Ventricle Surface Reconstruction from Cardiac MR Slices Using Deep Learning

Hao Xu[1(✉)], Ernesto Zacur[1], Jurgen E. Schneider[2], and Vicente Grau[1]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK
hao.xu@eng.ox.ac.uk

[2] Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK

Abstract. Reconstructing 3D ventricular surfaces from 2D cardiac MR data is challenging due to the sparsity of the input data and the presence of interslice misalignment. It is usually formulated as a 3D mesh fitting problem often incorporating shape priors and smoothness regularization, which might affect accuracy when handling pathological cases. We propose to formulate the 3D reconstruction as a volumetric mapping problem followed by isosurfacing from dense volumetric data. Taking advantage of deep learning algorithms, which learn to predict each voxel label without explicitly defining the shapes, our method is capable of generating anatomically meaningful surfaces with great flexibility. The sparse 3D volumetric input can process contours with any orientations and thus can utilize information from multiple short- and long-axis views. In addition, our method can provide correction of motion artifacts. We have validated our method using a statistical shape model on reconstructing 3D shapes from both spatially consistent and misaligned input data.

Keywords: Mesh reconstruction · Cardiac MRI · Deep learning

1 Introduction

Generating anatomically accurate 3D surface meshes is a key step in a wide range of applications including cardiac function analysis, interventional guidance and diagnosis [1–3]. Personalization of cardiac surfaces in 3D is also the first step required for computational simulations of cardiac electromechanics using the finite element method [4–6]. Cardiac MR (CMR) imaging provides accurate shape information of the heart non-invasively [2]. A standard clinical CMR study includes a stack of short-axis (SAX) slices, covering at least from the left/right ventricular (LV/RV) apex to the atrioventricular plane (base), plus at least two long-axis (LAX) views: horizontal long-axis (HLA, also known as 4 chamber view or 4CH) and vertical long-axis (VLA, also known as 2 chamber view or 2CH) [7]. Traditional isosurfacing algorithms cannot be directly used due to the sparsity

Y. Coudière et al. (Eds.): FIMH 2019, LNCS 11504, pp. 342–351, 2019.
https://doi.org/10.1007/978-3-030-21949-9_37

of the input data and because of the presence of motion artifacts (misalignment between slices caused by multiple breath holding and possible body movement during acquisition) [8], which make the task of reconstructing 3D structure from CMR data particularly challenging.

Reconstruction of 3D surfaces from CMR data is normally formulated as a 3D mesh adaptation problem to sparse contours or points [9–12], and solutions often incorporate shape priors during that process. The form of such prior could be a regular shape [10,11,13] or a statistical template with plausible variations [14], and deviations from regular and smooth geometries are penalized in a fitting process. These methods usually include an explicit smoothing term during fitting to further regularize the shape in addition to the use of shape priors, and both types of bias may compromise the accuracy of the reconstructed surfaces when handling pathological hearts [15,16]. In order to generate anatomically meaningful surfaces while also preserving the wide variability of the shape, reconstruction methods should incorporate a wide collection of plausible shape priors to refer to when fitting the input data. Deep learning algorithms have shown their capacity for storing a large number of accurate mappings between pair-wise data [17], and no explicit smoothing term is required by these methods, allowing the appearance of sharp edges and corners during reconstruction if necessary. However, for Convolutional Neural Networks (CNNs), input data are required to be highly structured (usually on a regular grid in 2D or 3D), and therefore most state-of-the-art deep learning methods for generating LV and RV 3D meshes using CMR data only take the SAX stack as input with the assumption that all SAX images are parallel [18,19]. This implementation limits the possibility of utilizing the LAX slices, which has been shown to improve the reconstruction of the ventricular meshes [12], as well as the possibility of incorporating out-of-plane geometric transformations for better correction of motion artifacts due to respiration [20,21].

In this paper, we propose to consider the sparse 3D information from contours in a volumetric form, and therefore tackle the problem similarly to volumetric image inpainting [22]. We transform the problem of mesh fitting from sparse input into a 3D volumetric mapping problem followed by isosurfacing from dense volumetric data. The sparse volumetric input can take contours in any positions within the volume, and therefore our method incorporates both SAX slices and LAX slices, and allows explicit out-of-plane motion artifact correction. Our method also has the capacity of reconstructing 3D meshes from misaligned input contours by itself.

1.1 Main Contributions

We have developed a novel bi-ventricular mesh reconstruction method through 3D volumetric mapping with a deep learning algorithm combining both SAX and LAX slices. To our best knowledge, this is the first deep learning method processing multiple views simultaneously without any assumption of image plane orientations.

We propose a method that can be used to reconstruct 3D meshes from both spatially consistent and misaligned input data, and our method is capable of processing intersecting slices with discrepancies among them, while most of the state-of-the-art methods rely on parallel contours.

2 Materials and Methods

Our 3D reconstruction method consists of three steps: 1. generating sparse volumetric input data from contours; 2. generating dense 3D volumetric predictions of LV myocardium and LV/RV cavities from input data using a variation of the 3D U-Net [23]; 3. generating 3D meshes from the predictions with an isosurfacing algorithm. The method is developed and evaluated by using synthetic data generated from a statistical shape model [14,24].

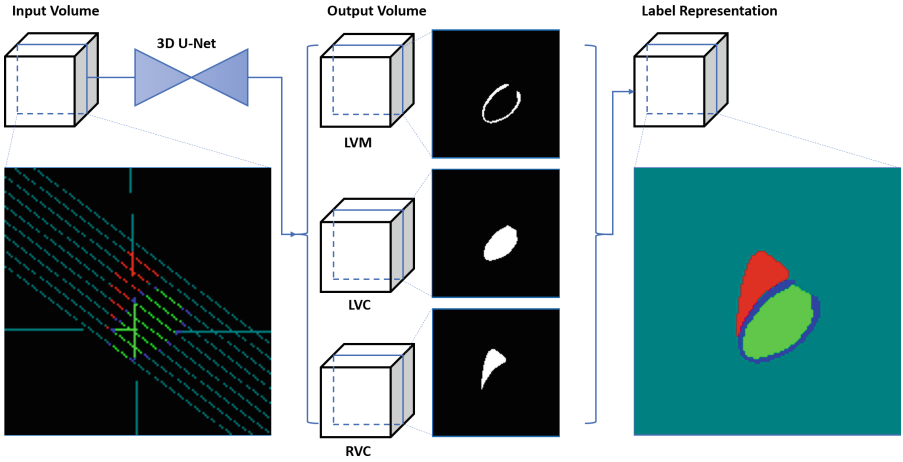

Fig. 1. Representation of the process carried out by the 3D U-Net. For illustration, cross-sectional views corresponding to specific slices (marked with dotted lines on the volumes) are shown; note that these do not correspond to original CMR slices. In the input volume, only voxels corresponding to the centre of CMR slices are assigned specific labels (2:red, 3:green, 4:blue and 1:cyan for RV cavity, LV cavity, LV myocardium and background, respectively), while all others are assigned the label "Unknown" (0). For each of the 3D output volumes, a value between 0 and 1 is assigned to each voxel representing the likelihood of belonging to the respective region. These 3D volumes are then combined to produce the label representation of the reconstructed volume, in which an average of 94.5% of the voxels are transformed from the original Unknown label. (Color figure online)

2.1 Data

We used the statistical shape model published by Bai et al. [14], in which the authors registered 1093 segmented hearts to a template space using rigid registration followed by the application of principal component analysis (PCA) to the surface meshes. The model is formed by labeled images with labels for background, LV myocardium and LV/RV cavities. We downloaded from the publicly available dataset (http://wp.doc.ic.ac.uk/wbai/data/) the mean shape model, the first 100 PCs and the corresponding eigenvalues or variances. We then used these to generate 120 different shapes simulating pathological and controlled cases from clinical data, limiting the variations to six standard deviations for any of the Principal Components.

To generate output references, we placed these shapes into a 3D volume with the size of $128 \times 128 \times 128$ and voxel size of $2 \times 2 \times 2\,\text{mm}^3$, by aligning the centre of the smallest sphere enclosing the corresponding contours and the centre of the 3D volume, and labeled voxels enclosed by surfaces accordingly. To simulate input sparse volumes in conditions similar to real clinical acquisition, we used real image planes from clinical datasets (consisting of both SAX stack and two LAX slices). To generate spatially consistent input datasets, we first aligned one set of image planes to each of the shapes, and the voxels located within 0.5 mm away from the planes were then given their original labels acccording to the shape model. Voxels located more than 0.5 mm away from the plane, or those with inconsistent labels between slices were assigned the "Unknown" label. To mimic the misalignment caused by motion artifacts, we kept fixed image planes and applied 3D rigid transformations to the model (random rotations no larger than 10° and random translations of no more than 4 mm) before the calculation of the labels for each plane. An example of input and output volumetric data with cross-sectional views is shown in Fig. 1.

2.2 Volumetric Mapping

We adapted a variation of 3D U-Net for volumetric mapping. The schematic diagram of the network architecture is shown in Fig. 2.

The network consists of an encoder and a decoder, with skip connections between feature maps with the same resolution. The encoder starts from the input layer with the size of $128 \times 128 \times 128 \times 1$ and has three max-pooling stages where each pooling layer has a window size of (2, 2, 2) and a stride of (2, 2, 2), giving feature maps with four different resolutions. The decoder then upsamples the feature maps back to the original resolution also in three stages using upsampling layers with factors of (2, 2, 2). The heavy duty of the computation within the network is carried out by convolutional blocks shown in Fig. 2, with each of them having two convolutional layers with kernel size of (3, 3, 3) and two batch normalization (BN) layers. Other than the output layer, which has sigmoid activation functions, rectified linear units (ReLU) are used throughout the network. The output layer has 3 convolutional kernels with the size of (1, 1, 1) giving the prediction in the form of $128 \times 128 \times 128 \times 3$ volume representing

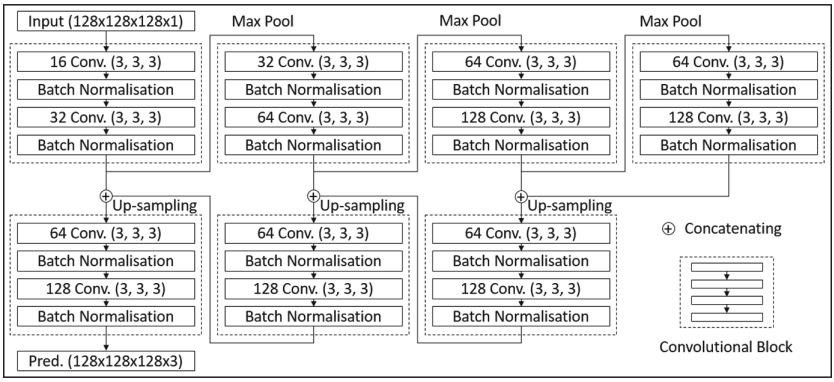

Fig. 2. A schematic diagram of our volumetric mapping network. For each convolutional layer, the number of filters is specified.

LV myocardium, LV cavity and RV cavity. The Dice coefficient loss was used during the training of the network to deal with imbalance between foreground and background voxels [25].

2.3 Isosurfacing

The prediction of the network was split into three $128 \times 128 \times 128$ volumes, and used to form the isosurfaces at 0.5 using marching cubes [26]. For each 3D volume, the largest object was then selected as the expected structure.

3 Experiments and Results

We performed two 4-fold cross-validation experiments with spatially consistent data and misaligned data respectively. For both experiments, the same number of bi-ventricular shapes were generated as described in Sect. 2.1, and they were randomly grouped into 4 sets of training data (80 shapes), validating data (10 shapes) and testing data (30 shapes) also in the same way, with the 4 testing datasets covering all generated shapes.

During the training phase of each cross-validation experiment, the neural network was randomly initiated and the network parameters were updated through back-propagation using the spatially consistent training data. The optimization was early stopped by evaluating the loss function of the validating data to prevent over-fitting. Once the training of the neural network for the spatially consistent cases was terminated, the learned parameters were used to initiate the network for reconstructing misaligned cases and fine-tuned using the misaligned training data. An overview of the training progress is shown in Fig. 3.

The testing data was only used for the evaluation of the method, and the predictions from the network for the testing data were then used to generate 3D meshes as described in Sect. 2.3. In both experiments, Dice coefficient and

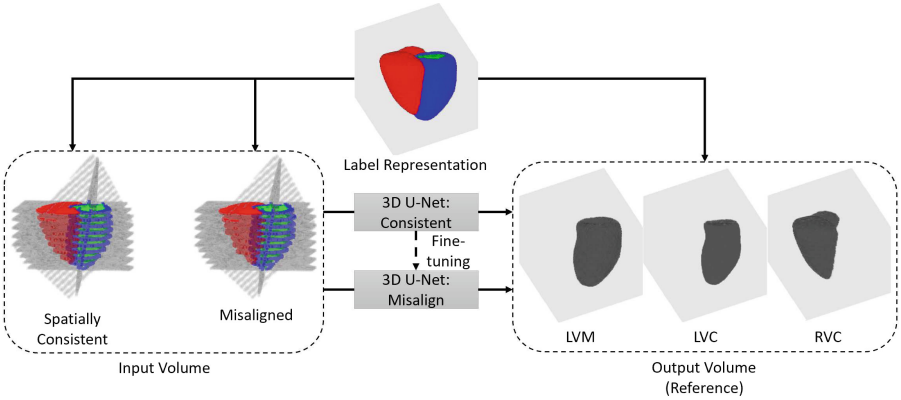

Fig. 3. Training progress overview. The blue, green and red objects respectively correspond to LV myocardium, LV cavity and RV cavity. For each case, both spatially consistent and misaligned input volumes are generated from the reference mesh according to Sect. 2.1. The output volume has voxels with values of 0 (light gray) and 1 (dark gray). (Color figure online)

Hausdorff Distance (HD) were the metrics used for quantitative analysis of the method. For the experiment with misaligned contours we registered the reconstructed shape to the reference shape with a rigid transformation before calculating the metrics. We also calculated the Euler characteristic for the generated meshes to evaluate the topology, and for all cases the value was expected to be equal to 2.

3.1 Statistical Shape Model

The Dice coefficient and HD results for experiments on spatially consistent and misaligned contours are shown in Table 1. For spatially consistent input data, the reconstructed LV/RV cavities shape matched the reference shape accurately with mean Dice coefficient values of 0.98. The LV myocardium has a more complicated shape with larger surface area, and therefore the Dice score is more sensitive to small changes, while the mean value is still 0.94. The voxel size is $2 \times 2 \times 2$ mm^3, and the mean values of HD are around 2 voxels. As the input data has only around 5.5% of voxels with known labels, the results suggest that our 3D U-Net stored a collection of plausible mappings between sparse input data and corresponding ventricular shapes during training and is able to utilize that information during the reconstruction of the testing cases.

The experiments on misaligned input data also achieved good accuracy, with above 0.93 mean Dice coefficient for LV/RV cavities and 0.83 for LV myocardium. As expected, comparing to the experiment on spatially consistent input data, the accuracy of these experiment is lower. This is unavoidable given the additional challenges involved in reconstructing the 3D shapes from misaligned input data, including discrepancies between slices, which results in the same set of input data

Table 1. Dice coefficient and Hausdorff Distance between reconstructed ventricular shape and reference ventricular shape for both spatially consistent and misaligned slices experiments. The Dice coefficient has values between 0 and 1 and Hausdorff Distance is in mm. The metrics are presented in the form of mean ± standard deviation.

	Consistent			Misaligned		
	LVM	LVC	RVC	LVM	LVC	RVC
Dice	0.94 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	0.83 ± 0.04	0.95 ± 0.01	0.93 ± 0.02
HD (mm)	4.56 ± 1.66	3.17 ± 1.07	3.87 ± 2.28	6.30 ± 5.21	4.94 ± 3.61	5.91 ± 2.55

mapping to a much wider range of plausible output shapes. It is important to highlight the limitations of the validation method itself. The misalignment of all slices, combined, will result in an overall rigid transformation of the underlying three-dimensional shape. This rigidly transformed shape is most likely to be reconstructed by our network. If we compare this reconstruction with the original shape, the accuracy will be significantly affected, even if the shapes are exactly equal. We attempted at compensating for this unwanted effect by performing a rigid registration, but this still leaves residual errors that are partly responsible for the decrease of accuracy shown in Table 1.

For more than 85% of the 3D meshes directly generated from network predictions, the Euler characteristic is 2, suggesting a robust network output with no isolated false predictions, holes within the object or handles attached to it. However, the post-processing described in Sect. 2.3 is still needed and increases the rate by 8%.

3.2 Real Contours

The method is directly applicable to real clinical contours. One example of reconstructed meshes from clinical contours is shown in Fig. 3. Our method reconstructs plausible 3D meshes for LV myocardium and LV/RV cavities.

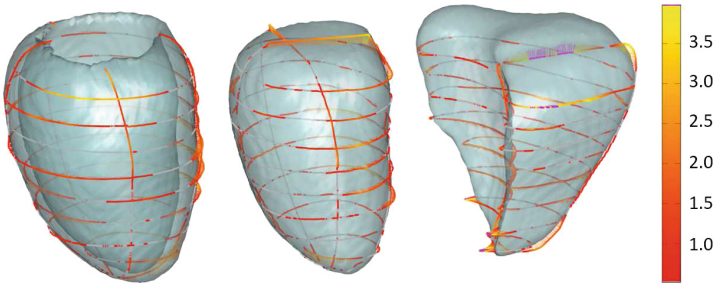

Fig. 4. Real contours aligned with the reconstructed meshes. The colours indicate the distance from contours to meshes. For distance smaller than 0.5 mm the colour is gray, and for distance larger than 4 mm the colour is purple. Colours corresponding to distances between 0.5 mm and 4 mm are shown in the color bar. (Color figure online)

The structure has no obvious distortions caused by misalignment, and the reconstructed meshes align well with the input contours. Quantitative validation in this case is complicated by the lack of gold standard 3D reconstructions (Fig. 4).

4 Conclusions

In this paper, we propose a bi-ventricular 3D mesh reconstruction method from CMR slices by transforming the mesh fitting problem into a volumetric mapping problem followed by isosurfacing. Our method takes advantage of a deep learning algorithm and is able to reconstruct anatomically realistic surfaces with a wide range of shapes from both spatially consistent and misaligned contours. The method has no constrains on the slice orientations and utilizes information from multiple SAX and LAX views simultaneously. It tolerates discrepancies between intersecting slices, and produces accurate 3D meshes from misaligned cases. We trained and evaluated our method using cases generated from a statistical shape model. We tested both the neural network predictions and reconstructed meshes using Dice coefficient and HD, and also evaluated topology with the Euler characteristic. A more detailed evaluation could be performed with more localized metrics, such as mesh quality metrics. We also applied our method to real contours and achieved good quality reconstructed meshes, suggesting the potential of the method to generate 3D meshes from CMR images directly. We showed that our network can compensate for motion artifacts such as the ones including out-of-plane misalignments. For future work, thanks to the flexibility of input slice orientations, our method can be developed to explicitly correct for image-plane rotation and translation in an iterative process to further minimize the disagreement between contours and reconstructed meshes.

Acknowledgments. We thank BHF Project Grant No. PG/16/75/32383 "Improving risk stratification in HCM through a computational anatomical analysis of ventricular remodelling" for support.

References

1. Vukicevic, M., Mosadegh, B., Min, J., Little, S.: Cardiac 3D printing and its future directions. JACC Cardiovasc. Imaging **10**, 171–184 (2017)
2. Suinesiaputra, A., et al.: Statistical shape modeling of the left ventricle: myocardial infarct classification challenge. IEEE JBHI **22**(2), 503–515 (2018)
3. Lehmann, H., et al.: Integrating viability information into a cardiac model for interventional guidance. In: Ayache, N., Delingette, H., Sermesant, M. (eds.) FIMH 2009. LNCS, vol. 5528, pp. 312–320. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01932-6_34
4. Zacur, E., et al.: MRI-based heart and Torso personalization for computer modeling and simulation of cardiac electrophysiology. In: Cardoso, M.J., et al. (eds.) BIVPCS/POCUS -2017. LNCS, vol. 10549, pp. 61–70. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67552-7_8

5. Arevalo, H., et al.: Arrhythmia risk stratification of patients after myocardial infarction using personalized heart models. Nat. Commun. **7**, 11437 (2016)
6. Deng, D., Zhang, J., Xia, L.: Three-dimensional mesh generation for human heart model. In: Li, K., Li, X., Ma, S., Irwin, G.W. (eds.) ICSEE/LSMS -2010. CCIS, vol. 98, pp. 157–162. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15859-9_22
7. AHA Writing Group on Myocardial Segmentation and Registration for Cardiac Imaging: Manuel D. Cerqueira, et al. "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association." Circulation **105**(4), 539–542 (2002)
8. Villard, B., Zacur, E., Dall'Armellina, E., Grau, V.: Correction of slice misalignment in multi-breath-hold cardiac MRI scans. In: Mansi, T., McLeod, K., Pop, M., Rhode, K., Sermesant, M., Young, A. (eds.) STACOM 2016. LNCS, vol. 10124, pp. 30–38. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-52718-5_4
9. Zou, M., Holloway, M., Carr, N., Ju, T.: Topology-constrained surface reconstruction from cross-sections. ACM Trans. Graph. **34**, 128 (2015)
10. Young, A., et al.: Left ventricular mass and volume: fast calculation with guide-point modelling on MR images. Radiology **2**, 597–602 (2000)
11. Medrano-Gracia, P., et al.: Large scale left ventricular shape atlas using automated model fitting to contours. In: Ourselin, S., Rueckert, D., Smith, N. (eds.) FIMH 2013. LNCS, vol. 7945, pp. 433–441. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38899-6_51
12. Villard, B., Grau, V., Zacur, E.: Surface mesh reconstruction from cardiac MRI contours. J. Imaging **4**(1), 16 (2018)
13. Lamata, P., et al.: An accurate, fast and robust method to generate patient-specific cubic Hermite meshes. Med. Image Anal. **15**, 801–813 (2011)
14. De Marvao, A., et al.: Population-based studies of myocardial hypertrophy: high resolution cardiovascular magnetic resonance atlases improve statistical power. J. Cardiovasc. Magn. Reson. **16**, 16 (2015)
15. Zhang, X., et al.: Atlas-based quantification of cardiac remodeling due to myocardial infarction. PloS One **9**(10), e110243 (2014)
16. Alba, X., et al.: An algorithm for the segmentation of highly abnormal hearts using a generic statistical shape model. IEEE TMI **35**(3), 845859 (2016)
17. Zhang, C., et al.: Understanding deep learning requires rethinking generalization. arXiv preprint arXiv:1611.03530 (2016)
18. Oktay, O., et al.: Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. IEEE TMI **37**(2), 384–395 (2018)
19. Duan, J., et al.: Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach. IEEE TMI (2019)
20. McLeish, K., Hill, D.L.G., Atkinson, D., Blackall, J.M., Razavi, R.: A study of the motion and deformation of the heart due to respiration. IEEE TMI **21**(9), 1142–1150 (2002)
21. Shechter, G., Ozturk, C., Resar, J.R., McVeigh, E.R.: Respiratory motion of the heart from free breathing coronary angiograms. IEEE TMI **23**, 1046–1056 (2004)
22. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, pp. 417–424. ACM Press/Addison-Wesley Publishing Co. (2000)

23. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
24. Bai, W., et al.: A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion. Med. Image Anal. **26**, 133–145 (2015)
25. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 4th 3DV, pp. 565–571 (2016)
26. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. In: ACM Siggraph Computer Graphics, vol. 21, no. 4, pp. 163–169. ACM (1987)