



Research article

Fully automated AI-based cardiac motion parameter extraction – application to mitral and tricuspid valves on long-axis cine MR images



Seung Su Yoon ^{a,b,*}, Carola Fischer ^{b,c,1}, Daniel Amsel ^a, Maria Monzon ^a, Solenn Toupin ^d, Théo Pezel ^{e,f}, Jérôme Garot ^e, Jens Wetzl ^b, Andreas Maier ^{a,b}, Daniel Giese ^{a,b}

^a Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

^b Magnetic Resonance, Siemens Healthcare GmbH, Erlangen, Germany

^c Technische Universität Berlin, Germany

^d Siemens Healthcare France, Saint-Denis, France

^e Institut Cardiovasculaire Paris Sud, Cardiovascular Magnetic Resonance Laboratory, Hôpital Privé Jacques CARTIER, Ramsay Santé, Massy, France

^f Université de Paris Cité, Service de Cardiologie, Hôpital Lariboisière – APHP, Inserm UMRS 942, 75010 Paris, France

ARTICLE INFO

Keywords:

CMR

AI

Mitral valve motion

Tricuspid valve motion

Longitudinal shortening

ABSTRACT

Purpose: In cardiac MRI, valve motion parameters can be useful for the diagnosis of cardiac dysfunction. In this study, a fully automated AI-based valve tracking system was developed and evaluated on 2- or 4-chamber view cine series on a large cardiac MR dataset. Automatically derived motion parameters include atrioventricular plane displacement (AVPD), velocities (AVPV), mitral or tricuspid annular systolic excursion (MAPSE, TAPSE), or longitudinal shortening (LS).

Method: Two sequential neural networks with an intermediate processing step are applied to localize the target and track the landmarks throughout the cardiac cycle. Initially, a localisation network is used to perform heatmap regression of the target landmarks, such as mitral, tricuspid valve annulus as well as apex points. Then, a registration network is applied to track these landmarks using deformation fields. Based on these outputs, motion parameters were derived.

Results: The accuracy of the system resulted in deviations of 1.44 ± 1.32 mm, 1.51 ± 1.46 cm/s, 2.21 ± 1.81 mm, 2.40 ± 1.97 mm, 2.50 ± 2.06 mm for AVPD, AVPV, MAPSE, TAPSE and LS, respectively. Application on a large patient database ($N = 5289$) revealed a mean MAPSE and LS of 9.5 ± 3.0 mm and 15.9 ± 3.9 % on 2-chamber and 4-chamber views, respectively. A mean TAPSE and LS of 13.4 ± 4.7 mm and 21.4 ± 6.9 % was measured.

Conclusion: The results demonstrate the versatility of the proposed system for automatic extraction of various valve-related motion parameters.

1. Introduction

Cardiac magnetic resonance imaging (CMR) is a versatile imaging modality that enables the assessment of both cardiac morphology and cardiac function. For time-resolved assessment of ventricular function, so-called cine images are typically acquired in several cardiac views including short axis, 2-chamber, and 4-chamber views. Clinically relevant parameters such as left ventricular (LV) and right ventricular (RV) ejection fractions can be extracted by segmenting the ventricles in diastole and systole. This segmentation task is often semi-automated using machine learning and/or deep learning approaches [1–4].

Nevertheless, cine images are known to provide much more information of potential diagnostic relevance, such as myocardial strain information [4] and motion-related parameters of the mitral and tricuspid valves (MV and TV, respectively). However, the wide-spread use of these parameters in clinical practice often remains hampered by the time and effort needed to locate and track these features.

MV plane motion provides structural and functional systolic and diastolic motion and has been shown to play an important role in the diagnosis of diastolic dysfunction [5]. Key metrics like mitral or tricuspid annular plane systolic excursion (MAPSE and TAPSE) or longitudinal shortening (LS) are routinely obtained in echocardiography

* Corresponding author at: Siemens Healthcare GmbH, Allee am Roethelheimpark 2, 91052 Erlangen, Germany.

E-mail address: seungsu.yoon@siemens-healthineers.com (S.S. Yoon).

¹ equal contribution.

[6–9]. Although their assessment using CMR would provide comparable diagnostic value [5,10], their wide-spread use remains limited for the above-mentioned reason.

Since manual tracking of mitral valve insertion points would be tedious and unpractical in clinical workflow, several semi-automated [11,12] and automated tracking algorithms [13–15] have been proposed over the recent years. Especially, the development of deep learning-based detection methods facilitates fast processing and can additionally overcome observer inconsistencies [14]. To enable a highly standardized detection task on diverse datasets, so-called dual-stage techniques have been proposed for segmentation or landmark detection tasks [14–17]. They typically consist of two stacked networks, one for the localisation of the area of interest and the second one for the detection of landmarks. For MV and TV annulus landmark detection, recently proposed networks are designed as 2D landmark coordinate regression tasks. These 2D convolutional based regression networks are applied on time-resolved images frame-by-frame and therefore do not exploit the coordinates' temporal correlation through explicit learning of temporal features. Further, as with other methods [18,19], a direct regression of pixel coordinates implies a highly nonlinear mapping from input images to pixel coordinates.

In this work, an AI-based method for tracking landmarks from long-axis cine images is proposed. Exploiting the dual-stage network approach, a localisation and a registration network are combined. In contrast to a coordinate detection task, an image-to-image mapping is developed that performs regression of heatmap representations rather than performing coordinate regression [13,20,21]. Such heatmaps encode the probability that a landmark is located at a specific pixel position [20,21]. This way, the network learns to generate high responses to positions near the target landmarks, while suppressing incorrect positions. While the localisation network [22] takes advantage of heatmap regression for target detection, the registration network utilises both spatial and temporal information for landmark tracking. Due to the versatile nature of the registration network [23], the network is readily applied to the automatic detection of MV, TV annulus as well as LV and RV apex points, without needing re-training. Based on the landmarks, several derived motion parameters including the atrioventricular plane displacements (AVPD), the atrioventricular plane velocities (AVPV), MAPSE, TAPSE, and LS are then automatically extracted. The performance of the method was evaluated on 245 semi-automatically annotated multi-center and multi-field-strength ground truth datasets. The system was then applied to retrospectively extract and analyse motion parameters from data acquired on 5289 patients.

2. Materials and methods

2.1. Data

Table 1 summarizes the data used in this work and [Supplementary Figure 1](#) shows the diagram of all datasets. All datasets used in this study are cine datasets acquired by retrospectively ECG-gated fast balanced steady-state free precession sequence with parallel imaging techniques.

The dataset A - CAP is composed of 2-chamber and 4-chamber view cine series from 83 patients originating from the Cardiac Atlas Project (CAP) [dataset] [24]. Data was acquired on 1.5 T and 3 T scanners from different vendors (GE Healthcare, Waukesha, Wisconsin, USA, Philips Healthcare, Best, The Netherlands and Siemens Healthcare, Erlangen, Germany) at two sites. Additionally, these datasets contain ground truth annotated landmarks of the MV and TV insertion points. This dataset was used for training the networks.

The Local ethics committee and IRB approvals were obtained for this dataset at the two host institutions. and a written informed consent was provided by all participants. The declaration of consent to participate in the study was signed by all participants.

The dataset B - ICPS, a subset ($n = 5289$) of the patient database acquired on a 1.5 T scanner (MAGNETOM Aera, Siemens Healthcare,

Table 1

Information about the population and acquisition parameters of different datasets used for training and analysis.

	A - CAP	B - ICPS	C - Multicenter
Number of patients	83	5289	169
Number of datasets (2-chamber view, 4-chamber view)	163	10,948	245
Age	61.8 ± 19.6	64.2 ± 12.2	61.2 ± 18.0
Gender	75% male	67.4% male	67% male
Heart rate [bpm]	62.7 ± 8.2	68.6 ± 14.2	67.8 ± 13.2
Field strength [T]	98% 1.5, 2% 3	100% 1.5	87% 1.5, 13% 3
Slice thickness	7.4 ± 1.2	6.0 ± 0.1	6.0 ± 0.0
Spatial resolution [mm ²]	1.5 ± 0.3	1.3 ± 0.1	1.4 ± 0.2
Temporal resolution [ms]	41.8 ± 9.9	36.5 ± 7.7	35.5 ± 9.0
Field of view [mm × mm]	353.3 ± 47.2 357.8 ± 42.1	341.4 ± 36.2 347.3 ± 28.9	338.4 ± 38.0 355.6 ± 23.2
Number of phases	23.2 ± 3.7	24.8 ± 1.8	25.2 ± 3.1

Erlangen, Germany) from January 2016 to September 2017 at a single site (ICPS, Massy, France) [dataset] [25] were used in this study. All patients were consecutively referred to a clinical vasodilator stress CMR examination. All cine series were acquired during rest. This dataset did not contain ground-truth annotations and used for inference of the network with subsequent analysis of the extracted parameters.

For this dataset, the study was assessed and approved by the local committee of the ICPS and was conducted in compliance with the 1984 Declaration of Helsinki. All participants were required to understand and give informed consent to participate.

The dataset C - Multicenter, a total of 169 patients, acquired on 1.5 T and 3 T scanners (MAGNETOM Aera and Prisma, Siemens Healthcare, Erlangen, Germany) consisted of 20 volunteers from the UK Biobank [dataset] [26], 20 volunteers acquired in-house and 205 datasets from the dataset B. In order to efficiently annotate all 245 datasets, two MV, TV insertion points and the apex points were semi-automatically annotated on every 4th frame. These points were then automatically propagated using spline interpolation. Following interpolation, a manual check by a trained cardiac MRI expert with more than 20 years of experience over all phases was performed and the landmarks were corrected, when needed. To allow quantification of inter-observer variability, 90 of these datasets were additionally annotated by a second expert (>3 years of cardiac MRI experience). Dataset C was used to analyse the accuracy of the method and compare it with expert annotations.

The local ethics committee of each center for the dataset C approved the study, and a written informed consent was provided by all subjects. The declaration of consent to participate in the study was signed by all subjects.

2.2. Network overview

The proposed dual-stage network is illustrated in [Fig. 1](#). It consists of a stepwise execution of two convolutional neural networks for localising and tracking pre-defined landmarks on 2-chamber or 4-chamber view cine series. First, the full-resolution original image is input to a 2D-Unet (Localisation network), which regresses heatmaps of the landmarks of interest (MV, TV insertion points and apex). Following localisation, the output is used to re-orient and crop all frames of the cine series to standardise the input to the second network (Registration network). The registration network is based on the VoxelMorph framework [27] and extracts the time-resolved deformation fields. These deformation fields are then used to track the landmarks over time. A final back transformation is used to map the tracked landmarks onto the original input images. The datasets used for training and validation the networks are

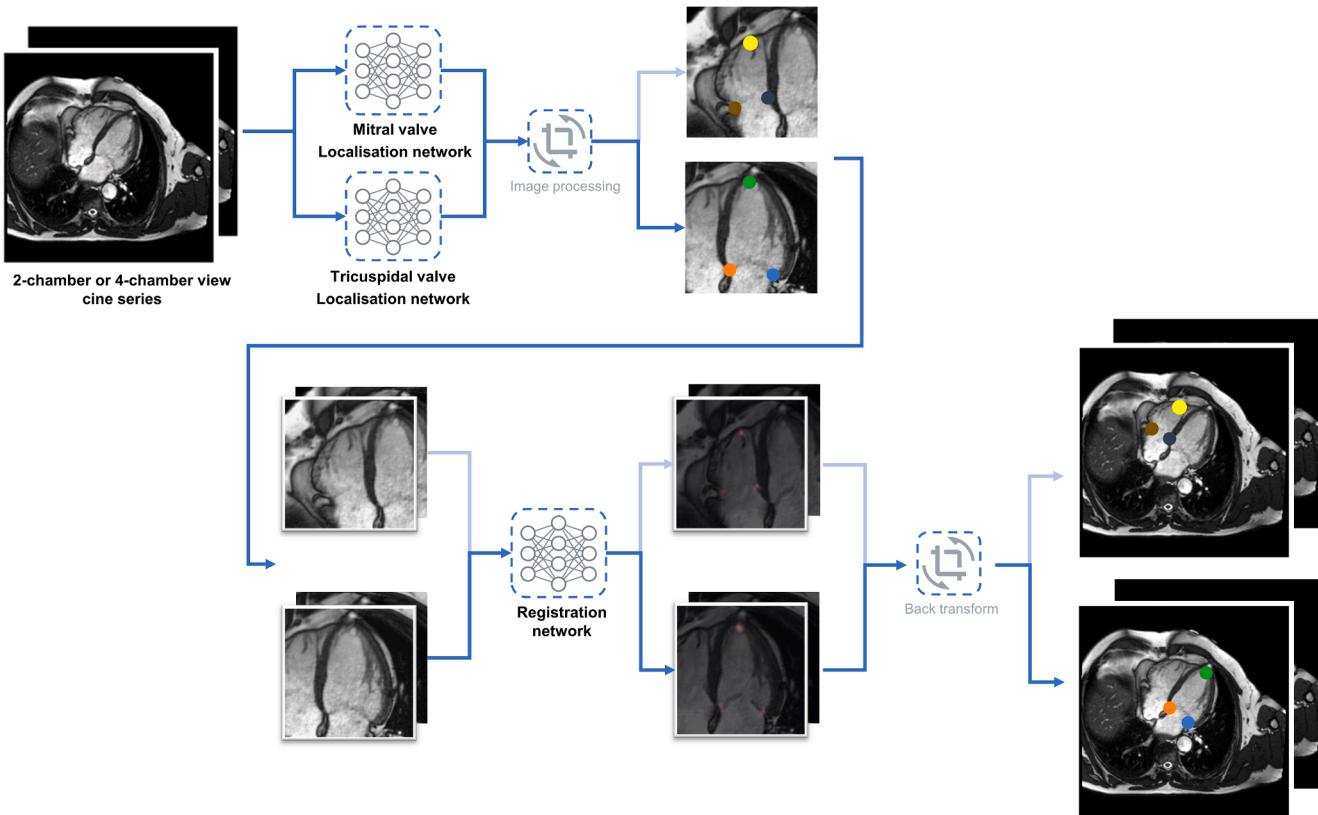


Fig. 1. Overview of the proposed system. By sequentially applying the proposed steps - localisation network, image processing, registration network and back transform - the landmarks are extracted. Blue, orange and green represent two MV insertion points and LV apex. Brown, black and yellow represents two TV insertion points and RV apex. An example of 4-chamber view cine is shown here, a 2-chamber view cine input excludes the TV localisation network.

split by 85% and 15%, respectively. For the hyperparameter tuning, the open-source framework, Optuna [28] was used for parameters, such as activation functions (ReLU, LeakyReLU), normalization layers (batch, instance normalization layer), number of output feature maps for the beginning convolution layer (16, 32, 64) and the learning rate (10^{-3} – 10^{-5}).

2.3. Localisation network

The input to the 2D U-Net is the first frame of a 2-chamber or 4-chamber view cine series, preprocessed to a fixed resolution of 1 mm^2 and center cropped to a fixed size of 384×384 pixels. The network is built using an encoder-decoder-based CNN architecture. An asymmetrically larger encoder is used to extract image features, and a smaller decoder is used to reconstruct the heatmap.

The encoder part consists of 9 ResNet blocks each consisting of two convolutions with batch normalization and ReLU activation, followed by additive skip connection. The downsampling is performed by strided convolutions by a factor of 2 while increasing the feature size by 2. Initial convolution is set to 16 filters with 3×3 filter size. Further convolutions have a size of 3×3 and the size of the final convolution of the encoder is 8 times smaller than the input. The decoder is built with a single block per each upsample level. Each block is used for upsampling, where the number of features is reduced by a factor of 2 using 1×1 convolutions and the spatial dimension is increased by using pixel shuffling, followed by summation of the encoder output of the identical level. In the final decoder part, the spatial dimension of the output is equal to the input image size and the number of features to the initial input feature size. A 1×1 convolution and a softmax activation function then splits the output into three channels. The loss function for training the network was a L_2 loss and an Adam optimizer with initial learning

rate of 10^{-4} and a batch size of 1 was used. The output represents either two MV insertion and the LV apex points, or two TV insertion and the RV apex points. A detailed architecture of the network is visualized in [Supplementary Data Figure 2](#).

2.4. Image processing

Based on the localised landmarks, the time-resolved cine series are rotated to horizontally align the MV or TV insertion points and the apex point pointing upwards. Furthermore, the series is centre-cropped to 256 cm^2 FOV and interpolated to 0.5 mm^2 resolution. In [Fig. 1](#), an exemplary case of a 4-chamber view is visualized after the localisation network and the image processing pipeline.

2.5. Registration network

The registration network takes the processed series as input and is trained to explicitly learn the temporal information. The method is built on the VoxelMorph framework consisting of a UNet, scaling and squaring layers, and a spatial transformer layer [23,27]. The network takes a pair of input images, a fixed reference image, and a moving image and computes the moved image, as well as the deformation field. Since the cine is time-resolved, the reference input is fixed to the first time-frame and the moving image corresponds to the subsequent time-frame until the last time-frame is reached.

In addition to the previously proposed method [23], the heatmap produced by the localisation network is warped using the deformation field to obtain the moving heatmap. A detailed architecture of the network is visualized in [Supplementary Data Figure 3](#).

2.6. Derived motion parameters

The extracted time-resolved coordinates were used to derive the following motion parameters (Fig. 2):

The landmark excursion over time (MV/TV annulus insertion or LV apex) was calculated as the distance relative to the first phase.

The atrioventricular plane displacement (AVPD) was calculated as the distance of the plane spanned by the MV or TV annulus points relative to the first phase.

The atrioventricular plane velocity (AVPV) was defined as the discrete temporal derivate of the AVPD.

The lateral/inferior and septal/superior MAPSE/TAPSE were calculated as the maximum points' excursion relative to the first phase. The global MAPSE/TAPSE was defined as the maximum AVPD across the cardiac cycle.

The longitudinal shortening (LS) was calculated as the percentage difference of the distance between the mid-valvular point and the apex point between end-systole and end-diastole. The end-systole and end-diastole time points were determined by the minimum and maximum distances between the apex and the midpoint of the valve, respectively.

2.7. Analyses

To evaluate the performance of the proposed system, the Euclidean distance error between the predicted and the ground-truth annotations and derived motion parameters were calculated from the entire dataset C. On 90 datasets subset from dataset C, the interobserver variability was assessed by computing the distance error's mean and standard deviations.

A subset of dataset C containing 20 healthy subjects (4-chamber view only) was used to compare the extracted parameters with literature.

Landmark coordinates and motion parameters were automatically extracted from the entire dataset B. The correlations of selected parameters (AVPD, MAPSE, and TAPSE) with patient age and gender (male 67.7%) were analysed ($n = 5628$). Statistical significance between gender groups was assessed using a two-sided two-sample T-test.

Furthermore, it was investigated if each age group differs significantly from the population mean using a two-sided one-sample T-test. P-values less than 0.05 were considered statistically significant.

3. Results

3.1. Network accuracy

The performance of the network compared to annotations is shown in Fig. 3, Fig. 4 and summarized in Table 2. Fig. 3 shows the mean and standard deviation of the distance error plotted over the cardiac cycle for each chamber view and every landmark (MV, TV and apex). Similarly, Fig. 4 shows time-resolved error plots of the derived motion parameters AVPD and AVPV as well as Boxplots of the errors in MAPSE, TAPSE and LS.

Table 2 summarizes the networks performance in which all landmarks' parameters were evaluated on average. Additionally, the interobserver variability is quantified.

3.2. Motion parameters

Motion parameters extracted from dataset B are shown in Fig. 5 and Fig. 6. Time-resolved AVPD and AVPV for the MV and the TV are shown, summarizing all 10,948 datasets, previously unseen by the network. Additionally, MAPSE, TAPSE and LS are shown as histograms.

For the MV, the system calculated a mean inferior/septal, anterior/lateral MAPSE and LS of 8.5 ± 2.6 mm, 10.6 ± 3.5 mm and 16.4 ± 4.1 % on the 2-chamber view and 10.4 ± 2.9 mm, 8.3 ± 2.5 mm and 15.4 ± 3.8 % on the 4-chamber views, respectively. For the TV, the system calculated a mean lateral, septal TAPSE and LS of 17.7 ± 6.5 , 9.0 ± 2.9 mm and 21.4 ± 6.9 % on the 4-chamber views.

For 20 healthy subjects, the mean inferior/septal, anterior/lateral MAPSE and LS of 11.8 ± 3.6 mm, 12.2 ± 1.7 mm and 18.7 ± 3.6 % for MV and a mean of lateral, septal TAPSE and LS of, 20.4 ± 3.2 , 11.0 ± 1.4 mm and for 20.0 ± 2.8 % for TV.

From dataset B, gender and age grouped AVPD of MV and TV are

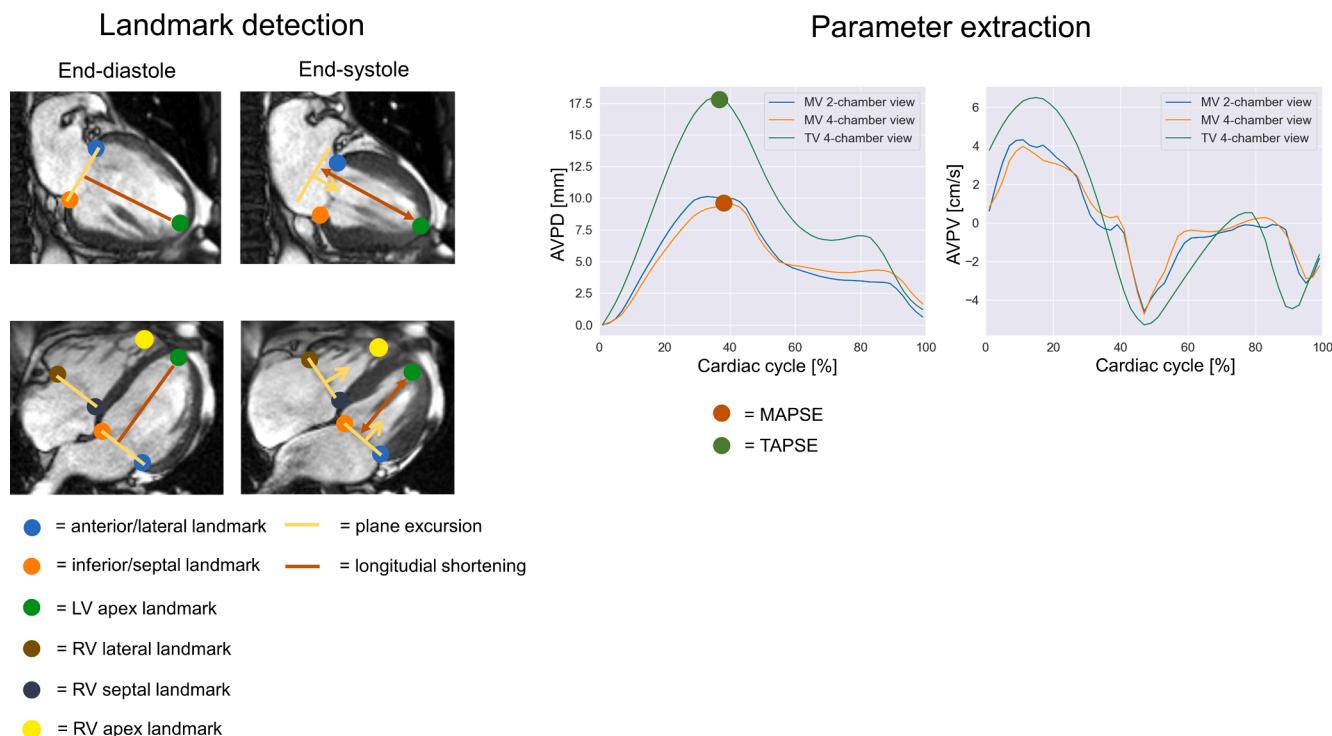


Fig. 2. Overview of the target landmarks and the parameter of interest.

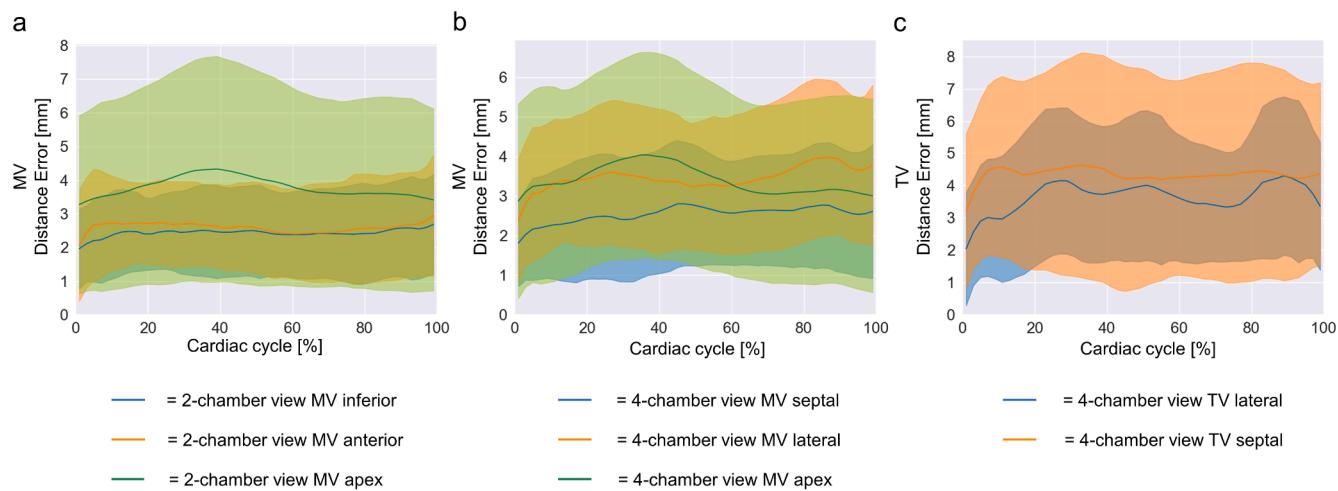


Fig. 3. The MV distance error of each landmark from 2-chamber view cine (a), the MV distance of each landmark from 4-chamber view cine (b), while on the right, the TV distance error of each landmark in 4-chamber view cine (c) are shown over the cardiac cycle, respectively.

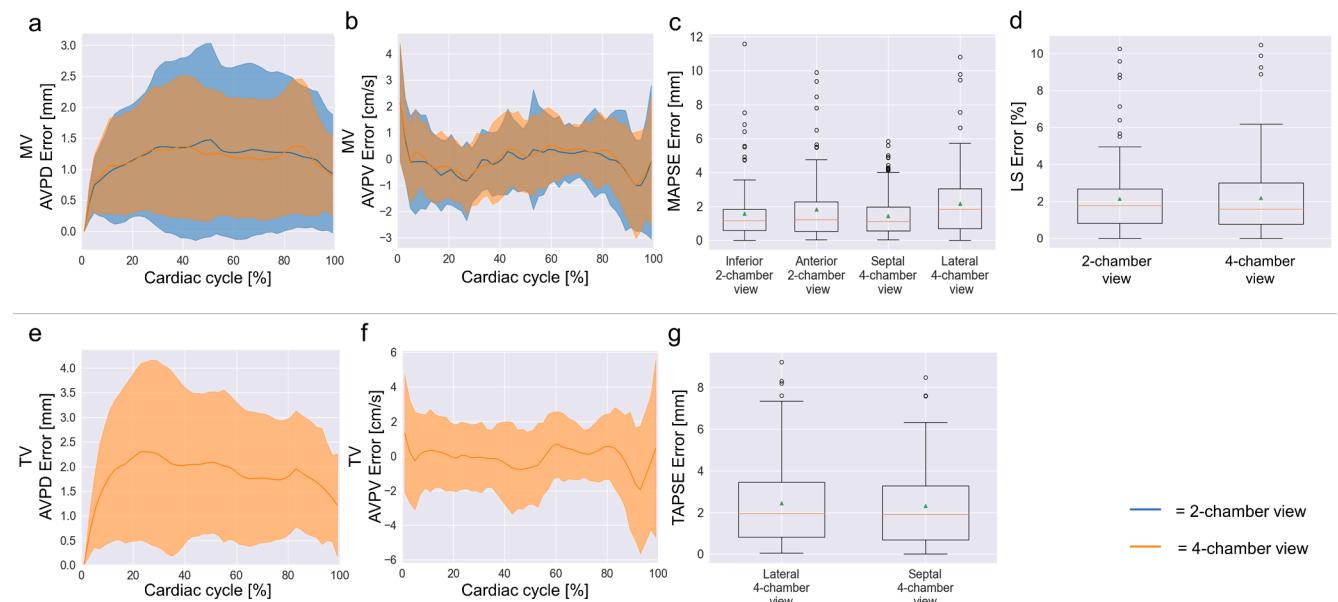


Fig. 4. The parameters are derived once from the system output and the manual annotations. The errors of individual parameter are presented in mean and standard deviation curves for AVPD (a,e) and AVPV (b,f) and in boxplots for the MAPSE (c), TAPSE (g) and LS (d) error values.

Table 2
Evaluation of the derived parameters and the interobserver variability.

	Euclidean distance error [mm]	AVPD deviation [mm]	AVPV deviation [cm/s]	MAPSE deviation [mm]	TAPSE deviation [mm]	LS deviation [%]
Dataset B, C						
2-chamber view,	3.53 ± 1.40	1.37 ± 1.40	1.38 ± 1.50	2.16 ± 1.90	-	2.55 ± 2.05
4-chamber view	3.79 ± 1.44	1.51 ± 1.23	1.63 ± 1.41	2.25 ± 1.72	2.40 ± 1.97	2.44 ± 2.06
Interobserver						
2-chamber view,	4.10 ± 1.41	1.86 ± 2.04	1.43 ± 1.68	3.12 ± 2.92	-	2.66 ± 2.51
4-chamber view	3.58 ± 1.35	1.30 ± 1.2	1.36 ± 1.50	2.10 ± 1.80	2.14 ± 1.6	2.75 ± 2.2

shown in Fig. 7. The *t*-test for AVPD, MAPSE and TAPSE revealed statistical significance in gender for all parameters on 2- and 4-chamber views.

The age group between 60 and 70 revealed no statistical significance compared to average whereelse for the other age groups, all MV-related parameters showed statistical significance compared to average on 2- and 4-chamber views.

4. Discussion

The proposed fully automated method performed robustly on large dataset in 2-chamber and 4-chamber views from long-axis cine CMR images. Mean AVPD errors of 1.37 mm and 1.51 mm for 2-chamber and 4-chamber views respectively are in-line with previously published work [13–15] and compare well with inter-observer variabilities of 1.86

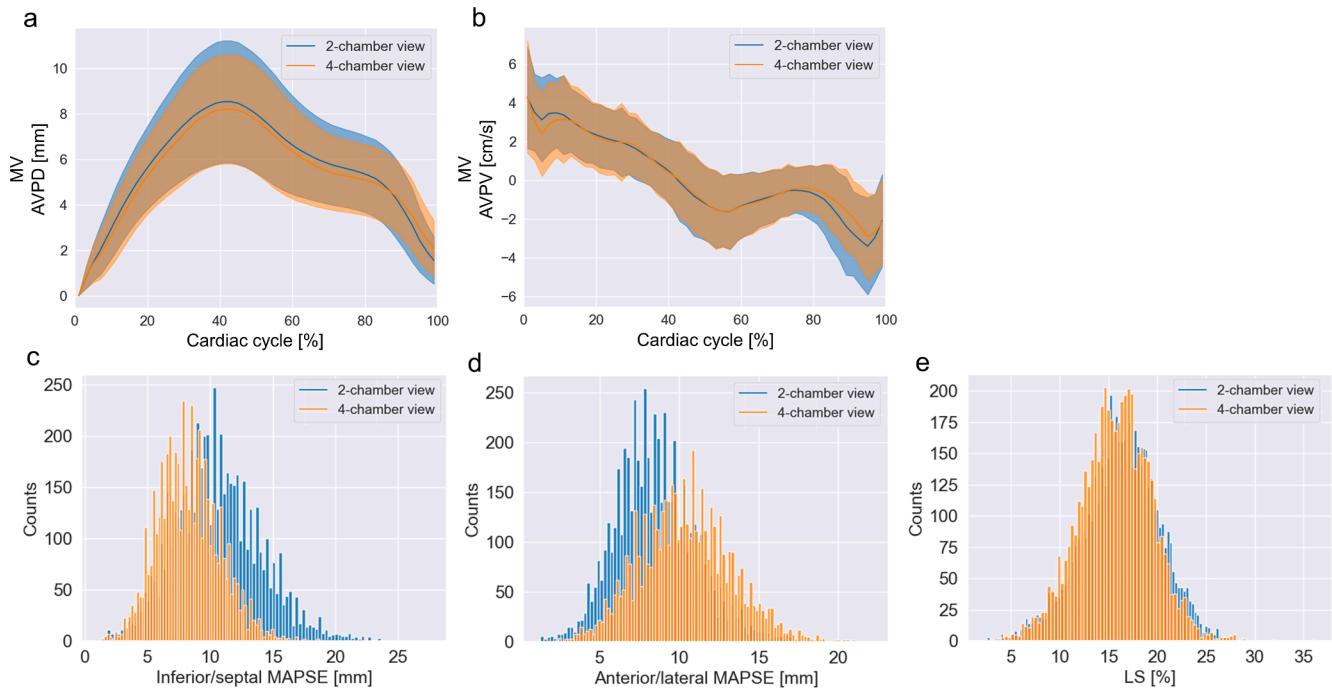


Fig. 5. Left ventricle: Mean and standard deviation of the MV AVPD (a) and AVPV (b) extracted from 10,948 unseen datasets are shown in the first row. Histograms of MAPSE (c, d) and LS (e), extracted from 2-chamber and 4-chamber views of these datasets are shown in the second row.

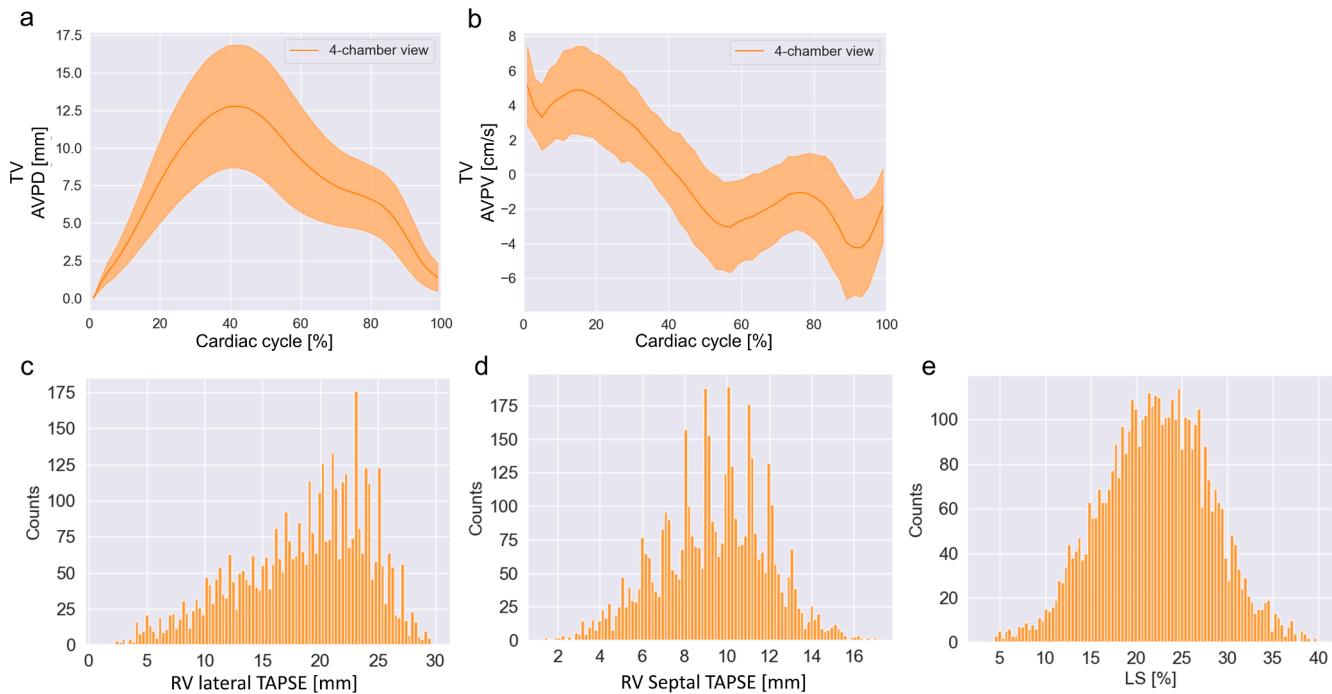


Fig. 6. Right ventricle: Mean and standard deviation of the TV AVPD (a) and AVPV (b) extracted from 10,948 unseen datasets are shown in the first row. Histograms of TAPSE (c, d) and LS (e), extracted from the 4-chamber views of these datasets are shown in the second row.

mm and 1.3 mm, respectively. The system therefore achieves human-level performance.

Compared to other proposed systems, the generalisation of the proposed network to other landmarks such as the TV was straight-forward. Only the first network of the dual-stage system was retrained for a broad localization while the second network remained unchanged, exploiting the full motion fields extracted by the registration network. This is a

major advantage over other methods, where the entire system requires to be retrained for each landmark [14,15]. Additionally, the current method was fully trained using only 163 datasets (as compared to 560 in [14] or 2329 in [13]). The proposed system also has the benefit of being able to process efficiently on GPUs. The average CPU and GPU (NVIDIA A100, memory usage was limited to 32Gb) execution time in this study was 3.23, 1.68 s, respectively. Further, the proposed system is not

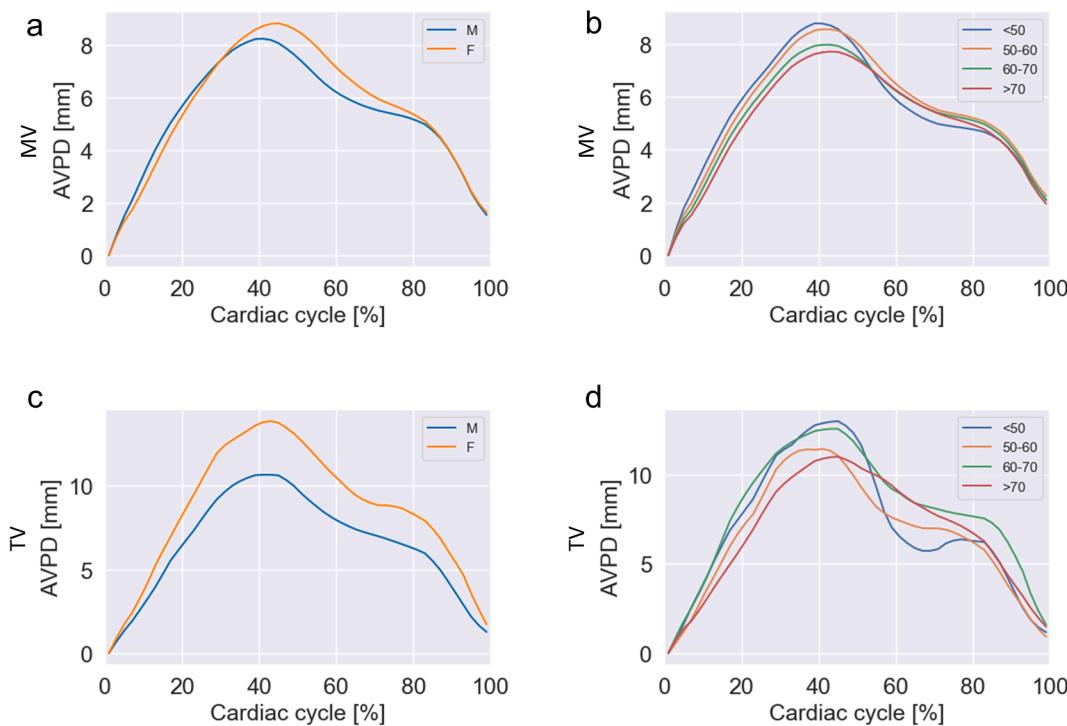


Fig. 7. Mean AVPD plots showing differences between gender and age for MV (a, b) and TV insertion points (c, d). T-test statistics are included in the text.

memory intensive as compared to 3D approach, as it has a 29-fold increase with similar accuracy when using the 3D network [23]. Nonetheless, the 3D approach for the registration can be further explored by using different architectures in future studies.

Compared with standard values for valve motion parameters, our findings agree with previously observed parameters: for instance, the AVPV clearly shows the e' and a' waves, well-known in echocardiography and representing the early and late diastolic filling. Extraction of further timing phases such as the isovolumetric relaxation times, the diastasis or the isovolumetric contraction times is straight-forward and could be of interest in the assessment of diastolic dysfunction.

Regarding plane excursion in healthy subjects, the findings in this work are in the expected range and are consistent with the fact that lateral MAPSE has a slower movement than lateral TAPSE and that plane excursion parameters decreases with age [28–30].

Similarly, the reported results regarding age and gender differences agree with previous findings [31]. Statistical significance of AVPD, MAPSE, TAPSE as a function of age was measured in all age groups except of the age group between 60 and 70 years. This is an expected finding as the average age of the used dataset, 64.2 years, falls under this category.

The analysis regarding gender showed statistically significant difference for MAPSE, TAPSE on different chamber views in this patient datasets. The AVPD are significantly lower in men than in women in our finding as shown in Fig. 7. These findings are in line with previously reported findings [32], in which the MAPSE and TAPSE were significantly lower in men than in women.

However, it is essential to note that all patients from the large dataset B were referred to a clinical vasodilator stress CMR. This dataset therefore contains patients with known or suspected ischemia and coronary artery disease with potential myocardial infarctions. As a result, a large variance in regional wall abnormalities is expected and also seen in the results (Fig. 5, Fig. 6). Therefore, a more in-depth study is needed in order to correlate and potentially predict clinical characteristics using the proposed network.

The method can be readily applied to fully automatically process and analyse large amounts of data retrospectively, potentially leading to interesting statistics when clustering and/or correlating with demographics, such as the recent performed work to manually define reference values on mitral valve diameters [33]. A correlation of these patients with clinical outcome and finding prognostic values is warranted in future work.

An in-line application of the proposed system could also aid to guide CMR measurements: for instance, the heart rest-period can be obtained or a prospective slice-following approach can be used in which the 2D slice follows the derived valve plane during the acquisition [34,35]. Finally, information on valve plane dynamics can be utilised in post-processing large multidimensional datasets, such as 4D flow datasets to improve valve flow assessment [12].

5. Conclusion

To automatically extract the valve-related motion parameters, the proposed AI-based system was developed and analysed on a large dataset. We investigated the robustness and feasibility of the system extensively on MV and TV related motion parameters. The system achieved human-level accuracy and can improve the workflow efficiency, automation and standardization of valve-related acquisitions or analyses.

CRediT authorship contribution statement

Seung Su Yoon: Conceptualization, Data curation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Carola Fischer:** Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft. **Daniel Amsel:** Conceptualization, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing. **Maria Monzon:** Conceptualization, Methodology, Software. **Solenn Toupin:** Data curation, Writing – review & editing. **Théo Pezel:** Data curation,

Writing – review & editing. **Jérôme Garot:** Data curation, Writing – review & editing. **Jens Wetzl:** Conceptualization, Data curation, Validation, Writing – review & editing. **Andreas Maier:** Supervision. **Daniel Giese:** Conceptualization, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare no conflict of interest and no financial disclosure in relation to this study. Carola Fischer receives a salary from Siemens Healthcare GmbH during her PhD. Seung Su Yoon, Solenn Toupin, Jens Wetzl, Daniel Giese are employees of Siemens Healthcare GmbH.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ejrad.2023.110978>.

References

- [1] C. Chen, C. Qin, H. Qiu, G. Tarroni, J. Duan, W. Bai, D. Rueckert, Deep learning for cardiac image segmentation: a review, *Frontiers in Cardiovascular Medicine* 7 (2020) 25.
- [2] E. Smistad, A. Østvik and others, “2D left ventricle segmentation using deep learning,” in 2017 IEEE international ultrasonics symposium (IUS), 2017.
- [3] J. Sun, F. Darbehani, M. Zaidi and B. Wang, “Saunet: Shape attentive u-net for interpretable medical image segmentation,” in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2020.
- [4] T. Chitiboi, B. Georgescu, J. Wetzl, I. Borgohain, C. Geppert, S. K. Piechnik, S. Neubauer, S. Petersen and P. Sharma, “Deep learning-based strain quantification from CINE cardiac MRI,” in ISMRM Annual Meeting, 2020.
- [5] J.G. Ramos, A. Fyrdahl, B. Wieslander, S. Thalén, G. Reiter, U. Reiter, N. Jin, E. Maret, M. Eriksson, K. Caidahl, P. Sörensson, A. Sigfridsson, M. Ugander, Comprehensive Cardiovascular Magnetic Resonance Diastolic Dysfunction Grading Shows Very Good Agreement Compared With Echocardiography, *J. Am. Coll. Cardiol. Img.* 13 (2020) 2530–2542.
- [6] S.F. Nagueh, O.A. Smiseth, C.P. Appleton, B.F. Byrd, H. Dokainish, T. Edvardsen, F. A. Flachskampf, T.C. Gillette, A.L. Klein, P. Lancellotti, et al., Recommendations for the evaluation of left ventricular diastolic function by echocardiography: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging, *Eur. J. Echocardiogr.* 17 (2016) 1321–1360.
- [7] S. Romano, R.M. Judd, R.J. Kim, H.W. Kim, I. Klem, J.F. Heitner, D.J. Shah, J. Jue, A. Farzaneh-Far, Left ventricular long-axis function assessed with cardiac cine MR imaging is an independent predictor of all-cause mortality in patients with reduced ejection fraction: a multicenter study, *Radiology* 286 (2018) 452.
- [8] S. Romano, R.M. Judd, R.J. Kim, H.W. Kim, J.F. Heitner, D.J. Shah, R.B. Devereux, P. Salazar, M. Trybula, R.C. Chia, et al., Prognostic implications of mitral annular plane systolic excursion in patients with hypertension and a clinical indication for cardiac magnetic resonance imaging: a multicenter study, *J. Am. Coll. Cardiol. Img.* 12 (2019) 1769–1779.
- [9] V. Rangarajan, S.J. Chacko, S. Romano, J. Jue, N. Jariwala, J. Chung, A. Farzaneh-Far, Left ventricular long axis function assessed during cine-cardiovascular magnetic resonance is an independent predictor of adverse cardiac events, *J. Cardiovasc. Magn. Reson.* 18 (2016) 1–10.
- [10] P. Thavendiranathan, C. Guetter, J.S. da Silveira, X. Lu, D. Scandling, H. Xue, M.-P. Jolly, S.V. Raman, O.P. Simonetti, Mitral annular velocity measurement with cardiac magnetic resonance imaging using a novel annular tracking algorithm: Validation against echocardiography, *Magn. Reson. Imaging* 55 (2019) 72–80.
- [11] V.P. Kamphuis, A.A.W. Roest, N.A. Marsan, P.J.V.D. Boogaard, L.J.M. Kroft, J. P. Aben, J.J. Bax, A.D. Roos, H.J. Lamb, J.J.M. Westenberg, Automated cardiac valve tracking for flow quantification with four-dimensional flow MRI, *Radiology* 290 (2019) 70–78.
- [12] J.J.M. Westenberg, S.D. Roes, N.A. Marsan, N.M.J. Binnendijk, J. Doornbos, J. J. Bax, J.H.C. Reiber, A.d. Roos, R.J.v.d. Geest, Mitral Valve and Tricuspid Valve Blood Flow: Accurate Quantification with 3D Velocity-encoded MR Imaging with Retrospective Valve Tracking, *Radiology* 249 (2008) 792–800.
- [13] H. Xue, J. Artico, R.H. Davies, R. Adam, A. Shetye, J.B. Augusto, A. Bhuva, F. Fröjd, T.C. Wong, M. Fukui, et al., Automated In-Line Artificial Intelligence Measured Global Longitudinal Shortening and Mitral Annular Plane Systolic Excursion: Reproducibility and Prognostic Significance, *J. Am. Heart Assoc.* 11 (2022) e023849.
- [14] R.A. Gonzales, F. Seemann, J. Lamy, H. Mojibian, D. Atar, D. Erlinge, K. Steding-Ehrenborg, H. Arheden, C. Hu, J.A. Onofrey, et al., MVnet: automated time-resolved tracking of the mitral valve plane in CMR long-axis cine images with residual neural networks: a multi-center, multi-vendor study, *J. Cardiovasc. Magn. Reson.* 23 (2021) 1–15.
- [15] R. A. Gonzales, J. Lamy, F. Seemann, E. Heiberg, J. A. Onofrey and D. C. Peters, “TVnet: Automated time-resolved tracking of the tricuspid valve plane in MRI long-axis cine images with a dual-stage deep learning pipeline,” in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2021.
- [16] Z. Jiang, C. Ding, M. Liu and D. Tao, “Two-stage cascaded U-Net: 1st place solution to BraTS challenge 2019 segmentation task,” in International MICCAI brainlesion workshop, 2019.
- [17] M. Monzon, S. S. Yoon, C. Fischer, A. Maier, J. Wetzl and D. Giese, “Fully automatic extraction of mitral valve annulus motion parameters on long axis CINE CMR using deep learning,” in International Society for Magnetic Resonance in Medicine (ISMRM) 29th Annual Meeting & Exhibition, 2021.
- [18] A. Toshev and C. Szegedy, “Deeppose: Human pose estimation via deep neural networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014.
- [19] T. Pfister, J. Charles and A. Zisserman, “Flowing convnets for human pose estimation in videos,” in Proceedings of the IEEE international conference on computer vision, 2015.
- [20] J.J. Tompson, A. Jain, Y. LeCun, C. Bregler, Joint training of a convolutional network and a graphical model for human pose estimation, *Adv. Neural Inf. Procs. Syst.* 27 (2014).
- [21] C. Payer, D. Stern, H. Bischof and M. Urschler, “Regressing heatmaps for multiple landmark localization using CNNs,” in International conference on medical image computing and computer-assisted intervention, 2016.
- [22] A. Myronenko, 3D MRI brain tumor segmentation using autoencoder regularization, in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4, 2019.
- [23] D. Amsel, S. S. Yoon, J. Wetzl and A. Maier, “Diffeomorphic Image Registration for CINE Cardio MR images using deep learning,” in Joint Annual Meeting ISMRM-ESMRMB & ISMRT 31st Annual Meeting, 2022.
- [24] C.G. Fonseca, M. Backhaus, D.A. Bluemke, R.D. Britten, J.D. Chung, B.R. Cowan, I. D. Dinov, J.P. Finn, P.J. Hunter, A.H. Kadish, et al., The Cardiac Atlas Project—an imaging database for computational modeling and statistical atlases of the heart, *Bioinformatics* 27 (2011) 2288–2295.
- [25] T. Pezel, P. Garot, T. Hovasse, T. Unterseeh, S. Champagne, M. Kinnel, S. Toupin, Y. Louvard, M.C. Morice, F. Sanguineti, et al., Vasodilatation stress cardiovascular magnetic resonance imaging: Feasibility, workflow and safety in a large prospective registry of more than 35,000 patients, *Arch. Cardiovasc. Dis.* 114 (2021) 490–503.
- [26] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, 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 Med.* 12 (2015) e1001779.
- [27] A.V. Dalca, G. Balakrishnan, J. Guttag, M.R. Sabuncu, Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces, *Med. Image Anal.* 57 (2019) 226–236.
- [28] A. Mayr, M. Paminger, M. Reindl, S. Greulich, S.J. Reinstadler, C. Tiller, M. Holzknecht, T. Nalbach, D. Plappert, C. Kranewitter, et al., Mitral annular plane systolic excursion by cardiac MR is an easy tool for optimized prognosis assessment in ST-elevation myocardial infarction, *Eur. Radiol.* 30 (2020) 620–629.
- [29] H. Bulluck, H. Ngamkasem, D. Sado, T.A. Treibel, M. Fontana, V. Maestrini, J. Moon, D.J. Hausenloy, A.S. Herrey, S.K. White, A simple technique to measure TAPSE and MAPSE on CMR and normal values, *J. Cardiovasc. Magn. Reson.* 16 (2014) 1–2.
- [30] M.M. Ochs, T. Fritz, F. André, J. Riffel, D. Mereles, M. Müller-Hennens, E. Giannitsis, H.A. Katus, M.G. Friedrich, S.J. Buss, A comprehensive analysis of cardiac valve plane displacement in healthy adults: age-stratified normal values by cardiac magnetic resonance, *Int. J. Cardiovasc. Imaging* 33 (2017) 721–729.
- [31] S. Cheng, V.R.S. Fernandes, D.A. Bluemke, R.L. McClelland, R.A. Kronmal, J.A. C. Lima, Age-related left ventricular remodeling and associated risk for cardiovascular outcomes: the Multi-Ethnic Study of Atherosclerosis, *Circulation Cardiovascular Imaging* 2 (2009) 191–198.
- [32] J.H. Riffel, R. Mayo, M. Mueller-Hennens, E. Giannitsis, H.A. Katus, F. Andre, Age-and gender-related reference values of cardiac morphology and function in cardiovascular magnetic resonance, *Int. J. Cardiovasc. Imaging* 37 (2021) 2011–2023.
- [33] F. Ricci, N. Aung, S. Gallina, F. Zemrak, K. Fung, G. Bisaccia, J.M. Paiva, M. Y. Khanji, C. Mantini, S. Palermi, et al., Cardiovascular magnetic resonance reference values of mitral and tricuspid annular dimensions: the UK Biobank cohort, *J. Cardiovasc. Magn. Reson.* 23 (2021) 1–13.
- [34] S. Kozerke, J. Schwitser, E.M. Pedersen, P. Boesiger, Aortic and mitral regurgitation: quantification using moving slice velocity mapping, *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 14 (2001) 106–112.
- [35] F. Seemann, E. Heiberg, M. Carlsson, R.A. Gonzales, L.A. Baldassarre, M. Qiu, D. C. Peters, Valvular imaging in the era of feature-tracking: A slice-following cardiac MR sequence to measure mitral flow, *J. Magn. Reson. Imaging* 51 (2020) 1412–1421.