

UC Irvine

UC Irvine Previously Published Works

Title

Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach

Permalink

<https://escholarship.org/uc/item/48q058fq>

Journal

Magnetic Resonance in Medicine, 78(6)

ISSN

0740-3194

Authors

Avendi, Michael R
Kheradvar, Arash
Jafarkhani, Hamid

Publication Date

2017-12-01

DOI

10.1002/mrm.26631

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Automatic Segmentation of the Right Ventricle from Cardiac MRI Using a Learning-Based Approach

Michael R. Avendi,^{1,2,3} Arash Kheradvar,^{1,2} and Hamid Jafarkhani^{3*}

Purpose: This study aims to accurately segment the right ventricle (RV) from cardiac MRI using a fully automatic learning-based method.

Methods: The proposed method uses deep learning algorithms, i.e., convolutional neural networks and stacked autoencoders, for automatic detection and initial segmentation of the RV chamber. The initial segmentation is then combined with the deformable models to improve the accuracy and robustness of the process. We trained our algorithm using 16 cardiac MRI datasets of the MICCAI 2012 RV Segmentation Challenge database and validated our technique using the rest of the dataset (32 subjects).

Results: An average Dice metric of 82.5% along with an average Hausdorff distance of 7.85 mm were achieved for all the studied subjects. Furthermore, a high correlation and level of agreement with the ground truth contours for end-diastolic volume (0.98), end-systolic volume (0.99), and ejection fraction (0.93) were observed.

Conclusion: Our results show that deep learning algorithms can be effectively used for automatic segmentation of the RV. Computed quantitative metrics of our method outperformed that of the existing techniques participated in the MICCAI 2012 challenge, as reported by the challenge organizers.

Magn Reson Med 000:000–000, 2017. © 2017 International Society for Magnetic Resonance in Medicine.

Key words: cardiac MRI; right ventricle; segmentation; deep learning; deformable models

INTRODUCTION

Compared with left ventricle (LV), the study of the right ventricle (RV) is a relatively young field. Although many recent studies have confirmed the prognostic value of RV function in cardiovascular disease, for several years its significance has been underestimated (1,2). Understanding the role of RV in the pathophysiology of heart failure traditionally has dawdled behind that of the LV. The RV used to be considered a relatively passive

chamber relating the systemic and pulmonary circulation until more recent studies revealed its importance in sustaining the hemodynamic stability and cardiac performance (3–5).

Cardiac MRI is the preferred method for clinical assessment of the RV (6–12). Currently RV segmentation is manually performed by clinical experts, which is lengthy, tiresome and sensitive to intra and interoperator variability (6,13,14). Therefore, automating the RV segmentation process turns this tedious practice into a fast procedure. This ensures the results are more accurate and not vulnerable to operator-related variabilities, and eventually accelerates and facilitates the process of diagnosis and follow-up.

There are many challenges for RV segmentation that are mainly attributed to RV anatomy. These can be summarized as: presence of RV trabeculations with signal intensities similar to that of the myocardium, complex crescent shape of the RV, which varies from base to apex, along with inhomogeneity reflected in the apical image slices, and considerable variability in shape and intensity of the RV chamber among subjects, notably in pathological cases (6).

Currently, only limited studies have focused on RV segmentation (6). Atlas-based methods have been considered in some studies (15–17), which require large training datasets and long computational times while their final segmentation may not preserve the mostly regular boundaries of the ground-truth masks (16). Nevertheless, it is challenging to build a model general enough to cover all possible RV shapes and dynamics (18). Alternatively, graph-cut-based methods combined with distribution matching (19), shape-prior (20) and region-merging (21) were studied for RV segmentation. Overall, these methods suffer from a low robustness and accuracy, and require extensive user interaction. Image-based methods have been considered by Ringenberget al (22) and Wang et al (23). They showed notable accuracy and processing time improvement over other methods while deformed RV shape and patient movement during the scan are the limitations of their method (22). Current learning-based approaches, such as probabilistic boosting trees and random forests, depend on the quality and extent of annotated training data and are computationally expensive (24–27).

Motivated by these limitations, we developed an accurate, fast, robust and fully automated segmentation framework for cardiac MRI. A convolutional neural network (28–31) is used to automatically detect the location of RV in the thoracic cavity and provide a region of interest (ROI). Afterward, a stacked autoencoder (stacked-AE) (32–37) is developed to automatically segment the RV and

¹The Edwards Lifesciences Center for Advanced Cardiovascular Technology, University of California, Irvine, California, USA.

²Department Biomedical Engineering, University of California, Irvine, California, USA.

³Center for Pervasive Communications and Computing, University of California, Irvine, California, USA.

Grant sponsor: an American Heart Association Grant-in-Aid Award; Grant number: 14GRNT18800013; Grant sponsor: Conexant-Broadcom Endowed Chair.

*Correspondence to: Hamid Jafarkhani, PhD, Center for Pervasive Communications & Computing, 4217 Engineering Hall, University of California, Irvine. E-mail: hamidj@uci.edu

Received 18 May 2016; revised 11 January 2017; accepted 11 January 2017

DOI 10.1002/mrm.26631

Published online 00 Month 2017 in Wiley Online Library (wileyonlinelibrary.com).

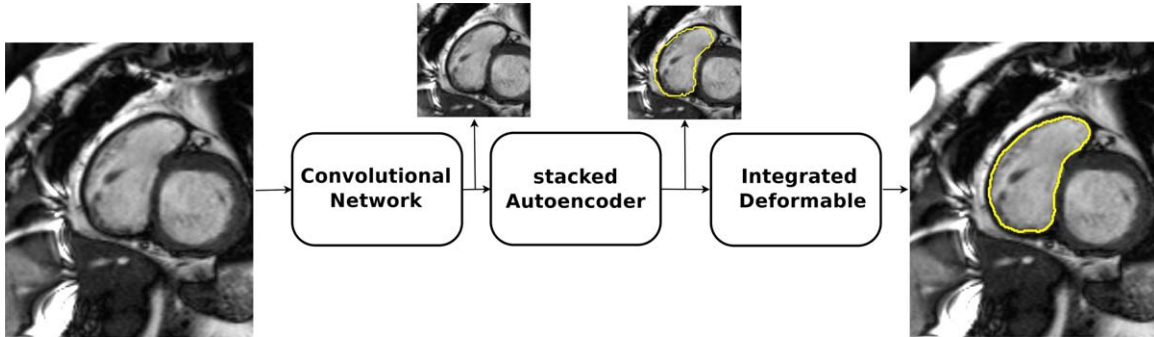


FIG. 1. Block diagram of the integrated deep learning and deformable model algorithm for RV segmentation.

provide an initial contour. Finally, a method is introduced that incorporates the initial contour into classical deformable models to provide an accurate and robust RV contour. The algorithm is successfully validated on the MICCAI 2012 RV database (6).

The developed deep learning algorithm is based on the supervised learning paradigm. In supervised learning, some example data and corresponding labels are required to train and develop the algorithm. In other words, the algorithm artificially mimics the function of a human annotator. As a result, the algorithm can perform as good as the human annotator. Therefore, to obtain good results, it is important to provide the algorithm with clean and accurate data and labels (38).

Our major contributions include: (i) designing a fully automatic RV segmentation method for MRI datasets; (ii) using deep learning algorithms, trained with limited data, for automatic RV localization and initial segmentation; and (iii) incorporating the deep learning output into deformable models to address the shrinkage/leakage problems and reduce the sensitivity to initialization. Finally, a better performance in terms of multiple evaluation metrics and clinical indices was achieved.

METHODS

We used the MICCAI 2012 RV segmentation challenge (RVSC) database (6) provided by the LITIS Lab, at the University of Rouen, France. The algorithms were developed in our research centers at UC Irvine. Then, the results were submitted to the LITIS lab for independent evaluations. The cardiac MRI datasets were acquired by a 1.5 Tesla Siemens scanner that includes 48 short-axis images of patients with known diagnoses. The database is grouped into three datasets namely: TrainingSet, Test1Set, and Test2Set. Each dataset contains 16 image sequences corresponding to 16 patients. Manual delineations of RV at the end-diastole (ED) and end-systole (ES) are included in TrainingSet only. A typical dataset contains nine images at ED and seven images at ES from base to apex. Image parameters are summarized as: slice thickness = 7 mm, image size = 256×216 (or 216×256) pixels with 20 images per cardiac cycle.

Our method requires square inputs; therefore, patches of 216×216 were cropped out of the original images and used during the training and testing procedures. We used images in TrainingSet to train our algorithm. After completion of training, the algorithm was deployed for

RV segmentation in Test1Set and Test2Set. The ground truth delineations of Test1Set and Test2Set are not publicly available and the LITIS Lab provided the independent assessment results upon receiving the automatic segmentations.

Algorithm Description

The method is carried out over three stages as shown in Figure 1. The algorithm receives a short-axis cardiac MR image as the input (Fig. 1). First, in Step 1, the ROI containing the RV is determined in the image using a convolutional network trained to locate the RV. Then, in Step 2, the RV is initially segmented using a stacked-AE trained to delineate the RV. The obtained contour is used for initialization and incorporated into deformable models for segmentation in Step 3. Each stage of the block diagram is individually trained during an offline training process to obtain its optimum values of parameters, as described in our previous publication on LV segmentation (39). After training, the system is deployed for automatic segmentation. Here, we have used our developed localization and segmentation algorithms jointly; however, the two can be applied independently. In other words, our segmentation algorithm can work in conjunction with other automatic localization techniques or even without localization. Each step is further explained as follows for completeness of the presentation.

Automatic Localization (Step 1)

The original images in the database have a large field of view, covering the RV chamber as well as parts of the other surrounding organs. In addition, direct handling of the images is not computationally feasible because of the large image size. As such, we first localize the RV and crop out a ROI from the original images such that the RV chamber is positioned approximately within the center of the images.

Figure 2 shows a block diagram of the automatic RV localization using convolutional networks. We use a down-sampled $m \times m$ image as the input to reduce complexity. Let us represent the pixel intensity at coordinate $[i, j]$ by $\mathbf{I}[i, j]$. Throughout the study, we represent the i -th element of vector \mathbf{x} by $\mathbf{x}[i]$ and the element at the i -th row and the j -th column of matrix \mathbf{X} by $\mathbf{X}[i, j]$.

Then, the filters ($F_l \in R^{a \times a}$, $b_0 \in R^k$, $l = 1, \dots, k$) are convolved with the input image to obtain k convolved

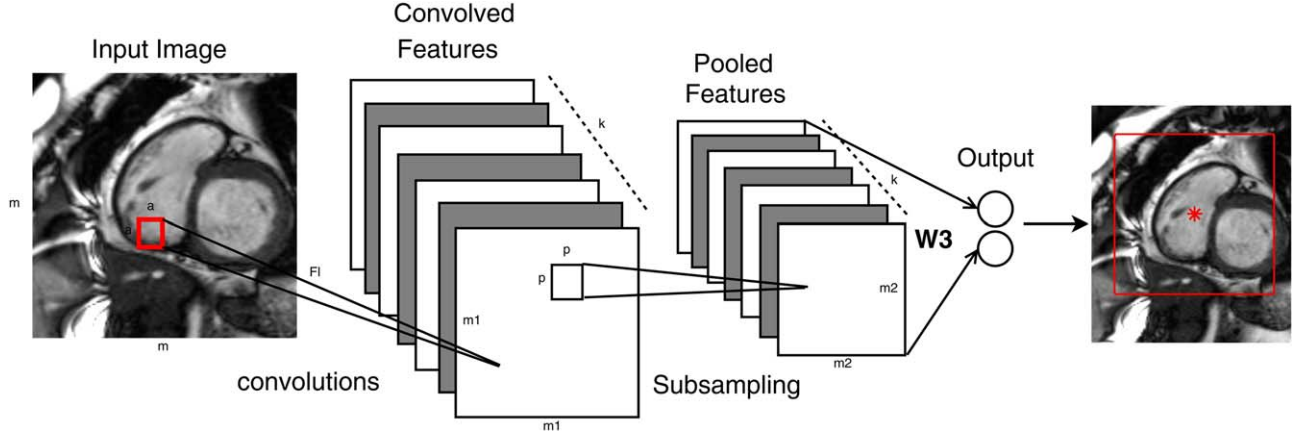


FIG. 2. Block diagram of the convolutional network for automatic localization.

feature maps of size $m_1 \times m_1$, computed as $C_l[i, j] = f(\mathbf{Z}_l[i, j])$ where

$$Z_l[i, j] = \sum_{k_1=1}^a \sum_{k_2=1}^a F_l[k_1, k_2] I[i + k_1 - 1, j + k_2 - 1] + b_0[l], \quad [1]$$

for $1 \leq i, j \leq m_1$, $l = 1, \dots, k$, and $m_1 = m - a + 1$.

As shown in Figure 2, the next step in automatic localization is average pooling. The goal of average pooling is to down-sample the convolved feature maps by averaging over $p \times p$ nonoverlapping regions in the convolved feature maps. This is done by calculating

$$P_l[i_1, j_1] = \frac{1}{p^2} \sum_{i=1+(i_1-1)p}^{i_1 p} \sum_{j=1+(j_1-1)p}^{j_1 p} C_l[i, j] \quad [2]$$

for $1 \leq i_1, j_1 \leq m_2$. This results in k reduced-resolution features $\mathbf{P}_l \in \mathbb{R}^{m_2 \times m_2}$ for $l = 1, \dots, k$, where $m_2 = m_1/p$ and p is chosen such that m_2 is an integer value. We set $m = 54$, $a = 10$, $m_1 = 45$, $p = 5$, $m_2 = 9$, $k = 100$ for an original 216×216 MR image.

The pooled features are finally converted to vector $\mathbf{p} \in \mathbb{R}^{n_2}$, where $n_2 = km_2^2$, and fully connected to a linear regression layer with two outputs. We train the network to find matrices $\mathbf{W}_1 \in \mathbb{R}^{2 \times n_2}$ and $\mathbf{b}_1 \in \mathbb{R}^2$ and compute $\mathbf{y}_c = \mathbf{W}_1 \mathbf{p} + \mathbf{b}_1$ at the output layer. Centered at the obtained output, a ROI with size M_{roi} is cropped from the original image to be used for the next stage. The image slices near the RV base require a larger region to cover the whole RV with respect to image slices at the apex. We group the contours into large and small, and set $M_{\text{roi}} = 171, 91$ for those, respectively. To optimize the performance of the automatic RV localization, the convolutional network is trained using the back-propagation algorithm (40) to obtain the parameter values $\mathbf{F}_l, l = 1, \dots, k$, $\mathbf{b}_0, \mathbf{W}_1$ and \mathbf{b}_1 .

Automatic Initialization (Step 2)

We use a stacked-AE to obtain an initial RV segmentation. As shown in Figure 3, in addition to the input and output layers, we have two hidden layers in the stacked-AE. The input vector, $\mathbf{x}_s \in \mathbb{R}^{n_1}$, is constructed by down-sampling and unrolling the sub-image obtained from the

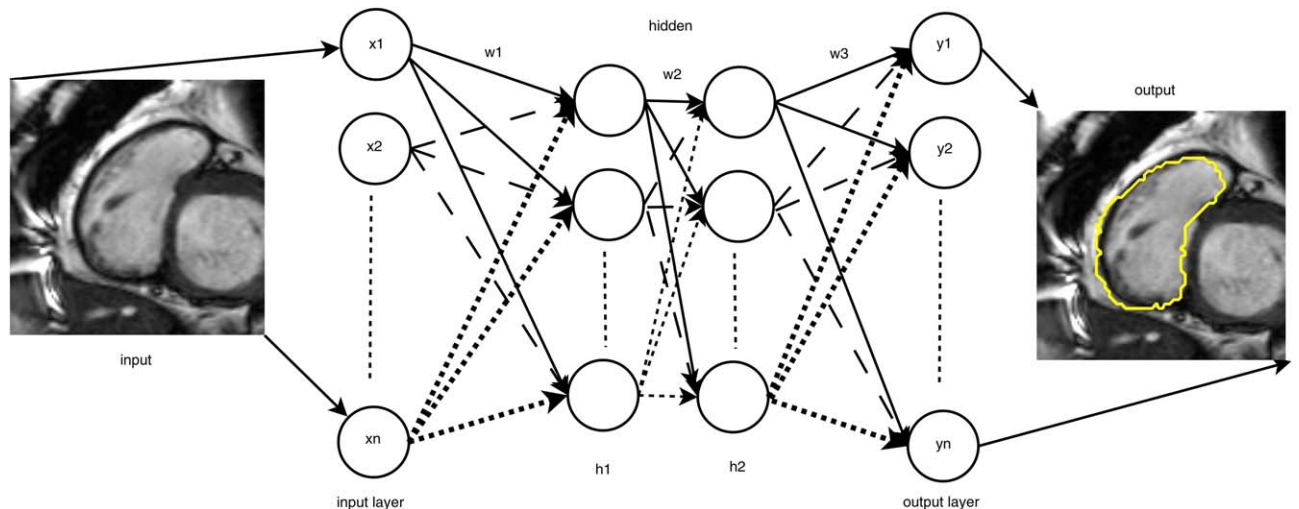


FIG. 3. Block diagram of the stacked-AE for initialization.

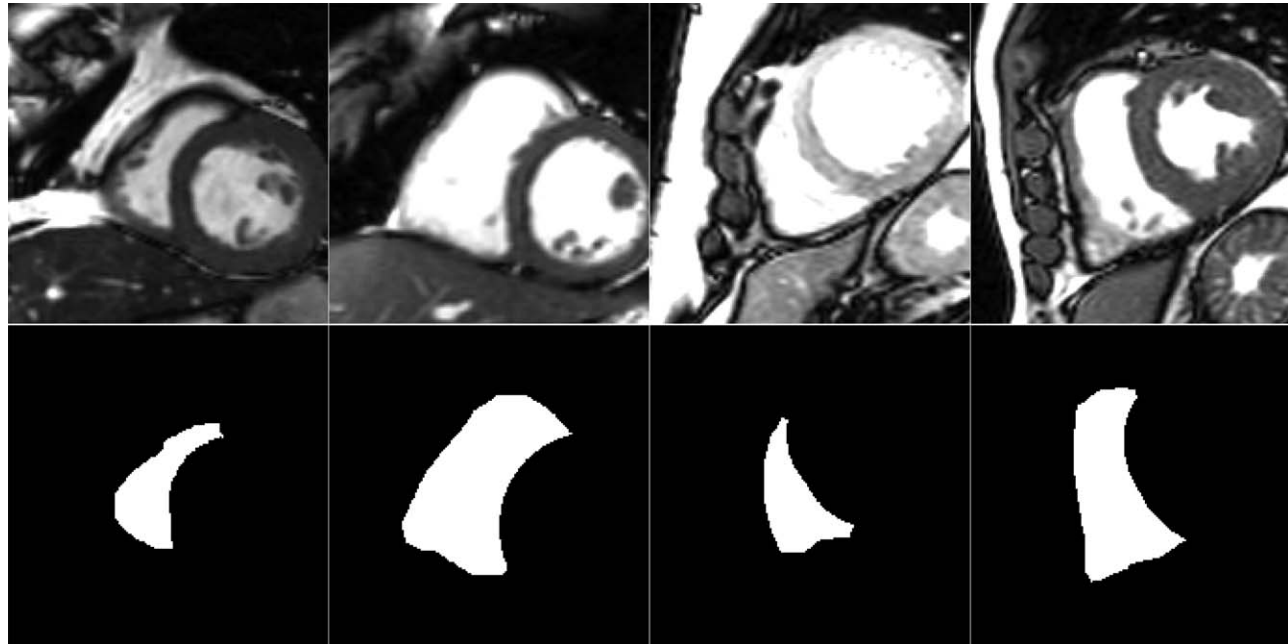


FIG. 4. Four examples of the training data for the stacked-AE, input (upper row) and labels (lower row).

automatic localization block. The hidden layers build the abstract representations by computing $\mathbf{h}_1 = f(\mathbf{W}_2 \mathbf{x}_s + \mathbf{b}_2)$ and $\mathbf{h}_2 = f(\mathbf{W}_3 \mathbf{h}_1 + \mathbf{b}_3)$. The output layer computes $\mathbf{y}_s = f(\mathbf{W}_4 \mathbf{h}_2 + \mathbf{b}_4)$ to produce a binary mask. The binary mask is black (zero) everywhere except at the RV borders, and can also be converted to a contour as shown in Figure 3. Here, $\mathbf{W}_2 \in \mathbb{R}^{n_2 \times n_1}$, $\mathbf{b}_2 \in \mathbb{R}^{n_2}$, $\mathbf{W}_3 \in \mathbb{R}^{n_3 \times n_2}$, $\mathbf{b}_3 \in \mathbb{R}^{n_3}$ and $\mathbf{W}_4 \in \mathbb{R}^{n_4 \times n_3}$, $\mathbf{b}_4 \in \mathbb{R}^{n_4}$ are trainable matrices and vectors that are obtained during the training process. We set the parameters as $n_1 = 3249$, $n_2 = 300$, $n_3 = 300$, $n_4 = 3249$. The output of this stage has a dual functionality; it is used as the initial contour for the segmentation step as well as a preliminary RV shape.

Training Stacked-AE

Although, an end-to-end supervised training can be used to train the stacked-AE, it does not lead to a good generalization due to the small size of the training data. For better generalization, we use an unsupervised layer-wise pretraining followed by an end-to-end supervised fine-tuning. Four typical examples of input images and labels are shown in Figure 4. The details can be found in Avendi et al (39).

RV Segmentation (Step 3)

As shown in Figure 1, the initial segmentation derived from the previous step is used as a preliminary contour in a deformable model. Deformable models are dynamic contours that eventually lie on the boundary of the object of interest. The evolution of the deformable models is aimed at minimizing an energy function. In conventional deformable methods, contours tend to shrink inward or leak outward because of the fuzziness of the cavity borders and presence of RV trabeculations. These issues are resolved by integrating the preliminary RV

shape obtained from the previous stage into the deformable models.

We define the level-set function $\varphi(x,y)$ with negative and positive values for the pixels inside and outside a contour, respectively. We also define the following energy function

$$E(\varphi) = \alpha_1 E_{\text{len}}(\varphi) + \alpha_2 E_{\text{reg}}(\varphi) + \alpha_3 E_{\text{shape}}(\varphi), \quad [3]$$

which is a combination of the length-based, region-based, and prior shape energy terms. The weights α_1 , α_2 , and α_3 are the combining parameters, empirically determined as $\alpha_1 = 1$, $\alpha_2 = 0.5$, and $\alpha_3 = 0.25$. The deformable method minimizes the energy function in Equation [3] to find the following unique contour:

$$\varphi = \underset{\varphi}{\text{arg min}} \{E(\varphi)\} \quad [4]$$

The solution φ^* will lie on the boundary of the object of interest. The optimization problem in Equation [4] can be solved using the gradient descent algorithm.

Implementation Details

Images of all cases in *TrainingSet* were collected and divided into the large-contour and small-contour groups. As such, there are approximately 128 and 75 images of size 256×216 or 216×256 in each group, respectively. We also artificially enlarged the training dataset by translation, rotation and changing the pixel intensities of our images based on the standard principal component analysis (PCA) technique explained by Koikkalainen et al (41). Accordingly, we augmented the training dataset by a factor of 10. Afterward, we built and trained two networks, one for the large-contour and one for the small-contour dataset.

Table 1
Quantitative Metrics and Mean Values (SDs) of DM and HD and Correlation Coefficient R^2 for RV Volumes

Phase	Test1Set (16 patients)			Test2Set (16 patients)		
	DM	HD (mm)	R^2_{Vol}	DM	HD (mm)	R^2_{Vol}
ED (step 2)	0.82 (0.18)	12.40 (8.60)	0.90	0.84 (0.80)	10.10 (6.90)	0.94
ED (all steps)	0.86 (0.11)	7.80 (4.26)	0.98	0.86 (0.10)	7.85 (4.56)	0.96
ES (step 2)	0.70 (0.25)	12.95 (8.70)	0.92	0.74 (0.26)	10.20 (9.20)	0.98
ES (all steps)	0.79 (0.16)	7.51 (3.66)	0.98	0.76 (0.20)	8.27 (4.23)	0.98

Over-fitting is a main issue in deep learning networks as many parameters should be learned. A poor performance on the test data is possible despite a well-fitted network to the training data. To overcome this issue, we adopted multiple techniques including artificial enlargement of the training dataset, performing layer-wise pre-training, l_2 regularization and sparsity constraints. The use of layer-wise pretraining greatly helped to mitigate the challenge of limited training data. To keep track of the number of parameters, the inputs were down-sampled and only two hidden layers were used in the networks. To monitor the overfitting problem during training, a four-fold cross-validation was performed. Accordingly, the original training data (16 subjects) was divided into four partitions, with three partitions in training and one partition for validation, in each fold. The average of the four-fold cross-validations is typically considered the final outcome of the model.

Our method was developed in MATLAB 2014a, performed on a Dell Precision T7610 workstation, with Intel(R) Xeon(R) CPU 2.6 GHz, 32 GB RAM, on a 64-bit Windows 7 platform.

RESULTS

The performance of our methodology was assessed based on comparing the accuracy of the automated segmentation with the ground truth (i.e., manual annotations by experts). *TrainingSet* of the RVSC database (6) was used for training only, and Test1Set and Test2Set were used for validation. Because the reference contours of Test1Set

and Test2Set are not publicly available, we submitted our automatic segmentation results to the LITIS Lab for independent evaluation.

Dice metric (DM) and Hausdorff distance (HD) were computed (6). DM is a measure of contour overlap, with a range between zero and one. A higher DM indicates a better match between automated and manual segmentations. Data augmentation improved the results related to DM for approximately 2.5%. HD measures the maximum perpendicular distance between the automatic and manual contours. Table 1 presents the computed DM, HD, and correlation coefficient R for RV volumes at the ED and ES. To test the effect of permuting the test and validation data, we used four-fold cross-validation, i.e., divided the Training dataset into four partitions. Then, we used three partitions (12 patients around 180 images) for training and one partition (4 patients, around 60 images) for validation in each fold. The DM results of the four fold cross-validations are 0.79, 0.80, 0.84, and 0.78. The average DM of the four fold cross-validations is 0.80. The results are slightly different for each fold. This is due to the fact that the model is trained with different sets of images and tested on different images in each fold. Two exemplary segmentations at the ED and ES are shown in Figures 5 and 6, respectively. Additional segmentation figures, including the results in Steps 2 and 3, can be found in Supporting Figures S2–S5, which are available online. These figures display images from the base to the apex for the best and worst DM results of Test1Set. The red and yellow contours correspond to

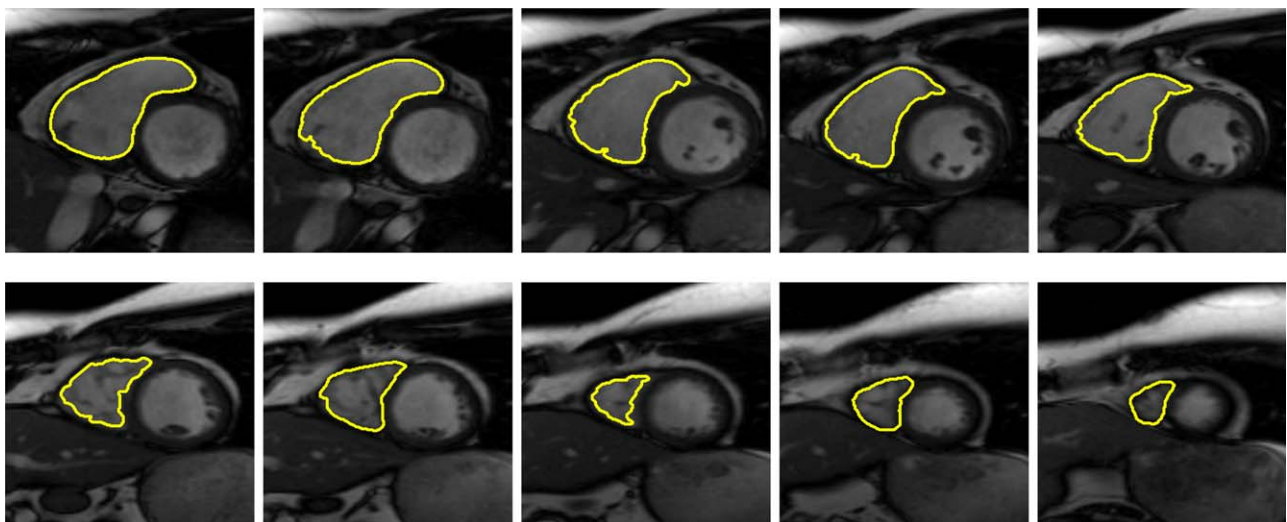


FIG. 5. Endocardial contours of RV at ED from base to apex (Patient #33 from Test2Set).

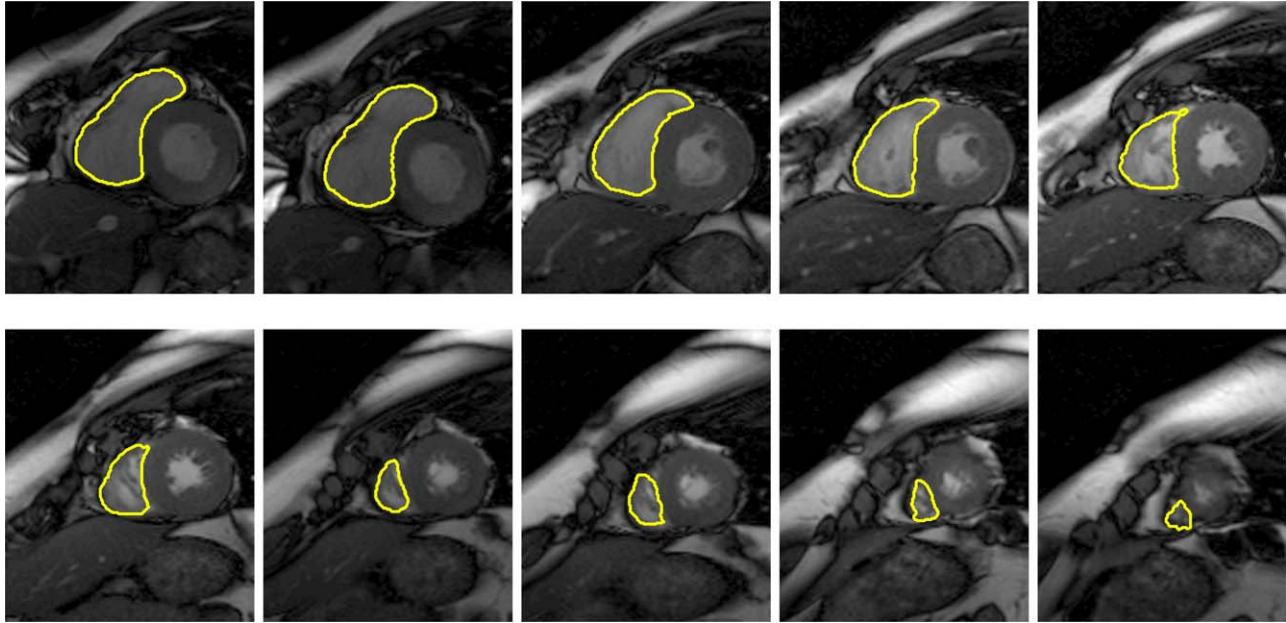


FIG. 6. Endocardial contours of RV at ES from base to apex (Patient #42 from Test2Set).

Step 2 and Step 3, respectively. The refinement in Step 3 leads to overall improvement in the average DM and volume calculations.

For clinical validation, end-diastolic volume (EDV), end-systolic volume (ESV), and ejection fraction (EF) were computed. Correlation and Bland-Altman plots were obtained to assess their agreement to the ground truth. Correlation plots are shown in Figures 7 and 8 and the remaining plots can be found in Supporting Figure S1.

The range of EDV, ESV and EF was (40 mL to 232 mL), (17 mL to 173 mL) and (21% to 67%), in Test1Set and (61 mL to 242 mL), (18 mL to 196 mL), and (19% to 71%) in Test2Set, respectively. The correlation with the ground truth contours of $R=0.99$, 0.99 , 0.96 and $R=0.98$, 0.99 , 0.93 for EDV, ESV, and EF were achieved, for Test1Set and Test2Set, respectively. No statistically significant difference in global EDV (Test1Set $P=0.96$, Test2Set $P=0.25$), ESV (Test1Set $P=0.12$, Test2Set $P=0.54$) and EF (Test1Set $P=0.1$, Test2Set $P=0.22$), was observed. The DM shows the average overlap between the manual delineations and the automatic

results. However, R^2 values correspond to the EDV, ESV, and EF. Obviously, a higher DM leads to a better volume estimation and higher R^2 values. This can be seen in Table 1, where both DM values and R^2 improve from Step 2 to Step 3.

The Bland-Altman analysis showed small biases for EDV (0.11 mL, -3 mL), ESV (0.12 mL, 1.1 mL), and EF (1.6%, 1.6%), in Test1Set and Test2Set, respectively. The level of agreement between the automatic and manual results was represented by the interval of the percentage difference between mean ± 1.96 standard deviation (SD). The confidence interval of the difference between the automatic and manual was measured as EDV (-18 mL to 18 mL), (-23 mL to 17 mL), ESV (-23 mL to 17 mL), (-12 mL to 15 mL), and EF (-8.7% to 5.5%), (-11% to 8.1%), for Test1Set and Test2Set, respectively. In addition, the coefficient of variation was measured as EDV (6.6%, 7.5%), ESV (9.1%, 10%), and EF (7.4%, 9.1%), for Test1Set and Test2Set, respectively.

Approximate elapsed times for training and test processes were obtained using the tic-toc command in

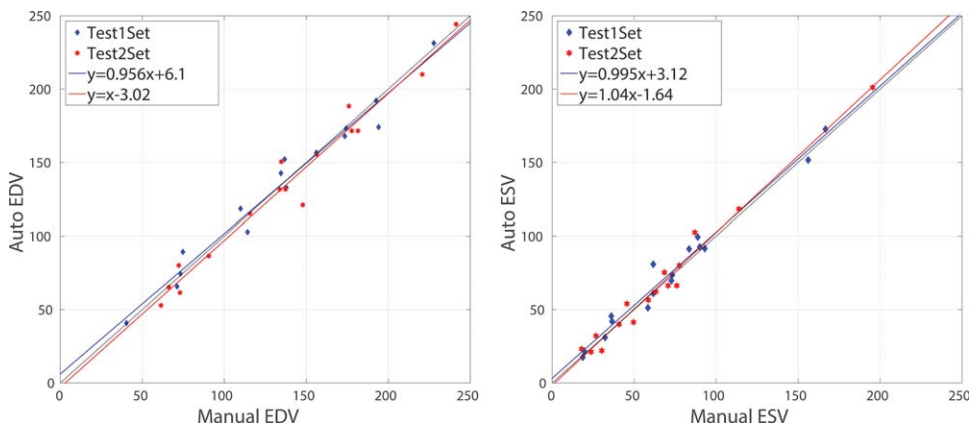


FIG. 7. Correlation plots for EDV and ESV of Test1Set and Test2Set.

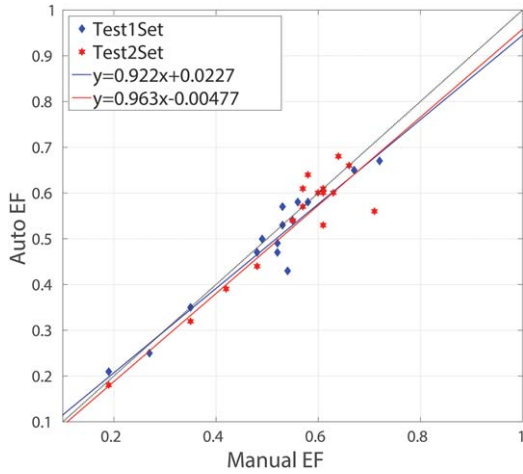


FIG. 8. Correlation plots for ejection fraction of Test1Set and Test2Set.

MATLAB and were as follows: training convolutional network: 7.8 h, training stacked-AE: 74 min. Once trained, the elapsed times for segmenting the RV in a typical MR image were as follows: ROI detection (convolution, pooling, and regression): 0.25 s, initialization (stacked-AE): 0.002 s, segmentation (deformable model): 0.2 s. The elapsed times during inference is the average of 10 tests. Also, the average elapsed time per patient was around 5 s assuming 10 image slices per patient at the ES or ED.

DISCUSSION AND CONCLUSIONS

Most of the challenges for RV segmentation are due to the complex anatomy of the RV chamber. These include RV trabeculations with signal intensities similar to the myocardium's, complex crescent shape of the RV that varies from base to apex, as well as significant variation of RV shape and intensity among the subjects (6). Due to these challenges, only limited studies have focused on RV segmentation. Among those, the state-of-the-art methods for RV segmentation suffer from several limitations such as leakage and shrinkage of contours due to the fuzziness of the RV borders and presence of trabeculations. Our learning-based method overcame these shortcomings and minimized shrinkage/leakage by integrating

the inferred shape into the deformable model. Figures 5 and 6 show that the RV was properly segmented from base to apex. Like many other methods in the literature (6), the large contours can be more accurately segmented compared with the small contours, and working with image slices in vicinity of the apex particularly at ES can be challenging due to the small size and irregular shape.

Computed metrics in Table 1 show that the contours at ED were more accurately segmented in terms of DM compared with the contours at the ES, with an overall accuracy of 86% in both testing sets. This is because the contours at ES are larger and easier to segment. Again, this is also a characteristic of other segmentation methods as reported in Petitjean et al (6).

Table 2 summarizes the computed quantitative metrics averaged over ED and ES. As can be seen from the table, our method outperforms the state-of-the-art methods. Mean DM improvements compared with the other methods range from 0 to 0.28 on Test1Set and 0 to 0.22 on Test2Set. Ringenberg et al (22) demonstrated a mean improvement of 0.01 on Test2Set. Our mean HD improvements range from 1.38–20.77 mm on Test1Set and 0.7–14.18 mm on Test2Set. The closest results to our method is the work by Ringenberg et al (22) with similar DM values. However, our method provides better HD values, i.e., 7.67 mm and 8.03 mm for Test1Set and Test2Set, respectively, compared with 9.05 mm and 8.73 mm reported by Ringenberg et al (22). The smaller HD values of our method indicates superiority of our method over Ringenberg et al.

Figures 7 and 8 show a high correlation for ESV, EDV, and EF (greater than 0.98 for RV volumes), denoting excellent match between the automatic and manual contours. The Bland-Altman analysis revealed negligible biases and a better level of agreement compared with that of the other methods. For example, the Bland-Altman diagrams related to EF showed a bias close to zero with the 95% limits of agreement (± 1.96 SD) close to ± 0.10 . This performance is similar to what reported by Caudron et al (42) for intraoperator variability values. A nonzero bias with the 95% limits closer to ± 0.2 exist for the other methods (6). Compared with the work by Ringenberg et al (22), that provides the closest results to ours in Table 2, our method provides a better R -value (correlation coefficient) for EDV, ESV and EF. For

Table 2
Quantitative Metrics and Mean Values (SDs) of DM and HD Average over ED and ES, for Our Method Compared to Other Techniques Evaluated on TEST1SET and TEST2SET of MICCAI 2012 RVSC Database (6)

Method	A/sA ¹	Test1Set (16 patients)		Test2Set (16 patients)	
		DM	HD (mm)	DM	HD (mm)
Our Method	A	0.83 (0.14)	7.67 (4.00)	0.82 (0.16)	8.03 (4.41)
Ringenberg et al 2014 (22)	A	0.83 (0.16)	9.05 (6.98)	0.83 (0.18)	8.73 (7.62)
Zuluaga et al 2013 (43)	A	0.78 (0.23)	10.51 (9.17)	0.73 (0.27)	12.50 (10.95)
Wang et al 2012 (23)	A	0.57 (0.33)	28.44 (23.57)	0.61 (0.34)	22.20 (21.74)
Ou et al 2012 (16)	A	0.55 (0.32)	23.16 (19.86)	0.61 (0.29)	15.08 (8.91)
Maier et al 2012 (44)	sA	0.80 (0.19)	11.15 (6.62)	0.77 (0.24)	9.79 (5.38)
Nambaksh et al 2013 (45)	sA	0.59 (0.24)	20.21 (9.72)	0.56 (0.24)	22.21 (9.69)
Bai et al 2013 (17)	sA	0.78 (0.20)	9.26 (4.93)	0.76 (0.23)	9.77 (5.59)
Grosgeorge et al 2013 (20)	sA	0.76 (0.20)	9.97 (5.49)	0.81 (0.16)	7.28 (3.58)

A = automatic, sA = semiautomatic.

example, for EF, our method provides $R=0.96$ and 0.93 for Test1Set and Test2Set, respectively, compared with $R=0.78$ and 0.91 reported by Ringenberg et al (22). The higher R -values in our statistical evaluation demonstrates a better performance. These observations show the potential clinical applicability of our framework for automatic RV segmentation.

The measured elapsed times show that the method can be trained within a relatively short time and off-line. The first stage, i.e., convolutional network, requires the longest computational time among the three stages. This is because the most time-consuming operation needed is the convolution of the filters with images. Nevertheless, these computational times can be reduced by developing the algorithms into GPU-accelerated computing platforms.

During the test, the average time to perform RV segmentation, in a typical image, was less than 0.5 s. Most of the computational time was spent by the convolution network and the integrated deformable model. Yet, the integrated deformable model converges faster than classical deformable models because of the initialization and integration with the inferred shape. Overall, our method needs 5 s per patient for the processing. Unfortunately, a fair comparison between computational-time related to different methods was not possible because the other methods have been developed over different platforms. Their reported computational times range from 19 s to 30 min per patient (6,16,17,19–23,43). In particular, the reported computational time reported by Ringenberg et al (22) on a similar workstation with Xeon processor is 19 s per patient, which is approximately four times more than the 5 s needed by our method.

As a limitation, the developed method may not perform as efficiently in patients with irregular RV shape, such as congenital heart defects. This is due to the fact that learning-based approaches are as good as their training data. A rich and diverse dataset for training will ensure the performance for various cases. In other words, to efficiently perform on patients with irregular shape RV, the training dataset should include some of those examples.

As discussed in our previous publication (39), a difficulty in applying deep learning approaches for cardiac MRI segmentation is the lack of enough data for training and eventually validation. For this work, we used a portion of the MICCAI 2012 RVSC dataset (6) and artificially enlarged that for training. Similar to LV segmentation, currently, no analytic approach exists to design hyperparameters in deep learning networks and they should be obtained empirically (39). Nevertheless, the results indicate that our automated method is accurate. Another limitation of this study is that the validation was performed on a dataset with a rather limited number of subjects and abnormalities. Also, because there is only one manual segmentation available from the MICCAI 2012 RVSC dataset, it was not possible to evaluate the intra- and interobserver variability. Testing our method on a larger set of clinical data with multiple manual segmentation, that currently we do not have access to, is subject of future research.

In prospect, manual segmentation is time-consuming and requires dedicated operator training that makes it

inefficient due to the extent of information in CMR images (46,47). Furthermore, because the traditional practice of manual segmentation is subjective, less reproducible and time-consuming, fully automatic 3D segmentation methods are highly desirable for computing functional parameters in patients, such as ejection fraction, cardiac output, peak ejection rate, filling rates among the other.

Our learning approach has the potential to be performed across the whole cardiac cycle. The method can also be extended to RV myocardial segmentation to provide additional clinical details. The current RV endocardial segmentation can be used as a preprocessing step to more accurately consider RV trabeculations. Comparison of RV segmentation results with that of LV segmentation, DM (94%), and HD (3.45 mm) (39), confirmed the difficulty of RV segmentation because of its complex shape variability. Nevertheless, further improvements of these metrics for RV segmentation to reach an accuracy similar to that of LV segmentation should be considered. Furthermore, the method can be considered for simultaneous multiple chamber segmentation by providing training labels that include multiple chambers.

In conclusion, we have developed a novel method for fully automatic RV segmentation from cardiac MRI short-axes. The method uses deep learning algorithms combined with deformable models. It brings more robustness and accuracy, particularly for challenging images with fuzzy borders. In contrast to the other existing automated approaches, our method is based on learning several levels of representations, corresponding to a hierarchy of features and does not assume any model or assumption about the image or heart. The method is simple to implement, and potentially more robust against anatomical variability and image contrast variations. The feasibility and performance of this segmentation method was successfully demonstrated through computing standard metrics and clinical indices with respect to the ground truth on the MICCAI 2012 RVSC dataset (6). Results indicate improvements in terms of accuracy and computational time compared with the existing RV segmentation methods.

ACKNOWLEDGMENTS

The authors thank Dr. Caroline Petitjean, at the University of Rouen, France, for providing the datasets and also evaluating our results compared with the ground truth. Without the dataset and her help this study would not be possible.

REFERENCES

- Haddad F, Doyle R, Murphy DJ, Hunt SA. Right ventricular function in cardiovascular disease, part II. Pathophysiology, clinical importance, and management of right ventricular failure. *Circulation* 2008; 117:1717–1731.
- Sanz J, Conroy J, Narula J. Imaging of the right ventricle. *Cardiol Clin* 2012;30:189–203.
- Mehta SR, Eikelboom JW, Natarajan MK, Diaz R, Yi C, Gibbons RJ, Yusuf SH. Impact of right ventricular involvement on mortality and morbidity in patients with inferior myocardial infarction. *J Am Coll Cardiol* 2001;37:37–43.
- Kevin LG, Barnard M. Right ventricular failure. *Continuing Education in Anaesthesia, Critical Care & Pain* 2007;7:89–94.

5. Vitarelli A, Terzano C. Do we have two hearts?. new insights in right ventricular function supported by myocardial imaging echocardiography. *Heart Fail Rev* 2010;15:39–61.
6. Petitjean C, Zuluaga MA, Bai W, Dacher JN, Grosgeorge D, Caudron J, Ruan S, Ayed IB, Cardoso MJ, Chen HC, et al. Right ventricle segmentation from cardiac MRI: a collation study. *Med Image Anal* 2015;19:187–202.
7. Yuan C, Kerwin WS, Ferguson MS, Polissar N, Zhang S, Cai J, Hatsukami TS. Contrast-enhanced high resolution MRI for atherosclerotic carotid artery tissue characterization. *J Magn Reson Imaging* 2001;15:62–67.
8. Lima JC, Desai MY. Cardiovascular magnetic resonance imaging: current and emerging applications. *J Am Coll Cardiol* 2004;44:1164–1171.
9. Frangi AF, Niessen WJ, Viergever MA. Three-dimensional modeling for functional analysis of cardiac images, a review. *IEEE Trans Med Imaging* 2001;20:2–5.
10. Tavakoli V, Amini AA. A survey of shaped-based registration and segmentation techniques for cardiac images. *Comput Vis Image Underst* 2013;117:966–989.
11. Heimann T, Meinzer H-P. Statistical shape models for 3D medical image segmentation: a review. *Med Image Anal* 2009;13:543–563.
12. Geva T. MRI is the preferred method for evaluating right ventricular size and function in patients with congenital heart disease. *Circ Cardiovasc Imaging* 2014;7:190–197.
13. Caudron J, Fares J, Vivier PH, Lefebvre V, Petitjean C, Dacher JN. Diagnostic accuracy and variability of three semi-quantitative methods for assessing right ventricular systolic function from cardiac MRI in patients with acquired heart disease. *Eur Radiol* 2011;21:2111–2120.
14. Bonnemains L, Mandry D, Marie PY, Micard E, Chen B, Vuissoz PA. Assessment of right ventricle volumes and function by cardiac MRI: quantification of the regional and global interobserver variability. *Magn Reson Med* 2012;67:1740–1746.
15. Zuluaga MA, Cardoso MJ, Modat M, Ourselin S. Multi-atlas propagation whole heart segmentation from MRI and CTA using a local normalised correlation coefficient criterion. *Funct Imaging Model Heart* 2013;7945:174–181.
16. Ou Y, Doshi J, Erus G, Davatzikos C. Multi-atlas segmentation of the cardiac MR right ventricle. *Proceedings of 3D cardiovascular imaging: a MICCAI segmentation challenge*. Nice, France; 2012.
17. Bai W, Shi W, O'Regan DP, Tong T, Wang H, Jamil-Copley S, Peters NS, Rueckert D. A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: application to cardiac MR images. *IEEE Trans Med Imaging* 2013;32:1302–1315.
18. Petitjean C, Dacher J-N. A review of segmentation methods in short axis cardiac MR images. *Med Image Anal* 2011;15:169–184.
19. Nambakhsh C, Rajchl M, Yuan J, Peters T, Ben-Ayed I. Rapid automated 3D RV endocardium segmentation in MRI via convex relaxation and distribution matching. *Proc MICCAI RV Segmentation Challenge*, 2012.
20. Grosgeorge D, Petitjean C, Dacher J-N, Ruan S. Graph cut segmentation with a statistical shape model in cardiac MRI. *Comput Vis Image Underst* 2013;117:1027–1035.
21. Maier OMO, Jimenez D, Santos A, Ledesma-Carbayo MJ. Segmentation of RV in 4D cardiac MR volumes using region-merging graph cuts. *Comput Cardiol* 2012;2012:697–700.
22. Ringenber J, Deo M, Devabhaktuni V, Berenfeld O, Boyers P, Gold J. Fast, accurate, and fully automatic segmentation of the right ventricle in short-axis cardiac MRI. *Comput Med Imaging Graph* 2014;38:190–201.
23. Wang C, Peng C, Chen H. A simple and fully automatic right ventricle segmentation method for 4-dimensional cardiac MR images. *Proc of 3D Cardiovascular Imaging: a MICCAI Segmentation Challenge*, 2012.
24. Lu X, Wang Y, Georgescu B, Littman A, Comaniciu D. Automatic delineation of left and right ventricles in cardiac MRI sequences using a joint ventricular model. *Funct Imaging Model Heart* 2011;6666:250–258.
25. Mahapatra D, Buhmann JM. Automatic cardiac RV segmentation using semantic information with graph cuts. *IEEE 10th International Symposium on Biomedical Imaging (ISBI)*. 2013;2013:1106–1109.
26. Moolan-Feroze O, Mirmehdi M, Hamilton M, Bucciarelli-Ducci C. Segmentation of the right ventricle using diffusion maps and Markov random fields. In *Medical image computing and computer-assisted intervention—MICCAI 2014*. New York: Springer; 2014. p 682–689.
27. Mahapatra D. Automatic cardiac segmentation using semantic information from random forests. *J Digit Imaging* 2014;27:794–804.
28. LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. In *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS)*. 2010. p 253–256.
29. Szegedy C, Toshev A, Erhan D. Deep neural networks for object detection. *Adv Neural Inf Process Syst* 2013;26:2553–2561.
30. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. Overfeat: integrated recognition, localization and detection using convolutional networks. *International Conference on Learning Representations (ICLR2014)*, 2014.
31. Krizhevsky A, Sutskever I, Hinton GE. Image net classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 2012:1097–1105.
32. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 2013;35:1798–1828.
33. Bengio Y. Learning deep architectures for AI. *Found Trends Mach Learn* 2009;2:1–127.
34. Vincent P, Larochelle H, Bengio Y, Manzagol P-A. Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th International Conference on Machine Learning*. 2008. p 1096–1103.
35. Baldi P. Autoencoders, unsupervised learning, and deep architectures. *ICML Unsupervised and Transfer Learning* 2012;27:37–50.
36. Deng L, Yu D. *Deep Learning: methods and applications*. Foundations and trends in signal processing. Hanover, MA: Now Publishers Incorporated; 2014.
37. Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P-A. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *J Mach Learn Res* 2010;11:3371–3408.
38. Goodfellow I, Bengio Y, Courville A. *Deep learning (adaptive computation and machine learning series)*. Cambridge: The MIT Press; 2016.
39. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal* 2016;30:108–119.
40. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Cogn Model* 1988;5.
41. Koikkalainen J, Töllli T, Lauerma K, Antila K, Mattila E, Lilja M, Lötjönen J. Methods of artificial enlargement of the training set for statistical shape models. *IEEE Trans Med Imaging* 2008;27:1643–1654.
42. Caudron J, Fares J, Lefebvre V, Vivier PH, Petitjean C, Dacher JN. Cardiac MRI assessment of right ventricular function in acquired heart disease: factors of variability. *Acad Radiol* 2012;19:991–1002.
43. Zuluaga MA, Cardoso MJ, Modat M, Ourselin S. Multi-atlas propagation whole heart segmentation from MRI and CTA using a local normalized correlation coefficient criterion. In: *Functional imaging and modeling of the heart*. New York: Springer; 2013. p 174–181.
44. Maier OMO, Jiménez D, Santos A, Ledesma-Carbayo MJ. Segmentation of RV in 4D cardiac MR volumes using region-merging graph cuts. *Comput Cardiol* 2012:697–700.
45. Nambakhsh CM, Yuan J, Punithakumar K, Goela A, Rajchl M, Peters TM, Ayed IB. Left ventricle segmentation in MRI via convex relaxed distribution matching. *Med Image Anal* 2013;17:1010–1024.
46. van Assen HC, Danilouchkine MG, Frangi AF, Ordás S, Westenberg JJ, Reiber JH, Lelieveldt BP. Spasm: a 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data. *Med Image Anal* 2006;10:286–303.
47. Kaus MR, von Berg J, Weese J, Niessen W, Pekar V. Automated segmentation of the left ventricle in cardiac MRI. *Med Image Anal* 2004;8:245–254.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article

Fig. S1. Bland-Altman plots for EDV, ESV, EF of Test1Set (top row) and Test2Set (bottom row).

Fig. S2. Endocardial contours of RV at ED from base to apex (Patient #21 from Test1Set). Our method resulted in best DM (0.93) for this case. Red and yellow correspond to Step 2 and Step 3 segmentation results, respectively.

Fig. S3. Endocardial contours of RV at ES from base to apex (Patient #21 from Test1Set). Our method resulted in best DM (0.93) for this case. Red and yellow correspond to Step 2 and Step 3 segmentation results, respectively.

Fig. S4. Endocardial contours of RV at ED from base to apex (Patient #29 from Test1Set). Our method resulted in worst DM (0.74) for this case. Red

and yellow correspond to Step 2 and Step 3 segmentation results, respectively.

Fig. S5. Endocardial contours of RV at ES from base to apex (Patient #29 from Test1Set). Our method resulted in worst DM (0.74) for this case. Red and yellow correspond to Step 2 and Step 3 segmentation results, respectively.