thus multi-sequence information is learned in the training stage. The proposed method achieves better results than the **uniGAN** method, which shows that the shape model and the shape discrepancy in the proposed method makes better use of the complementary information than the **cGAN,** which is designed for general purposes.

4.2. Comparison with related studies

As mentioned in the Introduction section, few studies have investigated myocardium segmentation from T2 CMR. Thus, we list the segmentation results from DE MRI reported in the literature in recent years for reference (Table 2). However, performing a com-

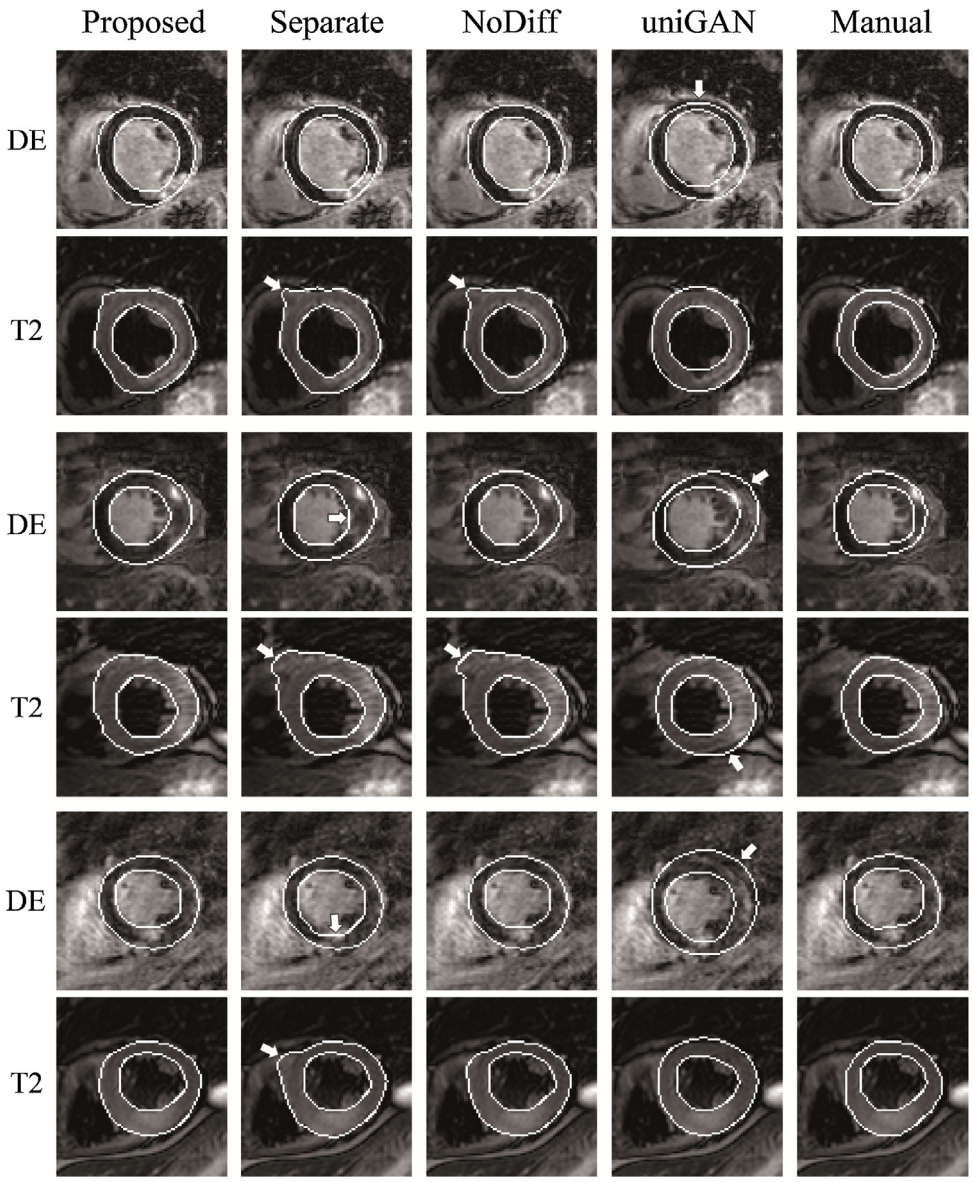

Fig. 6. The results of the proposed method and three other methods compared with the manually delineated ground truth. The corresponding T2 and DE slices of three subjects are shown: the results shown in rows 1 and 2, 3 and 4, and 5 and 6 correspond to same patient. The arrows indicate the major errors.

Table 2
Results reported in the literature. Note that Wei et al. (2013) and Tao et al. (2015) evaluated the results with two sets of manual segmentations. Endo, Epi and Myo represent the endocardium, epicardium and myocardium, respectively.

Reference	No. of patients	DSC (%)		
		Endo	Epi	Myo
Wei et al. (2013)	21	95.33 ± 3.63, 92.64 ± 4.36	96.88 ± 1.84, 94.35 ± 2.70	88.57 ± 4.75, 82.32 ± 5.59
Tao et al. (2015)	50	N/A	N/A	81 ± 7, 83 ± 9
Kurzendorfer et al. (2017)	26	85 ± 6	84 ± 6	N/A
Liu et al. (2017)	22	86.74 ± 5.82	90.40 ± 3.17	73.77 ± 5.56
Ours	32	87.91 ± 4.47	92.62 ± 2.40	78.13 ± 6.22

parison across different studies is difficult due to differences in the data sets. In the works of Wei et al. (2013) and Tao et al. (2015), the corresponding cine MRI is semi-automatically segmented and used as the prior information. In Wei et al. (2013), manual registration is also introduced when the automatic translational registration fails. In contrast, our approach is fully automatic. In the work of Kurzendorfer et al. (2017), no prior shape information is used, whereas only weak shape information of the myocardium thickness constraint is used in the work of Liu et al. (2017).

Zhuang (2016) also proposed a method to perform multi-sequence myocardium segmentation using three sequences: cine, T2 and DE CMR. This method corrects misalignments of the slice and the whole image volume within the segmentation framework and generates a single segmentation in the common space for all three sequences. However, this work did not consider the non-rigid

shape deformation of the heart across different sequences, which could limit further improvement of the segmentation accuracy. The T2 CMR results are $83.0 \pm 12.4\%$ for the endocardium, $91.9 \pm 2.3\%$ for the epicardium and $76.9 \pm 10.7\%$ for the myocardium, and the DE CMR results are $87.8 \pm 3.9\%$ for the endocardium, $90.3 \pm 2.7\%$ for the epicardium and $74.0 \pm 5.1\%$ for the myocardium.

4.3. Recommended parameter ranges and possible usage

In the proposed method, the parameters for the relative weighting of energy function terms υ, γ, μ_1, μ_2 and the weighting parameters of sparse regularization terms λ_1, λ_2 were jointly tested, respectively. The best DSC for DE and T2 sequences are achieved at slightly different parameter values. The recommended parameter ranges where both sequences can achieve relative good results are: 0.5–1.5 for μ_1, 0.5–1.5 for μ_2, 4–5 for υ, 0.1–1 for γ, 0.1–10 for λ_1, and 10–100 for λ_2. Since μ_1 and μ_2, are related to a particular sequence in the energy function, parameter μ_1 mainly affects the segmentation results of T2 images, while μ_2 affects DE results. The parameter λ_1 in the sparse representation based shape model controls the level of sparsity. When the value of λ_1 is too large, the number of effective templates is too small to represent different myocardium shapes. The parameter λ_2 controls the sparsity of registration error compensation term $\boldsymbol{e}$. When the value of λ_2 is too small, the registration error compensation term $\boldsymbol{e}$ is no longer a sparse vector, which can lead to the violation of shape constraint and reduce the amount of information shared across the two sequences. In the contrast, when the value of λ_2 is above the optimal range, the term $\boldsymbol{e}$ cannot efficiently model the shape discrepancy. When λ_2 is extremely large, this is the case of **NoDiff** method. Compared with DE images, the parameters υ, γ, λ_1 and λ_2 have less impact on the segmentation results of T2 images, which is agree with the observation that the T2 images have better tissue contrast and can provide more intensity information. The sparse representation based model is able to select the proper shape templates from the dictionary to model the myocardium shape. The threshold T is used to remove some irrelevant shape templates and reduce computation time. Since the shape repository used in this study is relative small, we choose a low threshold of 0.3. If the shape repository is large and has a good variety of shape characters, a higher value should be set.

This multi-sequence segmentation framework can be used for unified segmentation of DE and T2 CMR. The computation time is relative short. This is because the level set framework is relative effect in computation and the number of templates in the target slice-specific shape dictionary is relative small. The posterior probability term in the CLS framework can be replaced by other models for local texture extraction, which can generate a probability map. Since this term is separately calculated in each sequence, even different texture extraction methods can be designed for different sequences. Thus, this framework can also be used for the unified segmentation of DE CMR and the corresponding phase in cine CMR or for other CMR sequences with proper texture extraction methods.

The segmentation results can also be used for registration correction of multiple CMR sequences. Then, the fused images can be used for visualization and further analysis.

4.4. Future work

In this work, a shape discrepancy compensation term is introduced to address the shape discrepancy, and the experiments have demonstrated its effectiveness. However, this term also has the potential to violate the shape constraint and lead to unrealistic segmentation results. One potential way to improve this framework is to use multitask learning, where each sequence is associated with a separate shape model when the model parameters are related. Thus, the shape information from different sequences can be shared, and each sequence maintains a realistic shape constraint. Additionally, a simple empirical method is used in the construction of a target-specific dictionary, and better methods to build the dictionary with more collected templates can be explored. Because the slice thickness was large in our data, we performed the segmentation in a slice-by-slice manner in this work. We will consider introducing 3D regularization in future work.

5. Conclusion

In this work, we propose a novel approach for multi-sequence segmentation in CMR. The framework is designed based on the CLS method. We adopt the sparse representation-based shape model to build the cross-constrained shape, and the shape discrepancy that cannot be captured by registration is explicitly modeled. The GMM is used to extract useful intensity information from each sequence. The experimental results demonstrate that the unified framework can improve the segmentation of both sequences and that the shape discrepancy term can handle differences between the sequences and provide a better shape constraint. This pilot work may lead to useful tools and inspire new ideas in the field of CMR multi-sequence analysis.

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is partially supported by the National Key Research and Development Program (2016YFC0106200), the 863 National Research Fund (2015AA043203) and the Chinese NSFC research fund (61190120, 61190124 and 61271318).

References

Amado, L.C., Gerber, B.L., Gupta, S.N., Rettmann, D.W., Szarf, G., Schock, R., Nasir, K., Kraitchman, D.L., Lima, J.A.C., 2004. Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model. J. Am. Coll. Cardiol. 44, 2383–2389.

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

Bai, W., Shi, W., Ledig, C., Rueckert, D., 2015. Multi-atlas segmentation with augmented features for cardiac MR images. Med. Image Anal. 19, 98–109.

Beek, A.M., van Rossum, A.C., 2010. Cardiovascular magnetic resonance imaging in patients with acute myocardial infarction. Heart 96, 237–243.

Chen, F., Liu, J., Zhao, Z., Zhu, M., Liao, H., 2017. 3D feature-enhanced network for automatic femur segmentation. IEEE J. Biomed. Health Inform.

Chen, F., Ma, R., Liu, J., Zhu, M., Liao, H., 2018. Lumen and media-adventitia border detection in IVUS images using texture enhanced deformable model. Comput. Med. Imaging Graph 66, 1–13.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3D U-Net: Learning Dense Volumetric Segmentation From Sparse Annotation, Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer International Publishing, Cham, pp. 424–432.

Ciofolo, C., Fradkin, M., Mory, B., Hautvast, G., Breeuwer, M., 2008. Automatic myocardium segmentation in late-enhancement MRI. In: IEEE International Symposium on Biomedical Imaging (ISBI)., pp. 225–228.

Duan, C.J., Liang, Z.R., Bao, S.L., Zhu, H.B., Wang, S., Zhang, G.X., Chen, J.J., Lu, H.B., 2010. A coupled level set framework for bladder wall segmentation with application to MR cystography. IEEE Trans. Med. Imaging 29, 903–915.

Eitel, I., Desch, S., Fuernau, G., Hildebrand, L., Gutberlet, M., Schuler, G., Thiele, H., 2010. Prognostic significance and determinants of myocardial salvage assessed by cardiovascular magnetic resonance in acute reperfused myocardial infarction. J. Am. Coll. Cardiol. 55, 2470–2479.

El Berbari, R., Kachenoura, N., Frouin, F., Herment, A., Mousseaux, E., Bloch, I., 2009. An automated quantification of the transmural myocardial infarct extent using cardiac DE-MR images. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society., pp. 4403–4406.

Friedrich, M.G., Abdel-Aty, H., Taylor, A., Schulz-Menger, J., Messroghli, D., Dietz, R., 2008. The salvaged area at risk in reperfused acute myocardial infarction as

visualized by cardiovascular magnetic resonance. J. Am. Coll. Cardiol. 51, 1581–1587.
Hajiaghayi, M., Groves, E.M., Jafarkhani, H., Kheradvar, A., 2017. A 3-D active contour method for automated segmentation of the left ventricle from magnetic resonance images. IEEE Trans. Biomed. Eng. 64, 134–144.
Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)., pp. 5967–5976.
Kim, H.W., Farzaneh-Far, A., Kim, R.J., 2009. Cardiovascular magnetic resonance in patients with myocardial infarction: current and emerging applications. J. Am. Coll. Cardiol. 55, 1–16.
Kohl, S., Bonekamp, D., Schlemmer, H.P., Yaqubi, K., Hohenfellner, M., Hadaschik, B., Radtke, J.P., Maierhein, K., 2017. Adversarial networks for the detection of aggressive prostate Cancer. arXiv, 1702–08014.
Kuhl, F.P., Giardina, C.R., 1982. Elliptic Fourier features of a closed contour. Comput. Graph. Image Process. 18, 236–258.
Kurzendorfer, T., Brost, A., Forman, C., Maier, A., 2017. Automated left ventricle segmentation in 2-D LGE-MRI. In: IEEE 14th International Symposium on Biomedical Imaging (ISBI)., pp. 831–834.
Liu, J., Zhuang, X., Liu, J., Zhang, S., Wang, G., Wu, L., Xu, J., Gu, L., 2014. Myocardium segmentation combining T2 and DE MRI using multi-component bivariate Gaussian mixture model. In: IEEE 11th International Symposium on Biomedical Imaging (ISBI)., pp. 886–889.
Liu, J., Zhuang, X., Wu, L., An, D., Xu, J., Peters, T., Gu, L., 2017. Myocardium segmentation from DE MRI using multi-component Gaussian mixture model and coupled level set. IEEE Trans. Biomed. Eng., PP, 1-1.
Merino-Caviedes, S., Cordero-Grande, L., Rodríguez, M.T.P., Sevilla-Ruiz, M.T., Revilla-Orodea, A., Martín-Fernández, M., Alberola-Lopez, C., 2016. A variational method for scar segmentation with myocardial contour correction in DE-CMR images. In: IEEE 13th International Symposium on Biomedical Imaging (ISBI)., pp. 956–959.
Motwani, M., Kidambi, A., Greenwood, J.P., Plein, S., 2014. Advances in cardiovascular magnetic resonance in ischaemic heart disease and non-ischaemic cardiomyopathies. Heart 100, 1722–1733.
Mukhopadhyay, A., Oksuz, I., Bevilacqua, M., Dharmakumar, R., Tsaftaris, S.A., 2015. Unsupervised Myocardial Segmentation for Cardiac MRI, Medical Image Computing and Computer-assisted Intervention (MICCAI). Springer International Publishing, Cham, pp. 12–20.
Petitjean, C., Dacher, J.N., 2011. A review of segmentation methods in short axis cardiac MR images. Med. Image Anal. 15, 169–184.
Queirós, S., Barbosa, D., Heyde, B., Morais, P., Vilaça, J.L., Friboulet, D., Bernard, O., D'hooge, J., 2014. Fast automatic myocardial segmentation in 4D cine CMR datasets. Med. Image Anal. 18, 1115–1131.
Rehwald, W.G., Fieno, D.S., Chen, E.-L., Kim, R.J., Judd, R.M., 2002. Myocardial magnetic resonance imaging contrast agent concentrations after reversible and irreversible ischemic injury. Circulation 105, 224–229.
Ripley, D.P., Musa, T.A., Dobson, L.E., Plein, S., Greenwood, J.P., 2016. Cardiovascular magnetic resonance imaging: what the general cardiologist should know. Heart 102, 1589–1603.
Roth, H.R., Oda, H., Zhou, X.R., Shimizu, N., Yang, Y., Hayashi, Y., Oda, M., Fujiwara, M., Misawa, K., Mori, K., 2018. An application of cascaded 3D fully convolutional networks for medical image segmentation. Comput. Med. Imaging Graph. 66, 90–99.
Swierczynski, P., Papiez, B.W., Schnabel, J.A., Macdonald, C., 2018. A level-set approach to joint image segmentation and registration with application to CT lung imaging. Comput. Med. Imaging Graph. 65, 58–68.
Tan, L.K., Liew, Y.M., Lim, E., McLaughlin, R.A., 2017. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences. Med. Image Anal. 39, 78–86.
Tao, Q., Piers, S.R.D., Lamb, H.J., van der Geest, R.J., 2015. Automated left ventricle segmentation in late gadolinium-enhanced MRI for objective myocardial scar assessment. J. Magn. Reson. Imaging 42, 390–399.
Wagner, A., Mahrholdt, H., Holly, T.A., Elliott, M.D., Regenfus, M., Parker, M., Klocke, F.J., Bonow, R.O., Kim, R.J., Judd, R.M., 2003. Contrast-enhanced MRI and routine single photon emission computed tomography (SPECT) perfusion imaging for detection of subendocardial myocardial infarcts: an imaging study. Lancet 361, 374–379.
Wang, B., Gu, X., Fan, C., Xie, H., Zhang, S., Tian, X., Gu, L., 2015. Sparse group composition for robust left ventricular epicardium segmentation. Comput. Med. Imaging Graph. 46, 56–63.
Wei, D., Sun, Y., Ong, S.-H., Chai, P., Teo, L.L., Low, A.F., 2013. 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. 17, 685–697.
Zhang, S.T., Zhan, Y.Q., Dewan, M., Huang, J.Z., Metaxas, D.N., Zhou, X.S., 2012. Towards robust and effective shape modeling: sparse shape composition. Med. Image Anal. 16, 265–277.
Zhuang, X., 2016. Multivariate Mixture Model for Cardiac Segmentation From Multi-sequence MRI, Medical Image Computing and Computer-assisted Intervention (MICCAI). Springer International Publishing, Cham, pp. 581–588.