An Open-Source End-to-End Pipeline for Generating 3D+t Biventricular Meshes from Cardiac Magnetic Resonance Imaging

Joshua R. Dillon[1(✉)], Charlène Mauger[2], Debbie Zhao[1], Yu Deng[2], Steffen E. Petersen[3,4], Andrew D. McCulloch[5], Alistair A. Young[2], and Martyn P. Nash[1,6]

[1] Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand
joshua.dillon@auckland.ac.nz
[2] School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
[3] William Harvey Research Institute, Queen Mary University of London, London, UK
[4] Barts Heart Centre, Barts Health NHS Trust, London, UK
[5] Institute for Engineering in Medicine, University of California San Diego, San Diego, USA
[6] Department of Engineering Science and Biomedical Engineering, University of Auckland, Auckland, New Zealand

Abstract. Increased interest in digital-twin based healthcare has stimulated recent advancements in personalised biventricular modelling from cardiac magnetic resonance (CMR) imaging. However, there remains no publicly available end-to-end pipeline for generating structured meshes of the heart across the entire cardiac cycle. This paper presents a new pipeline and tests it on CMR data contributed from two centres. The proposed pipeline comprises view classification, segmentation of the left and right ventricular chambers and myocardium, contour generation, and model fitting in a fully automated sequence, requiring only an image directory as input. The use of 3D U-Nets was explored, and found to increase the temporal coherence of resulting meshes compared to 2D U-Nets when evaluated on 10 test cases. The pipeline is available to be deployed across various applications, including digital-twin based simulations, statistical shape and motion analyses, and clinical research. The pipeline—including code, models, and documentation—can be accessed at https://github.com/UOA-Heart-Mechanics-Research/biv-me.

Keywords: CMR · Biventricular modelling · Open-source

1 Introduction

Cardiac magnetic resonance (CMR) imaging is the gold standard imaging modality for cardiac chamber quantification, from which high-fidelity representations

R. Chabiniok et al. (Eds.): FIMH 2025, LNCS 15673, pp. 372–383, 2025.
https://doi.org/10.1007/978-3-031-94562-5_34

of biventricular anatomy can be created. Personalised modelling of biventricular geometry and motion is vital for the accurate quantification of chamber dimensions and kinematics, as well as the development of digital twin-based simulations of cardiac biophysics. Several automated pipelines have been reported to be capable of generating biventricular meshes from cine CMR. These pipelines apply a variety of approaches, including adapting biventricular shape templates to sparse contours generated by automatic segmentation [1,4,11,14], tetrahedralising point clouds generated from point or label completion networks [2,15,16], or by application of graph convolution networks [3]. Despite this diversity, existing methods for modelling biventricular shape and motion from CMR lack deployability in research or clinical settings, and they have not typically been made available for public use, which hampers the formation of a consensus, standardised approach to cardiac modelling. While U-Net-based segmentation models have demonstrated excellent performance on 2D cine CMR [8], their use relies on accurate view labelling and categorisation of often poorly-defined, heterogeneous image file systems. Furthermore, the vast majority of published pipelines to date have modelled the heart only at end-diastole (ED) and end-systole (ES), likely due to the time-cost of segmenting all frames of the cardiac cycle, and/or the lack of temporal coherence across individually segmented frames. Such exclusion of most of the cardiac cycle prevents more detailed characterisation of left ventricular (LV) and right ventricular (RV) dynamics, such as the calculation of rates of filling (based on derivatives of volume time curves), which may provide clinical value in the evaluation of diastolic dysfunction [7,10]. The RV myocardium is usually either excluded from modelling pipelines, or estimated by a constant fixed offset (often 3 mm) to the RV endocardial border, due to its slenderness relative to the spatial resolution of the imaging data. Available pipelines are often trained on single datasets, or limited to specific pathologies, further constraining their application as generalisable research tools.

This paper presents an open-source, deployable, end-to-end 3D U-Net-based pipeline for generating 3D+t biventricular meshes from CMR images. We show that this approach decreases modelling time and increases temporal coherence of the resulting meshes. We demonstrate the generalisability of the pipeline by evaluating it using two independent datasets with mixed demographics and pathologies. The pipeline—including code, models, and documentation—is available at https://github.com/UOA-Heart-Mechanics-Research/biv-me.

2 Methods

2.1 Data Description

The development of this pipeline was carried out using CMR images contributed from two datasets. The first dataset comprised images of 200 healthy participants from the UK Biobank [12]. This research has been conducted using the UK Biobank Resource under Application Number 2964. Overall ethical approval for UK Biobank studies was gained from the NHS National Research Ethics Service on June 17, 2011, and this was last extended on June 18, 2021. All study participants provided written informed consent.

The second dataset consisted of a mixed healthy/patient cohort of 160 participants from the CARDIOHANCE study [17], prospectively recruited and imaged at the Centre for Advanced MRI, University of Auckland. Ethical approval was granted by the Health and Disability Ethics Committee of New Zealand (17/CEN/226). Written informed consent was obtained from each participant. Patient diagnoses included coronary artery disease, myocardial infarction, left ventricular hypertrophy, hypertrophic cardiomyopathy, amyloidosis, and heart failure. The proposed pipeline was evaluated using 10 cases (5 from each dataset).

2.2 View Classification

Data Preparation - For each case, the CMR series were divided into separate view classes. Due to differences in imaging protocols between centres, there is variability in the cardiac image views available for building cardiac models. We therefore defined a comprehensive set of view classes that covered all image types available across the two datasets. These classes were short axis (SAX), long axis two-chamber left-heart (2ch), long axis three-chamber (3ch), long axis four-chamber (4ch), long axis two-chamber right-heart (2ch-RT), right ventricular outflow tract (RVOT), transverse right ventricular outflow tract (RVOT-T), left ventricular outflow tract (LVOT), short axis atria (SAX-atria), and short axis images that contain no ventricular structures (Other). Of these views, the 2ch-RT, RVOT and RVOT-T were exclusive to the CARDIOHANCE imaging protocol. All images for each of the 200 UK Biobank and 160 CARDIOHANCE cases were manually reviewed and divided into one of the above 10 view classes by a single expert observer (Fig. 1).

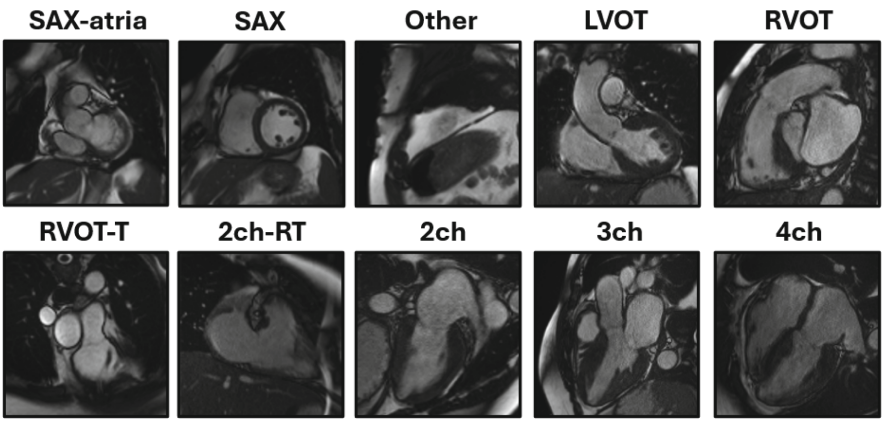


Fig. 1. Typical examples of each view class. Refer to Sect. 2.2 for view definitions.

Model Training - A convolutional neural network based on the ResNet50 architecture [5] was trained to automatically predict the view of a CMR image series. The model was trained for 50 epochs using a cross-entropy loss function with an Adam optimiser and an initial learning rate of 0.0001. Augmentations for orientation (random horizontal and vertical flipping, rotation) and translation (random translation, random crop) were carried out prior to model training. Cases were randomly split 80/10/10 for training, validation, and testing, yielding 288 cases for training, 36 cases for validation, and 36 cases for testing. This amounted to 4278 image series for model training, 533 for validation, and 520 for testing. Training, validation, and evaluation of view prediction performance on the test set were carried out using the first frame of each image series.

2.3 Segmentation

Data Preparation - The self-configuring deep learning framework nnU-Net [6] was used to train segmentation models for each of the views to automatically label the relevant cardiac structures within each image series. Notably, the RV myocardium was segmented for each view where present. Of the 360 cases from Sect. 2.1, 100 were randomly chosen to be segmented by a single expert at both ED and ES frames. 3–5 SAX series were segmented per case, along with 1–2 2ch, 3ch, 4ch, and RVOT series (where available). These training data consisted of segmentations from 52 UK Biobank cases and 48 CARDIOHANCE cases.

Generating Ground Truth Labels - Rather than manually segmenting all frames, we used an 'overfitting' approach that involved deliberately overfitting a default 2D nnU-Net model to its training data at ED and ES (without cross-validation or held-out test data). For each view, a model was trained for 750 epochs and then applied to the training cases to predict segmentations of all other (non ED or ES) frames. For occasions when the overfit model predictions generated disconnected segmentations, they were corrected manually. After correction, predictions across the entire cardiac cycle were concatenated along the time dimension to form new 2D+time (2D+t) reference segmentations for training.

Segmentation Models - The labels generated by the overfit model were subsequently used to train default 3D nnU-Net models with single-fold cross-validation and a 90/10 training/testing split for 500 epochs. A subset of 10 cases from the view classification test set (36 cases) was used for evaluating the predicted segmentations and model building pipeline. Henceforth, this set of 10 cases, unused for any model training, will be referred to as 'evaluation cases' or as 'the evaluation set'. For comparison, default 2D nnU-Net models were trained for 250 epochs (with the same training/testing split) using the original manual segmentations at ED and ES. The 2D and 2D+t segmentation models were applied separately on the evaluation set to generate labels (see Fig. 2) and contour sets (see Fig. 3) for each frame in the cardiac cycle. These contour sets include endocardial, atrial

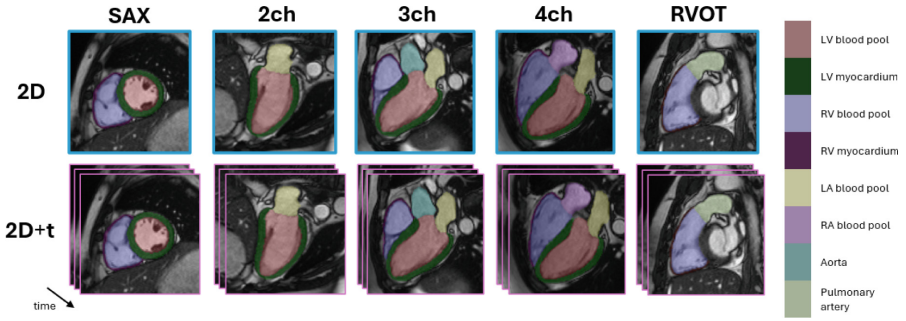


Fig. 2. Segmentations generated for a CARDIOHANCE evaluation case by the 2D and 2D+t segmentation models. Refer to Sect. 2.2 for view definitions.

and epicardial borders, and key landmarks (apex, RV inserts, mitral valve, tricuspid valve, aortic valve, pulmonary valve) generated from intersections of those borders.

2.4 Biventricular Mesh Fitting

Using these contour sets, biventricular meshes were fitted for each time frame using a previously described iterative diffeomorphic registration algorithm [9]. Briefly, a multi-class subdivision surface template mesh is refined to each contour set with an implicit linear least squares fit and an explicit diffeomorphic fit to yield a structured point-correspondent mesh across the cardiac cycle. Time-varying biventricular meshes were generated for the 10 evaluation cases using contour sets generated by either the 2D or 2D+t segmentation models.

2.5 Pipeline

View classification, segmentation, contour generation, and mesh fitting were carried out fully automatically as part of an end-to-end pipeline. The only input to the pipeline was a path to a folder/directory containing raw CMR DICOM images for each subject. The outputs of the pipeline were epicardial (2502 vertices), LV endocardial (1572 vertices), and RV endocardial meshes (1956 vertices) for each frame of the cardiac cycle for each case. Processing took between 20–35 minutes per case depending on the number of frames in the cine CMR series.

2.6 Evaluation

For each case, volumes and masses for the LV and RV were calculated from the generated biventricular meshes. The temporal coherence of each biventricular mesh was evaluated by the smoothness of the volume-time traces and the consistency of the mass-time traces. To calculate smoothness, a ‘smooth’ reference trace was constructed for each volume-time trace by applying a low-pass

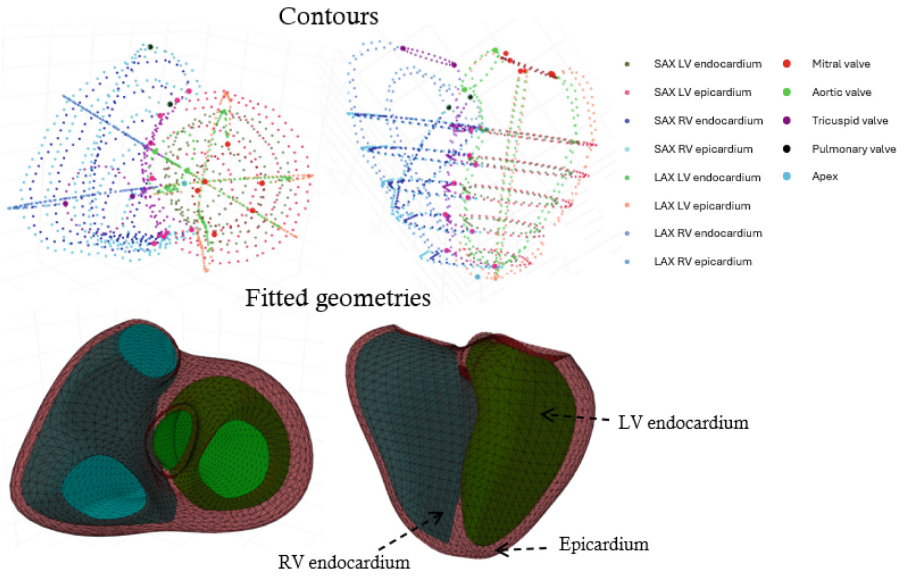


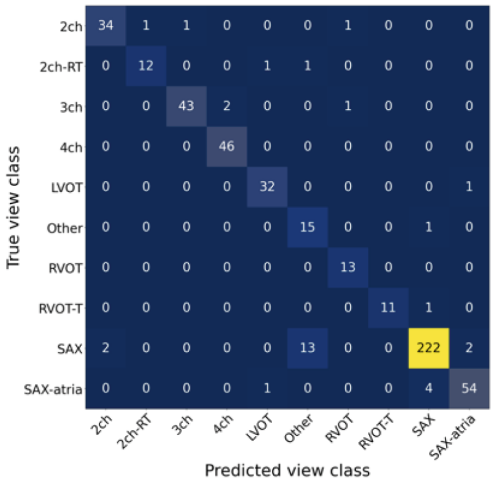
Fig. 3. Contour sets and fitted meshes generated for the first frame of a CARDIOHANCE evaluation case.

filter. Smoothness was calculated as the L2 norm of residuals between each original and filtered volume-time traces, whereby lower values represented smoother traces. The consistency of each mass trace was calculated as the L2 norm of residuals between the mass-time trace and the average mass, whereby lower values reflected greater mass consistency. Smoothness and mass consistency values were averaged across the 10 evaluation cases and compared between biventricular meshes fitted using the 2D and 2D+t segmentation models, with paired sample t-tests used for evaluating statistical significance of the differences.

3 Results

View Classification - As shown in Fig. 4, the view classification model achieved excellent predictive performance, with high F-scores and true positive rates across most classes. 15/239 (6%) SAX images mistaken for SAX-atria or Other images as the distinctions between those views are often subtle. We found that excluding the most apical and/or most basal SAX slice had little downstream impact on the generated biventricular meshes (data not shown). The overall accuracy across all view classes was 94%.

Segmentation - The performance of each segmentation model on the evaluation set was assessed by calculating the Dice coefficient (Table 1). Similar performance was observed between the 2D and 2D+t segmentation models, with superior accuracy for the latter when segmenting RVOT images (albeit for a



View	TPR	FPR	TNR	FNR	F-score
2ch	0.92	0.00	1.00	0.08	0.93
2ch-RT	0.86	0.00	1.00	0.14	0.89
3ch	0.93	0.00	1.00	0.07	0.96
4ch	1.00	0.00	1.00	0.00	0.98
LVOT	0.97	0.00	1.00	0.03	0.96
Other	0.94	0.03	0.97	0.06	0.67
RVOT	1.00	0.00	1.00	0.00	0.93
RVOT-T	0.92	0.00	1.00	0.08	0.96
SAX	0.93	0.02	0.98	0.07	0.95
SAX-atria	0.92	0.01	0.99	0.08	0.93

b)

Fig. 4. a). Confusion matrix representing the performance of the view classification model across 36 test cases. b). Summary of performance metrics for each view including True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR), and False Negative Rate (FNR).

small sample size of 4 images, since 1 case was missing the RVOT view). The 2D+t segmentation model was 4-fold more efficient than the 2D segmentation model, taking an average of 2 s per frame on an NVIDIA RTX 2000 8 GB GPU, compared to 8 s per frame for the 2D model. For the range of cases in the evaluation set (between 20–50 image frames per cycle), the 2D+t model was 2 to 5 min faster than the 2D model per case.

Temporal Coherence - Biventricular meshes generated using the 2D+t segmentation model exhibited higher temporal coherence than those from the 2D segmentation model, as evidenced by the significantly smoother LV volume-time traces (2D: 6.98 ± 2.46 vs 2D+t: 4.43 ± 1.62, $p = 0.004$) and RV volume-time traces (2D: 7.90 ± 3.40 vs 2D+t: 5.38 ± 1.99, $p = 0.030$). No significant differences in LV or RV mass consistency were observed.

4 Discussion

The proposed end-to-end biventricular meshing pipeline showed excellent performance in creating personalised parameterisations of cardiac shape and motion across the cardiac cycle. The modelling pipeline performs well for disorganised image file systems, due to the reliability of the view classification model (Fig. 4). There were negligible differences in segmentation performance (Dice coefficient) or smoothness of biventricular meshes between UK Biobank and

Table 1. Comparison of segmentation performance between the 2D and 2D+t segmentation models. For the 2D model, the Dice score was calculated with reference to the ground truth ED and ES labels, whereas for the 2D+t model, Dice was calculated with reference to the ground truth labels for all image frames for a heart cycle.

	Dice	
View	**2D**	**2D+t**
SAX (n = 50)		
LV blood pool	0.98	0.98
LV myocardium	0.94	0.93
RV blood pool	0.93	0.93
RV myocardium	0.74	0.76
2ch (n = 10)		
LV blood pool	0.98	0.98
LV myocardium	0.95	0.95
LA blood pool	0.96	0.96
3ch (n = 12)		
LV blood pool	0.97	0.98
LV myocardium	0.95	0.95
RV blood pool	0.95	0.96
RV myocardium	0.79	0.78
LA blood pool	0.98	0.97
Aorta	0.98	0.98

	Dice	
View	**2D**	**2D+t**
4ch (n = 10)		
LV blood pool	0.97	0.98
LV myocardium	0.95	0.96
RV blood pool	0.97	0.97
RV myocardium	0.81	0.82
LA blood pool	0.96	0.97
RA blood pool	0.97	0.97
RVOT (n = 4)		
RV blood pool	0.90	0.95
RV myocardium	0.67	0.77
Pulmonary artery	0.90	0.95

CARDIHOANCE cases (Fig. 5), indicating good generalisability across imaging centres, protocols, and subject demographics. As well as being readily deployable, a key contribution of this pipeline is a step towards improved modelling of biventricular motion. The 'overfitting' approach for generating high quality full cycle training data from just the ED and ES frames was a viable way to facilitate development of a 2D+t segmentation approach that was more efficient and resulted in smoother models of motion compared to a conventional (frame-by-frame) 2D segmentation approach. The improvements in temporal coherence and the reduced segmentation times from the 2D+t segmentation approach can enable robust estimates of kinematics, filling rates, strain rates, and other indicators of cardiac function that may improve clinical assessment of various cardiac conditions [10]. As the 'overfitting' approach utilises only the ED and ES frames, there is a possibility that the approach introduces a temporal bias; that is, that the predictions on the intermediate frames of the 2D+t reference labels are biased towards those key frames. Further improvements in temporal coherence could be achieved by incorporating additional key frames, e.g., at diastasis.

A valuable contribution of this work is towards characterisation of the RV. The RV myocardium is often excluded from cardiac modelling, or estimated by applying a constant (often 3 mm) offset from the endocardial RV free wall. The

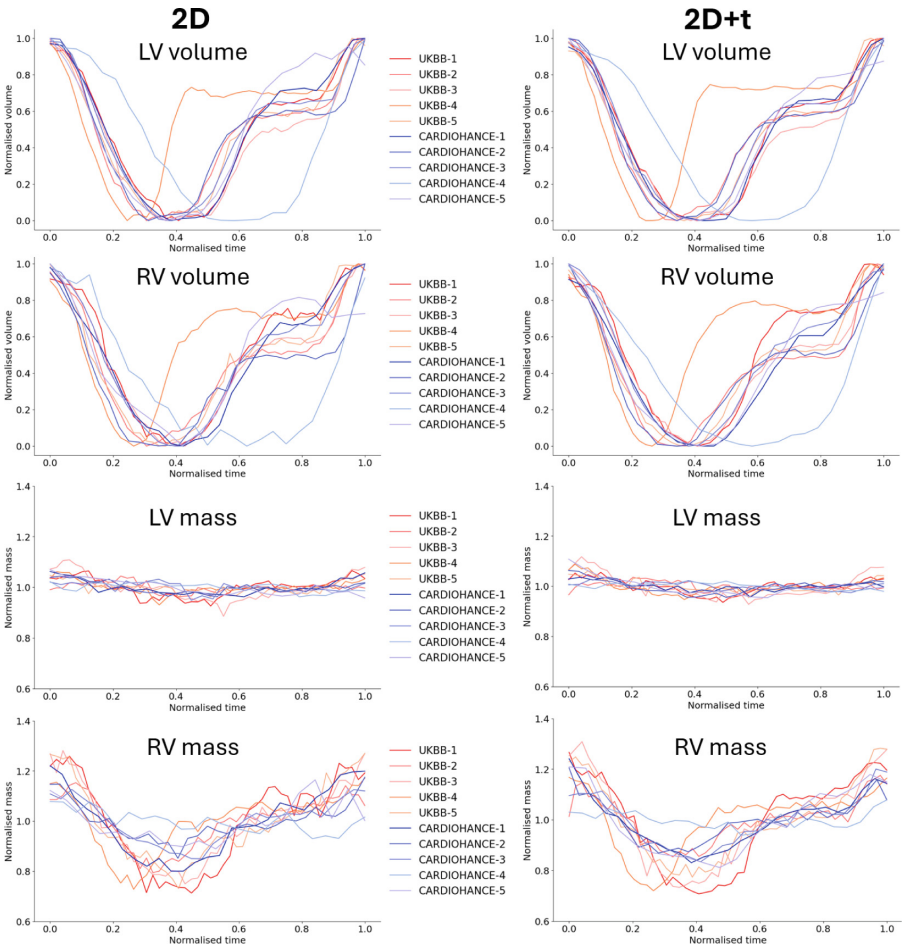


Fig. 5. Normalised volume and mass curves for all 10 evaluation cases generated by applying the 2D or 2D+t segmentation models. Volumes were normalised to lie between 0–1, whereas masses were normalised with respect to the average mass for each case.

inclusion of the RV myocardium in our segmentation models allows for regional and temporal variation of RV wall thickness, enabling future research into the clinical value of wall thickness and thickening rates in different parts of the RV. The standard cine image spatial resolution of 1.4–1.8 mm [12] is poor considering the slenderness of the RV myocardium. Nevertheless, our pixel-wise approach allows for representation of multiple states of wall thickness, which is superior to a constant wall thickness approximation. With respect to the Dice score, RV myocardium segmentation performance was relatively poor compared to other cardiac structures (Table 1). Improvements in performance can be expected with an increase in training data, given the small dataset employed for this study,

but an upper limit of segmentation performance is not known given the lack of available benchmarks. Given that Dice likely underestimates segmentation performance due to the slenderness and asymmetry of the RV myocardium, it may be more relevant to judge segmentation performance by physiology-based heuristics, e.g., the degree to which RV mass is conserved across the cycle. As RV mass traces were relatively inconsistent compared to LV mass traces (Fig. 5), this warrants further research into approaches to RV myocardium segmentation that optimally conserve RV mass.

Since the template-based diffeomorphic fitting approach provides point correspondence between each mesh, population-level statistical shape and/or motion analyses can be carried out without recourse to standardisation schemes such as a universal ventricular coordinate system (UVC) [13]. If required, UVC can be directly embedded in the structure of the meshes, allowing for a consistent mapping between meshes and easy integration into existing digital twin-based pipelines for cardiac biophysical simulation (such as electrophysiology and/or mechanics). While other pipelines have been developed for generating biventricular meshes from CMR images, this is the first published pipeline that is open-source, end-to-end, capable of generating temporally coherent point-correspondent meshes for the entire cardiac cycle, and with regional and temporal variation in RV wall thickness.

5 Conclusion

The biventricular meshing pipeline presented in this paper demonstrated excellent performance across data contributed from two different centres. The improvements in temporal coherence and parameterisation of the RV myocardium represents important advancements in personalised anatomical modelling of the heart. The pipeline is readily deployable for a range of applications, from digital-twin based simulations, to statistical shape and motion analyses, and to clinical cardiac research.

Acknowledgments. This study was funded by the Health Research Council of New Zealand (grants 17/608 and 23/527). We would like to thank Laura Del Toso, Esther Puyol-Anton, Sachin Govil, Brendan Crabb, Anna Mîra, and Devran Ugurlu for their contributions to the development of this pipeline. We gratefully acknowledge the study participants of the UK Biobank and CARDIOHANCE, and the staff at the Centre for Advanced MRI at the University of Auckland for their imaging expertise.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Code Availability. The pipeline code, models, and documentation are available at https://github.com/UOA-Heart-Mechanics-Research/biv-me.

References

1. Banerjee, A., et al.: A completely automated pipeline for 3D reconstruction of human heart from 2D cine magnetic resonance slices. Philos. Trans. Roy. Soc. A: Math. Phys. Eng. Sci. **379**(2212), 20200257 (2021). https://doi.org/10.1098/rsta.2020.0257
2. Beetz, M., Banerjee, A., Ossenberg-Engels, J., Grau, V.: Multi-class point cloud completion networks for 3D cardiac anatomy reconstruction from cine magnetic resonance images. Med. Image Anal. **90**, 102975 (2023). https://doi.org/10.1016/j.media.2023.102975
3. Deng, Y., et al.: ModusGraph: automated 3D and 4D mesh model reconstruction from cine CMR with improved accuracy and efficiency. In: Greenspan, H., et al. (eds.) MICCAI 2023. LNCS, vol. 14226, pp. 173–183. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-43990-2_17
4. Govil, S., et al.: A deep learning approach for fully automated cardiac shape modeling in tetralogy of Fallot. J. Cardiovasc. Magn. Reson. **25**(1), 15 (2023). https://doi.org/10.1186/s12968-023-00924-1
5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition (2015). https://doi.org/10.48550/arXiv.1512.03385
6. Isensee, F., Jaeger, P.F., Kohl, S., Petersen, J., Maier-Hein, K.H.: nnU-net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods **18**(2), 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z
7. Kawaji, K., et al.: Automated segmentation of routine clinical cardiac magnetic resonance imaging for assessment of left ventricular diastolic dysfunction. Circ.: Cardiovasc. Imaging **2**(6), 476–484 (2009). https://doi.org/10.1161/CIRCIMAGING.109.879304
8. Martín-Isla, C., et al.: Deep learning segmentation of the right ventricle in cardiac MRI: the M&Ms challenge. IEEE J. Biomed. Health Inform. **27**(7), 3302–3313 (2023). https://doi.org/10.1109/JBHI.2023.3267857
9. Mauger, C., et al.: An iterative diffeomorphic algorithm for registration of subdivision surfaces: application to congenital heart disease. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2018, pp. 596–599 (2018). https://doi.org/10.1109/EMBC.2018.8512394
10. Mendoza, D.D., et al.: Impact of diastolic dysfunction severity on global left ventricular volumetric filling - assessment by automated segmentation of routine cine cardiovascular magnetic resonance. J. Cardiovasc. Magn. Reson. **12**(1), 46 (2010). https://doi.org/10.1186/1532-429X-12-46
11. Miller, R., Kerfoot, E., Mauger, C., Ismail, T.F., Young, A.A., Nordsletten, D.A.: An implementation of patient-specific biventricular mechanics simulations with a deep learning and computational pipeline. Front. Physiol. **12** (2021). https://doi.org/10.3389/fphys.2021.716597
12. Petersen, S.E., et al.: UK Biobank's cardiovascular magnetic resonance protocol. J. Cardiovasc. Magn. Reson. **18**(1), 8 (2016). https://doi.org/10.1186/s12968-016-0227-4
13. Schuler, S., Pilia, N., Potyagaylo, D., Loewe, A.: Cobiveco: consistent biventricular coordinates for precise and intuitive description of position in the heart - with MATLAB implementation. Med. Image Anal. **74**, 102247 (2021). https://doi.org/10.1016/j.media.2021.102247

14. Xia, Y., et al.: Automatic 3D+t four-chamber CMR quantification of the UK biobank: integrating imaging and non-imaging data priors at scale. Med. Image Anal. **80**, 102498 (2022). https://doi.org/10.1016/j.media.2022.102498
15. Xu, H., Zacur, E., Schneider, J.E., Grau, V.: Ventricle surface reconstruction from cardiac MR slices using deep learning. In: Coudière, Y., Ozenne, V., Vigmond, E., Zemzemi, N. (eds.) FIMH 2019. LNCS, vol. 11504, pp. 342–351. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21949-9_37
16. Xu, Y., et al.: Improved 3D whole heart geometry from sparse CMR slices (2024). https://doi.org/10.48550/arXiv.2408.07532
17. Zhao, D., et al.: MITEA: a dataset for machine learning segmentation of the left ventricle in 3D echocardiography using subject-specific labels from cardiac magnetic resonance imaging. Front. Cardiovasc. Med. **9** (2023). https://doi.org/10.3389/fcvm.2022.1016703