Modelling Multi-phase Cardiac Anatomy with Generative Deep Learning

Thalia Eleni Seale[1], Blanca Rodriguez[2], Vicente Grau[1], Abhirup Banerjee[1]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, OX3 7DQ, United Kingdom
[2] Department of Computer Science, University of Oxford, Oxford, OX1 3QD, United Kingdom

Abstract

The motion of the cardiac anatomy has a considerable influence on cardiac function and disease development. Previous contributions to the analysis of cardiac anatomy have mostly focused on one or two phases from the cardiac cycle. However, a more complete understanding of cardiovascular disease could be achieved through the analysis and quantification of motion abnormalities in the anatomy over the full cardiac cycle. In this work, we propose a deep learning pipeline capable of reconstructing and generating continuous-time representations of the biventricular anatomy from a finite number of time points in the cardiac cycle. We demonstrate the proposed model provides interpretable quantification of geometric and motion characteristics of the anatomies. Our analysis over a dataset of 190 subjects shows that the reconstructions from the proposed model are accurate to sub-pixel resolution, with an average Chamfer distance of 1.71 ($\pm$ 1.13) mm.

1. Introduction

The diversity of the anatomy and function of the human cardiac anatomy between individuals has significant consequences for the efficacy of cardiovascular disease (CVD) prevention, diagnosis, and treatment. It also has implications for cardiac modelling such as for *in silico* trials [1]. There are known limitations to widely used clinical metrics *e.g.* left ventricular ejection fraction (LVEF), which reduce complex anatomy and motion to a single global score [2]. Modelling the full cardiac cycle could reveal diagnostically relevant information normally obscured.

Previous contributions to modelling the cardiac anatomy include a point cloud-based geometric deep learning network combined with a β-VAE framework to generate population-specific biventricular models from a single phase [3]. The work was extended in [4] by simultaneously modelling both the extreme, end-diastolic (ED) and end-systolic (ES), phases. Prior studies have focused only on specific phases of the cardiac cycle; however, MRI scans in clinical settings typically record 25-50 frames in the cardiac cycle. Modelling the motion of the full cycle would enable an extensive understanding of the cardiac anatomy and motion and, in turn, would facilitate investigations into pathological cardiac motion and the relationship between motion and electrophysiology.

In this article, we introduce a generative model for reconstructing the complete cardiac cycle, which transforms a discrete set of input phases to a continuous representation of the geometry across the full cardiac cycle. The proposed approach first applies principal component analysis (PCA) to extract the geometric features from the input point clouds. We then use a time variational autoencoder (TimeVAE) architecture [5] to input geometric features over time to obtain parameters for a continuous representation of the 3D shape changes over the full cycle. In an analysis of 190 subjects from the UK Biobank dataset [6], the proposed model is demonstrated to reconstruct point clouds through the entire cycle with an average Chamfer distance of 1.71 ($\pm$ 1.13) mm. We further illustrate the model's ability to disentangle meaningful geometric and motion characteristics of the cardiac cycle.

2. Methods

We begin with an overview of the dataset before describing the proposed model.

2.1. Dataset

The dataset consists of 190 subjects randomly chosen from the UK Biobank study [6], after excluding individuals with a history of CVD or obesity. Cine MRI acquisitions are processed using the pipeline from [7, 8] to obtain 3D point cloud representations of the biventricular anatomy, each containing 36,000 points. Points in each cloud are in correspondence and normalised.

2.2. Model

The model pipeline (see Fig. 1) consists of two steps:
1. **Geometric feature extraction.** A shared coordinate space is constructed using PCA. We project each frame

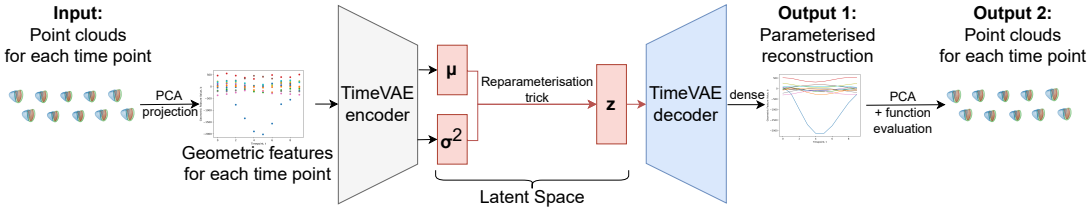

Figure 1. Schema of the proposed model pipeline.

onto the first k_{geom} principal components (PCs) to obtain a k_{geom}-dimensional time series.

2. **Temporal feature extraction.** A TimeVAE models the k_{geom}-dimensional time series. The output of the decoder is a set of parameters for k_{geom} d-degree polynomials. The polynomials are projected back into the point cloud space producing a continuous-time representation of the cardiac anatomy.

2.2.1. Geometric feature extraction

We first construct an interpretable, reduced-dimension coordinate space for the biventricular point clouds, irrespective of cardiac phases. PCA is used to find the linearly independent directions with highest variations in shape space.

Let $\mathbf{p}_i \in \mathbb{R}^3$ be a point in the biventricular point cloud. Each biventricular point cloud is a set of M points $Q = \{\mathbf{p}_1, \ldots, \mathbf{p}_M\}$. These are stacked into a feature vector

$$\mathbf{q} = [\mathbf{p}_1, \ldots, \mathbf{p}_M]^T \in \mathbb{R}^{3M}. \tag{1}$$

To obtain the principal components, we compute the eigenvectors $\mathbf{u}_i$ and eigenvalues λ_i of the covariance matrix. A reduced-dimension representation of the point clouds is obtained by selecting the first k_{geom} components explaining 98% of the variance. The eigenvectors $\mathbf{u}_i$ then form the basis of a reduced-dimension coordinate space for the point clouds. If $\mathbf{U}_{k_{\text{geom}}} = [\mathbf{u}_1|\mathbf{u}_2|\ldots|\mathbf{u}_{k_{\text{geom}}}]$ is a matrix with columns given by the first k_{geom} eigenvectors, then each point cloud can be represented in k_{geom} as

$$\mathbf{b} = \mathbf{U}_{k_{\text{geom}}}^T(\mathbf{q} - \bar{\mathbf{q}}), \tag{2}$$

where $\bar{\mathbf{q}}$ is the mean.

2.2.2. Temporal feature extraction

We next train a variational autoencoder to model the temporal variations in the geometric features extracted in the first step. We use the base encoder and decoder from the TimeVAE architecture [5]. The decoder outputs are transformed into a set of parameters for a polynomial regression that characterises the changes in contribution of each principal component over time, thereby providing a continuous representation of the cardiac motion.

The input to the encoder is $\mathbf{b} \in \mathbb{R}^{k_{\text{geom}} \times t_{\text{tot}}}$, where t_{tot} is the number of time points from the cardiac cycle provided as input. The base TimeVAE encoder outputs μ (mean), σ^2 (variance) $\in \mathbb{R}^{k_{\text{temp}}}$, which parameterise the latent space $z \in \mathbb{R}^{k_{\text{temp}}}$ of the model.

The base TimeVAE decoder deconvolves the latent space. It is then transformed by a dense layer into $\alpha \in \mathbb{R}^{k_{\text{geom}} \times (d-1)}$, which parameterises a set of polynomials $\hat{\mathbf{b}} : [0, 1] \to \mathbb{R}^{k_{\text{geom}}}$ each of degree d. t is a continuous variable representing the time in the cardiac cycle, with $t = \{0, 1\}$ both representing the end-diastolic phase. The ith polynomial gives a continuous reconstruction of the ith geometric feature. Each polynomial is given by

$$\hat{b}_i(t) = \alpha_{i,1} + t(t-1)\prod_{j=2}^{d-1}(t - \alpha_{i,j}), \tag{3}$$

ensuring that $\hat{\mathbf{b}}(0) = \hat{\mathbf{b}}(1)$.

The point cloud reconstruction at t, $\hat{Q}(t) = \{\hat{\mathbf{p}}_1(t), \ldots, \hat{\mathbf{p}}_M(t)\}$, is given by the unstacked projection of $\hat{\mathbf{b}}$ back into the point cloud coordinate space:

$$\hat{\mathbf{q}}(t) = \bar{\mathbf{q}} + \mathbf{U}_{k_{\text{geom}}}\hat{\mathbf{b}}(t). \tag{4}$$

2.3. Training and Loss

The loss function follows the design of the β-VAE with a reconstruction and a latent space loss balanced by parameter β.

$$L = L_{\text{recon}} + \beta L_{\text{KLD}}, \tag{5}$$

where L is the total loss, L_{recon} is the reconstruction loss, and L_{KLD} is the latent space loss given by the Kullback-Leibler divergence (KLD), which compares the latent distribution to an isotropic Gaussian distribution.

For the reconstruction loss, we use the average Euclidean distance between corresponding points in the ground truth point cloud Q_t and its reconstruction $\hat{Q}(t)$:

$$L_{\text{recon}} = \frac{1}{M}\sum_{i=1}^{M} ||\mathbf{p}_{i,t} - \hat{\mathbf{p}}_i(t)||_2. \tag{6}$$

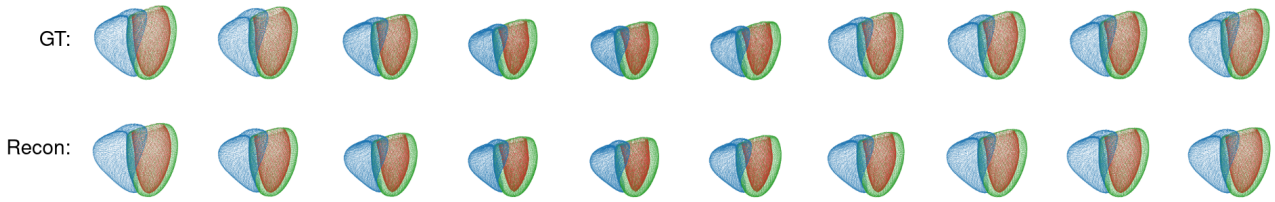

Figure 2. Qualitative comparison of the ground truth (GT) with reconstruction (Recon) results of the proposed method on one sample test case.

3. Results

In this section, we evaluate the reconstructive, predictive, and generative properties of the model as well as its explainability and interpretability.

We apply a train, validation, and test split of 160, 20, 10, respectively. Both train and validation sets have 10 available cardiac phases with frame numbers 0, 5, 10, 15, 20, 25, 30, 35, 40, and 45 from the MRI acquisition. The test set contains all 50 frames.

For geometric feature extraction, 98% of the variance is explained by the first 12 principal components. The degree of the polynomial is experimentally selected as $d = 6$. The network is trained over 5000 epochs; we select the model with the best L_{recon} with $L_{\text{KLD}} \leq 1.5$ for favourable reconstruction and generation trade-off.

3.1. Reconstruction and prediction

To assess the model's reconstruction and prediction performance, we select point clouds from the unseen test dataset as our gold standard, and compare the network's outputs to the ground truth using Chamfer distance.

Qualitative results are shown in Fig. 2. Point clouds at the 10 cardiac phases observed during training are reconstructed with average Chamfer distance of 1.58 ($\pm$ 0.32) mm; the unobserved 40 cardiac phases are predicted with average Chamfer distance 1.74 ($\pm$ 1.26) mm. The mean Chamfer distance for all reconstructed frames is 1.71 ($\pm$ 1.13) mm, which is below the voxel resolution of the underlying images ($1.8 \times 1.8 \times 8$ mm^3).

3.2. Full cycle generation

The generative performance of the model is demonstrated by randomly sampling from the latent space probability distributions. Figure 3 shows the point cloud sets generated from such samples. We observe both noticeable differences in cardiac shapes and sizes between samples and realistic deformations in the anatomy throughout the cardiac cycle, indicating the decoder can generate diverse anatomy and motion characteristics.

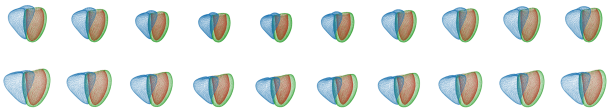

Figure 3. Generated point clouds from two randomly sampled latent space vectors.

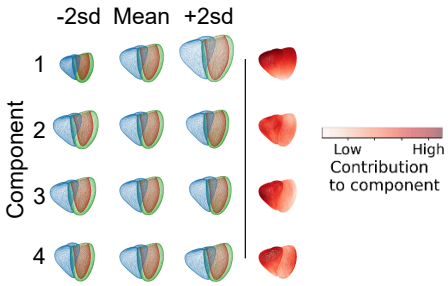

Figure 4. Plots of the first 4 PCs for the training set. *Left:* cardiac shape changes over 2 standard deviation (sd) for each component. *Right:* magnitude of contributions for each point from the cloud to each PC.

3.3. Geometric feature and latent space analysis

By first constructing a basis using the first k principal components, we find a set of interpretable time series as inputs and outputs for the TimeVAE. Figure 4 shows the interpretable changes to the anatomy by varying each component. For example, PC1 encodes mainly scale, whereas PC4 encodes the overall width of the anatomy.

We analyse the contributions of latent variables in the neural network to the generated functions by varying individual latent space components while keeping the remaining latent space constant, and passing the resulting vectors through the decoder to obtain the respective outputs.

Figure 5 demonstrates the gradual, interpretable changes to the biventricular shapes, sizes, and motions for two latent variables. Varying latent 15 shifts the position of the minimum achieved by PC1 (scale) which corresponds to

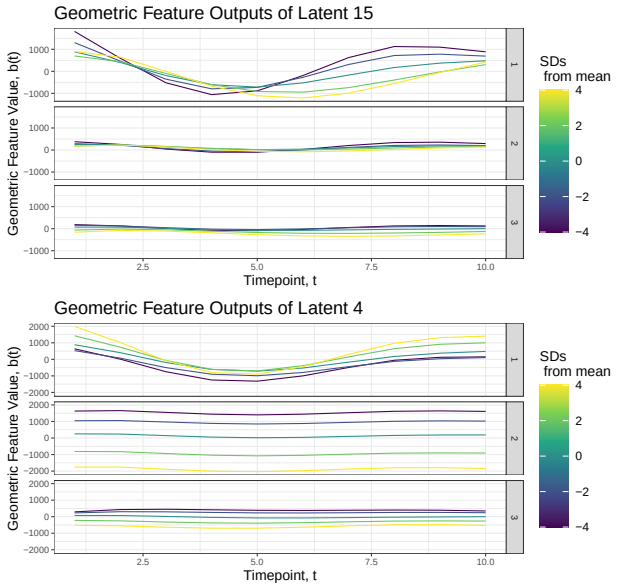

Figure 5. Plots demonstrating the interpretability of the latent space of the TimeVAE.

the end-systolic phase – a motion characteristic. Varying latent 4 leads to large changes in average value of PC2, corresponding to a geometric characteristic.

4. Discussion

The proposed model provides a novel quantification of geometric and temporal features within the cardiac cycle, which can be used to characterise differences in cardiac function across populations. From the example in Fig. 2, we see that the model can represent a range of anatomies.

The continuous-time representation allows for flexible outputs for use in downstream tasks such as *in silico* trials. Further work could include subject metadata in the latent space of the TimeVAE model; for personalised medicine, this could improve predictions for patients with minority status, and for *in silico* trials, the model would allow simulation of populations with known demographics.

The focus of this study was to establish a proof of concept before tackling the more complex pathological geometric variations. Future work will include a larger variety of anatomies such as those with CVD and validate the model against existing clinical metrics such as ejection fraction.

5. Conclusion

The current study presents a novel generative deep learning model for reconstructing and predicting a continuous-time representation of the biventricular anatomy over the full cardiac cycle. The network described uses PCA and extends TimeVAE and β-VAE architectures to create a generative deep learning model with interpretable latent features. The method provides the basis for modelling differences in the motion and geometry among populations, with the capacity to represent geometries with sub-voxel accuracy.

Acknowledgments

This research has been conducted using the UK Biobank Resource under Application Number '40161'. The authors express no conflict of interest. TS is supported by the EPSRC Centre for Doctoral Training in Health Data Science (EP/S02428X/1). AB is supported by the Royal Society University Research Fellowship (Grant No. URF\R1\221314). BR was funded by a Wellcome Trust Fellowship in Basic Biomedical Sciences (214290/Z/18/Z).

References

[1] Corral-Acero J, et al. The 'digital twin' to enable the vision of precision cardiology. European Heart Journal 2020; 41(48):4556–4564.

[2] Borlaug BA, Sharma K, Shah SJ, Ho JE. Heart failure with preserved ejection fraction: JACC scientific statement. Journal of the American College of Cardiology 2023;.

[3] Beetz M, et al. Generating subpopulation-specific biventricular anatomy models using conditional point cloud variational autoencoders. In Statistical Atlases and Computational Models of the Heart. Multi-Disease, Multi-View, and Multi-Center Right Ventricular Segmentation in Cardiac MRI Challenge. 2022; 75–83.

[4] Seale T, et al. Modelling multi-phase cardiac anatomy using point cloud variational autoencoders. In IEEE International Symposium on Biomedical Imaging (ISBI). 2024; 1–5.

[5] Desai A, et al. TimeVAE: A variational auto-encoder for multivariate time series generation. arXiv preprint arXiv211108095 2021;.

[6] Petersen SE, et al. UK Biobank's cardiovascular magnetic resonance protocol. Journal of Cardiovascular Magnetic Resonance 2015;18(1):1–7.

[7] Banerjee A, et al. A completely automated pipeline for 3D reconstruction of human heart from 2D cine magnetic resonance slices. Philosophical Transactions of the Royal Society A 2021;379(2212):20200257.

[8] Beetz M, et al. Multi-class point cloud completion networks for 3D cardiac anatomy reconstruction from cine magnetic resonance images. Medical Image Analysis 2023;90:102975.

Address for correspondence:

Thalia Eleni Seale
Old Road Campus Research Building, Oxford, OX3 7DQ, United Kingdom
thalia.seale@eng.ox.ac.uk