

A Convolutional Neural Network-based Deformable Image Registration Method for Cardiac Motion Estimation from Cine Cardiac MR Images

Roshan Reddy Upendra¹, Brian Jamison Wentz^{3,5}, Suzanne M Shontz^{3,4,5} and Cristian A Linte^{1,2}

¹ Chester F Carlson Center for Imaging Science, ² Biomedical Engineering, Rochester Institute of Technology, Rochester, NY, USA

³ Bioengineering Graduate Program, ⁴ Electrical Engineering and Computer Science, ⁵ Information and Telecommunication Technology Center, University of Kansas, Lawrence, KS, USA

Abstract

In this work, we describe an unsupervised deep learning framework featuring a Laplacian-based operator as smoothing loss for deformable registration of 3D cine cardiac magnetic resonance (CMR) images. Before registration, the input 3D images are corrected for slice misalignment by segmenting the left ventricle (LV) blood-pool, LV myocardium and right ventricle (RV) blood-pool using a U-Net model and aligning the 2D slices along the center of the LV blood-pool. We conducted experiments using the Automated Cardiac Diagnosis Challenge (ACDC) dataset. We used the registration deformation field to warp the manually segmented LV blood-pool, LV myocardium and RV blood-pool labels from end-diastole (ED) frame to the other frames in the cardiac cycle. We achieved a mean Dice score of 94.84%, 85.22% and 84.36%, and Hausdorff distance (HD) of 2.74 mm, 5.88 mm and 9.04 mm, for the LV blood-pool, LV myocardium and RV blood-pool, respectively. We also introduce a pipeline to estimate patient tractography using the proposed CNN-based cardiac motion estimation.

1. Introduction

Regional myocardial function assessment entails accurate characterization of several parameters such as myocardial motion, strain, torsion and wall thickness, which have been shown to help better understand myocardial diseases [1]. To this end, accurate cardiac motion estimation from CMR images is important, as it helps assess the kinematic and contractile properties of the myocardium, both of which are direct quantifiers of regional heart function.

Cardiac motion estimation involves finding an optical flow representation between consecutive 3D frames of a 4D cine CMR dataset and, therefore, can be formulated as an image registration problem. Deep learning algorithms have gained increased popularity in medical image registration. Qin *et al.* [2] proposed a joint deep learning net-

work for cardiac motion estimation and segmentation of 2D cine cardiac MR images. Qiu *et al.* [3] compared the performance of supervised and unsupervised training strategies for cardiac motion estimation using convolutional neural networks (CNN), performed in the 2D plane. Morales *et al.* [4] proposed an unsupervised CNN-based 3D deformable registration method for cardiac motion estimation; however, they do not account for the out-of-plane motion of the two-dimensional stack of the CMR images that leads to slice misalignment.

Here we propose a CNN-based 4D deformable registration technique for consistent motion estimation from 4D cine CMR images. We assess the performance of the proposed framework on the ACDC dataset [5] and also compare the effect of the gradient-based operator and the Laplacian-based operator as regularization terms in the loss function used to back-propagate the CNN (Fig. 1b).

2. Methods

2.1. Data

We used the 2017 ACDC challenge dataset [5], consisting of short axis cine CMR images for 150 subjects, acquired using two different MRI scanners of magnetic strength 1.5 T and 3.0 T. The dataset is evenly split into five disjoint subgroups - normal, dilated cardiomyopathy, hypertrophic cardiomyopathy, abnormal RV and subjects with prior myocardial infarctions.

2.2. Slice Misalignment Correction

To reduce artifacts caused by inherent slice misalignments during cine CMR image acquisition, we leverage the slice misalignment correction method presented by Dangi *et al.* [6]. We train a modified version of the U-Net model [7] inspired from Isensee *et al.* [8], to segment the cardiac chambers (LV blood-pool, LV myocardium and RV blood-pool) from 2D CMR slices. We use these predicted segmentation maps to crop the regions of interest (RoI) and to identify the centers of the LV blood-pool. The 2D

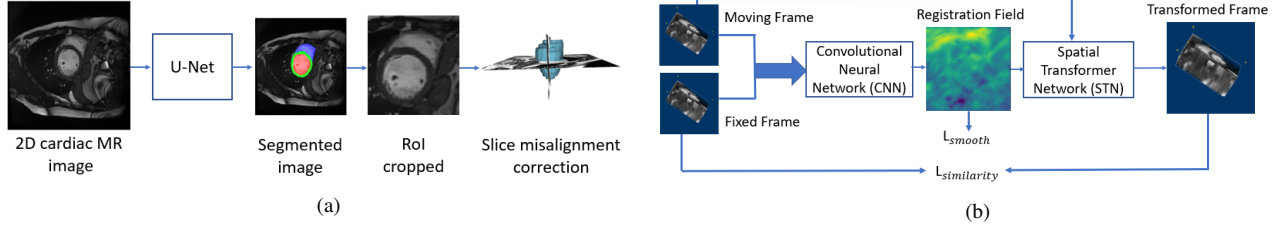


Figure 1: (a) Slice misalignment correction and (b) 4D deformable registration workflow.

slices are stacked such that the LV blood-pool centers are collinear, resulting in a slice misalignment corrected 3D CMR image (Fig. 1a). The U-Net model is trained on the ED and end-systole (ES) frames of the CMR data of 80 subjects and validated on 20 subjects.

2.3. 4D Deformable Registration

We employ the VoxelMorph [9] framework to find an optical flow representation between a sequence of 3D image pairs $\{(I_{ED}, I_{ED+t})\}_{t=1,2,3,\dots,N_T-1}$ where N_T is the total number of frames, and at each iteration, an image pair (I_{ED}, I_{ED+t}) is input to the CNN and a registration field ϕ is output. The registration field is fed to a spatial transformer network (STN) [10] along with the ED frame, I_{ED} , to produce a warped image, $I_{ED} \circ \phi$ (Fig. 1b).

To train the CNN, a loss function consisting of two components is used to optimize the network:

$$L = L_{similarity} + \lambda L_{smooth}, \quad (1)$$

where $L_{similarity}$ is the mean squared error (MSE) between the target frame I_{ED+t} and the warped ED frame $I_{ED} \circ \phi$:

$$MSE = \frac{1}{|\Omega|} \sum_{i \in \Omega} (I_{ED+t}(i) - [I_{ED} \circ \phi](i))^2, \quad (2)$$

where Ω is the spatial domain of the images and $i \in \mathbb{R}^2$ is the position of a point on the frame. The second term in the loss function (eq. (1)) is a smoothing loss function L_{smooth} that spatially smoothes the registration field ϕ and λ is the regularization parameter. In general, a diffusion regularizer on the spatial gradients of the registration field is used as the smoothing loss function and is given by

$$L_{smooth} = \sum_{i \in \Omega} \|\nabla \phi(i)\|^2. \quad (3)$$

Here we experiment with a Laplacian operator in the smoothing loss function, inspired from Zhu *et al.* [11]:

$$L_{smooth} = \sum_{i \in \Omega} \|\Delta \phi(i)\|^2. \quad (4)$$

Unlike the gradient operator, which only considers the local properties of the objective function $y = x^2$, the

Laplacian operator considers the global properties of the function $y = x^2$, i.e., it considers the slope magnitude and its trends when choosing a direction. [11].

We divide the total 150 CMR dataset into 110 for training, 10 for validation and 30 for testing. All the cropped input cine CMR frames are resampled to $96 \times 96 \times 16$ voxels with 1.5 mm isotropic resolution. We train the VoxelMorph CNN using the Adam optimizer with a learning rate of 10^{-4} , halved at every 10^{th} epoch for 50 epochs on a machine equipped with a NVIDIA RTX 2080 Ti GPU with 11 GB of memory; the regularization parameter λ is set to 10^{-3} .

3. Results

The manual segmentation labels of the LV blood-pool, LV myocardium and RV blood-pool for ED and ES frames of 100 subjects are provided in the ACDC challenge dataset. To evaluate the performance of the CNN-based deformable registration algorithm, we warp the segmentation map of the ED frame to ES frame using the estimated registration field, and compute the Dice score and Hausdorff distance (HD) between the segmentation map of ES frame and the warped segmentation map of ED frame. We refer to this as the “gold” standard comparison, as the segmentation maps used for comparison are manually annotated by experts. We also warp the segmentation map of the ED frame to all subsequent cardiac frames, and compute the Dice score and HD between the warped segmentation map of ED frame and the segmentation maps predicted by the modified U-Net model [8]. We refer to this as the “silver” standard comparison, as the segmentation masks used as reference were not annotated by experts, but rather were generated using techniques previously validated against expert annotations.

In Table 1, we show the mean Dice score and mean HD for LV blood-pool, LV myocardium and RV blood-pool before registration (post misalignment correction) and after registration on the test dataset, for both “gold” and “silver” standard comparisons. We also compare the effect of the gradient-based operator and the Laplacian-based operator on the VoxelMorph-based deformable registration method.

Our proposed method achieves a 83.04% Dice score and 8.46 mm HD for all cardiac chambers following regis-

Table 1: Mean Dice score and Hausdorff distance (HD) for LV blood-pool (LV), LV myocardium (MC) and RV blood-pool (RV), for both “gold” and “silver” standard comparisons. Statistically significant differences between the registration metrics *before* and *after* registration were evaluated using the Student t-test and are reported using * for $p < 0.05$ and ** for $p < 0.005$. The best evaluation metrics achieved are labeled in **bold**.

		Dice (%)			HD (mm)		
		LV	MC	RV	LV	MC	RV
ED to ES frames - Gold standard	Before registration	87.30	69.15	70.18	7.22	8.93	11.85
	VoxelMorph (gradient)	92.17**	79.39**	77.58*	5.59*	8.05	11.75
	VoxelMorph (Laplacian)	93.73**	80.59**	79.63*	5.11*	7.98*	11.62
ED to all frames - Silver standard	Before registration	81.29	80.15	77.32	3.13	6.08	8.61
	VoxelMorph (gradient)	94.67**	84.08**	82.73*	2.51	6.07	8.96
	VoxelMorph (Laplacian)	94.84**	85.22**	84.36**	2.74	5.88*	9.04

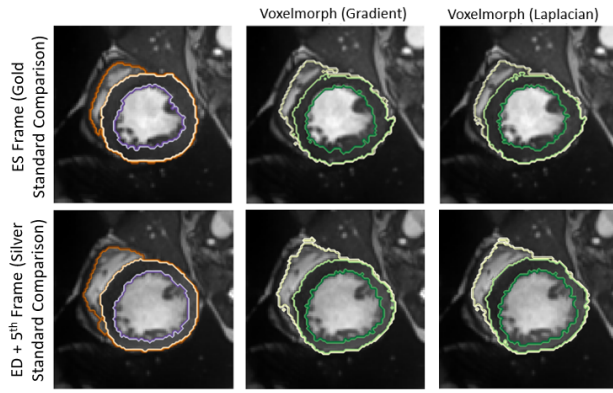


Figure 2: Panel 1-1: ES CMR slice with manually annotated segmentation contours of cardiac chambers overlaid on the slice; Panel 1-2: post registration contours using gradient-based operator as smoothing loss with segmentation contours of warped ED frame overlaid on ES frame (Dice: 81.32%, HD: 3.64 mm); Panel 1-3: post registration contours using Laplacian-based operator as smoothing loss (Dice: 83.56%, HD: 3.48 mm). Panel 2-1: ED + 5th frame CMR slice with segmentation contours obtained from U-Net model; Panel 2-2: post registration contours using gradient-based operator as smoothing loss (Dice: 92.36%, HD: 4.12 mm); Panel 2-3: post registration contours using Laplacian-based operator as smoothing loss (Dice: 92.42%, HD: 4.12 mm).

tration using the gradient-based smoothing loss function, and a 84.65% Dice score and 8.23 mm HD following registration using the Laplacian-based smoothing loss function, for our “gold” standard comparison evaluated at ES frames. Similarly, for our “silver” standard comparison, conducted across all frames, we report a 87.16% Dice score and 5.84 mm HD following registration using the gradient-based smoothing loss function and a 88.14% Dice score and 5.88 mm HD following registration using the

Laplacian-based smoothing loss function. Fig. 2 shows the cardiac chamber contours propagated using our registration from ED frame to the other cardiac frames.

4. Discussion and Conclusion

We present a deep learning-based 4D deformable registration method for cardiac motion estimation from 3D cine CMR images. The workflow also includes a slice misalignment correction step that alleviates the challenges associated with out-of-plane motion in the slice stack that would otherwise impact frame-to-frame image registration and motion extraction. In addition, we evaluate and compare the effect of the gradient-based operator and the Laplacian-based operator for smoothing the registration field on the performance of VoxelMorph-based registration network for cardiac motion estimation. We observe that the Laplacian-based smoothing loss function regularizes better than the gradient-based smoothing loss function. This can be attributed to the fact that the gradient operator only considers the local properties of the objective function $y = x^2$ and the Laplacian operator considers global properties more than the gradient.

In Fig. 3, we show our intended application of cardiac motion estimation, wherein the registration field obtained from our CNN-based 4D deformable registration is used to propagate the patient-specific anatomy and tractography information from the ED frame to its subsequent frames. The ex-vivo cardiac fiber architecture (CFA) was acquired in 21 diffusion-tensor magnetic resonance imaging (DT-MRI) scans and made publicly available by the CardioVascular Research Grid from John Hopkins University. The patient geometry and CFA are resampled isotropically, and an affine transformation is performed on the CFA while maintaining the integrity of the patient geometry. Afterwards, large deformation diffeomorphic metric mapping (LDDMM) registration is performed to create an invertible diffeomorphic map to warp and fit all 21 CFA scans to the patient geometry. The tractography reconstruction was

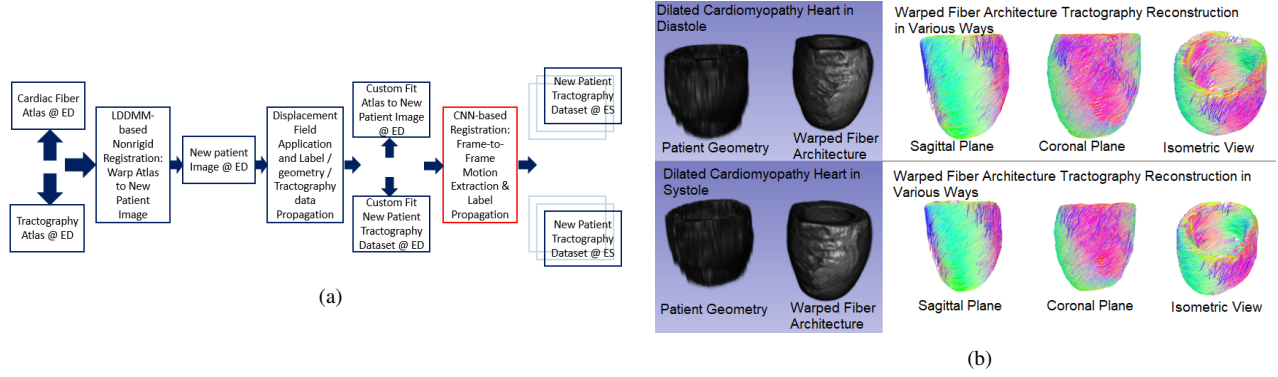


Figure 3: (a) Pipeline to estimate patient tractography using CNN-based frame-to-frame motion extraction. (b) Left: Volume renderings of segmented LV myocardium patient geometry with dilated cardiomyopathy and accompanying CFA acquired separately via *ex-vivo* DT-MRI warped to fit patient geometry. Right: Tractography reconstruction renderings of the warped CFA given the DT-MRI gradient table.

performed on the warped and fitted CFA using generalized Q-sampling imaging (GQI) with a diffusion sampling length ratio of 0.90 given its accompanying DT-MRI gradient table converted to a b-table using a b-value of 1500 s/mm^2 . This workflow is shown in Fig. 3a and patient-specific tractography examples are shown in Fig. 3b.

Ultimately, we intend to use this technique to build dynamic patient-specific myocardial models with associated fiber architecture for biomechanical cardiac simulations.

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] Shi W, Zhuang X, Wang H, Duckett S, Luong DV, Tobon-Gomez C, Tung K, Edwards PJ, Rhode KS, Razavi RS, et al. A comprehensive cardiac motion estimation framework using both untagged and 3-D tagged MR images based on nonrigid registration. *IEEE Transactions on Medical Imaging* 2012;31(6):1263–1275.
- [2] Qin C, Bai W, Schlemper J, Petersen SE, Piechnik SK, Neubauer S, Rueckert D. Joint learning of motion estimation and segmentation for cardiac MR image sequences. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018; 472–480.
- [3] Qiu H, Qin C, Le Folgoc L, Hou B, Schlemper J, Rueckert D. Deep learning for cardiac motion estimation: supervised vs. unsupervised training. In *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2019; 186–194.
- [4] Morales MA, Izquierdo-Garcia D, Aganj I, Kalpathy-Cramer J, Rosen BR, Catana C. Implementation and validation of a three-dimensional cardiac motion estimation network. *Radiology Artificial Intelligence* 2019;1(4):e180080.
- [5] Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng PA, Cetin I, Lekadir K, Camara O, Ballester MAG, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE Transactions on Medical Imaging* 2018; 37(11):2514–2525.
- [6] Dangi S, Linte CA, 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*, volume 10576. International Society for Optics and Photonics, 2018; 1057607.
- [7] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015; 234–241.
- [8] Isensee F, Jaeger PF, Full PM, Wolf I, Engelhardt S, Maier-Hein KH. Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. In *International Workshop on Statistical Atlases and Computational Models of the Heart*. Lecture Notes in Computer Science, volume 10663. Springer, 2017; 120–129.
- [9] Balakrishnan G, Zhao A, Sabuncu MR, Guttat J, Dalca AV. Voxelmorph: a learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging* 2019;38(8):1788–1800.
- [10] Jaderberg M, Simonyan K, Zisserman A, et al. Spatial transformer networks. In *Advances in Neural Information Processing Systems*. 2015; 2017–2025.
- [11] Zhu Y, Zhou Sr Z, Liao Sr G, Yuan K. New loss functions for medical image registration based on Voxel-morph. In *Medical Imaging 2020: Image Processing*, volume 11313. International Society for Optics and Photonics, 2020; 113132E.

Address for correspondence:

Roshan Reddy Upendra. RIT - Institute Hall (73) Rm 3130, Rochester NY 14467 USA. E-mail: ru6928@rit.edu