Fully automatic extraction of mitral valve annulus motion parameters on long axis CINE CMR using deep learning

Research Internship for MSc. in Medical Engineering

Maria Monzon[1], Seung Su Yoon[1], Andreas Maier[1] , Jens Wetzl[2], Daniel Giese[2]

[1]Fakultät für Pattern Recognition, FAU Erlangen-Nürnberg
[2]Magnetic Resonance, Siemens Healthcare GmbH, Erlangen, Germany
maria.monzon@fau.de, seung.su.yoon@fau.de

Abstract. Diastolic dysfunction is an important cause of cardiac insufficiency, defined as a malfunctioning filling of the heart during diastole. The analysis of mitral valve motion is known to be relevant in the diagnosis of cardiac dysfunction. Cardiac motion parameters can be extracted from Cardiac Magnetic Resonance (CMR) images. However, in clinical setting valve motion modeling usually needs a manual intervention to localize the valvular plane. We propose two chained Convolutional Neural Networks (CNN) for automatic tracking of mitral valve-annulus landmarks on time-resolved 2-chamber and 4-chamber CMR images. The first CNN is trained to detect the region of interest and the second to track the landmarks along the cardiac cycle. The presented deep learning system has high accuracy in terms of temporal landmark tracking and motion assessment. Furthermore, we successfully extracted several motion-related parameters thereby overcoming time-consuming annotation and allowing statistical analysis over a large number of datasets.

1 Introduction

Diastolic dysfunction is an important malfunctioning filling of the heart during diastole[1], potentially leading to premature cardiovascular death if not diagnosed at an early stage. Left ventricular diastolic dysfunction (DD) is estimated to affect from 27% to 43% in middle-aged adults[2]. Its diagnosis remains challenging as both cardiac functional and structural abnormalities need to be evaluated[3].

In clinical setting, DD diagnosis is derived from early diastolic mitral annular velocity (e') in echocardiography. Equivalent measurements can be obtained from CMR long-axis cine images, which represent the the rate of change in the left ventricle[4]. It is known that the analysis of the mitral valve annulus (MVA) throughout the cardiac cycle might act, amongst others, as a predictor for DD. Furthermore, CMR valvular flow imaging acquisition is specially challenging[5],

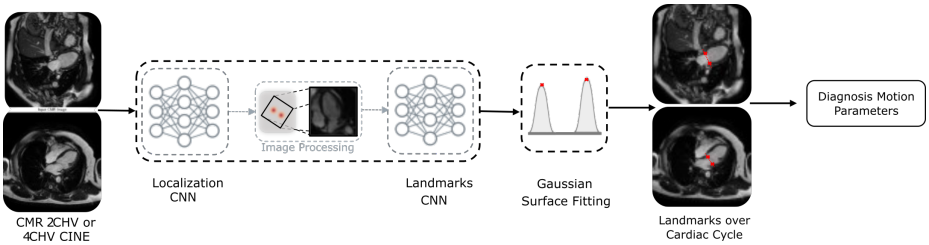


Fig. 1: Proposed system outline diagram. The long-axis Cardiac Magnetic Resonance (CMR) sequences are forwarded to chained Convolutional Neural Network (CNN) system. The first CNN localize the region of interest (ROI). The second CNN regress the time-resolved mitral valve annulus landmarks from Gaussian heatmaps and refifned by a surface post-processing step. Finally, the motion through the cardiac cycle is derived as the mitral atrio-ventricular plane displacement (MVAPD) and velocity(MVAPV), mitral valve annulus landmark displacement and diameter ratio parameters.

due to the through-plane motion of the heart valves through the cardiac cycle. In the literature, some works stating the advantages of prospectively gated slice-tracking PC-MRI can be found[6][7][8], by modelling the motion proportionally to slice offset of the valvular plane. However, prior motion assessment is clinically expensive, due to the need of a dedicated pre-scan and expert manual intervention. In everyday clinical practice, evaluation of CMR dynamics is often conducted by visual assessment and semi-automatic tools. An automatic prospective slice tracking algorithm would improve CMR valvular imaging and mitral-valve flow assessment.

In recent years, CNNs have achieved state-of-the-art performance in a variety of medical image analysis tasks. Furthermore, atrioventricular plane tracking was shown feasible using direct coordinate regression based on CNNs[9]. Nevertheless, no time-resolved feature information was extracted. In this work we present a robust, fully automated CNN system based on heatmap regression that tracks the MVA insertion points on 2-chamber (2CHV) and 4-chamber (4CHV) CINE CMR series along the cardiac cycle.

The presented work aimed to develop both a method for valve motion assessment for prospective slice tracking task in CMR image acquisition and retrospective diagnosis motion parameter extraction. The system accuracy is analyzed based on annotated data. Motion parameters extraction and system robustness are further validated retrospectively on N=1468 unlabeled datasets.

2 Materials and Methods

In this section, we introduce the details about the dataset and the motion assessment system. The proposed landmark tracking pipeline (Fig. 1) is composed of five main stages: region localization network, image processing step, temporal landmark tracking network, surface fitting post processing and automatic

Table 1: Cardiac Atlas Challenge Dataset and unlabeld test dataset details

Datasets	**Number**	**Field Strength**	**Pixel Spacing**	**Manufacturer**	
Train	59	1.5T (98.3%) 3.0T (1.7%)	1.55 ± 0.26 $\frac{\text{mm}}{\text{pixel}}$	GE Healthcare Philips Medical Systems Siemens Healthcare	17% 29% 54%
Validation	12	1.5T	1.12 ± 0.37 $\frac{\text{mm}}{\text{pixel}}$	GE Healthcare Philips Medical Systems Siemens Healthcare	92% 8% 0%
Test	12	1.5T	1.25 ± 0.26 $\frac{\text{mm}}{\text{pixel}}$	GE Healthcare Philips Medical Systems Siemens Healthcare	17% 50% 33%
Unlabeled	1468	1.5T	1.82 ± 0.01 $\frac{\text{mm}}{\text{pixel}}$	Siemens Healthcare	100%

parameter extraction motion-related parameters including displacements, velocities (e') and diameters are described.

2.1 Data

The dataset used in this work, images from 83 subjects were provided from the Cardiac Atlas Project landmark detection challenge[10]. The multi-vendor dataset (Tab. 1) is composed of 1.5T and 3T 2 chamber-view (2CHV) and four chamber-view (4CHV) CINE CMR series, with a mean in-plane resolution 1.48 ± 0.37 mm. All patients are accompanied by the MVA landmark coordinates annotated by an experienced analyst and one frame and semi-automatically annotated throughout the cardiac cycle.

All CINE series were temporally interpolated to 32 time-frames, flipped to show the apex in an upwards orientation and were normalized to 0-1 range values. Finally, the dataset split into 70% training, 15% validation and 15% testing data subsets.

For robustness qualitative validation, 1468 unlabeled CINE 4CHV and 2CHV datasets[11] (this research has been conducted using the UK Biobank Resource, access application 30769) acquired on 1.5T systems (MAGNETOM, Siemens Healthcare, Erlangen, Germany) was used.

2.2 Network Architecture

Our approach for temporal landmark localization is based on fully CNN that regress time-resolved heatmaps, which represent the landmark pseudo-probability of a being at a certain pixel over the cardiac cycle. Regressing heatmaps leads to a faster convergence and reduction of outliers compared to direct landmark coordinate regression[12].

The proposed CNN system resembles the human behaviour. First attempts to find a target structure visually, by detection the region of interest (ROI)

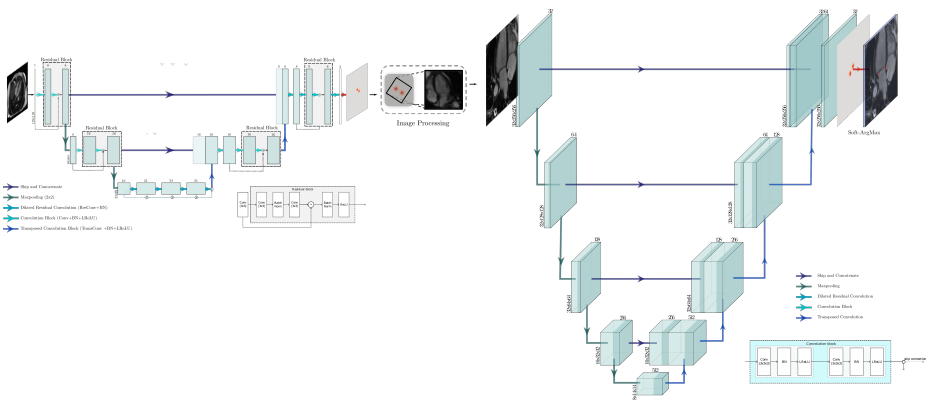


Fig. 2: 2-D Residual and 3-D Unet CNN network architectures and respective feature extraction block diagrams.

around the MVA insertion points. Subsequently the second CNN identify the target motion, as time-resolved MVA landmark tracking.

The first localization network, is Residual 2-D UNet[**?**] model with 3 decoder-encoder blocks and dilated convolution bottleneck, with a total of 49627 parameters. The feature extraction relies on spatial convolution (3x3) and followed by Batch Normalization[13] (BN) and Leaky Rectified Linear Units[14] (LReLU) activation layer.

As intermediate processing step, the images were rotated to have the same alignment based on the inferred landmarks and interpolated to a common 0.5 mm pixel-spacing.

The second 3-D UNet[15] network is used to track the valves along the complete cardiac cycle, composed of four encoder and decoder blocks. The convolutional blocks corresponds to the composite sequence of double Convolution (3x3x3)-BN-LReLU operations, which were used to extract both spatial and temporal features. For down-sampling asymmetrical max-pooling layers were applied into temporal and spatial dimensions (Fig. 2).

2.3 Training and inference procedure

The proposed framework was trained using Pytorch framework [16] using GPU-acceleration, on 2 Nvidia Testla V100 GPUs. Training was carried out for a fixed number of 200 epochs as a stopping criteria. The networks were trained using Adam optimizer[17] with momentum of $\beta = 0.9$ and learning rate $\lambda = 0.0001$ with weight decay regularization. The optimal learning rate was automatically tuned by a learning rate finder.

The presented model networks are trained from scratch using the Adaptive Wing Loss[18] between synthetic ground truth heatmaps and predicted output. The ground truth heatmaps were created for each training sample by rendering 2-D Gaussian blob around the landmark annotation stack along a separate image

channel for each landmark. The heatmap pixels near the landmark coordinate have values nearly to one, according to Eq. 1,

$$G_\sigma(x,y) = e^{-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2}\right)} \tag{1}$$

where x_0, y_0 represent the originalcentroid cordinate and σ the Gaussian standar deviation. The standard deviation is an hyper-parameter that needs to be fixed. Furthermore, following a similar approach to Teixera et al.[19], the heatmaps' standard deviation σ_x, σ_y is exponentially decreased over the training epochs:

$$\sigma_t = \sigma_0 \cdot 0.95^{N_e} \tag{2}$$

where σ_0 represents the initial standard deviation, N_e is the number of epochs.

In order to prevent overfitting due to the reduced dataset and improve the robustness of our approach, online data augmentation was performed. Based on our experiments, we included transforms in forms of shift, center cropping, rotation, Gaussian noise addition, contrast enhancement and Gaussian blurring and max-clipping to increase the variability of the dataset.

Although heatmap regression based on CNN is the current state-of-the-art method in landmark detection, it often suffers from quantization error. The maxima response of the centroid leads to inaccuracies due to the low resolution of the heatmap[20].

Therefore, in the inference phase, an additional iterative non-least-square Gaussian surface fitting algorithm was implemented in order to refine the centroid coordinate extraction (see Fig. 3). For two heatmaps $H(x,y)$ and $\hat{H}_k(x,y)$ it is defined as:

$$GS(d_x,d_y) = \sum_{i=-n_x}^{n_x} \sum_{j=-n_y}^{n_y} \Big(H(x+i,\, y+j) - \hat{H}_k(x+i-d_x,\, y+j-d_y)\Big)^2 \tag{3}$$

where d_1, d_2 are the optimization error difference and n_x, n_y the optimizing window. Although the quantitative improvement for getting subpixel maxima can be neglected, smoothness over time is crucial for slice tracking application.

2.4 Parameter Extraction

Several time-resolved motion curves were derived from the MVA landmark output (Fig 3), in order to asses the motion of the mitral valve:

- MVA plane displacement[8] (MVAPD) curve was defined as the time-resolved perpendicular distance of the MVA plane relative to the first acquisition cardiac phase.
- MVA plane velocity[8] (MVAPV) was derived as the MVAPD time-resolved discrete temporal derivate.
- The septal and lateral MVA landmark velocity curves[21] (SMVAV, LMVAV) were computed as the temporal derivative of each landmark displacement.

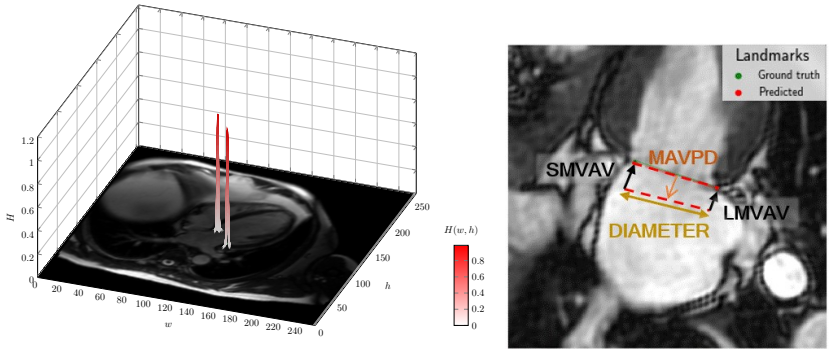


Fig. 3: 3-D representation of Heatmaps placed on landmark position (left) and representation of the MVA plane (in red) in systole end diastole (right) with representation of different MVA parameter vectors: MAVPD, SMVAV, LMVAV.

– The time-resolved diameter evolution throughout the cardiac cycle was derived as the variation of euclidean distance between both landmarks in mm.

Thereafter, the following motion diagnosis parameters are extracted from the time-curves:

1. MAV peak displacement (MVAPD-PD) was defined as the minimum point of MVAPD curve.
2. Early diastolic velocity (MVAPV-e') was computed as the central maximum of the MVAPV curve.
3. The total motion of the annulus (VAD) was quantified as the total displacement sum over all time-frames in mm.
4. The central maximum of each annular velocity curve (SMVAV and LMVAV) represents early annular diastolic velocity (MAVL-e') parameters.
5. Maximum diameter (MAMD) over cardiac cycle and the difference between maximum and minimum diameter (MADC) were extracted as contraction parameters.

3 Results

Network accuracy was evaluated by the root mean square difference on the test data subset composed of 13 CINE series. Although the training evaluation was computed on heatmaps, in testing only the landmark coordinates values were validated. For each test dataset, the mean for each time series of the mean over time of Root Mean Square Error (RMSE) loss for the $C = 2$ coordinates distance along temporal dimension $T = 32$ is computed:

$$E(l, \hat{l}) = \frac{1}{2} \sum_{c=1}^{2} \left(\frac{1}{32} \sum_{t=1}^{32} \sqrt{(l_x - \hat{l}_x)^2 + (l_y - \hat{l}_y)^2} \right) \tag{4}$$

Fig. 4: Mean euclidean distance for each test dataset predicted with the presented model

Bland-Altman Analysis		
Parameter	Value ± std	unit
MVAPD-PD	0.53 ± 2.44	mm
MVAPV-e'	0.08 ± 3.48	cm/s
VAD	15.39 ± 54.62	mm
MAVL-e'	0.12 ± 3.73	cm/s
MAMD	0.31 ± 3.66	mm
MADC	0.28 ± 3.12	mm

Table 2: Bland-Altman analysis of predicted motion parameters against ground-truth landmarks

where $l, \hat{l}$ represent the ground truth and predicted landmark coordinates respectively. Landmark coordinate mean errors of 1.75 ± 0.64 mm (2CHV) and 1.74 ± 0.72 mm (4CHV) were achieved. The mean error for each test CINE series is depicted in Fig.4. Further Bland-Altman analysis on extracted motion parameters (Sec.2.4) between ground truth and detected was also performed, summarized in Table 2.

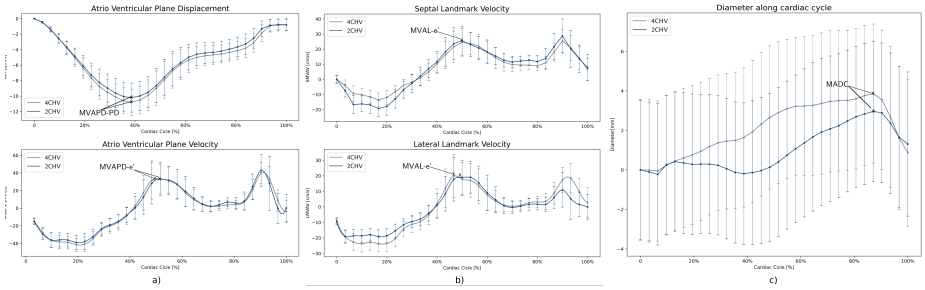
On 1468 unlabeled datasets successful inference was assessed by detecting unambiguous outliers. Every tracked series whose plane displacement is not temporally smooth (based on mean standard deviation) at any cardiac phase is discarded. The localization network fails to locate the ROI in less than 0.5% of unlabeled datasets and at least one time-frame was not smoothly tracked in 16,53% of the datasets. Finally, motion parameters were extracted from this data. The median modeled timecurves are shown in Fig. 5.

3.1 Discussion

The proposed system was shown to successfully track MVA landmarks along the cardiac cycle. 3-D convolution blocks are able to predicted temporal smoother landmarks with high accuracy although a comparatively small annotated dataset size (N=83) employed. This may be explained in heatmap regression technique which represents the probability of the landmark motion over the cardiac cycle. Furthermore, Heatmap regression avoids the need to learn the highly non-linear domain transfer, from pixel to coordinate space. A post-processing step with the non-linear least square fitting is able to refine the landmark location. The additional step outputs subpixel maxima and therefore predicts a smooth motion.

Extraction of derived parameters of interest was successful and showed good agreement with ground-truth data, based on Bland-Altman analysis.

Fig. 5: Extracted parameters from the N=1468 unlabeled datasets, represented in blue and gray for 2CHV and 4CHV respectively. The error bars represent the standard deviation in each plot : a) MAVPD derived MVAPV. b) Mean SMVAV and LMVAV curves. c) Mean diameter time-resolve difference and indication of MADC parameter



Qualitative results in an extensive dataset indicates that the network is generalizable to other acquisiton and multi-vendor datasets, despite the limited sized subsets employed for training.

4 Conclusion

In this work, we propose a cascaded convolutional neural network for automatic landmark tracking from 2CHV and 4CHV time-resolved cardiac MRI. The network architecture consists of a 2-D UNet to focus on the mitral valve annulus ROI to perform an alignment preprocessing and a 3-D UNet to incorporate temporal information . Our method is trained on CINE series from 83 patients extracted from the Cardiac Atlas Project Landmark Detection challenge dataset. The proposed system was shown to successfully track MVA landmarks with a mean error mean error of 1.75 ± 0.68 mm average RMSE for the testing stage, in the range of the data's resolution.

Future work will aim to extensively analysis motion parameters on larger numbers of categorized patient datasets and clinical validation. Furthermore, online integration of the approach in MRI scanner for automatic prospective slice tracking algorithm acquisition is planned, for improved morphological and/or flow measurements of the mitral valve.

References

1. Kossaify A, Nasr M. Diastolic Dysfunction and the New Recommendations for Echocardiographic Assessment of Left Ventricular Diastolic Function: Summary of Guidelines and Novelties in Diagnosis and Grading. Journal of Diagnostic Medical Sonography. 2019;35(4):317–325.

2. Nayor M, Cooper LL, Enserro DM, et al. Left ventricular diastolic dysfunction in the community: Impact of diagnostic criteria on the burden, correlates, and prognosis. Journal of the American Heart Association. 2018;7(11).
3. Gutierrez C, Blanchard DG. Diastolic heart failure: Challenges of diagnosis and treatment. American Family Physician. 2004;69(11):2609–2616.
4. Azarisman SM, Teo KS, Worthley MI, et al. Cardiac magnetic resonance assessment of diastolic dysfunction in acute coronary syndrome. Journal of International Medical Research. 2017;45(6):1680–1692. PMID: 29239257. Available from: https://doi.org/10.1177/030006051769826.
5. Menchón-Lara RM, Simmross-Wattenberg F, Casaseca-de-la Higuera P, et al. Reconstruction techniques for cardiac cine MRI. Insights into Imaging. 2019;10(1).
6. Kozerke S, Scheidegger MB, Pedersen EM, et al. Heart motion adapted cine phase-contrast flow measurements through the aortic valve. Magnetic Resonance in Medicine. 1999;42(5):970–978.
7. Kozerke S, Schwitter J, Pedersen EM, et al. Aortic and mitral regurgitation: Quantification using moving slice velocity mapping. Journal of Magnetic Resonance Imaging. 2001;14(2):106–112.
8. Seemann F, Pahlm U, Steding-Ehrenborg K, et al. Time-resolved tracking of the atrioventricular plane displacement in Cardiovascular Magnetic Resonance (CMR) images. BMC Medical Imaging. 2017;17.
9. Gonzales RA, Onofrey J, Lamy J, et al.. Time-resolved tracking of the atrioventricular plane displacement in long-axis cine images with residual neural networks; 2020. Available from: http://archive.ismrm.org/2020/2233.html.
10. Fonseca CG, Backhaus M, Bluemke DA, et al. The cardiac Atlas Project-an imaging database for computational modeling and statistical atlases of the heart. Bioinformatics. 2011;27(16):2288–2295.
11. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLoS Medicine, 12, 3. 2015;.
12. Tompson J, Jain A, LeCun Y, et al. Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation. CoRR. 2014;abs/1406.2984. Available from: http://arxiv.org/abs/1406.2984.
13. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. CoRR. 2015;abs/1502.03167. Available from: http://arxiv.org/abs/1502.03167.
14. Maas AL. Rectifier Non-linearities Improve Neural Network Acoustic Models; 2013. .
15. Çiçek Ö, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. CoRR. 2016;abs/1606.06650. Available from: http://arxiv.org/abs/1606.06650.
16. Paszke A, Gross S, Massa F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Wallach H, Larochelle H, Beygelzimer A, et al., editors. Advances in Neural Information Processing Systems 32. Curran Associates, Inc.; 2019. p. 8024–8035. Available from: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.
17. Kingma D, Ba J. Adam: A Method for Stochastic Optimization. International Conference on Learning Representations. 2014 12;.
18. Wang X, Bo L, Li F. Adaptive Wing Loss for Robust Face Alignment via Heatmap Regression. CoRR. 2019;abs/1904.07399. Available from: http://arxiv.org/abs/1904.07399.

19. Teixeira B, Tamersoy B, Singh V, et al. Adaloss: Adaptive Loss Function for Landmark Localization. ArXiv. 2019;abs/1908.01070.
20. Nibali A, He Z, Morgan S, et al.. Numerical Coordinate Regression with Convolutional Neural Networks; 2018.
21. Thavendiranathan P, Guetter C, Serafim da Silveira J, et al. Mitral annular velocity measurement with cardiac magnetic resonance imaging using a novel annular tracking algorithm: Validation against echocardiography. Magnetic Resonance Imaging. 2018 08;55.