Reconstructing 3D Cardiac Anatomies from Misaligned Multi-View Magnetic Resonance Images with Mesh Deformation U-Nets

Marcel Beetz
Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford
Oxford, OX3 7DQ, UK

Abhirup Banerjee
Division of Cardiovascular Medicine
Radcliffe Department of Medicine
University of Oxford
Oxford, OX3 9DU, UK

Vicente Grau
Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford
Oxford, OX3 7DQ, UK

Abstract

High-quality three-dimensional (3D) representations of cardiac anatomy and function are crucial for improving cardiac disease diagnosis beyond strictly volume-based biomarkers used in current clinical practice, as well as for the accurate simulation of cardiac electrophysiology and mechanics. However, current gold standard cardiac magnetic resonance imaging (MRI) protocols typically only acquire a set of 2D slices to approximate the true 3D morphology of the underlying heart. In this work, we propose a novel geometric deep learning method, the Mesh Deformation U-Net, to reconstruct 3D cardiac surface meshes from 2D MRI slices as the key part of a fully automatic end-to-end pipeline. Its architecture combines spectral graph convolutions and mesh sampling operations in a hierarchical encoder-decoder structure to enable efficient multi-scale feature learning directly on mesh data. A targeted preprocessing step approximately fits a template mesh to the sparse MRI contours, before the Mesh Deformation U-Net corrects for motion-induced slice misalignment by simultaneously utilizing information from multiple MRI views and the template-induced anatomical shape prior. We evaluate the Mesh Deformation U-Net on a large synthetic dataset of heart anatomies and obtain small reconstruction errors below the pixel size of the underlying image resolution for three different cardiac substructures. Furthermore, we apply the pre-trained Mesh Deformation U-Net as the key component of a 4-step reconstruction pipeline to cine magnetic resonance images of the UK Biobank and observe realistic heart reconstructions on both a local and global level. We calculate multiple widely used clinical metrics for the reconstructed meshes and obtain values in line with other large-scale population studies.

©2000 Beetz et al.
Beetz et al.

Keywords: Cardiac Surface Reconstruction, Mesh Deformation U-Nets, Cine MRI, Slice Misalignment Correction, Motion Artifact, Spectral Graph Convolutions, Mesh Pooling, Geometric Deep Learning

1. Introduction

Cardiac magnetic resonance imaging (MRI) is one of the gold standard medical imaging modalities for the diagnosis of cardiac anatomy and function (Stokes and Roberts-Thomson, 2017). The current clinical practice, mostly the cardiac cine MRI protocol, usually images the heart myocardium with high soft-tissue contrast in a collection of sparse and intersecting (a stack of short-axis and few long-axis) 2D image planes, with a typical $1-2 \times 1-2 \text{mm}^2$ in-plane resolution and $8-10 \text{mm}$ out-of-plane resolution. However, this can only present the underlying heart anatomy at a finite number of spatial locations and orientations and as a result, provides a sparse representation of the true 3D geometry of the human heart, limiting the accuracy of cardiovascular disease diagnosis (Corral Acero et al., 2022; Di Folco et al., 2022; O’Dell, 2019; Suinesiaputra et al., 2017).

In order to solve the sparsity issue of 2D cine MRI, many previous works have attempted to reconstruct the complete 3D cardiac surfaces from the acquired 2D slices (Banerjee et al., 2021a; Joyce et al., 2022; Mauger et al., 2019; Villard et al., 2018). One of the biggest challenges for 3D cardiac surface reconstruction from sparse 2D cine MRI acquisitions is the presence of slice misalignment artifacts induced by the respiratory or breathing motion, cardiac or heart-beating motion, and the movements caused by the patient or imaging device. Several previous approaches have aimed to first segment the acquired 2D images in the cardiac substructures of interest and then address the misalignment artifacts by optimising the consistency among the sparse heart contours, thus accounting mainly for the in-plane misalignment artifacts (McLeish et al., 2002; Su et al., 2014; Villard et al., 2017). A recent approach aims to fit a statistical shape model (SSM) (Bai et al., 2015) to the sparse heart contours in 3D space and then optimally align the contours on the SSM, accounting for both in-plane and out-of-plane misalignments (Banerjee et al., 2021b). The final 3D reconstruction usually relies on fitting a smooth surface mesh to the resulting contours in a per-case regularised optimisation procedure (Banerjee et al., 2021a; Joyce et al., 2022; Mauger et al., 2019; Villard et al., 2018).

Although these methods can accurately model the patient-specific 3D deformations for cardiac surface modeling, they suffer from extended execution times, making them unsuitable for population-level analysis. Recently, deep learning based algorithms have become the state-of-the-art approaches for 3D cardiac surface reconstruction tasks, primarily due to their faster execution times, easy scalability, high versatility, and good performance based on the pretrained models. However, these approaches lack validation on real datasets (Xu et al., 2019), can only process single image inputs (Wang et al., 2020; Zhou et al., 2019), rely on inefficient voxelgrid representations of anatomical surface data (Chen et al., 2021; Xu et al., 2019), have been evaluated on only a small number of real cases (Beetz et al., 2021; Beetz et al., 2021a), or require additional preprocessing and postprocessing steps that are complex and error-prone (Beetz et al., 2021a; Chen et al., 2021).

In this work, we propose the Mesh Deformation U-Net as a novel geometric deep learning approach to cardiac surface reconstruction. Its architecture follows a hierarchical encoder-
decoder design with U-Net-inspired (Ronneberger et al., 2015) long skip connections and is specifically designed to directly process anatomical surface mesh data in an efficient manner. This enables the network to overcome the considerable memory and execution inefficiencies of previous voxelgrid-based deep learning (Xu et al., 2019) and per-case optimisation approaches (Banerjee et al., 2021a, Joyce et al., 2022, Villard et al., 2018) respectively. In addition, a targeted template mesh fitting preprocessing step introduces an anatomical shape prior in a fast and straightforward manner which is particularly beneficial for the highly sparse anatomical input data. Furthermore, it allows the Mesh Deformation U-Net to retain the same vertex connectivity for all meshes in the dataset which is an important requirement for many follow-up 3D cardiac modeling tasks (Mauger et al., 2019, Beetz et al., 2021b c, Corral Acero et al., 2022, Beetz et al., 2022a b, Li et al., 2022).

2. Methods

2.1 Dataset and Preprocessing

We use two datasets in this work for method development and evaluation. The first one is a synthetic dataset derived from a 3D MRI-based statistical shape model (SSM) (Bai et al., 2015). Hereby, we first generate a population of 250 3D cardiac anatomies from the SSM which we consider as our ground truth 3D shapes for the reconstruction task, since they are based on high-resolution 3D MRI acquisitions (1.25 × 1.25 × 2 mm). Next, we slice each of the meshes at the common imaging planes of a typical cine MRI acquisition and then apply 10 different random rigid transformations to each slice to introduce slice misalignment that mimics real acquisitions and obtain the sparse, misaligned inputs for our method as point clouds. Since the Mesh Deformation U-Net requires 3D triangular meshes with the same vertex connectivity as its inputs, we design an approximate mesh fitting step to convert the sparse anatomy point clouds into dense meshes. To this end, we select the mean mesh of the SSM as our template mesh and deform it to fit the respective sparse point clouds. We determine the deformation based on the Earth Mover’s distance (EMD) between the sparse point cloud and the template mesh. The EMD provides a one-to-one mapping between points and vertices by minimising the average mapping distance between point cloud \( P_1 \) and mesh vertices \( P_2 \), as follows:

\[
EMD(P_1, P_2) = \min_{\phi: P_1 \rightarrow P_2} \frac{1}{|P_1|} \sum_{x \in P_1} \| x - \phi(x) \|_2. \tag{1}
\]

We move each vertex of the template mesh to the spatial location of the corresponding closest point of the point cloud as determined by the Earth Mover’s distance in our deformation step. This results in an approximate fit of the template mesh to the sparse contours which includes all anatomical information of the MRI segmentation in the initial fitted mesh, while maintaining the template-induced built-in shape prior in a dense representation of the cardiac surface.

The second dataset consists of 1000 real cine MR images of the UK Biobank study (Petersen et al., 2013) with a pixel resolution of 1.8 × 1.8 × 8.0 mm. In order to apply the Mesh Deformation U-Net to this real dataset, we develop the 4-step reconstruction pipeline described in Sec. 2.4.
2.2 Mesh Deformation U-Net

The architecture of the proposed Mesh Deformation U-Net utilises recent advances in geometric deep learning in a hierarchical encoder-decoder structure.

Its inputs and outputs are 3D surface meshes with the same vertex-connectivity, and the vertices are stored as $n \times 3$ vectors where $n$ refers to the number of vertices and 3 to the x,y,z coordinates of each vertex respectively. In our experiments, we apply the network to three cardiac substructures and use a different number of vertices, namely 1870, 2620, and 1706, for the LV endocardial, LV epicardial, and RV endocardial anatomy meshes, respectively. The network architecture consists of multiple blocks of graph convolution, rectified linear unit (ReLU) activation layers, and mesh sampling operations organised in a hierarchical structure to enable efficient multi-scale feature learning directly on mesh data. Similar to the U-Net (Ronneberger et al., 2015), the encoder and decoder follow a symmetric design, where corresponding levels have the same mesh resolution and number of feature maps and are connected by long skip connections to facilitate the information flow between earlier and later parts of the network. We use spectral graph convolutions for feature learning which we calculate via the Chebyshev polynomial approximation of order 5 (Defferrard et al., 2016), while the sampling operations rely on quadric error minimisation (Ranjan et al., 2018).

2.3 Network Training and Implementation

We select the vertex-wise mean squared error between the predicted meshes and ground truth meshes as the loss function for the Mesh Deformation U-Net. We choose the Adam optimiser (Kingma and Ba, 2014) with a learning rate of 0.001 and a batch size of 8 to train all networks on a CPU with 8 GB memory. We employ a dataset split of 75%/5%/20% for the training, validation, and testing dataset, respectively, and stop the training process when no loss improvement has been observed on the validation dataset for 10 epochs. We use the PyTorch (Paszke et al., 2019) and PyTorch Geometric (Fey and Lenssen, 2019) frameworks to implement our deep learning and geometric deep learning code, respectively.
2.4 Cardiac Anatomy Reconstruction Pipeline

In order to utilise the Mesh Deformation U-Net for cardiac anatomy reconstruction from real cardiac MRI acquisitions, we propose the 4-step pipeline depicted in Fig. 2.

In the first step, we segment the input MR images into four classes (LV endocardium and epicardium, RV endocardium, background) using three separate convolutional neural networks for the short-axis (SAX), 4 chamber long-axis (LAX), and 2 chamber LAX views, respectively (Bai et al., 2018). Next, we extract the contours of the segmentation masks for each cardiac substructure and place the combined information from each view in 3D space as a point cloud (Banerjee et al., 2021a). Then, we fit the template mesh of the SSM dataset to this sparse point cloud representation using the mesh fitting procedure described in Sec. 2.1 to obtain a dense anatomy mesh. Finally, a Mesh Deformation U-Net is applied to remove any slice misalignment from the mesh in the key step of the pipeline and output the final reconstructed 3D cardiac shape.

3. Experiments

3.1 Reconstruction From Synthetic MRI Contours

First, we want to assess whether the Mesh Deformation U-Net combined with the approximate mesh fitting step is capable of accurately reconstructing synthetic 3D cardiac shapes from sparse and misaligned contours. To this end, we first apply the mesh fitting step to all sparse anatomy point clouds of the SSM dataset (Sec. 2.1). Next, we train a Mesh Deformation U-Net using the fitted meshes and the ground truth meshes of the training dataset. Then, we select the fitted meshes of the test dataset and pass them through the trained Mesh Deformation U-Net to obtain its predicted mesh reconstructions on unseen data. We execute this process separately for the LV endocardium and epicardium and RV endocardium, and present the results of three sample cases for each substructure in Fig. 3.

We observe that the mesh fitting step has successfully incorporated the anatomical information of the sparse point clouds while maintaining a realistic shape despite the shared connectivity restriction. The predicted surface meshes align well with the respective gold standard ones on both a global and local level and for different input shapes and misalignments. The performance is consistent across all three cardiac substructures.
After a qualitative validation, we next aim to quantify the reconstruction performance of the Mesh Deformation U-Net on the SSM dataset. We select the median surface distance, the mean surface distance, and the Hausdorff distance between the respective predicted and gold standard meshes of the test dataset as our evaluation metrics and report the results separately for each cardiac substructure in Table 1.

<table>
<thead>
<tr>
<th>Cardiac structure</th>
<th>Median surface distance (mm)</th>
<th>Mean surface distance (mm)</th>
<th>Hausdorff (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LV endocardium</td>
<td>0.78 (±0.17)</td>
<td>0.98 (±0.23)</td>
<td>4.69 (±1.67)</td>
</tr>
<tr>
<td>LV epicardium</td>
<td>0.77 (±0.19)</td>
<td>1.00 (±0.25)</td>
<td>4.65 (±1.41)</td>
</tr>
<tr>
<td>RV endocardium</td>
<td>0.97 (±0.25)</td>
<td>1.25 (±0.32)</td>
<td>4.77 (±1.46)</td>
</tr>
</tbody>
</table>

Values represent mean/median (± standard deviation/quartile deviation).

We find surface distance values below the voxel size of the underlying 3D MRI acquisition (1.25 × 1.25 × 2 mm) for all cardiac substructures, with slightly lower scores for the left ventricular reconstruction than the right ventricular one.

### 3.2 Reconstruction Pipeline on Real MR Images

After the evaluation of the Mesh Deformation U-Net on synthetic MRI contours, we next assess the capability of our proposed 4-step pipeline to reconstruct 3D shapes from real MR images. To this end, we select the Mesh Deformation U-Net pre-trained on the SSM dataset and apply it together with the three other pipeline steps to the raw MRI acquisition of the UK Biobank dataset. We execute the pipeline separately for each cardiac substructure and visualise the point cloud contours, approximately fitted mesh, and the final predicted mesh of three sample cases in Fig. 4.

Similar to the results on the SSM dataset, we observe that the mesh fitting step is sufficiently well designed to accurately encapsulate information from the anatomy contours in the fitted mesh, even for cases with stronger misalignment. The final mesh predictions
Figure 4: Reconstruction results of three sample cases of the UKBB dataset. exhibit realistic cardiac surfaces that adequately reflect the underlying input anatomies without showing any major remaining misalignment artifacts.

Since no gold standard 3D meshes are available for the UK Biobank dataset, we evaluate our pipeline’s reconstructions with multiple widely used clinical metrics on a population level. To this end, we calculate the LV volume, LV mass, and RV volume of the reconstructed UK Biobank meshes separately for female and male subjects and compare the results with two other large-scale population studies of cardiac anatomy in Table 2. While Petersen et al. (2017) computed the metrics based on 2D MRI slices only using the modified Simpson’s rule, Bai et al. (2015) obtained the values directly from 3D representations of the heart.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>LV Volume (ml)</td>
<td>124 (±21)</td>
<td>138 (±24)</td>
<td>129 (±21)</td>
</tr>
<tr>
<td></td>
<td>LV Mass (g)</td>
<td>70 (±13)</td>
<td>96 (±16)</td>
<td>84 (±18)</td>
</tr>
<tr>
<td></td>
<td>RV Volume (ml)</td>
<td>130 (±24)</td>
<td>-</td>
<td>131 (±30)</td>
</tr>
<tr>
<td>Male</td>
<td>LV Volume (ml)</td>
<td>166 (±32)</td>
<td>178 (±36)</td>
<td>166 (±38)</td>
</tr>
<tr>
<td></td>
<td>LV Mass (g)</td>
<td>103 (±21)</td>
<td>128 (±24)</td>
<td>122 (±29)</td>
</tr>
<tr>
<td></td>
<td>RV Volume (ml)</td>
<td>182 (±36)</td>
<td>-</td>
<td>174 (±43)</td>
</tr>
</tbody>
</table>

Values represent mean (± standard deviation) in all cases.

We find that the values achieved by our pipeline generally lie between the ones reported by Petersen et al. (2017) and Bai et al. (2015), for all metrics and both subpopulations.

4. Discussion and Conclusion

In this work, we have developed the Mesh Deformation U-Net as the key component of a novel 3D cardiac surface reconstruction pipeline from cine MR images. Our experiments
on the SSM dataset have shown that the Mesh Deformation U-Net is capable of accurately reconstructing hearts with errors below the underlying image resolution. On the one hand, this indicates that the proposed mesh fitting step is able to successfully retain the important anatomy information of the inputs, while simultaneously maintaining the same vertex connectivity and overcoming the sparsity issue in the point cloud contours. On the other hand, it demonstrates that the hierarchical multi-scale architecture of the Mesh Deformation U-Net is adequately designed to capture relevant local and global shape features and to directly process high resolution 3D mesh data in an efficient manner. These conclusions are further corroborated by the findings that the high reconstruction accuracy is achieved for three separate cardiac substructures and for a variety of different cardiac shapes and sizes while using highly sparse contours with varying misalignments as inputs. Furthermore, we observe that the proposed pipeline reconstructs realistic 3D hearts on the real UK Biobank dataset with clinical metrics in line with other large-scale population studies. This shows not only that the SSM dataset is adequately designed to reflect real acquisition conditions but also that both the mesh fitting step and the Mesh Deformation U-Net are highly suitable for cardiac surface reconstruction from real MR images. Their straightforward integration into the multi-step pipeline also demonstrates a high degree of applicability and robustness, especially in light of the cross-domain transfer from synthetic to real data.

The proposed pipeline is also designed to simultaneously consider anatomy information from both SAX and LAX views, which is in contrast to many previous works focusing only on SAX slices (Lamata et al., 2014). This combined multi-view processing allows the network to utilise additional information and is especially beneficial in our highly sparse settings. It is also of particular importance in the apical and basal regions of the heart, where only very limited information is available from the SAX slices alone. However, due to the shape prior induced by the template mesh in our mesh fitting step and the direct mesh processing of our network, we hypothesise that our pipeline would still yield positive results even in the absence of LAX information or the presence of other slice information. Due to the deep learning basis of our method, it comes with considerable speed and memory advantages after training compared to case-specific reconstruction techniques (Lamata et al., 2014; Banerjee et al. 2021a,b), while retaining high levels of accuracy. These characteristics are crucial for an easy and robust scaling of the approach to large population cohorts. The benefits also still hold, although to a lesser extent, in comparison with previous voxelgrid-based deep learning approaches (Xu et al., 2019). Since the Mesh Deformation U-Net utilises geometric deep learning operations which can be directly applied to lightweight mesh representations, it can process the same 3D cardiac surface data with a fraction of the time and memory requirements of a voxelgrid-based approach. This allows an easier scaling to both larger numbers of cases and higher data resolutions. It also enables the training of our network on a standard CPU as opposed to the typically required high-powered GPUs. While previous point cloud-based deep learning approaches (Beetz et al., 2021a) also share these advantages, they require a complicated meshing step for many follow-up tasks and are not able to easily create meshes with shared vertex-connectivity that are required for many follow-up 3D cardiac modeling tasks. Furthermore, while the mesh fitting step required by our pipeline induces additional complexity and possible errors, we believe that its regularising role as a geometric shape prior as well as its non-linear deformation abilities are still advantageous, especially in our highly sparse settings.
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 is supported by the Stiftung der Deutschen Wirtschaft (Foundation of German Business). A. Banerjee is a Royal Society University Research Fellow and is supported by the Royal Society (Grant No. URF\R1\221314). The work of A. Banerjee and V. Grau is supported by the British Heart Foundation (BHF) Project under Grant HSR01230. The work of V. Grau is also supported by the CompBioMed 2 Centre of Excellence in Computational Biomedicine (European Commission Horizon 2020 research and innovation programme, grant agreement No. 823712).

References


Xiang Chen, Nishant Ravikumar, Yan Xia, Rahman Attar, Andres Diaz-Pinto, Stefan K Piechnik, Stefan Neubauer, Steffen E Petersen, and Alejandro F Frangi. Shape registration with learned deformations for 3D shape reconstruction from sparse and incomplete point clouds. *Medical Image Analysis*, page 102228, 2021.


