CNN-Based Cardiac Motion Extraction to Generate Deformable Geometric Left Ventricle Myocardial Models from Cine MRI

Roshan Reddy Upendra[1(✉)], Brian Jamison Wentz[3,5], Richard Simon[2], Suzanne M. Shontz[3,4,5], and Cristian A. Linte[1,2]

[1] Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, USA
ru6928@rit.edu
[2] Biomedical Engineering, Rochester Institute of Technology, Rochester, NY, USA
[3] Bioengineering Program, University of Kansas, Lawrence, KS, USA
[4] Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS, USA
[5] Information and Telecommunication Center, University of Kansas, Lawrence, KS, USA

Abstract. Patient-specific left ventricle (LV) myocardial models have the potential to be used in a variety of clinical scenarios for improved diagnosis and treatment plans. Cine cardiac magnetic resonance (MR) imaging provides high resolution images to reconstruct patient-specific geometric models of the LV myocardium. With the advent of deep learning, accurate segmentation of cardiac chambers from cine cardiac MR images and unsupervised learning for image registration for cardiac motion estimation on a large number of image datasets is attainable. Here, we propose a deep leaning-based framework for the development of patient-specific geometric models of LV myocardium from cine cardiac MR images, using the Automated Cardiac Diagnosis Challenge (ACDC) dataset. We use the deformation field estimated from the VoxelMorph-based convolutional neural network (CNN) to propagate the isosurface mesh and volume mesh of the end-diastole (ED) frame to the subsequent frames of the cardiac cycle. We assess the CNN-based propagated models against segmented models at each cardiac phase, as well as models propagated using another traditional nonrigid image registration technique. Additionally, we generate dynamic LV myocardial volume meshes at all phases of the cardiac cycle using the log barrier-based mesh warping (LBWARP) method and compare them with the CNN-propagated volume meshes.

Keywords: Patient-specific modeling · Deep learning · Image registration · Cine Cardiac MRI · Mesh warping

1 Introduction

To reduce the morbidity and mortality associated with cardiovascular diseases (CVDs) [3], and to improve their treatment, it is crucial to detect and predict the

D. B. Ennis et al. (Eds.): FIMH 2021, LNCS 12738, pp. 253–263, 2021.
https://doi.org/10.1007/978-3-030-78710-3_25

progression of the diseases at an early stage. In a clinical set-up, population-based metrics, including measurements of cardiac wall motion, ventricular volumes, cardiac chamber flow patterns, etc., derived from cardiac imaging are used for diagnosis, prognosis and therapy planning.

In recent years, image-based computational models have been increasingly used to study ventricular mechanics associated with various cardiac conditions. A comprehensive review of patient-specific cardiovascular modeling and its applications is described in [17]. Cardiovascular patient-specific modeling includes a geometric representation of some or all cardiac chambers of the patient's anatomy and is derived from different imaging modalities [8].

The construction of patient-specific geometric models entails several steps: clinical imaging, segmentation and geometry reconstruction, and spatial discretization (i.e., mesh generation) [12]. For example, Bello *et al.* [2] presented a deep learning based framework for human survival prediction for patients diagnosed with pulmonary hypertension using cine cardiac MR images. Here, the authors employ a 4D spatio-temporal B-spline image registration method to estimate the deformation field at each voxel and at each timeframe. The estimated deformation field was used to propagate the ED surface mesh of the right ventricle (RV), reconstructed from the segmentation map, to the rest of the timeframes of a particular subject. Cardiac MRI is a current gold standard to assess global (ventricle volume and ejection fraction) and regional (kinematics and contractility) function of the heart under various diseases. In particular, cardiac MRI enables the generation of high quality myocardial models, which can, in turn, be used to identify reduced function.

In this work, we propose a deep learning-based pipeline to develop patient-specific geometric models of the LV myocardium from cine cardiac MR images (Fig. 1). These models may be used to conduct various simulations, such as assessing myocardial viability. In our previous work [19], we introduced a preliminary, proof of concept, CNN-based 4D deformable registration method for cardiac motion estimation from cine cardiac MR images, using the ACDC dataset [4]. Here, we demonstrate the use of the CNN-based 4D deformable registration technique to build dynamic patient-specific LV myocardial models across subjects with different pathologies, namely normal, dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM) and subjects with prior myocardial infarctions (MINF). Following segmentation of the ED cardiac frame, we generate both isosurface and volume LV meshes, which we then propagate through the cardiac cycle using the CNN-based registration fields. In addition, we demonstrate the generation of dynamic LV volume meshes depicting the heart at various cardiac phases by warping a patient-specific ED volume mesh based on the registration-based propagated surface meshes, using the LBWARP method [15]. Lastly, we compare these meshes to those obtained by directly propagating the ED volume mesh using the CNN-based deformation fields.

2 Methodology

2.1 Cardiac MRI Data

We use the 2017 ACDC dataset that was acquired from real clinical exams. The dataset is composed of cine cardiac MR images from 150 subjects, divided into five equally-distributed subgroups: normal, MINF, DCM, HCM and abnormal RV. The MR image acquisitions were obtained using two different MR scanners of 1.5 T and 3.0 T magnetic strength. These series of short axis slices cover the LV from base to apex such that one image is captured every 5 mm to 10 mm with a spatial resolution of 1.37 mm^2/pixel to 1.68 mm^2/pixel.

2.2 Image Preprocessing

We first correct for the inherent slice misalignments that occur during the cine cardiac MR image acquisition. We train a modified version of the U-Net model [13] to segment the cardiac chambers, namely LV blood-pool, LV myocardium and RV blood-pool, from 2D cardiac MR images. We identify the LV blood-pool center, i.e., the centroid of the predicted segmentation mask and stack the 2D cardiac MR slices collinearly to obtain slice misalignment corrected 3D images [7,19].

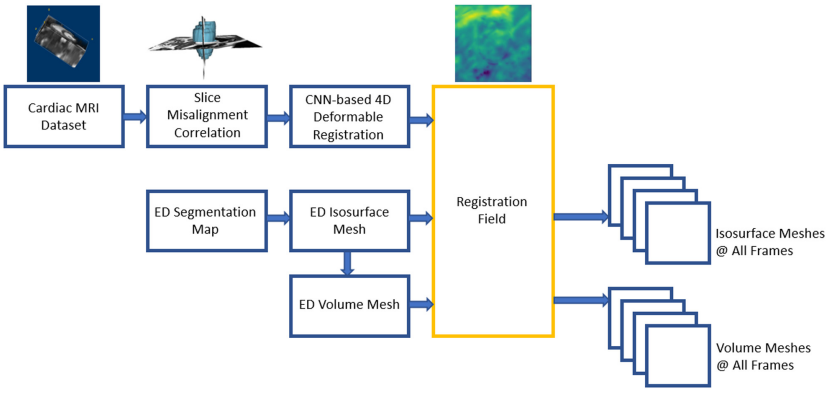

Fig. 1. Overview of the proposed CNN-based workflow to generate patient-specific LV myocardial geometric model.

2.3 Deformable Image Registration

CNN-Based Image Registration. We leverage our 4D deformable registration method described in [19] which employs the VoxelMorph [1] framework to

determine the optical flow representation between the slice misalignment corrected 3D images. The CNN is trained using the following loss function:

$$L = L_{\text{similarity}} + \lambda L_{\text{smooth}}, \tag{1}$$

where $L_{\text{similarity}}$ is the mean squared error (MSE) between the target frame and the warped frame, L_{smooth} is the smoothing loss function to spatially smooth the registration field, and λ is the regularization parameter, which is set to 10^{-3} in our experiments. Inspired by Zhu *et al.* [20], we use the Laplacian operator in the smoothing loss function as it considers both global and local properties of the objective function $y = x^2$ instead of the traditional gradient operator which considers only the local properties of the function $y = x^2$. A detailed comparison of both these smoothing loss functions with respect to cardiac motion estimation from cine MR images is found in [19].

The 4D cine cardiac MRI datasets are composed of 28 to 40 3D image frames that cover the complete cardiac cycle. For this discussion, we shall refer to the 3D images as I_{ED}, I_{ED+1},...,I_{ED+N_T-1} where I_{ED} is the ED image frame, and N_T is the total number of 3D images. We employ the fixed reference frame registration method, wherein the task is to find an optical flow representation between the image pairs $\{(I_{ED}, I_{ED+t})\}_{t=1,2,3,...,N_T-1}$.

During training, we use 110 of the total 150 MR image dataset for training, 10 for validation and the remaining 30 for testing. The CNN for cardiac motion estimation is trained using an Adam optimizer with a learning rate of 10^{-4}, halved at every 10^{th} epoch for 50 epochs. Both the U-Net model used for segmentation and slice misalignment correction, and the VoxelMorph network trained to estimate cardiac motion were trained on NVIDIA RTX 2080 Ti GPU.

Conventional Image Registration. We compare the performance of the VoxelMorph framework with that of the B-spline free form deformation (FFD) non-rigid image registration algorithm [14]. This iterative intensity-based image registration method was implemented using SimpleElastix [9,11], which enables a variety of image-registration algorithms in different programming languages. The FFD algorithm was set to use the adaptive stochastic gradient descent method as the optimizer, MSE as the similarity measure and binding energy as the regularization function. The FFD-based image registration was optimized in 500 iterations, while sampling 2048 random points per iteration, on an Intel(R) Core(TM) i9-9900K CPU.

2.4 Mesh Generation and Propagation

We use the manual segmentation map of the ED frame to generate isosurface meshes. The slice thickness of each MRI image slice is 5 mm to 10 mm; however, in order to obtain good quality meshes, the segmentation maps were resampled to a slice thickness of 1 mm. We use the Lewiner marching cubes [10] algorithm to generate the meshes from the resampled segmentation maps of the ED frames, on an Intel(R) Core(TM) i9-9900K CPU, and then simplification techniques, such as

r-refinement and edge collapse, were performed using MeshLab 2020.07 [5]. The simplification techniques are repeated multiple times to reduce the number of vertices until the mesh has been fully decimated while preserving the anatomical integrity and aspect ratio of the isosurface meshes.

Volume meshes of the initial surface meshes at the ED phases for four patients with various heart conditions were generated based on the decimated patient-specific surface meshes using Tetgen 1.6 [16]. In particular, a constrained Delaunay mesh generation algorithm was used to generate tetrahedral meshes based on the triangulated surface meshes. Steiner points were added within the boundary of the surface mesh so that the tetrahedra maintained a radius-edge ratio of 1.01 and a maximum volume of 9 mm^3 as needed for generation of valid meshes [16]. Volume mesh quality improvement was performed using the feasible Newton method in Mesquite [18]. This method iteratively minimizes the quadratic approximation of a nonlinear function and converges linearly toward a local minimum while performing an Armijo line search to ensure feasibility of the elements; feasibility in this case refers to a valid, non-inverted element. The volume mesh converged to the highest quality indicated by the minimum average scaled Jacobian of the elements in the mesh.

To demonstrate the VoxelMorph-based motion extraction and propagation to build patient-specific LV myocardial models, we generate two sets of volume meshes at each cardiac frame for each patient in each pathology group (Fig. 2).

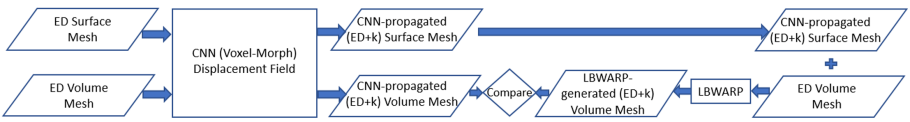

Fig. 2. Pipeline to generate dynamic volume meshes (at cardiac frames (ED + k)) by direct CNN-based propagation, as well as volume mesh warping based on dynamic boundary meshes.

The first set is produced by propagating the volume meshes at the ED frame to all the subsequent frames of the cardiac cycle using the deformation field estimated by the VoxelMorph-based registration method. For the second set, the ED volume mesh generated with Tetgen was used to generate the volume meshes corresponding to the other cardiac phases. We employed the LBWARP method [15] to deform the ED volume mesh onto the target surface mesh for the new cardiac phase (Fig. 3). The method computes new positions for the interior vertices in the ED volume mesh, while maintaining the mesh topology and point-to-point correspondence [15]. The simplification of the ED isosurface meshes, generation of the ED volume meshes and generation of the volume meshes corresponding to the other cardiac phases using the LBWARP method were performed on a machine equipped with AMD FX(tm)-6300 Six-Cores processor and a NVIDIA GeForce GTX 1050 Ti graphics card.

Briefly, LBWARP first calculates a set of local weights for each interior vertex in the initial ED volume mesh based on the relative inverse distances from each of its neighbors. The projected Newton method is used to solve the strictly convex optimization problems. For each set of local weights, a sparse system of linear equations is formed specifying the representation of each interior vertex in terms of its neighbors. Next, the boundary vertices in the ED surface mesh are mapped onto the new surface boundary. Finally, the interior vertices in the ED volume mesh are repositioned by solving the original system of linear equations with new right-hand side vectors to reflect the updated positions of the boundary nodes, while maintaining edge connectivity and point-to-point correspondence, and ultimately yielding the volume meshes that correspond to each new cardiac phase.

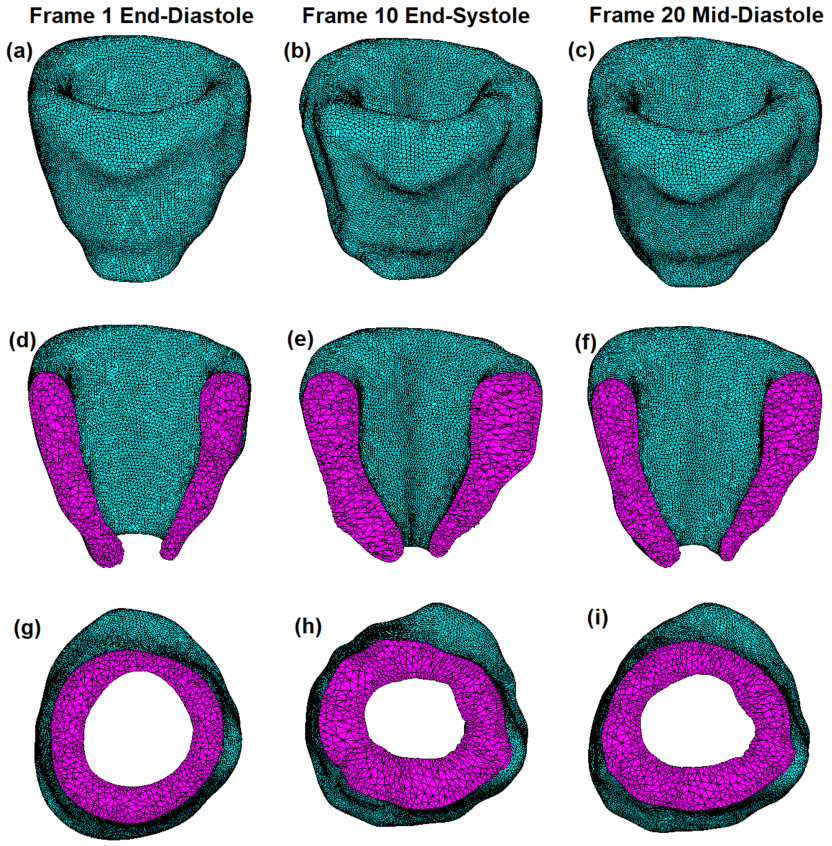

Fig. 3. Warped volume meshes for a patient with a healthy heart generated using LBWARP at three cardiac phases (a) end-diastole; (b) end-systole; and (c) mid-diastole; (d–f) Long axis cutaway view of volume meshes at the three cardiac phases, respectively; (g–i) short-axis cutaway view of volume meshes at the three cardiac phases, respectively.

3 Results and Discussion

To evaluate the registration performance, the LV isosurface (generated from the ED image segmentation map) is propagated to all the subsequent cardiac frames using the deformation field estimated by FFD and VoxelMorph. We then compare these isosurfaces to those directly generated by segmenting all cardiac image frames using a modified U-Net model [13] (Sect. 2.2), which we refer to as the "silver standard".

Table 1 summarizes the performance of the FFD and VoxelMorph registration by assessing the Dice score and mean absolute distance (MAD) between the propagated and directly segmented (i.e., "silver standard") isosurfaces.

Figure 4 illustrates the distance between the three sets of isosurfaces (segmented, CNN-propagated and FFD-propagated) for one patient randomly selected from each pathology. The MAD between the surfaces is less than 2 mm at all frames, with the CNN-propagated isosurfaces being closest to the "silver standard" segmented surfaces. Figure 5 illustrates the model-to-model distance between the FFD-propagated and CNN-propagated isosurface meshes at end-systole (ES) and mid-diastole frames for subjects from all four pathologies.

Since the CNN-propagated isosurfaces are in closer agreement to the "silver standard" segmented surfaces when compared to the FFD-propagated isosurfaces, we use the CNN-propagation method to generate the volume meshes at each phase of the cardiac cycle. As mentioned in Sect. 2.4 and shown in Fig. 2, we generate two sets of volume meshes at each frame of the cardiac cycle. Figure 6 shows the mean node distance between the two sets of volume meshes across all cardiac frames for one subject in each of the four pathologies. Figure 6 also shows the mean node distance between the two sets of volume meshes at each frame of the cardiac cycle for the four subjects. It can be observed that the two sets of volume meshes are in close agreement with each other, exhibiting a mesh-to-mesh distance within 0.5 mm.

Table 1. Mean Dice score (%) and mean absolute distance (MAD) (mm) between FFD and segmentation (FFD-SEG), CNN and segmentation (CNN-SEG), and FFD and CNN (FFD-CNN) results. Statistically significant differences were evaluated using the t-test (* for $p < 0.1$ and ** for $p < 0.05$).

	Normal		MINF		DCM		HCM	
	Dice	MAD	Dice	MAD	Dice	MAD	Dice	MAD
FFD-Segmentation	74.80	1.53	77.69	1.09	80.41	0.91	77.39	1.97
CNN-Segmentation	80.41**	1.15	81.21*	0.87	83.39*	0.91	82.46*	1.09
FFD-CNN	77.81	1.13	82.12	0.75	81.67	0.97	77.34	1.77

We also briefly investigated the effect of using initial-to-final frame vs. adjacent frame-to-frame registration to extract the cardiac motion throughout the

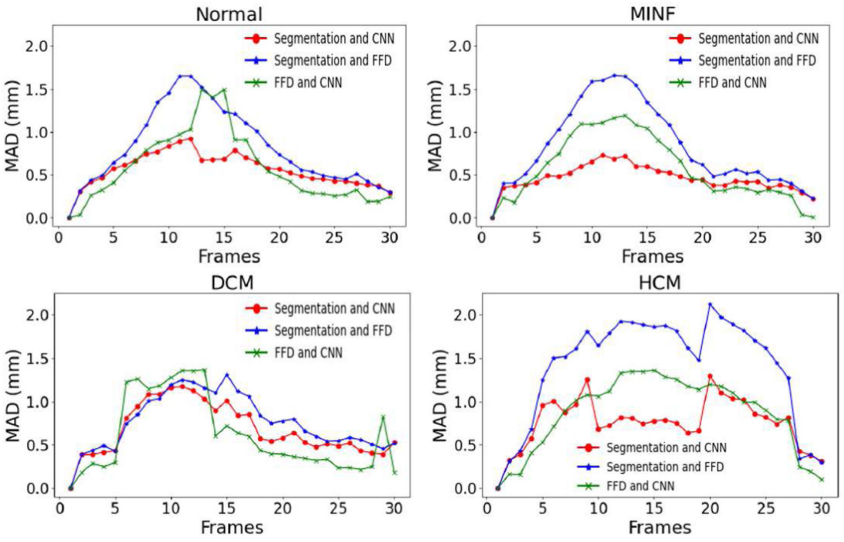

Fig. 4. MAD between FFD- and CNN-propagated, and segmented (i.e., "silver standard") isosurfaces at all cardiac frames for all patient pathologies.

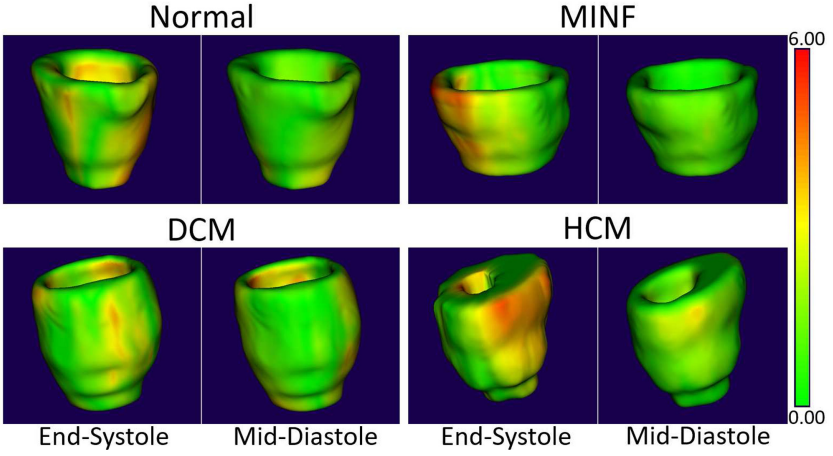

Fig. 5. Model-to-model distance between the isosurface meshes generated from FFD- and the CNN-propagation method for all patient pathologies at end-systole and mid-diastole frames.

cycle. Although the sequential registration method estimates smaller deformation between two consecutive, adjacent image frames compared to the larger deformations estimated by the initial-to-final frame registration, their concatenation across several frames accumulates considerable registration errors. As such, when using these concatenated registration-predicted deformation fields

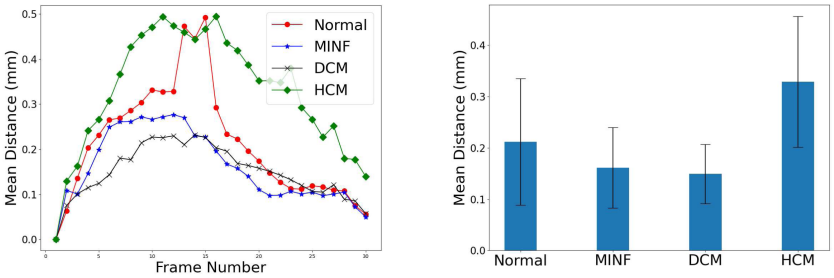

Fig. 6. Mean node-to-node distance at each cardiac frame between the CNN-propagated and LBWARP-generated volume meshes (left); mean (std-dev) node distance across all frames for each patient pathology (right).

to propagate the ED isosurfaces and volume meshes to the subsequent cardiac phases, the Dice score and MAD between the propagated and segmented geometries rapidly deteriorate, along with the quality of the propagated surface and volume meshes.

Following the generation of the dynamic, multi-phase meshes, we also assessed the quality of the ES meshes. One set of ES meshes was generated by propagating the ED mesh using the CNN-based extracted motion, while the other set of ES meshes was generated by warping the ED volume mesh based on the dynamic boundary meshes via the LBWARP approach. Unlike the starting ED phase meshes, the ES phase meshes contain some lower quality elements indicated by the lower minimum scaled Jacobian values, but are still suitable for use in simulations.

Moreover, although the proposed VoxelMorph-based cardiac motion extraction method can capture the frame-to-frame motion with sufficient accuracy, as shown in this work, our ongoing and future efforts are focused on further improving the algorithm by imposing diffeomorphic deformations [6]. This improvement will help maintain high mesh quality and prevent mesh tangling and element degeneration, especially for the systolic phases.

4 Conclusion

In this work, we show that the proposed deep learning framework can be used to build LV myocardial geometric models. The proposed framework is not limited to any pathology and can be extended to LV and RV blood-pool geometry.

Acknowledgments. This work was supported by grants from the National Science Foundation (Award No. OAC 1808530, OAC 1808553 & CCF 1717894) and the National Institutes of Health (Award No. R35GM128877).

References

1. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: VoxelMorph: a learning framework for deformable medical image registration. IEEE Trans. Med. Imaging **38**(8), 1788–1800 (2019)
2. Bello, G.A., et al.: Deep-learning cardiac motion analysis for human survival prediction. Nature Mach. Intell. **1**(2), 95–104 (2019)
3. Benjamin, E.J., et al.: Heart disease and stroke statistics-2017 update: a report from the American heart association. Circulation **135**(10), e146–e603 (2017)
4. Bernard, O., et al.: Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? IEEE Trans. Med. Imaging **37**(11), 2514–2525 (2018)
5. Cignoni, P., et al.: Meshlab: an open-source mesh processing tool. In: Eurographics Italian Chapter Conference. vol. 2008, pp. 129–136. Salerno, Italy (2008)
6. Dalca, A.V., Balakrishnan, G., Guttag, J., Sabuncu, M.R.: Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. Med. Image Anal. **57**, 226–236 (2019)
7. Dangi, S., Linte, C.A., Yaniv, Z.: Cine cardiac MRI slice misalignment correction towards full 3D left ventricle segmentation. In: Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling. vol. 10576, p. 1057607. International Society for Optics and Photonics (2018)
8. Gray, R.A., Pathmanathan, P.: Patient-specific cardiovascular computational modeling: diversity of personalization and challenges. J. Cardiovascular Transl. Res. **11**(2), 80–88 (2018)
9. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging **29**(1), 196–205 (2009)
10. Lewiner, T., Lopes, H., Vieira, A.W., Tavares, G.: Efficient implementation of marching cubes' cases with topological guarantees. J. Graph. Tools **8**(2), 1–15 (2003)
11. Marstal, K., Berendsen, F., Staring, M., Klein, S.: SimpleElastix: a user-friendly, multi-lingual library for medical image registration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 134–142 (2016)
12. Morris, P.D., et al.: Computational fluid dynamics modelling in cardiovascular medicine. Heart **102**(1), 18–28 (2016)
13. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
14. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging **18**(8), 712–721 (1999)
15. Shontz, S.M., Vavasis, S.A.: A mesh warping algorithm based on weighted Laplacian smoothing. In: 12^{th} International Meshing Roundtable, pp. 147–158 (2003)
16. Si, H.: Tetgen, a Delaunay-based quality tetrahedral mesh generator. ACM Trans. Math. Softw. (TOMS) **41**(2), 1–36 (2015)
17. Smith, N., et al.: Euheart: personalized and integrated cardiac care using patient-specific cardiovascular modelling. Interface Focus **1**(3), 349–364 (2011)

18. Trilinos Project Team, T.: The Trilinos Project Website (2020). https://trilinos.github.io. Accessed 12 Nov 2020
19. Upendra, R.R., Wentz, B.J., Shontz, S.M., Linte, C.A.: A convolutional neural network-based deformable image registration method for cardiac motion estimation from cine cardiac MR images. In: 2020 Computing in Cardiology, pp. 1–4. IEEE (2020)
20. Zhu, Y., Zhou Sr., Z., Liao Sr., G., Yuan, K.: New loss functions for medical image registration based on Voxelmorph. In: Medical Imaging 2020: Image Processing, vol. 11313, p. 113132E. International Society for Optics and Photonics (2020)