Predicting 3D Cardiac Deformations With Point Cloud Autoencoders

Marcel Beetz[1], Julius Ossenberg-Engels[1], Abhirup Banerjee[2,1][0000−0001−8198−5128], and Vicente Grau[1][0000−0001−8139−3480]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, UK

[2] Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DU, UK

Abstract. Mechanical contraction and relaxation of the heart play an important role in evaluating healthy and diseased cardiac function. Mechanical patterns consist of complex non-linear 3D deformations that vary considerably between subjects and are difficult to observe on 2D images, which impacts the prediction accuracy of cardiac outcomes. In this work, we aim to capture 3D biventricular deformations at the end-diastolic (ED) and end-systolic (ES) phases of the cardiac cycle with a novel geometric deep learning approach. Our network consists of an encoder-decoder structure that works directly with light-weight point cloud data. We initially train our network on pairs of ED and ES point clouds stemming from a mixed population of subjects with the aim of accurately predicting ED outputs from ES inputs as well as ES outputs from ED inputs. We validate our network's performance using the Chamfer distance (CD) and find that ED and ES predictions can be achieved with an average CD of 1.66 ± 0.62 mm on a dataset derived from the UK Biobank cohort with an underlying voxel size of $1.8 \times 1.8 \times 8.0$ mm [8]. We derive structural and functional clinical metrics such as myocardial mass, ventricular volume, ejection fraction, and stroke volume from the predictions and find an average mean deviation from their respective gold standards of 1.6% and comparable standard deviations. Finally, we show our method's ability to capture deformation differences between specific subpopulations in the dataset.

Keywords: Cardiac Deformation Prediction · Point Cloud Autoencoders · Cardiac Anatomy Reconstruction · Cardiac Contraction · Geometric Deep Learning · Cardiac MRI.

1 Introduction

Cardiovascular diseases are the most common cause of death in the world, accounting for 32% of all annual fatalities in 2019 [10]. A major driver behind this is the often insufficient understanding of cardiac pathologies, particularly the relationship between structural changes and cardiac function. Furthermore, the differentiation between subject subpopulations in terms of cardiac function

is often not taken into account. All of these affect predictions of clinical outcomes and consequently complicate the application of personalised treatment plans. In an attempt to address this issue, we aim to gain a more comprehensive understanding of the structure-function interactions in the heart. To this end, we propose a novel point cloud autoencoder network (PCN) to capture 3D cardiac deformation and model subpopulation-specific cardiac mechanical function. While deep learning on point cloud data has recently been applied to multiple problems in cardiac image analysis, such as segmentation [12], disease classification [4], and surface reconstruction [1], this work proposes, to the best of our knowledge, the first point cloud-based deep learning approach for cardiac deformation modelling. Grid-based deep learning techniques have previously been used for survival prediction [2], image registration [6], and motion-modelling for image sequences [5] but have not investigated subpopulation differences. We improve upon the closest prior work in this area [7] in several ways. Firstly, we utilize 3D instead of 2D data which enables us to capture real cardiac deformation and allows us to calculate clinical metrics based on volumetric changes in the cardiac cycle, such as ejection fraction (EF) and stroke volume (SV). Secondly, we work directly with point cloud data, rather than pixel information, which is more memory-efficient and has a straighforward expansion to 3D. Thirdly, the PCN is more efficient at working with surface models than grid-based networks, allowing us to scale to considerably higher surface densities. Finally, we employ the PCN to predict both phase directions, i.e. end-systolic (ES) cardiac shape conditioned on structural end-diastolic (ED) inputs as well as ED outputs from ES inputs, allowing us to model both cardiac contraction and relaxation.

2 Methods

We first provide a brief overview of our dataset before explaining the network architecture and training procedure of our proposed method.

2.1 Dataset

Our dataset consists of ∼500 female and ∼500 male subjects that were randomly chosen from the UK Biobank study [8]. We select cine MRI acquisitions at both ED and ES phases of the cardiac cycle for each subject and use the multi-step pipeline described in [1] to obtain corresponding 3D point cloud representations of the biventricular anatomy.

2.2 Network Architecture

Our method takes as input a multi-class point cloud that represents the biventricular anatomy at either end of the cardiac cycle (ED or ES) and is tasked to predict the corresponding deformed anatomical surface at the other extreme phase of the cardiac cycle. The network architecture consists of an encoder-decoder

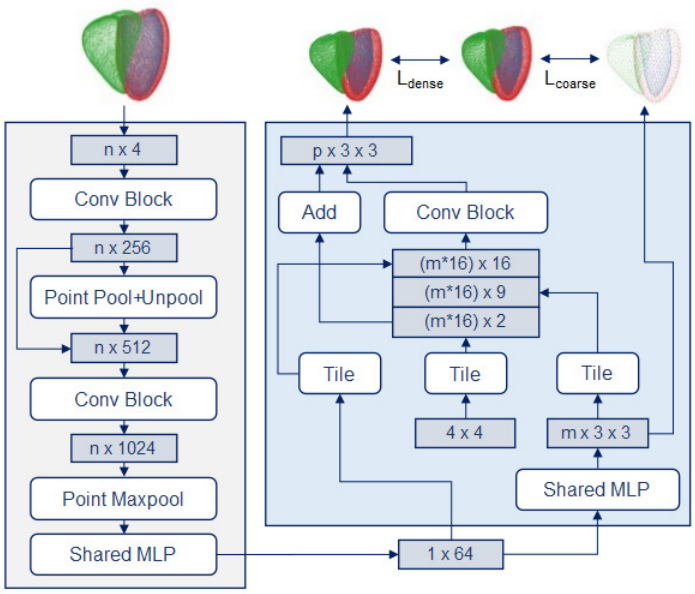

Fig. 1. Architecture of the proposed point cloud autoencoder network. The input point cloud consists of n points, each of which is stored as a 4-dimensional vector (x,y,z point coordinates and a class label). The network outputs both a coarse low-density point cloud and a dense high-resolution point cloud with separate 3D coordinates to represent class information.

structure inspired by the Point Completion Network [13] and its application to cardiac image analysis [1] (Fig. 1).

We use an extended version of the PointNet++ architecture [9] as our encoder by adding a conditional input to each point to indicate its cardiac substructure and multiple fully connected layers to facilitate the processing of the complex multi-class biventricular anatomy point clouds. The decoder first uses a shared multi-layer perceptron (MLP) to predict a low-resolution multi-class point cloud (coarse output) with the aim of capturing the global anatomical structure. This is followed by a FoldingNet [11] step, which outputs a high resolution multi-class point cloud (dense output) with accurate shapes both on a local and global level.

2.3 Training

The loss function of our network consists of two terms, one acting on the coarse prediction (L_{coarse}) and the other on the dense output (L_{dense}). In both cases, the predictions are compared to the gold standard point cloud using the Chamfer distance. A weighting parameter α is multiplied with the dense loss in the combined loss term to put more emphasis on generating a globally realistic shape at the beginning of training, while encouraging both high local and global accuracy

in later training stages. Accordingly, we start training with a low α of 0.01 and then gradually increase it during training until it reaches a value of 5.0.

3 Experiments and Results

For experiments in Sections 3.1-3.3, all networks were trained on ∼800 pairs of ED and ES point clouds, stemming in equal proportions from female and male subjects. At test time, ∼75 pairs of ED and ES point clouds were used as separate test sets for both female and male subjects.

3.1 Prediction Quality

In order to evaluate our network's prediction performance of both biventricular contraction and relaxation, two separate networks, one for each phase direction (ED to ES and ES to ED), were trained and tested jointly on both female and male subject data. We visualize the input, predicted, and gold standard point clouds of multiple sample cases from the unseen test dataset of the ES prediction task in Fig. 2 and of the ED prediction task in Fig. 3.

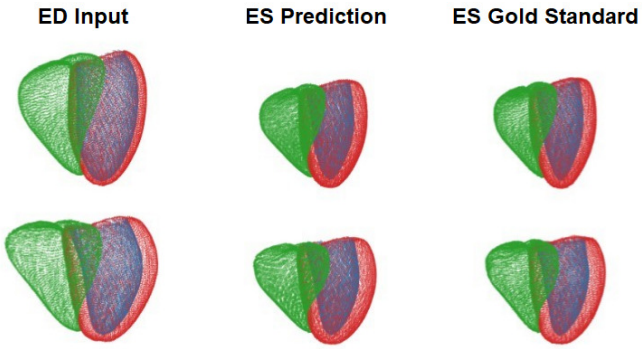

Fig. 2. ES prediction results of our method on two sample cases of the unseen test dataset.

Next, we quantify the performance of both the ED prediction and the ES prediction networks by calculating the Chamfer distances between the network predictions and the corresponding gold standard point clouds in the test set. We report the results separately for each sex and for the three ventricular substructures (left ventricular (LV) endocardium, LV epicardium, and right ventricular (RV) endocardium) to evaluate both overall and subpopulation-specific prediction quality, as well as enable a more localized assessment of deformation accuracy (Table 1).

Our method achieves Chamfer distances similar to the voxel resolution of the original MRI acquisition across all classes, sexes, and phases, while LV and ES prediction show slightly better results than RV and ED prediction.

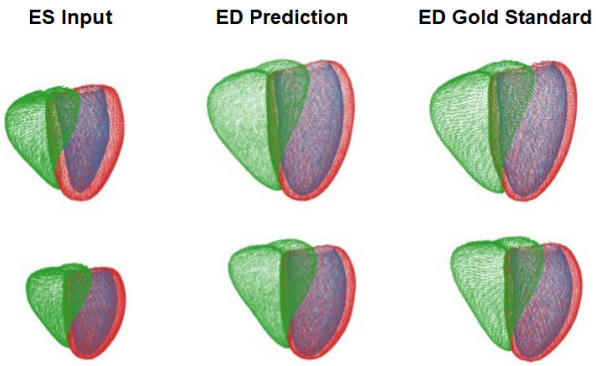

Fig. 3. ED prediction results of our method on two sample cases of the unseen test dataset.

Table 1. Prediction results of the proposed method on the test dataset.

Sex	Input Phase	Predicted Phase	Class	Chamfer Distance (mm)
Female	ED	ES	LV endocardium	1.35 (±0.65)
			LV epicardium	1.39 (±0.57)
			RV endocardium	1.69 (±0.77)
	ES	ED	LV endocardium	1.64 (±0.73)
			LV epicardium	1.61 (±0.81)
			RV endocardium	2.04 (±0.86)
Male	ED	ES	LV endocardium	1.37 (±0.47)
			LV epicardium	1.38 (±0.35)
			RV endocardium	1.75 (±0.50)
	ES	ED	LV endocardium	1.89 (±0.70)
			LV epicardium	1.68 (±0.43)
			RV endocardium	2.16 (±0.62)

Values represent mean (± standard deviation (SD)).

3.2 Cardiac Anatomy Prediction

In order to assess the predictive ability of our method from a clinical perspective, we first convert both the point clouds predicted by our method in Section 3.1 and the gold standard point clouds of the test dataset to multi-class surface meshes using the Ball Pivoting Algorithm [3]. We then calculate the LV volume,

LV mass, and RV volume for the predicted and gold standard meshes for both phase directions and report the results split by sex in Table 2.

Table 2. Cardiac volume metrics calculated from meshed point clouds predicted by our method.

Sex	Input Phase	Predicted Phase	Clinical Metric	Gold Standard	Prediction	% Diff
Female	ED	ES	LV ES volume (ml)	49 (±11)	49 (±9)	0.0
			RV ES volume (ml)	62 (±15)	62 (±12)	0.0
			LV mass (g)	86 (±14)	88 (±14)	3.4
	ES	ED	LV ED volume (ml)	123 (±23)	121 (±22)	1.6
			RV ED volume (ml)	146 (±25)	144 (±23)	1.4
			LV mass (g)	86 (±14)	85 (±14)	1.2
Male	ED	ES	LV ES volume (ml)	64 (±14)	63 (±11)	1.6
			RV ES volume (ml)	89 (±16)	86 (±14)	3.4
			LV mass (g)	121 (±24)	120 (±20)	0.8
	ES	ED	LV ED volume (ml)	153 (±29)	150 (±24)	2.0
			RV ED volume (ml)	187 (±28)	183 (±25)	2.1
			LV mass (g)	121 (±24)	121 (±22)	0.0

Values represent mean (± SD) in all cases.

Overall, we find good alignment between the predicted and gold standard scores in both means and standard deviations across all metrics, prediction tasks, and sexes. Differences tend to be greater for male hearts and ED predictions.

3.3 Cardiac Function Analysis

Since cardiac deformation is assessed using function-specific metrics in clinical practice, we also calculate ejection fraction (EF) and stroke volume (SV) of all test predictions and gold standard data for both chambers and phase directions, separating male and female subjects, to provide further validation of our method's performance in modelling biventricular deformations (Table 3).

We observe an average difference between the gold standard and predicted values of 1.72% indicating good overall prediction performance. Similar to the clinical anatomy metrics reported in Table 2, female hearts and ES predictions show slightly better scores.

Table 3. Cardiac function metrics calculated from meshed point clouds predicted by our method.

			ES Prediction		ED Prediction	
Sex	Clinical Metric	Gold Standard	Ours	% Diff	Ours	% Diff
Female	LV EF (%)	60 (±7)	60 (±6)	0.0	60 (±5)	0.0
	LV SV (ml)	75 (±17)	75 (±18)	0.0	73 (±14)	2.7
	RV EF (%)	65 (±23)	65 (±22)	0.0	63 (±19)	3.1
	RV SV (ml)	85 (±19)	84 (±19)	1.2	85 (±23)	0.0
Male	LV EF (%)	58 (±6)	59 (±6)	1.7	57 (±6)	1.7
	LV SV (ml)	90 (±21)	90 (±21)	0.0	86 (±17)	4.4
	RV EF (%)	59 (±18)	58 (±16)	1.7	56 (±19)	5.1
	RV SV (ml)	102 (±31)	104 (±18)	2.0	98 (±30)	3.9

Values represent mean (± SD) in all cases.

3.4 Subpopulation-Specific Cardiac Deformations

After evaluating our network's prediction performance on the entire test set, we also aim to analyze whether the PCN is able to extract features specific to certain subgroups in the data, enabling subpopulation-specific deformation modelling. To test this hypothesis, two separate networks were first trained on point cloud data pairs from exclusively female subjects, one in the ED to ES and the other in the ES to ED direction. They were then tested on data from both female and male subjects to investigate potential sex-specific differences in cardiac deformation prediction. Results are presented in Fig. 4 displaying the respective histograms of the CDs for each sex (different colors) and for both directions of deformation prediction (different subfigures). We find statistically significant differences in the prediction performance between male and female hearts for both prediction tasks (Kolmogorov-Smirnov (KS) test: p-value < 0.001).

4 Discussion

The results in Table 1 demonstrate that our method achieves high prediction accuracy with average CDs comparable to the voxel size of the underlying images ($1.8 \times 1.8 \times 8.0$ mm) [8]. Low standard deviations of approximately 0.5 mm indicate a robust prediction performance and show that our method is capable of accurately capturing the highly variable and complex non-linear 3D biventricular contraction and relaxation patterns. This is further corroborated by the qualitative prediction results in Fig. 2 and 3, which show good alignment between prediction and gold standard on both a global and local scale for both

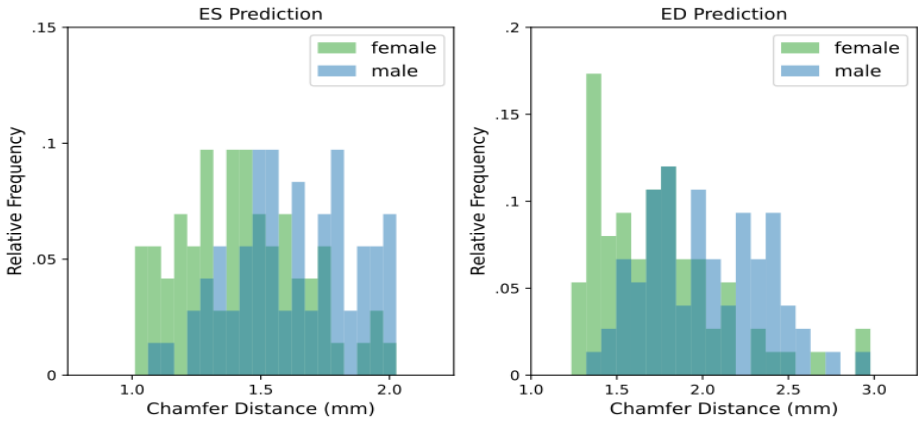

Fig. 4. Comparison of prediction performance on both female and male subjects using networks trained on only female cases. The Chamfer distances between prediction and gold standard point clouds were calculated for all female (green color) and male (blue color) cases in the test dataset. The resulting distributions of the Chamfer distances for each sex are presented as two histograms. The relative frequency (y-axis) indicates how often CD values fall within each of the pre-defined distance ranges (x-axis). Results are depicted for both ES prediction (left) and ED prediction (right).

phase directions. Since we work with lightweight point cloud data instead of the grid-based data structures used in previous approaches, we are able to achieve these results on surface representations with higher resolution and hence greater anatomical accuracy. This is only made possible by the PCN architecture which allows for fast and memory-efficient point cloud processing compared to previous deep learning approaches requiring regular grid structures. We observe that results are slightly better for predictions of ES from ED than vice versa and for female test subjects compared to males. We suspect that these lower overall CDs are likely caused by the fact that female hearts and hearts at ES are of smaller average size than male hearts and hearts at ED, while still retaining the same point cloud resolution. The difficulty in precisely locating the ES frame in an image sequence could also affect these results. When considering both volumetric cardiac anatomy and cardiac function metrics, our method's predictions result in very similar values to the reference values. This provides further evidence of the accuracy of the predicted shapes, both on their own and in relation to their respective input point clouds, and is crucial for the acceptance of more complex data-driven methods in clinical practice, as the results are in line with established clinical image-based biomarkers. From the histograms in Fig. 4, it can be seen that when training PCNs on just female point cloud data pairs and then testing on a mixed-sex population, the CDs for female subjects are significantly lower than those for male subjects. To better compare the distributions of CDs between test groups, the KS test was applied to assess histogram similarity. We find p-values below 0.001 in both phase directions, which indicate

statistically significant differences between the distributions. This supports the hypothesis that cardiac shape deformation is subpopulation-specific. When given one subpopulation to train on, the network extracts deformation patterns that are unique to that subpopulation in the data and cannot be found in the other, thereby performing better during test time.

5 Conclusion

We have shown in this work that point cloud autoencoder networks are highly capable of modelling global and local non-linear cardiac deformations and predicting relevant clinical metrics with a high level of accuracy. Their memory-sparse set-up makes them more efficient than grid-based methods, easily scalable to higher resolutions, and suitable for use in larger network architectures. Furthermore, we have observed that PCNs can be used to extract subpopulation-specific cardiac deformations, which opens up a range of future research avenues.

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 M. Beetz was supported by the Stiftung der Deutschen Wirtschaft (Foundation of German Business). The work of J. Ossenberg-Engels was supported by the Engineering and Physical Sciences Research Council (EPSRC) and Medical Research Council (MRC) [grant number EP/L016052/1]. The work of A. Banerjee was supported by the British Heart Foundation (BHF) Project under Grant HSR01230. The work of V. Grau was supported by the CompBioMed 2 Centre of Excellence in Computational Biomedicine (European Commission Horizon 2020 research and innovation programme, grant agreement No. 823712).

References

1. Beetz, M., Banerjee, A., Grau, V.: Biventricular surface reconstruction from cine MRI contours using point completion networks. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). pp. 105–109 (2021)
2. Bello, G.A., Dawes, T.J., Duan, J., Biffi, C., de Marvao, A., Howard, L.S., et al.: Deep-learning cardiac motion analysis for human survival prediction. Nature machine intelligence **1**(2), 95–104 (2019)
3. Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., Taubin, G.: The ball-pivoting algorithm for surface reconstruction. IEEE Transactions on Visualization and Computer Graphics **5**(4), 349–359 (1999)
4. Chang, Y., Jung, C.: Automatic cardiac MRI segmentation and permutation-invariant pathology classification using deep neural networks and point clouds. Neurocomputing **418**, 270–279 (2020)

5. Krebs, J., Mansi, T., Ayache, N., Delingette, H.: Probabilistic motion modeling from medical image sequences: application to cardiac cine-MRI. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 176–185. Springer (2019)
6. Krebs, J., Mansi, T., Mailhé, B., Ayache, N., Delingette, H.: Unsupervised probabilistic deformation modeling for robust diffeomorphic registration. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 101–109. Springer (2018)
7. Ossenberg-Engels, J., Grau, V.: Conditional generative adversarial networks for the prediction of cardiac contraction from individual frames. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 109–118 (2019)
8. Petersen, S.E., Matthews, P.M., Francis, J.M., Robson, M.D., Zemrak, F., Boubertakh, R., et al.: UK Biobank's cardiovascular magnetic resonance protocol. Journal of cardiovascular magnetic resonance **18**(1), 1–7 (2015)
9. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In: Advances in neural information processing systems. pp. 5099–5108 (2017)
10. WHO: Cardiovascular disease death rate (2019), https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
11. Yang, Y., Feng, C., Shen, Y., Tian, D.: Foldingnet: Interpretable unsupervised learning on 3D point clouds. arXiv preprint arXiv:1712.07262 (2017)
12. Ye, M., Huang, Q., Yang, D., Wu, P., Yi, J., Axel, L., Metaxas, D.: PC-U net: Learning to jointly reconstruct and segment the cardiac walls in 3D from CT data. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 117–126. Springer (2020)
13. Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M.: PCN: Point completion network. In: 2018 International Conference on 3D Vision (3DV). pp. 728–737 (2018)