

BIVENTRICULAR SURFACE RECONSTRUCTION FROM CINE MRI CONTOURS USING POINT COMPLETION NETWORKS

Marcel Beetz* Abhirup Banerjee[†]* Vicente Grau*

* Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, UK

[†] Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, UK

ABSTRACT

Many important cardiac biomarkers used in clinical practice describe cardiac anatomy and function in three dimensions (3D). However, common cardiac magnetic resonance imaging (MRI) protocols often only generate two-dimensional (2D) image slices of the underlying 3D anatomy and are susceptible to various types of motion artifacts causing slice misalignment. In this paper, we propose a deep learning method acting directly on point clouds to reconstruct a dense 3D biventricular heart model from misaligned 2D cardiac MR image contours. The method is able to reduce mild, medium, and strong slice misalignments (mean translation ~ 3.5 mm; mean rotation $\sim 2.5^\circ$) to a Chamfer distance below image resolution (1.25 mm) with high robustness (standard deviation 0.18 mm) on a statistical shape model dataset. It also manages to reconstruct smooth 3D shapes with accurate left ventricular volumes from cine MR images of the UK Biobank study.

Index Terms— Cardiac Surface Reconstruction, Cine MRI, Motion Correction, Point Completion Networks.

1. INTRODUCTION

Cardiac image segmentation is an important task in clinical practice as it enables the calculation of clinically important biomarkers from cardiac images. In recent years, deep learning methods, such as the U-net [1], have largely replaced previous manual or semi-automatic approaches in image segmentation research and achieved similar performance as human experts on a variety of datasets, including cine MRI [2].

However, whereas image segmentation quantifies structures in 2D, most clinically relevant anatomical biomarkers (e.g. volume, mass) are either inherently 3D or benefit significantly from the 3D extension of multiple 2D measurements. One way to obtain accurate 3D biomarkers is to increase the spatial resolution during image acquisition and directly generate 3D imaging data, as for example in 3D MRI. However, in current clinical practice, 3D MRI suffers from considerably lower temporal resolution and requires breath holds of up to twice as long as standard 2D cine MRI, making it infeasible for many patients. Another way to calculate 3D biomarkers is to estimate them directly from the 2D images of an MRI scan

with the modified Simpson’s rule. While simple and fast, this approach comes with considerable limitations in terms of accuracy, leading to potentially erroneous diagnostic and treatment decisions [3]. A third way tries to combine the advantages of both approaches by first building a 3D surface reconstruction based on the available 2D cardiac image contours and then calculating the required metrics directly from this 3D representation [4]. However, cardiac 3D surface reconstruction is a challenging problem due to two main reasons:

- *Data sparsity*: standard cine MRI has relatively low and anisotropic resolution, resulting in sparse 3D data.
- *Motion artifacts*: various types of motion during image acquisition lead to misalignments of anatomical structures between different image slices.

Considerable research has focused on correcting slice misalignment induced by respiratory motion between breath holds [5, 6] and on reconstructing 3D surfaces from sparse and noisy input data [7, 8]. In this paper, we propose a fast and fully automatic geometric deep learning method, capable of addressing both the sparsity and misalignment problem in a single model. It uses anatomical information from both short-axis (SAX) and long-axis (LAX) cardiac imaging planes, which are readily available as part of standard cardiac cine MRI acquisitions. Our main contributions are as follows:

- We develop an efficient deep learning method acting directly on point clouds for fully automatic biventricular surface reconstruction.
- We reinterpret the reconstruction task as a point cloud completion problem enabling the use of a point completion network for cardiac surface reconstruction.
- We evaluate our method on a dataset derived from a statistical shape model with different amounts of misalignment and on cine MR images from the UK Biobank study [9].

To the best of our knowledge, this is the first application of geometric deep learning to cardiac surface reconstruction from cine MRI contours and the first deep learning method overall to demonstrate successful biventricular surface reconstruction on real cine MR images. Previous deep learning approaches either focused only on the left ventricle [4], a different modality [10], or lacked validation on real data [11].

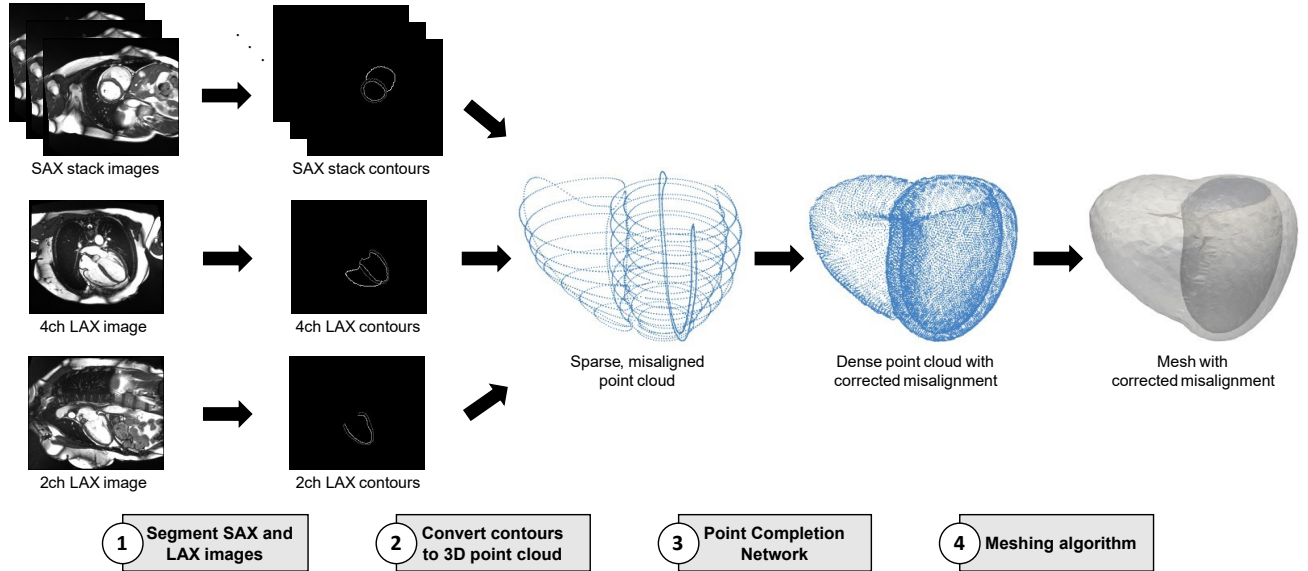


Fig. 1: Biventricular 3D surface reconstruction pipeline from cine MR images.

2. METHODS

This section describes the datasets used in this work (Sec. 2.1) and explains the proposed surface reconstruction pipeline (Sec. 2.2 and 2.4) and training procedure (Sec. 2.3).

2.1. Datasets

We use two datasets, one for training our method and providing a first validation and the other to evaluate our method.

We derive the first dataset from the biventricular statistical shape model (SSM) by Bai et al. [12]. The SSM was created based on 3D cardiac MRI scans with a resolution of $1.25 \times 1.25 \times 2$ mm of 1084 healthy volunteers, followed by image registration, segmentation, and principal component analysis. The 3D nature of the MRI acquisitions offers high spatial resolution both in-plane and between planes with drastically reduced slice misalignment and data sparsity compared to 2D acquisitions, which allows us to consider the shapes as our ground truth for method development.

The second dataset consists of 10 cases randomly sampled from the UK Biobank study [9]. Each case consists of a 4-chamber (ch) LAX, a 2ch LAX, and a stack of SAX cardiac cine MRI slices at the end-diastolic (ED) phase.

2.2. Surface Reconstruction Pipeline

We propose a 3D surface reconstruction pipeline that converts misaligned 2D cine MR images into dense biventricular 3D meshes (Fig. 1). The SAX image stack and the 2ch and 4ch LAX images from a standard cine MRI acquisition serve as inputs to our pipeline. First, the SAX and 4ch LAX images are segmented into 4 anatomical regions: left ventricular

(LV) cavity, LV myocardium, right ventricular (RV) cavity, and background; whereas the 2ch LAX image is divided into 3 areas: LV cavity, LV myocardium, and background. Next, we position the contours of each image slice in 3D space and convert the contours to point clouds by selecting each point along a given contour according to the in-plane resolution of the underlying image. The resulting point cloud is a sparse and misaligned 3D representation of the true biventricular surface. In the subsequent step, we employ a deep learning method to correct both misalignment and sparsity simultaneously (see Sec. 2.4). Our method outputs a dense biventricular point cloud with corrected misalignment, which we convert into a mesh using the Ball Pivoting Algorithm [13].

2.3. Training Data Generation

We use the SSM described in Sec. 2.1 to construct three datasets to analyze our method’s performance for different levels of misalignment. We first apply random deformations to the mean shape mesh along its first 31 modes of variation ($\sim 99.7\%$ of total variance in heart shapes) to generate 125 different biventricular meshes. For each mesh, we extract its vertices to form dense point clouds, which serve as the ground truth for network training. Next, we generate sparse and misaligned point clouds as network inputs. To this end, we first determine the LAX and SAX slice planes of each mesh by mimicking the procedure of a standard cardiac MRI acquisition [14]. We then artificially introduce slice misalignment caused by respiratory motion between breath holds at three different severity levels (mild, medium, strong) by rigidly transforming each slice. We sample the translation and rotation values for each level from separate normal dis-

tributions with zero mean. The respective standard deviations are selected to reflect realistic values observed in clinical practice [11] (translation: 1.5 mm, 2.5 mm, 3.5 mm; rotation: 0.5°, 1.5°, 2.5°). Finally, we slice the mesh at the randomly misaligned planes and sample points along each resulting contour to create the input point clouds. We apply 8 different sets of misalignment transforms to each of the 125 meshes to create a dataset of size 1000 for each level of misalignment.

2.4. Point Completion Network

Our surface reconstruction network has an encoder-decoder architecture similar to the Point Completion Network (PCN) proposed by Yuan et al. [15]. Two PointNet [16] layers are used as the encoder; whereas a multi-layer perceptron is followed by a FoldingNet [17] block in the decoder. We adjust the input and output layers of the PCN to fit the dimensionality of the biventricular point clouds and select a patch size for the FoldingNet block that enables a dense surface reconstruction. The loss function consists of two terms [15]:

$$L_{total} = L_{coarse} + \alpha * L_{dense} \quad (1)$$

The first loss term L_{coarse} compares the point cloud produced by the decoder before the FoldingNet block with the dense ground truth point cloud. Its goal is to force this sparse, intermediate point cloud to be a good overall representation of the global biventricular shape. The second loss term L_{dense} acts on the dense point cloud prediction of the FoldingNet block and enforces the desired smooth shape representation on both a global and local level. The weight α controls the importance of each of the two loss terms for the total loss and is steadily increased during training. We use the Chamfer distance for both loss terms.

3. EXPERIMENTS

For each of the three levels of misalignment severity, we split the respective dataset derived from the SSM into a training, validation, and test dataset of sizes 750, 50, and 200. We then train three different PCNs for each misalignment level and evaluate them on the corresponding test datasets.

Fig. 2 shows the sparse input point cloud, the network’s coarse and dense predictions, and the ground truth point cloud for a representative sample case of each misalignment level. Overall, our method is able to predict the global shape of the ground truth point clouds accurately for all levels of misalignment. The coarse predictions show a good coverage of the biventricular anatomy with points evenly spread out across the respective surfaces, and the LV myocardium, RV, and LV cavity are clearly discernible despite the low resolution. The dense prediction exhibits a smooth and mesh-like surface that captures shape variations on both a global and local scale.

In order to obtain quantitative results, we compare the dense model predictions with the ground truth point clouds

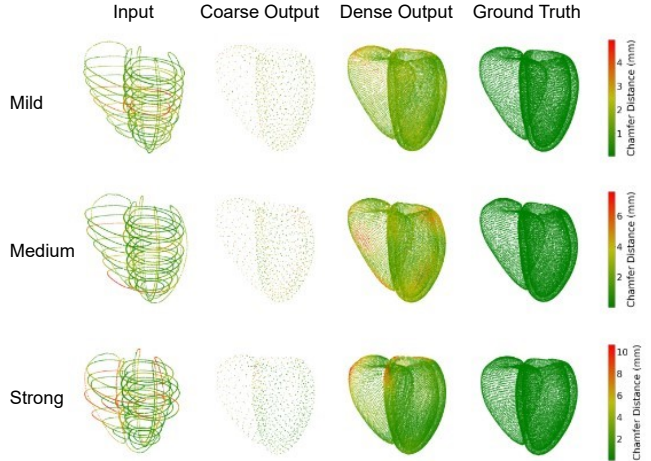


Fig. 2: Qualitative reconstruction results for each level of misalignment on the SSM-derived dataset. Point colors represent the Chamfer distances to the respective ground truth.

Table 1: Quantitative reconstruction results

	Misalignment level		
	Mild	Medium	Strong
Chamfer (mm) ¹	0.84 (±0.11)	1.01 (±0.17)	1.19 (±0.18)
Exec. time (ms) ²	24.10	23.78	24.50

¹ Chamfer: *mean* (± *SD*). ² Exec. time: *mean*.

of the test datasets for each level of misalignment using the Chamfer distance (Table 1). We find that our method achieves a mean Chamfer distance of 0.84 mm in case of mild misalignment with standard deviation (SD) of 0.11 mm. The Chamfer distance rises steadily when increasing the amount of misalignment, but remains below the underlying image pixel resolution of 1.25 mm for all misalignment levels. Standard deviations also increase with larger amounts of misalignment although to a lesser extent than the mean values. The execution time, i.e. the time required by the network to generate an output point cloud from an input, remains close to 24 ms for all misalignment levels.

Since no ground truth is available for the UK Biobank (UKB) dataset, we first conduct a visual analysis of the reconstructed shapes. We select the PCN trained on the SSM dataset with medium misalignment due to its closeness to real misalignment conditions and apply it directly to cine MRI segmentations of the UKB without further training.

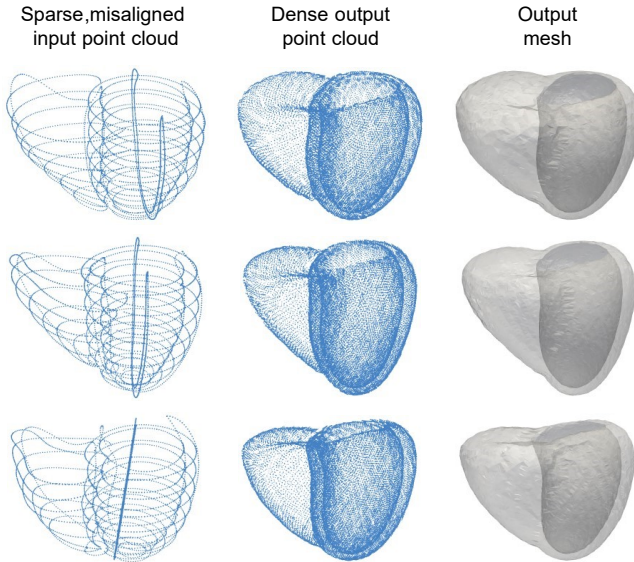
Fig. 3 shows the sparse and misaligned input point cloud, the reconstructed dense output point cloud with corrected misalignment, and the final biventricular mesh for three sample cases of the UKB dataset. Similar to our results on the SSM dataset, we find that the output point clouds exhibit an overall smooth structure with little to no remaining mis-

Table 2: Comparison of UKB results with references values

	Ours	Petersen et al. [18]
LV ED Volume (ml) ¹	112 (± 11)	124 (± 21)

¹ Values represent *mean* (\pm *SD*).

alignment and a high correspondence to the respective input shapes on both a local and a global level. The Ball Pivoting Algorithm [13] was able to quickly and accurately transform the point clouds into triangular meshes.

**Fig. 3:** Qualitative reconstruction results for three UKB cases.

For a quantitative validation of our method on the UKB dataset, we calculate the LV volume of 10 randomly selected female UKB cases, which were reconstructed with our method, and compare them to the clinical reference ranges, which were derived directly from the 2D SAX images of 432 female UKB participants [18] (Table 2). We find a difference in mean volume of about 12 ml, which is similar to previously reported error estimates [3].

4. DISCUSSION

We have shown in this work that PCNs can be highly effective in solving both the sparsity and misalignment problems in cardiac 3D surface reconstruction from cine MRI contours.

With a deep learning approach similar to Bai et al. [2] used for the segmentation task, our method is fully automatic and therefore does not suffer from inter-/intra-operator variability or other human errors. We utilize both SAX and LAX contours information, which gives our method access to slice intersection points and therefore decreases the negative effects of potential outliers. Once trained, our method is able

to generate dense point clouds with corrected misalignment in less than a tenth of a second (Table 1). This constitutes a vast reduction in processing time compared to previous iterative mesh-fitting methods (e.g. [8]), which need several hours to achieve results of similar quality. Our method acts directly on point clouds, which enables a fast, scalable, and memory-efficient surface reconstruction. This is in contrast to the grid-based representations (e.g. voxelgrids) required by classical deep learning approaches (e.g. [11]), which are inefficient at storing sparse surface data and have limited resolution leading to partial volume effects. This reduced spatial efficiency coupled with longer running times is especially disadvantageous for applications in clinical practice and large-scale cardiac physiology simulations.

Our method is able to correct even strong misalignments to Chamfer distances below image resolution, which demonstrates a high robustness and an ability to handle a wide variety of anatomical shapes, variations in image plane selection, and motion-induced slice misalignments. This is further corroborated by the low standard deviations of the Chamfer distances, which show the method’s low susceptibility to outliers. Furthermore, Chamfer distances evaluated on point clouds are known to overestimate the true surface-to-surface distances due to the discrete and sparse nature of point clouds. Therefore, we hypothesize that the true reconstruction quality might be slightly higher than the reported results.

We have successfully applied the PCN trained on the SSM dataset to the UKB dataset without any further training on the latter. This shows that we are able to extract realistically misaligned slices from the SSM meshes and that large-scale, high-quality datasets can be leveraged to train deep learning methods for applications on similar data. We have also observed a slight difference in mean LV volumes between our reconstructed meshes and the UKB reference values [18]. On the one hand, this is indicative of the plausibility of our results due to the closeness of values with established clinical precedent. On the other hand, the small difference might reflect the higher accuracy of volume calculations from dense 3D data compared to the 2D slices [3].

5. CONCLUSION

In this paper, we have developed and validated a deep learning method based on point completion networks that can generate high-quality biventricular 3D surface reconstructions from cardiac MRI contours in a fully automatic and fast process. Our method has been able to correct even strong amounts of misalignment and reconstruct meshes with accurate shapes and volumes from cine MR images of the UKB study.

6. COMPLIANCE WITH ETHICAL STANDARDS

This is a numerical simulation study for which no ethical approval was required.

7. ACKNOWLEDGMENTS

This research has been conducted using the UK Biobank Resource under Application Number ‘40161’. The authors express no conflict of interest. The work of A. Banerjee was partially supported by the CompBioMed 2 Centre of Excellence in Computational Biomedicine (European Commission Horizon 2020 research and innovation programme, grant agreement No. 823712) and the British Heart Foundation (BHF) Project under Grant HSR01230.

8. REFERENCES

- [1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.
- [2] Wenjia Bai, Matthew Sinclair, Giacomo Tarroni, Ozan Oktay, Martin Rajchl, et al., “Automated cardiovascular magnetic resonance image analysis with fully convolutional networks,” *Journal of Cardiovascular Magnetic Resonance*, vol. 20, no. 1, pp. 65, 2018.
- [3] Walter G O’Dell, “Accuracy of left ventricular cavity volume and ejection fraction for conventional estimation methods and 3D surface fitting,” *Journal of the American Heart Association*, vol. 8, no. 6, pp. e009124, 2019.
- [4] Dong Yang, Pengxiang Wu, Chaowei Tan, Kilian M Pohl, Leon Axel, and Dimitris Metaxas, “3D motion modeling and reconstruction of left ventricle wall in cardiac MRI,” in *International Conference on Functional Imaging and Modeling of the Heart*. Springer, 2017, pp. 481–492.
- [5] Benjamin Villard, Ernesto Zacur, Erica Dall’Armellina, and Vicente Grau, “Correction of slice misalignment in multi-breath-hold cardiac MRI scans,” in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2016, pp. 30–38.
- [6] Enzo Ferrante and Nikos Paragios, “Slice-to-volume medical image registration: A survey,” *Medical Image Analysis*, vol. 39, pp. 101–123, 2017.
- [7] Alireza Khatamian and Hamid R Arabnia, “Survey on 3D surface reconstruction,” *Journal of Information Processing Systems*, vol. 12, no. 3, 2016.
- [8] Benjamin Villard, Vicente Grau, and Ernesto Zacur, “Surface mesh reconstruction from cardiac MRI contours,” *Journal of Imaging*, vol. 4, no. 1, pp. 16, 2018.
- [9] Steffen E Petersen, Paul M Matthews, Fabian Bamberg, David A Bluemke, Jane M Francis, et al., “Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank-rationale, challenges and approaches,” *Journal of Cardiovascular Magnetic Resonance*, vol. 15, no. 1, pp. 46, 2013.
- [10] Meng Ye, Qiaoying Huang, Dong Yang, Pengxiang Wu, Jingru Yi, Leon Axel, and Dimitris Metaxas, “PC-U net: Learning to jointly reconstruct and segment the cardiac walls in 3D from CT data,” *arXiv preprint arXiv:2008.08194*, 2020.
- [11] Hao Xu, Ernesto Zacur, Jurgen E Schneider, and Vicente Grau, “Ventricle surface reconstruction from cardiac MR slices using deep learning,” in *International Conference on Functional Imaging and Modeling of the Heart*. Springer, 2019, pp. 342–351.
- [12] Wenjia Bai, Wenzhe Shi, Antonio de Marvao, Timothy JW Dawes, Declan P O’Regan, Stuart A Cook, and Daniel Rueckert, “A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion,” *Medical Image Analysis*, vol. 26, no. 1, pp. 133–145, 2015.
- [13] Fausto Bernardini, Joshua Mittleman, Holly Rushmeier, Cláudio Silva, and Gabriel Taubin, “The ball-pivoting algorithm for surface reconstruction,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 5, no. 4, pp. 349–359, 1999.
- [14] Andrew M Taylor and Jan Bogaert, “Cardiovascular MR imaging planes and segmentation,” in *Clinical Cardiac MRI*, pp. 93–107. Springer, 2011.
- [15] Wentao Yuan, Tejas Khot, David Held, Christoph Mertz, and Martial Hebert, “PCN: Point completion network,” in *International Conference on 3D Vision*. IEEE, 2018, pp. 728–737.
- [16] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas, “Pointnet: Deep learning on point sets for 3D classification and segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 652–660.
- [17] Yaoqing Yang, Chen Feng, Yiru Shen, and Dong Tian, “Foldingnet: Point cloud auto-encoder via deep grid deformation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 206–215.
- [18] Steffen E Petersen, Nay Aung, Mihir M Sanghvi, Filip Zemrak, Kenneth Fung, et al., “Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort,” *Journal of Cardiovascular Magnetic Resonance*, vol. 19, no. 1, pp. 18, 2017.