Article:
Hu, H, Pan, N, Liu, H et al. (4 more authors) (2021) Automatic segmentation of left and right ventricles in cardiac MRI using 3D-ASM and deep learning. Signal Processing: Image Communication, 96. 116303. ISSN 0923-5965

https://doi.org/10.1016/j.image.2021.116303

Automatic Segmentation of Left and Right Ventricles in Cardiac MRI using 3D-ASM and Deep learning

Huaifei Hu [a, b, c, 1], Ning Pan [a, b, c, 1], Haihua Liu [a, b, c, *], Liman Liu [a, b, c], Tailang Yin [d, *], Zhigang Tu [e,*], Alejandro F. Frangi [f, g, h, i *]

a. College of Biomedical Engineering, South Central University for Nationalities, Wuhan 430074, China;
b. Hubei Key Laboratory of Medical Information Analysis and Tumor Diagnosis & Treatment, Wuhan 430074, China;
c. Key Laboratory of Cognitive Science, State Ethnic Affairs Commission, Wuhan 430074, China;
d. Reproductive Medicine Center, Renmin Hospital of Wuhan University, Wuhan, Hubei 430060, China;
e. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote sensing, Wuhan University, 430079, Wuhan, China.
f. Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), School of Computing, University of Leeds, Leeds, UK
g. Biomedical Imaging Sciences Department, Leeds Institute for Cardiovascular and Metabolic Medicine (LICAMM), School of Medicine, University of Leeds, Leeds, UK
h. Medical Imaging Research Centre, Cardiovascular Sciences Department, KU Leuven, Leuven, Belgium
i. Medical Imaging Research Centre, Electrical Engineering Department, KU Leuven, Leuven, Belgium

* Corresponding author.
E-mail addresses: huaifeihu@mail.scuec.edu.cn (H. Hu), panning@mail.scuec.edu.cn (N. Pan), liuhh@mail.scuec.edu.cn (H. Liu), limanliu@mail.scuec.edu.cn (Li, Liu), reproductive@whu.edu.cn (T, Yin), tuzhigang@whu.edu.cn (Z. Tu), A.Frangi@leeds.ac.uk (A.F. Frangi).

[1] These authors contributed equally to this work.

ABSTRACT

Segmentation of the left and right ventricles in cardiac MRI (Magnetic Resonance Imaging) is a prerequisite step for evaluating global and regional cardiac function. This work presents a novel and robust schema for MRI segmentation by combining the advantages of deep learning localisation and 3D-ASM (3D Active Shape Model) restriction without any user interaction. Three fundamental techniques are exploited: 1) manual 2D contours are used to build distance maps to get 3D ground truth shape, 2) derived right ventricle points are employed to rotate the coarse initial shape for a refined bi-ventricle initial estimation, 3) segmentation results from deep learning are utilised to build distance maps for the 3D-ASM matching process to help image intensity modelling. The datasets used for experimenting the cine MRI data are 1000 cases from UK Biobank, 500 subjects are selected to train CNN (Convolution Neural Network) parameters, and the remaining 500 cases are adopted for validation. Specifically, cases are used to rebuild point distribution and image intensity models, and also utilised to train CNN. In addition, the left 500 cases are used to perform the validation experiments. For the segmentation of the RV (Right Ventricle) endocardial contour, LV (Left Ventricle) endo- and epicardial contours, overlap, Jaccard similarity index, Point-to-surface errors and cardiac functional parameters are calculated. Experimental results show that the proposed method has advantages over the previous approaches.

Keywords: Left and right ventricle segmentation; automatic initialisation; deep learning; statistical shape models.

1. INTRODUCTION

Being one of the top lethal factors [1], cardiovascular disease has received considerable concern in clinical practice. Thus quantitative analysis of cardiac function is a critical step for the better patient management, risk evaluation and therapy decision. To evaluate the clinical parameters of the heart, such as ejection fraction, myocardial mass, the volumes of the heart has to be computed. To calculate such volumes,

the primary step is to draw the contours of the heart based on MRI due to high discrimination among endocardium, epicardium, right ventricle and other tissues. In clinical operation, manual delineate task is not only dull, troublesome and introduces intra and inter- rater variability for a radiologist when facing large-scale cardiac images. For this purpose, cardiac segmentation has aroused extensive attention in medical image analysis.

In recent years, several challenges has been hold for cardiac segmentation, e.g., MICCAI2009 [2], MICCAI-STACOM2011[3], MICCAI2012 [4], ACDC[5]. These challenges have greatly promoted the development of medical image processing, a variety of semi-automatic/automatic cardiac segmentation methods have been exploited. These algorithms include image feature based method, atlas registration and learning-based methods, etc. For a detailed review of previous work, the reader can refer to recent topical literatures [6-9].

Image feature based methods perform image segmentation based on the attributes of the image itself, including, for instance, thresholding, region growing, and graph cuts[10-12]. Efficient and straightforward, segmentation methods based on image features are the most basic and widely used algorithms yet they are mostly only helpful aided with considerable manual intervention. Since the image feature method only depends on the shallow features of the image, in the actual process, the surrounding tissues with similar characteristics to the heart interfere with each other, and the segmentation result is susceptible to noise.

Atlas registration method uses atlas information to convert image segmentation into image registration and fusion [13]. It mainly includes three steps: atlas selection, registration and fusion. To reduce computation load and improve robustness, an spatial transformation is adopted to maximise similarity between float and fixed images. Due to the limited capacity of the atlas, this method is difficult to process complex shapes and time consuming.

The learning-based method [14-17] mainly uses deep learning algorithms, especially convolutional neural networks. Mimicking human visual information processing mechanisms, deep learning can

automatically learn multi-level image features and map images to a high-level feature space [18, 19]. Because of excellent feature extraction and expression capacity, deep learning is widely used in medical image segmentation [12]. However, high-level model-based information is not explicit owing to the low-level nature of the inputs and subsequent pooling operations, resulting in occasionally implausible segmentation results.

Different from the above mentioned algorithms, using a priori shape constraint to segment organs from medical images, statistical shape models have a widely application for 3D or 4D (3D+t). Methods adopting a priori knowledge can do a robust and accurate segmentation in medical image analysis. The shape constraint is called PDM (Point Distribution Model) which is deformed to outline an unknown object within an unknown image. When the Statistical Shape Models (SSMs) are utilised for cardiac segmentation, two elements are needed: a starting predict of the bi-ventricular position, and an appearance of the image called IIM (Image Intensity Model). For each point in the 3D shape belonging to the cardiac images, the 3D-ASM captures the image intensity information of the corresponding point from all the training shapes, and allows the image stack slices in the training set to intersect the 3D shape. By sampling at each side of the landmarks, perpendicularly to the boundary of the intercepted shape, the IIM can be trained by calculating the second order statistics for the normalised image gradients [18]. Under the joint action of the PDM and the IIM, the initial shape keeps approaching the target contour. After several iterations, a 3D contour for cardiac images can be finally produced.

The contributions of this work are three-fold. Firstly, we introduce a fully automatic algorithm to initialise bi-ventricle for cardiac MRI segmentation, by using deep learning model and complex transformation techniques to predict an initial position of the heart, hence an initial shape for both left and right ventricles can be created. Secondly, we invent distance map techniques by constructing a full CNN, and the distance maps are applied to help IIM in the cardiac segmentation by 3D-ASM. Third, we proposed a schema to combine CNN and 3D-ASM for left and right ventricles segmentation.

The rest of this paper is organised as follows. In the next section, we introduce our pipeline for cardiac image segmentation. Then we describe the data source used in this work. In the experiments section, we make comparisons to show the advantages of our method. At last, we make discussion and conclusion of our work.

2. Method

2.1. Overview

In this section, our pipeline exploits PDM, IIM reconstruction and automatic segmentation of left and right ventricles using 3D-ASM, here the statistical shape model adopted is SPASM (Sparse Active Shape Model) [19]. As described in Fig. 1, our approach includes several steps, i.e. initial shape optimisation, construction of PDM and IIM, CNN training and 3D-ASM modelling & cardiac quantification. Firstly, we organised the raw cardiac MRI subjects with ground truth according to the time frames per case. The position of the hear is initially guessed based on extracting selected landmarks and then providing an initial LV approximation. However, we need the bi-ventricular initial shape for both left and RVs segmentation. Right ventricle points from PDM are fit to the corresponding manual contour points [20]. With complex transformation and Procrustes analysis [21], a bi-ventricular initial shape can be derived for 3D-ASM segmentation. Second, we created distance maps for LV and RV by computing Euclidean distance in the manual contours, and this distance maps are merged into IIM to perform 3D-ASM segmentation. Thirdly, the results from 3D-ASM are utilised for PDM and IIM training. Fourthly, we applied CNN to train these organised subjects. Then, the test subjects are sent to the trained CNN for coarse segmentation. The masks for left and right ventricles can be obtained separately. To get the bi-ventricular initial shape, we use the same method in the PDM and IIM reconstruction. Instead of using manual contours, the CNN segmentation contours is utilised to obtain the initial shape for both left and right ventricles. Then we created distance maps for LV and RV by computing Euclidean distance in the CNN segmented masks, and this distance maps

are a penalty item of IIM for 3D-ASM segmentation. At last, the segmentations of 3D-ASM are employed for LV-RV quantification.

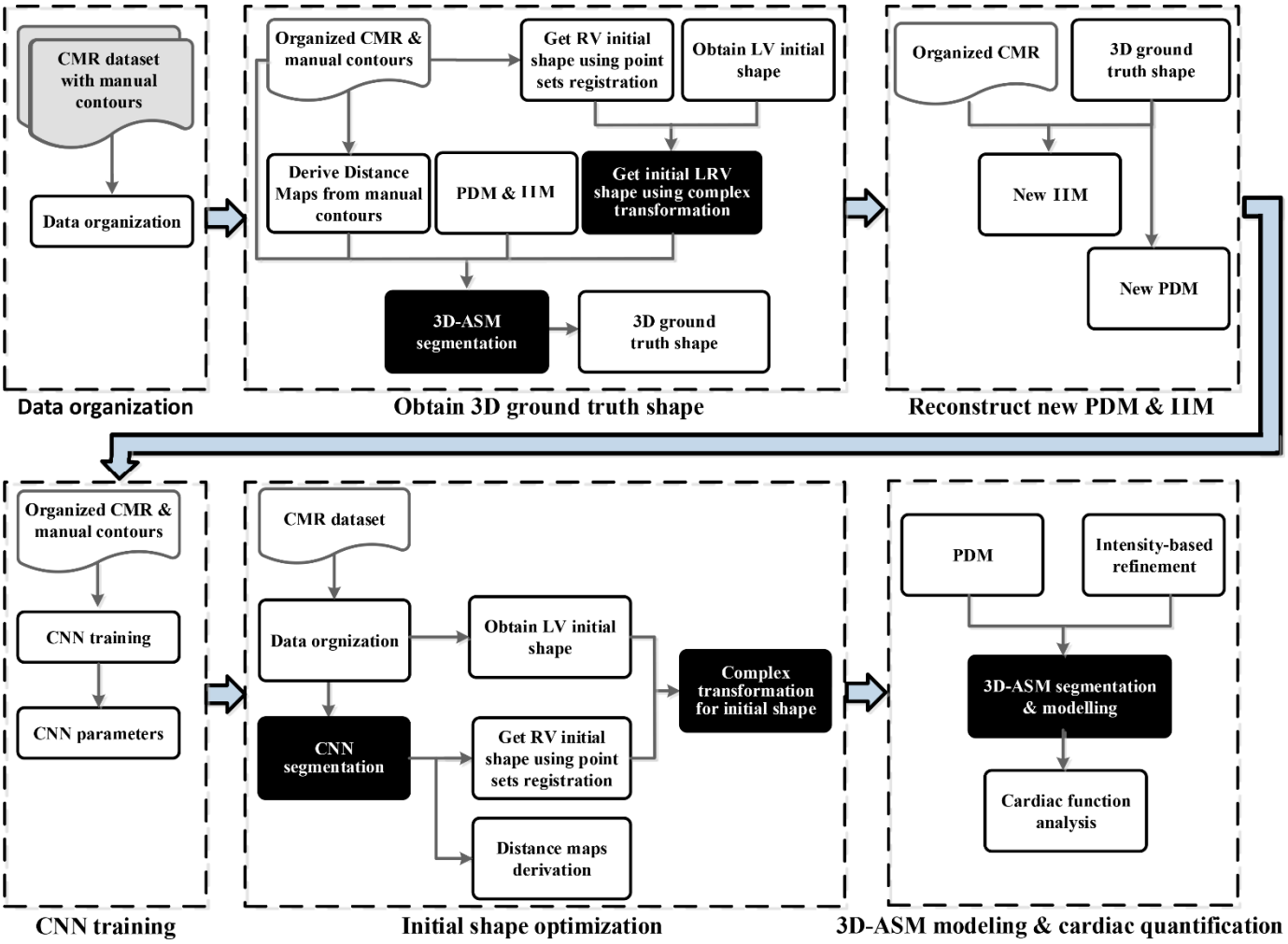


Fig 1. The pipeline of the proposed algorithm.

2.2. Step I: Initialisation of 3D-ASM

The initial shape for LV can be obtained by roughly scaling and locating the mean shape of the model. Three points are marked by the user, two points (AOTIA, MITRAL) are at the basal level, and a third one (LVAPEX) in the apical slice. In the mean shape, corresponding anatomical landmarks were defined by a skilled operator. Similarity transformation is applied to align the mean shape to the landmarks.

To obtain initial shape for 3D-ASM without manual intervene, we adopt the algorithm proposed by Albà et al. [22] to get the AOTIA, MITRAL and LVAPEX landmarks for LV. First , the location of LV is estimated by intersection among the 4CH (4 Chambers), 2CH (2 Chambers) views in LAX (Long-axis) and the views

in SAX (Short-axis). Then in 4CH view image, the intersection points from basal and apex level are utilised to train a random forest regressor using two feature descriptors (i.e., the Histogram of Oriented Gradients and Gabor Filters). At last, AOTIA, MITRAL and LVAPEX landmarks can be derived for LV initial shape (See Fig 2).

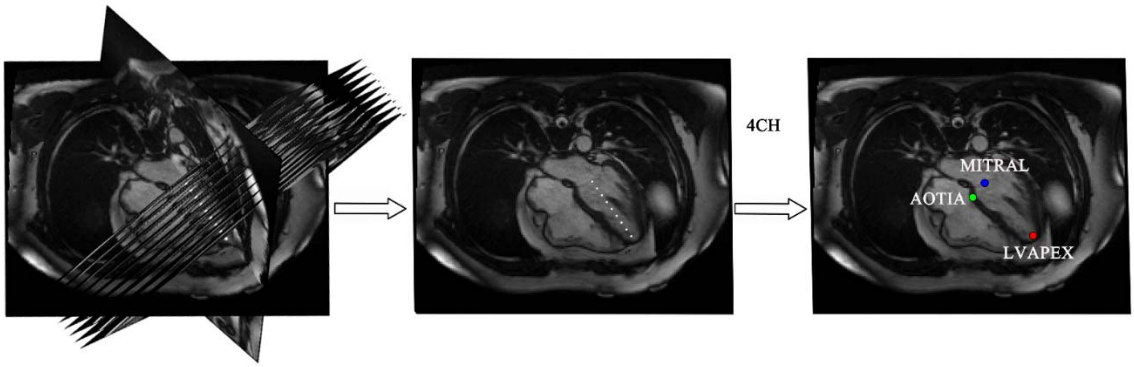


Fig 2. Definition of AOTIA, MITRAL and LVAPEX.

Then an initial shape for LV can be automatic obtained (See Fig. 3(a), (c)), Procrustes analysis [21] is then employed to get a bi-ventricular model initialisation (See Fig. 3(b), (d)). However, the RV points in the initial shape deviate too much from RV contour. Consider that shape for RV is more complex than that of LV, a complex transformation is needed to rotate the bi-ventricular initial shape to a proper location.

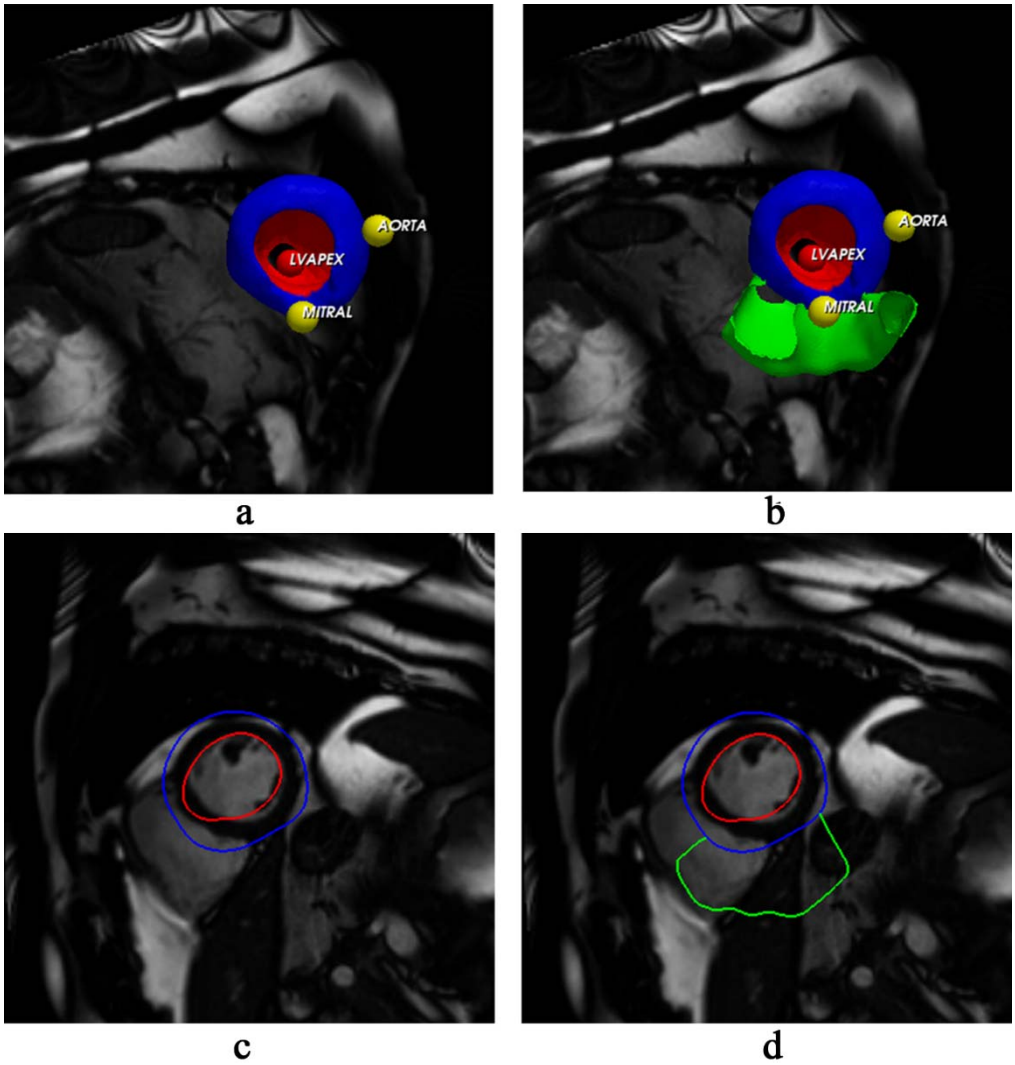


Fig 3. Bi-ventricular model initialisation. (a) LV initial shape and three points (AOTIA, MITRAL and LVAPEX), (b) Bi-ventricular initial shape using Procrustes analysis, (c) LV initial shape in short-axis view, (d) Bi-ventricular initial shape in short-axis view.

2.3. Step II: Complex transformation for Initial shape

To get a good location, point-set registration [23] is employed to align bi-ventricular initial shape to the manual contours. First, the manual contour points for RV in base slice are fitted to a plane (See Fig. 4(a)). Second, the manual contour points and points from PDM are rotated to make the fitted plane perpendicular to Z-axis (See Fig. 4(b)). Third, the RV points in the initial shape is extracted with the RV points from ground truth (See Fig. 4(c)). Consider shape similarity for the RV points from PDM and ground truth, some points for RV from PDM should be removed before point-set registration. In Fig. 4(c), supposing $Z_{\min}$ is the

smallest Z coordinate value obtained from the marked points in PDM at the basal level, all points from PDM with Z coordinate value bigger than $Z_{\min}$ is removed. Also, another fitting plane from apex slice points in ground truth is built, we remove points from PDM below the fitting plane in apex slice. At last, two point sets with similar shape are derived for point-set registration (See Fig. 4(d)).

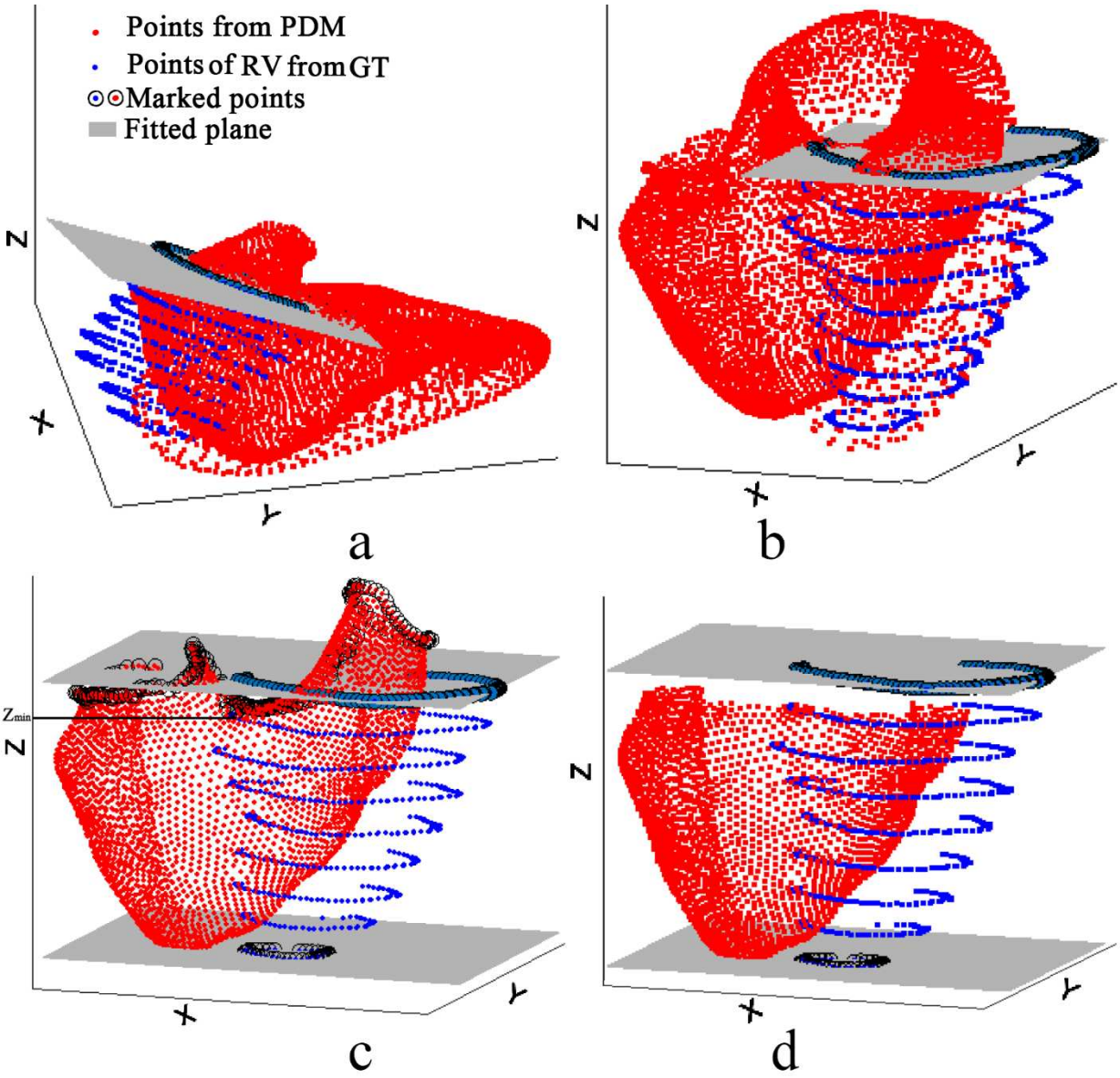


Fig 4. Initial shape rotation. (a) LV, RV points from PDM and RV points from ground truth at their original position, (b) Rotated points from PDM and ground truth, (c) RV point sets from PDM and ground truth, (d) RV point sets with similar shape from PDM and ground truth. GT: ground truth.

In Fig. 4(d), point-set registration is applied to align RV point sets from PDM with that from ground truth, where a transformed point sets can be obtained (See Fig. 5(a)). Here Procrustes analysis is applied again in

point sets of PDM from Fig. 4(d) and Fig. 5(a), and a transformed matrix T is derived to rotate LV points in the initial shape defined as follows:

$$\mathrm{T}=\begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

where α represents the rotation angle in short-axis view. We initiallise the SSM to coincide with the center of the LV point cloud. Then, the LV points are rotated α degrees around Z-axis. We also supplement the removed points in the RV of the initial shape, and a bi-ventricular initial shape can be obtained (See Fig. 5(b)). Then the bi-ventricular initial shape is rotated to its original position (See Fig. 5(c)), and a better location for RV can been seen in short-axis view (See Fig. 5(d)).

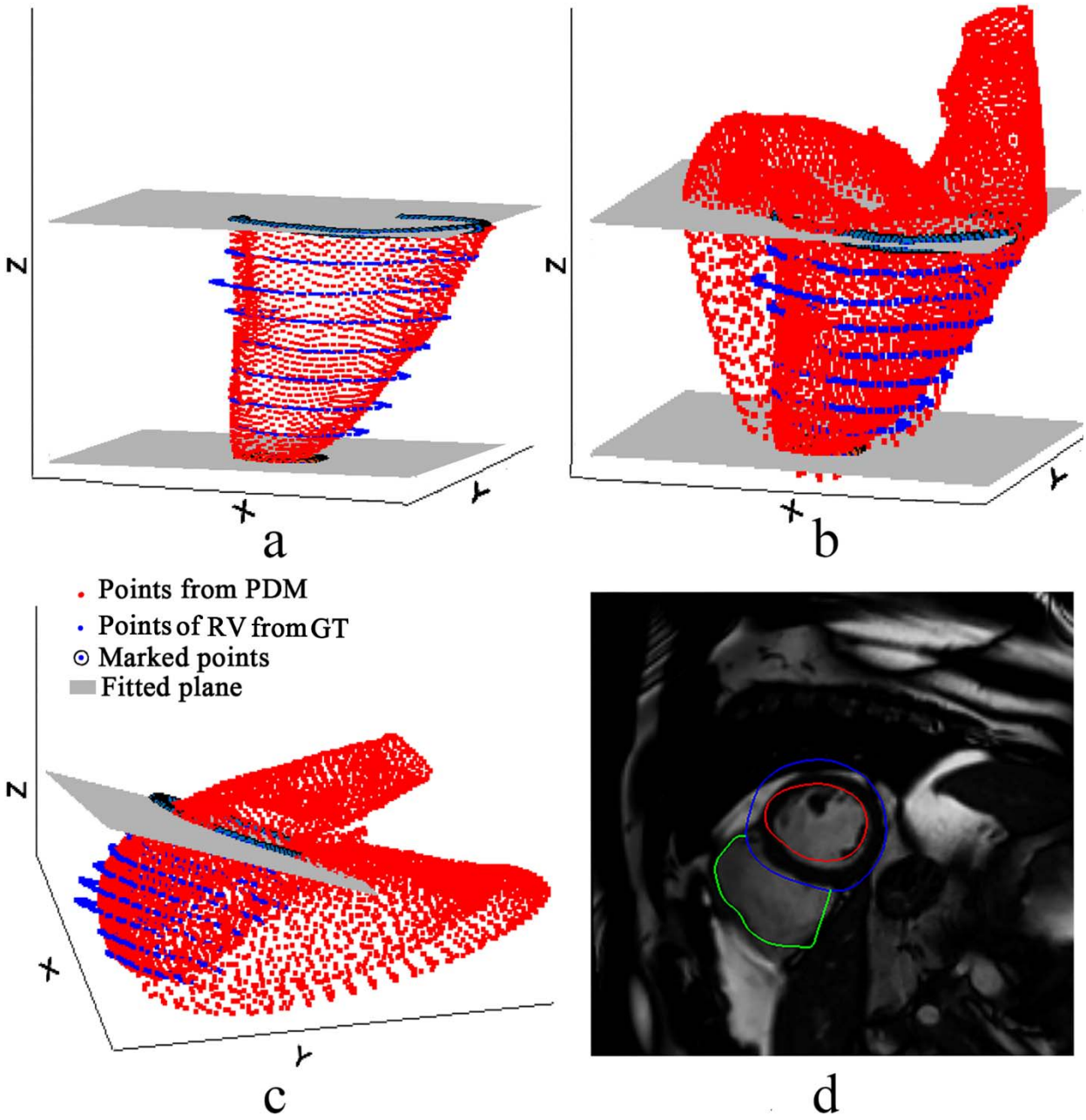


Fig 5. Complex transformation for initial shape rotation. (a) Registered points from PDM and RV contour points from ground truth, (b) LV and RV points using complex transformation, RV contour points from ground truth, (c) LV, RV points from PDM and RV points from ground truth at their original position, (d) Bi-ventricular initial shape in short-axis view. GT: ground truth.

2.4. Step III: Distance-restricted LRV refinement

After the bi-ventricular initial shape is built, 3D ground truth can be derived from the 2D manual contours by our proposed distance map techniques.

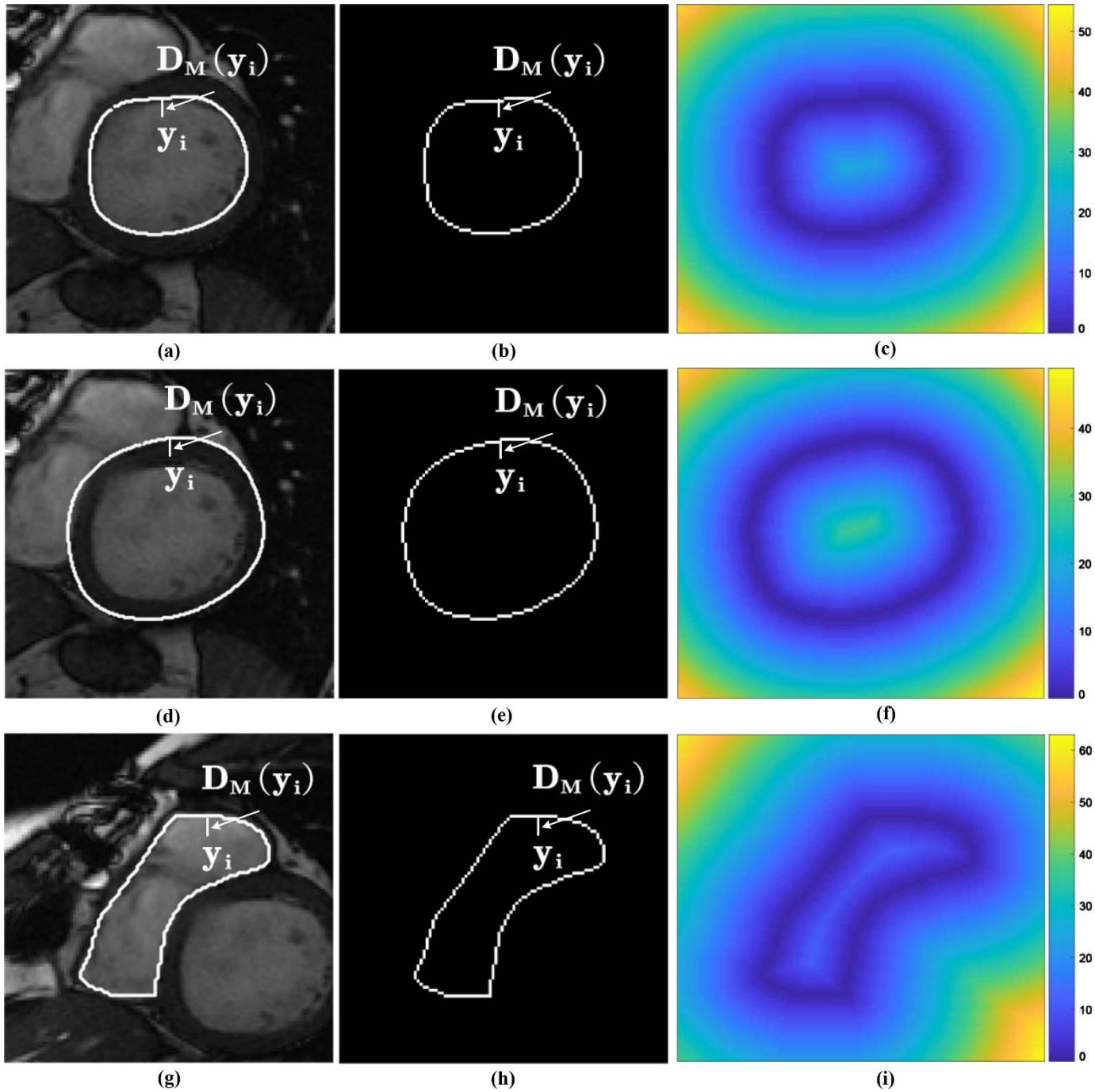


Fig 6. Distance maps for ground truth of LV and RV. (a) Original cardiac MR is depicting endocardial contour, (b) Binary image of endocardial contour, (c) Distance map for endocardial contour, (d) Original cardiac MR is depicting epicardial contour, (e) Binary image of epicardial contour, (f) Distance map of the epicardial contour, (g) Original cardiac MR is depicting RV contour, (h) Binary image of RV contour, (i) Distance map of the RV contour, (Cropped for better view).

Figure 6 shows three distance maps constructed from the RV endocardial, LV endocardial and epicardial contours derived from 2D ground truth, respectively. For a point y_i, the distant value $D_M(y_i)$ is defined

as follows:

$$D_M(y_i) = d_{min}(y_i, \partial M) \tag{2}$$

where $d_{min}(y_i, \partial M)$ is the minimal distance from point y_i to the ground truth contour ∂M.

In the following 3D-ASM segmentation, the above steps are executed for all short-axis cardiac images. The smaller the value of a pixel in the distance maps, the more likely the pixel belongs to the cardiac contours. Since the values of the manual contour in the distance maps are purposely made to be zero, so there exists a driven force that can lead the initial shape to the actual cardiac boundary. The detailed fitting process will be introduced in the next paragraph.

2.5. Step IV: Three-dimensional Image-driven Adaptation of ASM

Let us assume a training set with M shapes, and three-dimensional points described as $\mathbf{x}_j^i = (x_j^i, y_j^i, z_j^i)$ with $i = 1 \ldots M$ and $j = 1 \ldots N$. Let $\mathbf{s}_i = (x_1^i, y_1^i, z_1^i, \ldots, x_N^i, y_N^i, z_N^i)^T$ be the ith vector representing the shape of the ith LRV surface and $S = \left[\mathbf{s}^1, \ldots, \mathbf{s}^M\right]$ set all training shapes in matrix form. All nuisance pose parameters (e.g. translation, rotation and scaling) have been removed S using Generalised Procrustes Analysis [24]. Hence, the shape class mean of S, $\bar{\mathbf{s}}$ can be written as

$$\bar{\mathbf{s}} = \frac{1}{M}\sum_{i=1}^{M} \mathbf{s}_i \tag{3}$$

and the shape class covariance is

$$\mathbf{C} = \frac{1}{M-1}\sum_{i=1}^{M} (s_i - \bar{s})(s_i - \bar{s})^T \tag{4}$$

The shape class covariance is represented in a low-dimensional space or Principal Component Analysis (PCA). This produces l eigenvectors $\mathbf{\Phi}_l = [\boldsymbol{\varphi}_1 \boldsymbol{\varphi}_2 \ldots \boldsymbol{\varphi}_l]$, and corresponding eigenvalues $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_l)$ of the covariance matrix computed via Singular Value Decomposition (SVD). Under the constraint of multi-dimensional Gaussian probability distribution, any shape in the shape class can be represented as

$$\mathbf{s} \approx \bar{\mathbf{s}} + \mathbf{\Phi}_l \mathbf{b} \tag{5}$$

where $\mathbf{b}$ is the PDM parameters and restricted to $|b_i| \le \beta\sqrt{\lambda_i}$ fall within 99% of the shape class distribution if $\beta = 3$. The parameters that reconstruct a shape $\bar{\mathbf{s}}$ are estimated from

$$\mathbf{b} = \mathbf{\Phi}_l^T (\mathbf{s} - \bar{\mathbf{s}}) \tag{6}$$

The components of $\mathbf{b}$ are the projection coefficients of mean-centred shapes $(\mathbf{s} - \bar{\mathbf{s}})$ along the columns of $\mathbf{\Phi}_l$.

There are two key components for 3D-ASM segmentation, one is the shape constraint called PDM, and the other one is IIM. In 3D-ASM, the IIMs capture local intensity distribution along cardiac boundaries. In this research, 1D intensity profiles are sampled with a length size $m = 15$ pixels normal to the myocardial contours. For the ith landmark, the mean intensity profile $\bar{\mathbf{g}}_i$, and the corresponding image intensity covariance $\mathbf{S}_{\mathbf{g}_i}$ are estimated. In the process of 3D-ASM matching, the intersections of the 3D PDM with all imaging planes define a stack of 2D contours oriented in 3D space. For each landmark, the 3D-ASM runs to seek the best-matching location where the intensity profile is derived along the normal to the boundaries and over the imaging planes. To obtain the best-matching location or the candidate point, $\mathbf{y}_i$ for each landmark, the Mahalanobis distance is minimised between profile sampled at the candidate position, viz. $\mathbf{g}_i(\mathbf{y}_i)$, and the mean profile, $\bar{\mathbf{g}}_i$, according to:

$$\mathrm{y}_\mathrm{i}^\mathrm{o} = \underset{\mathrm{y_i}}{\mathrm{argmin}}\left((\mathrm{g}(\mathrm{y_i}) - \bar{\mathrm{g}}_\mathrm{i})^\mathrm{T} \mathrm{S}_{\mathrm{g}_i}^{-1} (\mathrm{g}(\mathrm{y_i}) - \bar{\mathrm{g}}_\mathrm{i})\right) \tag{7}$$

Consider the sparse property of CMR images, inevitably, no image slices in the stack can be found to intersect with the mesh triangles. In such situation, the points containing the triangles are updated by PDM instead of IIM, the updated mechanism is defined as follows:

$$\mathrm{w}(\mathrm{p},\mathrm{q}) = \exp\left\{-\frac{\|\mathrm{p}-\mathrm{q}\|^2}{2\sigma^2}\right\} \tag{8}$$

where q represents any image-driven point at a Gaussian kernel centred position and σ is the width of the kernel. p is a neighbouring point driven by equation (8) under the condition that no IIM can be available.

During the 3D-ASM segmentation, IIM adopts Eq. (9) to select candidate points. We optimise the process by employing the generated distance maps. The optimal candidate points can be derived using the following equation:

$$y_i^o = \underset{y_i}{\operatorname{argmin}}\left((g(y_i) - \bar{g}_i)^T S_{g_i}^{-1} (g(y_i) - \bar{g}_i) + \eta * D_M(y_i)\right) \tag{9}$$

where $\eta = 3$ is a penalty factor.

After Step IV, we can obtain the 3D ground truth of bi-ventricles, a surface mesh with the same topology containing LV endo-/epi- and RV endo-cardiums. Then the derived mesh with the image data are employed to retrain PDM and IIM. Technical details can be referred in [25].

2.6. Step V: CNN Distance Maps and 3D-ASM segmentation

In this study, a fully convolutional network (FCN) with 16 layers is adopted for automatically segmenting the LV myocardium, blood-pool and RV blood-pool for short-axis slices as depicted in Fig 7. In the architecture, batch normalisation and RELU follows each convolutional layer, max-pool layer is used to reduce or down sample the dimensionality of the input image. Helped by max-pool layer, feature detection is independent of noise and small changes like image rotation or tilting. The upscale layers which will up-sample the input image to a higher resolution. The softmax layer applies a softmax function to the input and converts the output of the last layer into a probability distribution.

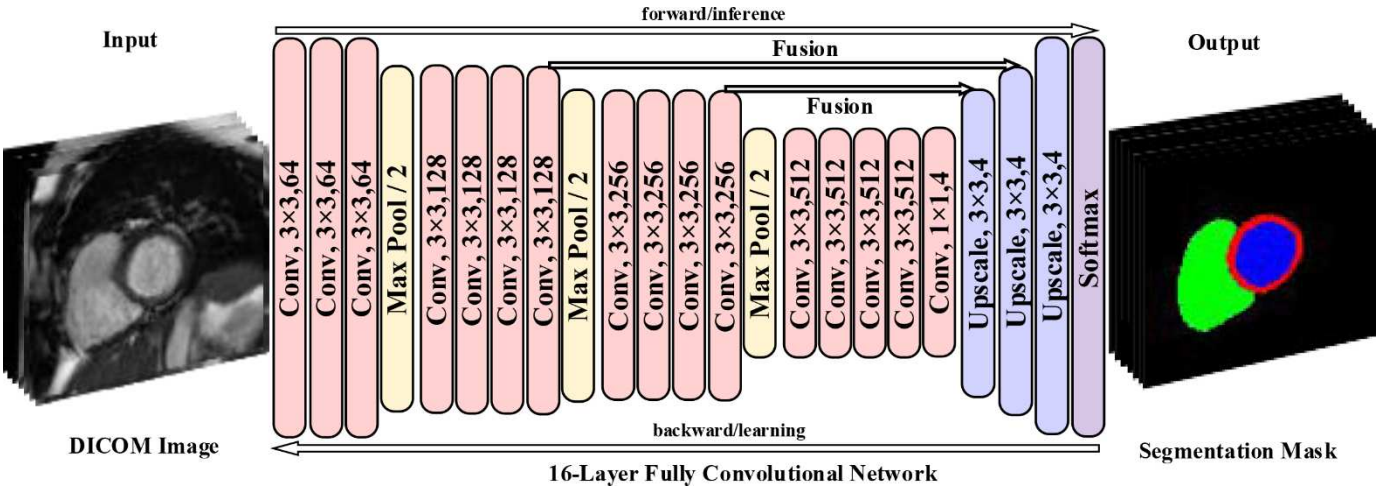


Fig 7. The architecture of the fully neural network.

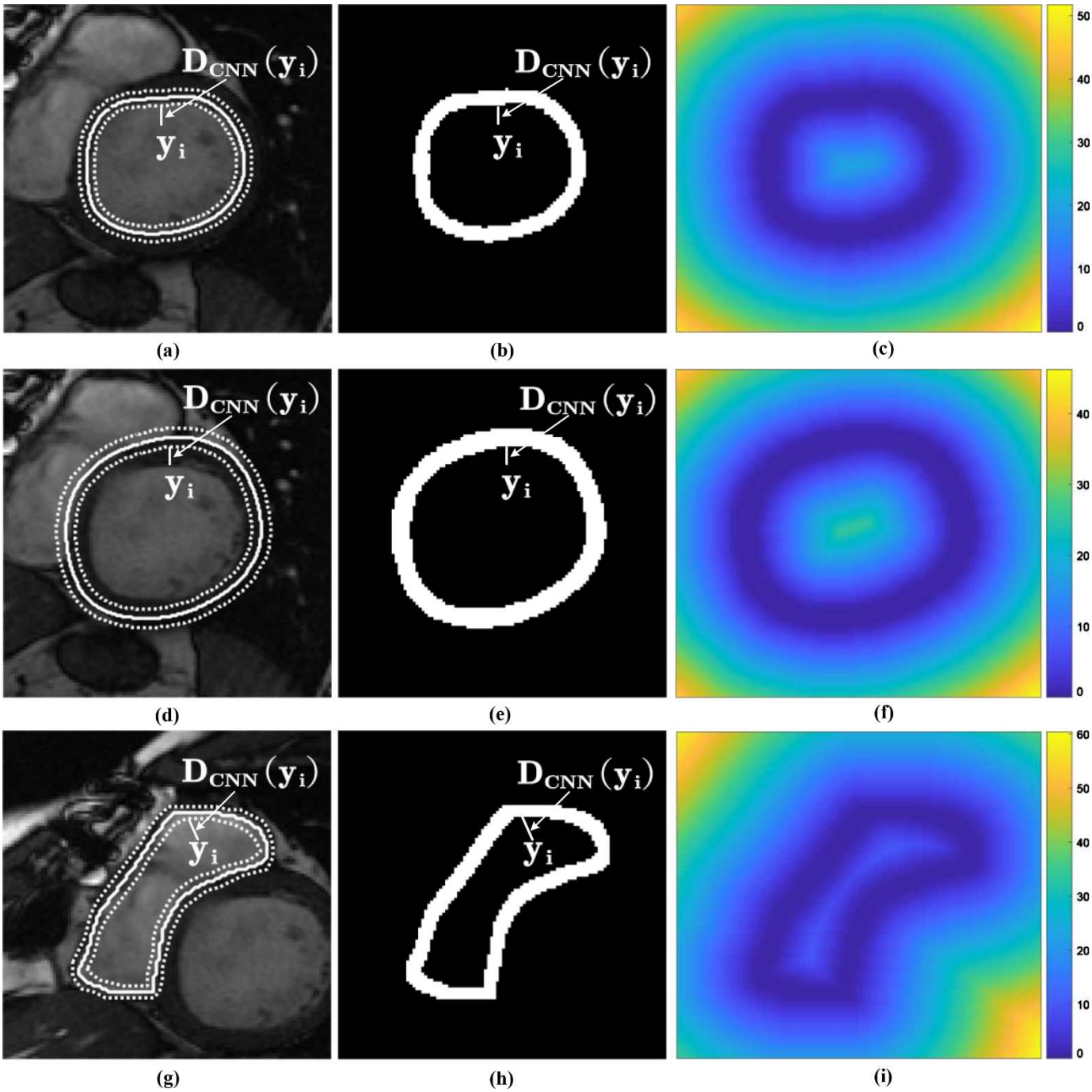


Fig 8. Distance maps for LV and RV contours. (a) Original cardiac MR is depicting endocardial contour, (b) Binary image of endocardial contour, (c) Distance map for endocardial contour, (d) Original cardiac MR is depicting epicardial contour, (e) Binary image of epicardial contour, (f) Distance map of the epicardial contour, (g) Original cardiac MR is depicting RV contour, (h) Binary image of RV contour, (i) Distance map of the RV contour. (Cropped for better view).

After the CNN operation, the contours of LV and RV can be obtained separately. Using the same techniques in Step II, the initial shape for bi-ventricles can be easily obtained. Then the LRV contours are

adopted to build distance maps for LRV segmentation. Fig. 8 shows three distance maps constructed from the narrow band for RV endocardial, LV endocardial, and epicardial contours derived from CNN. For a point y_i, the distant value $D_{CNN}(y_i)$ is defined as follows:

$$D_{CNN}(y_i) = d_{min}(y_i, \partial CNN) \tag{10}$$

where $d_{min}(y_i, \partial CNN)$ is the minimal distance from point y_i to the narrow band ∂CNN.

In the 3D-ASM matching, IIM uses Eq. (7) to select candidate points. We optimise the process by employing the generated CNN-based distance maps. To minimise the value of the distance map and the Mahalanobis distance between the sampled intensity profile and the mean intensity profile [26], the optimal candidate points can be derived via the following equation:

$$y_i^o = \underset{y_i}{\operatorname{argmin}} \left((g(y_i) - \overline{g}_i)^T S_{gi}^{-1} (g(y_i) - \overline{g}_i) + \eta * D_{CNN}(y_i) \right) \tag{11}$$

where $\eta = 3$ is a penalty item used for candidate points searching region. $D_{CNN}(y_i)$ denotes the value of the CNN distance map for the candidate point y_i. The matching algorithm is illustrated in Algorithm 1.

Algorithm 1: Matching Algorithm: SPASM-CNN

```
Input: InitialShape, ImageStack := Short axis images,
       DmapEndoStack := Distance maps for endocardial contours,
       DmapEpiStack := Distance maps for epicardial contours,
       DmapRVEndoStack := Distance maps for RV endocardial contours,
       MeanProfiles := trained intensity profiles per landmark
Output: BestFit
for Iteration < MaxIteration do
    if iteration:=1 then
        CurrentShape ← InitialShape
    else
        CurrentShape ← InitialShape
    end
    Function Intersect(ImageStack, CurrentShape)
        foreach imageplane ∈ ImageStack do
            2Dcontour ← intersection with CurrentShape
        end
        for All points in contourStack do
            interPoints ← find closest landmarks in CurrentShape
        endfor
        return interPoints
    End Function Function
     FindCandidates(interPoints, meanProfiles, DmapEndoStack, DmapEpiStack)
        foreach j ∈ interPoints do
            sampledProfiles ← sample perpendicular profiles
            if j ∈ Endocardialcontour then
                sampleddis ← sample distances from DmapEndoStack
            else if j ∈ Epicardialcontour then
                sampleddis ← sample distances from DmapEpiStack
            else if j ∈ RVEndocardialcontour then
                sampleddis ← sample distances from DmapRVEndoStack
            end
        end
        for Possible profle positions in search range do
            Mdis ← Mahalanobis(meanProfiles,sampledProfiles)
            NewMdis ← Mdis + η * sampleddis
            MinMdis ← minimal NewMdis
        endfor
        CanditatePoints ← store MinMdis candidate positions
        return CanditatePoints
    End Function
    Function Propagate(CanditatePoints)
        foreach p ∈ CanditatePoints do
            Forces ← calculate propagation to neighbouring nodes of p
        end
        DefShape ← apply Forces to CurrentShape
        BestFit ← best parameters from Eq.(6) to fit DefShape
        return BestFit
    End Function
endfor
```

3. Dataset

In this paper, the dataset, which consists of 1000 cardiac MRI cases from UK Biobank [27], is used to test our method's performance. Ground truth for left and RV contours delineated by experts are available for

CNN training and quantitative analysis of cardiac functions. Cardiac magnetic resonance (CMR) images (end-diastolic short-axis view) data from UK Biobank (UKB) was accessed under access application #11350 and used to train and validate the proposed method.

CNN parameters are learned from CMR images in SA view of 500 subjects, and the remaining 500 cases are used for validation. The CNN is trained to distinguish between background, LV blood-pool, myocardium, and RV blood-pool.

4. Results

In this section, some experiment results demonstrate that our proposed algorithm can get accurate and robust segmentation for LV and RV.

To show the accuracy of our method, comparisons are carried out between the proposed algorithm and other approaches. The overlap (Dice) and Jaccard similarity (Jac) indexes evaluate the overlap between the automated produced segmentation A and ground truth M. They are defined as below:

$$Dice = \frac{2*A \cap M}{A + M} \tag{12}$$

$$\mathrm{Jac} = \frac{A \cap M}{A \cup M} \tag{13}$$

Dice and Jac are between 0 and 1, and the higher values imply better agreement between the two segmentations.

We compare the results among three CNNs, Bai [14], U-net [28], and our adopted FCN. The overlap and Jaccard indexes can be seen in Table 1. In our adopted FCN, overlap and Jaccard index are 0.93 and 0.87 for LRV in ED (end-diastole) phase, respectively, they are 0.90 and 0.83 in ES (end-systole) phase. The FCN can get more accurate results than other CNNs.

Table 1. Overlap and Jaccard indexes for the 3 CNNs.

		ED			ES		
	Method	Bai [14]	U-Net [29]	Adopted FCN	Bai [14]	U-Net [29]	Adopted FCN
LVEndo	Dice	0.92±0.05	0.91±0.05	**0.93±0.05**	0.88±0.07	0.88±0.06	**0.89±0.07**
	Jac	0.88±0.06	0.85±0.05	**0.89±0.06**	0.81±0.07	0.81±0.06	**0.83±0.07**
LVEpi	Dice	0.94±0.03	0.94±0.03	**0.95±0.03**	0.93±0.04	0.93±0.03	**0.94±0.04**
	Jac	0.90±0.04	0.89±0.04	**0.90±0.04**	0.88±0.05	0.88±0.03	**0.89±0.05**
RVEndo	Dice	0.88±0.06	0.88±0.04	**0.90±0.06**	0.84±0.07	0.84±0.06	**0.86±0.07**
	Jac	0.81±0.07	0.80±0.05	**0.83±0.07**	0.75±0.08	0.74±0.07	**0.78±0.08**
LRV	Dice	0.91±0.05	0.91±0.04	**0.93±0.05**	0.88±0.06	0.88±0.05	**0.90±0.06**
	Jac	0.86±0.06	0.85±0.05	**0.87±0.06**	0.81±0.07	0.81±0.05	**0.83±0.07**

In Table 2, it can be seen in the ED phase that the overlap indexes from ours for LVEndo, LVEpi are 4.8% and 3.5% larger than those from Albà et al [22]. The Jaccard indexes from ours for LVEndo, LVEpi are 8.1% and 6.5% larger than those from Albà et al. Compared with Albà et al in ES phase, the corresponding values of overlap indexes for LVEndo, LVEpi from ours are 15.7% and 12.2% larger, they are 22.8% and 17.7% in the Jaccard indexes. A conclusion can be obtained that our schema can derive more accurate results than those from 3D-ASM adopted by Albà et al [22].

Table 2. Overlap and Jaccard indexes for the clinical cases.

		ED		ES	
	Method	**Proposed**	Albà's (2018)	Proposed	Albà's (2018)
LVEndo	Dice	**0.88±0.04**	0.84±0.06	**0.81±0.04**	0.70±0.11
	Jac	**0.80±0.05**	0.74±0.07	**0.70±0.05**	0.57±0.13
LVEpi	Dice	**0.89±0.03**	0.86±0.05	**0.83±0.05**	0.74±0.10
	Jac	**0.82±0.04**	0.77±0.07	**0.73±0.06**	0.62±0.12
RVEndo	Dice	**0.77±0.06**	Null	**0.69±0.06**	Null
	Jac	**0.67±0.07**	Null	**0.57±0.07**	Null
LRV	Dice	**0.85±0.04**	Null	**0.78±0.05**	Null
	Jac	**0.76±0.05**	Null	**0.67±0.06**	Null

The Point-to-surface errors measures the mean distance from automatic points P_A and manual points P_M to surfaces of ground truth S_M and automatic S_A. Point-to-surface errors (P2S) is defined:

$$P2S = \frac{1}{2|P_A|}\sum_{p\in P_A} d(p,S_M) + \frac{1}{2|P_M|}\sum_{p\in P_M} d(p,S_A) \tag{14}$$

where $d(p,S)$ denotes the minimal distance from point p to surface S. The smaller the distance metric, the better the match.

For Point-to-surface errors, the results can be seen in Figure 9. In Figure 9, Point-to-surface errors for LVEndo, LVEpi, and myocardial are calculated for our algorithm and 3D-ASM. It can be seen that the values from ours are smaller than the corresponding indexes from 3D-ASM. Considering that the image contrast in ES phase is weaker than in the ED phase, the Point-to-surface errors are higher in the ES phase than those in the ED phase for both the proposed and 3D-ASM results.

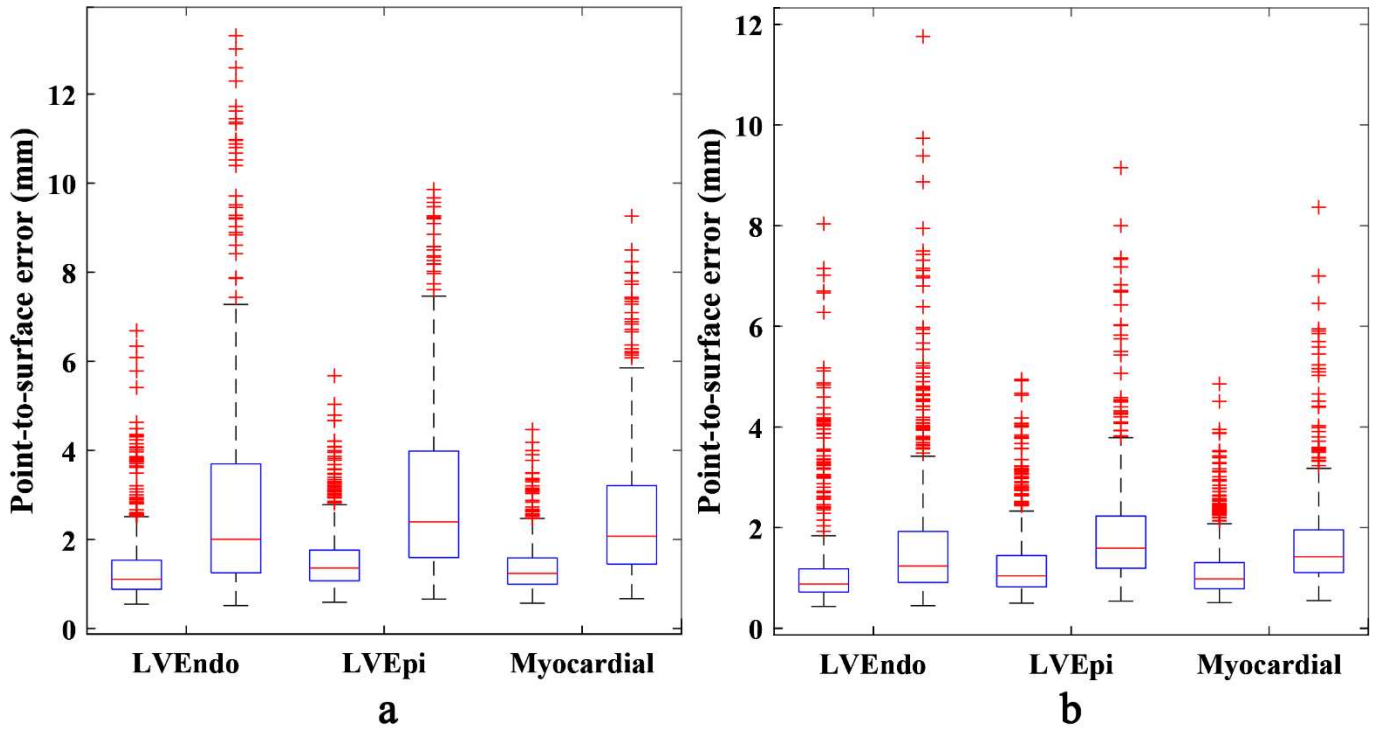


Fig 9. Boxplot of Point-to-surface errors for proposed method and 3D-ASM. (a) ED phase, (b) ES phase. (For each group: LVEndo, LVEpi and Myocardial, two subfigures are displayed, the left is from the proposed algorithm, and the right shows results from 3D-ASM adopted in Albà's (2018).

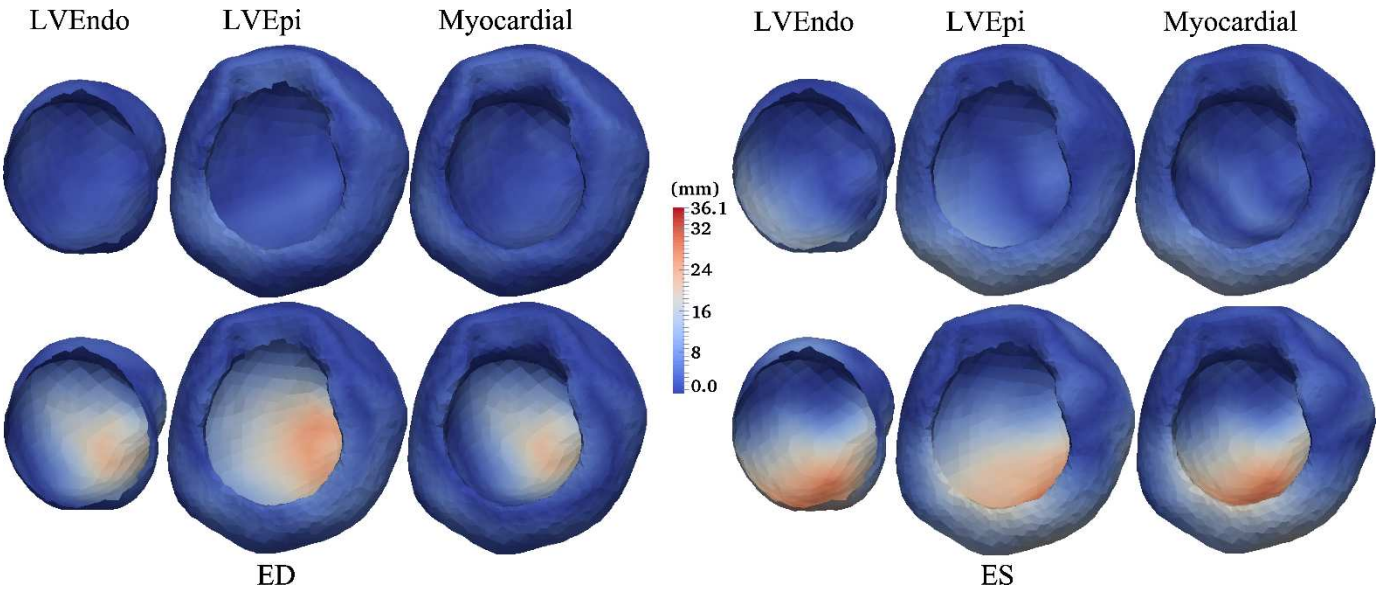
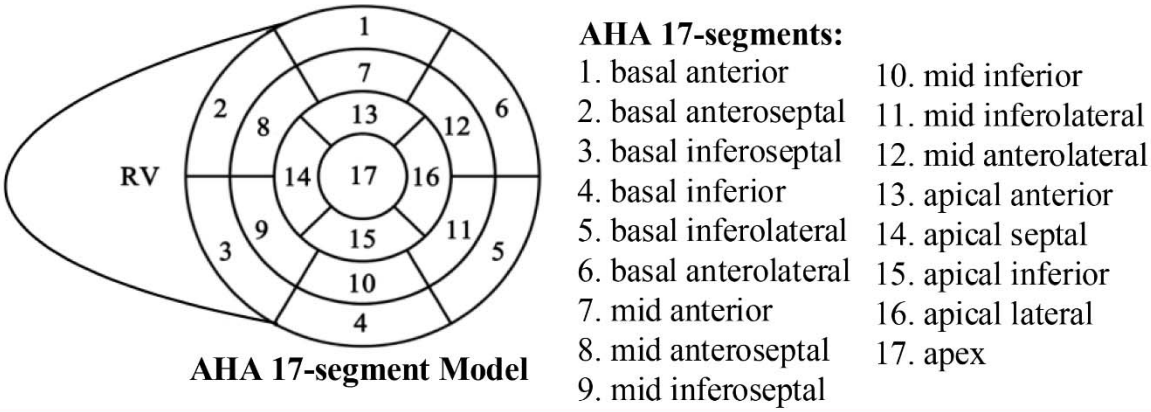


Fig 10. Point-to-surface errors for the proposed method and Albà's in ED and ES phases.

Figure 10 shows Point-to-surface errors for one case between our method and 3D-ASM. Since our algorithm combines the advantages of CNN and 3D-ASM adopted in [22], our results match better than those of only using 3D-ASM.

Figure 11 shows the mean and standard deviation values of Point-to-surface errors for the regional analysis between the proposed method and 3D-ASM adopted by Albà in a bulls-eye display of the AHA 17-segment model [30]. Compared with 3D-ASM, we observe that ours are closer to the ground truth in most regions in terms of the mean and standard deviation values, which confirmed the high quality of our proposed algorithm.



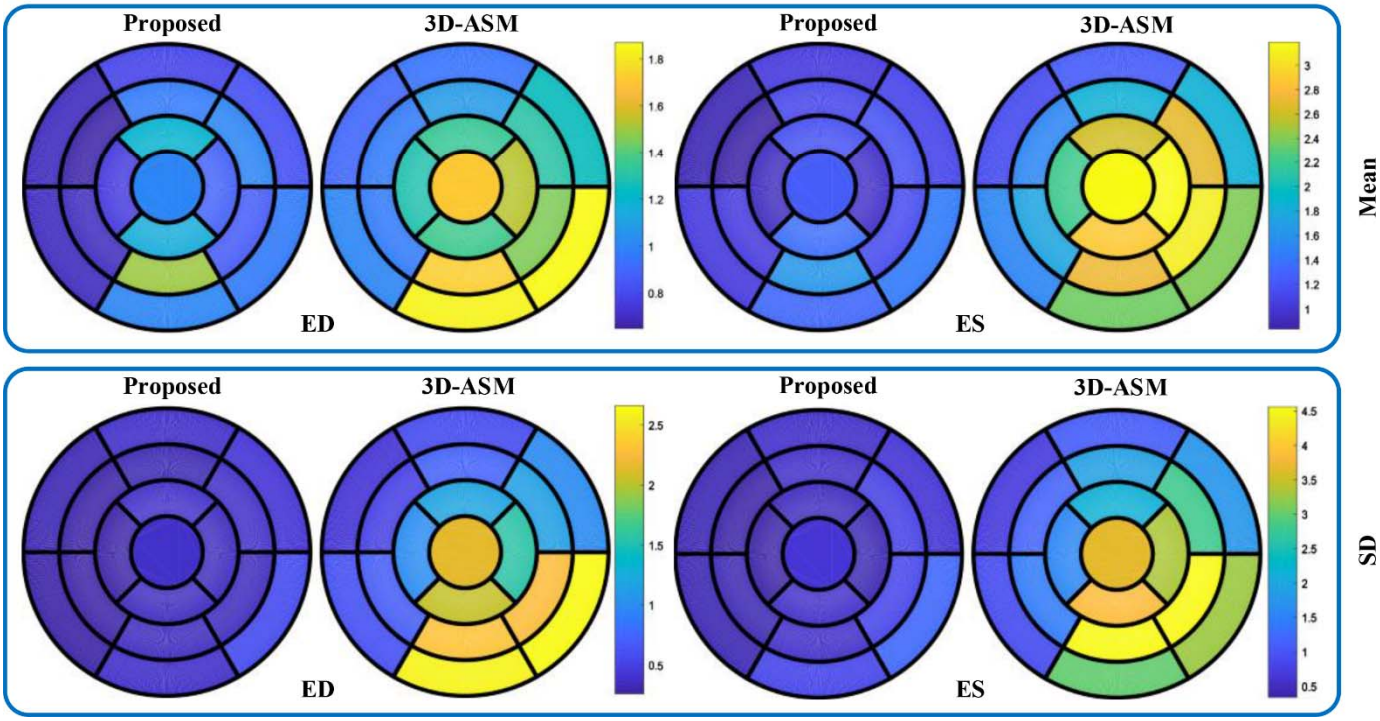


Fig 11. Point-to-surface errors for the proposed method and 3D-ASM by Albà's (2018) presented as bulls-eye displays. (SD: standard deviation).

We provide a visual comparison between the manual segmentation and the automatic approaches. Figure 12 displays good and bad segmentations of short-axis slices for two subjects.

Figure 12 shows segmentation of two cases among manual, the proposed method and 3D-ASM adopted by Albà's (2018) in ED and ES phases. Segmentation results from automatic approaches and clinical experts are in different styles. The contours from experts are red curves, whereas the outputs from automatic are green ones. The first and the third rows are results from our proposed algorithm, while the second and the last rows are from 3D-ASM. It can be seen that our method obtains better segmentation performance than Albà (2018) which used the 3D-ASM strategy. Since only the LV initial shapes can be obtained in Albà's

work, no RV segmentation can be seen in Albà's method.

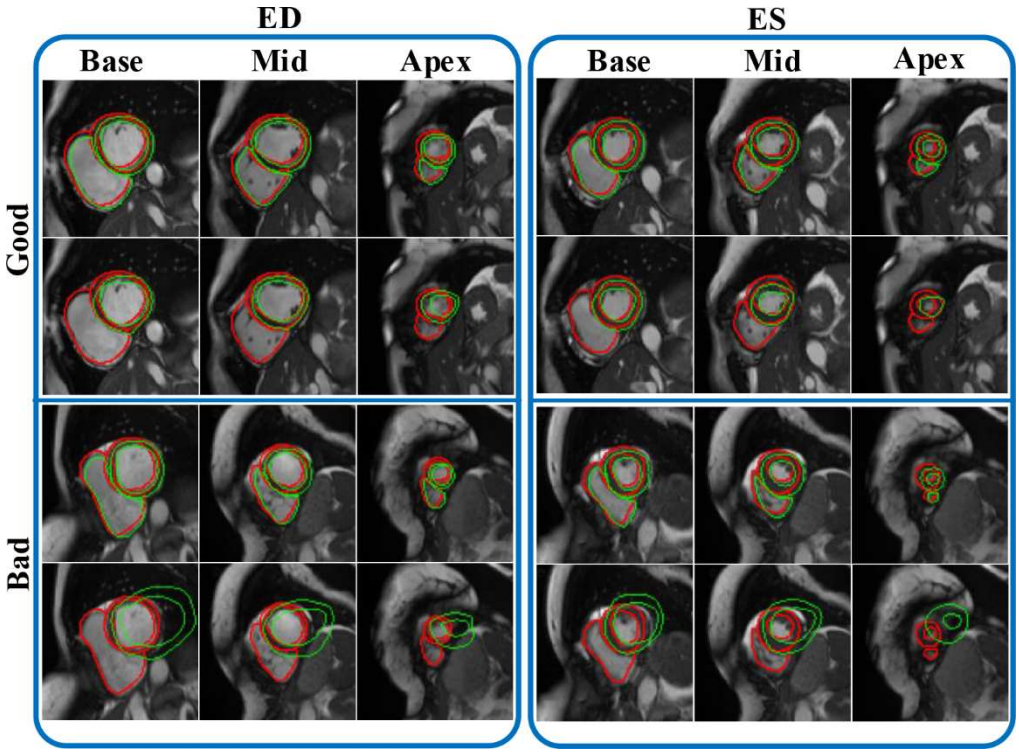


Fig 12. Short-axis slice segmentation of good and bad cases (Good case: good segmentation by ours and Albà's (2018); Bad case: ,wrong segmentation by Albà's (2018)) in ED and ES phases. For each case, the first row comes from the proposed algorithm results, while the second row is from Albà's (2018). The green curves indicate automatic contours, while the red ones are the ground truth (cropped for better viewing).

Table 3 shows cardiac functional parameters for both LV and RV between the manual and proposed methods. The results from ours are close to those from experts.

Table 3. Cardiac functional parameters.

	From experts	**Proposed**
LVEDV (ml)	143.92±33.77	**135.34**±31.91
LVESV (ml)	58.15±18.37	**54.60±17.02**
LVSV (ml)	85.77±19.94	**80.74±23.79**
LVM (g)	94.79±26.28	**100.86±25.57**
LVEF (%)	59.60±6.52	**59.66**±10.87
RVEDV (ml)	148.75±35.07	**139.94±36.73**
RVESV (ml)	72.98±24.62	**66.99±22.26**
RVSV (ml)	75.77±18.19	**72.95±20.04**
RVEF (%)	50.93±6.85	**52.13±10.89**

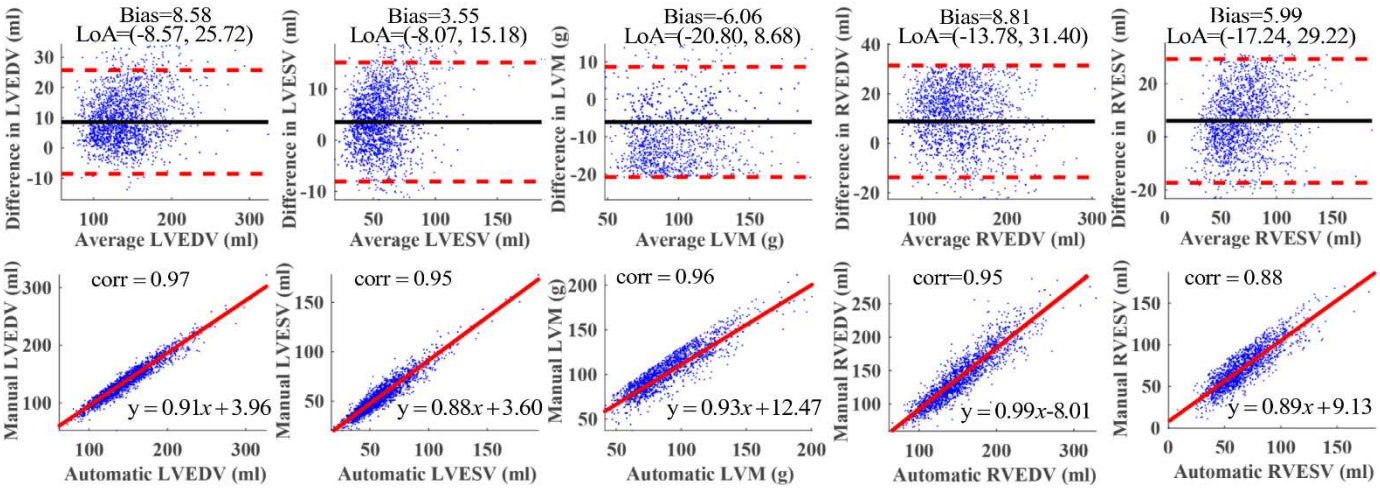


Fig 13 Plots of Bland-Altman and correlation of cardiac functional indexes between the manual and automated results. In the top row, the black horizontal lines reprent mean difference (i.e. bias), while the two red dashed lines are limits of agreement (LoA, i.e. ±1.96 standard deviations from the mean). The second row denotes correlation plots for the coresponding cardiac functional indexes coming from manual and proposed method.

The Bland-Altman and correlation plots are shown in the top and bottom rows in Figure 13 respectively. The Bland-Altman plot is an effective tool to measure agreement and bias between two techniques. It can be seen that strong agreement in the Bland-Altman plots, indicating that slight bias exists in clinical indexes computed by the proposed approach. The correlation coefficient (corr) is the specific measure that quantifies the strength of the linear relationship between two variables in a correlation analysis. Correlations of cardiac indexes range from 0.88 to 0.97, demonstrating a strong relationship between manual and automatic methods.

5. Discussion

We proposed a fully automatic segmentation for left and right ventricles in cardiac MRI. Our approach, which combines a bi-ventricular model initialisation, deep learning neural network, and a 3D-ASM segmentation, obtains outstanding performance. Three landmarks are automatically derived to guide the LV shape initialisation upon which the bi-ventricular model is initialised. Procrustes analysis is employed to get the coarse bi-ventricle estimate. However, the coarse initial LRV shape is doomed to get failure segmentation

due to irregular RV shape.

The image planeis rotated to be perpendicular to Z-axis. he coarse initial LRV shape and the RV points from CNN are rotated with the same transformation. Then the angle is calculated between RV points from coarse initial LRV shape and CNN points using thepoint-set registration method. At last, the coarse initial shape is refined by rotating with the obtained angle and transformed to its original position.

In general, 3D ground truth shape cannot be obtained by manual delineation. To build a 3D ground shape, 2D manual contours are used to construct distance maps for 3D-ASM segmentation. The segmented shapes and the corresponding images are employed to rebuild models of point distribution and image intensity.

In the segmentation process, there are still challenges for 3D-ASM. Firstly, there exists similar shapes and edge information between the LV endo- and epicardial contours. Secondly, other organs with robust edge information may divert model fitting owing to limitations in the IIM. Figure 12 shows poor results for 3D-ASM due to strong edges from other organs.

To overcome these difficulties, we invent distance maps by utilising CNN segmentation results. The distance maps are converted to a penalty item in the IIM to drive the refined initial shapes to the LRV position. Consequently, a 3D LRV shape is created, representing an accurate segmentation for the cardiac images.

6. Conclusion

This study introduces a hybrid schema that can automatically initialise bi-ventricle for 3D-ASM. 2D manual contours are employed to build distance maps to get 3D ground truth shapes. In the segmentation process, deep learning is used to refine the bi-ventricular initial shapes and build distance maps for the IIM. Results show that our algorithm can cope with technical difficulties and derive robust segmentation of left and right ventricles for cardiac MRI studies with subvoxel accuracy. Our approach still ssome aspects of being improved, such as the deep learning schema that can be optimised and how to utilise time constraint to enhance segmentation in different phases using 3D-ASM. We will test the improved method on the

complete set of UK Biobank CMR imaging study (>40k subjects) in the future.

Acknowledgement

This research has been conducted using the UK Biobank Resource under Applications 11350. The authors are grateful to all UK Biobank participants and staff. AFF acknowledges support from the Royal Academy of Engineering Chair in Emerging Technologies Scheme (INSILEX, CiET1819/19), Royal Society International Exchanges Programme (CROSSLINK, IEC\NSFC\201380) and EPSRC-funded Grow MedTech (CardioX, POC041) and Pengcheng Visiting Scholars Programme from the Shenzhen Government.

The National Natural Science Foundation of China supported this study (No. 62076257, 61773409 and 61976227), Fundamental Research Funds for the Central Universities (No. CZY20039), and Applied Basic Research Programme of Wuhan (No. 2020020601012239).

The authors thank Dr M Pereanez, Dr R Attar, and M Hoz from Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB) at the University of Leeds throughout this work.

REFERENCE

[1] M.K. Ali, M.K. Bullard, S. Saydah, G. Imperatore, E.W. Gregg, Cardiovascular and renal burdens of prediabetes in the USA: analysis of data from serial cross-sectional surveys, 1988–2014, Lancet Diabetes Endocrinol, 6 (2018) 392-403

[2] P. Radau, Y. Lu, K. Connelly, G. Paul, A. Dick, G. Wright, Evaluation framework for algorithms segmenting short axis cardiac MRI, The MIDAS Journal-Cardiac MR Left Ventricle Segmentation Challenge, 49 (2009)

[3] A. Suinesiaputra, B.R. Cowan, A.O. Al-Agamy, M.A. Elattar, N. Ayache, A.S. Fahmy, A.M. Khalifa, P. Medrano-Gracia, M.-P. Jolly, A.H. Kadish, A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images, Medical image analysis, 18 (2014) 50-62

[4] C. Petitjean, M.A. Zuluaga, W. Bai, J.-N. Dacher, D. Grosgeorge, J. Caudron, S. Ruan, I.B. Ayed, M.J. Cardoso, H.-C. Chen, Right ventricle segmentation from cardiac MRI: a collation study, Medical image analysis, 19 (2015) 187-202

[5] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P.-A. Heng, I. Cetin, K. Lekadir, O. Camara, M.A.G. Ballester, Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved?, IEEE transactions on medical imaging, 37 (2018) 2514-2525

[6] P. Peng, K. Lekadir, A. Gooya, L. Shao, S.E. Petersen, A.F. Frangi, A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging, Magnetic Resonance Materials in Physics, Biology and Medicine, 29 (2016) 155-195

[7] C. Petitjean, J.-N. Dacher, A review of segmentation methods in short axis cardiac MR images, Medical image analysis, 15 (2011) 169-184

[8] T. Heimann, H.-P. Meinzer, Statistical shape models for 3D medical image segmentation: a review, Medical image analysis, 13 (2009) 543-563

[9] R. Attar, M. Pereañez, A. Gooya, X. Albà, L. Zhang, M.H. de Vila, A.M. Lee, N. Aung, E. Lukaschuk, M.M. Sanghvi, Quantitative CMR population imaging on 20,000 subjects of the UK Biobank imaging study: LV/RV quantification pipeline and its evaluation, Medical image analysis, 56 (2019) 26-42

[10] M. Bernier, P.-M. Jodoin, O. Humbert, A. Lalande, Graph cut-based method for segmenting the left ventricle from MRI or echocardiographic images, Computerized Medical Imaging and Graphics, 58 (2017) 1-12

[11] S. Bouzidi, A. Emilien, J. Benois-Pineau, B. Quesson, C. Ben Amar, P. Desbarats, Segmentation of left ventricle on dynamic MRI sequences for blood flow cancellation in Thermotherapy, Signal Processing Image Communication, 59 (2017) 37-49

[12] H. Huaifei, N. Pan, J. Wang, T. Yin, R. Ye, Automatic Segmentation of Left Ventricle from Cardiac MRI via Deep Learning and Region Constrained Dynamic Programming, Neurocomputing, 347 (2019) 139-148

[13] R. Zhou, Z. Liao, T. Pan, S.A. Milgrom, C.C. Pinnix, A. Shi, L. Tang, J. Yang, Y. Liu, D. Gomez, Cardiac atlas development and validation for automatic segmentation of cardiac substructures, Radiotherapy and Oncology, 122 (2017) 66-71

[14] W. Bai, M. Sinclair, G. Tarroni, O. Oktay, M. Rajchl, G. Vaillant, A.M. Lee, N. Aung, E. Lukaschuk, M.M. Sanghvi, Automated cardiovascular magnetic resonance image analysis with fully convolutional networks, Journal of Cardiovascular Magnetic Resonance, 20 (2018) 65

[15] Q. Zheng, H. Delingette, N. Duchateau, N. Ayache, 3-D consistent and robust segmentation of cardiac images by deep learning with spatial propagation, IEEE transactions on medical imaging, 37 (2018) 2137-2148

[16] W. Zeng, J. Peng, S. Wang, Q. Liu, A comparative study of CNN-based super-resolution methods in MRI reconstruction and its beyond, Signal Processing Image Communication, 81 (2019).https://doi.org/10.1016/j.image.2019.115701

[17] Y. Guo, X. Jia, B. Zhao, H. Chai, Y. Huang, Multifeature extracting CNN with concatenation for image denoising, Signal Processing-image Communication, 81 (2020) 115690.https://doi.org/10.1016/j.image.2019.115690

[18] F.M. Sukno, S. Ordas, C. Butakoff, S. Cruz, A.F. Frangi, Active shape models with invariant optimal features: Application to facial analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, 29 (2007) 1105-1117

[19] H.C. Van Assen, M.G. Danilouchkine, A.F. Frangi, S. Ordas, J.J.M. Westenberg, J.H.C. Reiber, B.P.F. Lelieveldt, SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data, Medical Image Analysis, 10 (2006) 286-303

[20] H. Radu, F. Florence, Y. Manuel, D. Guillaume, Z. Jian, Rigid and articulated point registration with expectation conditional maximization, IEEE transactions on pattern analysis and Machine Intelligence, 33 (2011) 587-602

[21] D.G. Kendall, A survey of the statistical theory of shape, Statistical Science, (1989) 87-99

[22] X. Alba, K. Lekadir, M. Pereanez, P. Medranogracia, A.A. Young, A.F. Frangi, Automatic initialization and quality control of large-scale cardiac MRI segmentations, Medical Image Analysis, 43 (2018) 129-141

[23] A. Myronenko, X.B. Song, Point Set Registration: Coherent Point Drift, IEEE Transactions on Pattern Analysis and Machine Intelligence, 32 (2010) 2262-2275

[24] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active Shape Models-Their Training and Application, Computer Vision and Image Understanding, 61 (1995) 38-59

[25] H. Hu, N. Pan, T. Yin, H. Liu, B. Du, Hybrid Method for Automatic Construction of 3D-ASM Image Intensity Models for Left Ventricle, Neurocomputing, 396 (2019) 65-75

[26] V.B.B. Ginneken, A.F. Frangi, J. Staal, T.B.B.H. Romeny, M.A. Viergever, Active shape model segmentation with optimal features, IEEE Transactions on Medical Imaging, 21 (2002) 924-933

[27] S.E. Petersen, P.M. Matthews, F. Bamberg, D.A. Bluemke, J.M. Francis, M.G. Friedrich, P. Leeson, E. Nagel, S. Plein, F.E. Rademakers, Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank - rationale, challenges and approaches, Journal of Cardiovascular Magnetic Resonance, 15 (2013) 46

[28] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, inInternational Conference on Medical Image Computing and Computer-Assisted Intervention, 2015, pp. 234-241.

[29] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, inMed. Image Comput. Comput.-Assist. Intervent., 2015, pp. 234-241.

[30] X. Zhuang, Challenges and methodologies of fully automatic whole heart segmentation: a review, Journal of healthcare engineering, 4 (2013) 371-407

Figure Captions:

Fig 1. The pipeline of the proposed algorithm.

Fig 2. Definition of AOTIA, MITRAL and LVAPEX.

Fig 3. Bi-ventricular model initialisation. (a) LV initial shape and three points (AOTIA, MITRAL and LVAPEX), (b) Bi-ventricular initial shape using Procrustes analysis, (c) LV initial shape in short-axis view, (d) Bi-ventricular initial shape in short-axis view.

Fig 4. Initial shape rotation. (a) LV, RV points from PDM and RV points from ground truth at their original position, (b) Rotated points from PDM and ground truth, (c) RV point sets from PDM and ground truth, (d) RV point sets with similar shape from PDM and ground truth. GT: ground truth.

Fig 5. Complex transformation for initial shape rotation. (a) Registered points from PDM and RV contour points from ground truth, (b) LV and RV points using complex transformation, RV contour points from ground truth, (c) LV, RV points from PDM and RV points from ground truth at their original position, (d) Bi-ventricular initial shape in short-axis view. GT: ground truth.

Fig 6. Distance maps for ground truth of LV and RV. (a) Original cardiac MR is depicting endocardial contour, (b) Binary image of endocardial contour, (c) Distance map for endocardial contour, (d) Original cardiac MR is depicting epicardial contour, (e) Binary image of epicardial contour, (f) Distance map of the epicardial contour, (g) Original cardiac MR is depicting RV contour, (h) Binary image of RV contour, (i) Distance map of the RV contour, (Cropped for better view).

Fig 7. The architecture of the fully neural network.

Fig 8. Distance maps for LV and RV contours. (a) Original cardiac MR is depicting endocardial contour, (b) Binary image of endocardial contour, (c) Distance map for endocardial contour, (d) Original cardiac MR is depicting epicardial contour, (e) Binary image of epicardial contour, (f) Distance map of the epicardial contour, (g) Original cardiac MR is depicting RV contour, (h) Binary image of RV contour, (i) Distance map of the RV contour. (Cropped for better view).

Fig 9. Boxplot of Point-to-surface errors for proposed method and 3D-ASM. (a) ED phase, (b) ES phase. (For each group: LVEndo, LVEpi and Myocardial, two subfigures are displayed, the left is from the proposed algorithm, and the right shows results from 3D-ASM adopted in Albà's (2018).

Fig 10. Point-to-surface errors for the proposed method and Albà's in ED and ES phases.

Fig 11. Point-to-surface errors for the proposed method and 3D-ASM by Albà's (2018) presented as bulls-eye displays. (SD: standard deviation).

Fig 12. Short-axis slice segmentation of good and bad cases (Good case: good segmentation by ours and Albà's (2018); Bad case: ,wrong segmentation by Albà's (2018)) in ED and ES phases. For each case, the first row comes from the proposed algorithm results, while the second row is from Albà's (2018). The green curves indicate automatic contours, while the red ones are the ground truth (cropped for better viewing).

Fig 13 Plots of Bland-Altman and correlation of cardiac functional indexes between the manual and automated results. In the top row, the black horizontal lines reprent mean difference (i.e. bias), while the two red dashed lines are limits of agreement (LoA, i.e. ±1.96 standard deviations from the mean). The second row denotes correlation plots for the coresponding cardiac functional indexes coming from manual and proposed method.

Table Captions:

Table 1. Overlap and Jaccard indexes for the 3 CNNs.

Table 2. Overlap and Jaccard indexes for the clinical cases.

Table 3. Cardiac functional parameters.